DynPenSpain01. A micro simulation model for the Spanish pensions system

F. J. Fernandez-Diaz, C. Patxot, G. Souto, M. Spielauer

Index

1. Introduction

The pension system is one of the main components of the Spanish welfare state. At present, it is composed of two different type of benefits: non-contributory (or *Beveridgean*) and contributory (or *Bismarckian*). The former can be considered a system of minimums, as it provides, under some conditions, a minimum income to those individuals who cannot access the contributory level. The latter is the main part of the system.[1] It is organized on a pay-as-you-go (PAYGO) basis and it includes different kinds of pension benefits - retirement, disability and survival - for those individuals who meet the eligibility requirements regarding age and past contributions to the system.

Clearly, the contributory retirement pension is the most important social expenditure program. In 2007 it represented over 65% of all contributory pension expenditure, and over 5% of GDP. Moreover, a huge increase in its size can be expected in the face of demographic ageing: retirement pension expenditure will be more than 75% of contributory pension expenditure and almost 12% of GDP in 2060 according to the last projection exercise of the Economic Policy Committee (2009).[2] Therefore, the concern about the necessary reforms to make it sustainable in the long term is fully justified.

Reform proposals vary from a complete restructuring of the system –like a switch to a true or to a notional capitalization system[3]– to marginal adjustments of the legal parameters in the current system. Given the expected increase in the ratio of pensioners to contributors, all of them imply raising contribution and/or reducing pensions. Alternatively, it is often proposed an increase in the retirement age, as it might improve sustainability both on the expenditure and on the revenue side. Although the legal retirement age is 65 for most workers in most countries, the fact is that average retirement age is clearly lower. Governments, concerned about the importance of this phenomenon, are interested in incentivizing workers to delay retirement. Though average retirement is quite high in Spain for European standards – according to Eurostat it was 62.6 years in 2008, 1.1 years above the EU15 level-, Spain will be hit by one of the most pronounced ageing process starting around 2020, a bit latter than in other European Countries given that the baby boom in Spain occurred latter than in other countries.

The design and evaluation of reform measures claims for the availability of analytical tools. Simulation models have been developing specially in last decades thank to the availability of an increasing amount and quality of data bases. In the context we are dealing with, those simulation tools need to take both a macro and a micro perspective. The first is crucial if one aims at analyzing in a consistent way the sustainability o the pensions or any other welfare state transfer. The second is crucial when considering the adequacy of the benefit level in what respects to income redistribution. The EU in the joint report (EPC and SPC) has been recommending to attend to both dimensions in order to keep the Welfare State, one of the main achievements of Europe in the past century.

In this paper we present a simulation model of the contributory retirement pension system in Spain. The model is, first, micro based, thanks to the availability of the MCVL –an administrative data set published by the Social Security Administration since 2004. Second, given the nature of the pension policy the model is dynamic. Finally the model is case based and continuous time and has been implemented using a programming language developed by Statistics Canada and used in many other applications.

[1] The civil servants have their own system.

[2] This projection is in line with previous projections and also with those done in academic work.

[3] The possibility to switch to a funded system is limited by the so-called *transition problem*: the initial gift given to generations who did not contribute and received a pension, now needs to be offset by a double burden on current workers who need to hold the old PAYGO system and contribute to the new one. Some countries have opted for introducing some kind of notional funding, which is somehow a sophisticated way to introduce a defined contribution system.

The paper is organized as follows. Section 2 is devoted to give and overview of the programming language employed. Section 3 is devoted to the description of the Spanish Pensions system and the data set. Section 4 describes the model structure, the data employed as an input and the behavioral equations included in the model. Section 5 presents the results and section 6 makes some final remarks.

2. The ModGen language programming

The model is implemented in Modgen, a generic micro simulation programming language developed and maintained at Statistics Canada and widely used in social science micro simulation. It is freely available at the Statistics Canada web site together with extensive documentation including simple application examples. Modgen supports the creation, maintenance and documentation of most dynamic microsimulation model types, including both continuous and discrete time models as well as interacting and non-interacting populations. Modgen was created with the goal of automating as many aspects of the model implementation as possible. This includes highly efficient event queuing, a graphical user interface, the graphical display of individual life histories and powerful tabulation facilities. As a result, Modgen empowers researchers with only moderate programming skills – comparable to those necessary for statistical analysis using statistical software packages - to implement models themselves.

The model is developed in continuous time. Its starting population is read in from a data file based on administrative data after some depuration. Besides extensive table output, it also produces cross-sectional micro-data output containing expected pensions by year of retirement. This allows calculating incentives to postpone retirement. Together with observed retirement events, this information is used to estimate behavioral equations for retirement decisions.

Like all Modgen applications, this model is highly modular and features an attractive user interface allowing easy scenario management and navigation through parameter tables and model results. In the following we outline the basic elements in ModGen programming. A selective summary of the process of producing a new model is described, focusing on the relevant features of ModGen in relation to the application derived in this study.[4] As said above, the language is designed in order to be as flexible as possible and to minimize the programming task of the model developer or programmer and users. The programmer needs to write the code in two or more "mpp" files. When the model execution is called those files are combined with the corresponding internal files generated by compilation -"cpp" files. Other internal files contain the code common to any simulation model.

Once the program is written the compilation process produces an autonomous and user friendly application. Once created, this can be used independently to produce simulations by changing the input parameters. Both parameters and output tables are organized in tables and parameters can be changed in the application easily.

In the following, the content of the mpp files that need to be produced is outlined. A detailed explanation together with the way to define each element in the code can be found in the model developer guide.

First the project needs to be opened from VS2008 and this generates an automatically created "mpp" file, named as the model, containing the simulation engine of the new model. Second the PersonCore.mpp file is created defining the actor and its elements. The agent is usually a person, but one or more kind of actors can be defined.[5] As said above, some common features in all micro simulation models are internally modeled. In particular the evolution of time and the age of the actor are internally taken care of. This, together with an

---

[4] See http://www.statcan.gc.ca/microsimulation/modgen/modgen-eng.htm for details.

[5] See the example ModGen web site in which houses are also actors that age at the same time as their occupiers.

efficient way of defining *ranges* and *classifications* (grouped under the category "*types*"), eases the programming task. This modular structure easies any posterior modifications or extensions in the model.

Required	
ModelName.mpp	Model type (case or time based) Time (continuous of discrete) Language
PersonCore.mpp	Define ACTOR and elements - Simple and derived STATES - EVENTS - Data Members - Function Members
OutputTables.mpp	Defines output tables that will appear in the application
Optional	
Module1.mpp (one or more)	Groups together actor elements and any computations (classification, ranges, partitions) refereeing a particular part of the model, a behavioral or non behavioral
Module2.mpp (one of more)	Specific computations taken aside to clarify the program, like: - Input uploading or output writing - Otters

The last required mpp file is the OutputTables.mpp file where the output tables needed are stored. This programming of this module is important as only the information in this tables is stored along the simulation. The rest of information is deleted to save space.

All the code could be written in one file –basically PersonCore.mpp- although it might be advisable to build different behavioral modules following the main structure of the model. If this is the case, the PersonCore.mpp file contains only the general elements of the actor, while the specific elements are left to specific modules (Module1.mpp). The types (parameters and ranges) can be distributed in any way along the modules.

Similarly any kind of auxiliary module (Module2.mpp) can be created to put aside specific computations, like input uploading, output writing or complex computations.

3. The Spanish pension system

The Spanish Social Security system as such was introduced in 1967. Previously, workers were insured against disability and retirement by mutual societies[6]. A substantial share of those individuals who belonged to that previous system are still working and, because of that, they still conserve some special privileges in acceding to early retirement within the contributory system. Due to this particularity, we will refer to them as Old System (OS) workers henceforth.

The Social Security is organized on a PAYGO basis and under a defined-benefit scheme. To receive a pension, which will be determined in part by the past contributions of the worker, a minimum period of contribution is required. The so-called Bismarckian orientation of the system relies on these elements. Social Security covers other contingencies besides retirement, such as permanent disability and survival, but retirement is clearly the most important of them. Below we outline some legal features relevant to our analysis.

[6] This system was called *Seguro Obligatorio de Vejez e Invalidez* (SOVI).

a) Determining pension rights

Beneficiaries of retirement pensions are affiliated workers who meet the legally established conditions of i) having the legal age (generally 65) and ii) proving a minimum contribution period of 15 years - 2 of them should be included in the 15 years prior to retirement date.

The initial retirement benefit is determined according to a formula depending on retirement age, the number of contribution years and the amount of past contributions. Since the reform posed by the 24/97 Act the computation is as follows. First, a basic amount (*RB*) is computed as a mean of the level of contributions made in the 15 years prior to retirement[7]. Second, a percentage depending on the number of years of contributions – *p(n)* – is applied to the *RB* in this way: 50% for the first 15 years, an increase of 3% for each additional year between the 16th and the 25th, and an increase of 2% for each additional year from the 26th until it reaches 100%, at 35 years of contributions. Furthermore, when retirement age deviates from the legal retirement age – 65 - a correction coefficient – *cc(n)* – is applied depending also on the number of years of contribution. Thus, in general terms, the legal formula to calculate the initial retirement benefit (*IRB*) or entry pensions is:

$$IRB = p(n) \cdot RB \cdot cc(n) \qquad (1)$$

The coefficient *cc(n)* will be above (below) one for workers delaying (advancing) their retirement from age 65. In the first case an additional 2% will be applied for each year of delay, or 3% if the worker has 40 or more years of contributions – this premium either acts on *p(n)* increasing up to 100%, holding *cc(n)* equal to one, or increasing *cc(n)* above one if the worker already reached the 35 contribution years. In the second case, if the worker retires before age 65 an annual penalty between 6% and 8% is applied, depending on the years of contribution. [8] Apart from this general rule, there are also other possibilities to reduce the age of retirement: i) certain professional activities, ii) disabled workers, iii) special retirement at age 64 with no penalty and iv) partial and flexible retirement. In each of them different correction coefficients are also applied, as shown in the next section.

b) The retirement decision: Handling multiple retirement paths

There are several ways to access retirement in Spain, some of them depending on the labor status. Table 3.1 shows all those different possibilities or retirement paths. The ordinary retirement age – 65 – is not compulsorily established by law, but it is agreed in most collective wage settlements. In any case, at the moment 65 is the reference age for retirement rules and incentives.

Early retirement pathways have been developing gradually with this direct purpose, but also with an indirect objective of employment policies, like intending to improve the labor market access for youth. Regarding the latter, special retirement at age 64 and partial retirement were established in 1985 and 1999, respectively.[9] Regarding the former there are several rules. First, apart from the disabled, for some professional activities – especially those which are dangerous or which imply being away from home – there is a possibility of retiring at a fixed or at a reduced age. Second, early retirement was generally available from age 60 under certain requirements but with some penalty in the old system. Given that this would eventually disappear, in 2002 (35/2002 Act) new rules were established for new system workers, allowing early retirement from age 61 with similar but stronger penalties, and including an explicit requirement of being unemployed.

---

[7] The legal name is *base reguladora*. In the computation, the contributions are updated from the third year according to the evolution of the consumer price index.

[8] In particular, the 8% affects those that only reach the minimum eligibility requirement of 15 contribution years. This penalty is reduced gradually for those crediting enough years to move to the next contribution years scale: 31-34 (7,5), 35-37(7), 38-39 (6,5) and 40 plus (6).

[9] The former was introduced in 1985 (RD 1994/1985) and the latter in 1999 (RD 144/1999).

In parallel, and also due to the need to foster sustainability in the face of demographic ageing, some rules were introduced in the same act intending to delay retirement. On the one hand, the possibility of combining work and pension receipts was introduced by reformulating partial retirement and introducing flexible retirement (the latter allowing retirees to come back to the labor market as part-time workers). On the other hand, incentives to continue working full-time after age 65 were introduced in the regular retirement path, as detailed above. Finally, there is a possibility to exit the labor market through pre-retirement, i.e. through a special agreement with the firm that leads to private or public subsidies until the age of eligibility.

Table 3.1 Retirement paths in Spain according to labour status

Labour Status	**Retirement path**	**Eligibility requirements and rules determining benefits (2007)**
Disabled	Disability*	At age 65 disability pensions are converted into retirement pensions, but keeping the same benefit level
Unemployed	Back to work (all)	
	Early retirement from age 60 (Old system)	Minimum *n* = 30 8% penalty per year until age 65 (gradually reduced to 6% if *n≥40*)**
	Early retirement from age 61 (New system)	Minimum *n* = 30 7.5% penalty per year until age 65 (gradually reduced to 6% if *n≥40*)**
	Regular retirement at 65	(See conditions bellow)
Worker	Special retirement at age 64	No early retirement penalty Substitution contract in the same firm
	Early retirement from age 60 (Old system)	8% penalty per year until age 65
	Regular retirement from age 65 (includes delayed retirement)	<ins><65:</ins> Reduced age for special professional activities with no penalty <ins>Age 65:</ins> Minimum *n*=15 (2 in the last 15) <ins>>65</ins>: Increases beyond 100% of *RB* by 2% per year (3% if *n*≥40)
	Partial retirement**	From age 60 Minimum *n* =15 years Part-time work and proportional reduction of pension If age < 65 substituting contract No early retirement penalty
Retired	Flexible retirement	Part-time work and proportional reduction of pension

Notes: *n*= number of years of contribution; * Only disabled below age 65 might change state back to work; ** 7.5% (*n*=30-34), 7% (*n*=35-37), 6.5% (*n*=38,39), 6% *n≥40;* *** In 40/07 Act the minimum *n* was increased to 30 and 6 years of seniority in the same firm were required.

Table 3.2 Distribution of new entries by pathways (Spain 2002-2007)

Year / Retirement Pathway	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**
From disability	**6.86%**	**5.84%**	**4.48%**	**2.33%**	**1.92%**	**0.60%**
Early retirement	**29.42%**	**33.79%**	**33.86%**	**24.02%**	**28.28%**	**26.77%**
Old system: from age 60 on	25.44%	29.50%	27.95%	19.62%	22.57%	20.24%
From unemployment	12.58%	14.10%	14.18%	10.18%	11.21%	10.76%
From employment	12.86%	15.40%	13.77%	9.44%	11.36%	9.48%
New system: from age 61 and unemployment	0.51%	0.92%	1.60%	1.47%	1.88%	2.32%
Special retirement at age 64	2.35%	2.19%	3.40%	2.13%	2.57%	3.14%
Collective wage settlements	0.00%	0.05%	0.06%	0.10%	0.23%	0.27%
Pre-retirement	1.12%	1.13%	0.86%	0.70%	1.02%	0.81%
Partial retirement (from employment)	**3.45%**	**5.30%**	**8.10%**	**7.78%**	**11.80%**	**12.82%**
Flexible retirement (from retirement)	**0.24%**	**0.52%**	**0.30%**	**0.31%**	**0.31%**	**0.20%**
Ordinary retirement pensions (Including delayed)	**60.04%**	**54.55%**	**53.26%**	**65.56%**	**57.69%**	**59.60%**
<60	1.14%	1.11%	0.98%	0.81%	1.11%	1.25%
60	0.97%	0.49%	0.49%	0.40%	0.24%	0.20%
61-64	1.83%	1.07%	1.35%	1.09%	1.22%	1.05%
65	44.83%	39.84%	38.25%	43.36%	41.93%	45.43%
>65	11.26%	12.05%	12.19%	19.90%	13.19%	11.66%
Missing age	0.00%	0.00%	0.00%	0.00%	0.00%	0.01%
Total	**100%**	**100%**	**100%**	**100%**	**100%**	**100%**

Note: Pre-retirement includes only those which can be identified as receiving public subsidies.
Source: own elaboration from the MCVL.

Table 3.2 shows the distribution of new entries to retirement in Spain along six subsequent years. Clearly, ordinary retirement is the most common and includes both a small share – around 1% – of early retirement at fixed age and a sizeable share – above 10% – of delayed retirement.[10]

Table 3.2 shows the distribution of new entries to retirement in Spain along six subsequent years. Clearly, ordinary retirement is the most common and includes both a small share – around 1% – of early retirement at fixed age and a sizeable share – above 10% – of delayed retirement.[11]

Interestingly, while the share of pensioners opting for the *OS* early retirement rules is still 20.24% per cent, early retirement at age 61 has not been very attractive. To some extent, this fact is explained by the requirement of being unemployed, but one can think that workers have opted for other less costly ways to access early retirement, like partial retirement. In some cases they even use the flexible retirement situation to improve their benefits once retired (in 2006 only 12.38% of pensioners in this situation were below age 65).[12] Indeed, in practice, some of the measures intending to delay retirement have still been used as a way to improve the pension benefit by early retirees. 98.75% of new entries and 87.76% of pensioners in partial retirement situation in 2006 were below age 65. Flexible retirement has indeed been used to stay longer in the labour market, but its impact is very limited as only 0.20% of new entries in 2007 have chosen this option.

[10] The table classifies new pensions using the variables "type of pension" and "situation of the pension" in the MCVL. In the next section a description of the database is provided.

[11] The table classifies new pensions using the variables "type of pension" and "situation of the pension" in the MCVL. In the next section a description of the database is provided.

[12] With respect to partial retirement that could start before retirement but could in principle be extended beyond 65, only 6.56% of pensioners in 2006 were older than 65.

Despite the abovementioned reform efforts, the average retirement age for males has been quite stable in the last years. Interestingly, at present females retire later, perhaps due to a joint retirement decision or to the need to complete their shorter contribution histories. This would also explain some of the differences observed between male and female retirement probabilities. Despite the stability of the average retirement age, the evolution of the share in new entries by age and sex has undergone substantial changes which seem to be driven by cyclical movements. In fact changes in early and delayed retirement move in different directions. The share of those retiring after age 65 is only slightly affected by incentives to delay retirement introduced in the 35/2002 Act. One should bear in mind that most collective wage settlements deny workers the possibility of delaying retirement, so that incentives to work beyond age 65 might become inoperative.

4. The model structure

a) The basic structure

Technically the model is structured as explained in Section 3. The required modules –general and PersonCore- include the definition of agent and the main elements, while some additional modules contain the computations related to labor market transitions, earnings development, pension rights and output tables.

Required	
ModelName.mpp	Model type (case or time based) Time (continuous or discrete) Language
PersonCore.mpp	Define ACTOR and elements - Simple and derived STATES - EVENTS - Data Members - Function Members
OutputTables.mpp	Defines output tables that will appear in the application
Optional	
Earnings.mpp	Evolution of wages
Work.mpp	Labor market transitions
Pensions.mpp	Determining pension rights Retirement event
Other	Reading the input file (2007 situation) and writing output file is necessary
Mortality	

b) Data employed

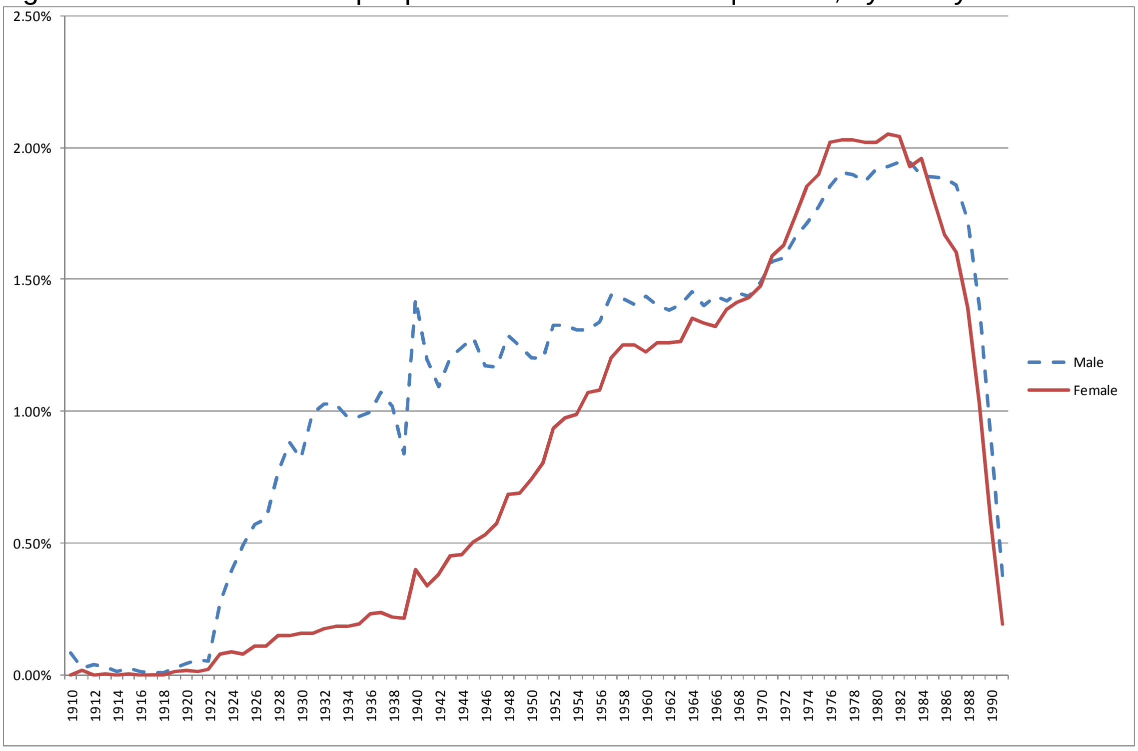
As said above the model starts from a subsample of individuals registered in Social Security in 2007 extracted for the 2007 wave of the MCVL. At this firs stage we stick to information given in the MCVL and no alignments to external data are made, unless it is strictly necessary. In the following, the main decisions taken regarding the MCVL data set and information used from external sources are summarized. Given that our main focus is retirement pensions the rest of events are modeled in the simplest way, given data availability. Agents experience first birth and second entry in the labor market. Third, other labor market transitions including changes in the qualification level and unemployment events- are derived using a behavioral (*B*) equation. Fourth, once agents attain the eligible retirement age, they start computing their expected pensions in each of the available pathways depending on their labor market status and make their retirement decisions –a behavioral equation is also implemented at this stage. Finally agents die according to exogenous age and gender specific mortally rates evolving in line with the one used in the standard population projections. Not all events are experienced by agents alive and working in 2007. In fact the model starts from individuals present in the 2007 wave of the MCVL as shown in Table 4.1.

Table 4.1. Initial simple from the 2007 MCVL

	The youngest in the labour market in 2007	Start retiring (age 59)	Finish retiring (age 75)	Maximum age 100
Birth year	1991	1948	1932	1907
Age in 2007	**16**	**59**	**75**	**100**
Start retiring (age 59)	2050	2007	1991	1966
Finish retiring (age 75)	2066	2023	2007	1982

Hence, new entries in the labor market from 2008 on and new births from 1991 need to be added in the model. In order to do so, we compare the number of people in the 2007 population and in the 2007 MCVL wave. Figure 4.1 below shows this relation for males and females.

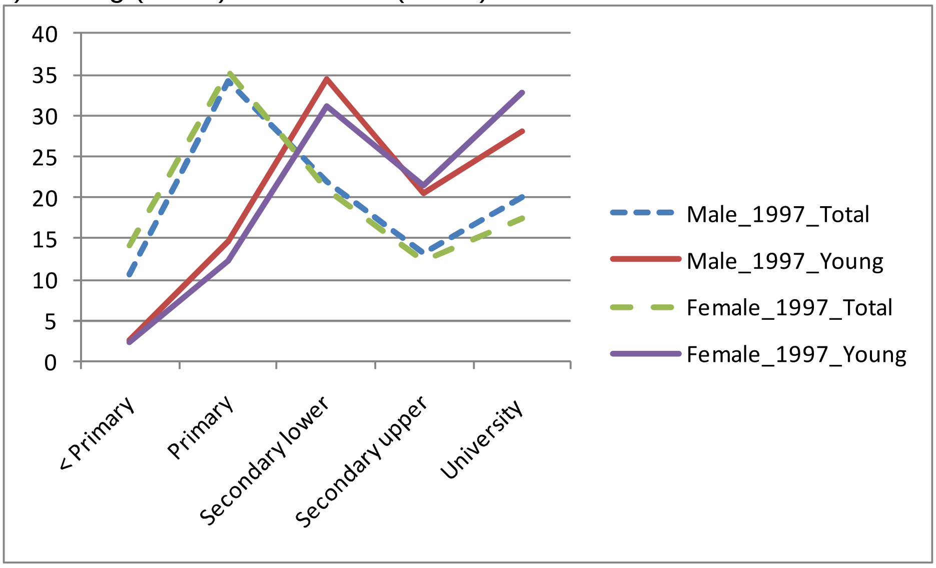
Figure 4.1 Ratio number of people in the MCVL/Total Population, by birth year.



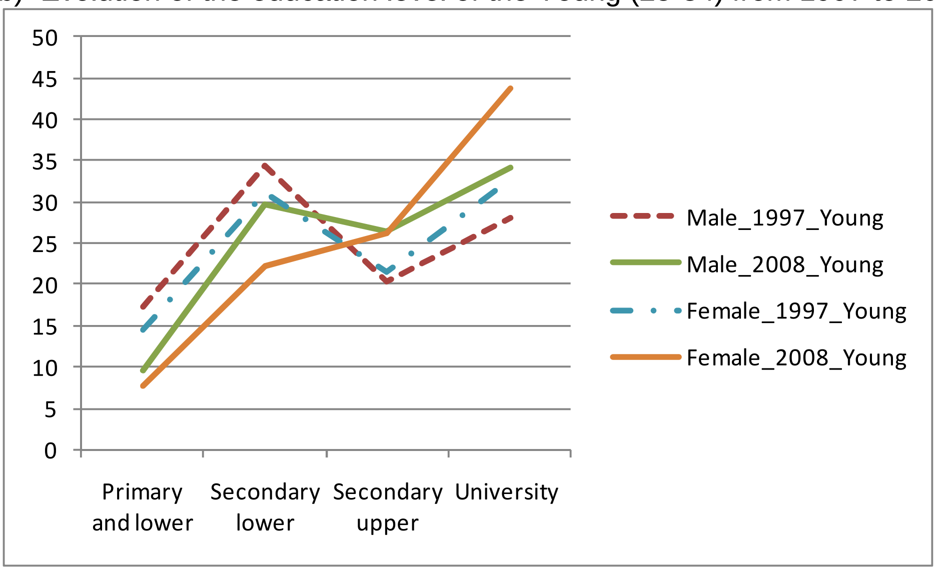
The distance between males and females for cohorts born from 1922 to 1960 is explained by low participation rates of females in those cohorts. Quite surprisingly, females overcome males for some particular cohorts. To avoid this temporary bias, we fix the weight given to males and female as the average of cohorts born between 1967 and 1971. Then, using standard population projections we add the resulting number of future newborns and entries in the labor market by gender to the 2007 sample.

Next step is assigning an educational level. Decisions on education level are minimized and simplified as much as possible. The MCVL contains the education level of the individuals. Nevertheless, this variable is collected from a different data set that it is not updated very frequently. As a result the education level is frequently missing or underestimated. For individuals registered in the MCVL we keep the value reported there or assign it as discussed in the following. For "future" individuals, born from 1991 on, the final education level is assigned randomly so as to reproduce the educational distribution observed for the Spanish population at MEC (2010). According to this publication the education level has grown substantially. In 1997 –Figure 4.2, panel a)- one can already observe that the young population was more educated that the total population. The improvement in the education level continues until 2008, as one can see by looking at Figure 4.2 panel b): the education level increases both for young for males and females from 1997 to 2008. Interestingly the education level is currently higher for young females that for males, inverting the initial situation. Note that we are assigning the final educational attainment without any attention to transitions from one education level to the next one, given the data constrains and the focus of our analysis.

Figure 4.2 Education level in Spain
a) Young (25-34) versus total (25-64) in 1997



b) Evolution of the education level of the Young (25-34) from 1997 to 2008



Source: Authors elaboration from MEC (2010).

Once the main characteristics of the individuals are assigned and they reach the age of 16 they are exposed to the risk of entering the labor market by age, gender, education and initial qualification level. This is obtained from the observation of the entry path of the last cohort, which had completed their incorporation in the labor market –those aged 36-40 in 2007 (see Table 4.2.). As shown below, most entrances are in the 5th contribution group (composed of part time workers) except for the most educated – more than ¼ of them enter once they obtain their degree. The patter is similar for females with slight differences.

Table 4.2. Entrance in the labour market by education and qualification level (CG)

	Initial qualification level (CG)					
	1	2	3	4	5	
Education level	Males					**Total**
1	0.83%	0.37%	0.72%	10.10%	87.98%	**100%**
2	6.95%	4.21%	3.77%	31.08%	53.98%	**100%**
3	25.17%	11.25%	9.98%	33.41%	20.19%	**100%**
	Females					
1	1.28%	0.94%	0.65%	26.93%	70.19%	**100%**
2	7.32%	7.54%	1.98%	53.14%	30.02%	**100%**
3	20.89%	15.53%	4.34%	43.72%	15.53%	**100%**

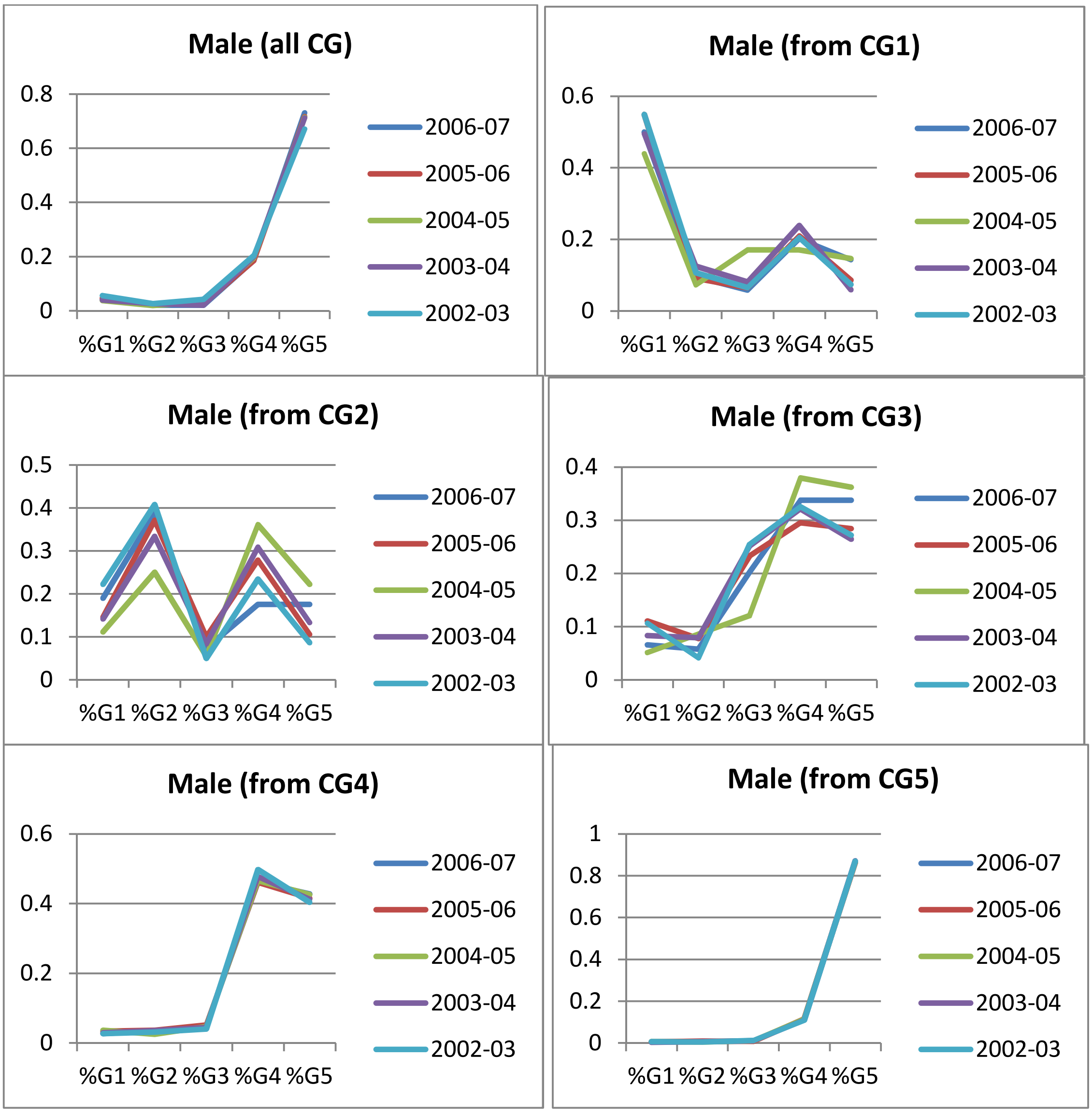
Once individuals enter in the labor market, they are exposed to the labor market transition –including unemployment- detailed in the next section. Finally, an exogenous age, gender and time specific mortality rate, coherent with demographic projection is applied.

Finally, regarding labor market transitions observed hazards are extracted from the MCVL. In particular, transitions between qualification levels within employment or between employment and unemployment are obtained by age and gender. To that effect, the 13 contribution groups in the general regime are grouped in five subgroups –those subject to the same contribution limits (thresholds). Figures 4.3 to 4.5 show the evolution of the transition hazards observed from 2002 to 2007.

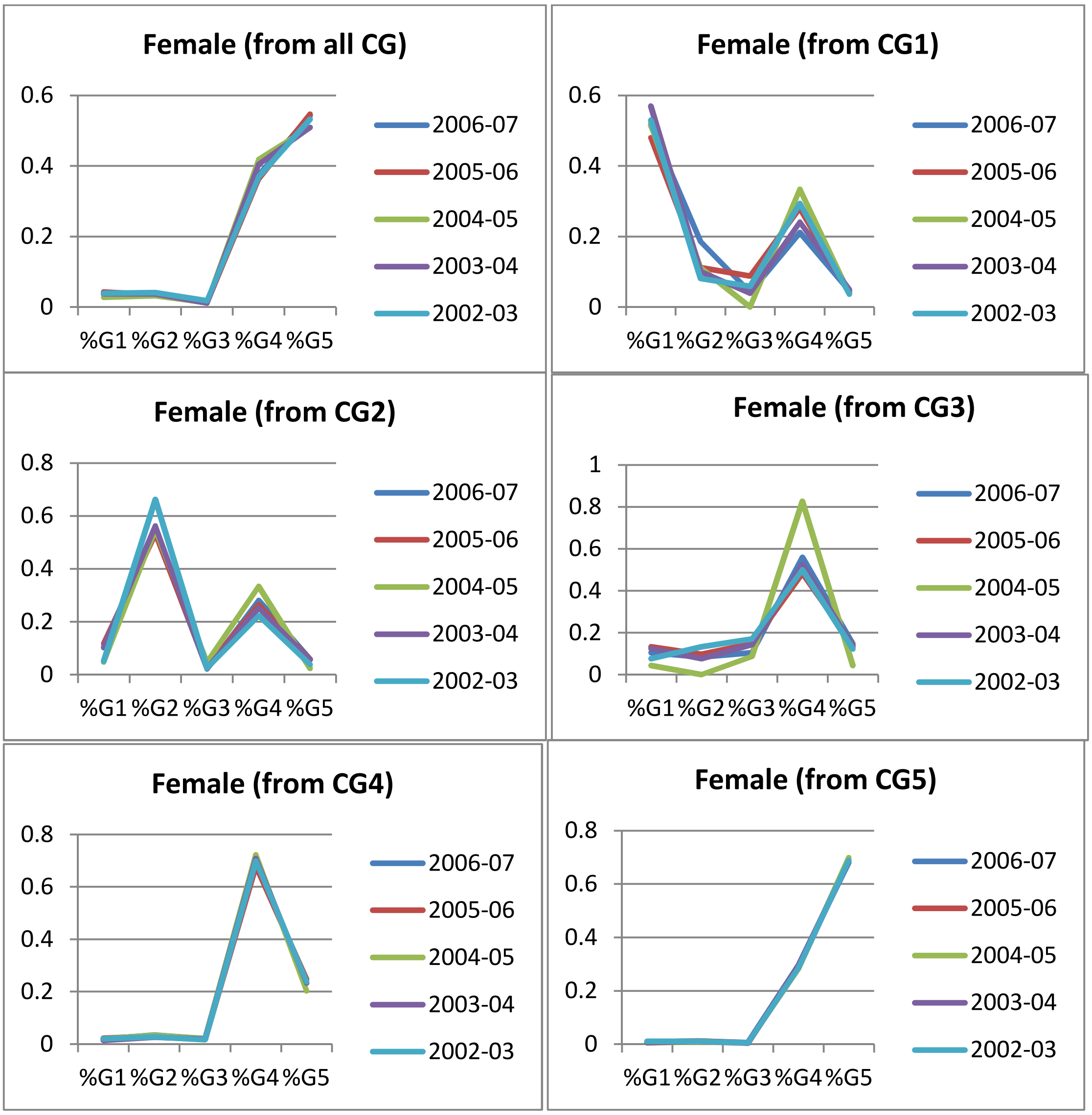
Figure 4.3 shows the changes in qualification level for workers reentering into employment from unemployment, both for male (panel a) and female (panel b). In all cases the transition hazards remain quite stable during the covered period. Only slight changes –more pronounced in the 2004-05- can be appreciated for male entering from groups 2 and 3 and females entering from group 3 to group 4.

Figure 4.3. Changes in qualification level from unemployment to employment

Panel a) Male

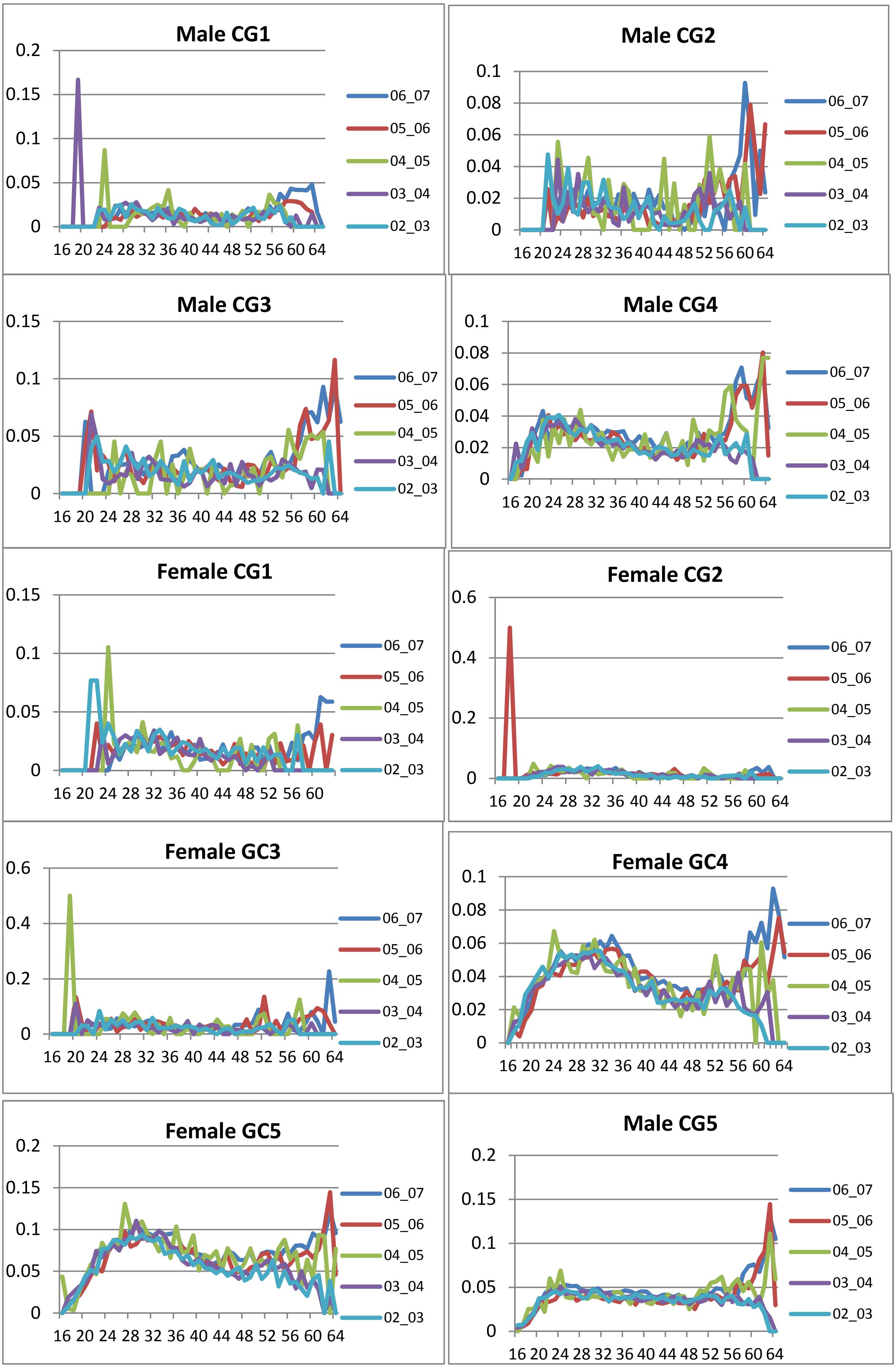


Panel b) Female



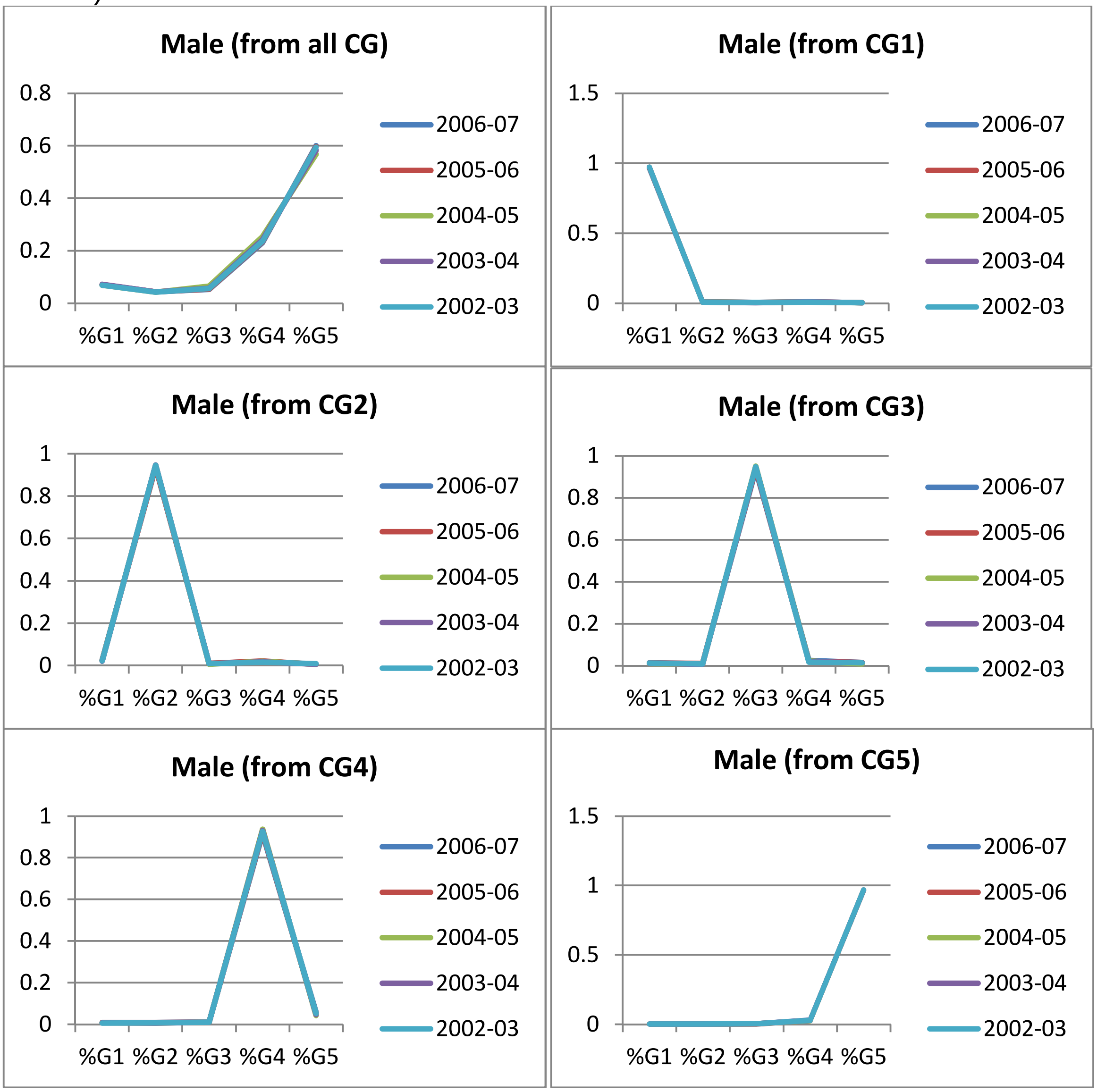
In Figure 4.4 the transition hazards from employment to unemployment are shown. In this case the stronger differences can be appreciated among the different years as the Figures show the detail by age. Nevertheless, the trend is also quite stable both for male and female.

Figure 4.4 Changes in qualification level from employment to unemployment

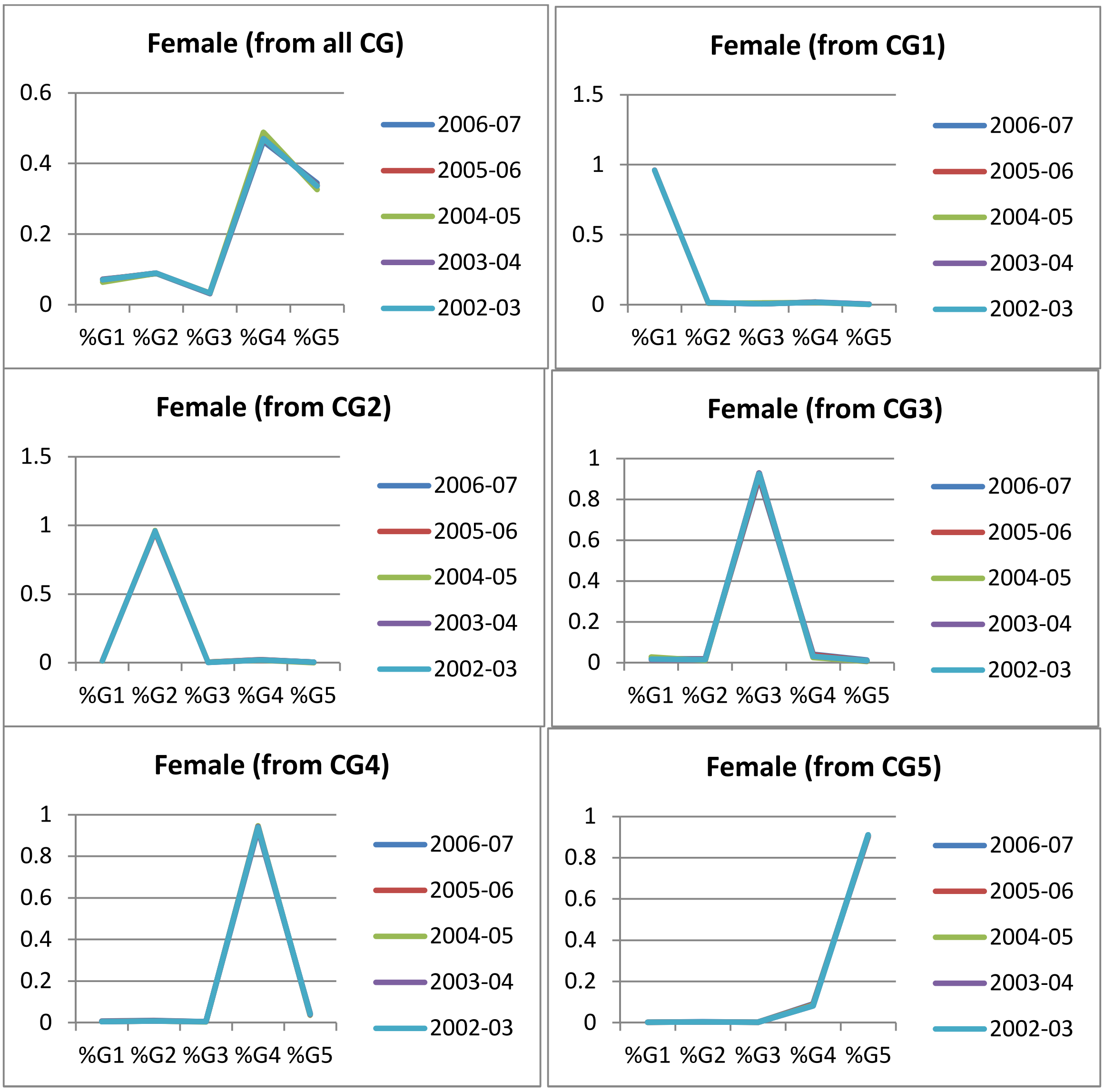


Finally, in Figure 4.5 hazards rate for transitions between different qualification levels (contribution groups) within employment can be observed. Again, no significant changes are appreciated for the different years analyzed.

Figure 4.5 Changes in qualification level within employment
Panel a) Male



Panel b) Female



As the transition hazards among the different qualification levels are quite stable during the observed period, the value of the last observed transition before the economic crisis (2006 to 2007) is taken and it is hold constant for the future. In Section 5.a), an explanation of how those hazards are altered during the crisis.

c) Behavioral analysis

In this section, we describe the behavioral equations introduced in the model so far. First, the retirement decision is modeled as it is the main focus of the analysis and the one is better reflected in the data base. Second, the way wage growth is modeled.

<u>The retirement decision</u>

Two main complications arise in the estimation of the retirement transition. First, the possibility of retiring through several different pathways, depending on the labor status, needs to be considered. Second the need to consider the effect of pension rights on the retirement decision implies more complexities in the kind of explanatory variables included. In the following we outline the solutions

Regarding the first, Table 3.1 details the different pathways to retirement, depending on the labor status. Ideally, a comprehensive consideration of such this complex retirement decision should allow for transitions between the four different states to all the retirement pathways allowed from them. The best approach would be estimating a multinomial model for all the pathways simultaneously.[13]

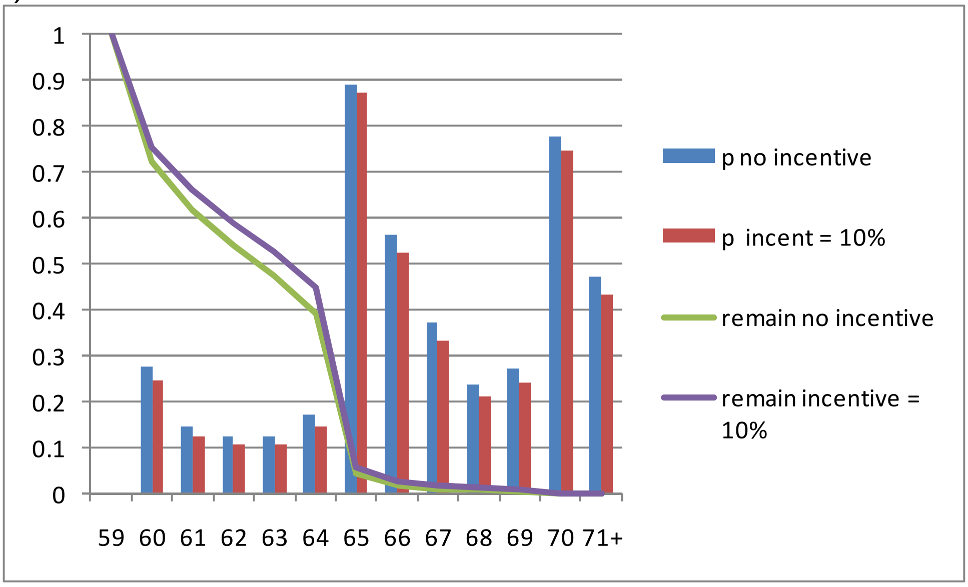
Second, as said above, in order to be able to behavioral response to reforms, it is necessary to add as explanatory variable a measure of the effect of pension rights on the retirement decision. Pension rights in turn, depend, first, on the whole labor history of the individual and, second, on the legal incentives to retire. The latter might depend on future pension rights and, hence, some prospects about the future need to be considered. In fact, structural models of retirement behavior take into account utility maximization along the life cycle and, hence the whole set of contribution and pension receipts from the pension system. In applied analysis using reduced form estimation of retirement decision this is usually summarized in the so-called Social Security Wealth (*SSW*). The latter takes the present value at the decision time of future contributions and pension rights. A given discount factor and a survival probability are used to give a weight to future payments/receipts. This way, reduced form estimations of the retirement decision include incentives to delay retirement based on an approximate measure of the *SSW* as explanatory variables. We opt for a more simplified approach given the difficulties in identifying the separate effects of retirement incentives and other explanatory variables, especially age. Gruber and Wise (2004), collect the results of a comparative study of reduced form estimation of retirement probability in this line. By considering the alternative retirement paths as perfect substitutes, a binary response model can be estimated and a weighted incentive measure of all the relevant pathways is used as an explanatory variable.[14] Although incentive measures turn out significant for most countries, results for Spain are an example of the so-called identification problem (See appendix B for details).

As said above we opt for a more simplified approach. When individual reaches the minimum initial retirement age, 59, a set of future expected pensions for each future year is obtained, conditional on the labor status. The expected pension for the following years is also obtained in order to capture incentives to delay retirement. Once the available retirement paths are filled up, the best pension is chosen. This way a vector of best pensions (and retirement paths) is obtained for each individual and used to estimated retirement probability as a function of age dummies and retirement incentives. The results shown below are obtained using age dummies, a constant and the increase in next year pensions as explanatory variables
Figure 4.6 shows the results of this estimation. Results turn out more significant for males that for females. This is not surprising given that the female cohort retiring in the observed period was not fully incorporated in the labor market. In any case the effect of an increase in expected future pension is very small. Hence, the results can be interpreted as the reaction of a myopic individual who mainly cares about the value of current pension, which in turn affects to the retirement path decision. These results will serve as a benchmark for future investigation.

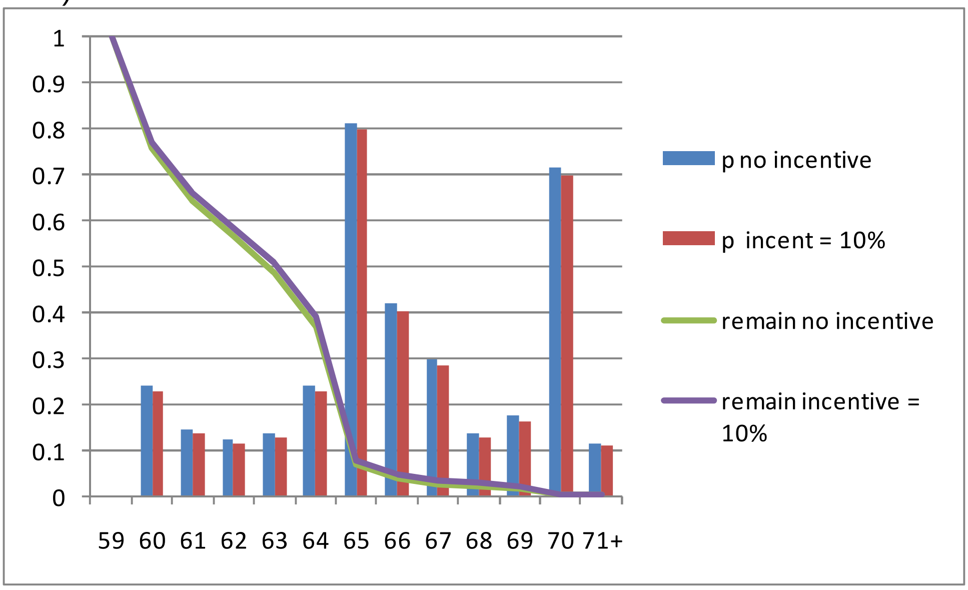
[13] See Maes, 2008a for a survey comparing multinomial logit to binary response models.
[14] See Maes 2008a for a survey.

Figure 4.6 Estimation results for retirement probability (and survival not retired) without vs. with incentive (set to 10%)

a) Males



a) Females



There is some previous work applied to Spain regarding retirement probabilities. First, Boldrin and Jiménez-Martín (2004) and Jiménez-Martín (2006), estimate retirement probabilities using an earlier preliminary version of the MCVL, not generally available. They mainly focused on employed individuals, while the alternative pathways to retirement were long-term unemployment benefits and disability included using as exogenous age and gender specific probabilities.[15] Second, more recently and using the MCVL, in the context of the employment search behavior García-Pérez et al. (2009) perform a reduced form estimation of the joint determination of the exit rate from employment and unemployment, using a duration model. They argue that the combination of generous unemployment benefits and strong early retirement penalties reduces the search effort of workers near retirement age. As a result, access to retirement through unemployment turns out to be the chosen path for a substantial share of low-skilled males. Besides the

[15] Argimón et al. (2007 an 2009) also estimated retirement probabilities. In the former retirement incentives are not considered among the explanatory variables. In the latter retirement incentives are included, but there is no clear way of controlling for the alternative pathways to retirement.

estimation method, we deviate from them, mainly by focusing on the choice between ordinary or delayed retirement.

Modeling earnings

Some decisions need to be taking in modeling wages. On the one hand, the initial value of wages for future workers needs to be decided. On the other hand, it is necessary to decide how wages are to be projected to the future both for 2007 workers and for those entering the labor market afterwards.

The projection of earnings is modeled taking into account they evolve with time, age and sex. The literature on this field highlights the fact that it is not possible to identify the separate effect of age, cohort (or birth year) and year, linked by a linear combination. Hence, a simplifying approach is taken. It is assumed that the growth rate of wages evolves as,

$$1 + g(e, s, gr, t) = [1 + g(t)] * [1 + dg(e, s, gr)] * [1 + u_t] \qquad (2)$$

being *g* the growth rate of wages, which depends on age (*e*), sex (*s*) and qualification group (*gr*) and *u* and error term. Fitzenberger et al., (2001) uses this specification to test the unified growth rate of wage hypothesis. Moral-Arce et al. (2009) apply this method to the Spanish case using the MCVL. The specification is not rejected which implies that an age profile exist though parameter do not allow to be identified. Hence we opt for obtaining $dg(e, s, gr)$ along a period and take the average. The latter implies that cyclical effects might be compensated. Although the cohort effect cannot be eliminated in this way -the data base is mainly observing the entry of the baby boomers in the labor market- it is the best we can do. To minimize errors, only full time wages are used to that purpose. For the moment only the deterministic terms of the equation are used to project -no error term is added.

For those working or contributing in 2007, the initial value of wages in 2007 is taken from the fiscal module of the MCVL, for those who wage could not be obtained, the contribution basis is taken otherwise. For future contributors the initial wage is assigned according to its education level, which, in turn, determines the contribution group, entry age and wage, as shown above. Similarly the error term observed in each cell is also used in order to ensure variability of initial wage.

Along the simulation, earnings (and contribution bases) are updated on a continuous time bases. To that purpose, both a current value and an accumulated value are maintained and updated in the following cases. First, earnings are updated at the beginning of the year, according to equation (2) –using and age, gender and group specific productivity growth rate- and adding the expected inflation rate. At the same moment contribution bases are also updated. Second, whenever a labor status transition occurs –both between CG within employment status or between unemployment and employment status- a change in wage is applied depending on gender and the original and final state. To that purpose we use the average change in wage observe.

Finally, each time one of the abovementioned changes occurs, total earnings (and contribution bases) functions are updated. This also happens at the end of the year, so that the annual flow of earnings and bc can be recovered and stored.

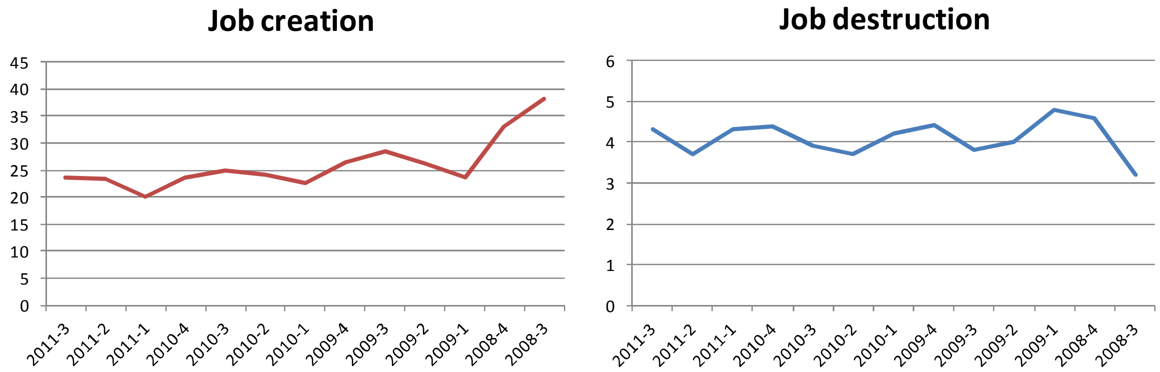
5. Results: An illustration of the dilemmas in dynamic micro simulation

In this Section the preliminary results of the simulation are shown. First the baseline situation is characterized by simulating the effect of the crisis. Second the impact of most of the reforms introduced in 2011 is analyzed.

a) The baseline situation and the effects of the crisis

The year 2007 is chosen as the baseline to avoid that projections are permanently affected by the current crisis. At the same time it is necessary to take into account the effect of the crisis, though it is still early for a thorough account of it. Hence we opt for a simplifying *ad hoc* simulation of a temporary increase (decrease) in the job destruction (creation) rate, in line to the evolution observed in the first years of the crisis, shown in Figure 5.1. Given the uncertainty on the duration of the current crisis the observed change in job destruction and job creation rates is applied to hazard rates observed in the MCVL (shown in Figure 4.4) from 2008 to 2014. As a result, the increase in unemployment rate shown in Figure 5.1 (panel a), is simulated. This increase in unemployment probably understated the impact of current crisis and the duration of its effects, but is enough to show the potentialities of the model. Panel b) in Figure 5.2. shows the effect of the current crisis on entry pensions (both number and level). The number of new entries increases at the beginning due to the increase in the unemployment rate, despite the reduction in the level of entry pensions.

Figure 5.1 Evolution of the job creation and destruction rates (2008-2011)

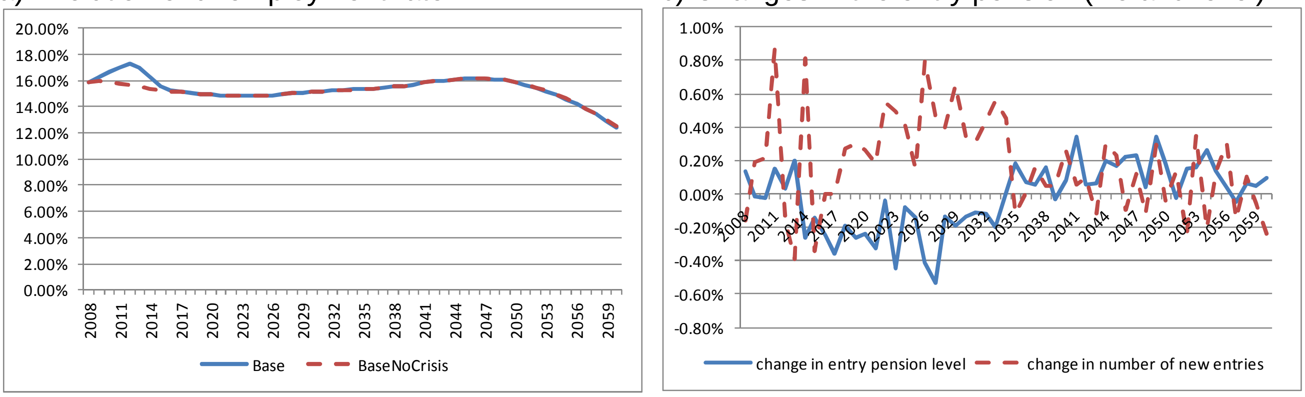


Source: Observatorio laboral de la crisis, (www.fedea.es/observatorio)

Figure 5.2 Effect of the current crisis

a) Evolution of unemployment rate

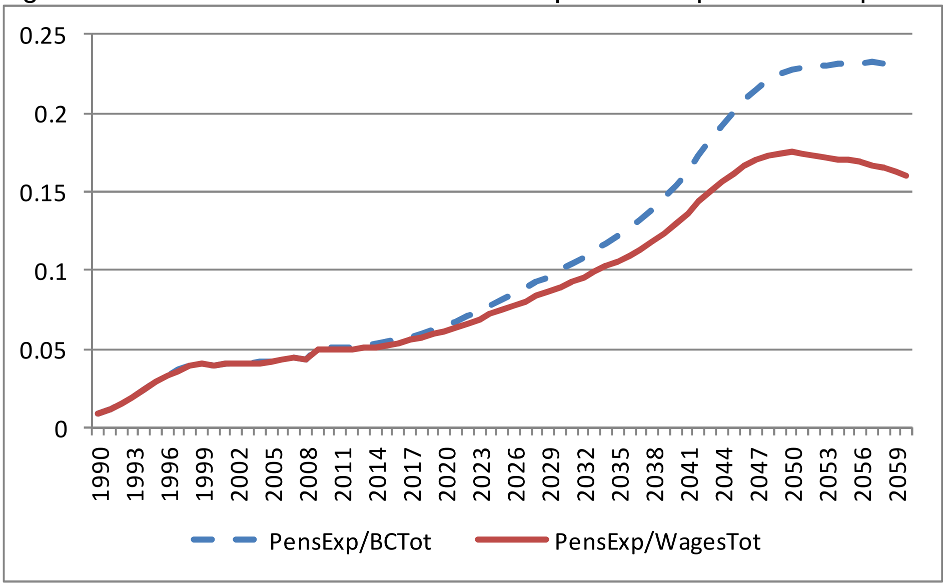
b) Changes in the entry pension (No and level)



Source: Authors elaboration.

We take as a baseline scenario the one that incorporates the effect of the crisis. Figure 5.3 shows the evolution of retirement pension expenditure as a share of both the wage bill and the contribution bases in the sample. The low values in the first years are not representative. It is interesting to note the increasing distance between the two ratios. This is reflecting the role of the contribution thresholds.

Figure 5.3 Evolution of the ratio retirement pension expenditure to potential revenue



b) The effects of the 2011 reform

Table 5.1 gives an overview of the reforms introduced in 2011. The first two measures aim at increasing contributiveness, or proportionality between contributions and pensions. The first increases the amount of past years contributions considered to compute the basic pension amount -*BR*- from 15 to 25 and the second modifies the share of it received as a pension –p(n)- to make it more linear and increasing the total number of years needed to obtain the 100% of BR. The last and most discussed measure modifies the reference retirement age from 65 to 67. The reform is implemented gradually from 2013 to 2027. In the following we will the simulation results and discuss the effects on pension expenditure sustainability and redistribution.

Figure 5.4 shows the effect of the first reform set: reforms affecting *BR* and *p(n)*. The change in p(n) has a sizable effect on the level of pension reaching a 2.5% then the reform is fully implemented. This average negative effect is probably hiding positive and negative effects for people with different working careers. The corresponding effect on the number of entry pensions is very small, due to the small effect of incentives on the retirement decision.

With respect the changes in the number of years used to compute BR the result is surprising at first glance. The implemented reform has on average an unexpectedly positive sizable effect reaching 2%. The expected effect of this measure depends on the shape of the lifetime real earning profile.[16] It this increasing, when BR takes more years in the past this implies a reduction in the level of wages considered and, hence, a cut in pension rights. Nevertheless the earnings career does not always grow to the same extent along the lifecycle. It is expected to grow more at the beginning, to stabilize around the 50s and then keep constant of eventually worsen if the working career is interrupted because of unemployment. Hence, the effect of this measure can be an increase in pension if wages are not growing in real terms. The fact that we obtain this result on average is probably due to the way wages are projected in the simulation. This might be overstating the effect of age. In order to test the validity of this result we simulate alternative measures in this line. In particular the effect of applying this measure in 2007 is implemented. This simulation has the advantage of running on observed past wages and the effect goes in the expected direction: Entry pensions are cut for the cohorts retiring between 2007 and 2020. From 2027 on then the effect is the same as the implemented reform. Finally, a continuation of the increase from 15 to 25 until 37 years is simulated.

[16] The relevant magnitude is real wages (and hence contributions) as the formula to compute BR uprates wages to inflation two years prior retirement.

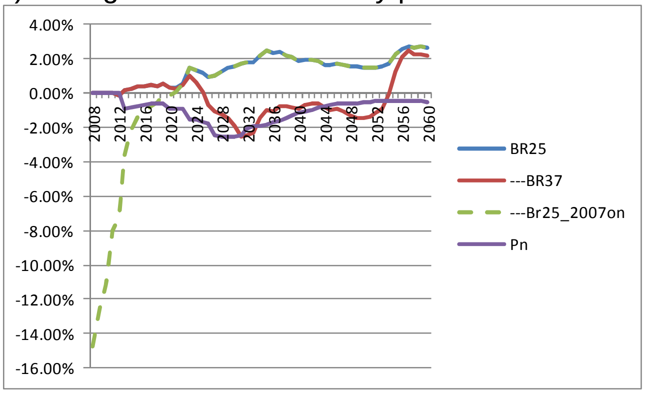
Interestingly, from 2027 the positive effect on pension starts decreasing and becomes negative. This confirms that earnings profiles are growing steeply in the first years of working career, so that going back to those years to compute BR reduces the pension rights. Overall these results require more investigation. Again projection of lifetime earnings profiles proves crucial to analyze redistribution effects.

	2007 Situación	**Reform agreement 2011**
p(n)	*p(n)*= 100% ; *n*=35 Three levels for n - First 15 years: 50% - 16 to 25: 3% / year - 26 to 35: 2% / year	*p(n)*= 100% *n*=37 Three levels for n - First 15 years: 50% - next 248 months: 0,19 per month - next months: 0,18 per month (gradual implementation 2013-2027)
BR$[bc_{t-15},\ldots bc_{t-1}]$	*bc* from the last 15 years	*bc* from the last 25 years (gradual implementation 2013-2022)
Retirement age	General 65 Minimum 61 (except Old system)	General 67 (65 if n≥38,5) Minimum 63 (61 if involutary unemployment)
	Retirement premium (*) - n<40: 2% / year - n≥40: 3% / year (*) There is a maximum limit. Early retirement	Delayed premium (*) - n<25: 2% / year - 25≤n≤37: 2,75% / year - n>37: 4% (*) Maximum limit kept Early retirement relaxed (not yet simulated)

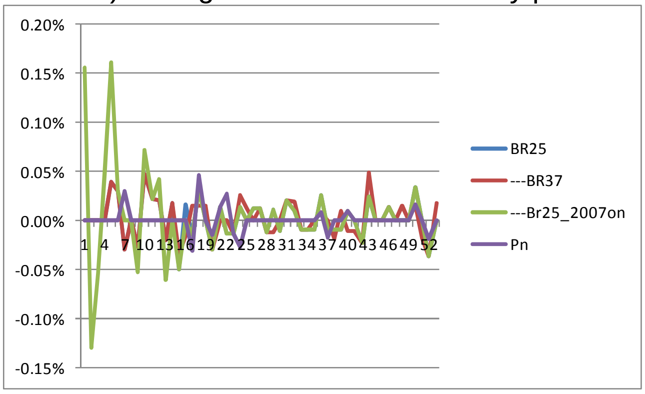
Finally Figure 5.5 shows the effect of delaying the normal retirement age. In this case the simulation needs to be done in three steps. First the effect on the pension level is computed. Second the age dummies are shifted. Third the premium to delay retirement is changed. Panel a) and b) in Figure 5.5 show that the main effect is driven by the age dummies. The number of entry pensions is substantially reduced during the implementation of the reform and this has a positive effect on pension level, probably because those who delay their retirement are the ones with lower pensions. Panel c-d) show the impact on the ratio expenditure to wage bill and on the retirement age. In this case the reform has a sizable impact. The observed average retirement age increases almost two years and this implies a cut in expenditure to wage bill of almost two points.

Figure 5.4 Effects of the reforms of BR and Pn (with respect to the baseline)

a) Change in the level of entry pension



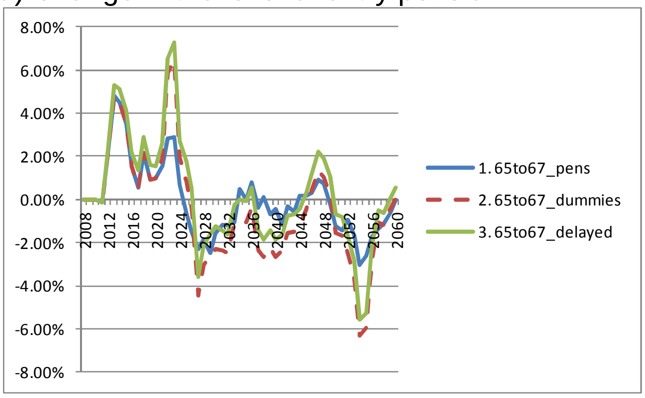
b) Change in the number of entry pension



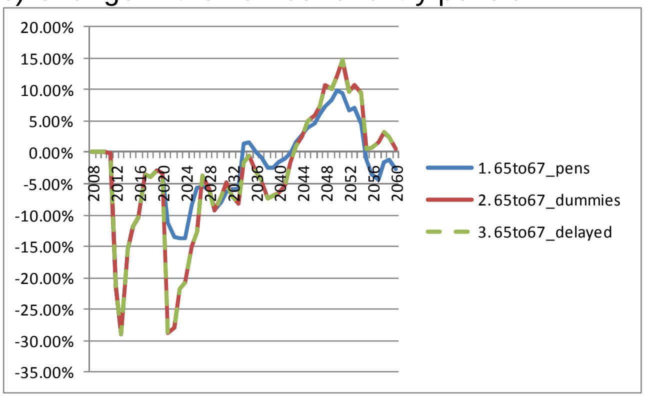
Source: Authors elaboration.

Figure 5.5 Effects of the reforms of retirement age (with respect to the baseline)

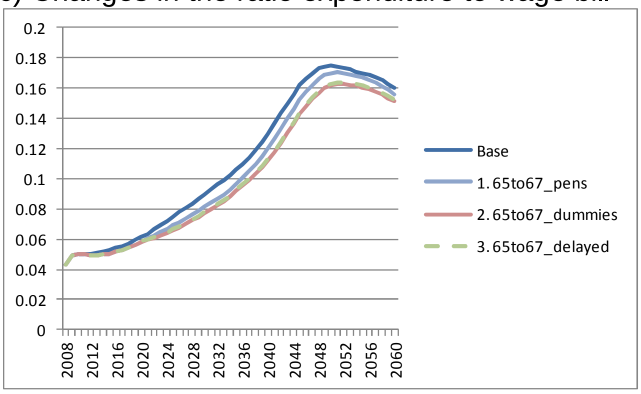
a) Change in the level of entry pension



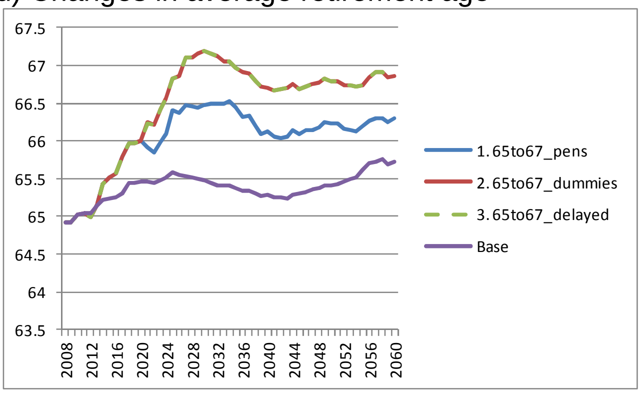
b) Change in the number of entry pension



c) Changes in the ratio expenditure to wage bill



d) Changes in average retirement age



Source: Authors elaboration.

6. Final remarks

In this paper we have shown a case based dynamic simulation model of the Spanish retirement pension system based on a sample drawn from the 2007 MCVL. Results show that the 2011 reform has had a substantial impact but does not seem to ensure sustainability. Further analysis is required in the following lines. First it is necessary to investigate the way wage grows and other assumptions affects results. Second alignment will be needed in order to measure the impact on sustainability from the selected sample. Finally, the implementation in the model of redistributive measures will allow measuring the effects of the reform not only on sustainability but also on pension's adequacy.

References

Argimón, I; C.I González; y R. Vegas (2007): Jubilación entre los 60 y 65 años. Algunas características. *Presupuesto y gasto público*, Nº 47, pp. 161-184

Argimón, I; M. Botella, C.I González; and R. Vegas (2009): "Retirement behavior and retirement incentives in Spain", *Documento de Trabajo, Banco de España* Nº 0913.

Benavides, A. (2006). *Análisis Práctico de las Prestaciones de la Seguridad Social*. Lex Nova, Valladolid.

Boldrin, M., S. Jimenez, and F. Peracchi (2004). "Micro-modeling of Retirement Behaviour in Spain", In *Social Security and Retirement around the World* (Jonathan Gruber & David Wise, eds.). NBER, University of Chicago Press, 2004, 499-567.

Börsch-Supan, A., R. Schnabel, S. Kohnz and G. Mastrobuoni (2004): "Micro-Modeling of Retirement Decisions in Germany". In *Social Security and Retirement around the World* (Jonathan Gruber & David Wise, eds.). NBER, University of Chicago Press, 2004, 499-567.

Eklof, M.-Hallberg,D.(2006), "Estimating retirement behaviour with special early retirement offers", Department of Economics Discussion Paper, Uppsala University.

European Policy Committee (2009). 2009 Ageing Report, Economic and Budgetary Projections for the EU-27 Member States (2008-2060). European Economy 2/2009.

García-Pérez, J.I., S. Jiménez-Martín and A. R. Sánchez-Martín (2010) Retirement incentives, individual heterogeneity and labour transitions of employed and unemployed workers, UPF WP 1239.

Gruber and Wise (2004): "Introduction and Summary", In *Social Security and Retirement around the World* (Jonathan Gruber & David Wise, eds.). NBER, University of Chicago Press, 2004, 499-567.

Herce, J.A. y J. Alonso (1998). *Los Efectos Económicos de la Ley de Consolidación de la Seguridad Social: Perspectivas Financieras del Sistema de Pensiones tras su Entrada en Vigor*. *Documento de Trabajo de FEDEA* 98-16.

Jimenez, S. (2006). "Evaluating the Labor Supply effects of Alternative Reforms of the Spanish Pension System", *Moneda y Credito*, 222, 271-307.

Jimeno, J.F. (2003). "La Equidad Intrageneracional de los Sistemas de Pensiones". *Revista de Economía Aplicada*, nº 31, pp. 5-32.

Maes, M. (2008a)"Survey of empirical research on retirement behaviour and retirement incentives", Working Paper 2008-18, Department of Economics, University of Brussels

Maes, M. (2008b)" Redistributive impact of reforming the old-age pension system in Belgium", Working Paper 2008-19, Department of Economics, University of Brussels

MEC (2010) Sistema estatal de indicadores de la educación, Instituto de evaluación, Ministerio de Educación.

Mincer, J. (1974). *Schooling, Experience and Earnings*. New York, Columbia University Press.

Monasterio, C., I. Sánchez, y F. Blanco (1996). *Equidad y Estabilidad del Sistema Español de Pensiones.* Serie Economía Pública, Fundación BBV.

MTAS (2003). *La Renovación del Pacto de Toledo*. MTAS, Madrid.

MTAS (2006a). *Acuerdo sobre Medidas en Materia de Seguridad Social*. MTAS, Madrid.

MTAS (2006b). *La Muestra Continua de Vidas Laborales*, Colección Informes y Estudios, Seguridad Social, Ministerio de trabajo y asuntos sociales (MTAS), Madrid.

Appendix A: Past and future reforms on the pension formula

This appendix summarizes the main features of past and future reforms affecting the pension formula. Past reforms in the pension formula refer to each one of the three factors Equation [1]. First, changes in RB basically relate to the amount of past BC used to compute the BR. When the system started only two years were considered. The number of years increased to 8 in 1985 (26/1985) and to 15 in 1997 (24/1997 Act). The current reform aims at increasing it to 25 still bellow the recommendation of the Toledo agreement of taking into accout the whole contribution history. Second, changes in the cc(n) have been varied and depend very much on the retirement path. Finally chages affecting *p(n)*, i.e. the weight attached to the number of contribution years, so as to compute the share of the *RB* received as a pension. Table I.1 summarizes the different legal changes undergone by this scale. According to the present one fixed by the 24/1997 Act, *p(n)* is decreasing after the minimum, so that the weight attached to the first years is higher – which results in a redistributive effect. The former scale gave a lower weight to the initial years. The following columns show further reforms in line with the Toledo Agreement proposal, fostering the Bismarckian nature of the system: specifically, full proportionality considering the present maximum of 35 years – that is 2.86% a year - or a maximum of 40 years – that is 2.5% a year. As shown in Table I.1, for the various scales considered, the weight attached to one particular year oscillates between 5% in the first ten years prior to the 1997 reform, to 2% during the last years of the current legislation.

Figure I.1 illustrates the potential effect of reforms on this legal parameter. On the one hand, the function *p(n)* is plotted for each of the legal scales considered. It is worth noting, first, that individuals who do not meet the minimum eligibility requirement would clearly benefit from a proportional rule. Second, it is clear that for those crediting between 15 and 35 years, both the previous and the present rules (26/1985 and 24/1997 Acts, respectively) are more generous than the two proportional rules.

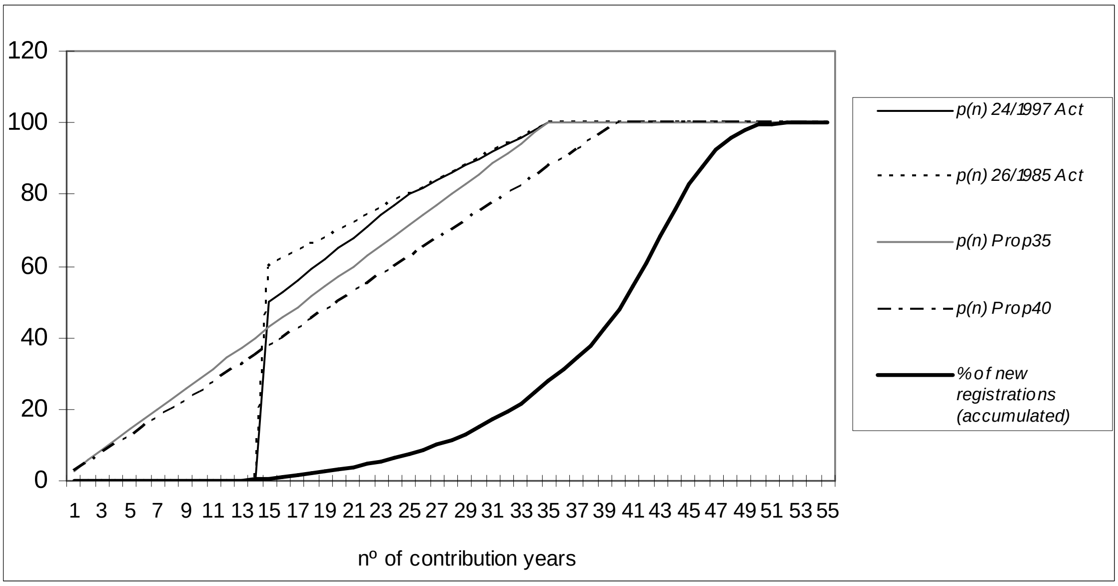
On the other hand, the cumulative distribution of new registration with respect to the number of contribution years, in the 2004 MCVL wave, is also shown. This highlights the share of individuals affected by each legal scale and hence its specific effect. First, note that most individuals – 72% – credit up to 35 contribution years and, as such, are affected neither by the legal changes already enacted nor by moving to a system of full proportionality with a maximum of 35 years. Second, we can see that the legal change introduced in 1997 only affected 6% of new pensions. Finally, it is interesting to note that an eventual change to full proportionality would affect almost 50% of individuals, which accounts for the highest effect obtained for this simulated legal change.

Table A.1. Weight attached to contribution years in the share of *RB* (several legal scenarios)

	Prior to 1985	**26/1985 Act**	**24/1997 Act**	**Total Proportionality**
Minimum eligibility condition	10 years	15 years	15 years	–
Contribution years	Total *p(n)* (per year)			
10	50% (5.0%)	–	–	–
15	(2.0%)	60% (*) (5.0%) (2.0%)	50% (3.3%)	In 35 years (2.86%)
16-25			(3.0%	In 40 years (2.50%)
26-35		(2.0%)	(2.0%)	

(*) 60%: according to the same previous scale, 50% from the first 10 years (5% a year) plus 10% from the next 5 years (2% a year).

Figure I.1. Average effect on pensions from fixing different functions of *p(n)*



Source: Authors' calculations using MCVL data and legal parameters.

Table A.2.: The 2011 reform in the pension formula

	2007 Situación	**Reform agreement 2011**
a) $p(n)$	*p(n)*= 100% ; *n*=35 Three levels for n - First 15 years: 50% - 16 to 25: 3% / year - 26 to 35: 2% / year	*p(n)*= 100% *n*=37 Three levels for n - First 15 years: 50% - next 248 months: 0,19 per month - next months: 0,18 per month (gradual implementation 2013-2027)
b) $BR(bc_{t-15},...,bc_{t-1})$	*bc* from the last 15 years	*bc* from the last 25 years (gradual implementation 2013-2022)
c) Retirement age	General 65 Minimum 61 (except Old system)	General 67 (65 if n≥38,5) Minimum 63 (61 if involutary unemployment)
d) $cc(n)$	<u>Delayed retirement(*)</u> - n<40: 2% / year - n≥40: 3% / year (*) There is a maximum limit. <u>Early retirement(**)</u> -Old system annual 8% before 65 - Unemployed: anual 7,5% before 65 (or lower if n≥35) - especial retirement at 64 with no reduction <u>Partial retirement (25-75% reduction)</u> - From age 61 to retirement age need to substitute the worker (minimum contribution base 65% of the old contract).	<u>Delayed retirement(*)</u> - n<25: 2% / year - 25≤n≤37: 2,75% / year - n>37: 4% (*) Maximum limit kept <u>Early retirement(**)</u> -Old system annual 8% before 65 -*n*<38,5: an additional 1,875% quarterly before legal retirement age -*n*≥38,5: an additional 1,625% quarterly before legal retirement age - It desapears (**) New possibility if stop working voluntarily (age≥63) or involuntarily (≥61), considering time to retirement age as contributed. <u>Partial retirement (25-75% reduction)</u> - = conditions (minimum contribution base 100% of the old contract)

Appendix B: The effect of retirement incentives on retirement decisions: The identification problem

Several problems arise in reduced from estimation to identify the effect of retirement incentives. First, constraints on the availability of longitudinal data imply that SSW is computed only for the last part of the life cycle in which the individual is observed. Hence SSW only contains the value of future expected pensions, avoiding the consideration of wages received before retirement. The Social Security Wealth of a worker (SSW_i)is calculated for each individual – the individual subscript is omitted here – as the expected present value of future pension benefits in case of retirement at age, *h*, higher than the actual age (*a*), as

$$SSW_h = \sum_{s=h+1}^{S} \gamma^{s-a} \pi_s B_s(h) \qquad \text{(B.1)}$$

Where S is the age of certain death, γ denotes the time discount factor, π_s is the conditional survival probability at age *s* for an individual alive at age *a*, and $B_s(h)$ is the pension expected at age *s>h* in case of retiring at age *h*.[17]

Four incentive measures (I) are usually derived from the measure of Social Security Wealth (*SSW*) defined in equation (B.1).[18] All of them measure gains in *SSW* from delaying retirement from now to a future age/year. First, the social security accrual (*SSA*) only considers the present age (*a*) and the next (*a+1*) and, hence, it is defined as the difference between SSW_{a+1} and SSW_a. Second, the implicit tax (*T*) measures the work disincentives as the ratio between *SSA* and the potential wage earned over the year, with a negative sign. Third, the peak value (*PV*) considers all the possible future ages(*h*)/years, up to maximum retirement age. Hence, it is defined as the maximum difference between SSW_h and the initial, given, SSW_a. Finally, the last incentive measure – the option value (*OV*) – is, on the one hand, similar to *T*, in that it takes into account future expected wages. On the other hand, similarly to *PV*, it measures the gain from postponing retirement to any future feasible age, but in terms of utility. In particular, following the specification in Gruber and Wise (2004), assuming that the utility function is linear in *SSW*, we can express the *OV* of retiring as a function of *PV* as follows,

$$OV_h = \sum_{s=h+1}^{S} \gamma^{s-a} \pi_s w^*(h) + zPV_h \qquad \text{(B.2)}$$

Being z the constant relation between utility and *SSW* in the utility function. The first term on the right-hand side in equation (B.2) is the present value of expected wages until retirement and the second term is proportional to *PV* – the gain in SSW from postponing retirement until the best option.[19]

Both I and SSW are usually included as explanatory variables in the estimations. The expected sign of the coefficient for I is negative, showing tha incentives to delay retirement reduce the probability of retirement. The coefficient for SSW is usually interpreted as a measure of the income effect the worker is experiencing when delaying retirement for one year and hence exchanging consumption for leisure. If leisure is a normal good, we should expect a positive coefficient. Some studies obtain the opposite sign. The reason might be that *SSW* is partially endogenous and is closely related to tastes for work, which might tend to have the opposite effect: the higher the income the lower the probability of retirement. Some studies try to palliate this problem by using instrumental variable estimates or by introducing a proxy for the taste for work as an explanatory variable. [20][21]

A non trivial issue in this setting is the specification of age as explanatory variable, given the strong effect of age in retirement decision. If single age dummies are specificities they capture most of the significance leaving a negligible effect for incentives measures. On the other hand, specifying age as a quadratic implies significant results but the age patter look unrealistic -See Moral-Arce et al. (2009) for details). [22]

Clearly there is an identification problem that also interacts with the effect of age. As Gruber and Wise (2004) point out, when age increases the retirement desire increases, but not necessarily linearly. The introduction of wage as an explanatory variable might help to capture heterogeneous tastes for work, but the problem is, then, that both age and wage enter the incentives calculation. Hence, the introduction of all these variables worsens the identification of the effect of the incentives. In fact, as seen above, the age dummies specification gives a very realistic shape or retirement probability by age but implies that age is the main explanatory variable in the regression. It is usually interpreted as the fact that they capture some of the legal settings and social norms not fully reflected by the incentive variables. Clearly the difficulties in capturing the role of the firm in the retirement decision are also behind this lack of significance.

---

[21] In a similar framework, Maes (2008b) obtains a positive sign for *SSW* and a negative one for lifetime wages, while Börsch-Supan et al. (2004) obtain a negative sign for *SSW* and a positive sign for estimated wage.

Appendix C: The MCVL data set and the sample extraction[23]

The MCVL is a sample extracted from Social Security administrative data. Four per cent of all individuals registered with the Social Security administration – both contributors and recipients of benefits – over the sampling year are selected and their entire life history in the social security records is included in the data set.[24] Thus, although it is not a pure panel, the data set is rich in longitudinal data. This feature, however, complicates the structure of the information as the registration unit varies substantially ranging from the person – in the personal data file – to the contract – in the affiliation file – or to the contract and year – in the contribution file. This structure also complicates the data selection. Furthermore, the quality of data is not homogenous, deteriorating the further back in time we go as more data are missing. The data collection itself was initiated at different points in time: data on pensions were first included around 1996; data on contributions around 1980; while some data on affiliation (contract registering) are available from as early as 1970. Clearly all these factors condition our analysis. We provide details of the data employed below. We focus primarily on the pension file whose registration unit is defined by the individual, the benefit and the year, but we also recover contributory data for those individuals in our sample.

Among the difficulties of dealing with such a large administrative data set – the sample size reaches about a million people in 2005 – the most challenging are dealing with empty contribution bases and relating contribution, affiliation and benefit data from the same individual, all defined with different registration units. In particular, in order to extract reliable data regarding contributions in a specific time unit, it is necessary to follow up all the contracts in which an individual has been involved, computing time and contribution separately so as to avoid an erroneous correspondence between working time and contribution per unit of time. Below we describe in detail how we dealt with this.

The model presented in this paper drives mainly on administrative data, extracted from the Social Security administration –the MCVL. seeks to examine the impact of certain reform measures on the probability of retiring. The MCVL allows for this kind of analysis because it contains data of the main factors included in the initial pension formula. Specifically, over the period covered by the data base, we are able to recover the number of working years, the life-cycle contributions of the individual and the retirement age, thus determining any penalizations for early retirement. It is also possible to recover these variables – except the not yet observed retirement age – to analyze future pension rights for potential pensioners – i.e., all the individuals in the sample who can opt for retirement. For pensioners, the total number of lifetime working years considered in computing the initial pension is also registered in the MCVL. Nevertheless, it is also necessary to obtain the annual working time in order to fill the gaps in the contribution data, in line with Spanish legislation.[25] Hence the annual contribution period or working time is obtained by recovering all the contracts signed by the individual for each year, taking into account part-time work as well as the possibility of contracts that ran simultaneously.

At the same time, the average hourly contribution is obtained. One of the main problems we faced was the existence of missing contribution data. This can occur either within a contract registered in the contribution file or due to a lack of correspondence between the affiliation data – starting long before 1980 – and contribution data – starting in 1980. For example, we might find, even after 1980, no recorded contribution

---

[23] See MTAS (2006b) for a detailed description of the *Muestra Continua de Vidas Laborales* (MCVL), available upon request at www.seg-social.es/Internet_1/Lanzadera/index.htm?URL=82.

[24] Both workers and pensioners are thus included and also individuals receiving unemployment benefits or benefits prior to early retirement. The latter can be identified by the type of relation they have with the Social Security.

[25] When the individual presents a non-contributory period within the last 15 years considered for computing the *RB*, the minimum contribution threshold rather than zero is considered to compute the *RB*.

for one specific worker, while data regarding affiliation showed the worker to be actually contributing.[26] An imputation process is developed to provide figures for the missing contributions. By tracking affiliation and contribution data, we treat missing values differently considering whether the individual is actually working – actual missing value – or not – if he is out of the labor force. In the former case, data from the same individual along the same year are used in order to recover absent contributions. If this is not available in the same year the individual is dropped from the sample. This implies keeping more than 60% of the 2007 sample. In the latter case, the gaps occurring in the last 15 years are filled with the minimum contribution threshold, according to Spanish legislation.

[26] Information regarding contributions was first gathered in 1980, but it is more reliable after 2001. The providers of the sample found that the share of contracts with missing data fell from 78% in 1984 to 94% in 1992 and to 99% in 2003.