Did the Intergenerational Solidarity Pact increase the employment rate of the elderly in Belgium? A macro-econometric evaluation.*

Catherine Smith
IRES - Université catholique de Louvain

April 9, 2014

Abstract

In December 2005, the Belgian government adopted the law on the Intergenerational Solidarity Pact (ISP) with the objective to increase the employment rate of the elderly. In order to meet that objective, several active ageing policies and reforms were taken. The aim of this paper is to investigate the overall effectiveness of the ISP in rising the elderly employment rate by gender. Two methods are used. Both rely on a macro-econometric model which explains the evolution of the elderly employment rate by the economic conditions. The first method uses forecasts of the macro-econometric model as an indicator of the value the employment rate would have taken in the absence of the policies. The second method tests for the presence of structural breaks after the introduction of the main policies of the ISP. The results of the first method suggest a positive impact of the policies on elderly employment rate which is slightly larger for men, and a negative impact on younger men's employment rate, suggesting a substitution effect. These effects are however too small to be statistically significant. Using the second method, no structural break is found.

JEL classification: J21, J26, H53, E32.

*I am grateful to Sophie Béreau, Vincent Bodart, Bart Cockx, Muriel Dejemeppe, Arnaud Dufays, William Parienté, Fatemeh Shadman and Bruno Van der Linden for their very useful comments. I also thank participants to the various seminars in Louvain-la-Neuve and the Belgian Day for Labour Economists in Leuven (December 2012) for their comments and suggestions. I finally thank the Belgian Federal Administration Science Policy for its financial support on the project "Age, women and employment: An evaluation" within the program Society and Future (CONTRACT NR TA/00/044).

1 Introduction

The challenge brought by the population ageing on the welfare system of many countries in the world is a well-known fact. In this respect, Belgium is particularly exposed because of policies introduced between the 70's and the 90's fostering early exit from the labour market. According to OECD statistics for 2005, in Belgium, the employment rate of men and women aged 55 to 59 years old was respectively 55.4% and 31.8% while they were 67.2% and 48.6% on average in the EU15. In 2012, the situation has somewhat improved with an employment rate of 63.9% and 48.4% for elderly men and women in Belgium compared to 73.2% and 59.4% on average in the EU15 countries. Under the impulse of the Lisbon strategy, Belgian policy-makers began to acknowledge the issue of population ageing and adopted a package of active ageing policies on December 2005, called the law on the Intergenerational Solidarity Pact (*Pacte de Solidarité entre les Générations*). The objective of the Intergenerational Solidarity Pact (ISP) was to increase the employment rate of elderly people. The main policies consist in, by order of importance, a permanent wage subsidy, an easy access to working time reduction with wage compensation and a pension bonus. The first two policies are targeted to people aged at least 50 years old. The pension bonus amounts to 2 euros per day of full-time work while being 62 years old and over.

The aim of this analysis is to determine if the increase in the elderly employment rate by gender observed after the introduction of the ISP is mainly due to the good economic conditions at that time or if the policies of the ISP could have played a role. The evolution of the employment rate of younger age groups is also analyzed to investigate possible substitution or complementarity effects of the ISP. The analysis stops in 2008q2 before the financial and economic crisis.

What are the mechanisms through which the policies of the ISP should have an impact on the elderly employment rate?

For Belgium, the best way to analyze the incidence of a permanent wage subsidy on the employment rate is through a simple wage-setting, price-setting framework. Indeed, in Belgium, wages are negotiated by unions and they have a particularly strong influence on elderly workers' wages as many sectoral collective wage agreements consider a wage premium based on age.[1] As a consequence, Belgium has an increasing wage profile (while for many countries it is bell-shaped). The impact of the subsidy increases with its amount and with the elasticity of the price-setting (or labour demand) and of the wage-setting curves. Concerning the elasticity of labour demand, Piekkola (2004) estimates a low value for people aged between 40 and 64 years old in Finland (-0.1 in the short and -0.2 in the long run). Then, as unions have a strong influence in Belgium, the wage-setting curve is also expected to be quite inelastic. The subsidy is granted to all workers, not only the newly hired, but the amount is quite low[2]. Since the elasticities of both the price- and wage-setting curves are expected to be small and as the amount of the subsidy is low, the effect on elderly employment is expected to be positive but small.

[1] This holds until 2009 when age was replaced by tenure but this paper does not go beyond 2008q2.

[2] The subsidy corresponds to 50 euros for someone aged 50 and 800 for someone aged 65.

Regarding the second policy, allowing a reduction of working time at the end of the career and compensating (partly) for the wage loss is believed to increase the employment rate of older people. Indeed, working hours are not completely flexible and people cannot always reduce their working time as they would like (Gustman and Steinmeier, 1983; Kahn and Lang, 1995; Stewart and Swaffield, 1997; Euwals, 2001; Bryan, 2007; Decicca, 2007). They might then choose to leave the labour market early. Consequently, a scheme that allows for more flexibility in the amount of hours worked could induce people to stay longer on the labour market. However, on the demand side, the policy could have positive (a better employee satisfaction) as well as negative (uncertainty about the return of the employee after the break, organizational difficulties) effects on employment (Devisscher and Sanders, 2007).

The third policy consists in a pension bonus received by people continuing to work after the age of 62. The expected positive effect is to increase the opportunity cost of early retirement and induce people to work longer (Maes, 2010). However, there is also an income effect working in the other direction: the bonus increases the expected revenue from early retirement and pushes people to consume more leisure (Maes, 2010).

At last, the ISP could have affected positively the perception of elderly people on the labour market which could potentially have a positive impact on both the demand and supply sides of the labour market.

In sum, the effect of the policies on the employment rate of elderly are expected to be positive but small.

A macro-econometric analysis is used to evaluate the impact of the ISP reforms taken as a whole. Indeed, such an evaluation is crucial for policy making and has never been rigorously completed for Belgium. The evaluation literature of labour market policies often take a micro-econometric approach. The motivation for taking a macro-econometric approach is to capture general-equilibrium effects while micro analysis only captures partial-equilibrium effects.

Two methods are used to identify a possible impact of the ISP. Both methods rely on a macro-econometric model which explains the evolution of the elderly employment rate by the economic conditions (economic fluctuations being the main determinant of the employment rate) separately for men and women over the period 1997 to 2007 and then evaluate the impact of the ISP over the period 2007 to mid-2008. The main hypothesis of the evaluation methods is that, in the absence of public policies, changes in the employment rate are closely linked to GDP growth rate. This hypothesis is quite standard in the empirical literature on aggregate employment equation which usually also allows labour productivity and labour cost to impact on employment (Bodart *et al.*, 2009; Mourre, 2006). The difference with this literature is that this analysis works by gender and age group which makes the use of those two variables impossible. Therefore, the employment rate is modeled as a function of GDP and its own past dynamics only. In this paper, the evaluation stops at the beginning of the financial and economic crisis of 2008 because anti-crisis employment policies were taken by the Belgian government at that time. Thus after mid-2008, the effects of the policies of the ISP cannot be disentangled from those taken to counteract the effect of the crisis. In principle, determinants of the

employment rate other than the ISP could have changed during that period. However, no important shock other than the ISP that could have had a major influence on the elderly employment rate at that time has been found. In the first method, the elderly employment rate is forecasted to build a "counter factual" employment rate over the period 2007-2008 in the absence of the policies of the ISP. The actual elderly employment rate is then compared to the forecasts. The second method tests for the presence of structural breaks in the relationship between the elderly employment rate and GDP using recursive Chow tests. Recursive Chow tests are used when one does not want to make hypothesis on the date of the break and are based on the recursive estimation of the model. For recursive Chow tests, an initial date is chosen and the Chow statistic is computed for the initial date and all subsequent ones.

The first method was previously used by Kopits (1978) to ascertain the employment effect of a temporary wage subsidy in France (incentive bonus for job creation) for the period 1959-1975. The author estimates an employment equation[3] by industry for the period before the subsidy was made available. Then, employment is predicted over two quarters and the difference between the actual values and the forecasts are used to evaluate the impact of the subsidy.

The evaluation method of this paper is also close to the one used in the papers of Dmitrijeva and Hazans (2007), Fahr and Sunde (2009), Hertweck and Sigrist (2013) and Klinger and Rothe (2012) which evaluate the effect of the Hartz I/II and III reforms in Germany on the the matching process between unemployed and vacant jobs. They add to their specification of the matching process a treatment dummy that indicates whether the package of Hartz reforms was in place or not (the dummy is equal to zero before implementation and one after). Doing so, they only test the effect of the reform on the constant of the model (the matching efficiency). Fahr and Sunde (2009) also run Chow tests to evaluate if there is a structural change in the matching process as a whole. Unlike those papers, this one does not use treatment dummy to test for an effect of the policies but rather uses forecasts as a counter factual and tests for breaks using recursive Chow tests. The advantage of using forecasts compared to using dummies is that it allows the effect of the policies to be more flexible and there are less parameters to estimate. Moreover, the strength of recursive Chow tests compared to regular Chow tests is that they do not make assumption on the time of the break.

A more common method for macro-econometric evaluation of labour market policies, that became popular in the 90's, uses variations in program scale (spending or number of participants) over time and/or across regional units or countries (de Koning, 1993; Lehmann, 1993; Calmfors and Skedinger, 1995; Hujer *et al.*, 2002; Jackman *et al.*, 1990; Layard *et al.*, 1991; OECD, 2011). For the evaluation of the ISP, this method cannot be replicated because of the large number of policies implied and data limitation.

As far as I know, for Belgium, there is only one paper, Bassilière *et al.* (2007), that investigates the macro-economic impact of one policy of the ISP which is the reduction in employers social security contributions for older workers. Bassilière *et al.* (2007) use simulations of a macro-econometric model with and without the reductions in social

[3]Employment by industry is explained by past production and the wage rate.

security contributions and compare the two scenarios. They find a positive impact on the employment of the elderly but a negative impact on low wage younger workers.

This paper proceeds by explaining the institutional framework of the employment policies targeted at the elderly in Belgium in section 2. Section 3 is devoted to the description of the data. Section 4 explains the methods used to evaluate the impact of the policies of the ISP. In section 5, the results are presented and interpreted. The final section concludes.

2 Institutional background and the policies of the ISP

This section is divided in two parts. The first gives a brief description of the organization of the social security and pension systems in Belgium. The second is devoted to the description of the policies taken by the ISP.

2.1 Institutional background

Belgium is a federal state that has decentralized some policies to regional authorities. The social security and pension systems are organized at the federal level. The payment of unemployment benefits (UB) and the issuing of sanctions in the case of non-compliance with the rules are federal competencies organized by the National Office for Employment. The regions, in turn, are responsible for the Public Employment Services which are in charge of counseling, job search assistance, intermediation services and training of unemployed and employed workers. The policies of the ISP where taken at the federal level.

Belgium has a pay-as-you-go public pension system. All employees and self-employed people pay compulsory contributions to the federal state funds for old age and invalidity pensions. The legal pension age is 65 years old for men and was brought from 60 years old in 1996 to 65 years old in 2009 for women. In this setting, no minimum years of contributions is imposed. Anticipated retirement is allowed up to 5 years before the normal retirement age with at least 35 years of contributions (including periods on unemployment or disability benefits). The basic state pension depends on the occupational status, the gross salary, the number of years worked and the family composition.

However, in Belgium, fours schemes allow an early withdrawal from the labour market. Those mechanisms are described as they were before the introduction of the ISP.

First, the status of unemployed with company bonus (*chômage avec complément d'entreprise*) entitles workers aged 60 years old and over who have lost their job to UB with a top-up complementary benefit paid by the previous employer. There is also a minimum years of contributions required, 35 years for men and 28 for women. People that benefit from this early retirement scheme are also exempted from job search and pension cover is continued during that period. The employer has to replace the early retired for a period of at least three years. The age threshold can be lowered up to 58 years old by means of collective bargaining agreements within industries. A lower age is possible in some sectors depending on more stringent career conditions and for company

recognized as being in economic difficulty.

The second scheme, and the most peculiar one, is known as the regime of old aged unemployed exempted from job search (*chômeur âgé non demandeur d´emploi*). Unemployed people aged between 50 and 57 years old with very long careers or aged at least 58 years old are exempted from job search but continue to be eligible for UB without any restriction in terms of availability on the labour market.

The third scheme is the full-time time-credit (*crédit temps à temps plein*). People of all age can interrupt their career for a period of one year if they have worked for the same employer for at least 12 months out of the 15 last months. The period of full-time time-credit can be extended up to 5 years by means of collective bargaining agreements within industries. People on full-time time credit receive an allowance from the National Office for Employment. Workers at the end of their career can then use this scheme right before being eligible for other routes toward early withdrawal.

Disability benefits constitute the last scheme leading to early withdrawal. It requires that the workers are not able to work.

2.2 The policies of the ISP

Even though Belgium has not waited the introduction of the ISP to act in favor of an increased activity of the elderly[4], the ISP marks a turning point through the number of policies taken on various aspects of older workers employability.

This section presents the policies introduced by the ISP and the major changes brought to existing policies. It begins by describing the policies and then looks at the magnitude of the policies in terms of number of participants and expenditure. The age threshold depends on the policies. Most of time they target people aged 50 years old and over and sometimes people aged 45 and over.

2.2.1 Description

The three main policies consist in a pension bonus, a reduction of employers' social security contributions (ESSC) and the strengthening of phased retirement. The pension bonus was introduced in January 2007. All workers and self-employed aged at least 62 years old with at least 44 years of career receive a pension bonus of 2 euros per day of full-time work. Then, the government introduced a reduction of ESSC for workers aged between 50 and 56 years old with a quarterly salary below 12,000 euros in April 2007. The quarterly reduction amounts to 50 euros for a worker aged 50 years old and increases by 50 euros every years until age 56. The reductions are permanent and concern the newly hired people as well as the existing labour force reducing the comparative cost of the elderly labour. For people aged 57 and over, the reduction is reset to 50 euros and an additional 50 euros per year worked over that age and it can be combined with ESSC reductions that existed before the ISP but only for people aged at least 58 years old since 2002 and 57 years old since 2004 and that amounts to 400 euros per quarter. The

[4]Some policies have been taken in 2002 and 2004.

maximum reduction is attained at 65 years old and amounts to 800 euros (400 euros due to the 2004 measure and 400 euros to the 2007 measure). The strengthening of phased retirement began in June 2007. It broadens the access to the end-of-career time-credit, which was introduced some years before. The end-of-career time-credit allows workers aged 50 years old and over to reduce their working time until retirement and benefit from generous allowance if they can prove 20 years of career and 3 years (5 years before June 2007) of tenure.

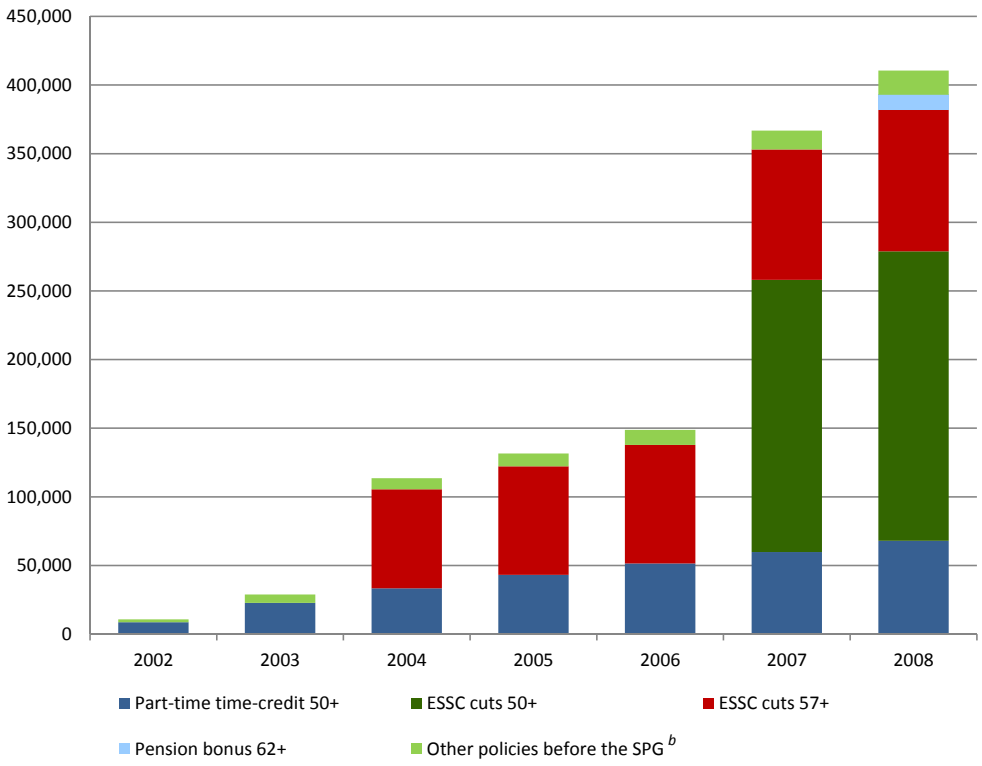
In 2007, the ISP also constrained people unemployed with company bonus to stay available for the labour market until the age of 58. Some policies were also introduced to promote the adaptation of the workplace to older workers but it was not widely used.

2.2.2 Participants and expenditure

This subsection focuses on the description of the data gathered on the number of participants and expenditure for policies targeted at older workers during the period 2002-2008. Although the ISP was adopted in December 2005, the graphs begin in 2002 to be able to see the changes the ISP brought as some policies existed before. The graphs stop before the economic and financial crisis of 2008.

Figure 1 shows the annual number of participants by measure. It reveals a small increase in the number of participants between 2003 and 2004 due to the introduction of the reduction of ESSC for workers aged 57 years old and over. This rise is "artificial" as the policy existed since 2002 for workers aged at least 58 years old but there are no data for the period 2002-2003. There is also a strong increase between 2006 and 2007 mainly due to the introduction of ESSC cuts for people aged at least 50 years old. The policies with the biggest number of participants are the cuts in ESSC (198,326 recipients aged 50-56 years old in 2007) and the part-time time-credit (59.741 mean payments per months in 2007).

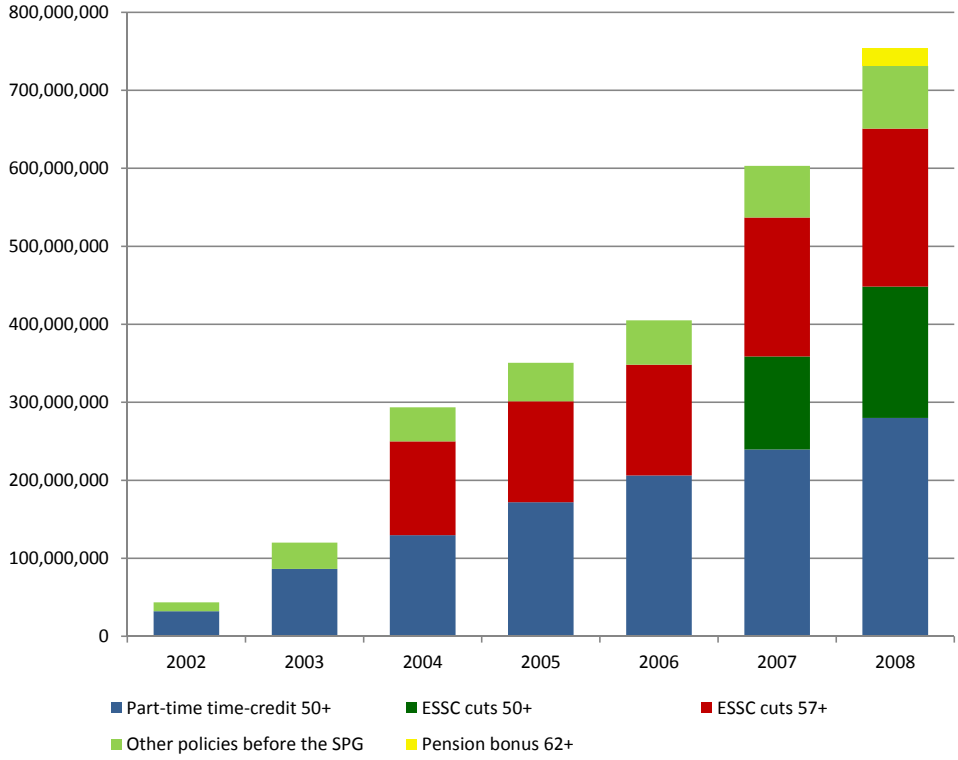
Figure 1: Recipients by policy from 2002 to 2008[a]



[a]The statistics are either measured by the number of recipients or by the number of payments (which does not always correspond to the number of recipients). For policies that deliver a permanent monthly or quarterly benefit, data are annualized by using the monthly or quarterly mean by year in order to avoid double counting.
[b]The category contains two policies: "plan activa" which is a hiring subsidy received by the employer who hires a long-term unemployed aged at least 45 years old and a hiring bonus (*complément reprise du travail*) paid to the newly hired insured unemployed aged 50 years old and over.

Figure 2 gives the yearly public expenditure in constant euros (base year 2010) for the policies of the ISP. The real expenditure went from 44 million euros in 2002 to 780 million in 2008. There is a rise in the expenditure between 2003 and 2004, which is again artificially (cf. above), and between 2006 and 2007 after the introduction of the cuts in ESSC for workers aged 50 years old and over.

Figure 2: Expenditure by policy from 2002 to 2008, constant euros, base year 2010



The year 2007 appears as pivotal for the ISP with the introduction of ESSC cuts for workers aged 50 years old and over. There is indeed a sharp rise in the number of participants and an increase in the spending at that time.

3 Data and sample

For this study, a national indicator of quarterly salaried employment rate by gender and age group in the private sector is built using data on national salaried employment and administrative data on the population by gender and age group. The focus is on the employment rate because it is the statistic that Belgian politicians consider as the objective of the policies of the ISP.

The quarterly salaried employment data by gender and age group from the National Social Security Office (NSSO) are available since the first quarter of 1997 (1997q1). The analysis is restricted to the private sector for two reasons. First, the private employment is more reactive to fluctuations in the economic activity than the public sector. Second, the private sector, as opposed to the public sector, is targeted by all policies of the ISP. The salaried private employment is available in fourteen age groups (<18; 18-19; 20-21; 22-24; 25-29; 30-34; 35-39; 40-44; 45-49; 50-54; 55-59; 60-64; 65>; missing). The data

are adjusted for seasonality[5] and corrected for a change in the classification that took place in 2003q1[6].

As the level of employment is influenced by the evolution of the population, the analysis is done using a measure of the salaried private employment rate referred as "the employment rate" for short. The employment rate is obtained by dividing the salaried private employment of a particular gender and age group by the population of that same gender and age group.[7]

The target of the ISP is the employment rate of the elderly defined, in this paper, as people aged between 50 and 59 years old. Employed people aged 60 years old and over are not taken into account because the main hypothesis of the methodology used in this paper, responsiveness to fluctuation in economic activity, does not hold for them. People aged 45 to 49 years old are also left apart because they are not concerned by all policies of the ISP. Younger age groups, people aged 35-44 years old and 18-44 years old, are, in turn, taken into account to investigate possible substitution or complementarity effects. The private salaried employment rate by gender for the age group 50-59 years old is depicted in figure 3 where the steady increase in the elderly employment rate since 2006 appears clearly for men while for women, the evolution is dominated by an upward trend.

In this paper, the employment rate is explained by an indicator of the economic activity. The obvious candidate is the gross domestic product (GDP) measured by the seasonally adjusted Belgian quarterly real GDP from the National Bank of Belgium.

As already mentioned, for the analysis, the sample is restricted to the period spanning from 1997q1 to 2008q2. The analysis stops at the beginning of the financial and economic crisis of 2008 because anti-crisis employment policies were taken by the government at that time. Thus after 2008q2, the effects of the policies of the ISP cannot be disentangled from those taken to counteract the effect of the crisis. Figure 4 shows the evolution of Belgian GDP for the period 1996q1-2008q2. Belgium underwent two

[5]For seasonal adjustment, I use the software developed by Eurostat, Demetra+, with the specification X12ARIMA RSA5C.

[6]During the first semester of 2003, the composition of employment measured by the NSSO changed due to harmonization between the different Belgian social security bodies. The dataset used in this paper is concerned by the introduction of two new categories of workers: disabled workers with more than 12 months of incapacity and people working in childcare. As a consequence, a level shift shows up in the employment data in 2003q1. To prevent those changes to have an impact on this analysis, the data are corrected before the statistical analysis. Because data on disabled workers and people working in childcare are not available by gender and age group for the period 1997-2002, the adjustment is made by replacing the permanent blip in the first difference of the employment by a linear interpolation and then rebuilding the level of employment before 2003 using this new first difference for 2003q1. Doing so, I use the method proposed by Nielsen (2004) for additive outliers.

[7]The data on the population by gender and age group are annual and measured as the resident population on the 1st of January of each year. The last data on the population by gender and age group is from January 2011. The population is available in 12 age groups (15-19; 20-24; 25-29; 30-34; 35-39; 40-44; 45-49; 50-54; 55-59; 60-64; 65>). As the employment data in the numerator of the employment rate are quarterly data, the annual population data are transformed in quarterly data. This is done by linear interpolation. The data of January for a given year are assigned to the fourth quarter of the previous year.

economic slowdowns, the first at the end of 2001 and the second at the end of 2002. Then, Since 2003, the Belgian GDP grew steadily.

Figure 3: Private salaried employment rate of people aged 50-59 years old. Panel A: men. Panel B: women.

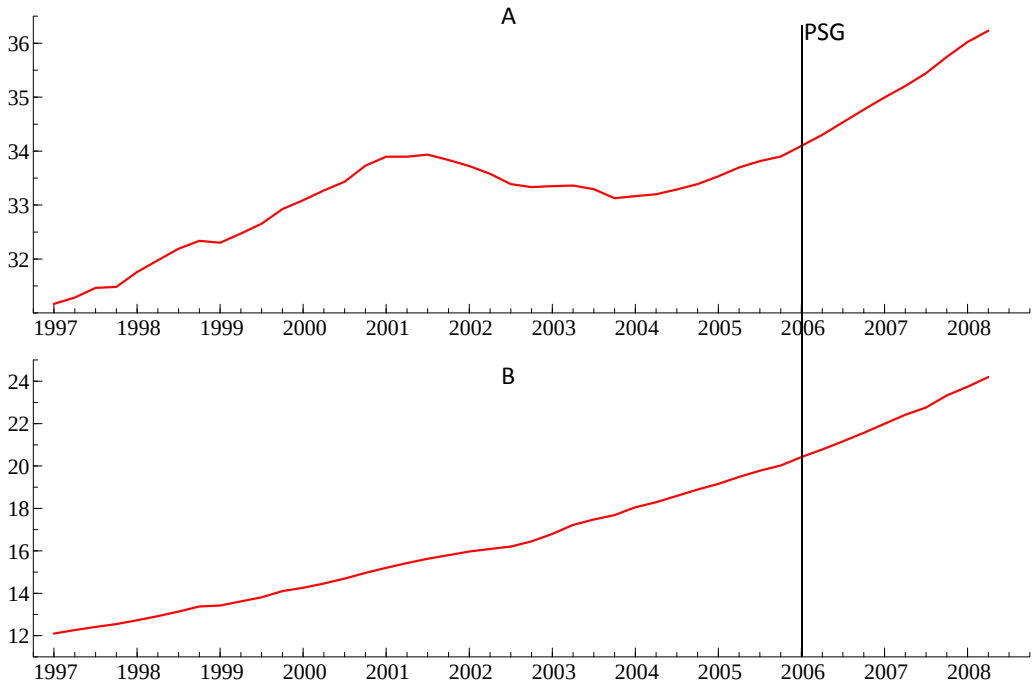
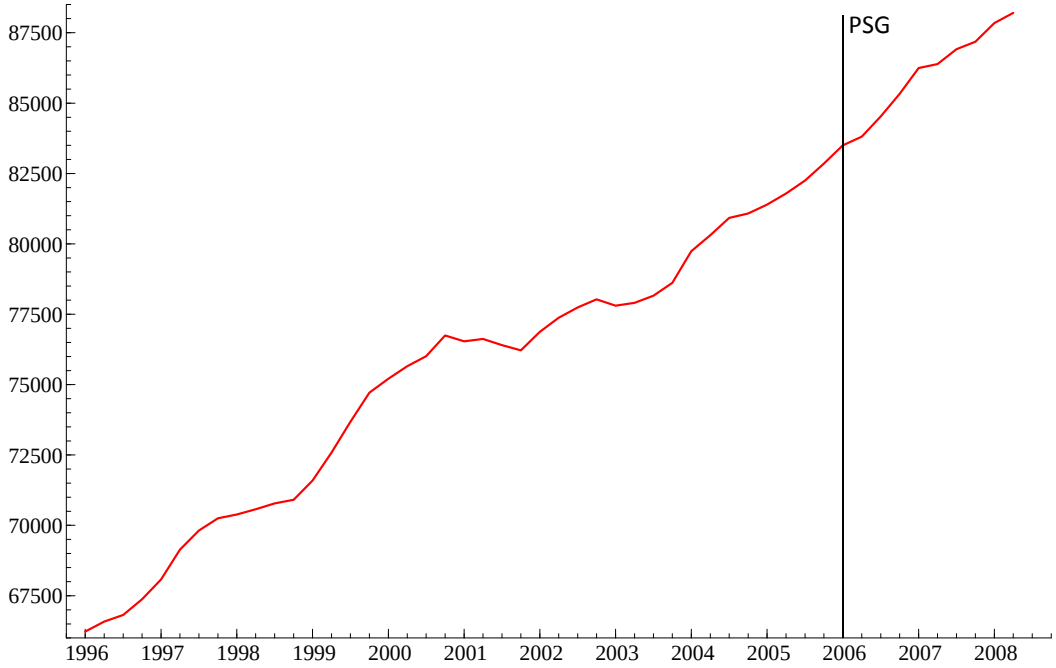


Figure 4: Belgian real GDP SA.



4 Evaluation methods and identification issues

Because the main hypothesis of the evaluation methods is that, in the absence of public policies, changes in the employment rate are positively linked to GDP growth rate, this section begins by describing this relationship.

As we know, most macro-economic variables are non-stationary series and indeed according to the Augmented Dickey-Füller (ADF) unit root tests, both the employment rate and GDP in the dataset are integrated of order one. The results of ADF tests are given in section 7.1 of the appendix.[8]

Regressions among variables integrated of order one are spurious unless they involve cointegrated variables. Could it be the case that there exist a long term equilibrium relationship between the employment rate and GDP? Due to the small sample size, a long term relationship between GDP and the employment rate in Belgium, if it exists, may be harder to identify. Being in the bivariate (two variables) case, the Engle and

[8]For the employment rate of men aged 35-44 years old, the ADF test cannot actually reject the unit root in the first difference but it is considered to be due to the small sample size and to the low power of the test, cf. section 7.1 in the appendix.

Granger (1987) test for cointegration is first used. Then as the Engle and Granger (1987) tests has some important defects[9], the robustness of the results is checked using the Johansen (1988) cointegration test. The Engle and Granger (1987) and the Johansen (1988) tests conclude both that there is cointegration between the employment rate and GDP for men aged 15-44 years old. However, they differ concerning men aged 35-44 years old. The Engle and Granger (1987) test finds cointegration while the Johansen (1988) test does not. For all other age groups and gender, no cointegration is found between the employment rate and GDP. On the basis of those results the assumption is made that there is no cointegration between the employment rate and GDP. The results of the Engle and Granger (1987) and the Johansen (1988) tests are reported in the appendix section 7.2.

As no cointegration relationship was found between the employment rate and GDP, the model is estimated in first difference. The analysis is based on an autoregressive distributed lag model with the first difference of the employment rate as dependent variable and the growth rate of GDP as the explanatory variable. Indeed, GDP could have a lagged effect on the employment rate as it is typically assumed that due to adjustment costs on the labour market, employment does not adjust immediately. The model selection is automated using Autometrics[10] on a general unrestricted model with four lags of the first difference of the employment rate and GDP growth rate:

$$\Delta \text{emp}_{g,a,t} = \alpha_0 + \alpha_1 t + \sum_{i=1}^{4} \phi_i \Delta \text{emp}_{g,a,t-i} + \sum_{i=0}^{4} \gamma_i \Delta \log[\text{gdp}]_{t-i} + \Delta \epsilon_t \qquad (1)$$

In equation (1), $\Delta \text{emp}_{g,a,t}$ stands for the first difference of the employment rate where the subscripts g corresponds to gender, a to age and t to time. $\Delta \log[\text{gdp}]_t$ is the growth rate of GDP at time t. For women aged 50-59 and 35-44 years old, a linear trend is added to the regression as the first difference of the employment rate has a trend. I expect the coefficients of the growth rate of GDP in equation (1) to be positive if the employment rate is to increase when GDP grows.

Bodart *et al.* (2009) who estimate an aggregate employment equation for Belgium for the period 1980-2002, do not find either a cointegration relationship between employment, GDP and the labour cost. They estimate their employment equation in first difference. However, for the euro area as a whole, Mourre (2006) finds a cointegration relationship between employment, GDP, labour productivity and labour costs on the period 1985-2002 and estimates an error correction model.

The remaining of this section is divided in four parts. The section begins with a description of the two evaluation methods. The first method, the counter factual

[9]The Engle-Granger methodology is limited to a single-equation test which is problematic in bivariate models as the equilibrium relationship could be expressed in two ways (the employment rate as a function of GDP or GDP as a function of the employment rate). Moreover, the Engle-Granger test relies on a two-step estimator which has a higher uncertainty than one-step estimators.

[10]Autometrics automates General-to-Specific multiple path simplification search of a data coherent, general unrestricted model. Data coherency means stable parameters and Gaussian, serially uncorrelated, homoscedastic errors.

analysis, is described in section 4.1 and the second, based on recursive break tests, in section 4.2. The third part discusses identification issues and the last, the possible endogeneity issues related to the model.

4.1 Counter factual analysis

The first method is inspired by micro-econometric evaluation methods where the objective is to retrieve the counter factual event in order to estimate the treatment effect. The counter factual corresponds to what would have been the outcome(s) in the absence of the policies.

The idea is transposed to this analysis using as counter factual the forecasts of the model in equation (1). The idea is to estimate the model of equation (1) on the period before the ISP reforms and use it to forecast the employment rate after the reforms. Using forecasts to build the counter factual, the employment rate is cleaned from the possible effects of the policies of the ISP and is as it should be based on its past dynamics and relationship to GDP. For this method, a starting date has to be determined for the forecasts. The forecasts starts in 2007q2 because as seen in figure 1 and 2, there is a strong increase in the number of beneficiaries and expenditure for the policies of the ISP after the introduction, in April 2007, of the reduction of ESSC for people aged at least 50 years old. Those figures also showed that the policies introduced before the ISP had a non-negligible number of recipients before 2007q2 and I acknowledge the fact that the estimated model on the period before the ISP also embodies the effect of those policies. Using the model in equation (1), dynamic forecasts[11] for the period 2007q2-2008q2 (5 points forecasted) are generated. The forecasts stop in 2008q2 because of anti-crisis policies taken after that date. As robustness check, the analysis is also done using as starting date 2006q2 which corresponds to the introduction of the first policy of the ISP.

The model is in first difference but the forecasts should be in level. To do so, the model of equation (1) is estimated for the period 1997q2-2007q1 with the identity equation defining the first difference of the employment rate. For each gender and age class separately, this simultaneous equation model is estimated by Full Information Maximum Likelihood (FIML). The model is given by:

$$\begin{cases} \Delta \text{emp}_{g,a,t} = \alpha_0 + \alpha_1 t + \sum_{i=1}^{p} \phi_i \Delta \text{emp}_{g,a,t-i} + \sum_{i=0}^{q} \gamma_i \Delta \log[\text{gdp}]_{t-i} + \Delta \epsilon_t \\ \Delta \text{emp}_{g,a,t} = \text{emp}_{g,a,t} - \text{emp}_{g,a,t-1} \end{cases}$$

Then, one can compare the counter factual employment rate based on forecasts to the realized employment rate. A 95% confidence interval around the previsions is used to assess the significance of the difference between the forecasts and the actual values. The significant difference between the counter factual employment rate (forecasts) and the realized employment rate represents the part of the employment rate that cannot be explained by the past dynamics and the evolution of GDP and could be interpreted as

[11] For the autoregressive component, the observations since 2007q1 and then the forecasts are used, while for GDP, the observed values are always used.

the impact of the policies of the ISP. If the realized employment rate is above (below) the upper (lower) bound of the confidence interval, there is a positive (negative) effect of the policies of the ISP on the employment rate.

4.2 Recursive break tests

The second method tests for the presence of structural breaks in the relation defining the employment rate in equation (1) after the introduction of the ISP (and especially between 2007q2 and 2008q2). Using small samples, this method is more adapted to detect abrupt changes in relations. If one expects the policies of the ISP to have a gradual effect on the employment rate, the tests will probably not be able to capture it due to the small sample size.

To point out structural changes, three types of Chow tests based on recursive estimation of the models in equation (1) are used and the forward recursive coefficients are used to determine from which component of the model the break (if it exists) could come from. If a structural break is found in the employment rate after the introduction of the ISP (2006q2), I can conclude that the policies of the ISP had a significant impact on the employment rate. Then, the direction of the effect has to be interpreted using the recursive coefficients.

A forward recursive analysis consists in choosing an initialization date $T_1 < T$, where T is the last date of the sample, and estimating the model on the baseline sample for the period $1, 2, ..., T_1$. Then, the model is re-estimated by adding the remaining observations one at a time $T_1+1, T_1+2, ...T$. The null hypothesis of recursive Chow tests is the stability of the relationship.The idea is to test whether two sub-samples give significantly different parameter estimates and if this is the case, which observation causes the difference.

The first test is the 1-step Chow test which evaluates whether a single observation is part of a regression derived from all prior observations. The test statistic is given by:

$$F_{\text{1-step Chow}} = \frac{RSS_t - RSS_{t-1}}{RSS_{t-1}(t-1-k)} \sim F(1, t-1-k)$$

Where k is the number of parameters in the model, RSS_t is the residual sum of squares from the regression on observations $1, 2, ..., t$ and RSS_{t-1} from observations $1, 2, ..., t-1$. The statistic is provided for each $t = T_1, T_1+1, T_1+2, ..., T-1$ (the sample size is increased by one observation sequentially).

The second test is the Break-point Chow test, also called the n-down Chow test, which evaluates how well the regression obtained from a sub-sample $1, ..., t-1$ observations explains the remaining $t, ..., T$ observations. The test statistic is given by:

$$F_{\text{n-down Chow}} = \frac{RSS_T - RSS_{t-1}}{RSS_{t-1}} \frac{T-t+1}{t-1-k} \sim F(T-t+1, t-1-k)$$

Where RSS_T is the residual sum of squares from the regression on the whole sample. The statistic is provided for each $t = T_1, T_1+1, T_1+2, ..., T-1$ (the sample size is increased by one observation sequentially).

The third test is the Forecast Chow test or n-up Chow test which does the same as the n-down Chow test but proceeding backwards. That is, the sub-sample going from $t+1, ..., T$ is compared to the entire sample. It evaluates how well the regression obtained from a sub-sample $t+1, ..., T$ explains the preceding observations $1, ..., t$. The test statistic is given by:

$$F_{\text{n-up Chow}} = \frac{RSS_T - RSS_{t+1}}{RSS_{t+1}} \frac{t}{T-t-k} \sim F(t, T-t-k)$$

Where RSS_{t+1} is the residual sum of squares from the regression on observations $t+1, ..., T$. Then, the sample size is increased backward by one observation sequentially and the statistic is provided for each $t = T_1, T_1 - 1, T_1 - 2, ..., 1$.

To be able to plot on one graph the Chow statistic and the 1% critical value for each subsample of the recursive analysis, the Chow statistic is normalized by the 1% critical value which becomes a straight line at unity on the graph. Nevertheless, those recursive Chow tests do not say which coefficient of the model is at the origin of the break, that is why they have to be interpreted together with the recursive coefficients.

4.3 Identification issues

Using macro-economic data, there are typically many factors affecting the time series that one cannot control for. This analysis explicitly controls for the impact of GDP on the employment rate and for trend and autocorrelation behavior of the employment rate. Even so, if a significant difference is found using the first method and a structural break is found using the second method during the period 2007(6)q2-2008q2, it could be due to any shock happening during this period. However, I argue that such significant effect can be attributed to the policies of the ISP as no other policies targeted at the elderly or events could have had a major impact on the employment rate at that time.

Moreover, with the second method, the effect cannot be quantified while for the first method a lower bound for the effect can be provided (the difference between the critical value at 5% and the realized value of the employment rate).

4.4 Endogeneity

There are two endogeneity issues. The first one is due to simultaneity as GDP determines the employment rate but the employment rate influences GDP as well. The second endogeneity problem comes from the relation that could exist between GDP and the policies of the ISP. Indeed, some policies involve government spending which contributes to rising GDP in which case there would be a positive link between GDP and the magnitude of the policies. The link could also be the other way around: during bad times, governments tend to increase policy efforts. Those two endogeneity issues are discussed one at a time in the following.

First, if there is no autocorrelation in the error terms of the models, the simultaneity problem arises only when the best specification includes the contemporaneous value of GDP growth rate and not when it includes only lagged values. Even when the relation

is contemporaneous, the fact of working by age group weakens the causality from the employment rate to GDP. This is particularly true for the main group of interest, the elderly workers, as they constitute only 16% of the total employment in Belgium (figures for 2007q1). As will be presented in the next section about the results, the contemporaneous value of GDP only appears in the regressions for men. For those regressions, two robustness checks for simultaneity issues are done. The first one simply consists in removing the contemporaneous value of GDP. The second robustness analysis for simultaneity uses the mean real GDP of five neighbouring countries (France, Luxembourg, Germany, the Netherlands and the U.K.) instead of the GDP of Belgium. As Belgium is a small open economy, the GDP of neighbourig countries has a great influence on Belgian economic activity and should then influence employment while being less subject to the simultaneity issue.

Concerning the second endogeneity problem, a regression of the number of policy recipients and spending on current GDP was run and there was no significant relationship between them.

5 Results

Let us now turn to the results. This section begins with the OLS estimation of the models in equation (1) of the previous section where the first difference in the employment rate is expressed as a function of GDP growth rate. The estimations stop in 2007q1 because the effect of the ISP are evaluated on the period 2007q2-2008q2. The results are given in table 1. The coefficients of the growth rate of GDP are significant at 1% or 5% level for each gender and age group and as expected, an increase in GDP has a positive impact on the employment rate. Interestingly, the coefficients are bigger for men than for women for all age groups and decrease with age.

All specification tests are respected except homoscedasticity for the age group 18-44 years old where the standard errors reported are robust to heteroscedasticity. There is no autocorrelation in the error terms so, as discussed in the previous section, the simultaneity problem arises only when the best specification includes the contemporaneous value of GDP growth rate (for men).

Table 1: Estimation of the employment rate by gender and age group as a function of GDP - OLS

Dep. var.	Δ Employment rate					
	50-59 years old		35-44 years old		18-44 years old	
	Men	Women	Men	Women	Men	Women
Δ Employment rate$_{t-1}$	0.533***		0.418***			
	(0.103)		(0.117)			
Δ log[GDP]$_t$	0.087***		0.127***		0.23***	
	(0.027)		(0.036)		(0.081)[a]	
Δ log[GDP]$_{t-3}$	0.115***	0.048**	0.153***	0.127***	0.381***	0.346***
	(0.028)	(0.021)	(0.039)	(0.041)	(0.09)[a]	(0.1)[a]
Constant	-0.071***	0.066**	-0.054	0.134**	-0.285***	-0.129
	(0.026)	(0.029)	(0.036)	(0.056)	(0.103)[a]	(0.078)[a]
Trend		0.006***		0.003*		
		(0.001)		(0.002)		
Nbr obs.	39	40	39	40	40	40
Adj. R^2	0.671	0.551	0.601	0.187	0.494	0.345
AR 1-4 test F-stat	0.674	0.395	0.329	1.264	0.32	0.571

[a] Heteroscedasticity consistent standard error of White (1980). *** significant at 1% level, ** 5% level, * 10% level. The model selection is automated using Autometrics.

This section proceeds by presenting alternately the results for the two evaluation methods developed in the previous section and the results from regressions taking into account the simultaneity problem.

5.1 Results for the counter factual analysis

For this first method which uses forecasts, the power to detect an effect of the policies of the ISP depends on the size of the confidence intervals around the previsions which is determined by the goodness of fit of the model (adjusted R^2) and the standard errors of the coefficients which are reported in table 1. For older workers, the model explains 67% of the variability for men and 55% for women. For younger workers, the R^2 is smaller and is the smallest for women aged 35-44 years old, for which the model only explains 19% of the variability. Then, due to the small sample size, the standard errors of the coefficients are also quite high. With low R^2 and high standard errors, the confidence intervals around the forecasts will probably be quite wide. Moreover, the effect of the policies of the ISP, if there is one, is expected to be small so that this first method is likely to lack statistical power.

The results are represented in figure 5. The solid line corresponds to the observed level of the employment rate and the dashed line corresponds to the dynamic forecasts after 2007q2 and to the fitted values of the model before that date. The bars corresponds to the 95% confidence interval around the forecasts. As expected, the confidence intervals around the forecasts are wide.

For people aged between 50 and 59 years old, the observed values are above the forecasts for both genders which goes in the direction of a positive effect of the ISP on the employment rate of the elderly. However, as one can see from the confidence interval on figure 5, the difference between the forecasts and the observed employment rate is not statistically significant at 95%. Neither is it significant at 90% (the 90% confidence interval is reported in Table 2). Table 2 compares the observed and predicted values in 2008q2. The observed employment rate for elderly men was 36.23% while the predicted one was 35.71%. The gap is only 0.52 percentage point which is 1.46% of the predicted value. For elderly women, the observed and predicted employment rates are respectively 24.20% and 23.94%. The gap is smaller for women, 0.26 percentage point which corresponds to 1.09% of the predicted value. The effect is bigger for elderly men compared to women. Even tough the difference is not statistically significant, this positive gap says that the evolution of GDP alone on that period cannot explain completely the rise in the employment rate. If there is a positive impact of the ISP, it is small and there is not enough statistical power in this setting to capture it.

For younger men from the age groups 35-44 and 18-44, the observed values are below the forecasts which could suggest a substitution effect. It means that solely based on the past dynamics of the employment rate and on the evolution of Belgian GDP on the period 2007q2-2008q2, the evolution of the employment rate for men from younger age groups should have been greater than it actually was. However, the difference is not statistically significant which does not allow us to conclude to the presence of substitution effects. The observed and predicted values are reported in table 3. For men aged 35-44 years old, the predicted value of the employment rate is 54.67% while the observed value is 54.04%. The gap is only 0.63 percentage point. For the broader age class of men aged 18-44, the difference between the observed and predicted values, respectively 47.05% and 47.34%, is even smaller, 0.26 percentage point.

At last, for younger women, there is nearly no difference between the observed values and the forecasts. The evolution of the employment rate during the forecast period is almost entirely explained by the past dynamics of the employment rate and by the evolution of Belgian GDP on the period 2007q2-2008q2.

Table 2: Observed versus predicted values of the employment rate and level of employment for elderly men and women in 2008q2

		Observed value	Predicted value (s.e.)	Gap in % points[a] (s.e.)
Employment rate				
Men aged 50-59		36.23%	35.71% (0.31)	0.52 (0.31)
	95% CI		[35.07 ; 36.34]	
	90% CI		[35.18 ; 36.24]	
Women aged 50-59		24.20%	23.94% (0.18)	0.26 (0.18)
	95% CI		[23.58 ; 24.30]	
	90% CI		[23.64 ; 24.24]	
Employment				
Men aged 50-59		258,178	254,439	3,739
Women aged 50-59		172,366	170,553	1,813

[a] The gap is the difference between the observed value and the predicted value.

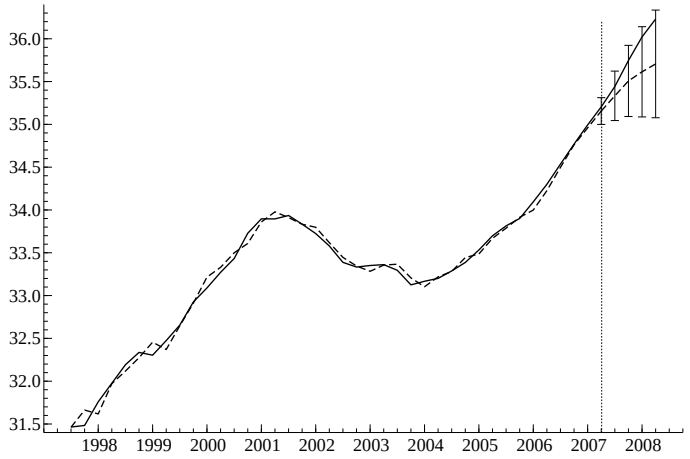
Table 3: Observed versus predicted values of the employment rate and level of employment for young men and women in 2008q2

		Observed value	Predicted value (s.e.)	Gap in % points[a] (s.e.)
Employment rate				
Men aged 35-44		54.04%	54.67% (0.36)	-0.63 (0.36)
	95% CI		[53,93 ;55,41]	
	90% CI		[54,06 ; 55,28]	
Men aged 18-44		47.05%	47.34% (0.47)	-0.29 (0.47)
	95% CI		[46,39 ; 48,28]	
	90% CI		[46,55 ; 48,12]	
Women aged 35-44		43.81%	43.91% (0.34)	-0.1(0.34)
	95% CI		[43,21 ; 44,61]	
	90% CI		[43,32 ; 44,49]	
Women aged 18-44		37.62%	37.56% (0.42)	0.06 (0.42)
	95% CI		[36,70 ; 38,42]	
	90% CI		[36,84 ; 38,27]	
Employment				
Men aged 35-44		430,012	431,650	-1,638
Men aged 18-44		1,010,271	1,014,139	-3,868
Women aged 35-44		339,339	337,532	1,807
Women aged 18-44		789,811	789,018	793

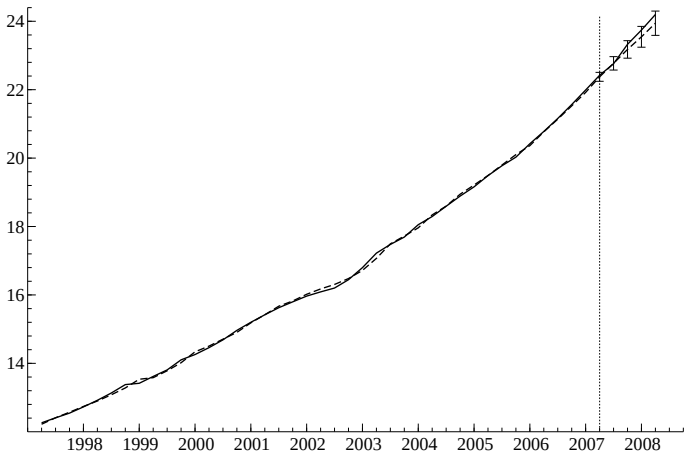
[a] The gap is the difference between the observed value and the predicted value.

In sum, the point estimates suggest that there could be a small positive impact of the ISP on the employment rate of elderly men and, on a smaller scale, of elderly women. Consistent with this finding, the results suggest a small negative impact for younger men which could be interpreted as a substitution effect. However, the observed values of the employment rate stay in the confidence interval around the forecasts for the period 2007q2-2008q2 which means that there is not enough statistical power to conclude to a significant impact of the ISP. For younger women, there is no evidence of an impact of the ISP.

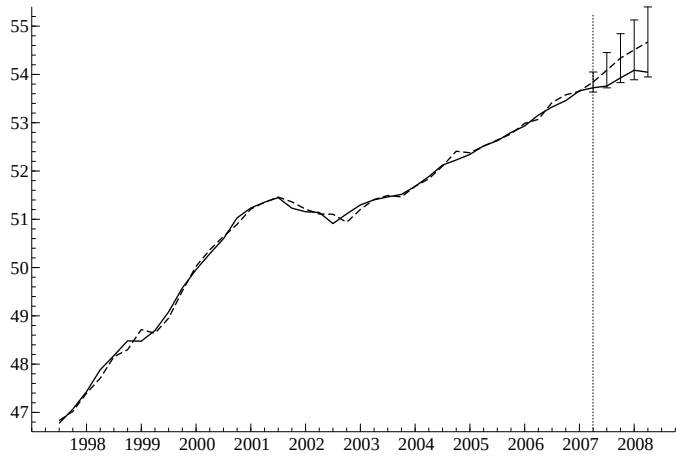
Because one could question the choice of 2007q2 for the starting date of the forecasts, the robustness of the results is tested by beginning the forecasts one year earlier, in 2006q2, just after the first policies of the ISP have been taken. The results of this robustness analysis are depicted in figure 6. For elderly people, the positive gap between the observed value and the forecasts of the employment rate is maintained but appears earlier, at the end of 2005. For younger men, the negative gap is maintained and also begins earlier, between the second and third quarter of 2006. These results suggest that the ISP could have had an impact from the beginning of its introduction. For younger women, there is still no difference between the forecasts and the observed values.



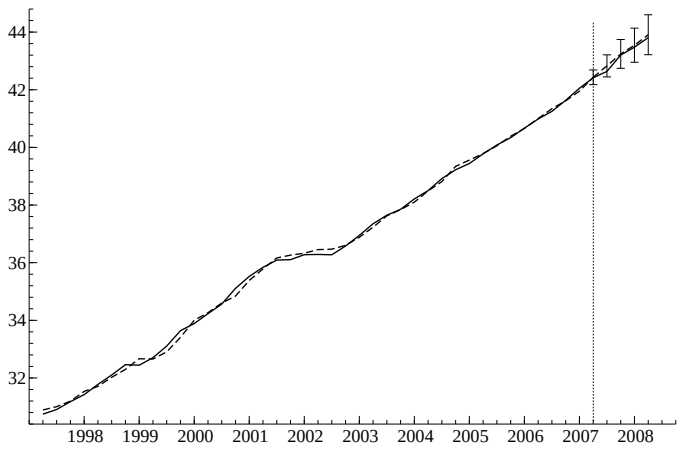
(a) Men aged 50-59 years old



(b) Women aged 50-59 years old

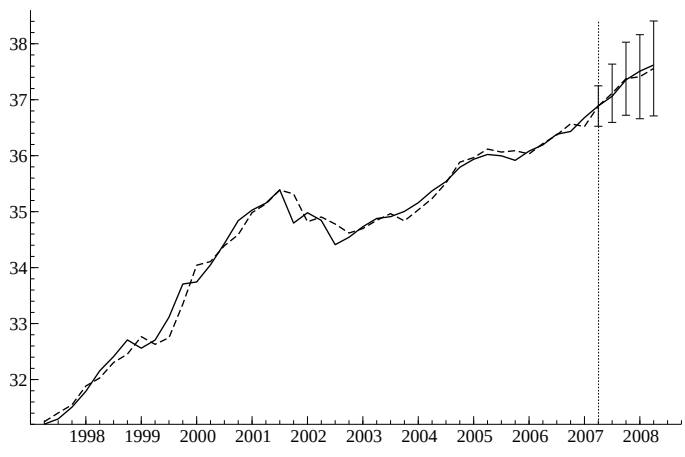


(c) Men aged 35-44 years old



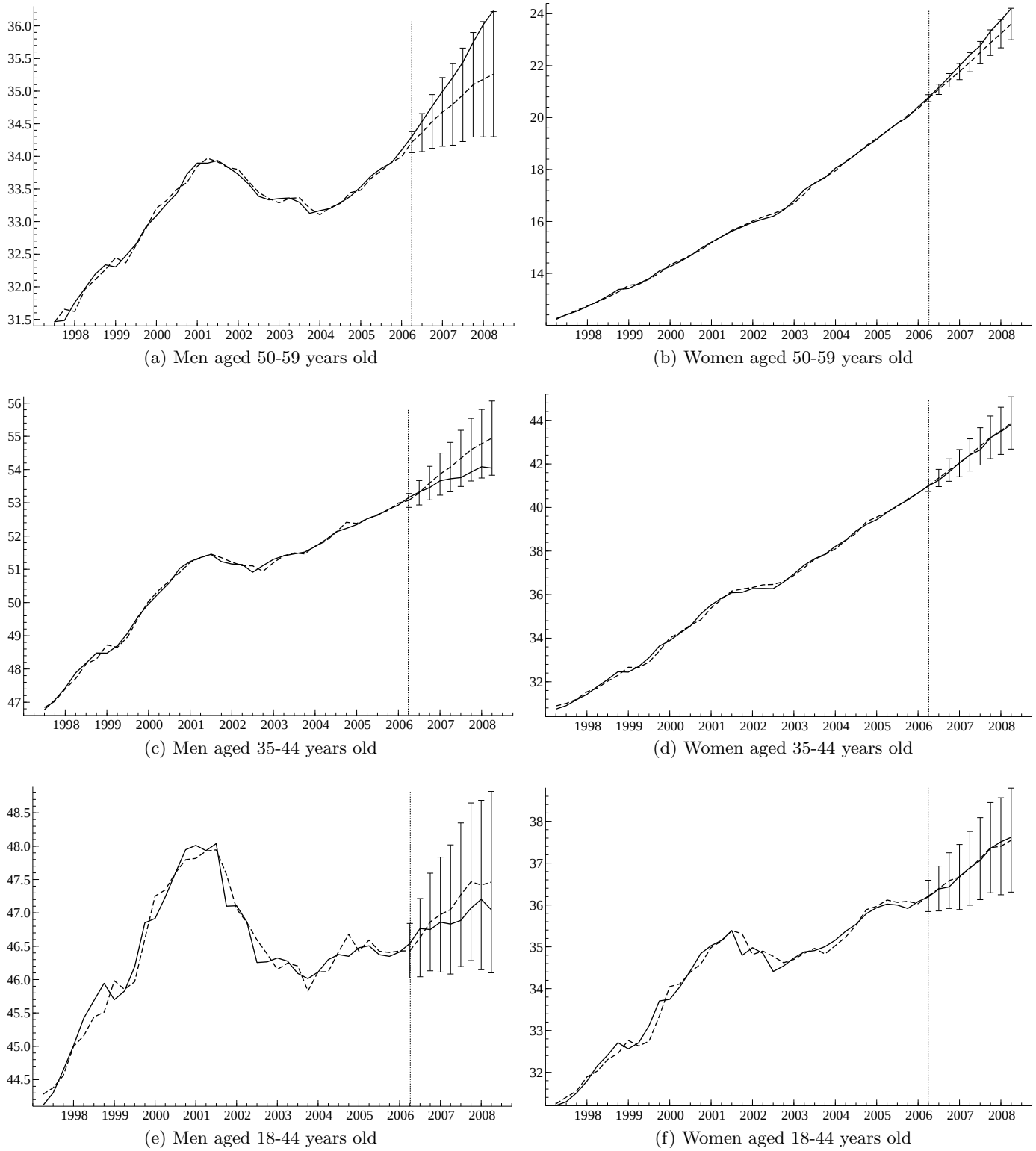
(d) Women aged 35-44 years old

(e) Men aged 18-44 years old



(f) Women aged 18-44 years old

Figure 5: Forecasts of the employment rate for the period 2007q2-2008q2.— observed series - - - forecasts ⊢⊣ 95% CI.



(a) Men aged 50-59 years old

(b) Women aged 50-59 years old

(c) Men aged 35-44 years old

(d) Women aged 35-44 years old

(e) Men aged 18-44 years old

(f) Women aged 18-44 years old

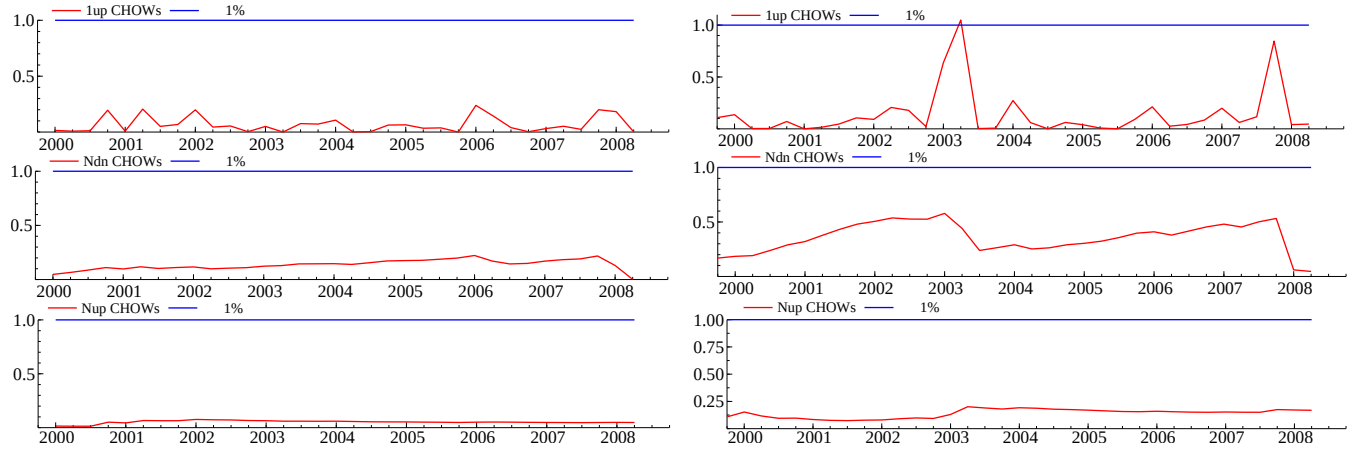
Figure 6: Forecasts of the employment rate for the period 2006q2-2008q2.— observed series - - - forecasts $\vdash\dashv$ 95% CI.

5.2 Results for the recursive analysis and break tests

The number of initial observations for the recursive analysis is set to 10. The results from the three types of recursive Chow tests for each gender/age group are depicted in figure 7. For each panel, the first graph gives the Chow statistic for the 1-step Chow test, the second for the n-down Chow test and the third for the n-up Chow test. The straight line at unity represents the 1% critical values. The forward recursive coefficients are presented in figure 8. At the beginning, the standard errors are large because the coefficients are estimated on a small number of observations.

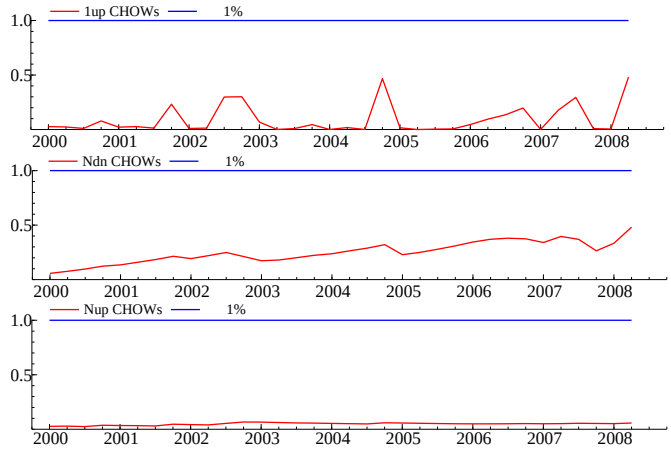
For the period 2007q2-2008q2, none of the Chow tests find a break in the employment rate for any age group and gender, which means that the policies of the ISP did not cause a structural break in the employment rate. Although, for elderly women, there is a spike in the 1-step Chow test statistic in 2007q3 but this it is not big enough to conclude in the presence of a break. For the period before 2007q2, two structural breaks are found, a first one in 2003q1 for elderly women and a second one in 2001q4 for men and women aged 18-44 years old.

For the recursive coefficients, as one could expect from the Chow tests, there is no large coefficient instability over the period 2007q2-2008q2. The non-significant spike in the 1-step Chow test in 2007q3 for women aged 50 to 59 years old is due to an upward spike in the constant of the model and to an upward shift in the coefficient of GDP in t-3. However, those features in the recursive coefficients are also present for women aged 35-44 years old so they are probably not due to the policies of the ISP. Still for elderly women, the break in 2003q1 is due to a decrease in the constant of the model and an increase in the trend. This break is apparently due to the fact that the method used to correct for the change in the measurement of employment introduced by the Belgian administration in 2003q1 is not perfect (cf. section 3). The break in 2001q4 for men and women aged 18-44 years old comes from a large shift in all coefficients of the model. This break could be due to the economic downturn in Belgium at that time. However, one has to be careful because in 2001q4, the coefficients are calculated using only 19 observations which could explain why the economic downturn can have such an impact on the coefficients.

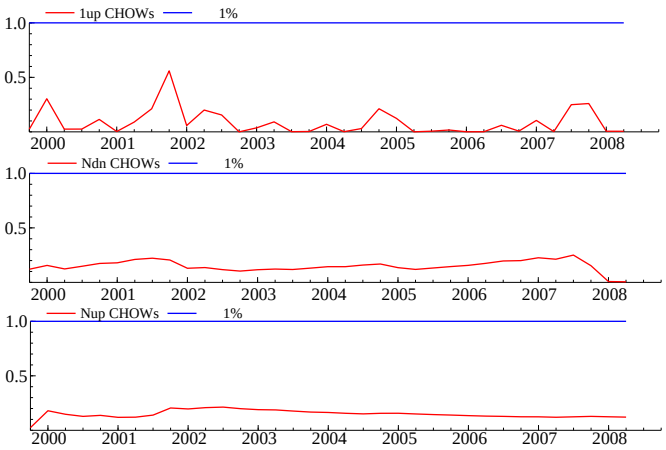


(a) Men aged 50-59 years old

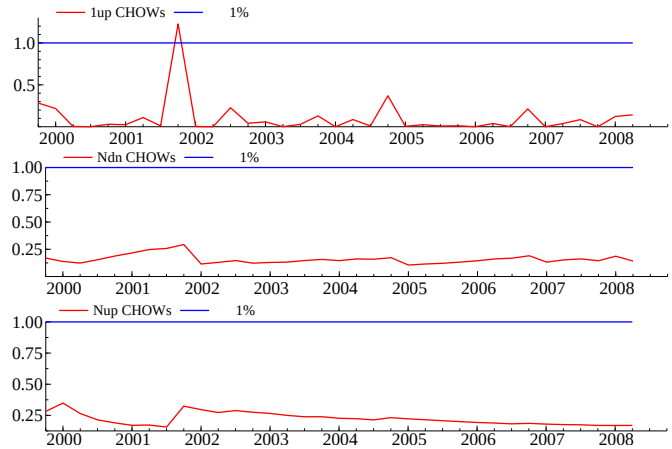
(b) Women aged 50-59 years old



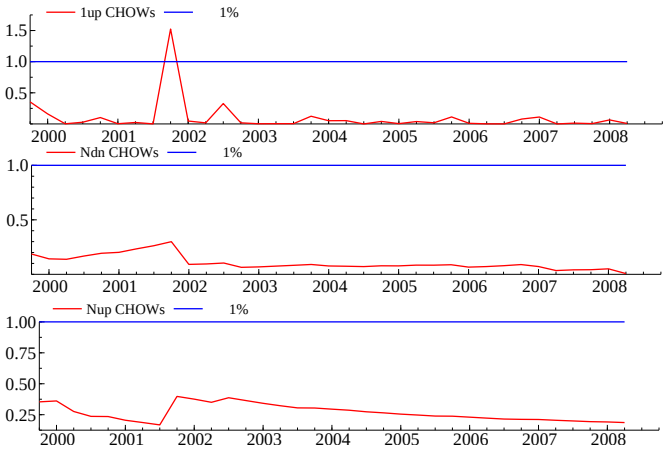
(c) Men aged 35-44 years old



(d) Women aged 35-44 years old

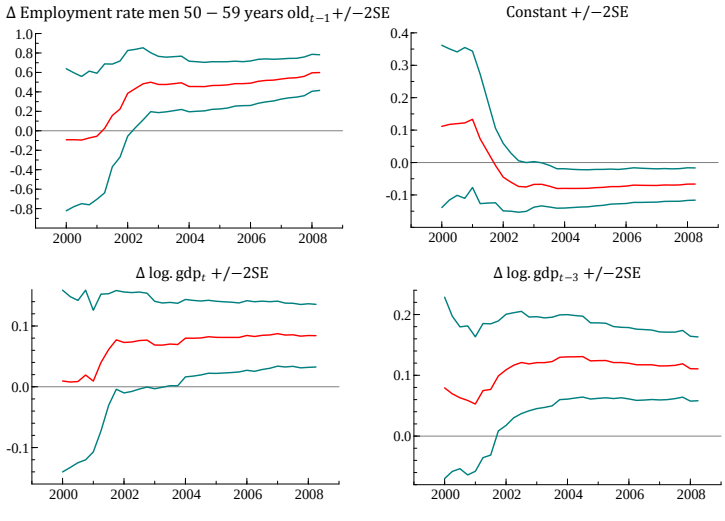


(e) Men aged 18-44 years old

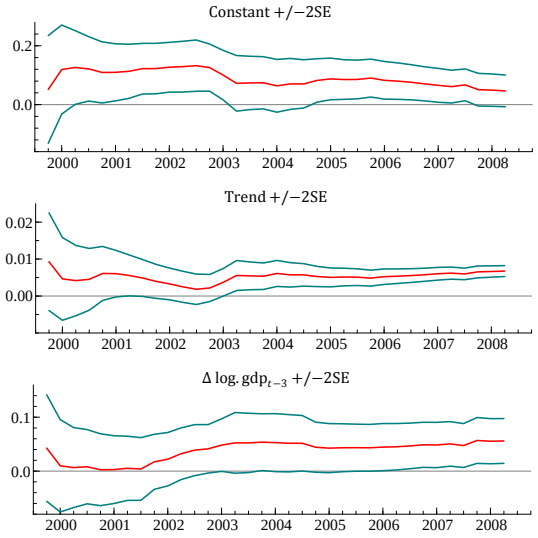


(f) Women aged 18-44 years old

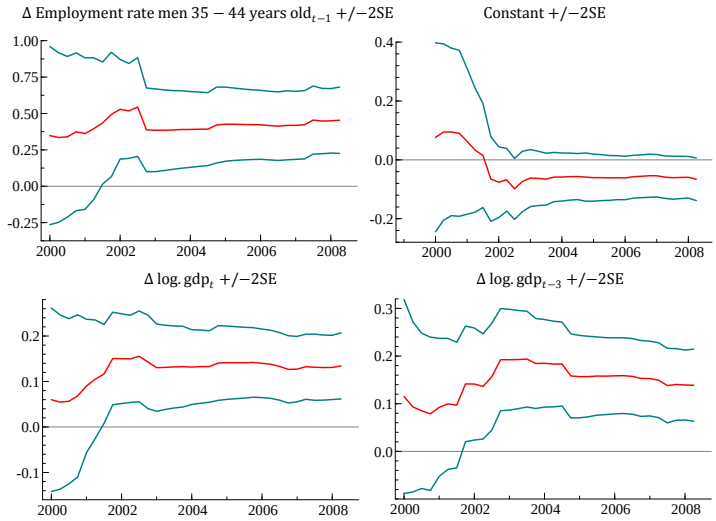
Figure 7: Recursive Chow tests for the models in first difference. 1-step Chow test, n-down Chow test and n-up Chow test.



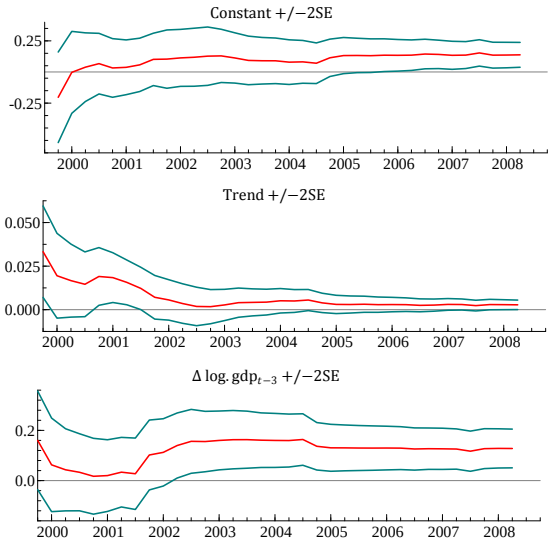
(a) Men aged 50-59 years old



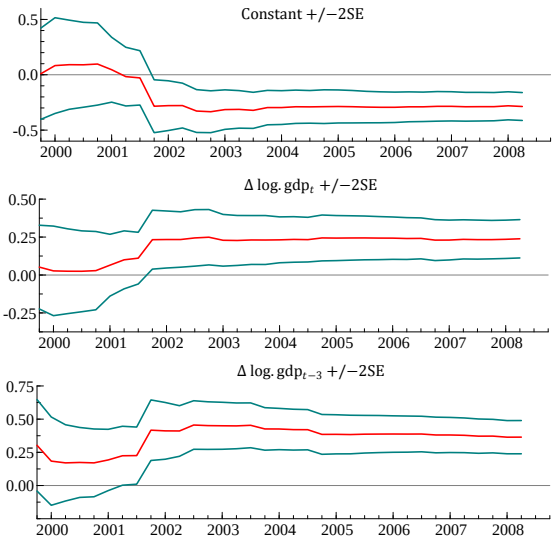
(b) Women aged 50-59 years old



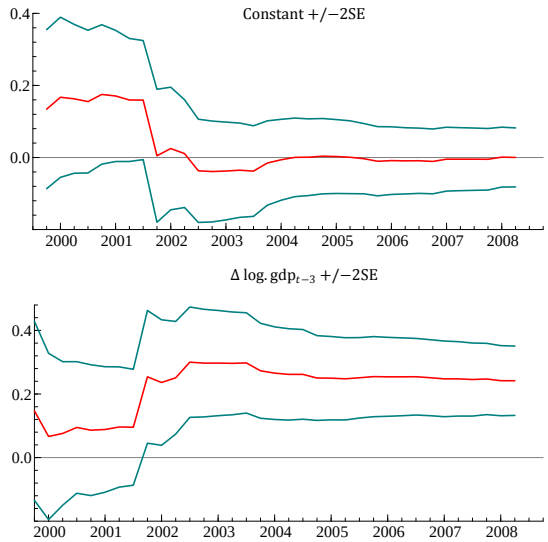
(c) Men aged 35-44 years old



(d) Women aged 35-44 years old



(e) Men aged 18-44 years old



(f) Women aged 18-44 years old

Figure 8: Recursive coefficients for the models in first difference.

5.3 Results taking into account the simultaneity issue

This section investigates whether the results are robust to the simultaneity issue. As discussed in the section about the method, the simultaneity bias arises only when the contemporaneous GDP growth rate appears in the regressors. This is the case for men of all age groups. Two robustness analysis are made. The first one consists in removing the contemporaneous value of GDP and the second takes the mean real GDP of France, Luxembourg, Germany, the Netherlands and the U.K. as a regressor instead of Belgian GDP.

Table 4 gives the OLS estimations of the two methods. The coefficients of the models without GDP at time t are very close to those including the contemporaneous value of GDP growth rate. Using the GDP of neighbouring countries, only the third lag is significant and the point estimates are again not too far from the baseline models. However, the R^2 decreases for all age groups, above all for men aged 35-44 and 18-44 years old.

Table 4: Estimation of the employment rate of men by age group as a function of GDP and GDP of neighbours - OLS

Dep. var.	Δ Employment rate Men					
	50-59 years old		35-44 years old		18-44 years old	
Δ Employment rate$_{t-1}$	0.553***	0.461***	0.387***	0.360**		
	(0.116)	(0.128)	(0.134)	(0.154)		
Δ log[GDP]$_{t-3}$	0.102***		0.142***		0.346***	
	(0.031)		(0.045)		(0.101)[a]	
Δ log[GDP neighbours]$_t$		0.031		0.053		0.115
		(0.033)		(0.046)		(0.099)[a]
Δ log[GDP neighbours]$_{t-3}$		0.132***		0.125**		0.223**
		(0.032)		(0.052)		(0.077)[a]
Constant	-0.015	-0.064*	0.03	-0.010	-0.129**	-0.173*
	(0.022)	(0.032)	(0.03)	(0.047)	(0.078)[a]	(0.092)[a]
Nbr obs.	39	39	39	39	40	40
Adj. R^2	0.583	0.627	0.474	0.418	0.345	0.136
AR 1-3 test F-stat	0.317	0.467	0.115	0.148	0.364	0.490

[a] Heteroscedasticity consistent standard error of White (1980).
*** significant at 1% level, ** 5% level, * 10% level.

The results for the counter factual and the recursive analysis are in the appendix section 7.3. As the R^2 is smaller, the confidence intervals are wider compared to the baseline analysis but the results are maintained. For the counter factual analysis, the point estimates indicate a possible small positive impact of the ISP for the elderly men and a small negative impact for younger men but the results are not statistically significant due to the lack of power. Table 5 compares the observed and predicted values of the employment rate of elderly men. As for the baseline regressions, the observed and

predicted values are very close and the observed values lie inside the 95% and 90% confidence interval. For the recursive analysis, no structural break is found for the period after the introduction of the ISP.

Table 5: Observed versus predicted values of the employment rate and level of employment for men aged 50-59 year old in in 2008q2

		Observed value	Predicted value (s.e.)	Gap in % points[a]
Regressions without the first lag of GDP				
Employment rate		36.23%	35.80%	0.43
	95% CI		[35.06 ; 36.53]	
	90% CI		[35.18 ; 36.41]	
Employment		258,178	255,087	3,091
Regressions using GDP of neigbours				
Employment rate		36.23%	35.84%	0.39
	95% CI		[35,19 ; 36,49]	
	90% CI		[35,3 ; 36,38]	
Employment		258,178	255,379	2,799

[a] The gap is the difference between the observed value and the predicted value.

6 Conclusion

The aim of this paper is to evaluate the macro-economic impact of the active ageing policies introduced by the Intergenerational Solidarity Pact (ISP) in December 2005 on the elderly employment rate by gender in Belgium. The impact of the reforms introduced by the ISP are evaluated over the period 2007q2-2008q2. The evaluation literature of labour market policies often take a micro-econometric approach. The advantage of the macro-econometric approach over the micro one is that it captures general-equilibrium effects. The motivation of this paper is the scarcity of empirical works on general equilibrium effects of labour market policies targeted at the elderly, especially in Belgium, although its importance for policy making.

The effect of the policies of the ISP on the employment rate of the elderly are expected to be small. Indeed, the main policy is a permanent wage subsidy and although it targets all workers aged above 50, it's amount is low compared to the cost of labour. Moreover, evidence points to a low wage elasticity of labour demand and supply. Because of these two elements, the policy is not expected to have a large impact on elderly employment.

Two identification methods are used to evaluate the impact of the ISP both relying on a macro-econometric model which explains the evolution of the elderly employment rate by its own dynamics and by the economic growth. The main hypothesis of the evaluation methods is that, in the absence of public policies, changes in the employment rate are closely linked to GDP growth rate. In the first method, a counter factual employment rate is build using forecasts of the macro-econometric model for the period

2007q2-2008q2. The counter factual employment rate is interpreted as the value the employment rate would have taken in the absence of the policies of the ISP. The actual elderly employment rate is then compared to the forecasts (counter factual). The second method tests for the presence of structural breaks in the macro-econometric model after the introduction of the ISP using recursive Chow tests. The two methods are also used to investigate possible substitution or complementarity effects of the ISP on the employment rate of younger age groups. In principle, significant differences between forecasts and observed values and the presence of breaks could also be attributed to changes in the determinants of the employment rate other than the GDP and the policies of the ISP during that period. However, to my knowledge, there is no important shocks other than the ISP that could have had a major influence on the elderly employment rate at that time.

As expected, the results from the first method show that if there is a positive impact of the ISP, it is too small to be statistically significant. For elderly men, the impact is only 0.52 (s.e. 0.31) percentage point, rising the employment rate from 35.7% to 36.2%, while it is even lower for elderly women, 0.26 (s.e. 0.18) percentage point, increasing the employment rate from 23.9% to 24.2%. Moreover, the results suggest a small negative effect on younger men's employment rate. For men aged 35-44 years old, the impact is -0.63 (s.e. 0.36) percentage point, decreasing the employment rate from 54.67% to 54.04%. For younger women, there is no sign of an impact of the ISP. Finding evidence for a substitution effect between younger and older labour force only for men is consistent with the finding of a stronger impact on elderly men compared to elderly women. However, there is not enough statistical power to conclude to any significant impacts. A relatively low sample size and goodness of fit of the models may explain the lack of statistical power to detect small expected effects. Using the second method, no structural break is found in the relationship between employment and GDP. I can conclude that the ISP did not induce brutal changes in the employment rate but could have had a gradual positive but small impact.

7 Appendix

7.1 Augmented Dickey Füller (ADF) test

The ADF test the null hypothesis of unit root process against the alternative of stationarity. When y_t follows an $AR(p)$

$$y_t = \alpha_0 + \alpha_1 t + \sum_{p}^{i=1} \rho_i y_{t-i} + \epsilon_t$$

The ADF test is as follows

$$\Delta y_t = \alpha_0 + \alpha_1 t + (\rho - 1) y_{t-1} + \sum_{p-1}^{i=1} \gamma_i \Delta y_{t-i} + u_t$$

where the null hypothesis is $\rho = 1$ and the alternative $\rho < 1$. When variables have a deterministic trend, it is added to the regression.

The table 6 below reports the ADF tests for the variables of the model.

Table 6: ADF tests - Lag selection via t-test

	Deterministic components	p-1	Coefficient ρ	t-statistic	Critical values 1%	5%	10%
Employment rate men 50-59 years old							
level	constant+trend	1	0.95606	-2.043	-4.141	-3.496	-3.178
Δ	constant	0	0.72435*	-2.882	-2.927	-3.574	-2.598
Employment rate women 50-59 years old							
level	constant+trend	1	0.98851	-0.9024	-4.141	-3.496	-3.178
Δ	constant+trend	0	0.28545***	-5.254	-4.141	-3.496	-3.178
Employment rate men 35-44 years old							
level	constant+trend	1	0.96904	-1.665	-4.141	-3.496	-3.178
Δ	constant	0	0.79011	-2.461	-2.927	-3.574	-2.598
Employment rate women 35-44 years old							
level	constant+trend	1	0.95464	-1.008	-4.141	-3.496	-3.178
Δ	constant	0	0.55383***	-3.878	-2.927	-3.574	-2.598
Employment rate men 18-44 years old							
level	constant	1	0.93034	-2.060	-2.927	-3.574	-2.598
Δ	constant	0	0.52493***	-4.026	-2.927	-3.574	-2.598
Employment rate women 18-44 years old							
level	constant+trend	1	0.92296	-1.697	-4.141	-3.496	-3.178
Δ	constant	0	0.32697***	-5.118	-2.927	-3.574	-2.598
LGDP							
level	constant+trend	1	0.90154	-2.681	-4.141	-3.496	-3.178
Δ	constant	1	0.4768***	-4.332	-4.141	-3.496	-3.178
Sample 1997q1-2010q4, 54 observations							

* reject the null of unit root at 10% level, ** 5% level, *** 1% level.

For men aged 35-44 years old, the unit root in the employment rate cannot be rejected at 10% confidence level. However, ADF unit root tests are not powerful tests in small sample and the t-statistic is quite close to the 10% critical value. In the analysis, the employment of men aged 35-44 years old rate is considered as stationary.

7.2 Cointegration tests

The formal definition of cointegration given by Enders (1995) is the following:

The component of a vector $x_t = (x_{1t}, x_{2t}, ..., x_{nt})$ are said to be *cointegrated of order b, d,* denoted $x_t \sim CI(d, b)$ if

1. All components of x_t are integrated of order d.

2. There exists a vector $\beta = (\beta_1, \beta_2, ..., \beta_n)$ such that linear combination $\beta x_t = (\beta_1 x_{1t}, \beta_2 x_{2t}, ..., \beta_n x_{nt})$ is integrated of order $(d - b)$, where $b > 0$.

The vector β is called the cointegration vector.

In this paper, the vector x_t is composed of two variables, the employment rate by age group and gender and GDP, that are integrated of order 1 and I test if there is a linear combination βx_t that is integrated of order 0.

7.2.1 Engle-Granger cointegration test

The Engle-Granger cointegration test can only detect one cointegration vector and is then better fitted to test cointegration between no more than two variables denoted y_t and z_t in the following.

The Engle and Granger (1987) methodology follows two steps (Enders, 1995):

- STEP 1: Pretest the variables for their order of integration. Cointegration necessitates that the variables be integrated of the same order. If variable are integrated of the same order proceed in step 2.

- STEP 2: Estimate the long-run equilibrium relationship. If the results of step 1 indicates that both y_t and z_t are $I(1)$, the next step is to estimate the long-run relationship as follows:

$$y_t = \gamma_0 + \beta z_t + e_t$$

where a deterministic trend can be added if necessary. In order to determine if the variables are cointegrated, test for the presence of unit-root in the $\hat{e}_t$. The $\hat{e}_t$ is the series of the estimated residuals of the long-run relationship. If the presence of a unit root in the $\hat{e}_t$ is rejected, then y_t and z_t are cointegrated of order(1, 1). The ADF test on the estimated residuals takes the following form:

$$\Delta \hat{e}_t = a_1 \hat{e}_{t-1} + \epsilon_t$$

There is no constant term as the estimated residual have zero mean by construction. The ADF t-statistics are valid although the ADF critical value are not. One has to correct the critical values for the fact that the $\hat{e}_t$ sequences are generated from a regression equation. The proper critical values are provided in MacKinnon (1991).

In this analysis, y_t is the employment rate by age group and gender and z_t is GDP. The first step was done in the previous section where it was found that all variables are $I(1)$.

For the second step, the long-term relationship is estimated with a deterministic trend because GDP and most employment rates by age and gender have one.

In table 7, where the estimated a_1 coefficient are reported, the t-statistic and the critical values provided by MacKinnon (1991).

Table 7: Engle Granger cointegration tests between employment rate by age group and gender and GDP - Lag selection for ADF via t-test

	nbr Lags	$a_1 + 1$[1]	t-statistic
Men 50-59 years old	1	0.95584	-2.094
Women 50-59 years old	1	0.92541	-2.104
Men 35-44 years old	1	0.75599**	-3.456
Women 35-44 years old	1	0.81152	-2.725
Men 18-44 years old	1	0.73765**	-4.038
Women 18-44 years old	1	0.76879	-3.242
Sample 1997q1-2010q4, 54 observations			

The critical value at 10% is 3.63066, at 5% is -3,9605 and at 1% is -4,62588. * reject the nul of no cointegration at 10% level, ** 5% level,*** 1% level.
[1] $a_1 + 1$ is the coefficient of $\hat{e}_{t-1}$ in $\hat{e}_t = (a_1 + 1)\hat{e}_{t-1} + \epsilon_t$.

A cointegration relationship is found for men aged 35 to 44 and 18 to 44 years old. If the estimated residuals $\hat{e}_t$ are used as the deviation from the long-term relationship in the ECM and estimate the following "near" VAR:

$$\Delta y_t = \gamma_1 + \alpha_y \hat{e}_{t-1} + \sum_{i=1}^{p} \Pi_{1,1} \Delta y_{t-i} + \sum_{i=1}^{p} \Pi_{1,2} \Delta z_{t-i} + \epsilon_{yt}$$

$$\Delta z_t = \gamma_2 + \alpha_z \hat{e}_{t-1} + \sum_{i=1}^{p} \Pi_{2,1} \Delta y_{t-i} + \sum_{i=1}^{p} \Pi_{2,2} \Delta z_{t-i} + \epsilon_{zt}$$

The number of lags for the VAR is selected based on AIC criterion. The results given below show that for the equilibrium relationship between GDP and the employment rate of men aged 35-44, only GDP adjust to the long-term equilibrium. On the contrary, for the age group 18-44 years old, it's the employment rate that adjust and not GDP.

Results of the VAR for the employment rate of men aged 35-44 years old

$\hat{e}_t$=	5.499	+emp$_t$	-0.536 log[gdp]$_t$	-0.001 t
Δemp$_t$=	-0.011	-0.013 $\hat{e}_{t-1}$	+0.623 Δemp$_{t-1}$	+0.102Δlog[gdp]$_{t-1}$ +$\epsilon_{emp,t}$
	(0.023)	(0.039)	(0.093)	(0.035)
Δlog[gdp]$_t$=	-0.111	+0.363$\hat{e}_{t-1}$	-0.294Δemp$_{t-1}$	+0.76Δlog[gdp]$_{t-1}$ +$\epsilon_{gdp,t}$
	(0.092)	(0.153)	(0.367)	(0.137)

Results of the VAR for the employment rate of men aged 18-44 years old

$\hat{e}_t$=	5.461	-0.532 $\log[\text{gdp}]_t$	+emp_t	-0.002 t
Δemp_t=	-0.058	-0.165$\hat{e}_{t-1}$	+0.253Δemp_{t-1}	+ 0.139$\Delta\log[\text{gdp}]_{t-1}$ +$\epsilon_{e,t}$
	(0.042)	(0.066)	(0.121)	(0.063)
$\Delta\log[\text{gdp}]_t$=	-0.1	+0.206$\hat{e}_{t-1}$	-0.336Δ emp_{t-1}	+ 0.737$\Delta\log[\text{gdp}]_{t-1}$ +$\epsilon_{gdp,t}$
	(0.099)	(0.155)	(0.283)	(0.145)

7.2.2 Johansen cointegration test

The starting point of this methodology is the vector autoregression (VAR) of order p given by

$$x_t = \gamma + A_1 x_{t-1} + ... + A_p x_{t-p} + \epsilon_t \tag{2}$$

Where x_t is a $n \times 1$ vector of variables that are integrated of order one and ϵ_t is an $n \times 1$ vector of innovations. In this case the vector is x_t is a 2×1 vector given by

$$y_t = \left| \begin{array}{c} \text{emp.}_{g,a,t)} \\ \log[\text{gdp}]_t \end{array} \right|$$

where $\text{emp}_{g,a,t}$ stands for the employment rate and the subscripts g corresponds to gender, a to age and t to time. $\log[\text{gdp}]_t$ is GDP at time t.

The VAR in equation (2) can be re-written as

$$\Delta y_t = \gamma + \Pi y_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta y_{t-i} + \epsilon_t$$

where $\Pi = \sum_{i=1}^{p} A_i - I$ and $\Gamma_i = -\sum_{j=1+1}^{p} A_j$

If the coefficient matrix Π has reduced rank $r < n$ (2 in this paper), then there exist $n \times r$ matrices α and β each with rank r such that $\Pi = \alpha\beta'$ and βy_t is stationary. The rank r corresponds to the number of cointegrating relationships, the elements of α are known as the speed of adjustment parameters in the vector error correction model and each column of β is a cointegrating vector. Johansen proposes two different likelihood ratio tests of the reduced rank of the Π matrix: the trace test and maximum eigenvalue test. The statistics are given by

$$J_{trace} = -T \sum_{i=r+1}^{n} ln(1 - \widehat{\lambda_i}) \text{ for the trace test and}$$
$$J_{max} = -Tln(1 - \widehat{\lambda_{r+1}}) \text{ for the max test}$$

Here T is the sample size and λ_i is the $i : th$ largest eigenvalue. The trace test tests the null hypothesis that the number of distinct cointegrating vectors is less than or equal to r against a general alternative. The maximum eigenvalue test, on the other hand, tests the null hypothesis of r cointegrating vectors against the alternative hypothesis of $r + 1$ cointegrating vectors.

The table 8 shows the result of the trace and max tests for each gender and age group. The choice of the number of lags in the VAR was made using AIC, SC and HQ tests.

Table 8: Johansen cointegration tests between employment rate by gender and age group and GDP

nbr lags	rank	eigenvalue	trace test [Prob]	max test [Prob]	trace test (T-nm)	max test (T-nm)
Men 50-59 years old						
2	0		8.79 [0.959]	6.24 [0.933]	8.14 [0.974]	5.78 [0.953]
	1	0.10912	2.55 [0.911]	2.55 [0.912]	2.36 [0.928]	2.36 [0.929]
Women 50-59 years old						
2	0		32.33 [0.006]**	22.40 [0.015]*	29.94 [0.013]*	20.74 [0.029]*
	1	0.33956	2.55 [0.911]	2.55 [0.912]	2.36 [0.928]	2.36 [0.929]
Men 35-44 years old						
2	0		26.65 [0.038]*	18.20 [0.072]	24.68 [0.068]	16.86 [0.113]
	1	0.28618	8.44 [0.224]	8.44 [0.224]	7.82 [0.275]	7.82 [0.275]
Women 35-44 years old						
2	0		20.28 [0.216]	17.06 [0.106]	18.78 [0.300]	15.80 [0.158]
	1	0.27087	3.22 [0.840]	3.22 [0.842]	2.98 [0.867]	2.98 [0.869]
Men 18-44 years old						
2	0		29.46 [0.015]*	23.24 [0.011]*	27.28 [0.031]*	21.52 [0.021]*
	1	0.34969	6.22 [0.443]	6.22 [0.444]	5.76 [0.501]	5.76 [0.502]
Women 18-44 years old						
	0		25.36 [0.056]	20.86 [0.027]*	23.48 [0.096]	19.31 [0.049]*
	1	0.32037	4.51 [0.672]	4.51 [0.673]	4.17 [0.718]	4.17 [0.719]
Sample 1997q1-2010q4, 54 observations						

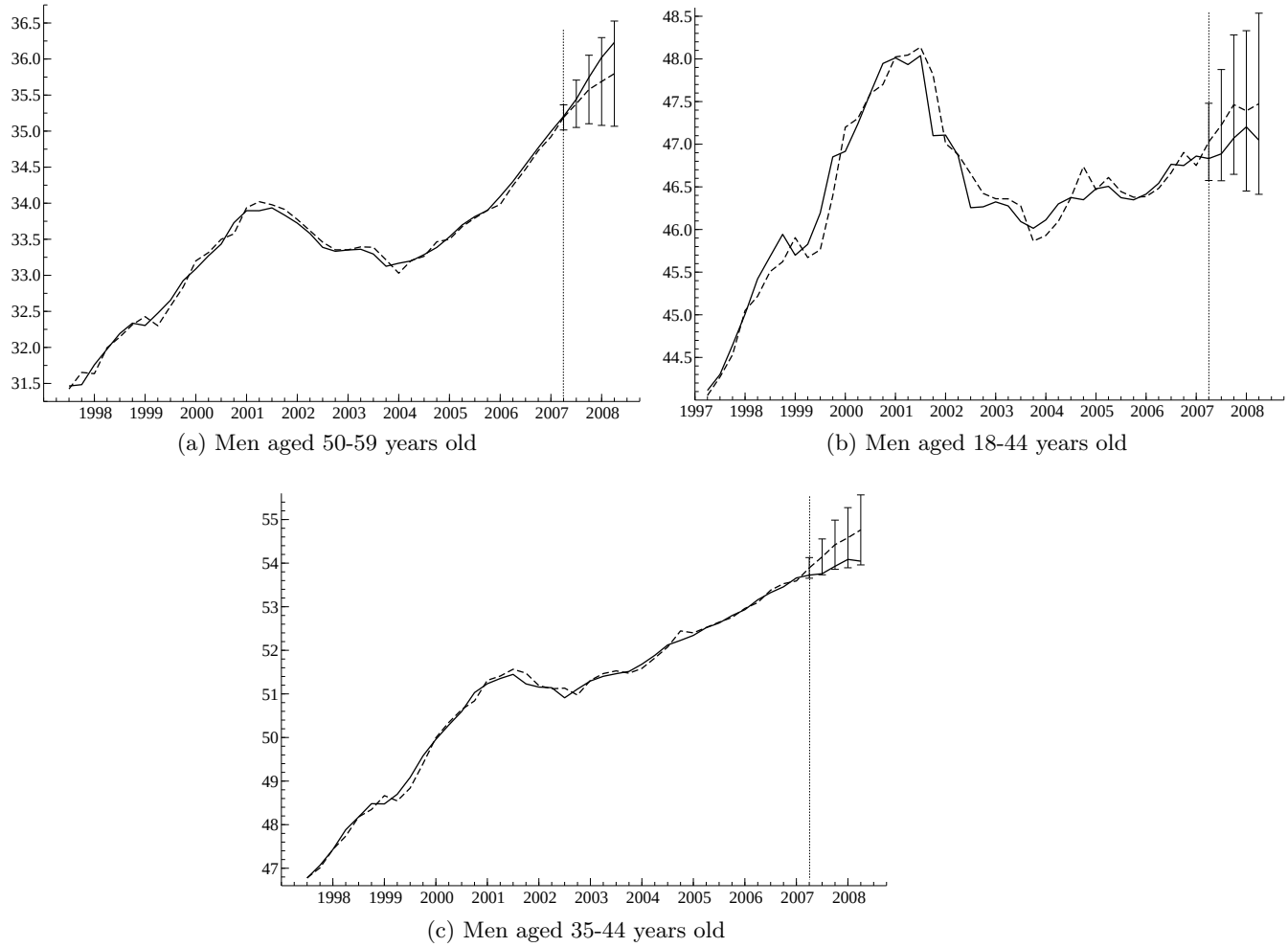
* reject the null of no cointegration at 5% level,** at 1% level

According to the Johansen test, there is a cointegration relationship between the employment rate of women aged 50 to 59 years old and GDP as well as between the employment rate of men aged 18 to 44 years old and GDP. Compare to the Engle-Granger test, no cointegration is found for men aged between 35 and 44 years old but there is one for elderly women. The Π matrix is reported in table 9 below. For the women of the age group 50-59 years old, there is no convergence to the long term relationship as the alpha and the beta are not of opposite signs. So the cointegration relationship only holds for men aged 18-44 years old.

Table 9: the Π matrix from Johansen cointegration tests

	beta	alpha
Women 50-59 years old		
Employment rate	1	-0.001
Log. GDP	-15.13	-0.001
Men 18-44 years old		
Employment rate	1	-0.208
Log. GDP	-0.639	0.22
Trend	0.003	-
Sample 1997q1-2010q4, 54 observations		

7.3 Robustness analysis for simultaneity issue

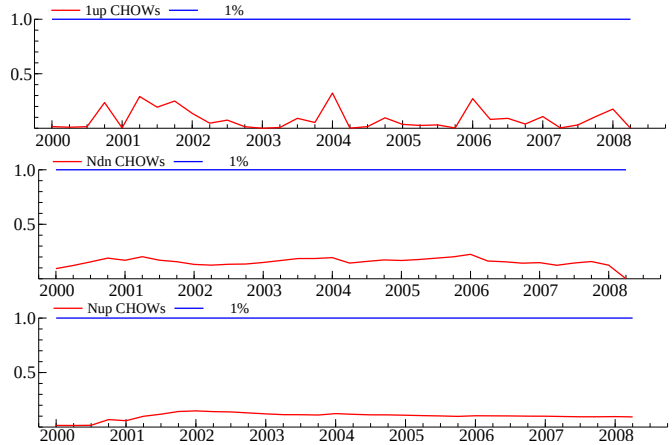


(a) Men aged 50-59 years old

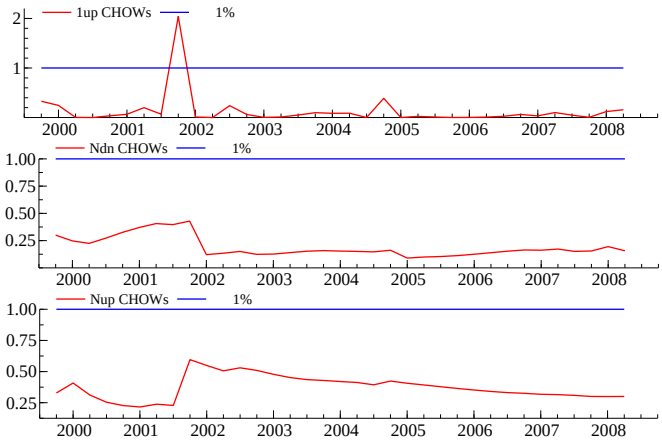
(b) Men aged 18-44 years old

(c) Men aged 35-44 years old

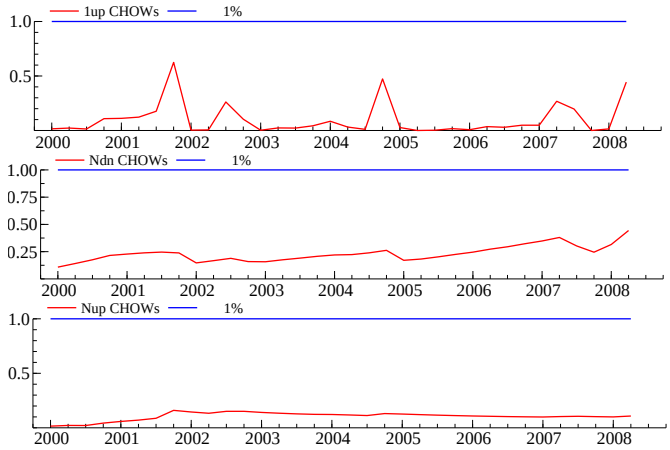
Figure 9: Forecasts of the employment rate for the period 2007q2-2008q2 robust to simultaneity.— observed series - - - forecasts ⊢⊣ 95% CI.



(a) Men aged 50-59 years old

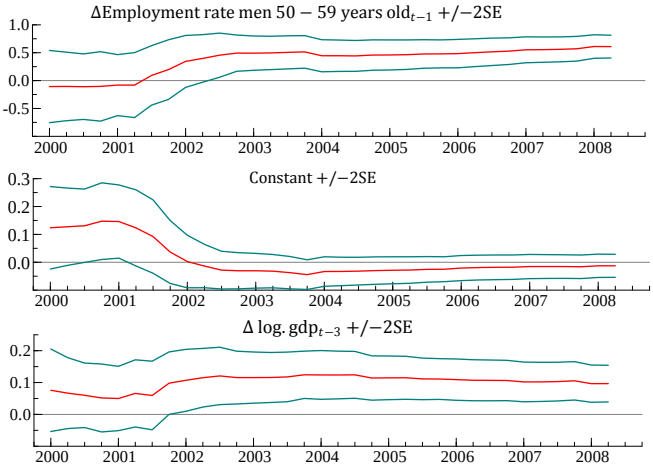


(b) Men aged 18-44 years old

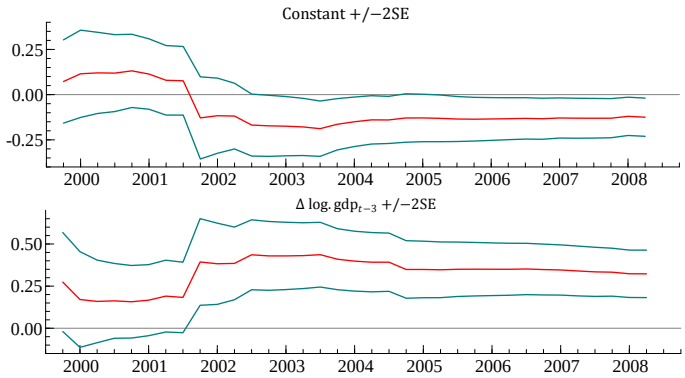


(c) Men aged 35-44 years old

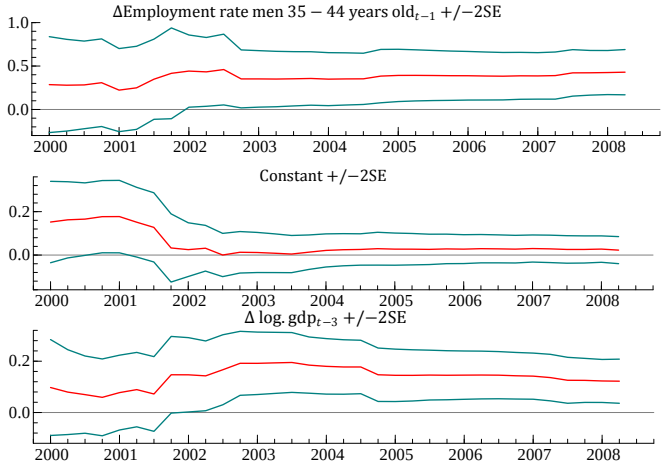
Figure 10: Recursive Chow tests for the models robust to simultaneity. 1-step Chow test, n-down Chow test and n-up Chow test.



(a) Men aged 50-59 years old

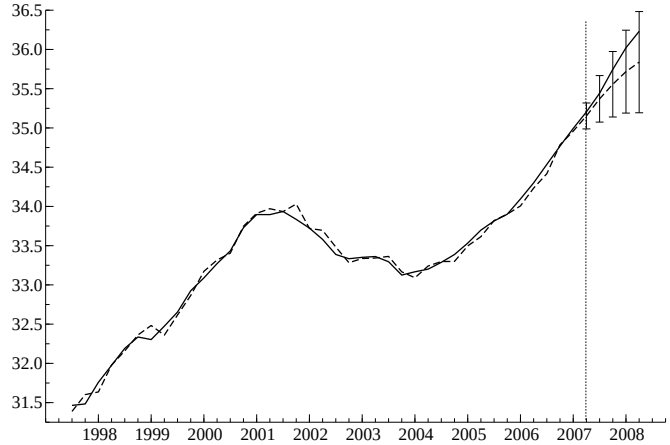


(b) Men aged 18-44 years old

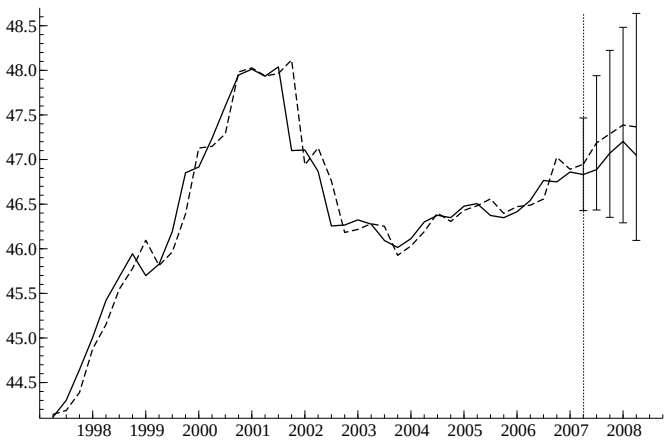


(c) Men aged 35-44 years old

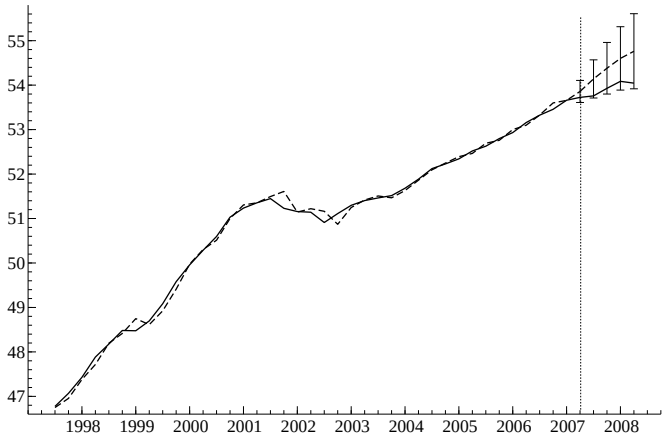
Figure 11: Recursive coefficients for the models robust to simultaneity.



(a) Men aged 50-59 years old

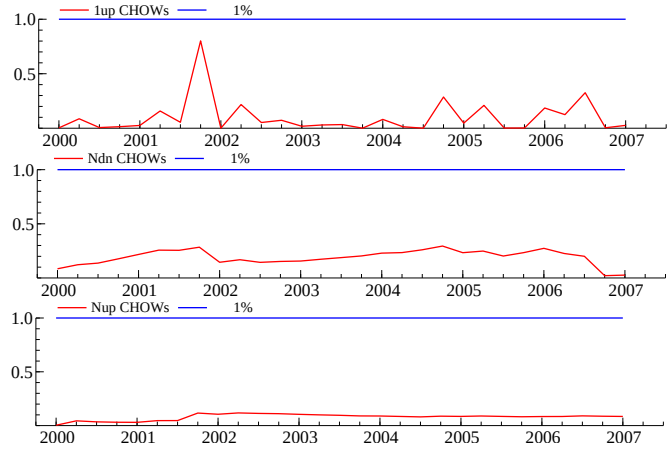


(b) Men aged 18-44 years old

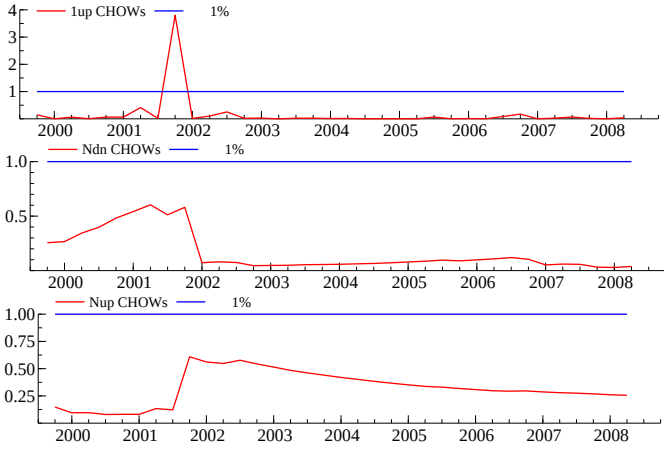


(c) Men aged 35-44 years old

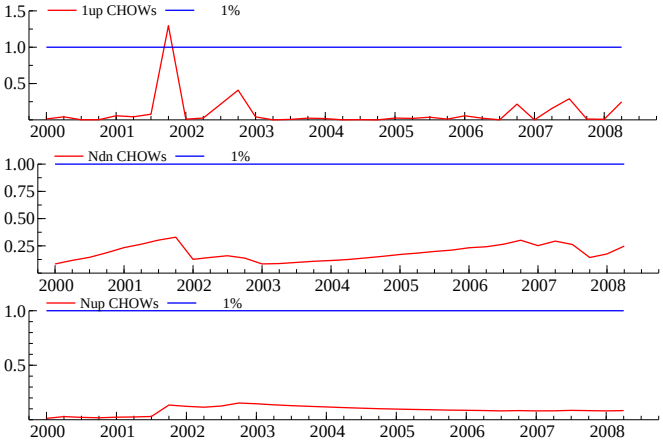
Figure 12: Forecasts of the employment rate for the period 2007q2-2008q2 using GDP of neighbours.— observed series - - - forecasts ⊢⊣ 95% CI.



(a) Men aged 50-59 years old

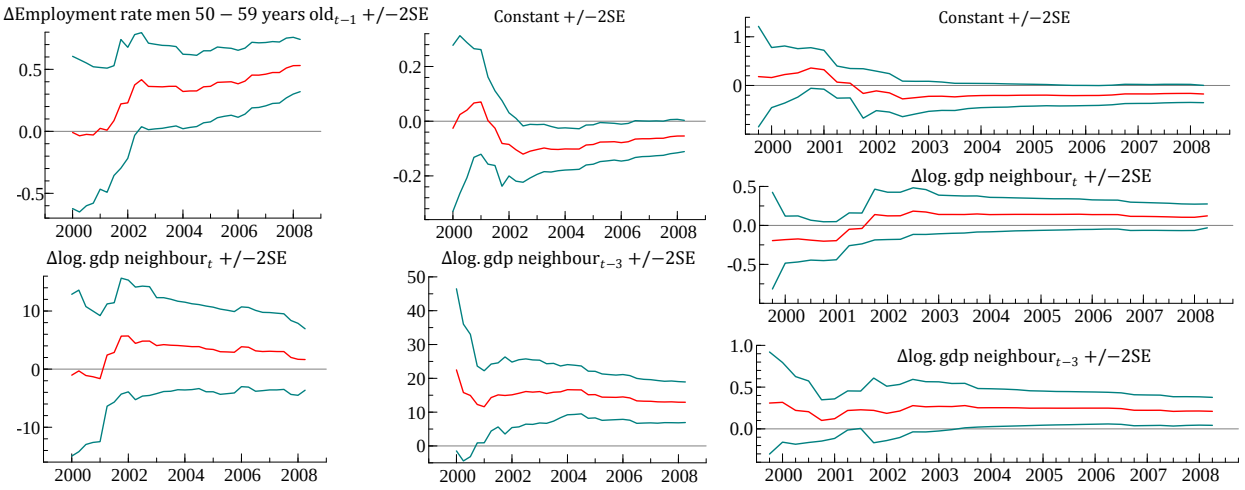


(b) Men aged 18-44 years old



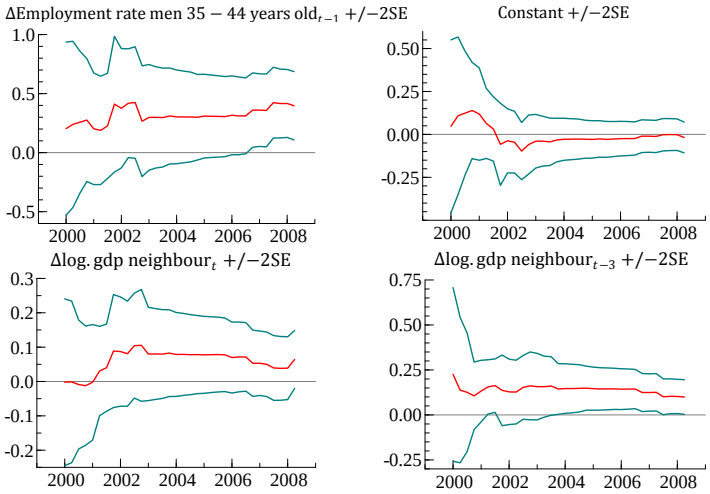
(c) Men aged 35-44 years old

Figure 13: Recursive Chow tests for the models using GDP of neighbours. 1-step Chow test, n-down Chow test and n-up Chow test.



(a) Men aged 50-59 years old

(b) Men aged 18-44 years old



(c) Men aged 35-44 years old

Figure 14: Recursive coefficients for the models using GDP of neighbours.

References

Bassilière, D., Bossier, F., Lebrun, I. and Stockman, P. (2007). *Le programme national de réforme de la Belgique. Effets macroéconomiques de réductions de charges sur le travail*. WORKING PAPER 11-07, Bureau fédéral du Plan.

Bodart, V., Ledent, P. and Shadman-Metha, F. (2009). *An Employment Equation for Belgium*. Discussion Papers (IRES - Institut de Recherches Economiques et Sociales) 2009016, Université catholique de Louvain, Institut de Recherches Economiques et Sociales (IRES).

Bryan, M. L. (2007). Free to choose? Differences in the hours determination of constrained and unconstrained workers. *Oxford Economic Papers*, **59** (2), 226–252.

Calmfors, L. and Skedinger, P. (1995). Does active labour-market policy increase employment? Theoretical considerations and some empirical evidence from Sweden. *Oxford Review of Economic Policy*, **11** (1), 91–109.

de Koning, J. (1993). Measuring the placement effects of two wage-subsidy schemes for the long-term unemployed. *Empirical Economics*, **18** (3), 447–468.

Decicca, C. (2007). Hours flexibility and retirement. *Economic Inquiry*, **45** (2), 251–267.

Devisscher, S. and Sanders, D. (2007). *Modernising Social Policy for the New Life Course*, OECD 2007, chap. Ageing and Life-course Issues: The Case of the Career Break Scheme (Belgium) and the Life-course Regulation (Netherlands), pp. 117–132.

Dmitrijeva, J. and Hazans, M. (2007). A stock-flow matching approach to evaluation of public training programme in a high unemployment environment. *Labour*, **21** (3), 503–540.

Enders, W. (1995). *Applied Econometric Times Series*. Wiley Series in Probability and Mathematical Statistics.

Engle, R. F. and Granger, C. W. J. (1987). Cointegration and error-correction: Representation, estimation, and testing. *Econometrica*, **55** (2), 251–276.

Euwals, R. (2001). Female labour supply, flexibility of working hours, and job mobility. *Economic Journal*, **111** (471), C120–34.

Fahr, R. and Sunde, U. (2009). Did the Hartz reform speed-up the matching process? a macro- evaluation using empirical matching functions. *German Economic Review*, **10** (3), 284–316.

Gustman, A. L. and Steinmeier, T. L. (1983). *Minimum Hours Constraints and Retirement Behavior*. NBER Working Papers 0940, National Bureau of Economic Research, Inc.

Hertweck, M. S. and Sigrist, O. (2013). *The Aggregate Effects of the Hartz Reforms in Germany*. Working papers 2013/01, Faculty of Business and Economics - University of Basel.

Hujer, R., Blien, U., Caliendo, M. and Zeiss, C. (2002). *Macroeconometric Evaluation of Active Labour Market Policies in Germany - A Dynamic Panel Approach Using Regional Data*. IZA Discussion Papers 616, Institute for the Study of Labor (IZA).

Jackman, R., Pissarides, C. A. and Savouri, S. (1990). Labour market policies and unemployment in the OECD. *Economic policy*, **5**, 450–490.

Johansen, S. (1988). Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control*, **12**, 231254.

Kahn, S. B. and Lang, K. (1995). The causes of hours constraints: Evidence from Canada. *Canadian Journal of Economics*, **28** (4a), 914–28.

Klinger, S. and Rothe, T. (2012). The impact of labour market reforms and economic performance on the matching of short-term and long-term unemployed. *Scottish Journal of Political Economy*, **59** (1), 90–114.

Kopits, G. F. (1978). Wage subsidies and employment: An analysis of the French experience. *Staff Papers - International Monetary Fund*, **25** (3), 494–527.

Layard, R., Nickell, S. and Jackman, R. (1991). *Unemployment: Macroeconomic Performance and the Labour Market*. No. 9780198284345 in OUP Catalogue, Oxford University Press.

Lehmann, H. (1993). *The Effectiveness of the Restart Programme and Enterprise Allowance Scheme*. CEP Discussion Papers dp0139, Centre for Economic Performance, LSE.

MacKinnon, J. G. (1991). Critical values for co-integration tests. In R. F.Engle and C. Granger (eds.), *Long-run economic relationship*, Oxford University Press, pp. 267–276.

Maes, M. (2010). *Le bonus de pension : un cadeau empoisonné*. Discussion Papers (IRES - Institut de Recherches Economiques et Sociales) 79, Université catholique de Louvain, Institut de Recherches Economiques et Sociales (IRES).

Mourre, G. (2006). Did the pattern of aggregate employment growth change in the euro area in the late 1990s? *Applied Economics*, **38** (15), 1783–1807.

Nielsen, H. B. (2004). Cointegration analysis in the presence of outliers. *Econometrics Journal*, **7** (1), 249–271.

OECD (2011). Pensions at a glance 2011. *OECD Publishing*.

Piekkola, H. S. E. (2004). *Labour Supply and Demand of Risk Groups: Young and Old*. Early intervention as an employment policy method, EU project/VS/2003/0678.

Stewart, M. B. and Swaffield, J. K. (1997). Constraints on the desired hours of work of british men. *Economic Journal*, **107** (441), 520–35.