UNIVERSIDAD CARLOS III DE MADRID

TESIS DOCTORAL

Three essays on the relative outcomes of minority groups in the Spanish labour market

Autor:
Raquel Vegas Sánchez

Director/es:
Juan Jose Dolado Llobregat

DEPARTAMENTO ECONOMIA

Getafe, Mayo 2011

TESIS DOCTORAL

Three essays on the relative outcomes of minority groups in the Spanish labour market

Autor: Raquel Vegas Sánchez

Director/es: Juan José Dolado

Firma del Tribunal Calificador:

Firma

Presidente: Sergi Jimenez

Vocal: Isabel Argimón

Vocal: Iliana Regio

Vocal: Libertad Gonzalez

Secretario: Ricardo Mora

Calificación:

Getafe, Mayo de 2009

THREE ESSAYS ON THE RELATIVE OUTCOMES OF MINORITY GROUPS IN THE SPANISH LABOUR MARKET

Raquel Vegas Sánchez

December 30

NOTA DE AGRADECIMIENTOS

Llegar hasta aquí ha sido difícil. Por ello, no quiero perder la oportunidad de agradeceros a todos los que durante estos años habéis participado de una u otra manera en esta etapa de mi vida. Muchas gracias a todos por vuestro apoyo y cariño que me han hecho crecer personal y profesionalmente.

Esta tesis no podría haber sido posible sin la ayuda de mi director de tesis, Juanjo Dolado. Gracias Juanjo por ser un guía decidido y paciente, y por confiar en mí más que yo misma.

Gracias también a George y Ricardo, por vuestro apoyo. Puede que ninguno de los dos seáis conscientes de cuanto habéis hecho por mí, pero vuestros ánimos y consejos me abrieron las puertas a una nueva etapa ilusionante en mi carrera académica.

Gracias a Noé y a Rebeca, por dedicarme generosamente vuestro tiempo y adaptaros a mis horarios en los días de la locura del primer año de doctorado y también por vuestra compañía y amistad. Gracias Rebeca por decirme las cosas tal y como son y por permanecer siempre a mi lado. Gracias a Riia, Ana y Maye, con quienes compartí lo peor del viaje. Gracias también a ti Fer, compañera y amiga de FEDEA, que llegaste en esta última etapa, pero que has conseguido escalar posiciones hasta un lugar privilegiado en la lista de personas a las que adeudo cariño. Gracias a vosotros también, Quique y Paula, por alimentar mi ilusión y escogerme como amiga.

Gracias a mis compañeros de trabajo en los dos años de prácticas que disfruté en el Servicio de Estudios del Banco de España por los ánimos que me habéis dado para seguir adelante con la tesis. Además de ser grandes profesionales, sois excelentes personas. A Paco, Lorena, Juan, Eva, Prado, Jaime, María, y Cristina, por todo lo anterior y por las risas, por los croissants a la plancha en el Panetto y los pinchos de tortilla a 80 céntimos, las comidas en Don Paco, las cenas de solsticio, y todos los momentos de complicidad que hemos compartido. A Aitor Lacuesta y a Mario Izquierdo, con quienes ha sido un placer poder trabajar y de los que tantas cosas he aprendido. Un agradecimiento especial va para Isabel Argimón por enseñarme tantas cosas, por darme a conocer la MCVL, una de mis herramientas fundamentales de trabajo en esta tesis, por tu cariño, por tu paciencia e impaciencia de los primeros tiempos, y por escucharme siempre.

Gracias al resto de mis compañeros de FEDEA, especialmente a Karla y Liliana, mis otras dos compañeras de despacho junto con Fer, de las que he recibido mucho más cariño del que creo merecer.

También deseo agradecerle a Sara de la Rica el haberme dado la oportunidad de trabajar con ella transmitiéndome su entusiasmo como investigadora de primera línea en Economía Laboral.

Dejo para el final a aquellos que siempre habéis estado y siempre estaréis a mi lado: A mis hermanas, Isabel y Mamen, por soportar estoicamente mis días de estudio y mis apuntes y libros desperdigados por todos los rincones de casa; a mi hermano Javi, por poner cara de pez y no hacerme nada de caso cuando me pongo pesada y, sobre todo, a mis padres que me habéis dado todo lo que soy.

Por último, un agradecimiento muy especial para Jose. Gracias por esperar pacientemente y, sobre todo, por estar siempre a mi lado y hacerme feliz.

THREE ESSAYS ON THE RELATIVE OUTCOMES OF MINORITY GROUPS IN THE SPANISH LABOUR MARKET

Introducción

Esta tesis se compone de tres capítulos donde se analiza el comportamiento de tres grupos de trabajadores, generalmente calificados como grupos minoritarios, pertenecientes al mercado laboral en España. Se trata de los inmigrantes, las mujeres y los trabajadores con edades próximas a la jubilación. Nos centramos en analizar temas relacionados con el flujo de rentas que reciben dichos trabajadores así como en sus trayectorias laborales. Para ello, se utilizan diferentes enfoques empíricos aplicados a datos procedentes de las únicas fuentes de datos en España que proporcionan información concerniente a los ingresos de los trabajadores: Muestra Continua de Vidas Laborales (MCVL), en el primer y tercer capítulo, y la Encuesta de Estructura Salarial (EES) en el segundo capítulo.

Los dos primeros capítulos se centran en el análisis de las diferencias salariales de dos grupos de trabajadores respecto a los que existe una amplia literatura económica: los inmigrantes (frente a los nativos) y las mujeres (frente a los hombres). Sin embargo, en ellos se adopta una perspectiva novedosa, analizándose elementos hasta ahora inexplorados, al menos, en la literatura dedicada a estos temas en el contexto de la economía española.

En concreto, en el Capítulo 1, titulado "*Asimilación salarial de los inmigrantes en España, un análisis longitudinal*" (conjunto con Mario Izquierdo y Aitor Lacuesta, y publicado en *Labour Economics*, vol 16, 2009) se analiza, el proceso de asimilación salarial de los inmigrantes que han llegado a nuestro país en los últimos 25 años, utilizando para ello datos de carácter longitudinal recientemente disponibles. Estos datos, permiten seguir a los inmigrantes que permanecen en España durante un determinado periodo de tiempo desde el inicio de su carrera laboral en nuestro país, lo que hace posible comparar sus patrones de evolución salarial con los de trabajadores españoles homólogos en cuanto a características socio-económicas observables.

En dicho capítulo, además de realizar la primera cuantificación existente en España del ritmo de asimilación salarial de los inmigrantes que han llegado a nuestro país durante la última década, se exploran algunos factores explicativos de dicho proceso de asimilación. En concreto, se cuantifica la importancia que tiene la movilidad laboral (espacial, sectorial y de empresas) de los inmigrantes en sus patrones de evolución salarial, mostrando sus efectos diferenciales por nacionalidades.

En el Capítulo 2, titulado "*Remuneración variable por rendimiento y Diferencial Salarial de Género en España*"(conjunto con Sara de la Rica y Juan J. Dolado), se usa la información detallada referida a los diferentes componentes que conforman el salario total de los individuos proporcionados por la Encuesta de Estructura Salarial en su ola de 2006. El objetivo de este capítulo es analizar qué papel juegan los pagos o remuneraciones variables en función del rendimiento del trabajador a la hora de explicar el diferencial salarial de género.

Este componente variable no anticipado es ofertado por cerca del 27% de las empresas consideradas en la Encuesta de Estructura Salarial 2006, afectando a un 23% de los trabajadores encuestados en dicha fuente de datos salariales. En concreto, se cuantifica la posible existencia de diferencias por género en el componente de remuneración por rendimiento (RR) y en qué medida estas diferencias contribuyen al diferencial total de género existente. Dado que la proporción que representa la remuneración por rendimiento sobre el salario total es pequeña y el porcentaje de personas que lo reciben también lo es, las diferencias en el componente del salario asociado al rendimiento tan sólo explican un 7% (en media) del diferencial salarial total para el conjunto de trabajadores de la muestra y en torno a un 18% (en media) para el conjunto de trabajadores que reciben remuneración por rendimiento.

En la segunda parte del capítulo se identifican algunas hipótesis útiles para explicar la existencia de diferencias de género en este componente del salario que, en teoría, debería determinarse de manera más competitiva que otros componentes fijados en convenios colectivos (salario base, remuneración de horas extraordinarias, etc.). A la vez, se intenta identificar cuál es, a la vista de los datos, la teoría más plausible para explicar diferenciales de género en este tipo de retribución salarial :(i) auto-segregación de las mujeres en empleos de en los que no se remunera por rendimiento, (ii) menor esfuerzo de las mujeres en el trabajo como consecuencia de su mayor implicación en las tareas del hogar, lo que se refleja en menor RR para las mujeres, (iii) aspectos de carácter monopsonístico que llevan a las empresas que remuneran por rendimiento a pagar menor RR a las mujeres que a los hombres de iguales características, basándose en la creencia de que las mujeres tienen una menor movilidad. La evidencia ofrecida por la información de la EES 2006 respalda la tercera hipótesis.

En el tercer capítulo, titulado "*Evaluando los incentivos para la jubilación anticipada en España*" (sin coautores), se analiza el impacto de las reglas establecidas por la Seguridad Social en materia de pensiones sobre las decisiones de los individuos en las edades próximas a su jubilación. Siguiendo la metodología de modelos de duración propuesta por Grubber y Wise (2004), se analiza la decisión de jubilación como una decisión dinámica discreta. Para ello, en base a la información contenida en la MCVL, se construyen las variables tradicionales de incentivos utilizadas en la literatura al respecto: riqueza de la seguridad social, o flujo descontado de rentas que se recibirán en forma de pensión de jubilación desde el momento de retiro hasta el final de la vida del individuo (Social Security Wealth, SSW), valor máximo de ese flujo de rentas en función del momento de jubilación elegido (Peak Value, PV) y tasa de acumulación (Accrual Rate; AR), que mide el incremento en valor descontado del flujo de rentas que recibirá el trabajador por retrasar un año la edad de jubilación. En la primera parte del capítulo, se estima un modelo de duración con el fin de determinar si bien las características personales, o bien los incentivos generados por el sistema de Seguridad Social, son los elementos que guían las decisiones de los individuos en lo referente a su jubilación.

En la segunda parte, teniendo en cuenta el resultado de que los incentivos generados por el sistema de Seguridad Social juegan un papel relevante en las decisiones de los individuos en las edades próximas a su jubilación, se plantean diferentes simulaciones relativas a los efectos sobre la edad de jubilación de tres cambios regulatorios alternativos en la legislación en materia de pensiones. En concreto, se analizan:

(i) Una reducción en la cuantía de las pensiones percibidas por los individuos a todas las edades, de forma que como máximo se reciba un 96% de la base reguladora, implicando una reducción en la carga financiera de la Seguridad Social,

(ii) Un incremento en el número de años de cotización exigidos para tener acceso a la pensión de jubilación contributiva, de 15 a 18 años, así como incremento de las bonificaciones si se alcanzan más de 35 años cotizados en el momento de la jubilación,

(iii) Un aumento en las bonificaciones de cada año adicional que el individuo permanezca trabajando una vez cumplidos los 65 años de edad.

Los tres cambios propuestos dan como resultado un leve incremento en la edad de jubilación y, consecuentemente, una reducción en la probabilidad de jubilación en cada una de las edades consideradas.

Introduction

This dissertation consists of three chapters which examine relevant issues related to the outcomes of three minority groups (immigrants, women and elderly workers) in the Spanish labor market. Throughout the dissertation, several features related to wage profiles and individuals' labor careers are analyzed from different economic perspectives. To do so, I use the only available sources in Spain of wage micro data, namely, the Continuous Sample of Working Histories (MCVL), in the first and third chapter, and the Earnings Structure Survey (EES) in the second chapter.

In the first two chapters, I analyze issues related to wage-gaps concerning immigrants (relative to natives) and women (relative to men. Both issues have been widely addressed in the labor economics literature. However, a novel perspective is adopted here to analyze some specific facts that had been unexplored to date, at least in the literature devoted to these issues in Spain.

In particular, in Chapter 1, entitled "*Wage Assimilation of immigrants in Spain, a longitudinal analysis*" (coauthored with Mario Izquierdo and Aitor Lacuesta, already published in *Labour Economics*, vol. 16, 2009) the wage assimilation process of immigrants that arrived to Spain over the past decades is analyzed, using a longitudinal data source recently available, MCVL. These data allow us to track those immigrants that have remained for a sufficiently long period of time in this country of destination. In this way, the assimilation wage patterns of immigrants are compared with those of natives

Moreover, I explore some potentially relevant factors that underlie the assimilation process. Specifically, I focus on the role of immigrants' labor mobility (across regions, sector and firms) as well as on their wage evolution, showing differential effects by nationality.

In Chapter 2, entitled "Pay *for Performance and Gender Wage Differential in Spain*" (coauthored with Sara de la Rica and Juan J. Dolado), I make use of detailed information about the individual wage components that make up total wages, available from the EES 2006. The goal of this chapter is to analyze the role played by the pay performance (PP) component of wages on the total gender wage gap observed in Spain. PP is a variable component of the total wage, determined by workers´ performance, which is offered by about 27% of companies considered in the EES 2006, applying to 23% of workers in that survey. Our focus lies on measuring whether there are gender pay differences in this variable wage component and to what extent these differences contribute to the overall gender gap in Spain. Since the share of PP over total wage and the proportion of workers receiving PP are small, gender differences in this wage component only explain 7% (on average) of the overall gender wage gap in the sample and around 18% (on average) of the gender wage gap of those workers who receive PP.

In the second part of the chapter, I test several hypotheses which could account for gender differences in PP, considering that, at least in theory, PP is bound to be determined in a more competitive fashion that other non-variable pay components (base wage, compensation for overtime work, etc.) : (i) self-segregation of women into jobs which do not entail PP schemes, (ii) women´s lower effort at marketplace due to greater involvement than men in housework and other household responsibilities, and (iii) monopsonistic features that lead firms to pay women less than men with the same observable characteristics, based on firms ´ beliefs about women´s lower job mobility. Evidence provides higher support to hypothesis (iii) than to (i) and (ii).

In the third chapter, entitled "*Evaluating the incentives to early retirement in Spain*" (single authored) , I examine the effect of Social Security rules on the retirement decisions of individuals whose age is close to legal retirement age. I followed the hazard model methodology proposed by Gruber and Wise (2004), where the retirement decisions are modeled as a dynamic discrete option, building on the information contained in the MCVL.

Incentive variables that are traditionally used in the literature -such as Social Security wealth, (discounted flow of revenues to be received in the form of pension from retirement, SSW), Peak Value (that compares this year's social security wealth to the maximum social security wealth that could be attained in the future; PV) and Accrual rate (the discounted increase in SSW from postponing retirement one year, AR)- are computed to estimate a duration model with the aim of determining whether individual characteristics or incentives generated by Social Security retirement rules are the key factors guiding retirement decisions in Spain.

In the second part of the chapter, given that incentives generated by the Social Security rules play an important role in explaining individual's retirement decisions, I simulate the effects of three alternative policy proposals:

(I) An overall reduction in the amount of the pension received at all ages that caps the pension to 96% of the regulatory base, therefore preventing individuals from receiving a pension equivalent to 100% of the regulatory base at any retirement age.

(II) An increase in the length of the contributions period required to claim a pension (18 instead of 15 years) and an additional increase in the pension received at the age of 66, if the worker has accumulated more than 35 contributory years over her/his professional career.

(III) Higher retirement benefits for each additional year beyond 65 that an individual remains working.

Under these three alternative scenarios, our basic finding is that there is a slight increase in the retirement age and, as a result, a reduction in the probability of retirement at each of the considered ages.

CHAPTER 1

ASSIMILATION OF IMMIGRANTS IN SPAIN: A LONGITUDINAL ANALYSIS *

Mario Izquierdo *; Aitor Lacuesta * & Raquel Vegas**

(*) Banco de España

(**) FEDEA & Universidad Carlos III

ABSTRACT

In this paper we use a panel to analyze the earnings assimilation of migrants in Spain. We show that immigrants reduce the wage gap relative to natives by 15pp during the first 5-6 years after arrival, but the earnings differential does not disappear completely. Earnings assimilation is not homogeneous across different nationalities, being faster for South-American and European (new accession countries to EU) immigrants compared to Africans. Finally, we show that human capital gains within the firm as opposed to higher mobility contribute the most to their assimilation process.

JEL codes: J31, J61

Keywords: *Immigration, Assimilation, Longitudinal data, Selection, Human Capital.*

(*) The authors would like to thank the contribution of the participants at the seminars at the Banco de España, INSIDE and EALE. The panel is publicly available in the website of the Social Security. We are in debt with Almudena Duran who solved all our questions regarding the data set. Moreover, we thank the National Institute of Statistics that provided the data on non-renewals. The opinions and views expressed in this paper do not necessarily reflect those of the Banco de España.

1. Introduction

Assimilation of immigrants in the host country is a broad concept. Some people emphasise the way immigrants adopt particular values and traditions of the host country with the underlying idea of keeping a cohesive society. Others are interested in how immigrants change their behaviour in terms of fertility or consumption as they spend more time abroad. A common feature of all these different ideas of behavioural adaptation is the difficulty of obtaining measures to analyse in quantitative terms the assimilation pattern of migrants. However, a much easier concept to evaluate is earnings differentials over time. It is a fact that the majority of immigrants face a wage gap differential relative to observationally equivalent natives upon arrival; however, this wage gap is seen to decrease over the years following migrants' arrival.

The literature on assimilation of wage earnings starts with Chiswick (1978), who found an assimilation rate of around 2% per year in the United States. This means that an initial wage gap of 30% completely disappears after 15 years' residence. Despite differences in quantitative terms, some degree of earnings assimilation has been also observed for many other countries.[1] However, the robustness of the previous empirical result was criticised because of the use of a single cross-section. In a single cross-section, migrants with different labour market experiences in the country of destination belonged to different entry cohorts. If there were a decrease in the quality of migrants that entered the country over time, as occurred in the United States (Borjas, 1999), the wage growth estimated would be an upward biased measure of the actual growth. This problem was traditionally solved by using repeated cross-sections and following the experiences of different individuals belonging to the same entry cohort (LaLonde and Topel, 1992). Despite solving the previous shortcoming, the use of synthetic cohorts was not free of problems if there was selective emigration. That is the reason why nowadays most of the studies on assimilation are using longitudinal data (Hu, 2000; Lubotsky, 2007; Constant and Massey (2002); or Edin et al., 2000). Longitudinal data track the wage growth of individuals who belonged to a particular entry cohort and stayed for a certain number of years in the country. In a sense, with longitudinal data one empirically estimates the assimilation profile of foreigners who stayed in the country for a certain period of time.[2] Once this type of information was considered, the assimilation rate slightly decreased in the United States (Lubotsky, 2007) and Sweden (Edin et al., 2000) to 1.5pp per year. Indeed, in Sweden there is no perfect convergence between migrants and natives. On the other hand, Constant and Massey (2002) did not find any change in their conclusions regarding what they found using a single cross-section for the case of "guest workers" in Germany.

[1] Adsera and Chiswick (2007) present evidence for most European countries. Evidence from Canada and Australia is found in Antecol et al. (2006), from Norway in Longva and Raaum (2001), from the Netherlands in Hartog and Winkelmann (2003), from Sweden in Bevelander and Nielsen (2000), and from the UK in Wheatly Price (1999).

[2] Notice that this is not an estimation of the assimilation profile of all foreigners who entered in a particular year.

We contribute to this literature by analysing the case of immigrant's assimilation in Spain. Notice that immigration in Spain has become a matter of huge relevance in recent years. The robust annual inflow of immigrants, on average more than 500,000 since 2000, pushed up the percentage of immigrants in the Spanish population from 2% in 2000 to 11% 6 years later. Moreover, Spain's case is interesting because it has particularly rigid labour market institutions compared to the US, Sweden or Germany. [3]

Analysis of the assimilation process of immigrants in Spain has been limited due to the lack of appropriate data for earnings. Most of the research was devoted to employment rates and the characteristics of the job held by migrants since their arrival. In particular, Amuedo Dorantes and de la Rica (2007) used the labour force survey to find that immigrants were increasing their employability very rapidly. If upon arrival they had a very high unemployment rate relative to natives, after 5 years this difference had practically disappeared. They also observe differences in the assimilation pattern by nationality. Indeed, South-Americans and immigrants from the EU Enlargement perform much better in terms of assimilation. Fernández and Ortega (2008) confirmed the results found in the foregoing paper using repeated cross-sections from the Labour Force Survey. However, they also found that immigrants did not improve their situation in terms of labour conditions (incidence of fixed-term contracts or over-education) relative to natives.

The recent availability of databases with information on immigrant's wages in Spain has enabled estimates of the wage differential between immigrants and natives to be made. Using the SES-2002, Simón et al. (2008) showed that the wage differential is about 30%. Using the NIS, Sanromá et al. (2008) analysed the difference in returns to human capital acquired abroad and in Spain. Their results tend to find that returns to human capital from abroad are very low. Indeed, this result is capital in the literature of assimilation. The fact that the portability of skills is low across countries explains why immigrants on arrival face an artificial depreciation of their human capital stock. In that case, a typical human capital model (Ben Porath, 1967) would predict a large incentive to invest in human capital upon arrival, which will consequently generate a steeper wage earnings profile over the migrants' life cycle compared with otherwise identical native workers. To the best of our knowledge, our paper is the first to conduct an analysis of the assimilation process of immigrants in Spain using a longitudinal dataset. As in other studies, we will distinguish between migrants of different nationalities.

On top of reporting some findings for the assimilation pattern of Spanish immigrants, the richness of the data will help us to explore whether the higher wage growth of immigrants occurs within the initial firm or whether it is due to changes of firm and province. Immigrants are usually found in more prosperous cities and growing sectors (Amuedo Dorantes and de la Rica, 2005; Banco de España, 2007) give some evidence in this regard for Spain. This is

[3] According to Antecol et al. (2006), a labour market with high EPL and with institutions that encourage real wage rigidities might produce a more rapid assimilation pattern for migrants in terms of unemployment than in terms of wages.

usually attributed to the fact that they are more mobile than natives, which enables them to take all possible opportunities in terms of earnings across firms and regions. Indeed, in a regulated labour market such as Spain's, there are many institutions that can contribute to non-market clearing wage differentials across regions or sectors (Izquierdo and Lacuesta, 2005). Therefore, the initial wage gap between migrants and non-migrants might disappear insofar as migrants reallocate themselves into jobs that offer better economic opportunities. If this were the case, assimilation would be explained outside a human capital framework, providing a completely different set of policy implications.

2 Data

Since 2004 a new source of information about the Spanish labour market is available for researchers: Continuous Sample of Working Histories (CSWH "Muestra Continua de Vidas Laborales", in Spanish[4]). This database provides a unique dataset with very rich information about labour market histories coming from Social Security Administrative records. It contains a representative sample of those individuals who have had an affiliation in the Social Security during the reference year. In this paper we use the information from the sample 2005B which contains information about 1.142.118 individuals (92% of those are natives), what amounts a 4% of the total population who had an affiliation in the Social Security during 2005.

The information contained in the CSWH allows to make a complete characterization of each individual in the sample, since it reports personal characteristics -such as nationality, date and country of birth (province if Spanish), gender, place of residence at the first time the individual entered into the social security system - some additional information about the composition of her household and, finally, labour market variables. Moreover, and this is the main distinctive feature of the CSWH, this information about the labour market career is available for the whole labour market career of the worker since her first entry in the Social Security records, including unemployment episodes.

Concerning employment periods, it gives information related to the type of contract, seniority in the firm, sector of activity, type of working day -part time vs. full time-, and wages (although monthly capped). Concerning unemployment, it gives information about the amount of benefits received, the length of the unemployment period receiving and not receiving benefits, and the type of unemployment benefits received. CSWH also contain relevant information about sick leaves and other contributed benefits received by individuals (retirement, disability, widows and orphans benefits).

When using the longitudinal aspect of the dataset, it is certainly important to note that the CSWH has been designed to be representative of the labour market in the current year. That is, the sample of workers in 2005 is a random sample of workers affiliated to the Social Security in 2005. However, for previous years if the exit of the labour market is not random, the sample loses representativeness, as we only have information of workers staying in the labour market in 2005. This is, of course, more relevant as we move far enough to the past and, for instance, for women entering and exiting from the labour market.

[4] Social Security Records is complemented with information from fiscal registers and the Municipal register of Inhabitants. A more detailed description about CSWH could be found in Durán (2007). In a recent article, García Perez (2007), the pros and cons of this dataset are described, particularly emphasizing its comparative advantages for analysing job mobility.

Ramos Muñoz (2007) showed that the CSWH is a good representation of the Spanish labour market. He compared aggregate results from this database with those coming from Labour Force Survey (LFS) and found very similar labour market composition across different variables in the current year. Main differences, although quite small, were found for youth, females and foreigners which can be related to employment relationships in the informal sector not registered in the Social Security records. That is the reason why our results only apply to the assimilation of legal immigrants into the Spanish labour market.

With respect to the sample selection we perform for our analysis, we restrict the original sample to males between 25 and 54 years old in order to avoid, as previously mentioned, the sample selection bias that is present in CSWH for collectives with lower participation rates. Since there is no earnings information before 1979 we keep individuals who start their labour career after that date. Moreover, we keep track of earnings only in the case of being employed in the Social Security General Scheme. That is, we cannot take into account earnings obtained as self-employed since in this case, the information available about contributions bases may not be closely associated to wages. We also drop unemployment insurance earnings because we want to be as close to the actual productivity of the worker.

In order to compute the days worked during a month or a year we consider all contracts in a particular month/year for each individual. Only full-time jobs are considered since information on part-time jobs does no allow computing hours worked.[5] A migrant is a person who has no Spanish nationality and we restrict the sample of migrants to be not members of the EU-15 since immigration from those countries has been traditionally very different to the immigration coming from the rest of the world (with a high share of retirees). Ideally we would have liked to do the project with the place of birth as a better approximation to the origin of migrants. The reason is that nationalization processes do not affect the country of birth. However, information about the place of birth is of poor quality, in terms of a high percentage of missing values[6] and the impact of nationalization processes in our results should be very small[7].

As for the definition of the main variables, we perform the analysis using two different measures of daily wages. Firstly, daily wages are computed as the

[5] In that sense, assimilation of earnings might come from working more hours instead of earning more per hour. There is some evidence that temporary migrants earn less per unit of time but working more hours (Dustmann and Schmidt, 2001). This could be a problem with the interpretation of cross section or synthetic cohort analysis but it is a less important problem for longitudinal data because we are only considering stayers. Nevertheless, it could be a problem if some stayers did plan to do a temporary migration at arrival.

[6] In particular, country of birth is missing for 36% of foreign born individuals. As we will document later, using this limited information instead of citizenship do not affect the aggregate results.

[7] These processes take, usually, a long period of time due to a quite complicated procedure. In principle someone needs 10 years of residence to apply for Spanish citizenship. In the case of South-Americans one could apply pretty rapidly, usually 2 years, but process of naturalization takes many more years. Thus, our results could only be affected by nationalizations in case of old entry cohorts in Spain and, as we will document later, assimilation process is estimated to be quite similar across different entry cohorts.

ratio between monthly earnings and the days worked in a particular month. However, since migrants are expected to rotate between employment and unemployment much more than natives, we might be capturing a self selected sample of immigrants (only those who are employed in a particular month). In order to solve this problem we check the robustness of the findings to a different computation of daily earnings: the ratio between the sum of the current year earnings and the days worked during the year. Both measures will lead to very similar results.

Experience in the Spanish labour market is computed sequentially from the moment the worker enters the sample regardless whether he worked in all years. Therefore, experience should be considered as years after the first entry in the labour market[8]. Many migrants entered Spain with some experience from abroad. However, we do not have information on that issue. In order to have information on experience abroad we estimate potential experience abroad by removing from the age of entry in Spain, the potential age of entry in the labour market in the origin country. Potential age of entry in the labour market in the origin country is 16 if the person has lower education than a university degree and 22 if the person received a university degree. On the other hand, the age of entry in Spain is the age of the person at the moment of the first contribution in the social security.

One evident problem with this estimation of potential experience is that our measure of education comes from Administrative registers in municipalities ("Padrón" in Spanish) and as it was commented above it does not appropriately depict actual education, and is not available for every one[9]. Many people present a lower education in the Padrón compared to their real level. This is the case because the administrative data is not updated unless someone changes residence into a different municipality. This problem should be smaller for immigrants since most of them arrive in Spain to work.[10]

An additional problem is that the age of entry into Spain does not necessarily depict the real age of entry, since many migrants, as it will be shown later, enter illegally into the country or work in the informal sector even being legal. Lubotsky (2007), who have alternative measures for the age of entry, showed that different measures affect slightly the quantitative results without affecting them qualitatively. Since we do not have alternative measures we need to stick to the abovementioned concept of year of entry. On this regard, the existence of several regularization processes in Spain is problematic. For an illegal migrant that is legalized, years since migration is an upward biased measure of the real years spent in Spain and the number of years abroad is a downward measure of the experience in the home country. OECD (2007) described all the processes of

[8] Since the study is restricted to males between 25 and 54 years old we do not think that exit and re-entry or long unemployment is an important issue.

[9] We consider as low educated those who do not report education.

[10] Even in that case, people might decide not to update the information, since could be inconvenient. For instance, being elected as a member of a public jury or of a polling station depends on the educational attainment.

regularization in the recent past[11]. In order to analyze whether our results are affected by those regularizations, we repeat all the analysis dropping from our sample all individuals who report to have had their first labour experience in those particular years.

[11] In 1985-86 there were 38181 legalized, in 1991 (110100), in 1996 (21300), in 2000 (163900), in 2001 (216400) and in 2005 (548700). The appendix provides some additional information regarding these processes.

3 Empirical strategy

Spaniards enter into the labour force with some specific skills, depending on their gender, education and birth cohort.[12] From that moment on, the wage is increased year by year by a function $\delta^{n}{}_{k}$ due to a process of on the job learning. The individual wage is finally determined by a macroeconomic shock μ_t and an idiosyncratic shock ε_{it}:

$$W^{n}{}_{it} = e^{(\mu_t+\varepsilon_{it})} W^{n}{}_{0} \prod_{k=1}^{t} \left(1+\delta^{n}{}_{k}\right)$$

$$\ln W^{n}{}_{it} = \ln W^{n}{}_{0} + \sum_{k=1}^{t} \delta^{n}{}_{k} = f^{n}(educ, birth\_cohort) + \sum_{k=1}^{t} \delta^{n}{}_{k} + \mu_t + \varepsilon_{it} \quad (1)$$

On the other hand, immigrants arrive to Spain with certain labour market experience in addition to their initial particular skills. The experience acquired abroad has certain value in the Spanish labour market say $\lambda^{I}{}_{k}$. Following Sanromá et al. (2008), we would expect some depreciation of the skills that migrants bring to the host country therefore we would expect $\lambda^{I}{}_{k} < \delta_{k}{}^{n}$. And once in the country, immigrants start to increase their human capital at a rate $\delta^{I}{}_{k}$ depending on their investment in human capital. Given that the individual enters at the age of ta, the wage is given by:

$$W^{I}{}_{it} = e^{(\mu_t+\varepsilon_{it})} W^{I}{}_{0} \prod_{k=1}^{t^a} \left(1+\lambda^{I}{}_{k}\right) \prod_{s=t^a}^{t} \left(1+\delta^{I}{}_{s}\right)$$

$$\ln W^{I}{}_{it} = \ln W^{I}{}_{0} + \sum_{k=1}^{t^a} \lambda^{I}{}_{k} + \sum_{s=t^a}^{t} \delta^{I}{}_{s} = f^{I}(nation, entry\_cohort) + \sum_{k=1}^{t^a} \lambda^{I}{}_{k} + \sum_{s=t^a}^{t} \delta^{I}{}_{k} + \mu_t + \varepsilon_{it} \quad (2)$$

Let's define an indicator of migrant status I. Pooling (1) and (2) we have a model of earnings for the pool of workers in this economy:

$$\ln W_{it} = \ln W^{N}{}_{0} + \sum_{k=1}^{t} \delta^{N}{}_{k} + \left(\ln W^{I}{}_{0} - \ln W^{N}{}_{0}\right) I + \sum_{k=1}^{t^a} \left(\lambda^{I}{}_{k} - \delta^{N}{}_{k}\right) I_i + \sum_{s=1}^{t} \left(\delta^{I}{}_{s} - \delta^{N}{}_{s}\right) I_i + \mu_t + \varepsilon_{it} \quad (3)$$

Earnings assimilation is the process by which an immigrant with the same labour market experience than a native has a higher wage growth. Therefore:

$$\delta^{I}{}_{s} - \delta^{N}{}_{s} > 0$$

Eventually, this higher growth wage could lead to an equal wage level for migrants and natives. The assimilation pattern of the previous equation could be estimated once we define a set of variables for experience in the host labour market, years since migration and years abroad.

$$\ln W_{it} = a_0 + \sum_{k=1}^{T} a_k 1(\exp_{it} = k) + \beta_0 I + \sum_{k=1}^{T-1} \gamma_k I_i 1(\exp\_abroad_{it} = k) + \sum_{k=1}^{T} \rho_k I_i 1(ysm_{it} = k) + \mu_t + \varepsilon_{it} \quad (4)$$

$$\text{Where } \rho_k = \delta^I{}_1 \quad \delta^N{}_1$$

The literature posed some problems associated to the estimation of Eq. (4) using a single cross-section. In a cross-section, migrants with different years of labour market experience have entered Spain at different points in time. Therefore, it is impossible to distinguish the wage growth associated with labour market experience from changes in the entry cohort quality (changes in $\ln WN_0$). Provided that there was a decrease in the quality of migrants that enter the country over time, as it happened in the United States (Borjas, 1999), the wage growth over the life cycle in the cross section is an upward biased measure of the actual average wage growth. This problem has been traditionally solved by using several cross-sections allowing us to follow the history of different individuals belonging to the same entry cohort.

Even in the case that the quality of migrants was the same over year of entry, we might face an additional estimation problem derived from selective emigration. Let's assume that migrants who decide to go back home after a while are those who perform the worse. In that case the actual wage growth would also be over-estimated in the cross section because those who present higher levels of experience are the best migrants of their cohort of entry. This shortcoming cannot be solved by using repeated cross sections. However, using longitudinal data, we observe information on wages in a retrospective way for every individual who enter in a particular year and have stayed in the country until 2005. This fact allows overcoming both problems. Notice that longitudinal data only identifies the assimilation profile of stayers. Finally, in order to incorporate time effects we include the NAIRU (HP filter on the original unemployment).[13]

[12] Including birth cohort is important given the notable measurement error in the variable educational attainment.

[13] Notice that we could not include experience, birth cohort and time effects at the same time. In that case we restrict macro shocks to be identical to the NAIRU. We chose the NAIRU since it is a sensible variable that might affect the power of unions in a bargaining setting and in principle, since we are talking about legal migrants, unions should affect the same way to natives and immigrants. Restricting time effects to certain macroeconomic variables has been widely used in the literature; see Beaudry and Lemieux (1999) for an example.

4 Empirical results for assimilation

Let's start by analysing the assimilation of all migrants coming from non-EU-15 countries. In order to have a visual description of the assimilation process in Spain, Table 1 shows raw wage differentials between migrants coming from non-EU15 countries and Spaniards as a function of the time spent in Spain by migrants. As a control group we choose a generation of Spaniards with a relatively long labour market career and a comparable number of years of potential experience.[14] The wages shown in the table are median wages. The first three columns show that initially immigrants face a wage penalty of 45%. However, over time this penalty decreases to around 20%.

Table 1: Wage differentials of immigrants by years since migration

	Spaniards	Non EU-15			
	1979-1982	1979-1982		1991-1995	2001-2005
	Euros/day	Euros/day	Differential	Differential	Differential
t=0	42.07	23.26	-44.69%	-30.06%	-20.53%
t=1	43.60	26.12	-40.09%	-28.87%	-19.93%
t=2	44.51	28.14	-36.78%	-30.23%	-18.48%
t=3	44.97	27.52	-38.79%	-29.78%	-17.84%
t=4	45.06	27.89	-38.10%	-28.33%	-15.81%
t=5	45.86	32.51	-29.11%	-26.90%	
t=6	46.61	33.74	-27.60%		
t=7	47.50	33.64	-29.18%		
t=8	48.51	38.22	-21.21%		
t=9	49.41	39.46	-20.13%		
t=10	49.94	40.94	-18.02%		

Source: CSWH 2005.

Note: Spaniards are constrained to the generation that entered the Social Security in 1979-1982 and t=0 corresponds to the wage after 12 years of labour market experience (27 years of age potentially). Each column of immigrants corresponds to a different year of entry in the social security and t=0 corresponds to the wage the first year in Spain.

This reduction of the wage differentials is also observed for other more recent entry cohorts of migrants. Notice that over time, the initial wage differential has decreased. This means that immigrants have experienced a quality upgrading over time.[15] As a consequence, the assimilation profile in the single cross section should be downward biased because of the increase in quality by migrant

[14] Since migrants entered in Spain between the age of 25 and 30, we estimate that they enter with a potential experience around 12 years, Therefore a comparable t=0 for a Spanish worker is an individual with 12 years of experience. In order to have long enough labour careers for Spaniards with 12 years of experience in table 1, the control group was chosen to have entered the Social Security between 1979 and 1982.

[15] The quality upgrading of more recent cohorts of Spaniards has not been as important as the immigrant upgrading. This is clear when plotting the life cycle profile of different cohorts of entry into the Social Security for Spaniards. Ideally we would have liked to use that information in constructing table 1 but we do not have long enough careers of Spaniards who entered in the 90's onwards.

cohort. Below, we will relate this increase in quality with the composition of the immigration cohorts and we will analyse how immigrants with different countries of origin follow different assimilation patterns.

In order to properly check whether there is assimilation on earnings we need to carefully estimate Eq. (4). Since the observed value of earnings is top-coded and the censored part is around 20% in the whole sampling period we will use median regressions for the dependent variable, being $\overline{\ln W_{it}}$ the salary cap:

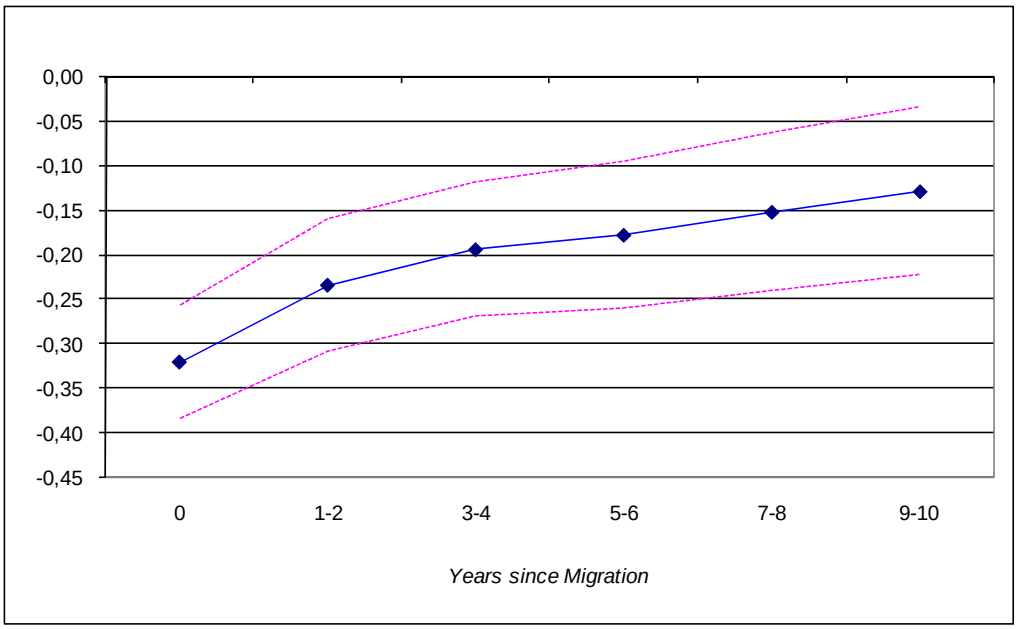
$$\ln W_{it}^{*} = \min\left(\ln W_{it}, \overline{\ln W_{it}}\right)$$

As in Lubotsky (2007) we use Powell (1984) semi-parametric censored least absolute deviation. We compute the standard deviation with a sandwich estimator (Koenker and Basset, 1978). In Table 2 we show the results for the regression pooling natives and immigrants coming from countries outside the EU-15. In the first three columns of that table the analysis is done with daily wages measured as the ratio between monthly earnings and the days worked in a particular month. The following three columns show the same results using annual wages and the days worked in a particular year. The first column of each block shows the wage model estimated using only the information in 2005, as it would be in a cross section. The coefficients in front of the dummies "*years since migration*" express the way the wage gap decreases over time in Spain. The results do show some assimilation over the time of residence (the first 8 years of experience in Spain reduced the wage differential by 17%). On the contrary, after that date, assimilation decreases notably.

The second column of each block shows the results using the panel and adding the cohort of entry for migrants. Indeed, the importance of introducing this variable is clear since in both cases there has been an upgrading of quality from 1996 onwards. Once we control for this fact, the results regarding assimilation change notably respect to what was observed in the cross section especially for very long spells. The coefficient in front of the dummy regarding the first 2 years of experience in Spain indicates that the initial wage gap decreases 8 pp after the first 2 years. The following 2 years the wage gap decreases 4 pp additionally (reducing the wage gap in 12pp). Finally, those who have stayed in Spain for longer than 8 years present a reduction in the initial wage gap of 20 pp. From that moment on, it does not appear to be further reductions, as can be seen in Fig. 1 where we show the evolution of wage differential for recent migrants with respect to a typical native. We have carried-out the same type of estimation for different cohorts of entry in order to test whether the assimilation process has varied over time. However, we find a qualitatively similar assimilation pattern across entry cohorts. These results increase our confidence in our choice of nationality to define a migrant since, if any, the impact of nationalizations should arise in older cohorts.[16]

[16] In addition, we have checked the robustness of our results to the definition of a migrant by country of birth instead of nationality. Despite the number of observations is lower in this case, main results reported in Table 2 remain. In particular, using the same specification as in column 2 of Table 2, assimilation profile is estimated to reduce wage gap between immigrants and natives by 9.4 pp in the first 2 years in the

Fig 1. Wage differential between migrants (exp abroad<5 and arrival 1996-2000) and natives (born 1964-1975).



The third column of each block presents the same type of regression excluding those years when regularization processes have taken place since we know that for those years experience in Spain is not well captured by the labour experience by the moment of first register in the Social Security. For other periods both concepts should be closely associated. Results in terms of assimilation are similar to those in column 2 and 4. Notice however, that experience abroad, as expected, is less valued in Spain than in the previous estimations. This is evidence in favour of the notable measurement error of the variable experience abroad.

Spanish labor market (8.6 pp in Table 2) or 14.3 pp after 4 years in Spain (12.7 pp in Table 2). Only for long working careers in Spain, the use of country of origin introduces a relevant change in our results estimating a higher assimilation profile, for instance, wage gap reduces by 26.5 pp after 12 years while it is estimated at 19.3 pp in Table 2. Detailed results are available upon request.

Table 2. Wage equation estimations at percentile 50.

	Monthly contribution over total days worked in a given month			Annual contribution over total days worked in a given year		
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*
Independent variables	**Coefficients**	**Coefficients**	**Coefficients**	**Coefficients**	**Coefficients**	**Coefficients**
	Std. Err	Std. Err	Std. Err	Std. Err	Std. Err	Std. Err
Total experience	**0.107**	**0.100**	**0.101**	**0.084**	**0.102**	**0.103**
	0.003	0.001	0.000	0.004	0.001	0.001
Total experience^2 (/10)	**-0.106**	**-0.095**	**-0.096**	**-0.074**	**-0.091**	**-0.092**
	0.005	0.002	0.000	0.005	0.002	0.002
Total experience^3 (/100)	**0.047**	**0.046**	**0.047**	**0.030**	**0.043**	**0.043**
	0.003	0.001	0.000	0.003	0.001	0.001
Total experience^4 (/1000)	**-0.007**	**-0.008**	**-0.008**	**-0.005**	**-0.007**	**-0.008**
	0.000	0.000	0.000	0.001	0.000	0.000
Illiterate	**-0.711**	**-0.764**	**-0.767**	**-0.716**	**-0.752**	**-0.757**
	0.003	0.001	0.000	0.003	0.001	0.001
Primary education	**-0.645**	**-0.689**	**-0.690**	**-0.650**	**-0.678**	**-0.681**
	0.003	0.001	0.000	0.003	0.001	0.001
Secondary education	**-0.331**	**-0.338**	**-0.336**	**-0.355**	**-0.353**	**-0.352**
	0.003	0.001	0.000	0.003	0.001	0.001
Birth Cohort <=1934		**-0.014**	**-0.014**		**-0.081**	**-0.079**
		0.013	0.000		0.012	0.011
Birth Cohort 1935-1944		**0.032**	**0.032**		**-0.042**	**-0.041**
		0.004	0.000		0.004	0.003
Birth Cohort 1945-1954	**0.120**	**0.084**	**0.083**	**0.108**	**0.023**	**0.022**
	0.008	0.002	0.000	0.009	0.002	0.002
Birth Cohort 1955-1964	**0.123**	**0.041**	**0.040**	**0.113**	**-0.004**	**-0.006**
	0.003	0.001	0.000	0.003	0.001	0.001
Birth Cohort 1965-1974	**0.052**	**-0.006**	**-0.007**	**0.042**	**-0.024**	**-0.024**
	0.002	0.001	0.000	0.002	0.001	0.001
NAIRU		**-0.021**	**-0.021**		**-0.025**	**-0.025**
		0.000	0.000		0.000	0.000
Non EU-15	**-0.305**	**-0.395**	**-0.349**	**-0.282**	**-0.406**	**-0.529**
	0.022	0.019	0.000	0.023	0.018	0.039
Years since migration, 1-2	**0.047**	**0.086**	**0.103**	**0.020**	**0.058**	**0.073**
	0.010	0.005	0.000	0.011	0.005	0.008
Years since migration, 3-4	**0.121**	**0.127**	**0.142**	**0.109**	**0.104**	**0.123**
	0.008	0.006	0.000	0.009	0.006	0.009
Years since migration, 5-6	**0.167**	**0.143**	**0.161**	**0.141**	**0.115**	**0.143**
	0.011	0.009	0.000	0.012	0.009	0.012
Years since migration, 7-8	**0.179**	**0.169**	**0.178**	**0.146**	**0.154**	**0.164**
	0.021	0.012	0.000	0.021	0.012	0.014
Years since migration, 9-10	**0.145**	**0.192**	**0.206**	**0.122**	**0.197**	**0.216**
	0.027	0.015	0.000	0.028	0.015	0.018
Years since migration, 10-12	**0.091**	**0.193**	**0.207**	**0.058**	**0.182**	**0.194**
	0.051	0.018	0.000	0.052	0.017	0.020
More than 13 years since migration	**0.148**	**0.201**	**0.204**	**0.112**	**0.183**	**0.212**
	0.024	0.022	0.000	0.024	0.021	0.025
Experience abroad, from 5 to 9 years	**0.077**	**0.093**	**0.057**	**0.078**	**0.093**	**0.067**
	0.022	0.013	0.000	0.023	0.013	0.016
Experience abroad, from 10 to 14 years	**0.075**	**0.063**	**0.025**	**0.072**	**0.055**	**0.025**

	0.022	0.013	0.000	0.022	0.013	0.016
Experience abroad, from 15 to 19 years	**0.048**	**0.009**	**-0.020**	**0.041**	**-0.008**	**-0.021**
	0.022	0.014	0.000	0.023	0.013	0.016
More than 25 years of Experience abroad	**0.040**	**0.004**	**-0.023**	**0.036**	**-0.041**	**-0.055**
	0.024	0.015	0.000	0.025	0.014	0.018
Arrival 1983-1985		**-0.100**	**-0.139**		**-0.220**	**-0.085**
		0.040	0.000		0.038	0.053
Arrival 1986-1990		**-0.012**	**-0.041**		**-0.045**	**0.095**
		0.022	0.000		0.020	0.039
Arrival 1991-1995		**-0.002**	**-0.048**		**-0.025**	**0.087**
		0.013	0.000		0.013	0.035
Arrival 1996-2000		**0.080**	**0.046**		**0.106**	**0.227**
		0.013	0.000		0.012	0.036
Arrival 2001-2005		**0.058**	**0.019**		**0.090**	**0.208**
		0.013	0.000		0.012	0.036
Constant	**8.526**	**8.763**	**8.760**	**8.569**	**8.758**	**8.755**
	0.009	0.004	0.000	0.010	0.004	0.004
R^2	0.14	0.14	0.14	0.126	0.122	0.123

Dependent variable: logarithm of daily wages
(1) and *(4)*. CSWH- 2005: Cross- Section 2005
(2) and *(5)*. CSWH- 2005: Longitudinal 1980-2005
(3) and *(6)*. CSWH- 2005: Longitudinal 1980-2005 without including those migrants whose arrival was during a period of Special Immigrant Legalization

.

5 The importance of the country of origin

Underlying the increase in the relative initial wage of recent immigrants there is a change in the immigrants' country of origin. This change could be attributed to a different composition in both entries and exits of immigrants over time. In our dataset we only observe the nationality of the stock of stayers (see Table 3). In the 80's and the beginning of the 90's, the weight of immigrants coming from Africa attained a percentage around 50%. However, over time, the weight of migrants coming from South-America and from countries of the latter European Enlargement has increased notably[17]. Notice that, both South-Americans and immigrants coming from the European Enlargement have some characteristics that make them more likely to have productivity above the average of immigrants from outside the EU-15, say the same language and a higher level of education respectively. As a consequence, they present a lower initial wage differential. This would lead to the observed result of quality upgrading over time.

Table 3: Stock of foreign workers in 2005 by year of entry in the Social Security and nationality

	Africa	Enlargement	South America	Other
1983-1985	40.51%	-	15.68%	43.81%
1986-1990	48.81%	3.79%	13.97%	33.43%
1991-1995	71.50%	5.85%	7.98%	14.67%
1996-2000	59.70%	9.41%	18.45%	12.44%
2001-2005	31.98%	20.48%	39.27%	8.27%

Source: CSWH 2005

Using alternative datasets we could confirm that both entries and exits contribute to the variation of the stock of stayers over time. The Survey on Residential Variation is a dataset that incorporates all entries in Spain[18]. The official numbers of entries from the rest of the world are broken by nationality in Table 4.

It is evident that, in the beginning of the 90's the percentage of immigrants coming from African countries increased, while South-Americans decreased. In the second half of the 90's, South-Americans and foreigners coming from countries of the European Enlargement started increasing its importance. From the 2000 onwards the latter group took over. This image on the evolution of entries resembles the variations on the stock that were observed in Table 3.

[17] The 2005 figure is biased due to the regularization process that legalized 600.000 individuals (mostly Ecuadorians and Romanians).

[18] We do not have microdata from years previous to 1988.

Table 4: Inflows of foreigners by year of registration (padrón) and nationality

	Africa	Enlargement	South America	Other
1986-1990	23.69%	6.67%	55.77%	13.87%
1991-1995	38.89%	5.51%	43.48%	12.12%
1996-2000	33.76%	9.51%	49.21%	7.52%
2001-2005	19.00%	24.68%	51.57%	4.75%

Source: Residential Variation Survey
Average of yearly percentages over the period

How selective emigration affect the initial wage differential is a more difficult issue to analyse since there are not credible registers of exits in Spain. However, we could rely on different studies that positively relate exits from Spain to the GDP per capita of the country of origin. Lacuesta and Puente (2009) used data on a renewal process of the municipal registers in 2005 to determine that immigrants were more likely to come back to countries with higher GDP per capita regardless the cyclical situation in the host country and at home. This result is also confirmed by other international studies.[19] In that sense, they show that immigrants coming from the enlargement and other richer countries presented a higher emigration rate than Africans and South Americans. This positive selective emigration in terms of country of origin could also contribute to the quality upgrading observed in the data.

Once we have seen the importance of the composition of immigrants by nationality, we might wonder whether assimilation is different depending on the country of origin. Assimilation profile could, in principle, be quite different across migrants with different characteristics. Thus, in Table 5, we show regression results similar to those in Table 2 but allowing for a different coefficient in the assimilation dummies for different nationality groups. We will specifically focus on Africans, Europeans from the new accession countries to the EU and South-Americans.

These results show a quite similar assimilation pattern in the initial years in Spain, that is, wage differential reduces by around 5 pp. in the first 2 years and around 10 pp. in the first 6 years. European and South-American migrants appear to perform relatively better than Africans, although differences are not high. For long periods in Spain, however, results tend to show more differences across nationalities with a lower assimilation profile for Africans migrants. These results could be indicating the impact of some characteristics as the common languages for South-Americans and the higher educational attainment for Europeans in the assimilation process. This result is not surprising given that a lower assimilation profile in unemployment rates for immigrants coming from African countries was also found by Amuedo Dorantes and de la Rica (2007).

[19] Borjas and Bratsberg (1996) in US, Dustman (2007) in UK, Bratsberg et al. in Norway (2007) and Edin (2000) for Sweden.

Table 5: Assimilation profile by nationality and years since migration

	Africa	Enlargement	South America	Others
Independent variables	**Coefficients**	**Coefficients**	**Coefficients**	**Coefficients**
	Std. Err	Std. Err	Std. Err	Std. Err
Years since migration, 1-2	**0.047**	**0.049**	**0.056**	**0.090**
	0.010	0.047	0.008	0.017
Years since migration, 3-4	**0.074**	**0.034**	**0.119**	**0.172**
	0.011	0.049	0.009	0.018
Years since migration, 5-6	**0.081**	**0.088**	**0.100**	**0.227**
	0.013	0.060	0.017	0.024
Years since migration, 7-8	**0.094**	**0.187**	**0.095**	**0.282**
	0.015	0.070	0.030	0.029
Years since migration, 9-10	**0.133**	**0.206**	**0.167**	**0.312**
	0.018	0.077	0.040	0.035
Years since migration, 11-12	**0.125**	**0.171**	**0.248**	**0.411**
	0.021	0.084	0.051	0.041
More than 13 years since migration	**0.113**	**0.147**	**0.334**	**0.417**
	0.025	0.105	0.054	0.044

Source: CSWH- 2005: Longitudinal 1980-2005
The regression is the same as in Table 2 with interactions of the immigration variables by nationality

6 Mechanisms underlying assimilation

In the previous sections we have shown that the wage gap between immigrants and natives reduces with the time spent in Spain at a relatively faster pace while this pattern appears to be somewhat different depending on the immigrant's nationality. Most of the literature related to assimilation of earnings has assumed that behind the reduction of wage differentials there is a conventional human capital accumulation story. As time goes on, immigrants adapt their knowledge and qualifications to the host country and acquire new skills and abilities that make them more prone to benefit from their previous skills. However, additional mechanisms could also play a relevant role in the assimilation process. For instance, labour mobility in most Mediterranean countries is quite reduced given that natives are reluctant to move across regions despite the big existing disparities of earnings.[20] Of course immigrants, who have a lower social attachment to particular regions, are much more prone to benefit from those differences (Amuedo-Dorantes and de la Rica, 2005; Banco de España, 2007). Moreover, the lower mobility of natives is not limited to regional movements but it is also evident in changes of firms. A higher incidence of permanent contracts with high dismissal costs across the native population makes them less likely to be willing to change a particular job and benefit from potential better productivity matches. Therefore differences in the degree of mobility would lead to a closing earnings gap between immigrants and natives regardless the human capital accumulation story.

In table 6 we observe how Spaniards and immigrants move across regions over time. For each nationality the first column describes the distribution of the population across Spanish regions in the data. The second column describes the distribution after five years of work.21 The table shows that the distribution of Spaniards is fairly constant over time. On the other hand, there are more notable differences across immigrants. Those differences are higher if we restrict the population to particular cohorts of entry22.

Even more striking are the changes in the distribution of the immigrant population by sectors over time.23 Table 7 shows a clear tendency between immigrants to move from low earnings sectors (construction and accommodation) to better paid service sectors.

[20] Bentolila (1997) argued that mobility decreases with high unemployment levels. Moreover, Antolín and Bover (1997) have seen housing market prices to be a deterrent of mobility.

[21] In this raw data we are not following a particular immigrant cohort of entry; therefore changes in the entry pattern could affect the table. However, we have done the analysis by particular generations of entry and the results were similar although the regional and by sector distributions of migrants were much more spiky. We decided to present the aggregate ones to increase the clarity of the argument.

[22] Notice that in the table we cannot distinguish changes in the entry pattern or real changes of residents. In particular, there are very few Europeans from the accession countries that entered before the year 2000 and stayed more than 5 years.

[23] Remember that the sectors belong to the general regime; therefore people working in agriculture, household activities and public administration are not completely captured.

Table 6: Regional distribution of workers by nationality

	Spain		Africa		Enlargement		South America	
	1st year	5 years later	1st year	5 years later	1st year	5 years later	1st year	5 years later
Andalucía	*17.48*	*14*	*7.93*	*5.16*	*9.28*	*0*	*6.77*	*4.98*
Cataluña	*16.31*	*18.78*	*37.7*	*34.7*	*23.71*	*12.07*	*19.07*	*19.78*
Madrid	*21.51*	*22.95*	*15*	*16.72*	*34.02*	*46.55*	*29.94*	*41.9*
Valencia	*9.91*	*9.91*	*7.44*	*10.44*	*13.4*	*12.07*	*9.97*	*5.76*
Galicia	*5.14*	*4.71*	*0.81*	*0.33*	*1.03*	*0*	*1.69*	*1.87*
Castilla y León	*3.91*	*3.56*	*2.02*	*1.59*	*4.12*	*6.9*	*2.95*	*1.4*
País Vasco	*4.74*	*5.24*	*2.51*	*2.45*	*1.03*	*3.45*	*2.34*	*1.56*
Canarias	*4.37*	*4.45*	*5.02*	*5.82*	*0*	*0*	*6.15*	*6.07*
Castilla-La Mancha	*3.4*	*3.04*	*3.96*	*3.11*	*0*	*8.62*	*4.1*	*3.74*
Murcia	*2.98*	*2.63*	*3.24*	*4.49*	*1.03*	*1.72*	*3.96*	*5.61*
Aragón	*1.88*	*2.31*	*4.05*	*3.44*	*3.09*	*1.72*	*2.37*	*1.56*
Extremadura	*1.46*	*1.4*	*0.32*	*0.26*	*0*	*0*	*0.54*	*0.62*
Asturias	*1.8*	*1.81*	*0.32*	*0.26*	*5.15*	*3.45*	*0.47*	*0.62*
Baleares	*2.5*	*2.31*	*5.34*	*7.2*	*4.12*	*3.45*	*5.15*	*2.18*
Navarra	*1.12*	*1.28*	*1.21*	*1.39*	*0*	*0*	*2.77*	*1.25*
Cantabria	*0.87*	*1*	*0.16*	*0.07*	*0*	*0*	*1.04*	*0.78*
La Rioja	*0.4*	*0.49*	*0.81*	*1.06*	*0*	*0*	*0.68*	*0.31*
Ceuta	*0.15*	*0.07*	*1.13*	*0.33*	*0*	*0*	*0.04*	*0*
Melilla	*0.08*	*0.07*	*0.89*	*1.19*	*0*	*0*	*0*	*0*

Source: CSWH 2005

In order to assess whether the catching up to the native born earnings levels could be explained as a result of the higher degree of migrants' mobility we are going to enlarge the earnings model in section 3 controlling by job to job changes in province and firm. With this distinction, we are trying to show how much of the change in the earnings differential is due to a higher labour mobility of immigrants. Of course, human capital gains could be underlying a particular change of firm or province. However, differential wage growth that occurs within the same firm cannot be related to mobility and should be more related to human capital gains.[24]

[24] An alternative interpretation of wage changes within the initial firm is a higher degree of discrimination by the employer. At the beginning, the employer used vague information to pay the employee because of

Table 7: Distribution of workers by sector and nationality

	Spain		*Africa*		*Enlargement*		*South America*	
	1st year	5 years later	1st year	5 years later	1st year	5 years later	1st year	5 years later
Agriculture	*3.33*	*1.62*	*3.71*	*1.6*	*3.55*	*1.01*	*1.6*	*1.1*
Extractive industries	*0.27*	*0.36*	*0.16*	*0.19*	*0*	*0*	*0.19*	*0.33*
Manufacturing industry	*22.79*	*23.73*	*16.09*	*11.06*	*9.22*	*15.15*	*11.06*	*11.91*
Supply of electrical energy, gas	*0.31*	*0.57*	*0*	*0.08*	*0.71*	*0*	*0.08*	*0*
Construction	*16.33*	*17.22*	*41.88*	*37.61*	*49.65*	*42.42*	*37.61*	*32.97*
Wholesale and retail trade	*20.01*	*19.39*	*11.4*	*14.05*	*9.22*	*13.13*	*14.05*	*12.35*
Accommodation	*9.56*	*5.93*	*9.76*	*13.45*	*7.09*	*4.04*	*13.45*	*14.33*
Transport	*3.23*	*5.04*	*1.91*	*4.61*	*5.67*	*8.08*	*4.61*	*5.51*
Financial and insurance	*1.57*	*2.48*	*0.05*	*0.19*	*0.71*	*0*	*0.19*	*0.55*
Real estate	*10.29*	*10.36*	*8.18*	*11.77*	*8.51*	*9.09*	*11.77*	*12.68*
Public Administration	*5.04*	*5.31*	*2.62*	*0.62*	*0*	*0*	*0.62*	*0.55*
Education	*2.13*	*1.97*	*0.71*	*0.46*	*2.13*	*0*	*0.46*	*1.21*
Health	*1.47*	*2.38*	*0.44*	*0.89*	*0*	*0*	*0.89*	*1.21*
Arts, recreation	*3.57*	*3.54*	*2.84*	*3.31*	*3.55*	*6.06*	*3.31*	*5.07*
Household activities	*0.09*	*0.11*	*0.27*	*0.08*	*0*	*1.01*	*0.08*	*0.22*
Extraterritorial	*0*	*0.01*	*0*	*0.03*	*0*	*0*	*0.03*	*0*

Source: CSWH 2005

Let's take equation (3) and define changes in log wages:

$$\Delta \ln W_{it+1} = \delta^{N}{}_{t+1} + \left(\delta^{I}{}_{t+1} - \delta^{N}{}_{t+1}\right) I_i + \quad v_{it+1} =$$
$$= \delta^{N}{}_{t+1} + \lambda_0 I_i + \lambda_1 prov\_mov + \lambda_2 firm\_mov + u_{it} \qquad (5)$$

By applying first differences in the data, all fixed effects vanish and the wage growth differential between natives and immigrants $\left(\delta^{I}{}_{t+1} - \delta^{N}{}_{t+1}\right)$ will be explained by different propensity to regional mobility λ_1, changes of firm λ_2 and human capital accumulation that yields a payoff within the initial firm λ_0.

Doing this analysis requires defining a main occupation for each worker in at least two different points in time. The data set provides information about all the legal labour relations that the workers have every day. However, the potential multiplicity of contracts in a certain moment of time generates an

the uncertainty on his productivity. Over time, the employer might learn rapidly on the real productivity of the employee and decide a higher salary. We are not going to be able to distinguish between the two hypotheses.

added difficulty. For a particular moment, if the individual has more than one job, we choose among all of them by preferring by this order the permanent one, the longer and the one that begins earlier. We repeat the process for the following years. By doing this, we get a sequence of movements of province and firms for each individual in the sample and, it is possible to compute the probability of changes from one year to other for each specific dimension considered.

Table 8 provides the probit coefficients for the regression of changes in labour status from one year to another. The table shows that non UE-15 workers in Spain are more occupationally mobile than natives regardless the nationality.

Table 8: Yearly change of status

	Dependent variable	
	Province change	Firm change
	(1)	(2)
Independent	**Coefficients**	**Coefficients**
variables	Std. Err	Std. Err
Africa	0.0136	0.0143
	(0.0006)	(0.0014)
Enlargement	0.0144	0.0196
	(0.0029)	(0.0064)
South America	0.0170	0.0210
	(0.0006)	(0.0014)
Others	0.0147	0.0003
	(0.0011)	(0.0025)
Age	0.0003	-0.0005
	(0.0000)	(0.0000)
Illiterate	0.0025	-0.0028
	(0.0003)	(0.0007)
Primary	0.0027	0.0005
	(0.0003)	(0.0006)
Secondary	0.0016	-0.0031
	(0.0003)	(0.0006)
Constant	0.0010	0.0734
	(0.0004)	(0.0008)

Source: CSWH 2005

To inspect the influence of labour choices - changes in province and firm - on wage growth, we estimate (5) on non top coded observations (table 9). The first column identifies that immigrants from outside the EU-15 experienced a wage growth that is 3 pp higher than comparable natives. The coefficient in front of the variable years since migration identifies the fact that this higher wage growth for immigrants vanishes over time and eventually disappears in 9 years. This result is consistent with what was obtained in section 4. Column 2 controls

for inter-provincial movements and column 3 controls for firm movements. As can be seen, both types of labour mobility are positively associated with wage growth, especially the second one although we should admit that an endogeneity problem arise since job movements are, likely, caused by wage changes. In any case, we want to emphasize that even after controlling for job mobility, the assimilation process mostly explains the decrease in wage gap. That is, as can be seen in our results in Table 9, controlling for job mobility only explains 1pp of the differential wage growth[25]. We can interpret this result, see table 10, as labour mobility only accounts for 32% of the higher wage growth of immigrants. In other words, around two thirds of the assimilation profile estimated in the Spanish economy can be related to the acquisition of country-specific human capital.

Table 9: Determinants of the increase in wages

	(1)	(2)	(3)	(4)	(5)	(6)
Independent	Coeff.	Coeff.	Coeff.	Coeff.	Coeff.	Coeff.
Variables	Std. Err	Std. Err	Std. Err	Std. Err	Std. Err	Std. Err
Non EU-15	0.031	0.029	0.021			
	0.004	0.004	0.004			
Years since migration	-0.004	-0.004	-0.003	-0.003	-0.003	-0.003
	0.001	0.001	0.001	0.001	0.001	0.001
Africa				0.023	0.020	0.012
				0.005	0.005	0.005
Enlargement				0.033	0.031	0.025
				0.013	0.013	0.013
South America				0.038	0.036	0.028
				0.005	0.005	0.005
Others				0.026	0.024	0.018
				0.007	0.007	0.007
Province change		0.037	0.011		0.037	0.011
		0.002	0.003		0.002	0.003
Firm change			0.033			0.033
			0.001			0.001
Total Experience	-0.010	-0.010	-0.010	-0.010	-0.010	-0.010
	0.000	0.000	0.000	0.000	0.000	0.000
Total experience²	0.000	0.000	0.000	0.000	0.000	0.000
	0.000	0.000	0.000	0.000	0.000	0.000
Time dummies	Yes	Yes	Yes	Yes	Yes	Yes
Constant	0.055	0.055	0.050	0.055	0.055	0.051
	0.010	0.010	0.010	0.010	0.010	0.010
N	1,289,468	1,289,468	1,289,468	1,289,468	1,289,468	1,289,468
R^2	0.006	0.006	0.007	0.006	0.006	0.007

Source: CSWH 2005

[25] The coefficient in front of the immigrant dummy decreases from 3.1 to 2.1

In the following columns the same exercise is performed by country of origin in order to check whether significant differences exist. It is clear that South-Americans and immigrants from the EU-Enlargement have a relatively higher wage growth respect to Africans, again consistent with the evidence shown in table 5. Regardless the country of origin, the contribution of mobility to this differential wage growth does not achieve 50%. The contribution is highest for African immigrants suggesting that the lower assimilation profile of immigrants from Africa is not driven by a different willingness to move. Instead, difficulties in the adaptability of skills or the acquisition of new ones could be the ultimate cause.

Table 10: Contribution of mobility in the differential wage growth

	Movements across provinces	Movements across provinces and firms
Non-EU-15	7.1%	32.3%
Africa	10.5%	46.1%
Enlargement	6.1%	23.0%
South America	5.5%	26.6%
Others	9.5%	32.2%

Source: CSWH 2005 with results in table 9

7 Conclusions

In this paper we analysed earnings assimilation of immigrants in Spain using longitudinal data. Immigration in Spain has become an important matter in the recent years since the percentage of immigrants in the Spanish population increased from 2% in 2000 to 11% six years later. We used a dataset of working histories to analyse how earnings evolve over time. We show that the initial wage differential with respect to natives of the same characteristics (around 30% for recent arrival cohorts) decreases with time spent in Spain. According to our results, however, assimilation of legal immigrants is pretty rapid, with a reduction of 15pp in the wage gap during the first 5-6 years, but the differential never vanishes completely. This result is very similar to the one obtained by Lubotsky (2007) in the US and Edin et al. (2000) in Sweden despite presenting less rigid wage bargaining systems.

The use of longitudinal data to analyse the Spanish case is an improvement respect to using single or repeated cross sections because it takes into account the fact that Spain has experienced important changes in the composition quality of immigrants over time. In this respect, estimations based on a single cross section would underestimate earnings assimilation for long spells. This is the case because in recent years, the weight of South-Americans and immigrants from the European Enlargement had grown notably respect to Africans and the former nationalities present higher initial wages than the latter.

Our estimates show a relatively different assimilation pattern across nationality groups with South-Americans and EU-Enlargement immigrants showing a faster assimilation process. This is also consistent with evidence by Amuedo-Dorantes and de la Rica (2007) in terms of unemployment rates.

There are two hypotheses behind the positive wage growth differential between migrants and natives. On the one hand, the initial low level of human capital (useful in the destination country) compared to the level faced by comparable natives, justifies a high incentive to invest in human capital upon arrival. On the other, in a regulated labour market such as the Spanish, there are many institutions that exacerbate economic differences across regions or municipalities. Migrants are more prone to benefit from those differences since they have lower social attachments than natives and lower severance payments rights. In order to separate out those two mechanisms we have analyses how wage growth differential between migrants and Spaniards evolve, first unconditionally, and second controlling for labour market mobility. According to our results, most of the assimilation pattern happens within the initial firm. Following this evidence, we conclude that human capital gains are important. However, notice that there is an alternative explanation based on the possibility of a change in the behaviour of the entrepreneur. At the beginning of the labour relationship the entrepreneur might be more reluctant to pay the uncertain productivity of an unknown input. However, over time the employer could be

discovering the real quality of that input and pay it accordingly. Further research should be devoted to the distinction of the two hypotheses. In any case, mobility seems to play a limited role. This result is observed regardless the nationality of origin. The highest contribution of mobility appears in the African group. This could be interpreted as a higher difficulty of this group in adapting or acquiring new human capital.

To conclude, our results could be indicating a positive contribution of immigration to the evolution of productivity in the Spanish economy over the coming years. Large immigration flows received in the recent past has, at least partly (Banco de España, 2006; Lacuesta et al., 2009), contributed to poor productivity performance. Our paper has shown that the initial wage, or productivity, gap between natives and immigrants rapidly decreases as immigrants acquire country-specific human capital. Hence, in the near future, while lower immigrant flows are expected, the assimilation profile of those migrants already in the Spanish labour market will tend to positively contribute to productivity growth. In any case, we should also recognize that much uncertainty remains about the behaviour of immigrants in the Spanish labour market since the immigration phenomena has taken place in a especially positive economic environment.

8 References

Adsera, A., Chiswick, B., 2007. *Are there Gender and Country of Origin Differences in Immigrant Market Outcomes across European Destinations*. Journal of Population Economics; 20; 495-526.

Amuedo-Dorantes, C., de la Rica, S., 2005. *Immigrants' Responsiveness to Labour Market Conditions and It's Implications on Regional Disparities: Evidence from Spain*. IZA DP; 1557.

Amuedo-Dorantes, C., de la Rica, S., 2007. *Labour Market Assimilation of Recent Immigrants in Spain.* British Journal of Industrial Relations; 45; 257-284.

Antecol, H., Kuhn, P., Trejo, S., 2006. *Assimilation via Prices or Quantities? Sources of Immigrant Earnings Growth in Australia, Canada, and the United States.* Journal of Human Resources; 41; 821–840.

Antolín, P., Bover, O., 1997. *Regional Migration in Spain: The Effect of Personal Characteristics and of Unemployment, wage and House Price Differentials Using Pooled Cross-sections.* Oxford Bulletin of Economics and Statistics; 59; 215-235.

Banco de España, 2006. *The Growth of the Spanish Economy: Some Basic Determinants*. In: Banco de España. Annual Report 2006; 35-50.

Banco de España, 2007. *La Contribución de la Población Extranjera a la Movilidad Geográfica del Trabajo.* In: Banco de España. Economic Bulletin. July-August 2007; 53-54.

Beaudry, P., Lemieux, T., 1999. *Evolution of the female Labour Force Participation Rate in Canada, 1976-1994.* Mimeo Centre Interuniversitaire de Reserche en Analyse des Organisations.

Ben-Porath, Y., 1967. *The Production of Human Capital and the Life Cycle of Earnings.* The Journal of Political Economy; 75; 352-365.

Bentolila, S., 1997. *La inmovilidad del Trabajo en las Regiones Españolas.* Working Paper Series Banco de España; 9718.

Bevelander, P., Nielsen, H.S., 2001. *Declining Employment Success of Immigrant Males in Sweden: Observed or Unobserved Characteristics?*. Journal of Population Economics; 14; 455-472.

Borjas, G., 1999. *Heaven's Door: Immigration Policy and the American Economy*. Princeton University Press.

Borjas, G., Bratsberg, B., 1996. *Who Leaves? The Outmigration of the Foreign-Born.* The Review of Economics and Statistics; 78; 165-176.

Bratsberg, B., Raaum, O., Sorlie, K., 2007. *Foreign-born Migration to and from Norway.* In: Ozden, C., Schiff,M., International Migration Economic Development & Policy. World Bank and Pallgrave MacMillan Ed. 2007; 259-290.

Chiswick, B., 1978. *The Effect of Americanization on the Earnings of Foreign-Born Men.* Journal of Political Economy; 86; 897-922.

Constant, A., Massey, D., 2002. *Self-selection, Earnings and Out-migration: A Longitudinal Study of Immigrants to Germany.* Journal of Population Economics; 16; 631-653.

Durán, A., 2007. *La Muestra Continua de Vidas Laborales de la Seguridad Social. Revista del Ministerio de Trabajo y Asuntos Sociales;* 1; 231-240.

Dustmann, C., Schmidt, C., 2001. *The Wage Performance of Immigrant Women: Full time Jobs, Part-time Jobs and the Role of Selection.* CEPR Working Paper; 2702.

Dustmann, C., Weiss, Y., 2007. *Return Migration: Theory and Empirical Evidence from the UK.* British Journal of Industrial Relations; 45; 236-256.

Edin, P.A.; LaLonde, R., Aslund, O., 2000. *Emigration of Immigrants and Measures of Immigrant Assimilation: Evidence from Sweden.* Swedish Economic Policy Review; 7; 163-204.

Fernández, C., Ortega, O., 2008. *Labour Market Assimilation of Immigrants in Spain: Employment at the Expense of Bad Job-Matches?.* Spanish Economic Review; 10, 83-107.

García Perez, J.I., 2007. *La Muestra Continua de Vidas Laborales (MCVL): una guía de uso para el análisis de transiciones*. Revista de Economía Aplicada; XVI; 5-28.

Hartog, J., Winkelmann, R., 2003. *Comparing migrants to non-migrants: The case of Dutch Migration to New-Zealand,* Journal of Population Economics; 16; 683-705.

Hu, W-Y., 2000. *Immigrant Earnings Assimilation: Estimates from Longitudinal Data.* The American Economic Review Papers and Proceedings of the One Hundred Twelfth Annual Meeting of the American Economic Association; 90; 368-372.

Izquierdo, M.; Lacuesta, A., 2005. *Heterogeneidad en los Mercados de Trabajo Regionales.* In: Banco de España. Economic Bulletin, October 2005; 102-111.

Koenker, R., Basset Jr, G., 1978. *Asymptotic theory of least absolute error regression.* Journal of the American Statistical Association; 73; 618-622.

Lacuesta, A., Puente, S., Cuadrado, P., 2009. *Omitted variables in the measure of a labour quality index: the case of Spain.* Working paper series Banco de España; 0835.

Lacuesta, A., Puente, S., 2009. *The Effect of Macroeconomic Fluctuations on the Inflows and Outflows of Migrants in Spain.* Principios de Economía Política; 14; 25-48.

LaLonde, R., Topel, R., 1992. *The Assimilation of Immigrants on the US labour market.* In: Borjas, G., Freeman, R., Immigration and the Work Force: Economic Consequences for the United States and Source Areas. The University of Chicago Press Ltd. London; 63-97.

Longva, P., Raaum, O., 2003. *Earnings Assimilation of Male Immigrants in England.* Journal of Population Economics; 14; 193-220.

Lubotsky, D., 2007. *Chutes or Ladders? A Longitudinal Analysis of Immigrant Earnings.* Journal of Political Economy. 115; 820-867.

OECD, 2007. *Trends in International Migration*, Annual Report 2005 Edition.

Powell, J., 1984. *Least Absolute Deviations Estimation for the Censored Regression Model.* Journal of Econometrics; 25; 303–25.

Ramos-Muñoz, B., 2007. *Comparación de las Magnitudes Estadísticas del Empleo según la Encuesta de Población Activa y la Muestra Continua de Vidas Laborales.* Mimeo presented in Jornadas de Usuarios de la Muestra Continua de Vidas Laborales, FEDEA (Madrid)

Sanromá, E., Ramos, R., Simón, H., 2008. *The Portability of Human Capital and Immigrant Assimilation: Evidence for Spain.* IZA DP; 3649.

Simón, H., Sanromá, E., Ramos, R., 2008. *Labour Segregation and Immigrant and Native-born wage Distributions in Spain: An Analysis using Matched Employer-Employee Data*, Spanish Economic Review; 10; 135-168.

Wheatly Price, S., 1999. *The Employment Adjustment of Male Immigrants in England.* Journal of Population Economics; 14; 193-220.

Appendix. Some additional information on regularization processes in the Spanish economy

Year	Number regularized	Primary policy target	Benefit granted	Conditions
1985-1986	38.181	All foreigners	Residence and work permit	Applied to unauthorized workers and residents
1991	110100	Workers	Residence and work permit	Illegal aliens working in Spain since May 15, 1991, rejected asylum seekers or those with asylum request pending
1996	21.300	All foreigners (13800 work permits and 7500 residence)	One year residence and work permit	Applied to those residing in Spain since January 1, 1996
2000	163.900	All foreigners	One year residence and work permit	Applied to those who had previously held or applied for either work of residence permits in the three years prior to 2000, or had filed an asylum application before 2000
2001	216.400	Workers	One year residence and work permit	Applied to those who could prove employment as well as social ties in Spain
2005	548.700	All foreigners	Six months residence and work permit	Applied to those who have an employment offer lasting for 6 months.

CHAPTER 2

PERFORMANCE PAY AND THE GENDER WAGE GAP IN SPAIN *

Sara De la Rica*, Juan J. Dolado** & Raquel Vegas***

(*) Universidad del País Vasco, FEDEA & IZA

(**) Universidad Carlos III & CEPR & IZA

(***) FEDEA & Universidad Carlos III

ABSTRACT

This chapter uses detailed information from a large wage survey in 2006 to analyze the gender wage gap in the performance–pay (PP) component of total hourly wages and how it contributes to the overall gender gap in Spain. Under the assumption that PP compensation is determined in a more competitive fashion than the other components of the wage, one would expect, in principle, to find a low gender gap in the first component. However, this is not what we find. After controlling for observable differences in individual and job characteristics as well as for non random selection, the estimated adjusted gender gap in performance pay is around 26 log points in favour of males. Further, there seems to be evidence of a "glass ceiling" pattern throughout the PP distribution. After examining alternative theories that could rationalize these findings, we conclude that monopsonistic features, possibly related to women´s lower mobility due to housework duties, may be more consistent with our results than other theories related to occupational segregation.

JEL Classification: J16 and J71.

Keywords: *performance pay, gender gaps, selection bias, quantile regressions*

* This paper was prepared for the 2010 FEDEA Monograph Annual Conference on Talent, Effort and Social Mobility, U. Pompeu Fabra, may 19-20, 2010. We are grateful to our discussant Ghazala Azmat for her suggestions and to Daniel Parent for useful comments on a preliminary version of the paper. The first two authors gratefully acknowledge financial support from the Spanish Ministry of Education (ECO2009-10818; SEJ200763098), MCINN (Consolider- Ingenio2010) and Consejería de Educación de la Comunidad de Madrid (Excelecon project).

1 Introduction

One of the cornerstones of the standard competitive model of the labour market is the well-known equilibrium condition stating that wage equals labour's marginal revenue product (MRP henceforth). Accordingly, the final wage distribution represents the equilibrium outcome of demand and supply forces. This straightforward implication of the competitive model has proven instrumental in the empirical analysis of how relevant phenomena, such as changes in the relative demand and supply of skills, have affected within and between-group wage inequality over the last few decades in economies with flexible labour markets, like e.g., UK and US (see, eg., Katz and Murphy, 1992). Yet, it is also well established that the competitive model can provide a rather misleading interpretation of how wages are actually determined in real-life labour markets when information is asymmetric or search frictions are present in the allocation of workers to jobs. In addition, a common feature of many existing labour market institutions (like unions and minimum wages) is that they tend to compress the wage distribution and thus reduce pay differences between more productive and less productive workers.[1]

While it is reasonable to acknowledge that the competitive paradigm often lacks realism in describing how wages are set, there are some specific forms of wage compensation that could be considered good proxies for the "wage equals MRP" condition. In particular, if compensation is paid at least partially as a function of *performance pay* – such as bonuses, comission or piece rates- it seems plausible to assume that this wage component becomes closer to worker´s productivity than the remaining components (e.g.,the base wage) that do not depend so closely on individual performance. Following this intuitive reasoning, Lemieux et al. (2009) have analyzed the impact of performance pay (PP hereafter) on wage inequality in the US . Their basic hypothesis is that, through a widespread reduction in the cost of gathering and processing information, PP growing incidence may have contributed to the increase in inequality, mainly at the top of the wage distribution. Indeed, their finding is that PP accounts for 25% of male wage inequality between the late 1970s and the early 1980s, what provides favourable support for this hypothesis. [2]

In this paper, we contribute to this line of research by making use of a recently available dataset on the detailed breakdown of total wage compensation for Spanish workers into its different components. We re-examine Lemieux et al. ´s (2009) hypothesis, but from a different angle which, as will me discussed below, is somewhat novel in this literature[3]. That is, unlike

[1] Empirical evidence, such as Beaudry and DiNardo (1991), Card (1996), DiNardo et al. (1996), Farber and Gibbons (1996), and Lemieux (1998) provide ample evidence about the different channels through which wages are not equalized to their MRPs.

[2] The existing literature has mainly focused on analyses of the incentive effects on productivity of PP arrangements; see, *inter ali*a, Booth (1999), Ewing (1996), Dohmen and Falk (2009) and Lazear (2000), among others.

[3] There is however a growing literature on gender differences in compensation for CEOs and top executives which shares element in common with performance pay (see, e.g., Bertrand and Hallock, 2001, and Bertrand et al., 2009)

these authors, we focus on the impact of performance pay gap on the overall observed gender wage gap. Since our data provide information about the precise amount of money that workers receive as payment to reward performance -Lemieux et al (2009) only know whether workers are under PP schemes or under alternative payment schemes, but they lack information about the amount of PP workers receive- we make a detailed analysis about the how gender wage gap in the PP component of wages evolves along the total wage distribution. Specifically, we analyze whether PP compensation differs by gender and the extent to which this component contributes to explaining the overall wage gender gap (gender gap hereafter) in Spain.

To our knowledge, the only paper available in the gender literature which deals with a similar issue to ours is Chauvin & Ash (1994). Specifically, these authors use wage micro data from a survey of business school graduates to examine how the gender gap structure changes across different pay components. Based on standard Oaxaca decompositions, their main finding is that there is a very significant reduction in the adjusted-by-characteristics gender wage gap once PP differences related to job performance are accounted for. In comparison to their work, our approach makes three contributions to this strand of the literature: i) first, whereas their database only provides information about a very small segment of the population, we use here information from a much larger sample of workers, as provided by the Earnings Structure Survey, (*Encuesta de Estructura Salarial (EES) in Spanish*) covering a representative sample of the Spanish employees in the construction and services sectors in 2006 (about 82.6% of the overall Spanish salaried employment that year); ii) secondly, whereas their focus lies exclusively on the mean gender gap, we analyze how this gap evolves throughout the entire wage distribution; and iii) finally, whereas they cannot identify firm fixed effects in their data, we have information not only at the individual level but also at the firm-job level. This last feature allows us to test among alternative explanations of the gender differences in PP within the same firms.

One could think of two alternative hypotheses regarding gender differences in PP. On the one hand, under the presumption that this component is determined in a more competitive fashion than the remaining components of the wage, the gender gap in PP between equally skilled men and women could be smaller than in the non PP components. In other words, since theoretically PP responds mostly to meritocracy, equally performing workers should receive the same PP irrespectively of their gender. Moreover, if women perceive some forms of (taste and/or statistical) discrimination in non-PP jobs relative to men, then they will seek intensively for PP jobs in order to ameliorate these disadvantages. However, against this hypothesis, one could claim that, to the extent that effort at the marketplace may be negatively affected by housework, PP could also provide a clear channel through which women´ s greater involvement in household tasks hinders their returns in the labour market and, therefore, it lowers their PP relative to men ´s[4]. For example, Amuedo-Dorantes and de la Rica (2006) find that variable wage complements in Spain, which can easily make up to 40 percent of men's wages, account for up to 80 percent of the

[4] See Becker (1985) and an stylized model with this flavour in Appendix 2.

aggregate gender gap due partially to women ´s lower availability to undertake long working hours.

On the other hand, even abstracting from the role of distorting labour market institutions, the assumptions of free access to PP jobs and/or the absence of search frictions in a competitive setup may not be suitable for female workers. First, as it has been stressed in the occupational segregation literature, women may select themselves into non-PP jobs (e.g. public sector jobs) because they anticipate that these positions are more compatible with their larger household responsibilities. Hence, in line with the so-called *mommy track* hypothesis (see Mincer and Polacheck, 1977), they may willingly opt for jobs entailing steadier and, possibly, lower pay in exchange for less penalties in case of career interruptions. Secondly, employers may be more reluctant to place women in fast-track jobs involving PP if they expect lower female work attachment even though they have the same ability distribution as their male colleagues (see Lazear and Rosen, 1990). Thirdly, statistical discrimination in the allocation of PP jobs may still prevail if employers invest on workers´ specific training and therefore try to minimize quits. Moreover, if women are aware of the existence of statistical discrimination in advance, this may discourage them from applying to these jobs leading to self-fulfilling equilibria (see Coate and Loury, 1993, and de la Rica et al., 2009). Lastly, the presence of some monopsonistic features in PP jobs, due to women´ s lower mobility or lack of alternative job offers, should not be discarded even if, contrary to the standard human capital explanation, this does not lead to lower productivity (see Booth et al., 2003, and Manning, 2003).

In view of the previous considerations, our goal in this paper is to dig deeper into the specific role played by PP as a determinant of the overall gender gap in Spain. Our data comes from the recently released 2006 wave of the Spanish Earnings Structure Survey which contains detailed micro-data information on the various components of the wage, such as the base wage, overtime pay and other wage complements. When compared to the longitudinal dataset used by Lemieux et al. (2009) - i.e., the interview years 1976-1999 of the PSID- our data suffers from a clear drawback since its cross-section nature prevents us from controlling for workers´ fixed effects. In exchange, however, it has the advantage of providing information about the precise amount of PP received by workers, in contrast to PSID which only reports qualitative information on whether employees receive a variable pay component as part of their total compensation at least once during their employment relationships (but not its amount). This implies that our data are less noisy than theirs and that we can focus specifically on the PP component rather than on jobs that pay PP, as Lemieux et al. (2009) do.

In the first half of the paper, we address the impact of PP on the observed gender gap in total pay both at the mean and throughout the wage distribution, since PP is bound to have substantially different effects at different percentiles. In effect, if PP is more concentrated at the higher quantiles, where bonuses and commissions are believed to represent a more important fraction of compensation, they may have a larger impact on the gender gap and therefore

help explain at least partly the so-called "glass ceiling" effect at the top of the wage distribution. The second part focuses exclusively on the PP wage component and explores whether there are potential selection issues in the fraction of employees receiving PP, and to what extent unadjusted gender gaps change once observable individual controls and selectivity biases have been accounted for. Additionally, we present evidence about adjusted gender PP gaps within-firms and within-occupations in order to disentangle the role played by different theories explaining the existence of sizeable adjusted gender gaps.

The rest of the paper is organized as follows. Section 2 describes the dataset and provides some basic descriptive statistics regarding the whole sample, the distribution and extent of PP, the differences between the characteristics of workers receiving and not receiving PP and the contribution of the gender gap in PP to the overall gender gap in raw terms. In Section 3, we test whether the PP wage component is set in a more competitive way than the non-PP component. Section 4 deals with adjusted gaps in PP jobs, once differences in individual and job characteristics across genders are accounted for; after addressing nonrandom selection problems among workers participating in PP jobs, we analyze which of the different explanations for gender gaps fits better with our evidence. Section 5 allows for different returns on observable characteristics by gender in order to identify which specific traits are differently rewarded in the market place. Finally, Section 6 concludes. Three appendices gather some analytical and further descriptive evidence that help us to interpret the results.

2 Data and descriptive statistics

Our data source is the third wave of the Spanish Earnings Structure Survey (*Encuesta de Estructura Salarial* or ESS 06 in short)[5]. The ESS is the outcome of a European Project aiming at the design of harmonized earnings databases for several European countries. The survey is based on two-stage random samples of workers from establishments in the manufacturing, construction and service industries. First, establishments are randomly selected from the Social Security General Register of Payments records, which are stratified by region and establishment size. In a second stage, samples of workers from each of the selected establishments are also randomly drawn. Overall, sample sizes are much larger than those provided by any other Spanish survey (see below for details). Besides wage compensation, EES collects individual information on workers' demographics (such as age and educational attainment) and job characteristics (including industry, occupation, contract type, type of collective bargaining, establishment's export activity, establishment size, and region).

The main advantages of EES 06 relative to its earlier waves are that: (i) establishments with less than 10 employees are included in this survey whereas only employees in larger establishments were previously interviewed; (ii) it includes a module where employers provide detailed information on the breakdown into fixed and variable components of the total annual wage compensation paid to their workers. This module allows us to identify PP, since data on annual bonuses and commissions related to productivity are specifically reported by each firm for their employee. However, the structure of EES does not enable us to construct a matched employer-employee dataset. One important shortcoming of the EES is that information is not available on either workers´ civil status, spouses´ characteristics or the number and age of children in the households. More concretely, besides reporting total monthly gross wages and working hours, EES 06 does provide information on both the ordinary (base wage and other complements due to shifts, tenure, job risks, etc.) and non-ordinary components of annual gross earnings. Regarding the latter category, the ESS 06 distinguishes between two different types of payments:

• ***Fixed Annual Non-ordinary Payments***. This payment "basically corresponds to extraordinary compensations at Christmas and summer vacations (in Spanish, known as *pagas por navidad y verano*) [6], the standard rates for overtime work and participation in firms´ normal profits". It is specifically stated that their amount is known in advance by the employee, typically established at the collective bargaining level, and that they do not depend on either workers´ or firms´ performance.

[5] The previous waves correspond to 1995 and 2002.

[6] This implies that the fixed part of the total annual gross wage is distributed into 12 ordinary installments and 2 extraordinary ones in june and december. This tradition dates back to the Francoist industrial relations during the dictatorship period.

• ***Variable Annual Non-ordinary Payments***. In contrast to the first category, these are "payments related to workers´ or firms´ performance whose amount is not established a priori since it depends on incentives, returns and extraordinary profits". It lumps together bonuses, compensations and piece rates.

Given this breakdown of total wage compensation, the PP component in the sequel will correspond to the *Variable Annual Non-ordinary Payments* whereas the non-PP components will be identified as the sum of the ordinary wage and the *Fixed Annual Non-ordinary Payments*.

2.1. Description of the dataset

Our sample consists of full-time workers aged 18-65 for whom the interview month (October) is an ordinary period regarding their labour status. Table 1 displays the weighted descriptive statistics for the male and female samples. They contain a total of 129,930 males (66.6%) and 65,223 females (33.4%) covering almost 18,000 establishments.

Inspection of workers´ demographic characteristics reveals the following stylized facts: (i) women´ s educational attainment is significantly larger than men´ s – e.g., the percentage of female workers with a university degree (32%) almost doubles men's (18%) whereas the fraction of women with at most primary education is 10 pp. smaller (18% vs. 28%) than men´s , (ii) women are about two years younger than men (from interpolation of the mid-points of the different age brackets), (iii) female job tenure is about 1.5 years shorter than males´ tenure, and (iv) the raw gender gap is about 21 log points in favour of male workers. As regards firms´ characteristics, on average, women work in larger establishments (> 200 employees) than men (a 9 pp. higher share), and enjoy a lower coverage by bargaining agreements at the firm level (3 pp. less).

Regarding total gross hourly wages, the gender gap in favour of men is about 21 log points, using differences in mean logged wages, and 23.1%, using the ratio of the wage levels.[7] Interestingly, the incidence of PP (22.7 %) is almost identical across genders which, *prima facie*, is consistent with our previous hypothesis that this kind of jobs are attractive to women because, in principle, they should be less subject to discriminatory practices. This statement, however, will need to be reconsidered later on, once we report further evidence on the distribution of women throughout the PP distribution.

2.2. Characterization of performance pay

Table 2a compares the sample characteristics of workers and firms in PP and non-PP jobs, distinguishing by gender. The main finding is that workers on

[7] Denoting the total annual gross wage by GAW, total hourly wages are defined as w= GAW /ORH+OVH), where ORH represents annual ordinary working hours set at the collective bargaining agreement (*jornada anual pactada)* and OVH denotes the overtime working hours completed in the month of the interview (october). The latter are annualized using the seasonal pattern of aggregate extra hours in the Spanish economy as of 2006.

PP jobs are more skilled than those in non-PP jobs (40% of women and 27% of men in the PP sample have a university degree against 29% and 15% in the non-PP sample). Likewise, they are older (about a 10 pp. larger share in the 31-50 age category), have longer tenure (about 2.5 years longer for women and 5 years for men), enjoy a higher rate of permanent contracts and work in larger establishments (typically less subject to centralized bargaining levels).

Table 2b, in turn, presents the incidence of PP jobs by industry and occupation. Regarding industries, Financial Intermediation (60%) and Education (9%) are the sectors where PP is most and least prevalent, respectively. As for occupations, the results confirm that PP incidence is much higher for the high-wage categories: 50% for managers and 30% for Professionals and Technicians.

Finally, Table 2c reports the share of female workers receiving PP throughout the distribution of this component of the wage, which we can compare to the average share of women receiving PP in our sample, i.e., about 23% (=14789/ 44249). The sharp decline in this proportion as we move upward in the PP distribution - from 41% at the bottom to 16% at the top- is seemingly inconsistent with the above-mentioned hypothesis about more skilled women being more likely to seek jobs entailing PP compensation, especially since, as reported above, they have higher educational attainments than men. By contrast, such evidence would be consistent with the implications of theories based on occupational segregation and/or lower mobility which predict a "glass "ceiling "pattern whereby well qualified females are less likely to get better paid positions than high qualified males.

2.3. The Contribution of Performance Pay to the Overall Gender Gap

We next analyze how important is PP, the size of the gender gap in this wage component and, finally, its contribution to the overall gender gap. As explained in Appendix 1, the computation of the respective contributions of the gender gaps in the PP and non-PP components to the overall gender gap is greatly facilitated by using a measure of the gap expressed in percent (i.e., the ratio between average male and female wage minus unity) rather than in log points (differences in average logged wages), as is customary in the literature. The first four columns in Table 3a present the total hourly wage compensation in PP jobs (expressed in €) across genders and the corresponding shares of total wages accounted by the PP component. Further, for comparison, the hourly wages in non-PP jobs appear in the last two columns. Table 3b reports similar evidence but this time referred to the two components of the wage compensation received by PP workers, i.e. its variable and non-variable components. Lastly, Figure 1 plots the three gender gaps in percentage terms (total hourly wage, w, PP/variable component, v, and other/fixed or predetermined components, f) both at the mean and throughout the wage distribution for the whole sample of workers in our sample.

As can be observed, workers on PP jobs earn more on average than workers on non-PP jobs (about 64% and 50% higher wages for men and women, respectively) in line with the evidence offered in Table 2a about the higher skills and longer job tenure of the former. Using individual information on the PP component of the wage, we can compute its unadjusted/raw gender gap. This yields a strikingly large gap of 74% (around 46 log points using the conventional measure of the gap based on the geometric mean) in favour of men. This gap turns out to be much larger than the raw gap in total hourly wages found for the whole sample (24% or 20.6 log-points). However, the share of the PP component in the total wage is rather low: 7% for women and 9% for men on average. Nonetheless, as expected, the PP share increases over the wage distribution, reaching 22% (men) and 17% (women) at the 90th percentile (P90th). Taken together, the low average PP share and the low incidence of PP jobs in our sample (22.7%) imply a rather tiny contribution of the gap in PP to the overall gender gap for the whole sample of workers: 1.7 pp. out of 24.0 pp. (about 7%), despite a mild increase at the upper part of the wage distribution where it reaches 3.8 pp. out of 33 pp. (11.5%) at P90th.

The next step is to examine whether PP plays a more relevant role when we consider the sample of PP jobs exclusively. Using this restricted sample, Figure 2 displays the raw gaps, again both at the mean and throughout the distribution, in the total hourly wage (w), the non-PP/fixed component (f) together with the estimated contribution of the PP component to the overall gap as explained in Appendix 1[8]. As expected, the contribution of PP to the aggregate hourly wage gap is now higher than in the whole sample reaching , on average, 5.7 pp. out of 32 pp. (i.e., about 18%) and 11.7 pp. out of 46.5 pp. (25%) at the top of the distribution.

In sum, two main conclusions stem from this preliminary evidence: (I) the gender gap in PP is much larger than in total hourly wage compensation, particularly at the top of the wage distribution where it can explain about one-fourth of the "glass ceiling" pattern observed at the higher percentiles, and (II) PP makes a dent at higher wages in line with the previous evidence that workers receiving this type of variable compensation have better observable characteristics.

In principle, several theories would be consistent with the above-mentioned results. First, as regards finding (I), it is likely that wages set in collective bargaining at the sectoral (provincial, industry-wide) level and actual wages are similar for non-college workers in less-skilled/blue-collar occupations, while bargained wages do not bind for college workers in high skill/ white collar occupations. There is evidence (see Dolado et al., 1997) pointing out that employers in Spain improve high-skill workers´ pay above compressed bargained wages through formal and informal agreements which are likely to involve variable PP arrangements. Therefore, insofar as unions compress the wage distribution and base wages respond more to occupational categories and tenure than to individual characteristics, like gender, it is likely that the raw non-PP gender gap would be quite smaller than the raw PP gap.

[8] The gap in the PP component (v) is the one displayed in Figure 1.

This is confirmed by the fact that the standard deviation of the (logged) fixed component of total hourly wages (0.61 and 0.60 for men and women, respectively) is less than one-half of the standard deviation of the (logged) PP component (1.41 and 1.34, respectively).

As for finding (II), it could be rationalized by either: (i) women exerting less effort in PP jobs due to disutility of housework, (ii) women self-selecting away from PP jobs where variable components represent a relevant share of total compensation, maybe due to systematic differences in risk preferences between men and women, or (iii) women receiving lower PP than men due to monopsonistic features elements in the PP segment of the labour market, possibly related to employers´ beliefs that women enjoy lower mobility than equally qualified men.

Appendix 2 offers a simple model of the main gender implications of jobs offering PP which illustrates the main implications of the above-mentioned theories throghout the sequel.

Disentangling which of the previous theories is more likely to operate in explaining the very large gender gap found for PP requires several steps. First, Lemieux et al.´s (2007) hypothesis stating that PP tends to be closer to MRP than non-PP compensation needs to be tested. Next, we also need to examine whether the pattern of the PP raw gap discussed above remains similar once it is adjusted for differences in individual and job characteristics across genders. In other words, it is only under the competitive labor market paradigm and under similar observable characteristics that the documented PP gap can be described as being "strikingly large". The next two sections are devoted to address these issues in detail.

3 Is PP determined in a competitive fashion?

Following the above-mentioned motivation, we devote this section to analyze whether PP jobs are "more attached to the worker" whereas non-PP jobs are more "attached to the job". The basic insight once more is that, if PP depends essentially on individuals´ endowed and acquired characteristics, MRP would be more transferable across firms and occupations, supporting the idea that PP is bound to be set in a more competitive fashion than non-PP jobs. If PP jobs pay more on the basis of workers´ productivity, then human capital variables – basically age, education and, to a lesser extent, tenure [9] -should have higher market returns in this kind of jobs than in other jobs not involving PP. Conversely, returns to job characteristics- such as firm size, sector, and tenure on the firm should receive a higher market reward in non-PP jobs.

To analyze this issue, Table 4 reports standard *mincerian* (logged) hourly wage regressions estimated by OLS where the returns (estimated coefficients) to job and human capital variables are displayed separately in the first two columns for PP and non-PP jobs, respectively. The third last column, in turn, shows the results from a pooled regression where interactions of human capital and job characteristics with an indicator of receiving PP are added to test for statistically significant differences between returns in the two samples. Thus, denoting the hourly wage of worker i in firm j as W_{ij}, individual and job characteristics as X_i and X_j, respectively, and an indicator (1/0) for receiving PP as D_i, the estimated model

$$\ln W_{ij} = \beta_0 + \beta_1 D_i + X_i \beta_2 + X_j \beta_3 + D_i X_i \phi_1 + D_i X_j \phi_2 + \varepsilon_{ij}$$

where we expect $\phi_1 > 0$ and $\phi_2 < 0$.

In line with the results by Lemieux et al. (2009), we find that the returns to characteristics attached to the worker are larger in PP than in non-PP jobs. For example, the returns to university and secondary education are 41% (0.304 vs. 0.215) and 60% (0.09 vs. 0.06) larger, respectively, than in non PP jobs. Likewise, the returns to age, as a proxy for potential experience, and to a lesser extent tenure follow the same pattern. By contrast, the returns to firm size and other characteristics of the job are significantly smaller in PP jobs, as is also the case for estimated coefficients on industry and occupational dummies, not reported in this table for brevity. This evidence suports the view that PP jobs pay wages closer to worker 's productivity than the rest. Yet, the fact that estimated returns on firm´s characteristics are, in gereneral, statistically significant points out that workers tend to be categorized by firms into jobs, albeit less so in the sample of workers receiving PP.

[9] The lower (in absolute terms) coefficient on the interaction of tenure and the PP indicator may reflect high union power in collective bargaining determining the non-PP components of the wage, where tenure is a key element in wage increases.

4 Adjusted gender gaps in the performance pay component

Once the pattern of the raw gender gap in PP has been described, we proceed next to compute their adjusted counterparts accounting for differences in observed individual and job characteristics. However, the fact that slightly less than one-fourth of workers in the whole sample receive PP and that these workers present different personal and job characteristics than non-PP workers, make us consider that non random selection of workers into PP jobs may be a relevant issue to address. This is particularly important if the selection process into PP is not exactly the same for males and females. In such a case, ignoring gender differences in selection will lead to biased estimates of the adjusted gap for the PP component.

4.1. Selectivity issues

Table 5 presents the results of estimating a probit model to explain participation in PP jobs (PP=1, non-PP=0). This model will be later used to compute the inverse Mills ratio in a conventional two-stage Heckman approach to control for selection in the estimation of log hourly wage regressions explaining the PP component. Given the lack of information regarding civil status or number and age of children in our sample, we use the availability of wage bargaining at the firm level (*Firm Agreement*) as the identifying variable in the participation equation. The insight for this choice is that jobs with this type of decentralized wage agreement are more likely to involve PP compensation than other jobs where wages are set at a more centralized bargaining level (sectoral/provincial or nationwide) and unions play a larger role. Further, the fact that the estimated coefficient on this variable is not statistically significant when included in the PP wage regression make us to trust on the validity of this exclusion restriction.

In the first column of Table 5, we present the estimates of the coefficients in the probit using the standard explanatory variables, where a *Female* indicator (1/0) captures gender differences in the probability of receiving PP compensation. As can be observed, women have a lower probability of getting PP than equally able men working in the same occupations. The remaining estimates are in line with the evidence presented in Table 2a: higher educational attainment, longer tenure and being in the 31-50 age intervals also raise this probability.

Thus, in principle, this evidence goes against the earlier hypothesis that, under the competitive labour market paradigm, equally productive men and women should not exhibit significant differences regarding paticipation in jobs offering PP compensation and that, if females anticipate non-competitive features in non PP jobs, they should be more prevalent in PP jobs. To examine whether these differences in participation can be related to women´s larger disutility in market work due to larger involvement in housework, as in Becker´s hypothesis, or rather to occupational segregation and/or lower mobility, as in the " mommy track" and "monopsonistic" hypothesis, the

second column in Table 5 reports the estimates obtained in an specification where interaction terms between the different age brackets and the *Female* dummy are added to the model. Under the first hypothesis, the main differences againt women shoud appear for those age groups more prone to bear household responsibilities since it is actual involvement in housework that hampers performance in market work. Lacking information on civil status and household composition, we choose to identify women aged 31-50 as those more prone to be heavily involved in child bearing, looking after elderly relatives, etc. Thus, conditioning on the remaining observable controls, one should expect lower probability of participation for women in this age group. This would correspond to a negative coefficient on the corresponding interaction terms between PP and 31-39 and 40-49 brackets indicators. By contrast, under the second hypothesis, the effect should be mainly captured by the female intercept since all women, irrespectively of their age, antipate career interruptions due to the above-mentioned reasons.

The basic finding is that the coefficients on these interaction terms are negative and highly significant, pointing out that, conditioning on all the remaining covariates, women in two above-mentioned age brackets have a lower probability of receiving PP than younger and older women, respectively. For example, the net coefficient of a woman aged 30-39 is -0.134 (=-0.103+0.016-0.047) whereas, for women below 30 or in the 50-59 interval, the corresponding net coefficients would be -0.103 and -0.041 (=-0.103-0.01+0.072), respectively. A chi-squared test rejects the null of equal coefficients across the previous age brackets with a p-value of 0.023. Further, when not only age dummies, but also education categories are interacted with the female dummy variable (third column in Table 5), the lower estimated probabilites of receiving PP for women in the two above-mentioned age brackets remain basically unaltered. Interestingly, the coefficient of the interaction between the female and university education dummies is negative. One possible interpretation of this result could be that women with higher educational attaiment face greater restrictions than less skilled women in adequatelly matching their skills and occupational requirements with their family responsibilities, and therefore shy away from PP. Overall, the above-mentioned evidence points out the "mommy track"/"monopsonistic" hypotheses are likely to play a joint role in explaning gender differences in receiving PP

4.2. Disentangling occupational segregation from monopsonistic features

4.2.1 Within- firms and within- occupations regressions

The next step is to analyze which of the two theories embedded in the second joint hypothesis is more likely to explain the PP gender gap: is it "occupational segregation" or " monopsonistic features"?. To try to discriminate between these two somewhat alternative explanations, we carry out the following exercise. Using the specification of a *mincerian* wage equation for the restricted sample of PP workers with a *Female* intercept and equal returns to individual and job characteristics across genders, we compare the

estimated coefficient on the *Female* indicator in a regression (augmented by the inverse Mills ratio obtained from the participation equation reported in the second column of Table 5 under four different specifications: (i) a pooled regression (*P*), (ii) within- occupations (*WO*),[10] (iii) within-firms (*WF*), and (iv) within-firms & occupations (*WFO*).

The insight for such comparison can be briefly described as follows. Let us denote the coefficient on the *Female* dummy in the four specifications above as β_P, β_{WO}, β_{WF} and β_{WFO}respectively. Then, under the "occupational segregation "hypothesis we should expect β_{WO} to be significantly smaller than β_p (since we are comparing men and women in the same occupation and firm) whereas β_{WF} should be similar to β_P. Conversely, under the "monopsonistic" hypothesis, the estimate of β_{WF} should be quite smaller than the estimate of β_p (since now the comparison is between men and women working in the same firm), while β_{WO} and β_P would be similar. Finally, if both theories play a role, then β_{WFO} should be below both β_{WO} and β_{WF} which, in turn, should be smaller than β_P.

Table 6 reports the estimates obtained under the alternative specifications where the OLS results (without selection correction) are also included in the first column for comparison. The following findings stand out. First, the adjusted average gender gap in the OLS pooled specification is about 41 log points against a raw gap of 46 log points. Second, once we control for selection bias in such specification, the gap increases slightly to 45 log points. The fact that this gap is larger than the OLS gap is explained by the highly significant positive sign on the coefficient of Heckman´s lambda which reflects strongly positive selection of workers in PP jobs. Since women have higher educational attainment than men in our sample, despite having lower tenure, this leads to a larger gap when selection is taken into account. Third, again controlling for selection biases, the estimate of the gap in the within-firm specification (34 log points) is quite smaller than the corresponding estimate in the within-occupation specification (43 log-points) which, in turn, is quite close to the gap estimated in the pooled specification (41 log points). Finally, the gap in the joint within-firm and occupation (29 log-points) is slightly lower than the gap in the within-firm and within-occupation model. Lastly, in Appendix 3(B) we report some evidence in favor of the existence of gender differences in risk attitudes whereby women tend to avoid jobs entailing PP more than men. However, the amount of self-segregation due to this feature seems to be too small in light of the similarity of the gaps in the pooled and WO specifications. Hence, we interpret all this evidence as seemingly yielding higher support to the "monopsonistic" hypothesis at the joint occupational-firm level than to the conventional "occupational segregation" hypothesis in explaining the large documented PP gender gap.

[10] We use the most disaggregated occupational classification available, i.e., 18 occupational categories.

One potentially important caveat about the previous results is that so far we have implicitly assumed that workers receiving PP and those entitled to receive PP were identical populations. However, that assumption may not hold if there is a fraction of workers, among those entitled to receive PP, who did not get it because of poor job performance or any other reasons (e.g., due to holding a different labor contract). In the absence of further information on PP entitlement, it seems natural to define an upper bound on the true size of this category of workers, by assuming that of all those individuals in our sample who did not receive PP ("PP non-recipients") but happen to work in the same firms and occupations as other co-workers who actually were awarded PP, belong to this group. In Appendix 3 (C), Tables A.8 and A.9, we provide a detailed discussion on the subsample PP non-recipients and their characteristics. The main conclusion to be drawn is that differences by gender among this category of workers are not noticeable across sectors and occupations. Further, to check if our main conclusions from Table 6 are altered when considering information about "PP non-recipients", we perform the following sensitivity analysis. In Table 7, we re-estimate the regressions reported in Table 6 considering now a selectivity correction concerning the probability of receiving PP among "PP recipients" and "PP non-recipients", namely the sample of workers which are assumed to be entitled to PP. As can be observed, the results in Tables 6 and 7 are rather similar supporting once more the "monopsonistic" hypothesis against the occupational segregation" hypothesis.

Nonetheless, the previous conclusion could be premature if it were to be the case that women exert less effort than men in the same occupation and firm because of their larger involvement in housework. In such a case, differences in effort could also be behind the gap in PP. Lacking a precise measure of productivity, it is difficult to test this hypothesis[11]. However, despite the absence of controls on household composition, the fact that our sample consist of full-time workers and that we control for age, education and tenure – all related to productivity- as well as that the overtime hours are similar for men and women receiving PP. (although, as discussed in Appendix 3(A) the proportion of men doing overtime in the overall sample of workers is higher than the corresponding proportion of women), make us believe that gender differences in effort do not play an important role in explaining the gap in PP. One possible test of differences in effort can be implemented by checking whether the proportion that PP represents over the total hourly wage is lower for women than for men, once we control for differences in characteristics and sample selection. The insight is simply that higher effort should give rise to a larger proportion of PP. As mentioned above, on average, these proportions are 9% for men and 7.2% for women. Though not reported for brevity, we have run a similar *mincerian* regression to the one shown in column (5) of Table 6, where the dependent

[11] In Appendix 3(A), we provide additional evidence to that reported in the main text about gender differences in effort exerted at the firm level not driving the gap in PP. We use information on overtime hours as a proxy for effort since the individuals considered in our sample are full-time workers.

variable is the *logit* transformation of the above-mentioned proportion.[12] We obtain that the Female indicator explains less than one-half (0.85 pp.) of the 1.8 p.p. gap (=9.0-7.2) gap in the proportions. Therefore, we conclude that such a difference is too small to be considered as a very relevant factor in explaining the gender gap in PP.

4.2.2 Quantile regressions

Further evidence on this issue can be obtained from comparing the relative pattern of the gender gap throughout the distribution on the PP component. Indeed the available theories on female segregation in slow-track jobs (*aka* non-PP jobs in our setup), like Lazear and Rosen ´s (1993), predict that gender gaps arise because women face lower probability of being assigned to PP jobs even if they are as skilled as men, not because they are subject to within-job discrimination. However, given that the ability standard for allocation to PP jobs is higher for women, it should be expected that the relatively few women who are at the top of the PP distribution should receive higher PP compensation than their male counterparts. In other words, conditioning on observable characteristics, the gender gap should be negative at the top percentiles of the PP distribution. By contrast, theories related to lower female mobility, like Booth et al. ´s (2003) "sticky floors" hypothesis, predict that women at all percentiles will be paid less than men since there is a higher rent to be earned by firms due to women having lower outside opportunities because employers perceive that they are less mobile than men.

To test which of the two previous implications is supported by the data, we use quantile regressions (QR) accounting for selectivity corrections under the within-firm & occupation specification. Following Buchinsky´s (1998) approach, the selectivity correction for workers receiving PP is based on a two-stage approach. First, a two-term series expansion of the inverse of the Mills ratio in Table 5 is used to obtain an estimate of a latent index that approximates the unknown quantile functions of the truncated bivariate distribution for the error terms in the wage and participation equations. The covariance matrix for the two-stage QR and the selectivity corrected estimates is obtained by bootstrapping the design matrix with 100 replications.

Table 8 reports the QR estimates of the coefficient on the *Female* dummy for a few relevant percentiles of the PP distribution. A clear "glass ceiling" pattern emerges with the gap evolving from 20 log-points at the bottom deciles to 43 log points at the top of the distribution. In line with our previous discussion, we interpret again this evidence as being supportive of the existence of monopsonistic features in the determination of PP.

---

[12] The *logit* transformation, $\ln(R/1-R)\in(-\infty,\ +\infty)$, achieves consistency with the support of the distribution of the error tem in the regression, where $R\in(0,\ 1)$ is the proportion of PP in the total hourly wage. Denoting by b the estimated coefficient in the regression, then the effect of the *Female* dummy, D, on R becomes $\delta R/\delta D = bR(1-R)$.

5. Decomposing the Gender Gap in PP

So far, the estimated models have assumed the same market returns (coefficients) to male and female characteristics, except the intercept. Since this assumption is rejected by the data (p=value 0.023), we next report results allowing for different remuneration to observed characteristics for workers in the PP sample.

Table 9 summarises the results of the slightly modified version of the Oaxaca-Blinder gender gap decomposition proposed by Gardeazabal and Ugidos (2004) when, as in our case, there are indicator variables in the hourly wage regressions which can take more than two categories (e.g.., education and age). The reported results correspond to the *WFO* specification. In general, the results indicate that the contribution of differences in returns to explain the PP gender gap (46 log-points) is much larger (88%) than the contribution of differences in characteristics (12%). Among the former, the largest components are the differences in constant terms (26 log points) and in the returns to age. Though we only report the aggregate contribution for all age categories, it is worth noticing that the two specific categories where differences in returns are larger are the 30-39 and 40-49 groups which jointly account for 5.67 log-points out of the 8.46 log- points contributed by age. This result somewhat points out that typical ages where individuals incur in child bearing or other household tasks involves a "marriage premium " for males and a "child/elderly parent penalty" for women, in line with many studies of the gender gap in Spain (see, e.g., de la Rica et al. , 2008). Interestingly, albeit not large, the differences in returns to tenure favour women, in agreement with our previous result that firms may find it optimal to offer steeper wage-tenure profiles to women than to men in order to retain them. Finally, the fact that the female intercept accounts for 26 log points of the overall gender gap when in the pooled WFO regression it accounted for 29 log-points may just reflect that the lack of variables in our dataset capturing civil status and household composition may still biasing upwards the size of this coefficient.

All in all, the result in this section do not change our previous hypothesis that the gender gap in PP and the corresponding "glass ceiling" may be well due to monopsonistic features in PP jobs, whereby female lower mobility leads to a rate of exploitation by firms even when women acquire higher education than men to signal their commitment to job stability.

5 Concluding remarks

In this paper we have used a large cross-section dataset for Spanish workers in 2006 to examine whether the gender performance-pay (PP) gap differs from the gender gaps in other components of wage compensation. We have found evidence that PP responds more to workers´ performance and that women in PP jobs have several observable characteristics which are better than men ´s (e.g., educational attainments). Yet, our main result is that the gender gap in PP is much higher, both in raw terms and adjusted for observable characteristics, than the gap in non-PP compensation, and that there are clear signs of a "glass ceiling" effect (higher gaps and lower female participation of women in the upper parts of the PP distribution).

According to our empirical results the most likely explanation for these findings is the existence of monopsonistic features in the PP segment of the labour market, related to women´s lower mobility due to their attachment to household tasks, and to a lesser extent on alternative theories explaining women ´s segregation in different occupations than men. Among the latter, we have also considered gender differences in attitudes towards risk. We find some evidence supporting the hypothesis that women are more risk averse than men, which leads them to prefer industries and occupations with more stable payment schemes where, therefore, PP represents a lower proportion of the overall pay. Nonetheless, the similarity of the estimated gender gaps in the pooled and within-occupations specifications lead us to conclude that this self-segregation phenomenon is not too relevant and that, as stated before, monopsonistic features in the PP segment of the labour market seem to provide a more plausible explanation of the documented gender gap in PP. .

References

Amuedo-Dorantes, C. & de la Rica, S. (2006), "*The Role of Segregation and Pay Structure on the Gender Wage Gap: Evidence from Matched Employer-Employee Data for Spain*", Contributions to Economic Analysis and Policies, Berkeley Electronic Press Journals, 1.

Beaudry, P. & DiNardo, J. (1991), "*The effect of implicit contracts on the movement of wages over the business cycle: Evidence from Micro Data*", Journal of Political Economy, 99(4):665-688

Becker, G. (1985), *"Human Capital, Effort and the Sexual Division of Labor",* Journal of Labor Economics, 3.

Bertrand, M. & K. Hallock (2001), "*The* ***Gender Gap in Top Corporate Jobs",*** Industrial and Labor Relations Review,55, 3-21.

Bertrand,M., Goldin, G. & L. F. Katz (2009), " *Dynamics of the gender gap for young professionals in the corporate and financial sectors" NBER W.P. Series* 14681.

Booth, A. & J. Frank (1999), "*Earnings, Productivity and Performance-Pay*", Journal of Labor Economics, 17 (3).

Booth, A., Francesconi, M & J. Frank (2003), "*A sticky floors model of promotion, pay and gender",* European Economic Review, 47, 295-322.

Buchinsky, M.(1998), "*Recent Advances in Quantile Regression Models: A Practical Guideline for Empirical Research.*" Journal of Human Resources, 33(1), 88-126.

Card, D. (1996), "*The Effect of Unions on the Structure of Wages: A Longitudinal Analysis.*" Econometrica 64, 957-979**.**

Chauvin, K. W., & Ash, R. A. (1994), "*Gender earnings differentials in total pay, base pay, and contingent pay*". Industrial and Labor Relations Review, 47: 634–649

Coate, S. & G. Loury (1993), *"Will affirmative-action policies eliminate negative stereotypes ",* American Economic Review, 83, 1220-40.

de la Rica, S., Dolado, J.J. & V. Llorens (2008), "*Ceilings or floors ?. Gender wage gaps by education in Spain",* Journal of Population Economics, 21, 755-776.

de la Rica, S., Dolado, J.J. & C. García-Peñalosa (2009), "*On gender gaps and self-fulfilling expectations: Theory, policies and some empirical evidence*"**,** IZA Discussion Paper No. 3553.

DiNardo, J., Fortin, N & T. Lemieux (1996), "*Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach,*" Econometrica, 64(5), 1001-44.

Dolado, J.J. , J. Jimeno & F. Felgueroso (1997), "*Minimum Wages, Collective Bargaining and Wage Dispersion: The Spanish Case*", European Economic Review, 41, 713-725 .

Dohmen, T. & A. Falk (2009), *"Performance Pay and Multi-dimensional Sorting Productivity, Preferences and Gender",* Forthcoming in American Economic Review.

Ewing, B. (1996), *"Wages and performance-based pay: Evidence from NLSY",* Economic Letters 51, 241-246.

Farber, H. & R. Gibbons (1996), *"Learning and Wage Dynamics"*, The Quarterly Journal of Economics, 1007-47 .

Gardeázabal, J & A. Ugidos (2004) *" More on identification in detailed wage decompositions"*, Review of Economics and Statistics, 86, 1034-36.

Katz, L.F. & K.M. Murphy (1992), *"Changes in relative wages, 1963-1987: Supply and demand factors"* Quarterly Journal of Economics, 107, 35-78.

Lazear, E. & S. Rosen (1990), *"Male and female differentials in job ladders",* Journal of Labor Economics, 8, 106-123.

Lazear, E. (2000): *"Performance Pay and Productivity,"* American Economic Review, 90(5),1346–1362.

Lemieux, T. (1998), "*Estimating the Effects of Unions on Wage Inequality in a Panel Data Model with Comparative Advantage and Non-Random Selection*" Journal of Labor Economics 16, April 1998, 261-291

Lemieux, T. & D. Parent (2009), *"Performance Pay and Wage Inequality",* Quarterly Journal of Economics 124(1), 1-49.

Manning, A. (2003), "*Monopsony in motion: Imperfect competition in Labor Markets*", Princeton University Press.

Mincer, J. & S. Polacheck, 1977, *"Women´s earnings re-examined",* The Journal of Human Resources, 13 (1), 118-134

APPENDIX 1: Contribution of PP gender gap to the total gender gap

The total hourly wage (w) is defined as the sum of the non-PP/fixed component (f) and the PP/variable component (v). Since $\ln(f+v) \neq \ln(f)+\ln(v)$, it is difficult to compute the gender gap in log points, as is customary in the literature. , by means of the difference between the averages of logged wages for males (M) as females (F) i.e., as $\overline{\ln w_M} - \overline{\ln w_F}$, and then proceed to decompose the overall gap into the respective gaps of the fixed and variable components.

However, such decomposition is straightforward if we express the gaps in percent rather than in logged points. In effect, given that:

$$w = f + v$$

wage averaging for each gender yields:

$$\overline{w}_M = \overline{f}_M + \overline{v}_M \tag{A.1.1}$$

$$\overline{w}_F = \overline{f}_F + \overline{v}_F \tag{A.1.2}$$

Therefore, the following decomposition of the total gender gap (in percent) in terms of the two gender gaps of the components (also in percent) holds exactly for the restricted sample of PP workers:

$$\left(\frac{\overline{w}_M - \overline{w}_F}{\overline{w}_F}\right)_{PP} = \alpha\left(\frac{\overline{f}_M - \overline{f}_F}{\overline{f}_F}\right)_{PP} + (1-\alpha)\left(\frac{\overline{v}_M - \overline{v}_F}{\overline{v}_F}\right)_{PP} \tag{A.1.3}$$

where $\alpha = \frac{\overline{f}_F}{\overline{w}_F}$, while for the whole sample of PP and non-PP workers, we have:

$$\left(\frac{\overline{w}_M - \overline{w}_F}{\overline{w}_F}\right)_{ALL} = \alpha\left(\frac{\overline{f}_M - \overline{f}_F}{\overline{f}_F}\right)_{ALL} + (1-\alpha)\varphi\left(\frac{\overline{v}_M - \overline{v}_F}{\overline{v}_F}\right)_{PP} \tag{A.1.4}$$

such that $\varphi = \frac{N_{PP}}{N_{ALL}}$, where N_{PP} and N_{ALL} are the respective number of observations in the restricted and whole samples. Hence, the second terms in the RHS of (A.3) and (A.4) are interpreted as the contributions of the gender gap in the PP component to the overall gender gap in the two samples.

APPENDIX 2: An illustrative model of the gender implications of PP

(A) Competitive wages

Let us assume that a PP worker of (exogenous) skill δ receives a wage W per unit of output produced and that firms incur a fixed C of monitoring the worker which in a competitive market is paid by the worker. Denoting effort by e, output is assumed to be $(\delta + e)$. Effort produces a disutility cost $c(e)$ which is increasing and convex. We assume the functional form $c(e) = e^{1+\gamma}/(1+\gamma)$ with $\gamma > 0$. Given women´s higher involvement in housework, their disutility of effort is higher than for men, namely, $\phi e^{1+\gamma}/(1+\gamma)$, where ϕ >1. Therefore, we can write down the utility of men (M) and women (F) in PP jobs as follows:

$$U_M^{PP} = W(\delta + e) - C - e^{1+\gamma}/(1+\gamma) \tag{A.2.1}$$

$$U_F^{PP} = W(\delta + e) - C - \phi\, e^{1+\gamma}/(1+\gamma) \tag{A.2.2}$$

Regarding non PP jobs, let us assume that the worker produces a minimum level of output, say $\bar{\delta}$, which can be monitored by the firm at no cost and does not involve any effort. After all, it is painful to produce output and, in the absence of monitoring, the worker can get away without producing any more than $\bar{\delta}$. This implies that the utility for both men and women of this type of job is simply given by:

$$U^{NPP} = W\bar{\delta} \tag{A.2.3}$$

The workers´ effort decision in PP jobs is simply obtained by equating the marginal revenue from exerting effort to its marginal cost. From (A.2.1) and (A.2.2), it yields $W = e_M^{\gamma}$ and $W_F = \phi e_F^{\gamma}$, whereby $e_M^{\gamma} = \phi e_F^{\gamma} > e_F^{\gamma}$. Substituting these two expression into (A.2.1) and (A.2.2), implies that worker $i(= M, F)$ will choose PP for $\delta > \delta_i^*$, where

$$U_M^{PP}(\delta_M^*) = \frac{\gamma}{1+\gamma} e_M^{1+\gamma} + \delta_M^* e_M^{\gamma} = \frac{\gamma\phi}{1+\gamma} e_F^{1+\gamma} + \varphi\delta_M^* e_F^{\gamma} - C = W\bar{\delta} = U^{NPP} \tag{A.2.4}$$

$$U_F^{PP}(\delta_F^*) = \frac{\gamma\phi}{1+\gamma} e_F^{1+\gamma} + \delta_F^* e_F^{\gamma} - C = W\bar{\delta} = U^{NPP} \tag{A.2.5}$$

Comparing both expressions, we get that $\delta_F = \phi\delta_M > \delta_M$. Thus, assuming that the skills distribution is identical across genders, we should expect fewer women in PP jobs and, conditionally on receiving PP, higher ability among female workers than among male workers.

Further, if women are aware of discrimination in non- PP jobs where, say, they get paid $U_F^{NPP} = \alpha W\bar{\delta}$, with $0 < \alpha < 1$, whereas $U_M^{NPP} = W\bar{\delta}$, then obviously

they will have higher preference for PP jobs than before. Moreover, in this case women will even be more prominent than men in PP jobs if $\phi = 1$.

(B) Predetermined wages and job attachment uncertainty

A slightly different model where wages are no longer given but set by employers in order to avoid career interruption can be written using a slight adaptation of Lazear and Rosen ´s (1992) model of assignment of workers to slow and fast- jobs. Let us assume that individuals in PP jobs work for two periods and are endowed with the same ability δ which is known to the firm. In the initial period, they produce δ and receive a wage W_1. As a result of longer tenure, their productivity in period 2 raises to $\mu\delta$, where $\mu > 1$ and gets paid W_2. It is assumed that, workers receive a disutility shock, ω, in *both* periods which may force them to quit the job (say, for family duties). The ω shock is an *i.i.d.* random variable, independent across periods, with c.d.f. F(ω) which is revealed to the worker after the wage in either period has been set by the firm. Thus, wages are predetermined and workers will stay in the firm both periods as long as $W_{ti} - \omega \geq 0$, $t=1, 2$ and $i=F, M$.

The key difference between men and women is that the c.d.f. for men, $F_M(\omega)$, is stochastically dominated by the c.d.f. for women $F_F(\omega)$, namely $F_M(\omega) > F_F(\omega)$ for $\omega > 0$. This assumption captures the fact that women are more likely to be affected by the shock than men. To simplify matters, and without loss of generality in terms of the qualitative results, we will assume that dF(.) are uniform distributions, such that the density functions verify: $f_M(\omega) = U[0, \varepsilon_M]$ and $f_F(\omega) = U[0, \varepsilon_F]$, with $\varepsilon_F > \varepsilon_M$.

To solve for both wages, we proceed backwards in time. Under the assumption that the wage in period 2, W_{2i} (i=f, m), is offered before ω is realized, employers will choose W_{2i} in order to maximize expected profits in period 2, subject to the participation constraint in this period and conditional on the probability of staying in the firm during period 1 (equal to W_{1i}/ε_i under a uniform distribution), namely:

$$\max_{W_{2i}} \frac{W_{1i}}{\varepsilon_i} \int_0^{W_{2i}} \left(\mu\delta - W_{2i}\right) d\omega = \max_{W_{2i}} \frac{W_{1i}}{\varepsilon_i^2} [\mu\delta W_{2i} - W_{2i}^2], \quad i = M, F \qquad \text{(A.2.6)}$$

Whereby the first-order condition (f.o.c.) w.r.t. W_{2i} implies that the wage paid in equilibrium to male and female workers is identical:[13]

$$W_2^* = \mu\delta / 2 \qquad \text{(A.2.7)}$$

and by replacing W_2^* into the bracketed term in (A.2.6) , the firm ´s profit in period 2 (Π^*_{2i}) is given by $\Pi^*_{2i} = W_{1i}^2 (\mu\delta)^2 / 4\varepsilon_i^2$.

[13] This is just the average of the worker´ s productivity´ and the outside wage which is assumed to be zero. The weight ½ in the average is due to the choice of the uniform distribution in the illustration. Alternative distributions will give rise to a weighted average with unequal weights.

Going back to period 1, firms will choose W1i to maximize the sum of expected profits in both periods, subject to the participation constraint in that period, i.e.

$$\max_{W_{1i}}\{\frac{1}{\varepsilon_i}\int_0^{W_{1i}}(\delta - W_{1i})d\omega + \frac{W_{1i}^2}{4\varepsilon_i^2}(\mu\delta)^2\}, \quad i = M, F \tag{A.2.7}$$

which implies that:

$$W_{1i}^* = \frac{\delta}{2} + \frac{(\mu\delta)^2}{8\varepsilon_i} \tag{A.2.8}$$

Since $\varepsilon_F > \varepsilon_M$, it follows that $W_{1M}^* > W_{1F}^*$. Given that, for the same δ, $W_{2M}^* = W_{2F}^*$, it follows that the return to tenure $(W_{2i}^* - W_{1i}^*)$ is higher for women than for men.

Non-PP jobs can be interpreted in terms of this model as implying that $\mu = 1$, that is, a flatter wage profile. From (A.2.7) and (A.2.8) with $\mu = 1$, we get that wages in non-PP jobs are lower than in PP jobs.

APPENDIX 3: Additional gender differences and their effect on the PP gap

(A) Gender differences in exerted effort (measured by overtime hours)

Gender differences in PP compensation could be due to differences in "effort" exerted by men and women, respectively, in similar jobs. Since effort is hard to monitor, firms are bound to interpret it broadly, including workers´ willingness to work irregular hours and to avoid long parental leaves. Women may be pushed to exert less "effort" and to have "less job attachment" due to their larger involvement in housework and motherhood responsibilities, specially while childrearing. Understanding "effort" in such a broad way may possibly lead employers to perceive mothers as less committed workers and, as a result, firms may be more reluctant in providing access to "fast track" jobs to these women.

This type of firms´ behavior makes it unlikely that women with young children would return to the occupation and wage patterns they enjoyed before their children were born, leading to high persistence in the observed PP gender gap. Further, if women anticipate this type of employers´ behavior, they may segregate themselves into specific occupations/sectors where their family/personal chores become less penalized.

The aim of this Appendix is precisely to shed some light on the role played by "differences in effort" in explaining the documented gender gap in PP. Measuring individuals' effort is a difficult task, and even harder is to identify whether there are differences by gender in workers´ effort within the firm. Indeed, the lack of information on civil status and number of children of workers in our database prevents us from testing properly for differences in effort between men and women with children, or between women with children and those without them. In view of these shortcomings, , the only issue we can address is whether there are gender differences in terms of "effort" broadly understood as discussed in the sequel.

The length of the workday could be considered as an imperfect measure of the level of effort that a worker exerts in the labor market and therefore of job engagement. However, since we have focused exclusively on full-time workers in this paper, this implies that even if the women are more involved in housework than men are, they have already decided not to reduce their labor market attachment by working part-time. Therefore, at least a priori, we should expect similar degree of commitment across genders except for their disposition to work from overtime. In other words, if women in our sample decide to put less "effort" in the job due to housework and/or motherhood burden, they will work less overtime. In our sample, the proportion of men working overtime (9.36%) nearly trebles the corresponding female proportion (3.35%). Table A.1 shows the percentage of workers doing overtime by gender, occupation and payment scheme (PP and no PP workers). As can be observed, the proportion of men working overtime is higher than women's in some occupations- such as managers, professionals, agriculture and fisheries, and operators and assemblers.

Table A.1: Incidence of overtime by gender, occupation and Payment scheme

	Non PP Jobs		PP Jobs	
	Males	*Females*	*Males*	*Females*
Managers	1.20%	1.10%	1.00%	0.20%
Professionals	1.80%	1.80%	1.00%	1.20%
Technicians	4.70%	5.60%	1.70%	2.30%
Clerks	6.10%	6.60%	2.30%	3.00%
Personal Services	9.00%	7.10%	2.20%	2.40%
Agriculture and Fisheries	5.20%	1.40%	2.80%	0.00%
Craftsmen	10.10%	19.50%	8.10%	9.80%
Operators and Assemblers	12.40%	19.00%	8.90%	10.50%
Laborers, non-qualified operators	11.00%	13.60%	6.10%	5.80%

To check if the observed gender differences in overtime are statistically significant, we proceed to test two null hypothesis: i) the same proportion of workers doing overtime by gender; and, ii) gender equality in average overtime hours worked within a specific job and within a payment scheme, for those working overtime. The alternative hypotheses considered are one-sided (i.e., women exert less "effort" than men) implying: i) lower female participation in overtime and, ii) lower average overtime hours for those women who work overtime. The results (percentage of men and women doing overtime by occupation, t values and p-values for the above mentioned hypothesis tests) are provided in Tables A.2 to A.5 and can be summarized as follows. On the one hand, Hypothesis (i) is rejected against the alternative of higher participation by males in all the occupations (but agriculture and fisheries) at 5% level, both considering the whole sample and the restricted sample of PP workers. On the other hand, Hypothesis (ii) is also rejected at 5% level against the corresponding one-sided alternative in all occupations (when all workers are pooled to perform the test) but agriculture and fisheries. However, when only PP workers are considered, we cannot reject the null of equal overtime hours in most cases.

Table A.2.Hypothesis Test: Mean Differences in Percentage of Workers doing Overtime by job and gender. All workers

	% Working Overtime		*Hypothesis Test*		
			Ho: diff =0	Ha: diff > 0	
	Male	**Female**	t	Pr(T > t)	
All occupations	**9.40%**	**3.40%**	**48.008**	**0.000**	***
Managers	1.20%	0.60%	1.878	0.030	***
Professionals	1.80%	1.00%	4.649	0.000	***
Technicians	5.00%	1.90%	13.76	0.000	***
Clerks	6.20%	2.40%	15.028	0.000	***
Personal Services	8.60%	2.30%	19.236	0.000	***
Agriculture and Fisheries	4.60%	2.30%	0.725	0.234	
Craftsmen	11.70%	8.40%	5.916	0.000	***
Operators and Assemblers	13.60%	9.20%	8.883	0.000	***
Laborers, non-qualified operators	11.40%	6.00%	13.077	0.000	***

Note: diff = mean (Male) - mean (Female)
Rejection criteria: (***) Accept the alternative at 5% level; (**) Accept the alternative at 10% level;
(*) Accept the alternative at 15% level

Table A.3.Hypothesis Test: Mean Differences in Percentage of Workers doing Overtime by job and gender. PP workers

	% Working Overtime		*Hypothesis Test*		
			Ho: diff =0	Ha: diff > 0	
	Male	**Female**	t	Pr(T > t)	
All Occupations	**10.70%**	**3.20%**	**27.24**	**0**	***
Managers	1.10%	0.20%	2.22	0.013	***
Professionals	1.80%	1.20%	1.797	0.036	***
Technicians	5.60%	2.30%	7.689	0	***
Clerks	6.60%	3.00%	6.661	0	***
Personal Services	7.10%	2.40%	6.73	0	***
Agriculture and Fisheries	1.40%	0.00%	0.338	0.368	
Craftsmen	19.50%	9.80%	5.216	0	***
Operators and Assemblers	19.00%	10.50%	6.354	0	***
Laborers, non-qualified operators	13.60%	5.80%	6.351	0	***

Note: diff = mean (Male) - mean (Female)
Rejection criteria: (***) Accept the alternative at 5% level; (**) Accept the alternative at 10% level;
(*) Accept the alternative at 15% level

Table A.4.Hypothesis Test: Mean Differences in Overtime Hours Worked by job and gender. All workers

	Overtime Hours Worked		*Hypothesis Test*		
			Ho: diff[2]=0	Ha: diff[2]> 0	
	Male	Female	t	Pr(T > t)	
All Occupations	15.54	14.03	1.293	0	***
Managers	14.35	32.43	-1.854	0.945	
Professionals	17.18	14.86	1.149	0.126	*
Technicians	14.71	12.65	2.019	0.022	***
Clerks	15.38	11.56	4.037	0.000	***
Personal Services	23.87	12.68	8.634	0.000	***
Agriculture and Fisheries[1]	13.48	16	-	-	
Craftsmen	14.39	14.65	-0.28	0.610	
Operators and Assemblers	14.1	14.2	-0.154	0.561	
Laborers, non-qualified operators	19.28	16.13	3.741	0.000	***

Note: [1]There is only 1 female working overtime and the test is therefore not performed.
[2]diff = mean(Male) - mean(Female)
Rejection criteria: (***) Accept the alternative at 5% level; (**) Accept the alternative at 10% level;
(*) Accept the alternative at 15% level

Table A.5.Hypothesis Test: Mean Differences in Overtime Hours Worked by job and gender. PP workers

	Overtime Hours Worked		*Hypothesis Test*		
			Ho: diff =0	Ha: diff > 0	
	Male	Female	t	Pr(T > t)	
All Occupations	14.27	12.815	1.994	0.023	***
Managers	13.96	16.000	-	-	
Professionals	20.64	21.813	-0.240	0.595	
Technicians	13.45	11.422	1.133	0.129	
Clerks	12.02	9.770	1.440	0.076	***
Personal Services	17.49	9.940	3.246	0.001	***
Agriculture and Fisheries[1]	16	-	-	-	
Craftsmen	14.09	15.085	49.171	0.664	
Operators and Assemblers	14.42	13.951	0.306	0.380	
Laborers, non-qualified operators	14.89	13.714	0.528	0.299	

Summing up, the above-mentioned results confirm that the proportion of men working overtime is higher than the corresponding female proportion across different occupations both in the whole sample and in the restricted sample of those receiving PP. By contrast, in general we do not find significant mean differences by gender in the number of overtime hours performed by workers receiving PP.

To identify if differences in effort play a role in explaining the observed gender PP gap, the exercise performed in the previous section is repeated here (see Table A.6) including as additional variables an indicator (1/0) for those working Overtime, interacted additionally with Female, plus the ratio between overtime hours worked and regular time in job. As can be observed, working overtime does not have significant effects on the gender PP gap. Hence, we interpret this preliminary evidence as further support for the "monopsonistic hypothesis".

Table A.6. Estimates of alternative specifications of log hourly wage equation for PP workers corrected for selectivity

Dependent variable: Log PP Hourly Wage					
	(1)	(2)	(3)	(4)	(5)
	OLS	Heckman selection	WF	WO	WFO
Female	**-0.425*****	**-0.469*****	**-0.358*****	**-0.430*****	**-0.244*****
	(0.014)	(0.017)	(0.016)	(0.015)	(0.017)
Overtime	**-0.294*****	**-0.296*****	**-0.184*****	**-0.174*****	**-0.049**
	(0.032)	(0.032)	(0.028)	(0.031)	(0.030)
Overtime*Female	**0.054**	**0.050**	**0.070**	**-0.033**	**-0.031**
	(0.065)	(0.063)	(0.053)	(0.062)	(0.055)
Overtime/Regular Time	**-5.909****	**-5.782****	**-2.350**	**-5.119****	**-3.810***
	(2.503)	(2.464)	(2.177)	(2.416)	(2.277)
Age 30-39 (ref:<30)	**0.252*****	**0.328*****	**0.218*****	**0.283*****	**0.164*****
	(0.019)	(0.023)	(0.024)	(0.020)	(0.024)
Age 41-49	**0.325*****	**0.370*****	**0.276*****	**0.303*****	**0.173*****
	(0.022)	(0.026)	(0.021)	(0.022)	(0.022)
Age 50-59	**0.315*****	**0.339*****	**0.250*****	**0.283*****	**0.140*****
	(0.026)	(0.030)	(0.022)	(0.025)	(0.024)
Age >60	**0.580*****	**0.514*****	**0.370*****	**0.386*****	**0.163*****
	(0.044)	(0.051)	(0.040)	(0.043)	(0.043)
College (ref: Primary)	**0.767*****	**1.294*****	**0.649*****	**0.889*****	**0.389*****
	(0.020)	(0.063)	(0.128)	(0.053)	(0.132)
Secondary	**0.105*****	**0.361*****	**0.139****	**0.287*****	**0.096**
	(0.018)	(0.035)	(0.064)	(0.029)	(0.067)
Tenure	**0.023*****	**0.059*****	**0.041*****	**0.061*****	**0.041*****
	(0.002)	(0.005)	(0.009)	(0.004)	(0.009)
Tenure_sq	**-0.001*****	**-0.001*****	**-0.001*****	**-0.001*****	**-0.001*****
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Indefinite contract	**0.451*****	**0.531*****	**0.415*****	**0.468*****	**0.411*****
	(0.020)	(0.024)	(0.028)	(0.021)	(0.029)
Exporting firm	**0.119*****	**0.258*****		**0.275*****	
	(0.016)	(0.024)		(0.020)	
Firm Size: 50-199	**-0.108*****	**0.234*****		**0.307*****	
Ref: <50)	(0.018)	(0.044)		(0.037)	
Firm Size: >199	**-0.192*****	**0.351*****		**0.483*****	
	(0.017)	(0.067)		(0.056)	
Firm Agreement	**-0.179*****				
	(0.016)				
Inv. Mills Ratio		**1.541*****	**0.156**	**1.882*****	**0.354**
		(0.167)	(0.379)	(0.141)	(0.398)
Constant	**-1.065*****	**-4.123*****	**-1.471****	**-4.360*****	**-1.584****
	(0.028)	(0.333)	(0.601)	(0.279)	(0.629)
Observations	44249	195163	44249	44249	44249

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

(B) Women segregation in no PP jobs as a result of gender differences in risk attitudes

So far, the "monopsonistic" and "mommy track" hypotheses have been considered as competing plausible explanations for the gender PP gap. However the economic literature considers additional potential rationalizations of this gender wage gap related to the so-called "soft factors", comprising differences in attitudes, preferences and traits between men and women (for instance differences in competitiveness, negotiating skills, etc.). Among these alternative (complementary) determinants of the gender gap in PP, differences in attitudes towards risk between men and women appear as a particularly relevant factor. In effect, the economic experimental literature provides evidence in favor of gender differences in risk preferences (see, e.g., the surveys by Croson and Gneezy, 2009, and Eckel and Grossman, 2008) where women are portrayed as being more risk adverse than men.

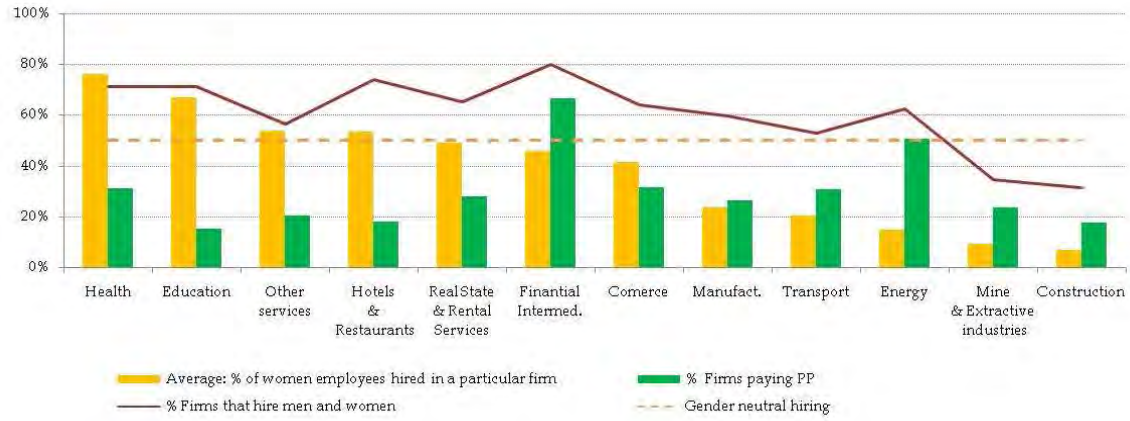
In our specific setup, risk-related factors underlying pay differences can be analyzed by considering the differences between base pay and PP, since the process determining the former component is very different from that determining PP compensation. The comparison of the gender gaps in these two pay components pay will allow us to analyze the extent to which the gender pay gap varies with the amount of risk in pay borne by the individual. The rationale for gender differences in variable pay is as follows. On the one hand, if firms consider that women are more likely to quit the firm than men due to family or other non-job- related reasons, they may find it optimal to offer women labour contracts where they bear the risk of career interruptions. In such a case, women will exhibit a higher proportion of PP in their overall pay. On the other hand, by contrast, if women are more risk averse than men, it is likely that they will select themselves into jobs and firms where pay is more stable. Hence, under this second hypothesis, women will exhibit a lower proportion of PP in their overall pay.

Figures A.1 to A.4 show how women and men are distributed across industries, according with the information available in EES 2006. Most of firms hire both men and women in all industries (red line). However, there are "*male industries*" such as mining and construction, where most firms hire exclusively men (the average percentage of women hired over total workers in these two sectors is 7% and 9% respectively). Likewise, it is interesting to stress that "*Female industries*" (i.e., those with higher feminization rates- average % of women over total workers in a firm – like Health, Education, Hotels & Restaurants and Other Services) are the ones which have a lower percentage of firms offering PP. Nevertheless, segregation of women in industries with less incidence of PP is not obvious from this Figure, since the Financial Intermediation sector, where the percentage of firms paying PP is highest, employs a nearly equal number of men and women (its feminization rate is around 46%).

Once we focus exclusively on firms paying PP that hire both male and female workers, we observe that women tend to concentrate more in those

industries with a lower proportion of firms paying PP. Additionally, the incidence of PP is higher among men than among women in all industries. When considering occupations, the same pattern emerges: the incidence of PP among women is higher in those occupations where the feminization index is lower (except in those "male" occupations, where both PP and feminization are low). Thus, these patterns are seemingly in line with the hypothesis related to gender differences in risk aversion, namely, being more risk averse than men, women tend to segregate into jobs where the incidence of PP is lower. However, once we restrict the analysis to men and women working in firms where PP is paid, we observe that the proportion of women receiving PP is higher than men´s, which would be consistent with our first hypothesis about firms offering women compensation schemes rewarding effort and penalizing quits.

Figure A.1 PP firms and feminization rates by sector. All firms in ESS 2006



Note: Industry Feminization rate is defined as the average ratio within a sector of the number of women in a firm/ Total employees in a firm. PP firms are defined as those firms that pay at least one of its workers PP.

Figure A.2 Performance Pay incidence and Feminization. Only firms hiring both men & women in ESS 2006.

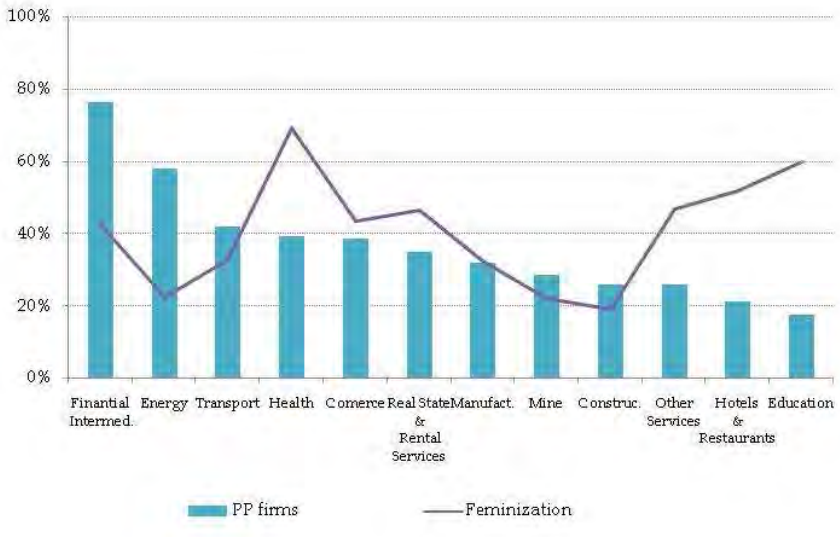


Figure A.3 Performance Pay incidence, by gender. Only firms hiring both men & women in ESS 2006

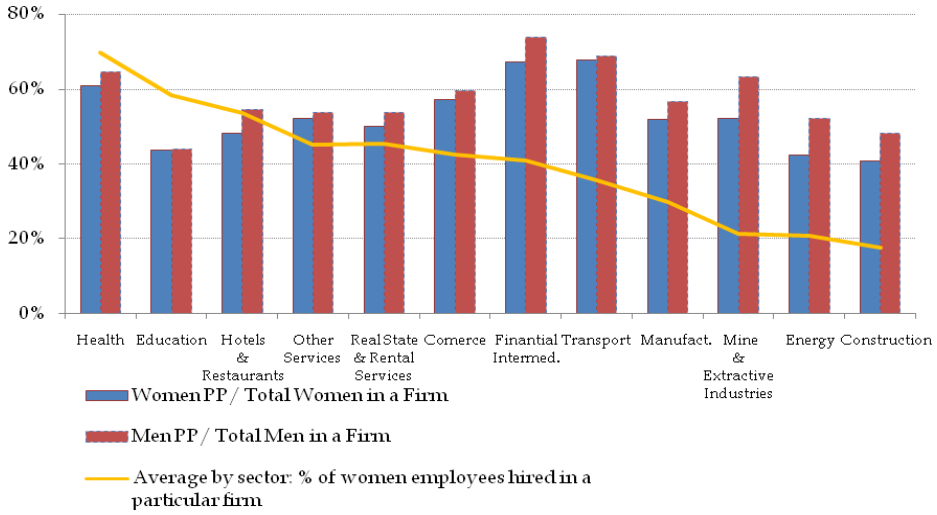


Figure A.4 Performance Pay incidence and feminization, by occupation. Only firms hiring both men & women in ESS 2006

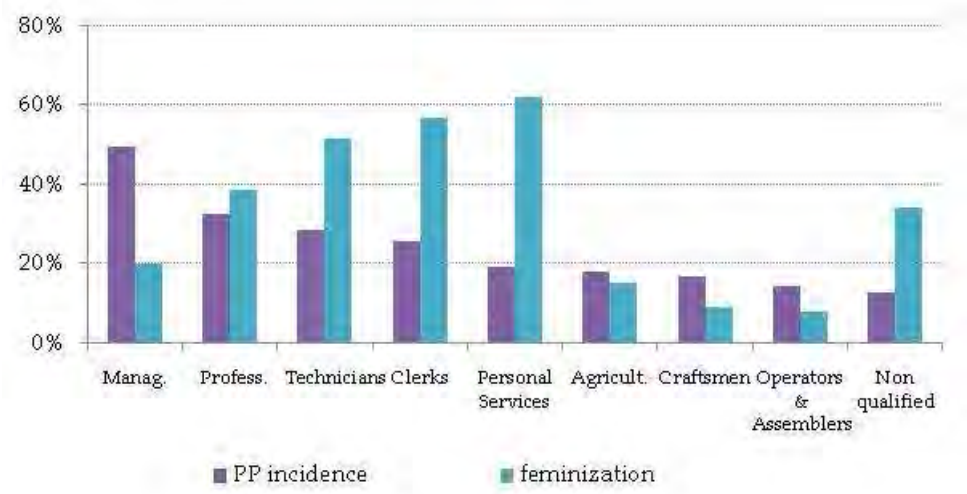
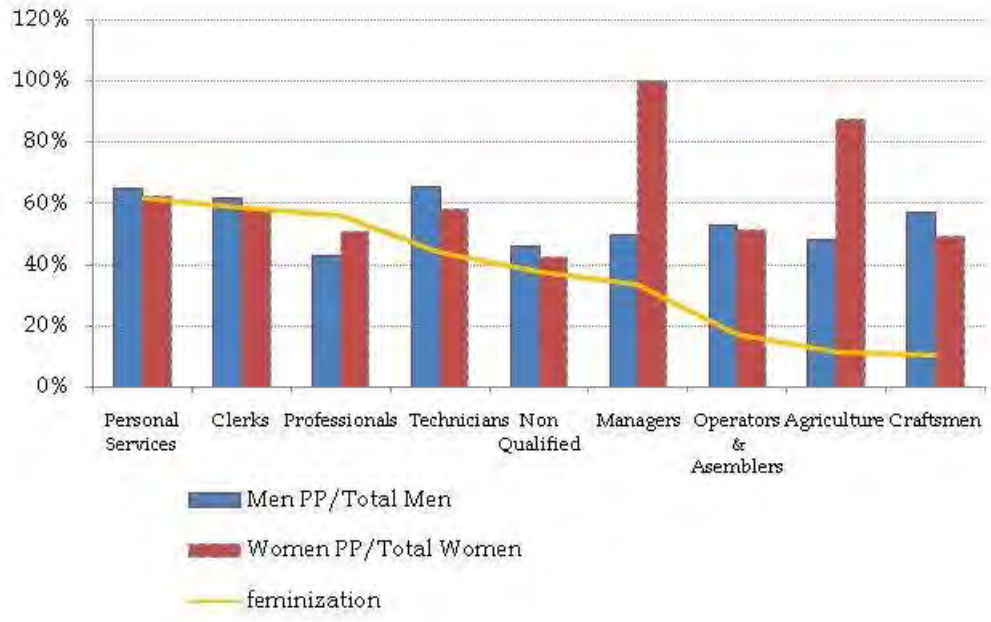


Figure A.5 Performance Pay incidence, by gender. Only firms hiring both men & women in ESS 2006



To shed some light on the above-mentioned arguments, we analyze how the gender pay gap varies across pay components for those individuals that receive PP within narrowly defined occupations. As described in section 2, we can distinguish among different types of payments: wage ordinary (base wage and other complements due to shifts, tenure, job risks, etc.) and non-ordinary (Fixed Annual Non-ordinary Payments and Variable Annual Non-ordinary Payments -performance pay-) components of annual gross earnings.

Specifically, in this section, we complement the QR estimates of the coefficient on the Female dummy throughout the PP distribution reported in Table 8 by providing equivalent QR estimates for Total hourly wage and Total hourly ordinary wage. These are presented in Tables A.7 and A.8, respectively. Notice that the Fixed Annual Non-ordinary Payments component of non-ordinary pay is omitted from this comparison since it is established at the collective bargaining level and therefore is bound to neither depend on workers´ or firms´ performance nor on differences in gender attitudes towards risk.

Table A.7 Adjusted Gender Wage Gaps – Quantile Regressions (with selection correction)

Dependent Variable: Log Hourly Total Wage

	(1) P10th	(2) P25th	(3) P50th	(4) P75th	(5) P90th
Female	-0.197***	-0.192***	-0.201***	-0.218***	-0.235***
(WFO)	(0.007)	(0.005)	(0.004)	(0.005)	(0.007)

Note: s.e´s. in parentheses. Estimations also control for the whole set of covariates (age, education, tenure, type of contract) as well as firm & occupational fixed effects (WFO).

Table A.8 Adjusted Gender Wage Gaps – Quantile Regressions (with selection correction)

Dependent Variable: Log Hourly ordinary Wage

	(1) P10th	(2) P25th	(3) P50th	(4) P75th	(5) P90th
Female	-0.086***	-0.126***	-0.148***	-0.175***	-0.199***
(WFO)	(0.005)	(0.004)	(0.005)	(0.007)	(0.009)

Note: s.e´s. in parentheses. Estimations also control for the whole set of covariates (age, education, tenure, type of contract) as well as firm & occupational fixed effects (WFO).

When considering the QR for the hourly ordinary wage, a statistically significant gender gap emerges at all percentiles (ranging from 8.6% in the 10th percentile to almost 20% in the 90th percentile). However, the gender gap in the distribution of hourly ordinary wages is quite lower than the gap in the distribution of PP hourly wage at all percentiles. The same holds when QR estimates of the coefficients of female dummies of the total hourly wage are compared with equivalent coefficients for the PP hourly wage.

Our results seemingly point out that, since women are less willing to bear risk than men, they tend to sort into occupations with more stable earnings. These occupations, due to compensating wage differentials in environments with risk-averse agents, tend to pay less on average, explaining in this fashion the lower gender gap observed in ordinary hourly wages. Moreover since women are more risk averse than men, they tend to negotiate wages with higher ordinary wages and lower PP, which would explain why the gender gap in total hourly wage is higher than the gender wage gap in ordinary wage which, in turn, is lower than the gender wage gap in PP hourly wages. Therefore, a preliminary conclusion from these results could be that differences in attitudes towards risk and attitudes towards negotiation cannot be discarded as additional explanatory factors behind the observed gender wage gaps.

(C) Misclassification of PP jobs. A sensitivity analysis

Table A.9 reports the proportions of non receivers across gender occupations and industries. Differences by gender are not noticeable when comparing these proportions by sector and occupation (by sector, the incidence of not receivers is higher among women in commerce transport and financial intermediation whereas, by occupation, the incidence is higher among men in clerks and agricultural workers)

A clearer pattern emerges when the probability of receiving PP by gender and education is compared with: i) the probability of working in a PP firm-job (that is within a firm and a job where at least one worker receives PP) and ii) the probability of being a non receiver. Table A.10 reports predicted probabilities by gender and educational attainment. Gender differences increase as the educational attainment increases, both for the probability of receiving PP and of working in a PP firm. Gender differences in the probability of being a non receiver has a U-shaped form when it is depicted by educational level. This indicates that: i) women are less likely to work in firms and jobs in which PP is paid (women segregation in jobs/firms), and ii) once they are working in PP-jobs firms, the probability of not receiving any PP is higher for them.

To check if our main conclusions in section 4.2 are altered when considering information about "non receivers", we perform the following sensitivity analysis: We re-estimate , (i) pooled (P), (ii) within- occupations (WO), (iii) within-firm (WF), and (iv) within-firm & occupation (WFO) regressions including in the sample of all workers (receiving or not PP) in firms and occupations where at least some one worker has obtained PP in 2006. Results in Table A.11 support once more the "monopsonistic" hypothesis against the conventional "occupational segregation" hypothesis.

A.9. Percentage of individuals working in PP jobs but who did not receive PP, by occupation and sector of activity

Male									
	Managers	Professionals	Technicians	Clerks	Personal Services	Agriculture	Craftsmen	Operators	Non Qualified
Mine	0.00%	5.90%	7.40%	0.00%	0.00%	50.00%	11.50%	0.00%	15.00%
Manufactures	6.10%	9.20%	15.40%	3.70%	2.90%	0.00%	7.30%	6.40%	4.30%
Energy	16.00%	14.90%	22.30%	7.70%	0.00%	0.00%	12.50%	13.50%	23.10%
Construction	5.00%	13.80%	12.90%	1.60%	5.30%	0.00%	13.50%	0.00%	2.90%
Commerce	9.40%	5.70%	16.80%	9.70%	12.40%	0.00%	3.40%	10.40%	6.60%
Hotels & Restaurants	6.70%	9.80%	5.70%	8.40%	7.40%	0.00%	3.80%	0.00%	6.90%
Transport	9.30%	14.50%	12.70%	11.00%	7.50%	-	14.80%	8.60%	7.10%
Financial Intermediation	14.30%	12.30%	22.00%	19.30%	0.00%	-	0.00%	0.00%	0.00%
Real State & Rental Services	8.80%	16.90%	17.70%	9.10%	9.90%	0.00%	20.40%	17.00%	7.90%
Education	1.70%	8.40%	10.10%	8.40%	1.60%	0.00%	6.30%	0.00%	7.30%
Health	0.00%	16.10%	6.90%	7.70%	5.70%	0.00%	0.00%	7.70%	3.00%
Other Services	4.20%	8.90%	6.50%	8.80%	4.20%	0.00%	3.00%	1.40%	6.90%
Female									
	Managers	Professionals	Technicians	Clerks	Personal Services	Agriculture	Craftsmen	Operators	Non Qualified
Mine	1.50%	6.50%	7.90%	5.30%	0.00%	0.00%	8.40%	4.40%	1.30%
Manufactures	4.80%	9.80%	12.10%	6.10%	1.70%	4.00%	6.40%	7.00%	3.20%
Energy	10.40%	16.70%	23.20%	9.10%	5.00%	0.00%	10.10%	12.40%	13.10%
Construction	3.00%	11.70%	8.90%	5.10%	4.40%	0.00%	8.50%	4.50%	3.30%
Commerce	7.20%	8.60%	15.10%	8.50%	11.60%	0.00%	6.30%	4.30%	5.70%
Hotels & Restaurants	1.40%	11.50%	7.90%	7.20%	8.90%	0.00%	3.80%	0.00%	4.70%
Transport	3.70%	7.10%	12.90%	7.50%	2.60%	0.00%	5.20%	4.80%	5.60%
Financial Intermediation	13.40%	10.50%	21.00%	19.30%	0.00%	-	0.00%	0.00%	0.00%
Real State & Rental Services	9.00%	18.90%	17.90%	15.40%	11.80%	3.20%	9.70%	12.80%	9.30%
Education	2.10%	10.00%	10.60%	7.80%	0.00%	0.00%	2.50%	0.00%	5.50%
Health	0.00%	15.00%	6.00%	15.90%	5.80%	0.00%	1.80%	3.60%	5.10%
Other Services	6.00%	6.00%	8.60%	10.80%	7.70%	8.20%	4.40%	4.80%	7.50%

Table A.10 Predicted probability of being a PP worker, Predicted probability of being working in a PP firm-job and Predicted probability of being a non receiver

	Predicted Probability of receiving PP		Predicted Probability of working in PP firms		Predicted Probability of being a non receiver	
	Men	Women	Men	Women	Men	Women
Primary	0.147	0.142	0.212	0.216	0.308	0.339
	(0.093)	(0.088)	(0.129)	(0.130)	(0.102)	(0.106)
Secondary	0.229	0.218	0.311	0.307	0.262	0.289
	(0.135)	(0.128)	(0.174)	(0.170)	(0.099)	-0.106
University	0.345	0.29	0.469	0.414	0.276	0.321
	(0.172)	(0.168)	(0.203)	(0.208)	(0.119)	(0.125)

Table A.11. Dependent Variable: log PP wage component

	(1) OLS	(2) Heckman selection	(3) Within Firm	(4) Within Occupations	(5) Within Firm-Occupation
Female	-0.407***	-0.425***	-0.320***	-0.387***	-0.207***
	(0.014)	(0.017)	(0.014)	(0.015)	(0.015)
Age 30-39 (ref:<30)	0.257***	0.370***	0.092**	0.303***	0.024
	(0.019)	(0.030)	(0.041)	(0.025)	(0.043)
Age 41-49	0.334***	0.429***	0.174***	0.346***	0.061
	(0.022)	(0.031)	(0.037)	(0.025)	(0.039)
Age 50-59	0.326***	0.396***	0.178***	0.328***	0.059*
	(0.026)	(0.033)	(0.032)	(0.027)	(0.034)
Age >60	0.601***	0.623***	0.363***	0.500***	0.159***
	(0.044)	(0.053)	(0.037)	(0.043)	(0.039)
College (ref: Primary)	0.793***	0.888***	0.505***	0.350***	0.174***
	(0.020)	(0.030)	(0.039)	(0.027)	(0.043)
Secondary	0.109***	0.200***	0.018	0.063***	-0.057
	(0.018)	(0.027)	(0.036)	(0.022)	(0.039)
Tenure	0.023***	0.054***	0.003	0.048***	-0.001
	(0.002)	(0.006)	(0.011)	(0.005)	(0.011)
Tenure square	-0.001***	-0.001***	0.000	-0.001***	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Permanent Contract	0.450***	0.592***	0.254***	0.507***	0.239***
	(0.020)	(0.034)	(0.052)	(0.028)	(0.054)
Firm Size: 50-199	-0.126***	-0.325***		-0.322***	
Ref: <50)	(0.018)	(0.037)		(0.031)	
Firm Size: >199	-0.209***	-0.495***		-0.480***	
	(0.017)	(0.046)		(0.038)	
Firm Agreement	-0.188***				
	(0.016)				
Exporting firm	0.114***	0.153***		0.142***	
	(0.016)	(0.020)		(0.016)	
Inv. Mills Ratio		1.546***	-1.666***	1.506***	-1.690***
		(0.263)	(0.520)	(0.223)	(0.543)
Constant	-1.107***	-2.098***	-0.025	-1.612***	0.208
	(0.028)	(0.169)	(0.385)	(0.143)	(0.402)
No. Obs.	44249	61615	44249	44249	44249
R-squared	0.165	0.166	0.109	0.086	0.065

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Figures and Tables (Main text)

Figure 1. Gender wage gaps (Total, Non-PP and PP components)

- Whole Sample (in percent)-

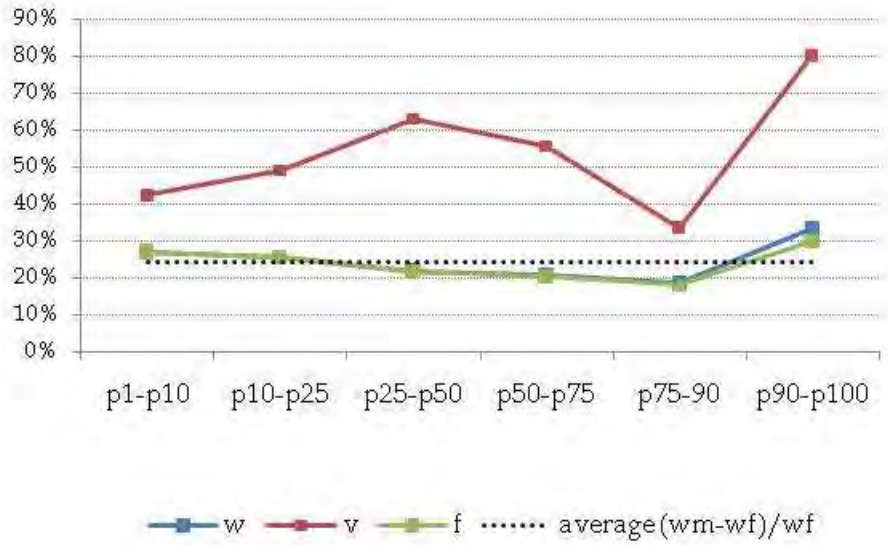
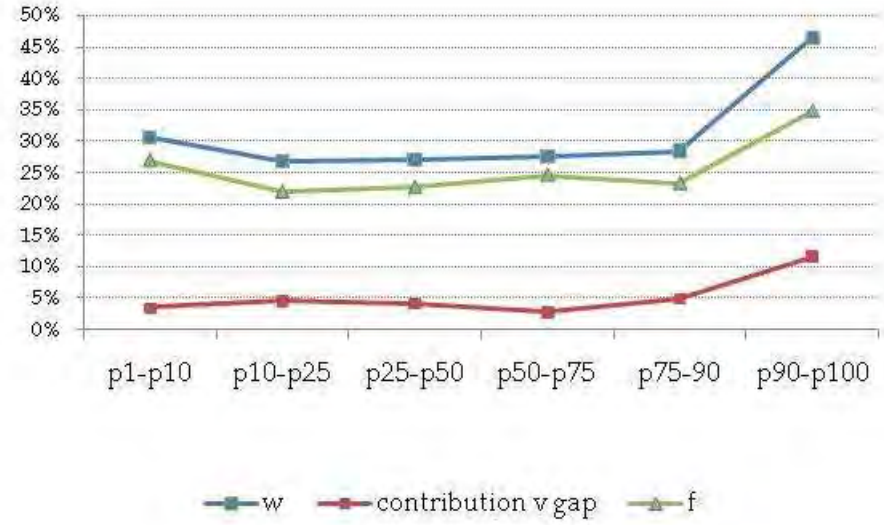


Figure 2. Gender wage gaps and the contribution of PP component

- Sample of PP workers (in percent)-



Note: Graphs depicted in Figure 2 stand for the decomposition of the gender gap described in Appendix 1.

Table 1 – Sample characteristics (Full-time workers between 18-65 years)

Variables	Women (65,233)	Men (129,930)	Ho: $\mu^W = \mu^M$
	Mean	Mean	p-value test
Individual Characteristics			
Education			
Primary or less	0.176	0.275	0.000
Secondary	0.508	0.545	0.000
University	0.316	0.180	0.000
Age			
Less 30 years	0.257	0.200	0.000
31-40	0.354	0.323	0.000
41-50	0.245	0.265	0.000
>50	0.143	0.212	0.000
Tenure (years)	7.41	8.867	0.000
Permanent Contract	0.727	0.768	0.000
Wages			
Total Hourly Wage (logs)	2.185	2.391	0.000
Performance Pay			
% PP job	0.227	0.227	0.000
Total Hourly PP (only PP workers)	1.127	1.958	0.000
Firm Characteristics			
Size			
<50 workers	0.339	0.403	0.000
51-200 workers	0.265	0.288	0.000
>200 workers	0.396	0.309	0.000
Firm Bargaining Agreement.	0.163	0.194	0.000
Exporting firms	0.163	0.194	0.000

Source: EES (2006)

Note: The hypothesis of equal means (proportions) for men and women is tested. P-values of the tests are reported in third column. The equal variance hypothesis is rejected for tenure, total hourly wages and total hourly PP. Differences are statistically significant for all variables considered

Table 2a: Characteristics of workers and firms by type of job and gender

	Performance Pay Job			Non-Performance Pay Jobs		
	Women	Men	Ho:	Women	Men	Ho:
	(14.789 obs)	(29.460 obs)	$\mu^{WPP}= \mu^{MPP}$	(50.444 obs)	(100.470 obs)	$\mu^{WPP}= \mu^{MPP}$
Variables	Mean	Mean	p-value $\mu^{WPP}\neq\mu^{MPP}$	Mean	Mean	p-value $\mu^{WPP}\neq\mu^{MPP}$
Education						
Primary or less	0.107	0.178	0.000	0.196	0.304	0.000
Secondary	0.494	0.545	0.000	0.512	0.545	0.000
University	0.399	0.277	0.000	0.292	0.151	0.000
Age						
Less 30 years	0.204	0.149	0.000	0.273	0.215	0.000
31-40	0.381	0.313	0.000	0.346	0.325	0.000
41-50	0.265	0.294	0.000	0.239	0.257	0.000
>50	0.150	0.244	0.000	0.141	0.203	0.000
Tenure (years)	9.281	12.037	0.000	6.861	7.938	0.000
Permanent Contract	0.814	0.862	0.000	0.741	0.741	0.000
Firm Characteristics						
Size						
<50 workers	0.201	0.235	0.000	0.380	0.452	0.000
51-200 workers	0.239	0.297	0.000	0.272	0.285	0.000
>200 workers	0.560	0.467	0.000	0.348	0.262	0.000
Collective Bargaining: ref: Industry level						
Firm Collective Bargaining	0.193	0.288	0.000	0.154	0.167	0.000
Firm Market (ref: International Market)						
Local or Nat. Market	0.181	0.239	0.000	0.135	0.153	0.000

Note: The hypothesis of equal means (proportions) for men and women is tested. P-values of the tests are reported in third column. The equal variance hypothesis is rejected for tenure, total hourly wages and total hourly PP both in PP workers and no PP workers. Differences are statistically significant for all variables considered

Table 2b: Incidence of PP by industry and occupation

	Mean	Std. Dev.	No. Obs.
Industries			
Financial Intermediation	0.598	0.49	10475
Energy	0.324	0.468	4627
Transportation	0.324	0.468	12710
Health	0.287	0.452	14178
Retail trade	0.241	0.427	17131
Manufactures	0.205	0.404	74332
Real State and Res. Serv.	0.194	0.395	16342
Mine & Extractive Ind.	0.188	0.391	2919
Other Services	0.146	0.353	9040
Construction	0.127	0.333	17096
Hotels and Restaurants	0.123	0.328	8315
Education	0.092	0.289	7998
Occupations			
Managers	0.497	0.5	6190
Technicians	0.326	0.469	30184
Professionals	0.288	0.453	20295
Clerks	0.257	0.437	24761
Personal Services	0.196	0.397	17528
Operators and Assemblers	0.18	0.384	34822
Craftsmen	0.169	0.375	37918
Agriculture and Fisheries	0.146	0.353	542
Laborers, non-qualified operators	0.127	0.333	22923

Table 2c: Share of women throughout PP distribution

	% Women in percentile
[p^{1th}-p^{10th}]	40.6%
[p^{11th}-p^{25th}]	40.3%
[p^{26th}-p^{50th}]	39.5%
[p^{51th}-p^{75th}]	32.0%
[p^{76th}-p^{90th}]	24.8%
[p^{91th}-p^{95th}]	18.6%
[p^{95th}-p^{100th}]	15.9%

Table 3a: Hourly Wages in PP and Non-PP Jobs

	Pay Performance Jobs				Non-Performance Pay Jobs	
	Women		Men		Women	Men
	Total Hourly Wage(€)	Ratio PP/Total Wage (%)	Total Hourly Wage(€)	Ratio PP/Total Wage(%)	Total Hourly Wage(€)	Total Hourly Wage(€)
Average	14.503	7.164	19.144	9.012	9.678	11.665
P10th	6.060	0.976	7.801	0.009	3.721	4.689
P25th	8.577	2.087	10.804	0.025	5.884	7.308
P50th	12.479	4.657	16.051	0.060	8.126	9.826
P75th	18.800	9.491	23.546	0.127	12.048	14.192
P90th	24.842	16.684	33.127	0.217	17.795	20.162

Table 3b: Hourly wage components for workers receiving PP (in €)

	Women		Men	
	PP wage	Other wage components	PP wage	Other wage components
Average	1.127	13.376	1.958	17.186
P10th	0.102	5.648	0.132	7.031
P25th	0.225	7.850	0.324	9.675
P50th	0.539	11.679	0.89	14.758
P75th	1.311	17.509	2.198	21.694
P90th	2.594	22.828	4.493	29.597

Table 4: Log hourly wage regressions

Dependent Variable: Log Hourly Total Wage

	(1) PP Workers	(2) Non-PP Workers	(3) Pooled
PP Job			0.208***
			(0.009)
Female	-0.223***	-0.212***	-0.219***
	(0.004)	(0.003)	(0.004)
Age 30-39 (ref:<30)	0.139***	0.098***	0.095***
	(0.006)	(0.003)	(0.004)
Age 41-49	0.199***	0.116***	0.114***
	(0.007)	(0.004)	(0.004)
Age 50-59	0.227***	0.161***	0.161***
	(0.008)	(0.004)	(0.005)
Age >60	0.262***	0.155***	0.158***
	(0.014)	(0.007)	(0.008)
College	0.277***	0.223***	0.215***
(ref: Primary)	(0.007)	(0.005)	(0.004)
Secondary	0.077***	0.063***	0.060***
	(0.006)	(0.003)	(0.003)
Tenure	0.044***	0.042***	0.043***
	(0.001)	(0.000)	(0.000)
Tenure sq.	-0.001***	-0.001***	-0.001***
	(0.001)	(0.000)	(0.000)
Permanent Contract	0.282***	0.313***	0.312***
	(0.006)	(0.003)	(0.003)
Firm Size: 50-199	0.067***	0.095***	0.094***
(Ref: <50)	(0.005)	(0.003)	(0.003)
Firm Size: >199	0.118***	0.166***	0.164***
	(0.005)	(0.003)	(0.003)
Firm Agreement	0.011	0.015*	0.013*
	(0.008)	(0.008)	(0.007)
Export market	0.027***	0.035***	0.045***
	(0.005)	(0.003)	(0.003)
Interactions with PP			
Female*PP			-0.007
			(0.005)
Age 30-39*PP (ref:<30)			-0.047***
			(0.007)
Age 41-49*PP			-0.029***
			(0.008)
Age 50-59*PP			0.072***
			(0.010)
Age >60*PP			0.101***
			(0.016)
College*PP			0.089***
(ref: Primary)			-0.007
Secondary*PP			0.030***
			(0.006)
Tenure*PP			0.011***
			(0.004)

Permanent Contract*PP			-0.025***
			(0.007)
Firm Size: (ref<50)			-0.027***
50-199*PP			(0.006)
Firm Size: >199*PP			-0.042***
			(0.006)
Firm Agreement*PP			-0.006**
			(0.003)
Export. Firm *PP			-0.014***
			(0.006)
No. Observations	44249	150914	195163
R sq.	0.605	0.511	0.573

Note: S.e´s. in parentheses. Estimations also control for industry dummies (11) and occupational dummies (8).

Table 5: Probit estimation

Dependent Variable: Receiving Performance Pay (1/0)

	(1)	(2)	(3)
Female	-0.047***	-0.103***	-0.066***
	(0.008)	(0.016)	(0.022)
Age 30-39 (ref:<30)	0.052***	0.016**	0.032***
	(0.010)	(0.008)	(0.012)
Age 40-49	0.032***	0.002	0.018
	(0.011)	(0.013)	(0.013)
Age 50-59	0.015	-0.01	-0.010
	(0.013)	(0.015)	(0.015)
Age >60	-0.076***	-0.099***	-0.137***
	(0.023)	(0.026)	(0.025)
University	0.260***	0.262***	0.543***
(ref: Primary)	(0.013)	(0.013)	(0.012)
Secondary	0.164***	0.161***	0.238***
	(0.009)	(0.009)	(0.010)
Tenure	0.030***	0.031***	0.032***
	(0.001)	(0.001)	(0.001)
Tenure square	-0.001***	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)
Permanent Contract	0.037***	0.035***	0.152***
	(0.010)	(0.010)	(0.009)
Firm Size: 50-199	0.295***	0.301***	0.287***
(Ref: <50)	(0.009)	(0.009)	(0.009)
Firm Size: >199	0.485***	0.478***	0.532***
	(0.008)	(0.009)	(0.008)
Firm Collective Agreement	0.096**	0.120***	0.095***
	(0.009)	(0.009)	(0.008)
Exporting firm	0.122***	0.120***	0.097***
	(0.009)	(0.011)	(0.009)
Interactions with female			
Age 30-39*Female		-0.047***	-0.048***
Ref:<30)		(0.019)	(0.019)
Age 40-49*Female		-0.031***	-0.032***
		(0.011)	(0.012)
Age 50-59*Female		0.072***	0.074***
		(0.024)	(0.024)
Age >60*Female		0.062	0.064
		(0.052)	(0.051)
University * Female (ref:			-0.084***
Primary)			(0.022)
Secondary * Female			0.059***
			(0.020)
No. Observations	195163	195163	195163
Pseudo R^2	0.111	0.123	0.110

Note: S.e´s. in parentheses

Table 6: Estimates of alternative specifications of log PP hourly wage equation corrected for selectivity (PP vs. all workers)

Dependent Variable: log PP hourly wage component

	(1)	(2)	(3)	(4)	(5)
	OLS	Heckman selection	Within Firm	Within Occupations	Within Firm-Occupation
Female	-0.407***	-0.453***	-0.343***	-0.434***	-0.290***
	(0.014)	(0.018)	(0.011)	(0.019)	(0.011)
Age 30-39 (ref:<30)	0.257***	0.337***	0.181***	0.281***	0.143***
	(0.019)	(0.024)	(0.018)	(0.024)	(0.017)
Age 41-49	0.334***	0.381***	0.252***	0.313***	0.192***
	(0.022)	(0.026)	(0.019)	(0.027)	(0.019)
Age 50-59	0.326***	0.353***	0.238***	0.298***	0.184***
	(0.026)	(0.031)	(0.022)	(0.031)	(0.021)
Age >60	0.601***	0.532***	0.416***	0.395***	0.311***
	(0.044)	(0.052)	(0.037)	(0.053)	(0.036)
College (ref: Primary)	0.793***	0.350***	0.362***	0.636***	0.280***
	(0.020)	(0.064)	(0.128)	(0.053)	(0.062)
Secondary	0.109***	0.380***	0.010**	0.213***	0.043
	(0.018)	(0.035)	(0.064)	(0.029)	(0.035)
Tenure	0.023***	0.061***	0.019***	0.060***	0.0173***
	(0.002)	(0.005)	(0.005)	(0.005)	(0.006)
Tenure square	-0.001***	-0.001***	-0.001***	-0.001***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Permanent Contract	0.450***	0.534***	0.389***	0.443***	0.362***
	(0.020)	(0.024)	(0.021)	(0.024)	(0.021)
Firm Size: 50-199	-0.126***	0.236***		0.317***	
Ref: <50)	(0.018)	(0.045)		(0.048)	
Firm Size: >199	-0.209***	0.366***		0.505***	
	(0.017)	(0.068)		(0.0726)	
Firm Agreement	-0.019				
	(0.016)				
Exporting firm	0.114***	0.262***		0.263***	
	(0.016)	(0.024)		(0.023)	
Inv. Mills Ratio		1.628***	1.693***	1.984***	1.513 ***
		(0.170)	(0.198)	(0.141)	(0.474)
No. Obs.	44249	195163	44249	44249	44249
R sq.	0.186	0.175	0.115	0.089	0.125

Note: S.e´s. in parentheses. Coefficients in (1) are derived from an OLS regression over the overall sample of workers. Coefficients in (2) are derived from Heckman estimation, performed to correct for selection into PP jobs. Coefficients in (3) to (5) correspond to within firm, within occupations and within firm and occupations fixed effect estimation, respectively. Inverse Mills ratio derived from estimates in (2) are included in the three last columns as an additional covariate to correct for selectivitys. Estimations in (1) and (2) also control for industry dummies (11) and occupational dummies (8).

Table 7. Estimates of alternative specifications of log PP hourly wage equation corrected for selectivity (PP vs. PP entitled workers)

Dependent Variable: log PP hourly wage component

	OLS	(2) Heckman selection	(3) Within Firm	(4) Within Occupations	(5) Within Firm-Occupation
Female		-0.425***	-0.320***	-0.387***	-0.207***
		(0.017)	(0.014)	(0.015)	(0.015)
Age 30-39 (ref:<30)	0.257***	0.370***	0.092**	0.303***	0.024
	(0.019)	(0.030)	(0.041)	(0.025)	(0.043)
Age 41-49	0.334***	0.429***	0.174***	0.346***	0.061
	(0.022)	(0.031)	(0.037)	(0.025)	(0.039)
Age 50-59	0.326***	0.396***	0.178***	0.328***	0.059*
	(0.026)	(0.033)	(0.032)	(0.027)	(0.034)
Age >60	0.601***	0.623***	0.363***	0.500***	0.159***
	(0.044)	(0.053)	(0.037)	(0.043)	(0.039)
College (ref: Primary)	0.793***	0.888***	0.505***	0.350***	0.174***
	(0.020)	(0.030)	(0.039)	(0.027)	(0.043)
Secondary	0.109***	0.200***	0.018	0.063***	-0.057
	(0.018)	(0.027)	(0.036)	(0.022)	(0.039)
Tenure	0.023***	0.054***	0.003	0.048***	-0.001
	(0.002)	(0.006)	(0.011)	(0.005)	(0.011)
Tenure square	-0.001***	-0.001***	0.000	-0.001***	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Permanent Contract	0.450***	0.592***	0.254***	0.507***	0.239***
	(0.020)	(0.034)	(0.052)	(0.028)	(0.054)
Firm Size: 50-199	-0.126***	-0.325***		-0.322***	
Ref: <50)	(0.018)	(0.037)		(0.031)	
Firm Size: >199	-0.209***	-0.495***		-0.480***	
	(0.017)	(0.046)		(0.038)	
Firm Agreement	-0.188***				
	(0.016)				
Exporting firm	0.114***	0.153***		0.142***	
	(0.016)	(0.020)		(0.016)	
Inv. Mills Ratio		1.546***	-1.666***	1.506***	-1.690***
		(0.263)	(0.520)	(0.223)	(0.543)
Constant	-1.107***	-2.098***	-0.025	-1.612***	0.208
	(0.028)	(0.169)	(0.385)	(0.143)	(0.402)
No. Obs.	44249	61615	44249	44249	44249
R-squared	0.165	0.166	0.109	0.086	0.065

Note: S.e´s. in parentheses. Coefficients in (1) are derived from an OLS regression over the overall sample of workers. Coefficients in (2) are derived from Heckman estimation, performed to correct for selection into PP jobs from PP entitled jobs. Coefficients in (3) to (5) come from within firm, within occupations and within firm and occupations fixed effect estimation respectively. Inverse Mills ratio derived from probit estimates of the probability of receiving PP among the subsample of PP entitled workers are included in the three last columns as an additional covariate to correct for selectivity. Estimations in (1) and (2) also control for industry dummies (11) and occupational dummies (8).

Table 8: Adjusted Gender Wage Gaps – Quantile Regressions (with selection correction)

Dependent Variable: Log PP Hourly Wage

	(1) **P10th**	(2) **P25th**	(3) **P50th**	(4) **P75th**	(5) **P90th**
Female (WFO)	-0.267**	-0.378***	-0.498***	-0.545***	-0.548***
	(0.022)	(0.020)	(0.014)	(0.016)	(0.019)

Note: s.e´s. in parentheses. Estimations also control for the whole set of covariates (age, education, tenure, type of contract) as well as firm & occupational fixed effects (WFO).

Table 9: Oaxaca Decomposition of (log) hourly Gap in PP

Unadjusted Gender Wage Gap in PP: 46 log-points

Variables	Absolute [relative] Contribution of Diff. in *Characteristics* $(X_m-X_f)*\beta_m$	Absolute [relative] Contribution of Diff. in *Returns* $(\beta_m-\beta_f)*X_f$
Sample Selection	0.62	-1.56
	[1.34%]	[-3.39%]
Tenure	3.44	-1.93
	[7.47%]	[-4.19%]
Education	-2.82	3.83
	[-6.13%]	[8.33%]
Age	1.35	8.46
	[2.93%]	[18.39%]
Type of Contract	1.02	2.42
	[2.22%]	[5.26%]
Occup. and Firm Effects	1.81	3.36
	[3.93%]	[7.30%]
Constant		26.0
		[56.52%]
Total	5.42	40.58
	[11.8%]	[88.2%]

Note: Decomposition based on WFO estimation separately for each gender. Similar qualitative results are found with a joint WFO estimation for men and women allowing for different returns by gender.

EVALUATING THE INCENTIVES TO EARLY RETIREMENT IN SPAIN*

Raquel Vegas Sánchez *

(*) Universidad Carlos III & FEDEA

Abstract

In this paper I analyze the influence that incentives play in the timing of the transition to retirement in Spain. I use the information contained in a recent new data source about the Spanish labour market, Continuous Sample of Working Histories 2006 (CSWH "Muestra Continua de Vidas Laborales", in Spanish) to construct incentive measures stemming from the Social Security provisions in relation to retiring at old age as well as to investigate the role played by such incentives and other socio-economic variables on the retirement hazard. I compute the effects of the pension system reform that took place in 2002, which made stricter the requirements to access an old pension. By means of a duration model, I carry out a dynamic reduced-form analysis of the retirement decisions of men aged between 60 and 70 years in 2006 who are entitled to receive retirement benefits. The results show that both the pension wealth and the substitution effects play a significant role in retirement decisions, but that, after the reform, the latter effects become less important.

JEL Classification: J14, J26

Keywords: older workers employment, retirement, public pensions

*Corresponding author: Raquel Vegas Sánchez. FEDEA C/Jorge Juan, 46. 28001 Madrid, Spain Tph: +34 914 359 209 Fax: +34 91 577 9575 E-Mail: rvegas@fedea.es

1 Introduction

The sustainability of pay-as-you-go systems is a matter of growing concern in ageing economies, such as Spain, since they were designed at a time when demographic structures exhibited a much lower life expectancy and higher fertility rates[1]. Therefore, both the number of people and the number of years that these people were alive to receive a public pension was much lower when the system was launched than nowadays[2] , cf. Herce et al (2009).

On top of such demographic developments, in recent years there has been a decline in the labour force participation of older workers. While in the seventies, the participation rates of males over 55 exceeded 50%, in 2008 it does not reach 30%. The combination of both factors has led to the prediction that the old age dependency ratio, which currently is 24.1% in Spain, will reach 59.1% in 2060[3]. Hence, a smaller proportion of people than what is currently the case will be providing the revenues to be transferred to older people in the form of pensions. In this sense, the behaviour of older workers is reinforcing the negative impact that demographic factors have on the sustainability of the pension system.

The role of pension benefits rules in relation to labour market participation of the elderly is regarded as central in many countries, as not only they may be too generous in providing income support, but they may also create incentives to early retirement. In this sense, three issues are relevant here: the amount of the pension that the system provides, the pattern of benefits associated with each age of retirement, and the entitlement rights that define the conditions to be met to be able to claim a pension at each age. Parametric changes for the Social Security system are discussed in Spain under the so-called Toledo Agreement, while some countries have already implemented large reforms[4].

Many changes have been directed to reducing incentives to early retirement embedded in the pension system and to increase incentives to leave the labour market at a later age. In particular, up to 2002, pension regulations implied total or substantial withdrawal from any form of employment requiring affiliation to the Social Security System to be able to receive an old-age pension. In 2002, partial retirement was regulated, so that employment and old-age pensions could be simultaneously enjoyed, while the mandatory retirement age at 65 was effectively abolished and bonuses on pension benefits were introduced for those workers than extend their labour market participation beyond 65. In 2006, additional requirements were imposed on

[1] The forecast that 32.3% of total Spanish population will be older than 65 in 2060 is mostly driven by the decline in fertility rates (European Commission (2008)).

[2] In fact, life expectancy at 65 for males was around 12 in the middle of the 20th century; it has grown up to 17 in 2008, and it is forecasted to increase to 22 years in the middle of this century, according to Eurostat, EUROPOP 2008 projections.

[3] Ratio calculated as 65+/(15-64). European Commission (2008)

[4] See Whitehouse et al.,2009 for a recent review of pension reforms in industrialized countries

partial retirement restricting it to workers older than 61 and with more than 6 years of seniority in the firm. Different governments have also promoted active labour force policies that should stimulate the demand for elderly workers, thus contributing to raising labour market participation of the elderly. Therefore, the effects of pensions on retirement compound labour supply and demand forces, so that their magnitude is an empirical issue.

The goal of this paper is to quantify the role that Social Security provisions for old-age pensions play in the retirement decision of Spanish workers. In previous work, there is already an assessment about the impact of individual demographic characteristics on the likelihood of retirement at each age under the Spanish retirement legislation but without considering economic incentives, cf. Argimón, González and Vegas (2007a). Here, I propose to fill this gap by estimating a reduced-form model for retirement in order to capture the effects of pension incentives on the timing of the pension claim. In particular, I analyse the probability of retiring at a given age, given that the person has not retired yet, as a function of individual demographic and economic incentive variables underlying the retirement legislation. Thus, I compare the retirement rates of workers with different levels of social security pension wealth at different ages, holding demographic variables constant, since decisions on retirement can be made within a certain age interval.

I contribute to the large body of research that examines the economic incentives created by the pension system in Spain, by adopting a different methodological approach to that traditionally followed in the literature (see, e.g., Boldrin, Jimenez-Martín & Peracchi, 1999, 2004, Jiménez-Martín, 2006, and Moral-Arce et al., 2009). First, instead of using a point in time estimation, I analyse retirement patterns and the effects of economic incentives on the probability of retirement by means of a logistic specification of a duration model. This approach allows to analyze the role played at each age by incentives stemming from social security rules for early retirement. In this fashion, the longitudinal information contained in Continuous Sample of Working Histories (*CSWH, "Muestra Continua de Vidas Laborales", MCVL in Spanish*), wave 2006, a recently released database is fully exploited in this paper[5]. This database is particularly useful since it contains detailed information registered in Spanish Social Security records and in the Census, that allows to take time dependence into account, since workers' decisions are tracked over time. The advantage of using CSWH relative to other databases used in some previous empirical work[6], is that its sample design is publicly known and, therefore, it allows for a better and broader understanding of my main results. I regard a duration model as an improvement over other straightforward binary dependent regression models since survival analysis is defined in terms of three attributes which are relevant in the context of

5 Moral-Arce et al (2009) also use the CSWH to evaluate the economic incentives generated by the Social Security System in Spain, but, unlike my work, they carry out such an analysis restricted to a given point in time, 2007.

6 A previous sample of a similar database with information up to 1995 and whose description can be found in Martínez (1999), has been used in Boldrin et al (1999, 2004), Boldrin and Jiménez-Martín (2002),and Jimenez-Martín et al. (2000, 2006). The main drawbacks of such sample are that the detailed sample design has not been disclosed and that it is not publicly available.

retirement decisions: (1) individuals evolve along a finite number of states, (2) decision changes may occur at any time point, not just at a given point in time, and (3) factors that have an effect on the event analyzed could be either time varying or constant over time. Point-in-time estimates do not allow considering differences in the length of time in which each person is at risk of experiencing the transition, namely, time dependence. These types of models ignore the structure of the dynamic optimization problem that underlies the workers decisions when they retire. Taking into account the inter-temporal correlations in the panel appears to be very important. Boldrin et al (1999, 2004) and Jiménez-Martín (2006) use a binary dependent model to estimate retirement probability. This procedure relies on the assumption that, conditional on covariates, the probability of the duration until a transition is adopted (retirement) does not change over time given that the units of observation have survived to time *t* (the period since the individual is eligible to retire). This technical limitation of point-in-time binary specifications can be overcome by considering that a transition occurred within some pre-specified interval of time. However, a large amount of information contained in the panel correlations is still lost. By including age dummies it would be possible to get some information about how retirement probability increases with age, but other individual panel correlations would not be taken into account. Additional problems would arise when time-varying covariates would be considered in the estimation. Instead of assuming that the hazard rate (that is, the instantaneous rate at which workers retire) is flat (in other words, the duration dependence does not exist) survival analysis enables to use serial correlation - duration dependence- in the estimation, which leads to an unbiased estimation of coefficients when hazard rate depends on survival time.

The second contribution of this paper is to quantify the effect on retirement decisions of the Retirement Amendment that took place in 2001 in Spain (*Acuerdo para la Mejora y el Desarrollo del Sistema de Protección Social, April 2001*) and that was fully implemented in 2002 (Law 35/2002, July). The reform implied a shift towards a system closer to an individually capitalized one in order to both increase labour force participation of older workers and foster retirement beyond 65.

Specifically, in this paper I focus on the impact of the reform on retirement decisions that imply a full withdrawal from the labour market. To do so, I select a sample of men aged from 60 to 70 years old in 2006, contributing to the Social Security General Scheme and entitled to receive retirement benefits. The sample period spans 10 years, from 1996 to 2006.

During this period, there was also another important pension reform in 1997 which increased the number of years of contribution applied to compute the regulatory base (from 8 to 15), established the revaluation of retirement benefits according with CPI and created a reserve fund to pay pensions. Subsequently, the 2002 reform was primarily aimed to the establishment of a system of gradual and flexible retirement and to encourage an increase in older workers participation rates.

Next, for convenience, I provide a brief outline of the changes implemented by the 2002 pension reform. Further details are given in Appendix 1.

- After 2002 pension reform, access to retirement was possible even if the employee has not reached the statutory retirement age. To take early retirement, the worker must have reached the age 61 (those who worked before 1967 are entitled to take early retirement at 60). Additionally, the worker must show a minimum of 30 years of contribution to Social Security (at least 2 within the last 15), and his last employment termination should be involuntary.

- The 2002 Reform modified further the early retirement penalty increasing the linkages between the contributive effort exerted by workers and the pension received. After the 2002 Amendment, workers who retire before 65 and after 61 years old are charged with a penalty that decreases with age and total years of contributions.

  Each full year of employment beyond statutory retirement age (65), implies a 2% increase in regulatory base to compute retirement benefits, such that those workers older than 65 who continue working and who can show a minimum of 35 years of contributions can receive a pension that exceeds 100% of the regulatory base value. It becomes possible to receive a partial pension from the age of 60 whilst continuing to work part-time. The corresponding pension is calculated by applying the rate of the reduction of working hours to the amount of pension which the individual would normally received given his contributions history at the time of retirement.

- No compulsory termination of unemployment benefits for workers older than 52 when they get the requirements needed to claim retirement benefits. Firms receive payment exemptions from Social Security contributions when workers are older than 65 years old and have contributed during 35 years to Social Security. At the same time, those firms hiring workers older than 60 using indefinite-term contracts receive a 50% bonus that increases by 10% each year, to reach 100% of the amount of common contingencies contributions when workers reach 5 years seniority in the company.

- Firms that dismiss older workers are punished: they must pay *Special Agreements* to finance contributions to the workers who have been massively dismissed (until the workers reach 61) provided the company is not bankrupt. From then onwards, the worker must necessarily pay his own contributions upon the *Special Agreement* until he decide to retire.

The rest of the paper is organized as follows. Section 2 contains a literature survey of the main determinants of retirement for elderly workers, placing special emphasis on those related to the Spanish pension system. Section 3 describes the dataset used in the empirical analysis. Section 4 contains

a detailed discussion on the estimated empirical model whereas Section 5 summarises the main results. Section 6 contains the conclusions. Additionally, three appendixes are included in the paper. Appendix 1 contains a brief description of the variables used in the empirical analysis. Appendix 2 provides a short explanation about the main features of the Spanish Social Security pension system, and Appendix 3 includes some robustness checks and additional comments to further clarify the analysis. Specifically, issues such as the possible biases stemming from the existence of people whose pensions are capped, some tax considerations, and the potential endogeneity of some of the control variables are properly addressed in the third Appendix.

2 An overview of empirical evidence on the role of Social Security on retirement decisions.

Population ageing is a matter of growing concern in modern developed countries, where the increase in population life expectancy, the decrease in fertility rates and the delay in the average age of entry of young people to the labour market, have jeopardized the sustainability of public pensions systems. These demographic phenomena have fostered the development of a vast economic literature both in the U.S. and Europe -and generally in most developed countries- concerning retirement during the last 30 years. The interest on this topic has lead to the publication of more than one thousand papers concerning retirement and pension related issues during the last decades (for a good survey, see Smith, 2004).

Empirical analyses of retirement behaviour are limited by the data available for estimation and also by the complexity of the retirement decision. As a result, most econometric studies report estimates obtained from reduced-form models. Nonetheless, since this type of models are subject to the Lucas' critique and also to potential biases arising from either omission of relevant variables or to individual working histories being correlated with changes in public pension schemes (see, e.g., Lumsdaine, Stock, Wise, 1992), there has been a shift to structural models in response to this shortcomings. However, structural models do come with their own limitations since, due to their complexity, their estimation has usually been restricted to model the retirement decision as a binary (yes/no) option[7].

Conducting a comprehensive literature review on retirement issues is beyond the scope of this paper. The next paragraphs will be devoted to reviewing some of the more relevant empirical work, with a special emphasis on the Spanish evidence gathered so far. As mentioned in the Introduction, labour force participation rates among older workers and retirement patterns have been extensively analyzed in the literature; cf. Peracchi and Welch (1994); Costa (1998); Blöndal and Scarpetta (1998); Quinn (2000); Gustman & Steinmeier (2000), and Alba (1997); Boldrin et al (1999); Ahn and Mira (2000); Gutiérrez-Domenech (2006) for Spain. This stream of research documents changes in retirement trends and pension retirement patterns during the last decades that reflect a tendency towards early retirement and lower participation rates among older workers.

One of most prolific branches in the retirement economic literature is the one that analyzes demographic influences on retirement decisions of workers – including papers which address the influence of gender, marital status, race or ethnicity, and/or other demographic characteristics on the individual (couples)

[7] Several papers (Lumsdaine et al., 1992, Spataro, 2002 and Butler et al., 2002) have tried to assess the predictive behaviour of structural and reduced form models of retirement. Spataro (2002) seems to find a duration model preferable to the structural model. Butler et al.(2002) notices that the duration model imposes less assumptions than the structural model, but it is less robust when one wants to make policy simulations of major changes in the programme.

retirement: cf. Poterba, Venti, and Wise (1995; 1999; 2000); Pienta (1994; 1996; 1997; 1998; 2003); Blau (1998), Maestas (2001).

The interrelations between age, productivity and retirement have also been extensively analyzed in Lazear (1979) and Kim (2003). The study of the interrelations between health and retirement and on the influence of health care costs on retirement decisions has also been extensively analyzed (for an international overview, see, e.g., Krause et al.,1997, Gustman & Steinmeier, 2004, Coile,2003). Additionally, labour supply has also been shown to be affected by disability schemes (see Börsh-Supan, 2007), an issue that has also been studied in Spain by Blanco (2000) and Jimenez-Martinez et al. (2006).

Other papers analyze pros and cons of defined benefit pensions systems as compared to defined contribution type plans and make simulations and projections about the consequences of shifting from one system to the other. cf. Gokhale et al (2001), Hansen (2000), Gustman and Steinmeier (2003). An important line of research in Spain is the one devoted to analyze the sustainability of the retirement pension system and the reforms needed to guarantee its sustainability. This debate has become more intense recently because of the negative consequences of the ongoing economic crisis; cf. Barea (1996); Herce y Alonso Messeguer (1999); Jimeno and Licandro (1999); Jimeno(2000); García et al (2005); Messeguer (2006); Saez & Taguas (2006); Balmaseda et al (2006); Díaz-Giménez and Saavedra (2009); de la Fuente and Domenech (2009). The analysis is usually formulated in terms of overlapping generation models. Overall, the available evidence on the effects of delaying the normal retirement age a number of given years shows that it has a favourable impact on the sustainability of the Spanish pension system; cf. Montero Muñoz (2000); Sánchez-Martín (2003); Boldrin and Jiménez-Martín (2003); Díaz-Giménez and Díaz-Saavedra (2008); Sánchez-Martín and Sánchez-Marcos (2008).

Regarding the specific aim of this paper, there is quite a lot of research concerning the relationship between the incentives embedded in the social security rules and retirement decisions; e.g. Duval (2003); Gruber & Wise (1999, 2004, 2007); Dorn & Sousa-Poza (2005, 2010); Martín and Moreno (1990); Lopez García (1990); Gómez & Hernández de Cos (2004). These studies focus on financial determinants to retire (the different rights acquired through age, gender, contributed earnings, years of contribution, pension replacement rates, etc) and the implicit incentives created by the system. The available micro-econometric evidence for Spain shows that the early retirement provisions play an important role in the modal age of retirement and its pattern in different ages (see, e.g., Boldrin et al. ,1999, 2004) and that, in general, labour force transitions of elderly men depend on Social Security regulations –see Alba (1997); García-Pérez & Sánchez- Martín (2008a; 2008b)[8], and are correlated with health considerations- see Blanco (2000); Prieto et al (2002); and Jiménez et al (1999).

[8] See García Pérez & Sánchez Martín (2008a, 2008b) provide some results on the links between unemployment, retirement and their associated public insurance programmes calibrated with data from the MCVL.

One of the most revealing evidence of the dependency between withdrawal from labour force and pension regulations in Spain is the presence of spikes in the benefit claiming moment around the ages of 60, the earliest age a pension can be, in general, claimed and 65, the ordinary retirement age. Such pattern is also observed in most western countries (see, e.g., Gruber and Wise, 1999, 2004) where eligibility ages prescribed in country-specific provisions also play a major role in defining the observed pattern of retirement by age. As mentioned earlier, there are few papers that attempt to evaluate the impact of pension Social Security incentives on labour force participation in Spain. In particular, Boldrin et al. (1999, 2004) and Jiménez-Martín (2006) follow a regression-based approach based on reduced-form behavioural equations to model the effect on the decision to retire of pension wealth, the incentives embedded in the pension system and individual demographic characteristics. After estimating retirement hazard rates by means of a probit model for a sample of individual work histories randomly drawn from the historical files of Social Security affiliates, they conclude that, while economic and financial measures of retirement incentives play some role in explaining retirement behavior, a substantial portion of the latter still remains unexplained and cannot be explained by Social Security factors. Further research by García Pérez and Sánchez Martín (2008b) find favourable evidence on the relevance of the social security incentives in explaining transitions from unemployment for older workers. Likewise, using a sample of individuals aged between 56 and 70 from the European Union Household Panel (PHOGUE), wave 7 (2000), Utrilla de la Hoz and Ubago (2005) find that a replacement rate (pension over total income) below 80% reduces the probability of retiring.

As for the impact of minimum pensions, this time based on the estimation of the behavioural parameters of a structural model, Jiménez-Martín & Sánchez (2000, 2006) show that the existence of minimum pension's regulations has an impact on early retirement decisions. They find that the combination of age penalties and minimum pensions generate large incentives to early retirement for those workers with low wages and short labour histories. They conclude that there is a threefold increase in retirement at 60 with respect to the economy without minimum pensions and total early retirement (before or at 60) is almost 50% larger.

A very intuitive procedure to analyze retirement incentives was provided by the option value model introduced by Stock & Wise (1990). In this model workers decide whether or not to retire by comparing the expected value of retiring at some period with the expected value of continuing working and retiring at any future date. In other words, workers compute the option value of postponing retirement. Gruber & Wise (1999, 2004, 2007) apply this approach across different individual countries while Scarpetta (1998) studies the importance of early retirement in OCDE countries. Coile & Gruber (2001) and Diamond & Gruber (1999) also follow this approach using USA data. In the case of Spain, Cairó (2009) analyses partial retirement incentives and its impact on the average age of retirement using the option-value framework.

The effect of the successive pension reforms that have taken place in Spain has been empirically addressed in a series of papers. The results point out, in general, to its effectiveness in lowering retirement. Jiménez (2006) carries out a simulation exercise that computes the effect of the Spanish old age pension reforms that took place in 1997 and 2002 and concludes that they reduce the hazard rates. A similar qualitative result is found in Gutiérrez-Domenech (2006), where, using a longitudinal survey of catalan population, it is shown that the 2002 reform contributed to the increase in the staying-on employed probabilities of the individuals older than 60. On the other hand, Sánchez- Martin (2005), using a calibrated overlapping- generations model finds that the overall effect of the 2002 reform is a clear drop in the average retirement age, as younger cohorts of low income workers benefit from the opportunity of leaving the labour force early.

3 The data

3.1 The CSWH

The database *Continuous Sample of Working Histories* (CSWH 2006) includes all the electronically recorded information that the Social Security administration keeps regarding the employment and contributory pension history of more than one million individuals, whose anonymity is preserved. It is a 4% random sample from a reference population, which is composed of all the people who at any time during 2006 had a registered record with the Social Security system, either because they were contributing or because they were receiving a pension. Therefore, it does include those that were unemployed for the whole 2006, either if they received a contributory pension or if they received unemployment benefits, as in the latter case, the National Employment Service pays their social contributions to the Social Security Funds. The sample does not contain any information relating to the scheme called *Régimen de Clases Pasivas* that covers public sector employment, so that most civil servants cannot be included in the analysis. It must be pointed out that as a consequence of the definition of the reference population, a large proportion of non-working females is included in CSWH, mostly as they receive a survival (widowhood) pension[9].

Most of the labour, contributory and pension history of the over one million individuals has been recovered, so that their employment history can be reconstructed. The data contain, for each employment spell, information on *covered earnings*, which are the amount of the earnings that the Social Security takes into account for the computation of pension rights. *Covered earnings* can be regarded as a good proxy for actual earnings, although they are subject to a ceiling and a floor: on the one hand, a minimum contribution must be paid over earnings, independently of the actual amount received, so that there is a minimum *covered earning* associated to it; on the other hand, earnings above a given ceiling are not subject to contribution and therefore do not generate further rights. *Covered earnings* are used in the empirical analysis as a proxy to wages.

Moreover, for each employment spell, CSWH 2006 also provides information on length and type of contract, worker's *Social Security Scheme*, the so-called *contribution group* (associated with the job and that determines the size of the workers Social Security contributions), as well as information about the firm, such as its activity sector and location (province). Available data also include some personal characteristics such as sex, age, place and year of birth. The CSWH has been matched with information coming from the Census. In the Census' module "Co-inhabitants" there is information about the number of people living with the person in the CSWH dataset, their age and sex, but not their working status. As far as social transfers are concerned, the database contains information about periods and amounts enjoyed for old-age and

[9] A detailed description of the sample can be found at Seguridad Social (2006) "La muestra continua de vidas laborales. 2004" and an overview in Argimón and González (2006)

disability pensions, and survivors' pensions such as, orphanage, widowhood and family help. There is no data on other sources of individual wealth or other sources of income. The CSWH has also been matched with Personal Income Tax data, corresponding to the reference year. In that sense, this data module provides information on additional sources of income, but not in a longitudinal dimension.

3.2 The subsample

The analysis performed here is referred to a selected subsample of the people affiliated to Social Security and whose registers are recorded in CSWH-2006.

The subsample considered here includes the available information on men who have contributed to Social Security at least once in their lifetime, have not collected an old age pension before 1997, and were born between 1936 and 1946, so that in 2006 they were 60-70 years old. Moreover, I restricted the sample to those men whose longest contributory relationship with the Social Security took place in the Social Security General Scheme, and who have not claimed disability benefits. I also dropped some individuals who have collected an old pension after 1997, but whose eligibility to access a retirement pension, in terms of having at least fifteen contributory years, could not be proved with the available data, as their pensions could not be estimated with the available information

Several reasons justify restricting the analysis to the specified subsample. First, the sample only refers to men because women often have quite different labour histories. Since women often experience a greater number of career interruptions than men during their potential working life -due for instance to maternity leaves- including women in the sample could lead to biased results due to different gender-linked behaviour in labour participation. Furthermore, women may react differently to economic incentives. The literature that examines retirement's decisions within households shows that women respond significantly to the employment status of their husbands, bringing forward her retirement decisions to coincide with their partner's retirement. Data about socio demographic characteristics, such as marital status are poorly recorded in the CSWH in the sense that information about civil status is not directly available. Although civil status could be somewhat approximated using information on age and gender of people living in the same household, the available data about people with whom the individual is living correspond only to the period when the administrative records was registered by first time and, in any case, there is no reference whatsoever to the spouse's working status, labour history and registered Social Security contributions. Hence, due to these information shortcomings, it is not possible to consider the joint decision to retire using this data.

Secondly, given that the aim here is to analyze retirement decisions taken by the elderly, I restrict the initial sample to people that are aged from 60 to 70 in 2006 (i.e. those born between 1936 and 1946) and that have already become entitled to a pension benefit, defined in terms of being able to prove at least 15

contributory years. It means that those for which the Social Security system does not record any contributory life, those that have retired before their 60th birthday[10] and those who started receiving an old age pension before 1997, when a large pension reform was introduced[11], are excluded from the analysis. Further exclusions involve those that receive an old-age pension coming from disability, as the determinants to claim such kind of pension are most likely linked to health, which is not comparable with the rest of the determinants. Moreover, the transition to the old-age retirement scheme is deterministic, so that disability pensions are converted into retirement pensions once the beneficiaries become 65 years old. Additionally, to ensure homogeneity in pension rules, the sample is limited to those whose longest recorded labour relation has taken place in the Social Security General Scheme, because it is the one that gathers the largest proportion of workers. (See Appendix 2 on legislation). Finally, some individuals with incomplete recorded contributory information are also excluded from the analysis, so that the sub-sample I use for the analysis comprises 35,853 men, whose distribution by year of birth and retirement age is reflected in Table 1. The administrative nature of the data source explains that a limited number of individuals had to be deleted from the sample as the available information for them did not seem consistent.

Table 1

Distribution by year of birth and retirement age. Sample of men born between 1936 and 1946 having worked in the General Regimen and with a relation with the Social Security in 2006

Born							*Not retired*	*Total*
	60	**61**	**62**	**63**	**64**	**65**		
1936		205	199	152	295	892	189	1,932
1937	880	201	162	149	295	866	195	2,748
1938	817	136	159	153	285	819	170	2,539
1939	652	103	135	134	243	732	190	2,189
1940	948	168	225	232	401	1,225	314	3,513
1941	704	140	187	192	332	981	893	3,429
1942	607	186	170	188	350		1,956	3,457
1943	600	213	231	205			2,671	3,920
1944	554	200	220				2,963	3,937
1945	546	209					3,449	4,204
1946	388						3,597	3,985
Total	6,696	1,761	1,688	1,405	2,201	5,515	16,587	35,853

[10] Early retirement before the age of 60 is only possible for dangerous and unhealthy jobs such as air pilots, some miners, railways workers, bullfighters and artists.

[11] Very few records relating to pensions awarded before 1997 are available.

4 Empirical framework

The earliest empirical work in this area considered reduced-form models of the retirement decision as a function of Social Security wealth and pension level. While there were differences in the estimation strategies, based mainly on the nature of available data, the results consistently suggested a role for Social Security. The main limitation of this type of studies was that they considered Social Security effects at a point in time and therefore cannot account for the impact on retirement decisions arising from the time pattern of social security wealth accruals.

In order to address this shortcoming, different approaches have been followed in subsequent analysis. The accrual of Social Security or other more forward-looking incentive measures has been developed and their effect has been analysed with the estimation of reduced-form models[12]. Alternatively, a different approach has been to consider structural models where workers face a budget constraint discontinuous or kinked[13]. Another line of research, the "option value" model of retirement postulated that not only the level of pension wealth or its increases with one additional year of work are important determinants of the retirement decision, but also the evolution of future wealth and work are relevant. So retirement decisions are modelled as a function of the difference between the utility of retirement at the current date and at the date that maximizes one's utility; cf. Stock and Wise (1990). Given the difficulties in the implementation of this technique in structural form models, numerous authors -Samwick (1998); Gruber and Wise (1999); Hakola (2002); Blundell et al. (2002)- have used the option value in reduced form models. More recently, applied general equilibrium models are also being used to explore the pension issue -for instance, in Imrohoroglu et al. (1999)- so that they need to be calibrated to be able to produce numerical results[14].

This paper follows the hazard-model approach in Grubber and Wise (2004), and Blundell et al. (2002), among others, to capture the effects of changes in social security wealth and other variables on retirement. Specifically, I analyze here the incentives to early retirement stemming from the Social Security rules introduced in the 2002 Retirement Amendment, using a reduced-form approach by means of a duration model. This procedure is particularly useful since it allows assessing the impact of relevant variables on the retirement decision taking into account time dependence.

Although the option value model is theoretically the most intuitive model and a structural model should provide more insight into the issue, I

[12] Spataro, 2005 proposes a set of alternative measures that feature the forward-looking aspect and applies them to Italian data.

[13] The lifetime budget constraint is analogous to the standard labor-leisure budget constraint, with annual hours replaced by years of labour force participation, and annual earnings replaced by cumulative lifetime compensation. The kinks are produced by changes in the accrual rates (the rise in retirement income entitlement caused by continuing to work for one more year). See for instance (Burtless, 1986).

[14] See Jimeno et al (2006) for a survey of the features of different approaches used in the literature to study the effects of population aging on Social Security expenditures.

choose the simplicity of this reduced form technique because of the computational complexity of the alternative approaches. Moreover, the chosen reduced-form specification is partially forward looking, as it allows for continuous updating of information as individuals grow older. That is, for an individual who complies with the requirements to retire at age *t*, the probability of retiring at age *t+1* is modelled in terms of the ratio of annual wage earnings over pension benefits, public pension accruals and labour situation at time *t*.

Admittedly, analysis of retirement behaviour that use reduced form models are potentially subject to the Lucas Critique whereby estimated coefficients may represent a convolution of deep parameters and changes in policy rules, being therefore subject to structural breaks. The reason for not being particularly worried about the Lucas´ critique in the specific setup of this paper is that the period under analysis (2002-2006) is quite short and close to 2002 when the reform took place. Therefore, it is unlikely that, in such a short period, the new rules regarding early retirement would have altered significantly the expectations formed by the individuals. These expectations are to a large extent driven by the parameters which define individuals´ utilities which, in turn, are captured in the estimation by the variables proxying the economic incentives provided by Social Security included in the analysis. As a result, the changes in individuals´ behavior during such a short period since the reform are bound to be well captured by the interactions of the previous three variables with the 2002 reform dummy variable *r2002*.

The retirement decision is analysed in this paper following a duration model approach that treats it as a dynamic discrete choice. The duration variable (T) is defined as the period from the age the person becomes entitled to receive a retirement pension until the age that she claims the benefit. It is treated as a discrete variable, defined in years, as a varying variable that changes as time goes by[15]. Let C_i be the maximum number of years that we could observe the individual in the sample, which is constant since it only depends on his year of birth and the moment he became eligible. Thus, for example, if a person born in 1941, becomes eligible in 2001 (when he is 60), and only claims a pension benefit in 2003 (when he is 62), will have T=3 and C=5 in 2004. In the estimation C_i is restricted to be at least 1 and no greater than 6 under the assumption that the decision to retire only takes place between 60 and 66, since only very few people retire after that age, or do not retire at all. The dataset contains some individuals that either can only be observed before they take the decision of interest (i.e. claim pension benefit) because I only observe the individuals up to the year 2006 or did not claim a retirement pension before the age of 66. In both cases, data is right censored. For these individuals we can only observe that T>C, so that they remain in the current situation for a larger number of years than the ones that the sample allows us to observe them. That is, we observe claiming at T only if T=<C. Otherwise, we observe that T>C. It is assumed that T is independent of C. It must be reminded that for some cohorts we can only have observations of people retiring after a given age (60 for those born in 1936) or before a given age (for instance, at most

.

at 60 for those born in 1946) given that this is the age when the sample was extracted, as shown in Table 1.

The hazard function (i.e. the probability of an individual retiring precisely at time *t*, given that he has not retired before *t*), is defined as:

$$\phi(t_i \mid \mathrm{X}_i) = prob\ (T = t_i \mid T >= t_i) =$$

$$prob\ (T = t_i) / prob\ (T >= t_i) = F[\theta_0(t_i) + \theta_1(t_i)x(t_i)]$$

where X_i is a vector of explanatory variables, some of which are age dependent and others are constant and independent of the age, F is a logistic cumulative density function and t_i is the duration observed for individual *i*. The hazard gives the probability of retirement defined over the surviving population at each time. In other words, it gives the probability of retiring at t for those eligible persons that have not retired.

A discrete duration model can be regarded as a sequence of binary choice equations (with cross equation restrictions) defined on the survival population at each duration. We are, therefore, interested in the conditional distribution of t in relation to variables x: $F(t_i \mid x_i)$

Let' us define c_i=1 if the individual has retired and c_i=0 if she continues working. Therefore the likelihood can be expressed as:

$$L = \prod_{i=1}^{n} \left\{ \phi_{t_i}^{c_i} \prod_{s=0}^{ti-1} (1 - \phi_s) \right.$$

where ϕ is the estimated hazard function, C is the maximum number of years that we could observe the individual in the sample, n_j is the number of people that retires when T=J and r_j is the number of people that are eligible to retire at j. The likelihood can be alternatively rewritten as:

$$L = \prod_{j=0}^{J} \phi_j^{\,n_j} 1 - \phi_j^{\,r_j - n_j}$$

The dependent variable in the model is the probability of retiring (claiming a retirement pension) at a specific age, given that the person has become eligible to do so[16] and has not retired the preceding years. The explanatory variables that have been included in the specification (X_i) can be grouped in three main categories: social security regulations, personal characteristics, and labour characteristics. A detailed description of these variables can be found in Appendix 1. For a complete derivation of the log-likelihood function see Jenkins (2005).

I build up traditional measures of incentive mechanisms that are standard in this literature. In particular, I use the Replacement Rate (RR_t), which is the ratio of the expected pension benefits over wages, the Social Security Wealth (SSW_t), that is, the present discounted value of the future stream of pension benefits, the Accrual Rate (SSA_t) which measures the discounted change in SSW from postponing retirement one year and the Peak Value (PV_t), that compares this year's social security wealth to the maximum social security wealth that could be attained in the future. They are constructed under the assumptions that: (i) the age of death is certain, (ii) no changes in social security regulation are expected by individuals, and (iii) it excludes any tax considerations.

Table 2

Descriptive values of Social Security incentives and working status. Sample of men born between 1936 and 1946, having worked in the General Regimen and with a relation with the Social Security in 2006

	All period considered 1996-2006			Before 2002			After 2002		
	median	mean	*s.d.*	median	mean	*s.d.*	median	mean	*s.d.*
SSW_t (1)	199.30	**222.50**	*96.83*	195.83	**211.84**	*87.02*	201.35	**228.17**	*101.21*
SSA_t (1)	6.89	**6.73**	*12.95*	10.12	**9.79**	*12.20*	4.97	**5.09**	*13.04*
PV_t (1)	12.65	**18.84**	*26.73*	21.68	**26.77**	*28.03*	8.32	**14.63**	*25.01*
RR_t (2)	57.61	**67.36**	*35.50*	50.77	**57.61**	*26.56*	61.93	**72.55**	*38.45*
l_{t-1} (2)		**57.49**			**62.67**			**54.74**	
u_{t-1} (2)		**24.28**			**27.54**			**22.55**	
N		**115,532**			**40,123**			**75,409**	
n		**35,853**			**16,212**			**28,259**	

N: Number of observations; n: people

(1) In thousands of euros

(2) In percentage

Table 2 provides information about the descriptive values of the incentive Social Security variables. It shows that, while the average stock of wealth was higher after the 2002 reform than before, the incentives linked to it were lower, and that RR increased after this reform. It also contains information related to working status, distinguishing between those observations corresponding to a working situation (l) previous to the decision moment and those corresponding to an unemployment situation (u). The rest of the observations correspond to either an already retired situation or where no work is involved. Additional data concerning the rest of the variables can be found in Table A1 of the Appendix 1. Both the level of social security wealth and the different incentive variables enter the equation. The level captures wealth effects: the larger the value of wealth, the larger the demand of all goods, including leisure, if leisure is a normal good. The incentive variables capture a substitution effect: the higher the price for leisure, the lower its demand, so that if there is a larger financial incentive to additional years of work, then individuals will retire later.

When tax considerations are addressed to compute economic incentive variables, the results previously described are slightly stronger, but the overall conclusions drawn in the main text remain similar. See Appendix 3 for further

details regarding the potential negative implications of excluding tax considerations.

The specification chosen (a logistic duration model) also allows testing whether the individual leaves his current situation during the first year he is entitled to collect a pension benefit. Finally, the impact of the regulatory change introduced from 2002 is also analysed, as it could have affected the average probability of retiring and through its effects on the incentive mechanisms.

As for the labour characteristics, the control variables are the individual's labour situation the year preceding retirement, distinguishing between those being working and being unemployed, the industry where the individual works and some measures of labour mobility or precariousness. In particular, there have been considered two variables as a proxy measure of the degree to which job insecurity and precariousness in employment (contingent work and temping, fixed-term contracts, involuntary part-time work, moonlighting) have an important effect in the individual's employment history. The number of different episodes of contribution, *numrel* (that includes both labour and unemployment periods), and their average length in years, *meanlenght*, are included as covariates.

Finally, some standard demographic controls such as age, education and health status are included. I also control for the collection of other benefits as they may interact with the old age pension. Finally, regional dummy variables and GDP growth are also included to control for the macro environment.

The duration dependence of the hazard rate is captured in two ways. On the one hand, following Bover et al. (2002), instead of imposing a specific functional form for duration, I introduce additive dummy variables for each of the possible discrete values of the duration variable. Durations of more than 6 years (which would necessarily imply that the person is at least 66) are treated as censored at 6, due to the relatively small number of observations under such a circumstance. On the other hand, interaction of certain independent variables with the duration are included to test if the variable effects change with the number of years that a person takes to retire. Appendix 1 contains a summary and a brief description of the variables used in the analysis.

5 Results

The qualitative impact of the variables on the hazard are discussed in terms of the sign and statistical significance of the estimated coefficients, which are reported in detail in Table A 2 in Appendix 1. The Table A 2 shows 3 different specifications each one contains two different sets of Social Security incentive variables. So, under model A I present the estimated coefficients obtained for the basic specification, distinguishing between the Accrual Rate (SSA) as incentive variable, and the Peak Value (PV). Model B tries to capture the impact of the 2002 reform through the inclusion of dummy variables that test for its relevance and the effect on the response to the social security variables, while model C, also includes dummy variables that capture the effect on the retirement hazard rate of being entitled to receive the minimum or the maximum pension benefit at each age.

Table 3a

Quantitative effects of Social Security measures on the retirement probability [(*)]

	MODEL A		***MODEL B***		***MODEL C***	
	SSA	PV	SSA	PV	SSA	PV
SSW_t	0.0060	0.0060	0.0072	0.0072	0.0072	0.0072
	0.0029	*0.0029*	*0.0034*	*0.0034*	*0.0033*	*0.0033*
SSA_t	-0.0007		-0.0004		-0.0006	
	-0.0003		*-0.0002*		*-0.0002*	
PV_t		-0.0007		-0.0004		-0.0003
		-0.0003		*-0.0001*		*-0.0001*
RR_t	0.0003	0.0002	0.0008	0.0009	0.0004	0.0006
	0.0001	*0.0001*	*0.0003*	*0.0004*	*0.0001*	*0.0003*

[(*)]Quantitative effects are computed as the discrete differences of the logistic function evaluated at a 10% increase in the variables' values with respect to the logistic function evaluated at the observed variables' values. The effects are estimated for the period beginning in 2002.

Results are obtained from the regressions presented in Table A2 where:
Model A is the basic model
Model B includes Model A variable and a dummy for the change in 2002 regulation as control variable
Model C includes all variables in Model B adding income level as control variable.

Median values are in italics

The size of the effect of the Social Security incentive variables on the probability of retiring is reported in Table 3a. In particular, for the variables that are continuous, the impact is measured as the effect of a change in the Social Security variable on the predicted hazard. For the qualitative ones, its impact is computed from the change from 0 to 1 in the independent variables, so that coefficients can be interpreted as the direct effect of having such characteristics on the probability of retiring. The more detailed quantitative impact of a change in all the variables included in the specification distinguishing by age is reported in Table A 3 of Appendix 1.

Table 3b shows the corresponding percentage change in probability equivalent to those reported in Table 3a. The main finding is that, according with predictions of Model C, a 10% increase in the PV decreases the retirement probability by 0.0003 points, implying a -0.2% average hazard increase.

Table 3 b
% Quantitative effects of Social Security measures on the retirement probability (*)

Model C

Period: 2002-2006

	SSA				PV			
		% Δ in predicted hazard :				% Δ in predicted hazard :		
Age	Predicted Hazard	10% Δ SSW	10% Δ SSA	10% Δ RR	Predicted Hazard	10% Δ SSW	10% Δ PV	10% Δ RR
60	0.152	5.5%	-0.6%	0.2%	0.153	5.4%	-0.5%	0.4%
61	0.058	7.8%	-0.7%	0.3%	0.057	7.7%	-0.4%	0.6%
62	0.072	7.8%	-0.5%	0.3%	0.071	7.8%	-0.3%	0.6%
63	0.079	8.0%	-0.4%	0.4%	0.078	7.9%	-0.2%	0.6%
64	0.173	6.1%	-0.3%	0.3%	0.172	6.1%	-0.1%	0.6%
65	0.723	1.6%	-0.2%	0.1%	0.723	1.5%	-0.1%	0.2%
Average	0.157	4.6%	-0.4%	0.2%	0.157	4.5%	-0.2%	0.4%

*Quantitative effects are reported as the % change in the predicted hazard. Probability values are reported in Table 3a

Economic incentives and regulation

The results in Table A 2 show that all the coefficients on social security variables (SSW, SSA and PV) are statistically significant with the expected sign. Increases in the total present value of the flow of pensions that a person will receive from the year she retires to the year she dies, i.e. a rise in SSW, increases the hazard. Increases in the difference of this amount derived from postponing the retirement (either one or more years) reduce the hazard, irrespectively of whether SSA or PV are used to capture the substitution effects.

Table 4

Relative incidence of Social Security measures on retirement decisions for all the period considered [(*)]

	SSA				PV			
	AIC	*BIC*	*ln Likelihood*		*AIC*	*BIC*	*ln Likelihood*	
(Mc): Overall Model C	***65659.72***	**66210.18**	**-32772.86** ***(57)***		**65648.87**	**66199.34**	**-32767.44** ***(57)***	
	AIC	*BIC*	*ln Likelihood*	$pr>\chi(k)^2=\chi^2(k,Mc-M_i)$	*AIC*	*BIC*	*ln Likelihood*	$pr>\chi(k)^2=\chi^2(k,Mc-$
Exclusión restriction								
(M_1): SSW_t	66186.56	66717.71	-33038.28 *(55)*	*0.000*	66164.28	66695.43	-33027.14 *(55)*	*0.000*
(M_2): SSA_t/PV_t	65694.62	66225.77	-32792.31 *(55)*	*0.000*	65694.62	66225.77	-32792.31 *(55)*	*0.000*
(M_3): RR_t	65659.82	66190.97	-32774.91 *(55)*	*0.128*	65656.79	66187.94	-32773.39 *(55)*	*0.003*
(M_4): SSW_t and SSA_t/PV_t	66212.33	66724.17	-33053.17 *(53)*	*0.000*	66212.33	66724.17	-33053.17 *(53)*	*0.000*
(M_5): SSW_t and RR_t	66201.02	66712.86	-33047.51 *(53)*	*0.000*	66188.33	66700.17	-33041.17 *(53)*	*0.000*
(M_6): SSA_t/PV_t and RR_t	65703.37	66215.21	-32798.69 *(53)*	*0.000*	65703.37	66215.21	-32798.69 *(53)*	*0.000*
(M_7): SSW_t, SSA_t/PV_t and RR_t	66226.46	66718.98	-33062.23 *(51)*	*0.000*	66226.46	66718.98	-33062.23 *(51)*	*0.000*
(M_8): r2002	65787.63	66299.47	-32840.81 *(53)*	*0.000*	65792.13	66303.96	-32843.06 *(53)*	*0.000*

[(*)]Results are obtained from regressions in columns 3 and 6 of Table A 2. Model parameters in parenthesis. k stands for the difference parameters betwen overall Model C and the alternative model considered

In order to provide an assessment of the model specification and, in particular, of the goodness of fit resulting from the inclusion of the different economic incentive variables whose statistical significance has been tested, I propose using the Akaike (AIC) and the Bayesian Information criteria (BIC)[17]. Each row in Table 4 shows AIC and BIC criteria for the 8 alternative specifications considered (M_1-M_8). These alternative models enforce an exclusion restriction in the overall model C, regarding the economic incentives variables that are specified in the first column of the Table. By providing a criterion to choose among nested models, I can compare the relative explanatory power of the different variables to affect the probability of retiring. The specification preferred should be the one including all the economic incentive variables, since it is the one with the lowest AIC and BIC. On pure likelihood grounds, overall Model C is also the preferred one. The only exception is encountered in the specification tested in M_3 that does include the accrual value. The specification dominates the one with the more myopic incentive measure in the unrestricted model.

The SSW variable and the other economic incentive variables are individually significant both in the SSA model as in the PV model, as shown in their respective t-ratios in Table A 2. Moreover, joint significance of the incentive measures has been tested using a Wald test. Based on the p-value associated with a chi-square of 128.81 (145.72) with 3 degrees of freedom for the accrual (peak value) specification, I am able to reject the null hypothesis indicating that the coefficients for SSW, SSA/(PV) and replacement are not simultaneously equal to zero. Thus, inclusion of these variables improves the overall fit of the model.

In spite of these effects being statistically significant, they are not very large. As the figures in Table 3 show, on average, a 10% rise in SSW increases the probability of retiring between 60 and 65 years of age by around 4.6 percentage points. Moreover, these probabilities show a U-shaped form with age (Table A3), reaching the highest impact at 65, so that the closer a worker is to that age, the more responsive to changes in SSW becomes.

As for the incentive variables, I find that increasing by 10% the difference between what a worker would receive if she retired now and what she would receive if she retired one year later (increasing SSA) decreases the average probability of retiring between 60 and 65 by an amount ranging between 0.3 and 0.7 pp and by a similar amount if the 10% increase would be in PV. The effects on the retirement probability of the SSA and PV incentives are also U-

[17] AIC=-2*ln(likelihood)+2*k, and BIC = -2*ln(likelihood) + ln(N)*k . N is the number of observations and k the number of parameters

.

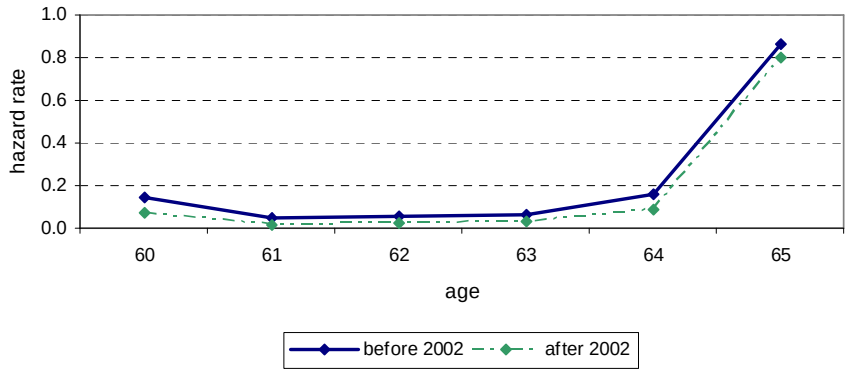
shaped in relation to age (Table A3), so that they are larger at 60 and 65 and lower for the intermediate ages in that age range.

On the other hand, the replacement rate shows the expected positive sign for the whole sample. Yet, it is not statistically significant, a finding which is confirmed by the results on the AIC and the BIC criteria shown in Table 4.

The main conclusion from the estimates presented in Table 4 is that the wealth variable plays a larger role in the determination of the hazard rate than any of the remaining incentive variables. This could imply that changing the amount of money that is transferred through pensions is more relevant for retirement decisions than changing the built-in incentives in the regulation.

The predicted probability of retiring before 2002 between 60 and 65 is higher than after 2002 and this relationship holds for every age in the range from 60 to 65 (Figure 1). The role of the 2002 reform on this change is tested by including a dummy variable that takes value 1 from 2002 onwards and the interaction of this dummy variable with the Social Security incentives and pension wealth.

Figure 1 Predicted hazard rates before and after 2002 by age



The evidence from the signs and t-ratios of both the dummy and the interacted variables that are used to test for effect of the changes that were introduced in 2002[18], point out that the reform did not change the pattern of response to the wealth variable, but that, in general it changed the response to the incentive variables, reducing its impact on the timing of retirement (Table A 2). This implies that we need a larger value of SSA (or PV) from 2002 to reach the same impact on the hazard, as shown in Table A 4, where a decomposition of the effects of the reform on the incidence of the different incentive variables is presented. This result could be explained as a consequence of the fact that one of the changes carried out under the reform improved the treatment for those that had more than 30 years of contribution, increasing for them the amount of the pension to be received at each age. The fact that most workers (64%) already had at least 30 contributory years (in particular, nearly one third had more than 34 years of contributions), may explain the finding that the reform reduced the incidence of the incentive structure. Additionally, the replacement rate becomes

[18] See Norton et al (2004) for a discussion of such test.

statistically significant for this period, so that a more myopic approach seems to hold. Therefore, the observed reduction in the hazard rate does not seem to stem from the new regulation, which has reduced its incidence, but from a variety of other factors that are captured by the dummy variable.

Since the Social Security legislation establishes a maximum and a minimum level of monthly contributions. I will refer in the sequel to the individuals who contribute at those levels as "*low (upper) topped*". Regarding this people, an interesting finding is that those who are *low topped* have a higher retirement probability at 60 than those who are not (Table A 4). It could be argued that the minimum pension mechanism offsets the effect of early retirement penalties so that it creates a strong incentive for low income earners to retire, which is especially relevant at 60. From 61 onwards, being low topped in wealth reduces the probability of retirement. In fact, the older the worker is, the higher is the reduction in the probability to retire. This result could be consistent with the idea that people might choose to carry on working in order to build up more pension rights, given the built-in incentives. The latter arise from the higher dependency between the amount of the pension to be received and the latest wages that the worker receives compared to future wages which, if they were increasing in age, could lead to a larger pension in the future.

The results also show that high earners (those who have their pensions capped) have a lower probability of retiring both at 64 and 65[19] (Table A 4), in line with the findings in Villagarcía (1995), Jiménez-Martín and Sánchez (1999), Blanco (2000) and Labeaga (2008). This finding could result from the fact that, for those workers, financial incentives are not a good proxy for the marginal utility from working. A lower potential wage rate for a elderly workers is likely to be associated with a lower probability of labour force participation as, other things equal, a lower wage rate represents a lower opportunity cost of leisure and a higher replacement rate for government pensions.

Appendix 3 explores possible biasing effects on the estimation of the Social Security economic incentive variable coefficients resulting from pooling both (upper and lower) capped workers and non-capped workers in the same regression. Results remain similar when analysis is restricted only to those who are not capped as can be seen in Table A 4.

Duration variables and cycle

Given that individuals enter the sample as soon as they satisfy the requirements to claim a pension, we can interpret the significance of the dummy coefficient for duration one period (g1) as a test for the relevance of becoming entitled to retirement benefits[20]. If the preferences for retiring were high enough, becoming eligible would be a main determinant of the decision to retire, and people will retire as soon as the regulation would entitle them to do so. However, the coefficient on g1 is statistically not significant (Table A 2).

[19] People from all ages can be low topped, but only people that are 64 or older can receive the maximum pension

[20] See Appendix 2 for a description of these conditions.

Therefore, it can be concluded that becoming eligible by itself is not a relevant ground for retirement. The results show also non-monotonic duration dependence.

Early-retirement may result from business cycle fluctuations. In the economic literature, the effect of business cycles on retirement is ambiguous. The ambiguity arises as a result of the existence of two opposite forces in retirement decisions in response to changes in the business cycle. On the one hand, finding employment becomes more difficult during recessions. Thus, it is likely that in downturns those workers who become unemployed at ages where retirement is possible experience an increase in the probability of leaving the labour market and therefore retire. On the other hand, recessions often involve huge reductions in assets values owned by workers and this negative wealth effect may discourage them from early retirement. Our results show that, during the period analyzed, the propensity to retire is pro-cyclical, so that the hazard retirement rate is higher during expansions. Pro-cyclicality of early retirement encountered in Spain is in line with previous results by Montizaan, Cörvers and de Grip (2007) for the US. Asset price changes are highly correlated with the evolution of the business cycle. If individuals rely on their investments to fund their consumption during retirement, on top of what they can get from the retirement pension, they are particularly vulnerable to asset market downturns. That could be the reason why in periods of economic prosperity, prospective retirees are more optimistic about the evolution of their other sources of income and therefore decide to retire earlier. However, pooling EPA data, Muñoz (1995) provides evidence that individuals retire earlier in recessions than in expansions. The difference between Muñoz´s (1995) results and mine might be explained by the growing development of investment funds and retirement plans during the sample period I have considered in the analysis. While in 1993 there were only 91 pension plans in Spain, these amounted to more than 1500 in 2003 managing 55,764,768 thousand Euros (compared to 152,837 thousand Euros managed in 1993)[21]. Furthermore although pension plans based on fixed income securities are the most common in Spain, equities pension plans have gained considerable importance (equities has grown from 3.7% of total investments in 1993 to more than 21 % in 2003).

Age

In the specification, I also include age dummies to account for differential effects related to this variable. These dummies should capture the effect of growing older "per se", and not through the different coefficients in the calculation of pension rights that are age dependent. The results show that, even when controlling for eligibility criteria and Social security variables, 65 is a prevailing retirement age.

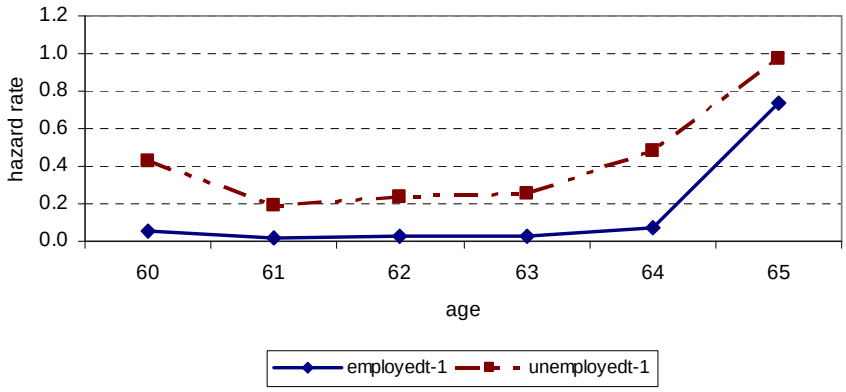
[21] Dirección General de Seguros y Fondos de Pensiones (DGSFP) y Asociación de Instituciones de Inversión Colectiva y Fondos de Pensiones (INVERCO)

Working status, industry and other labour history related variables

The labour market status prevailing during the year before the retirement decision is taken could be relevant for the retirement hazard. In particular, the sample allows us to distinguish among four labour status: working, receiving unemployment benefits and thus also contributing to Social Security, contributing without working nor perceiving unemployment benefits (special agreement), and not contributing.

The results show that a person working in a particular year has a lower probability of retiring the following year than a person who was not working, even when I condition on social security incentives. They also show that being unemployed the preceding year increases the probability of retiring. Such findings may only reflect the predominance of special early retirement programs that exist for unemployed old people[22]. I also find that the size of these effects varies with age, as can be seen in Figure 2, showing a u-shaped form.

Figure 2 Predicted hazard rates according to previous working status and age



The industry where the individual works (sector of activity) is also included as a covariate, allowing for the distinction between those working in the service sector and in the industry sector. The estimated coefficient is, however, not statistically significant, in contrast to other works where the sector of activity is found to be playing a significant role (Conde-Ruiz and García, 2004, Blanco, 2000, Villagarcía,1995, and Muñoz,1995).

The next couple of variables try to capture the quality of the labour relations, through the number of contracts that have been recorded for a given work history up to the eligibility moment and the average length of these contracts. The results show that job turnover (proxied by the total number of labour relations that a person has had) has a negative impact on the retirement hazard but that impact that fades away as duration increases. The sign of the coefficient suggests that the higher the labour turnover is the lower the probability of retiring between 60 and 65 becomes. This result is in line with the

[22] See García Pérez and Sánchez Martín (2008b) for an analysis of the transitions from unemployment for older people.

findings that workers with a firm-specific training history retire earlier than workers with a general training background (see Montizaan, Cörvers and de Grip, 2007).

Regarding the average length of the contract, it seems to play a significant role in the decision to retire only as time unfolds. Notice, however that, although in general, the longer the contract the more stable the individual's working life has been, this it is not a sophisticated measure of a stable career as the same average may result from quite different job histories.

Individual characteristics and other

The results obtained for the negative role of higher education on the probability of retiring is in agreement with the available findings in Villagarcía (1995) and Gutiérrez-Domènech (2006)[23]. One explanation for such a finding may result stem from the theoretical argument stating that low-ability workers are induced to retire early because of the intra-generational redistribution built in early retirement provisions via the utility derived from leisure (Conde-Ruiz and Galasso, 2003). Further, the effect of education is reinforced by duration, so that the probability of retiring of a higher educated worker becomes larger the longer the time it has elapsed since becoming eligible.

As for the health status, the results show a lower hazard for those receiving disability benefits the year before retirement. Such counterintuitive finding may just reflect the fact that those receiving disability benefits, besides having a poorer health, will probably be the ones receiving retirement disability pensions when they turn 65, the only age in which this type of pensions can be awarded and which I am not including in the analysis.

Receiving public transfers other than unemployment or disability benefits reduces the retirement probability, which may be a consequence of liquidity constraints.

Dummy variables for the region (Autonomous Community) where the worker initially registered are included in the specifications, as a way to capture other differences in the economic environment. Coefficients are not reported but are available under request.

In sum, most of the covariates that I have included in the estimation provide evidence in line with previous empirical findings in the literature. The only noticeable differences lie in the pro-cyclicality of retirement and the absence of differential retirement patterns by sector of activity. The former can be due to the growing importance of complementary equity-based pension plans during the last decade, which is precisely the period analyzed in this paper. The latter discrepancy can be due to the restrictive characteristics of the sample selected here (only men working in the Social Security General Scheme)

[23] Muñoz (1995), on the other hand, finds evidence that the education has a quadratic effect, so that those individuals with little or with a lot of studies retire later than those with an intermediate level of education.

which is narrower than those used in other available studies that analyze not only employees in public and private sector but also self-employed workers.

Sensitivity analysis. Alternative regulatory schemes

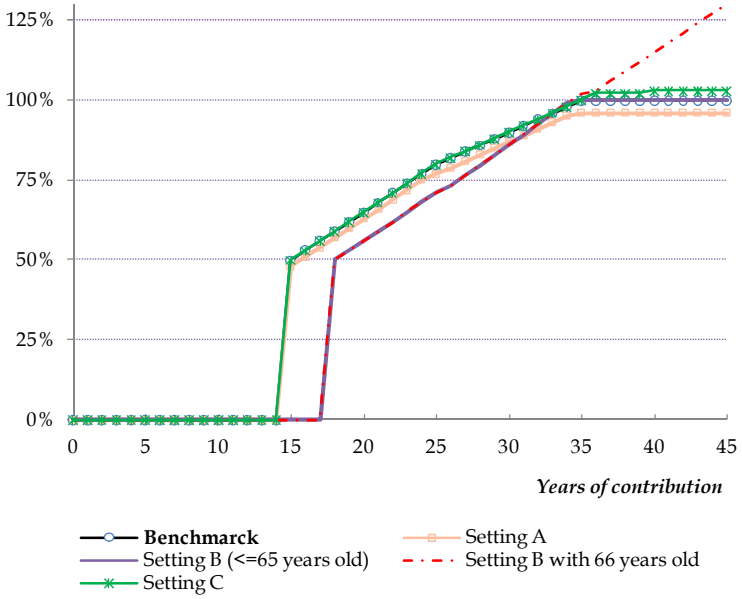
Finally, I have performed a few sensitivity exercises with the aim of shedding light on how the estimated retirement probabilities respond to small changes in the economic incentive mechanisms underlying the Spanish public pension system. The reason for considering only small changes of the baseline rules is to minimize the Lucas' critique, and also to check whether these changes have enough sizeable effects or more drastic changes are needed in which case a structural approach to evaluate their impact would probably be required. Following a similar argument to that given earlier against the validity of the Lucas´ critique in the specific setup of this paper, the sensitivity analysis in this section should also be interpreted as being only valid during a short period in which hat individuals have not time to change their expectations. As a result, the results in this section should only be considered as a counter- factual exercise that provides a quantification of what would have happened if alternative retirement reforms were proposed, given the observed behavior of individuals after the pension reform in 2002.

Specifically, I have computed the pension wealth that a person would receive under 3 alternative regulatory schemes implying very small changes over the retirement rules established in the 2002 pension regulatory framework. Recall that this framework was characterized by requiring at least 15 years of contribution so that a pension reaches 50% of a proxy of gross average lifetime earnings (regulatory base), raising to 100% after 35 years of contribution. I have computed the SSW and the incentive measures that result from the assumed alternative settings for each individual in the sample to obtain counterfactual predicted retirement probabilities. These are compared with the ones predicted from my estimates, which act as a benchmark. The new probabilities are obtained using the specification in column 1 of Table A 2. In particular, I compute the changes in relation to the population that in the sample are subject to the rules prevailing in 2002. Each of the three alternative regulatory schemes tries to change only one item in the pension rules.

The first sensitivity scenario (setting A) would consist of an overall reduction in the amount of the pension received at all ages. It caps the pension at 96% of the regulatory base, not allowing individuals to get the 100% of it at any retirement age. Under this scenario, with 15 contributory years, the pension amount is only 48% of the regulatory base. Moreover, up to 25 years of contribution, there would be an extra 3% increase in each year is added, and from 26, the increase would be 2% per year up to 96% of the regulatory base at 35 or more years of contribution. Even if there were no changes in the number of people retiring, such scenario implies a reduction of the pension burden in relation to the benchmark.

The second scenario (setting B) affects mostly the incentive structure as it requires more years of contribution to claim a pension (18 years instead of 15), so that the increase ladder becomes steeper. Moreover, at the age of 66, if the person has more than 35 contributory years, an additional increase in the amount of the pension is added. Under this scenario, there is a 3% increase up to 25 contributory years. Beyond that age and with a longer contributory period, the increase is 3.22%. Only 2.6% of people at 60 have a labour history shorter than 18 years. Therefore, under this scenario, we cannot initially conclude whether the burden rises or not.

Figure 3: Percentage of the Regulatory Base linked to the number of years of contribution



The third scenario (setting C) proposes changes in the short-term incentives to stay beyond 65, in the sense that it only introduces higher retirement benefits for each additional year beyond 65 that an individual remains not retired. It reproduces part of the changes introduced by the "Acuerdo sobre Medidas en Materia de Seguridad Social", of 13 July 2006, which were implemented two years later. In particular, at 66 with at least 35 years of contribution, a 2% increase in the pension is added, raising it to 3% for those individuals with at least 40 years of contribution.

Table 5 shows the computed counterfactual predicted retirement probabilities for these three scenarios, the average predicted retirement age over the age interval 60-65 and the changes in the number of retirees in the same interval and its disaggregation by ages. All the changes analysed result in an extremely small increase on the average retirement age and a reduction in the number of people who retire between 60 and 65 which ranges from 0.08% under setting C to 1.6% under setting B.

Table 5

Effects of alternative settings on retirement between ages 60 and 65

	Benchmark	Setting A	Setting B	Setting C
Predicted retirement probability	77.08	76.03	76.82	76.95
Predicted retirement age	62.92	62.93	62.93	62.92
Change in number of retired people	-	-1.53	-1.58	-0.08
by age:				
60	-	-1.80	-1.63	0.00
61	-	-2.45	-2.09	0.00
62	-	-2.48	-2.10	0.00
63	-	-2.59	-2.26	0.00
64	-	-2.06	-1.90	0.00
65	-	-0.60	-1.09	-0.20

Estimations are derived from Model C, Column 1

(1) In percentage

(2) The effects are estimated for the period begining in 2002

Setting A: Capping the pension to 96% of the Regulatory Base at 35 years of contribution and to 48% at 15 years of contribution

Setting B: 18 years of contribution to claim a pension. Retiring after 65 implies 3% increase

Setting C: Retiring after 65 implies 2% increase and 3% if 40 years of contribution

Under the first two scenarios, retirement rates at all ages are affected by the proposed changes and the effects are increasing with age up to 63, when the largest impact is recorded. The smallest effect is at 65, probably reflecting that this age is regarded as normal retirement age.

Under setting C, there is a very small impact on the probability of retiring between 60 and 64 compared to the benchmark case irrespectively of the use of a specification where only changes in SSA (Model C, column 1) or only changes in PV (Model C, column 2) are included. In either case, the decrease in the number of people that choose to retire at a later age is rather small (0.20 %), as only the incentive changes, but not the social security wealth at that age. In fact, those whose entire contributory life is below 35 years, experience a reduction in the amount they receive at 65 in relation to the benchmark. Notice, however, that when the specification with the PV is chosen in setting C, the retirement pattern predicted by ages turns out to be different. Since PV considers information about the entire time horizon, individuals take into account earlier in their lives changes that will affect them after 65, in contrast with what happened when predictions are derived from the specification with SSA. Despite these differences, predicted retirement age is similar under both specifications.

Table 6

Effects of alternative settings on retirement between ages 60 and 65

	Benchmark	Setting A	Setting B	Setting C
Predicted retirement probability [1]	77.08	77.23	76.99	77.76
Predicted retirement age	62.92	62.93	62.93	62.93
Change in number of retired people[2]	-	-1.49	-1.63	-1.63
by age:				
60	-	-1.02	-1.24	0.65
61	-	-2.86	-2.88	-0.52
62	-	-3.07	-3.00	-0.67
63	-	-3.38	-3.10	-0.80
64	-	-2.50	-2.28	-0.41
65	-	-0.49	-0.84	0.00

Estimations are derived from Model C, Column 2

[1] In percentage

[2] The effects are estimated for the period begining in 2002

Setting A: Capping the pension to 96% of the Regulatory Base at 35 years of contribution and to 48% at 15 years of contribution

Setting B: 18 years of contribution to claim a pension. Retiring after 65 implies 3% increase

Setting C: Retiring after 65 implies 2% increase and 3% if 40 years of contribution

6 Conclusions

There is some agreement that generous early retirement provisions account for a large proportion of the drop in the labour force participation of elderly workers that had been observed in Spain during the nineties. This paper aims at quantifying the impact of these provisions under the Spanish Social Security System.

The evidence shows that, in general, the economic incentives stemming from Social Security regulations on old age pensions in Spain seem to have the expected effect on retirement. The present value of the future flow of pensions has a positive impact on the probability of retiring, with larger pensions shortening the span between becoming eligible for retirement and actually claiming the retirement pension. Therefore, all measures taken to reduce the present value of such flow at early ages may have the desired effect of reducing early retirement. Moreover, it seems that the built-in incentives in the system discouraging early retirement have a non negligible effect on old-age retirement, so that they are effective in retaining people in the labour market. The higher flow of pensions that workers receive for staying at work one additional year compensate for both the loss of leisure that they experience for the additional year that they keep contributing and the wage they receive at work. The quantitative size of such effect is statistically significant. Moreover, small variations in the incentives measures have a sizeable effect on early retirement. For the period beginning in 2002, a 10% increase in SSW (SSA, PV, replacement) results in a 4.6% (-0.4%, -0.2% and 0.4%) increase in the hazard according with estimates in Model C (see Table 3b). Therefore, from a policy perspective, there is a need to reinforce such effects.

It has been also found that the new scheme implemented since 2002 has reduced the probability of retiring at each age, in spite of the fact that the substitution effects captured through the incentives measures seem to have also reduced their incidence on retirement decisions. Moreover, there is evidence pointing out to a more myopic behaviour of workers regarding social security incentives. The changes in the regulation that have taken place may explain such results.

Any new change in the incentive structure of pensions should take into account the longer work histories that younger people have when becoming eligible. In fact, the counterfactual results show that a small change in the incentive structure has a small impact on the number of people retiring. It seems that to increase the number of people staying beyond 65 requires more than a tiny push. It is therefore necessary to combine economic incentives with other institutional constraints in order to effectively increase the retirement age.

There is an urgent need to address the consequences of an ageing population on the Social Security accounts. Minor changes in the rules that define the amount received at each age in relation to the years effectively

contributed may have a positive impact on the pension accounts, through its impact on the probability to retire at each age. Yet, as the evidence is this paper shows, this is not enough. Prospective amendments in retirement rules should be oriented to link the possibility of retiring and the benefit rights not only to contribution efforts but also to life expectancy.

7 References

Ahn, N., & Mira, P. (2000). "Labor Force Participation and Retirement of Spanish Older Men: Trends and Prospects". (FEDEA, Ed.) FEDEA Working Paper (Working paper 2000-25).

Alba, A. (1997). "Labour Market Participation and Transitions of Older Workers in Spain". (Universidad Carlos III, Working Paper Nº 97-39.).

Argimón, I., & González, C. I. (Mayo de 2006). "La Muestra Continua de Vidas Laborales de la Seguridad Social". (B. d. España., Ed.) Boletín Económico Banco de España.

Argimón, I., & González, C. I. & Vegas, R.(Abril de 2007a). "La situación laboral en las edades próximas a la jubilación. Una explotación de la Muestra Continua de Vidas Laborales". (B. d. España, Ed.) Boletín Económico Banco de España.

Argimón, I., & González, C. I. & Vegas, R. (2007b). "Jubilacion entre los 60 y los 65 años. Algunas características". Presupuesto y Gasto Público, 47-(2/2007) , pp 161-184.

Ayuso, J., Jimeno, J., & Villanueva, E. (2007). "The effects of the introduction of tax incentives on retirement savings". Banco de España, Documento de Trabajo nº 0724 .

Balmaseda, M., Melguizo, A., & Taguas, D. (2006). "Reforming the Spanish contributory pensions system". *MPRA Paper 19574, University Library of Munich, Germany, revised 01 Mar 2006.*

Barea, J. (1996). "Escenarios de evolución del gasto público en pensiones y desempleo en el horizonte 2020". *Documentos de Economía Pública, Fundación BBVA*.

Blanco, A. (2000). "The decision of early retirement in Spain". FEDEA, EEE 76.

Blau, D. M., & Gilleskie, D. B. (1998). "Labor Force Dynamics of Older Married Couples". Journal of Labor Economics , 16 (3), pp 595-629.

Blondal, S. (1998). "The retirement decision in the OECD countries". OECD Economics Dept. Working Paper No. 202.

Blondal, S., & Scarpetta, S. (1998). "Falling Participation Rates Among Older Workers in the OECD Countries: The Role of Social Security Systems". OECD Economics Department, Working Papers.

Blundell R., Meghir. C. and Smiths, S. (2002). "Pension incentives and the pattern of early retirement". *The Economic Journal* (112), pp 153-170.

Boldrin, M., Jimenez-Martín, S. and Peracchi, F.. (1999). "Social security and retirement in Spain". NBER papers in: *Social Security Programs and Retirement around the World* (pages. 305-353). En J. a. Gruber (Ed.), National Bureau of Economic Research, Inc.

Boldrin, M., Jimenez-Martín, S. and Peracchi, F... (2004). "Micro-modelling of retirement behaviour in Spain". In *Social Security programs and retirement around the world. Micro-estimation* (pages 499-578). J. G. Wise (eds.), The University of Chicago Press for the NBER.

Boldrin, M. y.-M. (2002). "Evaluating Spanish pension expenditure under alternative reform scenarios". *UPF Working Paper 652*.

Börsch Supan, A. (2003). "Incentive Effects of Social Security Under an Uncertain Disability Option",. En D. Wise (Ed.), *Themes in the Economic Aging,* (págs. pp 281-311). Chicago.

Börsch Supan, A. (2007). "Disability Insurance Enrolment and Health: A Comparartive European Perspective. En D. A. Wise, & D. Cutler (Edits.), *Work Disability.* University of Chicago.

Bover, O., Arellano, M., & Bentolila, S. (2002). "Unemployment duration, benefit duration and the business cycle". *The Economic Journal, 112. April* , pp. 223-265.

Buckley, J., and Westerland, C. (2004). "Duration dependence, functional form, and correct standard errors: Improving EHA models of state policy diffusion". *State Politics and Policy Quarterly, 4(1):94–113, 2004.*

Burtless, G. (1986). "Social Security, unanticipated benefit increases and the timing of retirement". *The Review of Economic Studies, 53.* , pp. 781-805c.

Cairó-Blanco, I. (2010). "An empirical analysis of retirement behaviour in Spain". *Spanish Economic Association and Fundación SEPI* , DOI 10.1007/s13209-009-0009-4.

Coile, C. (2000). "Social Security incentives for retirement" . *NBER DP 7651, April.*

Coile, C. (2003). "Health Shocks and Couples' Labor Supply Decisions". *Center for Retirement Research, Working Paper 2003-08*.

Conde-Ruiz, J. I., & Galasso, V. (2003). "Early retirement". *Review of Economic Dynamics, 2.* , pp 12-36.

Conde-Ruiz, J. I & García, E. (2004). "Demografía y empleo de los trabajadores próximos a la jubilación en Cataluña". *FEDEA Documento de Trabajo 2004-29.*

Costa, D. L. (1998). "*The Evolution of Retirement: An American Economic History 1880 - 1990*". The University of Chicago Press, Chicago, IL.

De La Fuente, A., & Domenech, R. (2009). "Convergencia real y envejecimiento:retos y propuestas". *BBVA Working Papers nº 0906*.

Diamond, P., & Gruber, J. (1999). "Social security and retirement in the U.S.". En J. Gruber, & D. Wise (Edits.), *Social Security programs and retirement around the world.* Chicago: University of Chicago Press.

Díaz-Gimenez, J., & Díaz-Saavedra, J. (2009). "Delaying retirement in Spain". *Review of Economic Dynamics , 12* (1), pp. 147-167.

Díaz-Gimenez, J., & Díaz-Saavedra, J. (2009). "Las pensiones públicas en España: una reforma que no funciona". Documento de Trabajo Unicaja, Premio de Investigación Economica Unicaja.

Dorn, D. and Sousa-Poza, A. (2005) "Early retirement: free choice or forced decision?" *Cesifo Working Paper no. 1542. Category 4: labour markets. September*

Dorn, D. and Sousa-Poza, A. (2010) "Voluntary and involuntary early retirement: an international anlysis" *Applied Economics. Vol 45. Issue 4.*

Duval R. (2003) "The retirement effects of old-age pension and early retirement schemes in OECD countries" *OECD Economics Department Working Papers No. (2003)24. Paris: OECD.*

García Pérez, J. I., & Sánchez Martín, A. R., (2008a). "Social Security and the search behaviour of workers approaching retirement". *Working Paper. Universidad Pablo Olavide. Econ 08 03*.

García Pérez, J. I., & Sánchez Martín, A. R. (2008b). "Jubilación y búsqueda de empleo a edades avanzadas". *Actual 33. Fundación Centro de Estudios Avanzados. Centro de Estudios Andaluces*.

Gokhale, J., Kotlikoff, L. J., & Warchawsky, M. J. (March, 2001). "Life Cycle Saving, Limits on Contributions to DC Pension Plans, and Lifetime Tax Benefits". *NBER Working Paper 8170*.

Gómez, A., & Hernandez de Cos, P. (2004). "An introductory analysis of the incentives for old workers to abandon the labour market".Mimeo.

Gruber, J., & Wise, D. (1999). *Social Security programs around the world.* Chicago: University of Chicago Press.

Gruber, J., & Wise, D. (2004). *Social Security programs and retirement around the world. Micro-estimation.* Chicago: Chicago, University of Chicago Press.

Gruber, J., & Wise, D. (2007).”Social security programs and retirement around the world: Fiscal implications”. The Chicago University Press.

Gustman, A. L. (1999). "What people don't know about their pensions and social security: An analysis using linked data from the Health and Retirement Study". (W. G. Gale, Ed.) *NBER Working Paper 7368 (Sept) Public Policies and Private Pensions.*

Gustman, A. L. & Steinmeier, T. L. (2002) "The social security early entitlement age in a structural model of retirement and wealth". *Journal of Public Economics , 89* (Issues 2-3), pp 441-463.

Gustman, A. L. & Steinmeier, T. L. (2004). "Personal Accounts and Family Retirement". *NBER Working Paper 10305*.

Gutiérrez-Domènech, M. (2006). “The employment of older workers” . *Working Paper No. 04/2006. La Caixa* .

Hakola, T. (2002). “Economic incentives and labour market transitions of the aged Finnish workforce”. VATT research report.

Hansen, F. (2000)."New Trends in plans". *Compensation and Benefits Review (32: 2)* , pp. 17 - 21.

Herce, J. A., & Alonso Messeguer, J. (1999). "Creación de Empleo y déficit del sistema de pensiones". *Cuadernos Económicos del ICE, 65*.

Herce, J. A.(2001). "La privatización de las pensiones en España". *Documento de Trabajo de FEDEA 2001-01*

Herce, J. A. García, E., & Jimeno, J. F. (2005). "La reforma de las pensiones: el papel de los mercados financieros". Fundación Caixa Galicia-CIEF.

Herce, J. A., Fernandez Pérez, J. L., Martín Barragán, E., Chuliá Rodrigo, E., Macmanus, P., Alonso Talon, P., y otros. (2009). "*Los retos socio-económicos del envejecimiento en España Resumen y conclusiones*". Analistas Financieros Internacionales. Madrid: AFI.

Imrohoroglu, A., Imrohoroglu, S., & Joines, D. (1999). "Computing models of social security". En R. Marimón, & A. Scott (Edits.), *Computational methods for the study of dynamic economies.*

Stock. J.H., & Wise, D. A. (1990). “Pensions, the option value of work and retirement”. *Econometrica, 58, ,* pp 1151-1180.

Jenkins, S. (1995). "Practitioner's Corner: Easy Estimation Methods for Discrete-Time Duration Models". *Oxford Bulletin of Economics and Statistics 57(1)* , pp 129-138.

Jenkins, S. (2005). "Survival Analysis" the Lecture Notes manscript, mimeo. available on http://www.iser.essex.ac.uk/iser/teaching/module-ec968.

Jimenez-Martín, S. (2006). "Evaluating the labor supply effects of alternative reforms of the Spanish pension system " . *Moneda y Crédito nº 222,* pp. 271-312.

Jimenez-Martín, S. & Sánchez A. (2000). "Incentivos y reglas de jubilación en España". *Cuadernos Económicos de ICE. Nº 65* , pp. 45-88.

Jimenez-Martín, S. & Sánchez A. (2006). "An evaluation of the life-cycle effects of minimum pensions on retirement behaviour: Extended Version,". *Working Papers 06.23, Universidad Pablo de Olavide, Departamento de Economía.*

Jimenez-Martín, S. ,Labeaga, J.M., & Vilaplana Prieto, C. (2006a). "A sequential model of older workers' labor force transitions after a health shock". *Health Economics,15* , 1033-1054.

Jimeno, J.F. (2001). "El sistema de pensiones contributivas en España: Cuestiones básicas y perspectivas en el medio plazo" *Nuevas Fronteras de la Política Económica, 2000, Teresa Garcia-Milá editor; CREÍ*

Jimeno, J. F. ;Rojas, J., & Puente, S. (2006). "Modelling The impact of ageing in Social Security expenditures". *Banco de España. Documento Ocasional 0601.*

Kim, S. (2003). "The Impact of Research Productivity on Early Retirement of University Professors". *Industrial Relations , 42*, pp. 106-125.

Krause, N. (1999). "Assessing Change in Social Support During Late Life". *Research on Aging, 21* (4), pp. 539-569.

Labeaga, J. (2008). "Estudios sobre las decisiones de retiro, la sostenibilidad del sistema de pensiones en España y la viabilidad financiera del Sistema Nacional de Dependencia". *Colección Estudios Económicos 04-08 Cátedra Nuevos Consumidores / FEDEA-BB*.

Licandro, O., & Jimeno, J. F. (1999). "La Tasa interna de rentabilidad y equilibrio financiero del sistema español de pensiones de jubilación". *Investigaciones Económicas XXIII (1)* , 129-143.

Lopez Garcia, M. (1990). "Efectos de las pensiones de Seguridad Social sobre la oferta de trabajo en España: un comentario". *Investigaciones Económicas , Volumen XXIV, 2*, 305-310.

Lumsdaine R. L., Stock, J. H and Wise, D.A. (1992) "Three Models of Retirement: Computational Complexity Versus Predictive Validity" , in Wise, D.A. (Ed.), *Topics in the Economics of Aging*. University of Chicago Press, Chicago, pp. 21–60.

Maestas, N. (2001). "Labor, Love & Leisure: Complementarity and the Timing of Retirement by Working Couples". *Working Paper UC- Berkeley, Department of Economics (2001)*

Martín Marcos, A., & Moreno Martin, L. (1990). "Efectos de las pensiones de Seguridad Social sobre la Oferta de trabajo en España: un comentario". *Investigaciones Económicas ,Volumen XXIV, 2*, 305-310.

Martínez, F. (1999). "Historiales de la Seguridad Social". Universidad Carlos III. Mimeo.

Messeguer, J. (2006). "La contribuitividad y el desequilibrio actuarial de las pensiones de jubilación en España". En *La Reforma de las Pensiones: Europa, Latinoamerica y España* (Vol. 4 , págs. 76-86). Panorama Social, Segundo Trimestre.

Montero Muñoz, M. (2000). "Estructura demográfica y pensiones un análisis de equilibrio general aplicado a la economía española". Investigaciones Economicas , XXIV (2), 297-327.

Montizaan, R., Cörvers, F., & De Grip., A. (2007). Training and early retirement". *Research Centre for Education and the Labour Market. ROA-RM-2007/3, July 07. The Netherlands.*

Moral-Arce, I., Patxot, C., & Souto, G. (2009). "Fostering the contributory nature of the Spanish Retirement Pension System: an arithmetic microsimulation exercise using the MCVL. Madrid: Instituto de Estudios Fiscales.

Muñoz, J. (1995). "Un análisis de los determinantes de la edad de jubilación mediante modelos de duración". *Documento de Trabajo 9523, CEMFI.*

Norton, E., Wang, ,. H., & Ai, C. (2004). "Computing interaction effects and standard errors in logit and probit models". *The Stata Journal, 4, number 2* , pp. 154-167.

Peracchi, F. a. (1994). "Trends in labor force transitions of older men and women". *Journal of Labor Economics* (12), pp 210-42.

Pienta, A. M. (1997). "Older Couples: An Examination of Health and Retirement within the Context of the Family". *Penn State University Population Research Institute, Working Paper Series 97-03*.

Pienta, A. M. (1999). "Partners In Marriage: An Analysis of Family Effects and the Likelihood of Retirement". *Penn State University Population Research Institute, Working Paper Series 98-05*.

Pienta, A. M. (2003). "Partners In Marriage: An analysis of husbands' and wives' retirement behavior". *Journal of Applied Gerontology* (22), pp. 340-358.

Poterba, J. M., Venti, S. F., & Wise, D. A. (1995). "Lump-Sum Distributions from Retirement Saving Plans: Receipt and Utility". *NBER Working Paper 5298*.

Poterba, J. M., Venti, S. F., & Wise, D. A. (2000). "Saver Behavior and Retirement Wealth". *American Economic Review , 90*, pp. 297- 302.

Poterba, J. M., Venti, S. F., & Wise, D. A. (1999). "Implications of rising personal retirement saving". *NBER working paper 6295*.

Prieto Romero, J., & Álvarez, S. (2002). "Estado de salud y participación laboral de las personas mayores". *Papeles de Trabajo 15/02. Instituto de Estudios Fiscales. Madrid.*

Quinn, J. F. (2000). "New Paths to Retirement," in *Forecasting Retirement Needs and Retirement Wealth*, ed. Olivia S. Michell, P. Brett Hammond, and Anna M. Rappaport (Philadelphia: University of Pennsylvania Press, 2000): 13-14.

Quinn, J. F. (2000). *Retirement Trends and Policies to Encourage Work Among Older Americans.* Boston College Working Papers in Economics 436, Boston College Department of Economics.

Saéz, M. J., & Taguas, D. (2006). La Reforma de las Pensiones. En *La Reforma de las Pensiones: Europa, Latinoamérica y España* (Vol. 4 Segundo Trimestre, págs. 46-60).

Samwick, A. (1998). "New evidence on pensions, social security, and the timing of retirement". *Journal of Public Economics, 70* , pp 207-236.

Sánchez Martín, A. R. (2003). "Jubilación endógena y sistema público de pensiones en España". *Universidad Carlos III WE 072515*.

Sánchez Martín, A. R. (2005). "Endogenous retirement and public pension system reform in Spain". *Mimeo.*

Sánchez Martín, A. R., & Sánchez Marcos, V. (2008). "Demographic change, pension reform and redistribution in Spain". DOCUMENTO DE TRABAJO FEDEA 2008-14 Serie Capital Humano y Empleo.

SEGURIDAD SOCIAL (2006). "*La muestra continua de vidas laborales*". Colección informes y estudios. Serie Seguridad Social, 24.

Smith Brothers, L. A. "*Retirement literature: An interdisciplinary review of preretirement influences on attitudes and decisions regarding retirement,* University of Wisconsin - Madison mimeo.

Spataro, L. (2002). "New tools in micro-modelling retirement decisions: overview and applications to the Italian case". *Centre for Research on Pensions and welfare policies. WP 28*.

Spataro, L.(2005). "Social security incentives and retirement decisions in Italy: an empirical insight". *Research in Economics, 59;* , pp 223-256.

Stock, J. H., & Wise, D. A. (1990). "The Pension Inducement to Retire: An Option Value Analysis", in *Issues in the Economics of Aging.* In: Issues in the Economics of Aging, (pages 205-230) National Bureau of Economic Research, Inc Chicago: University of Chicago Press.

Utrilla De La Hoz, A., & Ubago, Y. (2005). "Efectos redistributivos del sistema de pensiones de la Seguridad Social y factores determinantes de la elección de la edad de jubilación. Un análisis por Comunidades Autónomas". *Papeles de Trabajo 3/05 Instituto de Estudios Fiscales. Madrid.*

Villagarcía, T. (1995). "Análisis econométrico del tránsito a la jubilación par trabajadores en edad avanzada". *Investigaciones Económicas. , Vol. XIX* (1), pp.65-81.

Whitehouse, E, D'addioa, A.,Chomika, R. and Reilly A. (2009) "The Future of Pensions and Retirement Income Two Decades of Pension Reform: What has been Achieved and What Remains to be Done" The Geneva Papers 34, 515–535.

APPENDIX 1: DATA

This appendix contains the definition of the variables included in the different specifications. As already been mentioned in the text, the data source is the Continous Sample of Working Histories (CSWH) 2006, "*Muestra Continua de Vidas Laborales*" in Spanish), a sample of administrative data gathered by the Social Security.

The main descriptive statistics for each variable are presented in Table A1. Notice that not all individuals in the sample are either working or have claimed an old-pension during 2006 since some are unemployed and receive benefits, some do not receive benefits but are also contributing, and some do not contribute at all. Moreover, some go through different labour situations before claiming the pension; cf.Argimón, González and Vegas (2007a).

(A) Economic Incentive Variables:

To calculate the Social Security benefits to which individuals in the sample are entitled, I make use of the Social Security covered earnings histories of individual in the CSWH 2006.

SSW $_{it}$*:* Value of Social Security Wealth of individual i at time t, at 2006 prices:

$$SSW(r)_{it} = \sum_{s=t}^{s=L} [B_i(s,r)[p(s \mid t)_i /(1+\rho_i)^{s-t}]],$$

indicates Social Security Wealth at time t (at age t) if retiring at age r, L is the maximum life length, $B_i(s,r)$ is the pension benefit in period s (at age s) if retiring at r, p(s | r) is the conditional probability of an individual at time t to be alive at time s where s>r, ρ_i is the individual discount rate.

$$SSW(r, r+j) = SSW(r+j)/(1+\rho)^j - SSW(r)$$

To calculate the pension I make use of data on *covered earnings* and from it I have built the Regulatory Base which has been computed, following the regulation , as a moving average of the contribution bases in the 15 years immediately before retirement. The minimum base that has been used to complete job careers has been the one corresponding to contributory group 5, senior administrative ("*oficial administrativo*"), the group with the largest volume of population. On the other hand, the maximum base has been taken to be the one corresponding to group 1, Engineers and Graduates ("*Ingenieros y Licenciados*"), the group with the highest base for all the years. The maximum life length (T) is assumed to be 98 years; ρ_i, the individual discount rate is assumed to be fixed at 3%, p(s | r), the conditional probability of an individual aged r to be alive at age s, has been taken from the National Statistics Institute (INE) demographic projections (scenario 2), based upon 2001 Census data.

Pensions are assumed to increase 2% yearly from 2006. Upper and lower limits on the pension benefits are applied to compute retirment pension (the minimum one corresponds to a worker with a dependent spouse). Not considering heterogeneity on survival probability among individuals with different income levels, as well as on discount rates, could be understood as a limitation on the assumptions behind the incentive variables used in the estimation. However, heterogeneity on official estimates of life expectancy by socioeconomic level are not available in Spain, and though it were possible to estimate them, it is presumable they would not differ too much, since the country is not characterized by such remarkable levels of wealth inequality. Maybe additional research would be needed to analyze the impact of considering different discount rates, also linked to income level. Assumed discount rates, and minimum and maximum pension, are homogeneous to ensure simplicity in the analysis.
In order to calculate the different incentive measures, I need to project SSW for the future. Two different situations arise, depending on the age of the individual and whether or not he has retired. For those that are not 66 before 2006, I need to project their pension and their SSW beyond this year. To do so, I assume that their salary and, therefore, their contributory base will be increasing at a 2% rate every year. For those that have retired before 2006, I project their salaries for the years before 2006 assuming that they keep the purchasing power of their last observed salary (or the following one), so that the contributory base increases by the same amount as the December over December CPI

SSA $_{it}$*:* The accrual rate,

$$SSA(r)_{it} = 1/(1+\rho_i)SSW(r+1)_{it} - SSW(r)_{it}$$

and we let

$$SSW(r+1)_{it} = \frac{SSW(r+1)_{it+1}}{(1+\Pi_t)}$$

A limitation of this index is that it does not take into account the comparison that the individual can make between pension benefits and the level of his/her income. It could be argued that the leisure preference is such that wages can fully compensate for the forgone leisure enjoyment from postponing retirement.

PV $_{it}$*:* Peak Value computed between the ages of 60 and 65 is defined as,

$$PV_i(r) = \max(SSW(r,r+1)_i, SSW(r,r+2)r_i, SSW(r,r+3)_i ..., SSW(r,66)_i)$$
$$= SSA_i, otherwise$$

I follow Coile and Gruber (2000) and restrict the peak value to be equal to the accrual rate, if the individual works beyond the highest value for his social security wealth. They also normalize the peak value by the expected stream of wages over the period between the maximum year and the current year. Hence their actual index measures the benefits of continued work relative to the social security wealth earnings in the same period. They call this measure the tax/subsidy rate. This normalization can also be made for both the accrual and the option value.

RR $_{it}$: Replacement rate, $RR_i(r)$ is the ratio of the expected pension benefits $B_i(r)$ at time t over wages $w_i(r)$ received at time t-1 for individual *i* at age *r*, if the person retires at age *r*.

$RR_i(r) = E_r(B_i(r)/w_i(r))$ where E is the expectation operator.

(B) Other variables

failurejubi $_{it}$: Dummy variable that takes value 1 if the person retires at time t and 0 otherwise. It is the dependent variable.

disab $_{it-1}$**:** Dummy variable that takes value 1 at time t if the person was receiving any disability benefit while he was a year younger (at time t-1) and zero, otherwise.

univ $_i$: Dummy variable that takes value 1 if the contributory group ("*grupo de cotización*") of the longest contributory relationship with the Social Security system is the one with the highest academic qualifications (group 1: "Engineers and Graduates"), and zero, otherwise.

numrel $_i$: Number of contributory labour relations that have been recorded by the Social Security before becoming entitled to an old age pension and that include those involving the perception of unemployment benefits.

Regional Government (Comunidad Autónoma) where the worker initially registered: Group of 19 dummy variables, each one corresponding to a CA, plus one for Ceuta and one for Melilla, that records the initial worker's registration. (IIccaa-)

serv $_i$**.** Dummy variable that takes value 1 if the longest job a person has held has taken place in the following CNAE sector classifications: Trade (50 to 52), Restoration (Hostelería) (55), Transport (60 to 64), other services, including education y health (65 to 67, 70 to 74, 80, 85 and 90)

u $_{it-1}$**.** Dummy variable that takes value 1 at time t if the person was receiving unemployment benefits, either as a subsidy or a contributory transfer, while he was a year younger (at time t-1), and zero, otherwise. That is, people whose relationship with Social Security is coded as a *TRL 751-756* in the administrative files.

l $_{it-1}$. Dummy variable that takes value 1 at time t if the person was working and contributing to Social Security while he was a year younger (at time t-1), and value zero, otherwise.

g_k$_{it}$**:** Dummy variables that take value 1 if the person is at time t in the k–th period decision and cero otherwise, where k=[1,6]. That is, **g_k** takes value 1 if the value of the length of the spell from the year the person becomes entitled to a retirement pension is **k**.

age_k$_{it}$**:** Dummy variables that take value 1 if the person is k years old at time t and cero otherwise, where k=[60,65].

cycle $_t$ **:** Spanish GDP real growth rate (for years 1997 to 2006)

r2002: Dummy variable that takes value 1 if the year of the observation is greater than 2001.

otherben $_{it-1}$**:** Dummy variable that takes value 1 at time t if the person was receiving any Social Security benefit other than disability, old age or unemployment while he was a year younger (at time t-1) and zero, otherwise.

meanlength $_i$**:** Average number of years for the spells that the individual i has had before becoming entitled to a pension.

t:a linear time trend

Table A 1
Descriptive values. Sample of
men born between 1936 and 1946, having worked in the General Regimen, with a relation with the Social Security in 2006

	All period considered								
	1996-2006			Before 2002			After 2002		
	median	mean	*s.d.*	median	mean	*s.d.*	median	mean	*s.d.*
SSW_t[(1)]	199.30	**222.50**	*96.83*	195.83	**211.84**	*87.02*	201.35	**228.17**	*101.21*
SSA_t[(1)]	6.89	**6.73**	*12.95*	10.12	**9.79**	*12.20*	4.97	**5.09**	*13.04*
PV_t[(1)]	12.65	**18.84**	*26.73*	21.68	**26.77**	*28.03*	8.32	**14.63**	*25.01*
RR_t[(2)]	57.61	**67.36**	*35.50*	50.77	**57.61**	*26.56*	61.93	**72.55**	*38.45*
l_{t-1}[(2)]		**57.49**			**62.67**			**54.74**	
u_{t-1}[(2)]		**24.28**			**27.54**			**22.55**	
numrel		**10.770**	*25.866*		**9.547**	*19.763*		**11.420**	*28.565*
meanlenght[(3)]		**7.237**	*7.243*		**7.763**	*7.436*		**6.957**	*7.122*
time since eligible[(3)]		**2.677**	*1.565*		**2.275**	*1.390*		**2.892**	*1.609*
univ		**0.106**	*0.307*		**0.101**	*0.302*		**0.108**	*0.310*
serv		**0.364**	*0.481*		**0.356**	*0.479*		**0.367**	*0.482*
disab		**0.126**	*0.332*		**0.024**	*0.152*		**0.181**	*0.385*
otherben		**0.134**	*0.341*		**0.030**	*0.172*		**0.190**	*0.392*
low60		**0.086**	*0.281*		**0.101**	*0.302*		**0.078**	*0.268*
low61		**0.047**	*0.211*		**0.034**	*0.181*		**0.053**	*0.224*
low62		**0.029**	*0.168*		**0.015**	*0.123*		**0.036**	*0.187*
low63		**0.018**	*0.134*		**0.008**	*0.091*		**0.024**	*0.152*
low64		**0.011**	*0.104*		**0.005**	*0.067*		**0.014**	*0.120*
low65		**0.008**	*0.086*		**0.003**	*0.053*		**0.010**	*0.100*
top64		**0.001**	*0.037*		**0.001**	*0.025*		**0.002**	*0.042*
top65		**0.003**	*0.058*		**0.001**	*0.034*		**0.005**	*0.067*
age61		**0.219**	*0.413*		**0.236**	*0.425*		**0.209**	*0.407*
age62		**0.175**	*0.380*		**0.158**	*0.365*		**0.184**	*0.387*
age63		**0.135**	*0.342*		**0.110**	*0.313*		**0.148**	*0.355*
age64		**0.100**	*0.300*		**0.068**	*0.252*		**0.118**	*0.322*
age65		**0.065**	*0.246*		**0.027**	*0.162*		**0.085**	*0.278*
cycle[(4)]		**3.667**	*0.679*		**4.262**	*0.688*		**3.350**	*0.406*
r2002		**0.653**	*0.476*		**0.000**	*0.000*		**1.000**	*0.000*

(1) In thousands of euros

(2) In percentage

(3) In years

(4) In percentage variation of GDP

Table A 2

Logit estimates of the effects of pension incentives on retirement behaviour between 60 and 65 years of age. Males born between 1936 and 1946, having worked in the General Regime with a relation with the Social Security in 2006(*)

	MODEL A		**MODEL B**		**MODEL C**	
	ACCRUAL	PEAK VALUE	ACCRUAL	PEAK VALUE	ACCRUAL	PEAK VALUE
	LR chi2(45)=38014.62	LR chi2(49)=38187.03	LR chi2(56)=38597.09	LR chi2(45)=38030.83	LR chi2(49)=38213.54	LR chi2(56)=38607.93
	Pseudo R2=0.365	Pseudo R2=0.367	Pseudo R2=0.371	Pseudo R2=0.365	Pseudo R2=0.367	Pseudo R2=0.371
	Log likelihood = -33064.09	Log likelihood = -32977.89	Log likelihood = -32772.86	Log likelihood = -33055.97	Log likelihood = -32964.63	Log likelihood = -32767.44
	115.532 number of observations					
	ECONOMIC INCENTIVES AND REGULATION					
SSW_t	**2.979**	**2.969**	**2.462**	**2.401**	**2.322**	**2.289**
	(23.08)***	(23.31)***	(11.69)***	(11.81)***	(10.78)***	(10.83)***
SSA_t	**-7.993**		**-12.941**		**-10.510**	
	(-7.07)***		(-7.78)***		(-5.69)***	
PV_t		**-4.284**		**-6.730**		**-5.363**
		(-8.11)***		(-9.26)***		(-6.92)***
RR_t	**0.046**	**0.043**	**-0.083**	**-0.144**	**-0.085**	**-0.139**
	(1.2)	(1.14)	(-1.25)	(-2.14)*	(-1.28)	(-2.06)*
SSW_t * r2002			**1.109**	**1.166**	**1.272**	**1.252**
			(4.76)***	(5.16)***	(5.48)***	(5.53)***
SSA_t * r2002			**8.821**		**4.018**	
			(4.50)***		(2.04)*	
PV_t * r2002				**4.451**		**3.265**
				(4.94)***		(3.61)***
RR_t * r2002			**0.232**	**0.300**	**0.151**	**0.252**
			(3.04)**	(3.89)***	(1.97)*	(3.26)**
r2002			**-0.763**	**-0.843**	**-0.689**	**-0.787**
			(-10.37)***	(-11.04)***	(-9.31)***	(-10.24)***
low60					**0.205**	**0.224**
					(4.45)***	(5.01)***
low61					**-0.035**	**-0.001**
					(-0.46)	(-0.01)
low62					**-0.587**	**-0.554**
					(-5.68)***	(-5.38)***
low63					**-0.922**	**-0.878**
					(-6.42)***	(-6.14)***
low64					**-1.885**	**-1.899**
					(-9.73)***	(-9.80)***
top64					**-0.546**	**-0.383**
					(-2.26)*	(-1.61)
top65					**-1.509**	**-1.415**
					(-11.52)***	(-11.07)***
	WORKING STATUS, SECTOR OF ACTIVITY AND OTHER LABOUR HISTORY RELATED VARIABLES					
l_{t-1}	**-1.577**	**-1.577**	**-1.551**	**-1.556**	**-1.575**	**-1.577**
	(-39.51)***	(-39.52)***	(-38.65)***	(-38.77)***	(-39.05)***	(-39.10)***
u_{t-1}	**0.942**	**0.934**	**0.965**	**0.952**	**0.954**	**0.947**
	(25.21)***	(24.98)***	(25.72)***	(25.33)***	(25.31)***	(25.09)***
numrel	**-0.022**	**-0.022**	**-0.021**	**-0.021**	**-0.023**	**-0.023**
	(-11.82)***	(-11.74)***	(-11.11)***	(-11.06)***	(-11.74)***	(-11.66)***
t x numrel	**0.004**	**0.004**	**0.004**	**0.004**	**0.004**	**0.004**
	(9.03)***	(8.97)***	(8.68)***	(8.65)***	(9.10)***	(9.04)***
meanlenght	**0.002**	**0.002**	**0.003**	**0.003**	**0.005**	**0.005**
	(0.75)	(0.73)	(0.96)	(1.01)	(1.57)	(1.58)
t x meanlenght	**0.003**	**0.003**	**0.003**	**0.003**	**0.002**	**0.002**
	(3.85)***	(3.92)***	(3.43)***	(3.40)***	(2.56)*	(2.61)**
serv	**-0.003**	**-0.001**	**0.020**	**0.022**	**0.023**	**0.024**
	(-0.07)	(-0.02)	(0.50)	(0.54)	(0.58)	(0.60)
t x serv	**-0.020**	**-0.022**	**-0.024**	**-0.026**	**-0.027**	**-0.028**
	(-1.67)	(-1.81)*	(-2.01)*	(-2.13)*	(-2.22)*	(-2.30)*

(cont…)

(cont…)

PERSONAL CHARACTERISTICS & OTHERS						
univ	**-0.729**	**-0.751**	**-0.706**	**-0.716**	**-0.779**	**-0.794**
	(-10.46)***	(-10.82)***	(-10.10)***	(-10.28)***	(-10.95)***	(-11.18)***
t x univ	**-0.078**	**-0.068**	**-0.086**	**-0.079**	**-0.038**	**-0.029**
	(-4.02)***	(-3.57)***	(-4.39)***	(-4.15)***	(-1.89)	(-1.49)
disab	**-1.578**	**-1.586**	**-1.468**	**-1.485**	**-1.484**	**-1.498**
	(-8.28)***	(-8.34)***	(-7.64)***	(-7.74)***	(-7.70)***	(-7.78)***
otherben	**-2.981**	**-2.970**	**-2.978**	**-2.972**	**-3.016**	**-3.015**
	(-17.55)***	(-17.54)***	(-17.39)***	(-17.41)***	(-17.54)***	(-17.55)***
age61	**-0.015**	**-0.037**	**-0.069**	**-0.095**	**0.098**	**0.069**
	(-0.06)	(-0.15)	(-0.29)	(-0.40)	(0.40)	(0.28)
age62	**-1.047**	**-1.076**	**-1.055**	**-1.087**	**-0.767**	**-0.791**
	(-3.70)***	(-3.80)***	(-3.70)***	(3.81)***	(-2.48)*	(-2.56)*
age63	**-0.710**	**-0.739**	**-0.725**	**-0.759**	**-0.098**	**-0.134**
	(-2.50)*	(-2.59)**	(-2.54)*	(-2.65)**	(-0.32)	(-0.44)
age64	**-0.712**	**-0.759**	**-0.686**	**-0.739**	**0.247**	**0.213**
	(-2.46)*	(-2.63)**	(-2.36)*	(-2.55)*	(0.78)	(0.67)
age65	**2.813**	**2.788**	**2.865**	**2.834**	**3.271**	**3.272**
	(10.94)***	(10.84)***	(11.09)***	(10.96)***	(12.34)***	(12.33)***
cycle	**0.241**	**0.246**	**0.111**	**0.119**	**0.125**	**0.131**
	(16.89)***	(17.18)***	(5.94)***	(6.42)***	(6.69)***	(7.03)***
DURATION VARIABLES						
g1	**-0.514**	**-0.540**	**-0.591**	**-0.615**	**-0.331**	**-0.370**
	(-1.96	(-2.06)*	(-2.24)*	(-2.33)*	(-1.23)	(-1.37)
g2	**-1.703**	**-1.733**	**-1.697**	**-1.723**	**-1.539**	**-1.572**
	(-6.76)***	(-6.88)***	(-6.74***	(-64.85)***	(-6.02)***	(-6.15)***
g3	**-0.522**	**-0.557**	**-0.538**	**-0.567**	**-0.442**	**-0.483**
	(-2.06)*	(-2.20)*	(-2.11)*	(-2.22)*	(-1.64)	(-1.79)
g4	**-0.858**	**-0.904**	**-0.856**	**-0.891**	**-1.094**	**-1.128**
	(-3.52)***	(-3.72)***	(-3.50)***	(-3.65)***	(-4.32)***	(-4.45)***
g5	**0.054**	**0.027**	**0.028**	**0.016**	**-0.494**	**-0.519**
	(0.22)	(0.11)	(0.12)	(0.07)	(-1.95)	(-2.05)*
cons	**-2.328**	**-2.264**	**-1.422**	**-1.306**	**-1.731**	**-1.635**
	(-3.82)***	(-3.71)***	(-2.30)*	(-2.12)*	(-2.76)**	(-24602)**

(*) Notes:

Dependent variable is 1 if person retires and 0 otherwise, conditioned on not having retired before. Monetary values are in 1,000,000 euros, prices 2006. Z-values are in parenthesis.

Statistical significance: * p<0.05; ** p<0.01; *** p<0.001. Coefficients of Autonomous Communities not resported. Prob > chi2=0.

Model A represents the basic model. Model B includes Model A and changes in 2002 regulation as control variables. Model C includes Model B and income levels as control variables.

Table A 3
Quantitative effects of pension incentives and other variables on the average hazard rate by age

	ACCRUAL							PEAK VALUE						
	All ages	**Age 60**	**Age 61**	**Age 62**	**Age 63**	**Age 64**	**Age 65**	**All ages**	**Age 60**	**Age 61**	**Age 62**	**Age 63**	**Age 64**	**Age 65**
Economic Incentives and Regulation														
SSW_t	0.00604	0.00717	0.00377	0.00462	0.00515	0.00878	0.00910	0.00604	0.00716	0.00377	0.00462	0.00514	0.00875	0.00912
	0.0029	*0.0043*	*0.0013*	*0.0019*	*0.0022*	*0.0064*	*0.0089*	*0.0029*	*0.0043*	*0.0013*	*0.0019*	*0.0023*	*0.0064*	*0.0089*
SSA_t	-0.00075	-0.00115	-0.00048	-0.00046	-0.00041	-0.00071	-0.00147							
	-0.0003	*-0.0006*	*-0.0002*	*-0.0002*	*-0.0002*	*-0.0003*	*-0.0011*							
PV_t								-0.00072	-0.00145	-0.00052	-0.00045	-0.00031	-0.00038	-0.00080
								-0.0003	*-0.0007*	*-0.0002*	*-0.0002*	*-0.0001*	*-0.0002*	*-0.0006*
RR_t	0.00026	0.00027	0.00014	0.00018	0.00022	0.00043	0.00055	0.00024	0.00025	0.00013	0.00017	0.00020	0.00039	0.00052
	0.0001	*0.0001*	*0.0000*	*0.0001*	*0.0001*	*0.0003*	*0.0005*	*0.0001*	*0.0001*	*0.0000*	*0.0001*	*0.0001*	*0.0002*	*0.0005*
Individual Characteristics														
numrel	-0.00078	-0.00186	-0.00069	-0.0005716	-0.00036	-0.0001854	0.0002756	-0.00077	-0.00184	-0.00068	-0.0005661	-0.00036	-0.0001858	0.0002722
	-0.0002	*-0.0009*	*-0.0002*	*-0.0002*	*-0.0001*	*-0.0001*	*0.0002*	*-0.0002*	*-0.0008*	*-0.0002*	*-0.0002*	*-0.0001*	*-0.0001*	*0.0002*
l_{t-1}	-0.13012	-0.17677	-0.07608	-0.09285	-0.10215	-0.19861	-0.15740	-0.13028	-0.17694	-0.07612	-0.09294	-0.10232	-0.19897	-0.15764
	-0.1160	*-0.1831*	*-0.0699*	*-0.0870*	*-0.0969*	*-0.2147*	*-0.1470*	*-0.1162*	*-0.1836*	*-0.0697*	*-0.0870*	*-0.0971*	*-0.2151*	*-0.1476*
u_{t-1}	0.07919	0.10601	0.04691	0.05679	0.06200	0.11593	0.10557	0.07862	0.10523	0.04654	0.05634	0.06152	0.11515	0.10498
	0.0594	*0.0885*	*0.0310*	*0.0392*	*0.0440*	*0.1076*	*0.1097*	*0.0589*	*0.0880*	*0.0304*	*0.0388*	*0.0438*	*0.1070*	*0.1092*
uni	-0.05719	-0.07249	-0.03245	-0.03880	-0.04222	-0.07833	-0.10884	-0.05799	-0.07355	-0.03281	-0.03922	-0.04265	-0.07925	-0.11126
	-0.0281	*-0.0390*	*-0.0126*	*-0.0164*	*-0.0187*	*-0.0516*	*-0.1128*	*-0.0285*	*-0.0398*	*-0.0127*	*-0.0166*	*-0.0190*	*-0.0526*	*-0.1173*
serv	-0.00182	-0.00233	-0.00113	-0.00132	-0.00142	-0.00247	-0.00284	-0.00181	-0.00232	-0.00112	-0.00131	-0.00142	-0.00246	-0.00283
	-0.0010	*-0.0014*	*-0.0005*	*-0.0006*	*-0.0007*	*-0.0018*	*-0.0030*	*-0.0010*	*-0.0014*	*-0.0005*	*-0.0006*	*-0.0007*	*-0.0018*	*-0.0030*
disab	-0.09491	-0.11417	-0.04653	-0.05607	-0.06077	-0.12213	-0.26254	-0.09540	-0.11485	-0.04667	-0.05618	-0.06087	-0.12247	-0.26505
	-0.0379	*-0.0526*	*-0.0166*	*-0.0212*	*-0.0240*	*-0.0666*	*-0.3103*	*-0.0379*	*-0.0532*	*-0.0165*	*-0.0213*	*-0.0242*	*-0.0673*	*-0.3141*
otherben	-0.14350	-0.15770	-0.06152	-0.07475	-0.08122	-0.17335	-0.52006	-0.14352	-0.15826	-0.06153	-0.07457	-0.08095	-0.17288	-0.52009
	-0.0542	*-0.0686*	*-0.0224*	*-0.0290*	*-0.0324*	*-0.0886*	*-0.5875*	*-0.0546*	*-0.0696*	*-0.0225*	*-0.0289*	*-0.0324*	*-0.0886*	*-0.5882*

NOTES:

Quantitative effects of continuous variables are computed as the discrete differences of the logistic function evaluated at a 10% increase in the variables' values with respect to the logistic function evaluated at the observed variables' values. Results are obtained under Model A in Table A 2. Median values are in italics. The effects are estimated for the period beginning in 2002.

Table A 4

Quantitative effects of pension incentives and other variables on the average hazard rate for ages between 60 and 65

	MODEL B				MODEL C			
	ACCRUAL		PEAK VALUE		ACCRUAL		PEAK VALUE	
	mean	***median***	**mean**	***median***	**mean**	***median***	**mean**	***median***
Economic Incentives and Regulation								
SSW_t	0.00718	*0.0034*	0.00717	*0.0034*	0.00725	*0.0033*	0.00715	*0.0033*
SSA_t	-0.00038	*-0.0002*			-0.00060	*-0.0002*		
PV_t			-0.0003802	*-0.0001*			-0.00035	*-0.0001*
RR_t	0.00082	*0.0003*	0.00086	*0.0004*	0.00036	*0.0001*	0.00061	*0.0003*
r2002 overall change	-0.02257	*-0.0097*	-0.02225	*-0.0100*	-0.01929	*-0.0086*	-0.01953	*-0.0087*
r2002 change								
low60					0.01627	*0.0088*	0.01781	*0.0097*
low61					-0.00268	*-0.0014*	-0.00005	*0.0000*
low62					-0.04063	*-0.0189*	-0.03861	*-0.0181*
low64					-0.10067	*-0.0382*	-0.10122	*-0.0385*
top64					-0.03790	*-0.0178*	-0.02744	*-0.0132*
top65					-0.08607	*-0.0345*	-0.08238	*-0.0337*
Individual Characteristics								
numrel	-0.00067	*-0.0001*	-0.00066	*-0.0001*	-0.00071	*-0.0002*	-0.00071	*-0.0001*
l_{t-1}	-0.12357	*-0.1088*	-0.12392	*-0.1088*	-0.12525	*-0.1094*	-0.12556	*-0.1097*
u_{t-1}	0.07886	*0.0588*	0.07767	*0.0577*	0.07751	*0.0566*	0.07698	*0.0561*
uni	-0.05481	*-0.0264*	-0.05498	*-0.0265*	-0.05596	*-0.0268*	-0.05638	*-0.0271*
serv	-0.00035	*-0.0002*	-0.00032	*-0.0002*	-0.00030	*-0.0002*	-0.00032	*-0.0002*
disab	-0.08750	*-0.0345*	-0.08822	*-0.0347*	-0.08763	*-0.0345*	-0.08829	*-0.0348*
otherben	-0.13931	*-0.0509*	-0.13919	*-0.0512*	-0.13911	*-0.0507*	-0.13920	*-0.0509*

NOTES:

Quantitative effects of the continuous variables are computed as the discrete differences of the logistic function evaluated at a 10% increase in the variables' values with respect to the logistic function evaluated at the observed variables' values.

Results are obtained under Models B and C in Table A2. The effects are estimated for the period beginning in 2002.

APPENDIX 2: LEGISLATION

The institutional framework

The labour-market based social security is mandatory for workers in Spain. Old-age public pensions are mainly provided through three different schemes: Social Security General Scheme (*Régimen General de la Seguridad Social (RGSS)*), Social Security Special Schemes (*Regímenes Especiales de la Seguridad Social (RESS))* and Civil Service Scheme for government employees *(Régimen de Clases Pasivas (RCP)).* The RGSS and RESS are administered and managed by the Social Security as a joint pay-as-you-go system. The RCP is administered and managed by the Central Government. Around 72% of social security contributions are obtained from the RGSS. The pension regulations for RESS, within which the self-employed are assigned, and RCP do not, in general, allow for early retirement, so that the focus of the analysis will be on the RGSS.

The main changes that have taken place in Social Security regulation in recent years that affect the period covered by the sample correspond to the reforms introduced in 1997 and 2002 in relation to the framework set in 1985.

The normal retirement age in Spain, that is the age when a person becomes eligible for the full pension benefit, is 65. In fact, in some sectors retirement at 65 has been compulsory for some years. However, since 2002, incentives have been built in the regulation so as to promote retirement beyond the age of 65. The regulation increased the amount of the pension to be received if the worker remained employed and payment of social contributions by employers and employees with indefinite-term contacts were waived. Early retirement is possible from the age of 60, under some specific conditions and it imply a reduction in the amount of the pension to be received. The early retirement penalty is defined by a reduction coefficient that is detailed below.

Entitlement criteria for RGSS:

A payroll tax defined for both employers and employees and levied on earnings, with a minimum contribution and a maximum pensionable earning, finances the Social Security System.

Up to 1997 only 8 contributory years were required to be entitled to a retirement pension. The change introduced in 1997 set a timetable to extend this period to fifteen years, one every year, so that in 2002, 15 years were required to be able to receive a pension at the age of 65, the ordinary retirement age. An additional requirement introduced in 1997 was that two of these contributed years had to have taken place during the last eight years.

As for early retirement before the age of 65, there are three different cases:

On the one hand, those who, before January 1, 1967, contributed to the labour mutual funds system that preceded the current Social Security system are entitled to retire from the age of 60 if the total contributory years are at least 15.

For those whose first year of contribution to Social Security was after 1967 and only since the 2002 amendment, the earliest retirement age is 61. For them, the minimum number of contributory years amounts to 30. Moreover, in order to be able to claim the pension, they have to have spent at least six months involuntarily unemployed and registered as job seekers in the Public Employment Service Offices, during the period immediately preceding the pension claim. Years spent unemployed and receiving unemployment benefits add as contributory years towards an old-age pension. In fact, unemployed workers aged 52, and older, can receive unemployment benefits that turn into subsidies until they are eligible for early or ordinary retirement.

Retiring at the age of 64 is also possible and is subject to different rules: from 2002 no previous period of unemployment is required, but just the minimum 15 contributory years. However, in this case, the firm needs to hire another worker for a minimum period of a year (substitution contract) to replace the retiree, if full benefits (as at the age of 65) are to be guaranteed.

Regulation also allows for early retirement for professions, especially those involving dangerous or unhealthy activities or some instances of those affected by industrial restructuring regulated by special legislation. Moreover, a new regulation came into force in 2002 allowing for partial retirement that can be simultaneously enjoyed with a part-time job. Workers can partially retire starting at the age of 60 if the firm replaces the retiree with another worker (relief contract) to compensate for the retiree's reduction in work-time

The database does not provide information about contributions dating from before the seventies. On the other hand, all men born between 1936 and 1946 could potentially have been working by 1967, as the youngest would have started working at the age of 21. Therefore, the assumption in the empirical part is that all people in the sample contributed to the labour mutual funds system, so that they only need 15 years of contribution to be entitled to retire[1]. In fact, in the sample 34% of those I observe retiring, do so at 60 and 40% do so with less than 30 registered years of contribution.

The pension amount:

The amount of the old age pension is defined by the interaction of different elements. On the one hand, the Regulatory Base (Base reguladora, BR) that defines the amount upon which to calculate the pension rights is directly related to wages received, but subject to lower and upper caps. Different caps have been in place for different types of workers depending on their group of contribution (*grupo de cotización*) associated with type of job and education level The minimum and maximum contributory periods to be considered in its calculation and the inflation correction mechanism to obtain its present value are regulatory defined. Reduction coefficients for early retirement and for less than 35 years of contribution are also defined. There is a special treatment for those that contributed to the system before 1967 (mutual funds contributors or "mutualistas") Minimum and maximum pensions are yearly defined and its amount depends on marital status and number of economically dependent people who depend on the person receiving the benefit. Pensions are up-rated by the Consumer Price Index. Up to 2002, any additional year contributed beyond 35 did not add to the amount of pension received.

A person retiring between 1985 and 1996 with at least 15 contributory years at year t had a regulatory base defined as:

$$BR_t = \frac{1}{112}\left(\sum_{j=1}^{24} w_{t-j} + \sum_{j=25}^{96} w_{t-j}\frac{I_{t-25}}{I_{t-j}}\right)$$

where w_{t-j} are *covered earnings* for the jth month before retiring at t and I_{t-j} is the price index for the jth month before retirement, so that only eight years were taken into account to define *BR*. Notice that the weighted average was divided by 112 as pensions are paid in fourteen annual instalments.

Since 2002, the regulatory base is defined as:

$$BR_t = \frac{1}{210}\left(\sum_{j=1}^{24} w_{t-j} + \sum_{j=25}^{180} w_{t-j}\frac{I_{t-25}}{I_{t-j}}\right)$$

so that 15 years are taken into account. A transitory period was set from 1997 to 2002 such that each year, one additional year was included in the up-rated part of the weighted average and, therefore the fifteen years were finally accounted for the definition of the regulatory base in 2002 [2]. As social contributions are paid 14 months a year, the effective number of years taken into account to compute the regulatory base was 6.8 up to 1997 and 12.9 since then. The 2006 agreement proposes to rise the effective contributory years to 15, without taking into account the 14 monthly payments.The relation between the first monthly pension received at time *t* (B_t) and the regulatory base (BR_t) calculated at *t* can be expressed as $B_t = \alpha_{nt}^T \cdot BR_t$ where $\alpha_{nt}{}^T = \alpha_{nt}{}^y \alpha_{nt}{}^a$, so that

$\alpha_{nt}^{\ y}$ depends only on contributory years (n), and $\alpha_{nt}^{\ a}$ depends on the age of retirement.

If retirement age is equal or larger than 65 then, and up to 1997, $\alpha_{nt}^{\ a} = 1$ and $\alpha_{nt}^{\ T}$ is expressed as:

$$\alpha_{nt}^{T} = \begin{cases} 0 & if\ n < 15 \\ 0.6 + 0.02(n-15), & if\ 15 \le n < 35 \\ 1, & if\ 35 \le n \end{cases}$$

The reform introduced in 1997 modified the number of years to define the contributory base and the substitution rate ($a^T{}_n$) if age of retirement was equal or larger than 65, so that:

$$\alpha_{nt}^{T} = \begin{cases} 0, & if\ n < 15 \\ 0 \cdot 5 + 0.03(n-15), & if\ 15 \le n < 25 \\ 0 \cdot 8 + 0.02(n-25), & if\ 25 \le n < 35 \\ 1 & if\ 35 \le n \end{cases}$$

That new scheme thus implied a more progressive approach to full benefits. For early retirement, regulation also sets a penalization system linked to age. *Mutualistas* that retire early were subject to a reduction coefficient equivalent to 8% for each year in advance of 65 that he/she retires, so that $\alpha_t^a = 1 - 0.08(65 - r)$ *where* $r >= 60$. The 1997 reform reduced the reduction coefficient to 7% for those with more than 40 contributory years, when claiming the pension. This coefficient should be jointly applied with the one corresponding to contributory years.

The 2002 reform changed the penalization mechanism, so as to make the age coefficient (α_t^a) more linked to the number of contributed years, so that:

$$\alpha_{nt}^{\ a} = \begin{cases} 0 & if\ r < 61 \\ 1 - k(65 - r), & if\ 61 \le r < 65 \\ 1, & if\ r \ge 65 \end{cases} \quad where \quad k \begin{cases} 0.08 & if\ n = 30 \\ 0.075 & if\ 31 \le n \le 34 \\ 0.07 & if\ 35 \le n \le 37 \\ 0.065 & if\ 38 \le n \le 39 \\ 0.06 & if\ n \ge 40 \end{cases}$$

where r is retirement age.

Moreover, it introduced a premium for late retirement, so that the pension was increased by 2% per additional year if the worker credited more than 35 years of contribution The 2006 Agreement proposed raising the premium to 3% for those with more than 40 contributory years. The partial retirement regulation introduced with the 2002 reform, established that no correction coefficient for age would be used for those claiming this type of pension. The reform agreed in 2006 aims at rationalizing this type of retirement, requiring six years of seniority in the firm before retiring, 30 contributory years (instead of the current 15), and

changing the maximum and minimum labour day reduction to 75% and 25%, respectively from the current 85% to 15%. Full implementation was in 2010.

$$\alpha_n^T = 1 + 0.02(r - 65) \; if \; r > 65 \; and \; n \geq 35$$

APPENDIX 3. ROBUSTNESS CHECK

(A) Censoring in contributions and earnings

To check whether there is any bias resulting from the fact that the estimation includes individuals whose contributions are capped (and therefore not equal to their productivity -wages-), I perform here an additional estimation exercise where only those observations that are not upper or lower capped are considered. As can be observed in Table A.3.1, the effect of economic incentives goes in the same direction as when considering all observations. Coefficients and marginal effect from the estimations are provided in the Tables below. As can be observed, increases in the total present value of the flow of pensions that a person will receive from the year she retires to the year she dies (increases in SSW) increases the hazard while an increase in the difference of this amount from postponing retirement, SSA, decreases the hazard. Thus, the results remain similar.

Table A3.1 Logit estimates of the effect of pension incentives on retirement behavior between 60 and 65 years of age. Males born between 1936 and 1946, having worked in the Social Security General Scheme with a relation with the Social Security in 2006 (*) and whose contributed earnings are not capped

	MODEL A		MODEL B	
	ACCRUAL	PEAK VALUE	ACCRUAL	PEAK VALUE
	LR chi2(45)=31064.40	LR chi2(45)=31028.47	LR chi2(49)=31112.89	LR chi2(45)=31075.08
	Pseudo R2=0.3643	Pseudo R2=0.3639	Pseudo R2= 0.3649	Pseudo R2= 0.3645
	Log likelihood = -27099.731	Log likelihood = -27117.7	Log likelihood = -27075.487	Log likelihood = -27094.391
	92.033 number of observations			
ECONOMIC INCENTIVES AND REGULATION				
SSW_t	**2.891*** **	**2.782*** **	**3.057*** **	**2.944*** **
	(0.144)	(0.143)	(0.235)	(0.233)
SSA_t	**-9.359*** **		**-7.256*** **	
	(1.392)		(2.011)	
PV_t		**-1.908*** **		-2.073**
		(0.617)		(0.820)
RR_t	**-0.124**	**-0.042**	**-0.278**	**-0.251*** **
	(0.141)	(0.040)	(0.373)	(0.072)
SSW_t * r2002			**-0.002**	**0.009**
			(0.270)	(0.268)
SSA_t * r2002			**-4.099***	
			(2.273)	
PV_t * r2002				**-0.623**
				(0.975)
RR_t * r2002			**0.208** **	**0.282*** **
			(0.084)	(0.082)
r2002			**-0.253*** **	**-0.330*** **
			(0.095)	(0.095)
WORKING STATUS, SECTOR OF ACTIVITY AND OTHER LABOUR HISTORY RELATED VARIABLES				
l_{t-1}	**-1.586*** **	**-1.584*** **	**-1.593*** **	**-1.591*** **
	(0.044)	(0.044)	(0.044)	(0.044)
u_{t-1}	**0.921*** **	**0.933*** **	**0.920*** **	**0.929*** **
	(0.041)	(0.042)	(0.042)	(0.042)
numrel	**-0.026*** **	**-0.026*** **	**-0.025*** **	**-0.025*** **
	(0.003)	(0.003)	(0.003)	(0.003)
t x numrel	**0.005*** **	**0.005*** **	**0.005*** **	**0.005*** **
	(0.001)	**(0.001)**	**(0.001)**	**(0.001)**
meanlenght	0.006*	0.005	0.005	0.004
	(0.003)	**(0.003)**	**(0.003)**	**(0.003)**
t x meanlenght	0.002**	0.002**	0.002**	0.002**
	(0.001)	**(0.001)**	**(0.001)**	**(0.001)**
serv	0.133***	0.135***	0.142***	0.142***
	(0.046)	**(0.046)**	**(0.046)**	**(0.046)**
t x serv	**-0.062*** **	**-0.063*** **	**-0.064*** **	**-0.064*** **
	(0.013)	(0.013)	(0.013)	(0.013)

PERSONAL CHARACTERISTICS & OTHERS				
univ	**-0.749*** **	**-0.789*** **	**-0.748*** **	**-0.785*** **
	(0.074)	(0.074)	(0.074)	(0.074)
t x univ	**-0.049** **	**-0.032**	**-0.054** **	**-0.036***
	(0.021)	(0.021)	(0.021)	(0.021)
disab	**-1.080*** **	**-1.086*** **	**-1.066*** **	**-1.071*** **
	(0.233)	(0.233)	(0.234)	(0.234)
otherben	**-3.455*** **	**-3.463*** **	**-3.480*** **	**-3.487*** **
	(0.209)	(0.209)	(0.210)	(0.210)
age61	**-14.912*** **	**-14.923*** **	**-14.919*** **	**-14.943*** **
	(0.850)	(0.851)	(0.854)	(0.855)
age62	**-16.004*** **	**-15.993*** **	**-15.998*** **	**-15.993*** **
	(0.793)	(0.794)	(0.799)	(0.801)
age63	**-15.162*** **	**-15.148*** **	**-15.178*** **	**-15.179*** **
	(0.775)	(0.775)	(0.780)	(0.781)
age64	**-15.077*** **	**-15.050*** **	**-15.074*** **	**-15.062*** **
	(0.748)	(0.748)	(0.752)	(0.754)
age65	**-11.806*** **	**-11.717*** **	**-11.816*** **	**-11.751*** **
	(0.650)	(0.651)	(0.655)	(0.657)
cycle	**0.174*** **	**0.180*** **	**0.104*** **	**0.106*** **
	(0.016)	(0.016)	(0.021)	(0.021)
DURATION VARIABLES				
g1	**-15.440*** **	**-15.485*** **	**-15.462*** **	**-15.523*** **
	(0.649)	(0.651)	(0.653)	(0.656)
g2	**-1.623*** **	**-1.657*** **	**-1.623*** **	**-1.667*** **
	(0.549)	(0.548)	(0.550)	(0.548)
g3	**-0.301**	**-0.346**	**-0.298**	**-0.364**
	(0.456)	(0.455)	(0.457)	(0.457)
g4	**-1.116*** **	**-1.159*** **	**-1.089** **	**-1.146*** **
	(0.423)	(0.422)	(0.424)	(0.423)
g5	**-0.255**	**-0.292**	**-0.249**	**-0.301**
	(0.370)	(0.370)	(0.371)	(0.370)
cons	**13.345**	**13.269**	**13.747**	**13.763**
	(0.000)	(0.000)	(0.000)	(0.000)

(*) Notes:

Dependent variable is 1 if person retires and 0 otherwise, conditioned on not having retired before. Monetary values are in 1,000,000 euros, prices 2006. std errors are in parenthesis.

Statistical significance: * p<0.05; ** p<0.01; *** p<0.001. Coefficients of Autonomous Communities not resported. Prob > chi2=0.

Model A represents the basic model. Model B includes Model A and changes in 2002 regulation as control variables. Model C includes Model B and income levels as control variables.

Table A3.2. Quantitative effects of pension incentives and other variables on the average hazard rate by age

	ACCRUAL							PEAK VALUE						
	All ages	**Age 60**	**Age 61**	**Age 62**	**Age 63**	**Age 64**	**Age 65**	**All ages**	**Age 60**	**Age 61**	**Age 62**	**Age 63**	**Age 64**	**Age 65**
Economic Incentives and Regulation														
SSW_t	0.006	0.007	0.004	0.005	0.006	0.009	0.009	0.007	0.008	0.005	0.006	0.006	0.010	0.009
	0.003	*0.004*	*0.001*	*0.002*	*0.002*	*0.007*	*0.009*	*0.004*	*0.005*	*0.002*	*0.002*	*0.003*	*0.008*	*0.009*
SSA_t	-0.001	-0.002	-0.001	-0.001	-0.001	-0.001	-0.002							
	0.000	*-0.001*	*0.000*	*0.000*	*0.000*	*-0.001*	*-0.001*							
PV_t								-0.001	-0.001	0.000	0.000	0.000	0.000	0.000
								0.000	*-0.001*	*0.000*	*0.000*	*0.000*	*0.000*	*0.000*
RR_t	0.000	0.000	0.000	0.000	0.000	-0.001	-0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	0.000	*0.000*	*0.000*	*0.000*	*0.000*	*0.000*	*-0.001*	*0.000*	*0.000*	*0.000*	*0.000*	*0.000*	*0.000*	*0.000*
Individual Characteristics														
numrel	-0.001	-0.002	*-0.001*	-0.001	*0.000*	0.000	*0.000*	-0.001	-0.002	-0.001	-0.001	0.000	0.000	0.000
	0.000	-0.001	*0.000*	0.000	*0.000*	0.000	*0.000*	*0.000*	*-0.001*	*0.000*	*0.000*	*0.000*	*0.000*	*0.000*
l_{t-1}	-0.131	-0.168	*-0.076*	-0.099	*-0.110*	-0.211	*-0.143*	-0.143	-0.188	-0.086	-0.110	-0.120	-0.226	-0.143
	-0.118	-0.173	*-0.071*	-0.095	*-0.106*	-0.233	*-0.134*	*-0.130*	*-0.184*	*-0.077*	*-0.102*	*-0.113*	*-0.241*	*-0.137*
u_{t-1}	0.077	0.098	0.045	0.058	0.064	0.118	0.095	0.084	0.109	0.051	0.065	0.070	0.127	0.095
	0.057	*0.079*	*0.030*	*0.041*	*0.046*	*0.112*	*0.098*	*0.064*	*0.086*	*0.033*	*0.045*	*0.050*	*0.118*	*0.097*
uni	-0.056	-0.069	-0.032	-0.040	-0.044	-0.081	-0.102	-0.062	-0.076	-0.036	-0.045	-0.048	-0.087	-0.109
	-0.027	*-0.035*	*-0.012*	*-0.017*	*-0.020*	*-0.054*	*-0.102*	*-0.031*	*-0.040*	*-0.014*	*-0.020*	*-0.022*	*-0.059*	*-0.123*
serv	0.006	0.008	*0.004*	0.005	*0.005*	0.009	*0.009*	0.007	0.008	0.004	0.005	0.006	0.009	0.009
	0.003	0.005	*0.002*	0.002	*0.003*	0.007	*0.010*	*0.004*	*0.005*	*0.002*	*0.003*	*0.003*	*0.008*	*0.010*
disab	-0.070	-0.083	-0.037	-0.046	-0.051	-0.097	-0.159	-0.074	-0.089	-0.040	-0.050	-0.054	-0.103	-0.162
	-0.030	*-0.039*	*-0.014*	*-0.019*	*-0.021*	*-0.059*	*-0.187*	*-0.033*	*-0.042*	*-0.015*	*-0.020*	*-0.023*	*-0.062*	*-0.193*
otherben	-0.156	-0.157	-0.065	-0.084	-0.092	-0.198	-0.593	-0.163	-0.163	-0.068	-0.088	-0.095	-0.204	-0.595
	-0.060	*-0.073*	*-0.026*	*-0.036*	*-0.041*	*-0.109*	*-0.660*	*-0.061*	*-0.072*	*-0.026*	*-0.036*	*-0.041*	*-0.112*	*-0.662*

NOTES:

Model B: Quantitative effects of continuous variables are computed as the discrete differences of the logistic function evaluated at a 10% increase in the variables' values with respect to the logistic function evaluated atthe observed variables' values. Results are obtained under Model A in Table A 2. Median values are in italics. The effects are estimated for the period beginning in 2002 and for those workers whose contributions are not upper or lower capped.

(B) The potential negative implications of excluding tax considerations

Tax- induced changes in labor supply behavior has been receiving an enormous amount of effort in economic literature. Such studies have focused on the impact of taxes on hours of work and labor force participation rates. However, due to lack of appropriate microdata gathering all the relevant elements in labor supply decisions, this type of analysis is rather scarce in Spain and the available literature has only focused on partial dimensions of the decision problem, like the extensive margin of labor market participation or the determination of wages.

The effect of personal income taxes has not yet been integrated in the analysis of retirement decisions to obtain conclusions about the effect of taxes in older labor force participation in Spain.

CSWH provides information about personal income taxes and earnings, but only regarding the reference year of the sample, that is 2006. That was the reason why in the main text, I stated that "... it excludes any tax considerations".
However in light of the apparent sensitivity of retirement decisions to taxes, this issue definitely requires further inquiry. In the future, when further CSWH waves will become available, it will be possible to analyze properly the effect of income taxes on retirement decisions, taking into account individual longitudinal information. Here, with the aim of getting some insight about the size of the effect of income taxes on retirement decisions, I use the fiscal information referred to the year 2006 and compute and economic incentive variables under the assumption that, during 2002-06 period, individuals will remain classified in the same income tax group where they were classified in 2006.
To define income taxation groups I classify individuals according their real yearly contributions into five different groups according to income taxation rules in 2006, namely, no income taxation for those whose yearly income is lower than 9000 €/year, 24% income tax for those whose income is between 9000-17360 €/year, 28% income tax for those whose income is 17361-32360 €/year, 37% income tax for those whose yearly income is over 32361-52360 €/year, and 43% income tax for those whose income is higher than 52360 €/year.
Given limitations in the available information, this procedure provides a rough identification mechanism on the degree of heterogeneity in the individuals´ responses to economic incentives taking into account tax

considerations. As can be observed in Table A.3.3 , the effect of net economic incentives is slightly stronger than when no tax information is considered, but the overall conclusions drawn in the main text remain similar.

Table A.3.3. Logit estimates of the effects of pension incentives on retirement behavior between 60 and 65 years of age. Economic incentives measured net of income taxes.

	MODEL A		MODEL B		MODEL C	
	ACCRUAL	PEAK VALUE	ACCRUAL	PEAK VALUE	ACCRUAL	PEAK VALUE
Net SSW_t	**4.376*** **	**4.307*** **	**4.405*** **	**4.215*** **	**4.130*** **	**4.000*** **
	(0.183)	(0.181)	(0.285)	(0.277)	(0.286)	(0.280)
Net SSA_t	**-11.578*** **		**-19.322*** **		**-16.748*** **	
	(1.497)		(2.154)		(2.480)	
Net RR_t	**0.104** **	**0.108** **	**0.070**	**0.024**	**0.070**	**0.036**
	(0.043)	(0.042)	(0.073)	(0.075)	(0.073)	(0.075)
Net PV_t		**-5.751*** **		**-9.115*** **		**-7.298*** **
		(0.708)		(0.973)		(1.054)
r2002			**-0.581*** **	**-0.656*** **	**-0.502*** **	**-0.595*** **
			(0.077)	(0.080)	(0.078)	(0.081)
Net SSW_t* r2002			**0.438***	**0.549** **	**0.598*** **	**0.638*** **
			(0.231)	(0.226)	(0.232)	(0.226)
Net SSA_t* r2002			**10.272*** **		**5.463*** **	
			(2.015)		(2.045)	
Net RR_t* r2002			**0.106**	**0.162***	**0.023**	**0.112**
			(0.084)	(0.086)	(0.085)	(0.086)
Net PV_t* r2002				**4.636*** **		**3.406*** **
				(0.942)		(0.946)

(C) The inclusion of other benefits as controls

The variable other benefits is included to take into account personal/ family circumstances that are not captured in the other covariates included in the estimations. It takes value 1 at time t if the person is receiving any Social Security benefit different from disability, old age or unemployment while he was a year younger (that is, widowhood benefits, orphanhood benefits and other type of benefits whose entitlement come from relatives' contributions). Unlike what would happen to a sample of women, the incidence of a possible endogeneity problem from including this variable is expected to be of scarce magnitude (1.2% of people in sample receive widowhood benefits, only 0.02% receive orphanhood benefits and only 0.01% receive benefits from family contributions),, since the sample is restricted to men.