Netspar
Network for Studies on Pensions, Aging and Retirement

Netspar DISCUSSION PAPERS

Marjan Maes

Estimation of a Structural Labour Supply Model for Belgium

Application to the Earnings Test for Pension Recipients

DP 09/2014-037

Estimation of a structural labour supply model for Belgium: application to the earnings test for pension recipients

Marjan Maes[1]

Abstract

In several OECD countries pension benefits are taxed away if retirees continue working and receive earnings that cross a threshold. Recently governments started to increase these earnings thresholds for working pensioners. On the basis of Belgian administrative data covering the years before and after such a reform, we show that bunching at a convex kink of the budget constraint tracks these tax rule changes. To account also for income effects along other parts of the piecewise-linear budget constraint, we estimate a structural labour supply model. A removal of the earnings test would lead to an increase of 2.44hours worked per week but causes a budgetary deficit as well: pension benefits will be claimed earlier (as the deferral rate for pensions equals zero in Belgium) while social contribution revenues fall back. The intuition is that the income effect which leads high-earners to work less partly offsets the substitution effect which stimulates low-earner (previously at the convex kink) to work more.

It took the OECD, the IMF and influential research institutions about two decades to convince policymakers that actuarially unfair early retirement programs stimulated older workers to retire early, that is before the statutory retirement age. Finally, the prospect of financially unsustainable pension schemes led policymakers to dismantle early retirement schemes, albeit facing popular resistance. By contrast, one of the few pension reforms that is by now widely implemented even before any research evidence is available about its budgetary or distributional consequences, let alone whether it is socially desirable, relates to the earnings test on pension benefits. As Table 1 demonstrates, several OECD-countries used to tax away pension benefits in case retirees continue working in the 90's. But several governments currently increase or abolish earnings test thresholds on retirees *past* the statutory retirement age. This is often accompanied by raising deferral rates on pension benefits as well. According to some theory of justice it may indeed seem unfair to "plague" those who set a good example by working at an old-age in countries characterized by massive early retirement.

[1] The author gratefully acknowledges financial support from NETSPAR Research grant and the University of Tilburg (CENTER). The author is grateful to Prof. Arthur Van Soest and Prof. Monika Buetler for comments and suggestions.

Table 1: Earnings[2] test rules in OECD countries in 1995 and 2013 (for single workers without dependent children; threshold as % of average earnings) (OECD(2013), Blöndal-Scarpetta(1998), ISSA(2003))

	1995					2013				
	Before statutory age		After statutory age			Before statutory age		After statutory age		
	Threshold	Withdrawal rate	threshold	Withdrawal rate	Deferral rate	Threshold	Withdrawal rate	Threshold	Withdrawal rate	Deferral rate
AU	91	full	29.8	Full	No	91	Full	/	/	4.2
CAN	n.a.[3]	n.a	160	15	No	n.a.	n.a.	150	15	7.2
DK	n.a.	n.a.	50	60	no	n.a.	n.a.	75	30	+/-5.8
FR[4]	/	/	/	/	no	/(after 41years contributions)	/	/(after 41years contributions	/	5
GER	15	33/50/66	/	/	no	15	33/50/66	/	/	6
GR	12000/?	Full	116	Full	No	60	70	60	70	3
IRE	n.a.	n.a.	0	Full	No	n.a.	n.a.	/	/	no
ITA	0	full	23	100	Yes?	/	/	/	/	+/-5.5
JAP	n.a.	n.a.[5]	<17 17-90 >90	20 60 100[6]	Yes	<21 21-93 >93	0 60 100	10	50[7]	8.4
NLS	n.a.	n.a.	/	/	No	n.a.	n.a.	/	/	No
NOR	n.a.	n.a.	20	40	9	/	/	/	/	Yes
POR	0	Full	0	Full	No	/	/	/	/	6
ESP	0	Full	0	Full	No	?	?	/	/	2-4
SWE	[8]		/	/	7.2	/	/	/	/	+/-5.5
UK	n.a.	n.a.	/	/	7.4	n.a.	n.a.	/	/	10.4
US	24	50	38	33	Yes	24	50	/	/	8

In a standard welfarist approach however, Cremer et al.(2009) have shown that an earnings test may be part of an optimal tax system if earnings are more unequally distributed after the statutory age of retirement than before and if labour supply elasticity is less after than below the statutory retirement age: "One way or the other, it appears that the empirical studies showing that an earnings test has an adverse effect on labor supply are not sufficient evidence to advocate their suppression."(p. 17). In contrast to a continued interest in its labour supply impact, the implications of earnings test tax changes for revenue raising are barely studied. Furthermore, the

[2] The earnings test should not be confounded with an assets or means-test as is usual in targeted pension schemes.
[3] Not applicable refers to old-age pension schemes for which the earliest age of claiming corresponds to the statutory retirement age. This typically relates to a basic pension.
[4] Access to old-age pension is conditional on ending the employment contract in the current job.
[5] We take 60 as statutory age in 1995 and 65 in 2013.
[6] For the earnings-related pension. It is 0% for the basic pension.
[7] 0% for the basic pension and 50% for the earnings-related pension.
[8] Partial pension can be combined with earnings (1/2 or 4/5 workweek)

labour supply impact itself of the earnings test has always been quite of a puzzle: how come that removing the marginal tax on work from more than 100% to zero leads to a very small increase in aggregate labour supply?

To address these issues, we argue that a small aggregate labour supply impact may result from the fact that there are substantial income effects for some individuals and substantial substitution effects for other people that partly offset each other in the aggregate. To demonstrate this we estimate a structural labour supply model that explicitly distinguishes between substitution and income effects like Friedberg(2000) and whose estimates will be used to simulate the budgetary impact of a removal of the earnings test in Belgium. We find a wage rate elasticity and income elasticity with respect to hours of respectively 0.02 and -0.13 which is comparable to the corresponding literature. A removal of the earnings test threshold for those aged 65 or older increases median hours worked per week with 2.44 hours per week. It would however lead to a loss of social contribution revenues of 22%. It may seem surprising that the increased labour supply generates a budget deficit: this is because the income effect mainly affects high earners and the substitution effect mainly stimulates low earners.

The situation in Belgium is extremely well-suited to identify the impact of this reform. As considerable increases in the earnings test threshold took place in Belgium in the last decade, we exploit a panel of Belgian administrative data (1999-2009) that cover years before and after this reform. Interestingly, this reform mainly raises the threshold for people of 65 and above and not for people younger than 65. In addition, in contrast to the UK and US pension scheme, the Belgian old-age benefit scheme does not compensate individuals in the event of deferred claiming past the age of 65 by increasing their future benefit streams. That would have complicated the identification of income effects and hence labour supply elasticities. Hence the nature of the Belgian reform offers possibilities for a clean identification strategy. Notwithstanding we are, to our knowledge, the first to analyse the earnings test in Belgium. Surprisingly, the empirical literature on the labour supply impact of an earnings test is in fact confined to studies on Canada, the US and the UK, the typical liberal welfare states. A recent exception is Hernaes-Zhiyang(2012) for Norway and Shimizutani,S.(2013) for Japan.

In principle, different models might be a candidate to estimate the labour supply impact of an earnings test reform. The first and most common approach is to estimate difference-in-difference models (see Disney-Smith(2002), Song-Manchester(2007), Haider-Lochran(2008),…). While such a reduced-form approach has the advantage to identify causal reform effects if adequate control and treatment groups are available, it is unable to disentangle income and substitution effects on labour supply.[9] That is why we did not follow this approach. Alternatively, we could have followed the emerging literature that estimates compensated elasticities of reported earnings with respect to the marginal tax rate on the basis of bunching behaviour at a convex kink of the budget constraint (Saez(2010), Gelbers et al.(2013),…)). This restriction to compensated elasticities might be explained by the fact that income effects usually seem to be insignificant for earnings of male prime-age workers. By contrast, when considering elderly workers, income effects will turn out to be quite important and are related to the receipt of old-age pension benefits.[10] A third approach would be to opt for the popular discrete choice model (like part-time, full-time and no retirement). However we want to compare our results to those of similar studies like Friedberg(2000). The gross wage elasticities usually calculated on the basis of discrete choice models are not directly comparable to the elasticities calculated in continuous models: "It is the post-tax wage that is used the conventional approach (such as the Hausman model) and is therefore the main focus in that part of the literature, whereas the discrete choice literature (including the present study) often report gross wage elasticities" (Dagsvik(2014), p.146). Secondly, discrete choice models often apply to prime-age workers f.e. van Soest-Das-Gong (2002), van Soest (1995), Wagenhals (2011),.. This explains perhaps why they typically do not emphasize the role of social security benefit entitlements like old-age pension rights as part of non-labour income. Yet, this is precisely the focus of our paper and what will drive for a large part what we call the income effect. Social benefit entitlements in case of part-time work absorb part of the income effect. It is well-known that discrete choice models are unable to reproduce well the peaks at part-time employment, unless dummy variable are introduced: "the ad hoc practice of introducing dummy variables into the conventional discrete labor supply model to improve the fit to the data has the unfortunate implication that the

[9] We also estimated a difference-in-difference model in Maes(2014).

[10] Gruber-Saez(2002) exploited tax returns data to uncover both income and substitution effects using 2SLS regression. However as their panel covers a period during which several marginal tax rates changed as well as tax exemption amounts, affecting all individuals in the income distribution, their instrumental variables turn out to be very adequate. That would unfortunately not be the case with the earnings test tax change we consider.

model will no longer be structural. As a consequence, it will be problematic how counterfactual policy simulations should be interpreted." (Dagsvik(2014),p. 139). In a third place, discretization of hours intervals implies incomplete use of information and rounding errors as noted by van Soest(1995). Discrete choice models incorporate the practice that employees may be forced to choose between a limited range of hours, like part-time or full-time. But we will show below that the distribution of hours of work per week in our sample is rather smooth, suggesting that individuals indeed have substantial freedom of choice

The structure of this paper is as follows. A first section discusses the earnings test rules in Belgium before and after the reform and presents a simple framework to understand the core argument (section 1). The second section presents our sample of longitudinal administrative micro-data (section 2). Descriptive evidence gives us a feel of the extent of bunching below the earnings test threshold before and after the reform (section 3). In section 4 we present estimation results of a structural labour supply model that allows us to disentangle income and substitution effects. The estimates of this structural model are used for policy simulations, i.e. the impact of a removal of the earnings test on labour supply and the governmental budget (section 5).

1. Belgian earnings test rules before and after a reform in a simple framework

Figure 1 displays the typical budget constraint of a Belgian older worker eligible for pension benefits before any reform took place (full black line). The horizontal axis shows annual hours of work and the vertical axis total net income that may result from earnings (=wage per hour times hours of work) and non-labour income (pension benefits, capital income). The pension beneficiary can obtain pension benefits in addition to earnings up to a certain threshold. For earnings above that threshold, corresponding to a number of hours h* in Figure 1, pension benefits are taxed at τ =100%. For earnings above 115% of the threshold, corresponding to hours of work h** in Figure 1, pension benefits are completely suspended. In this way the earnings test generates a piecewise linear budget constraint that consists of three segments, corresponding to earnings below to the first threshold, a horizontal segment, and earnings above 115% of the threshold. Graphically speaking, the wage $W(1-\tau)$ represents the slope of the first and third segments while virtual non-labour incomes (Yv1, Yv2 and Yv3) represent the intercepts corresponding to each extended or linearized segment. Under the hypothesis of normally

behaved preferences, we expect workers to bunch just below the threshold (corresponding to a convex kink). We also expect that no individuals are located in the non-convex kink h**. We will show below that that is indeed the case for the individuals in our sample.

Suppose now that the government raises the threshold successively as indicated in table 2: this occurred in 2002, 2004, 2006, 2007 and 2008 for pensioners reaching the statutory retirement age of 65 and older while it happened only once for those younger than the statutory age of 65.[11]

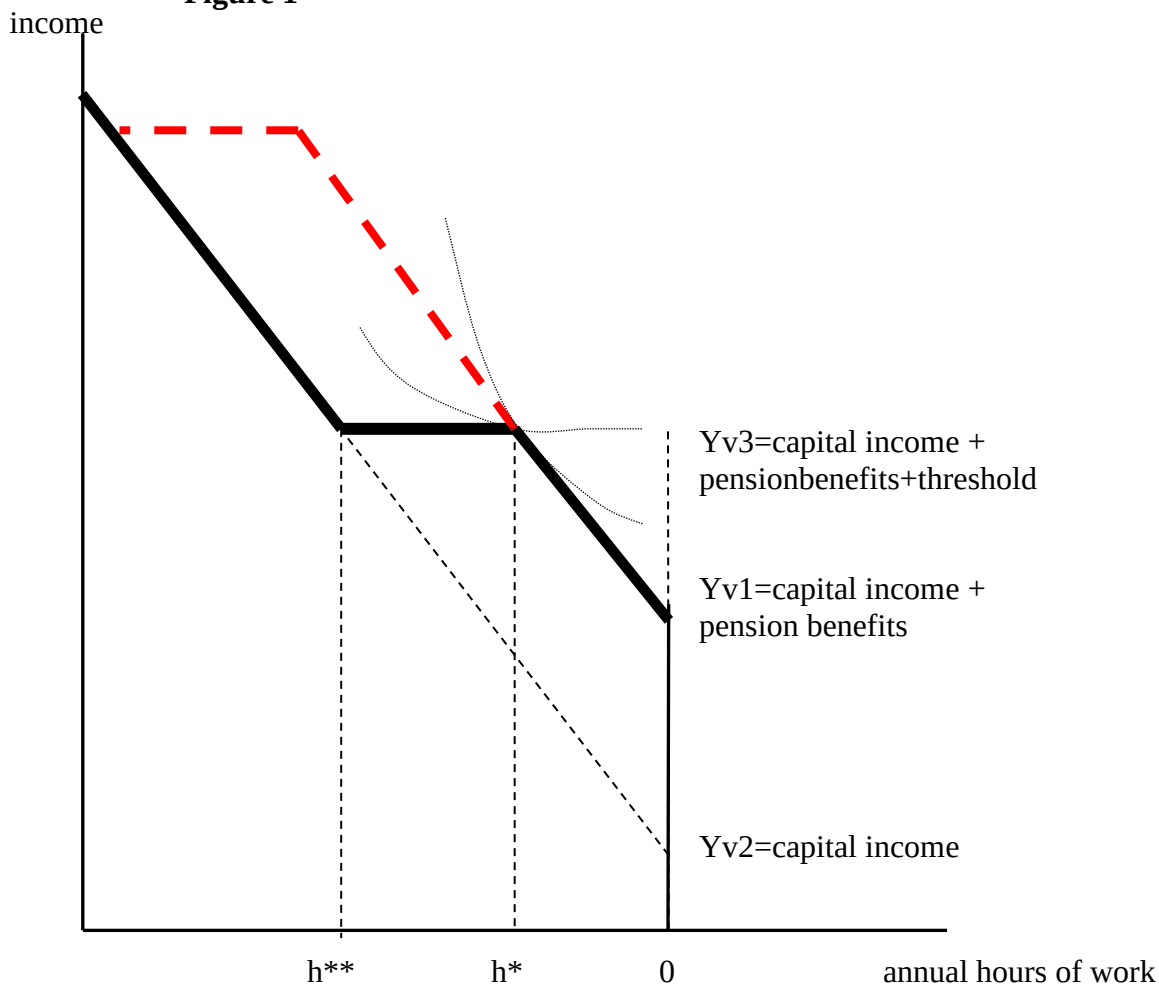
Table 2: Evolution earnings test thresholds (nominal amounts, in Euros) in Belgium for employees

	people below age 65	people of age 65 or older
1999	6857	6857
2000		
2001		
2002	7421	10845
2003		
2004		13556
2005		
2006		15590
2007		17149
2008		21436
2009		

For individuals aged 65 and older the budget constraint would change as indicated by the red dashed line in Figure 1. On the one hand, those who are located in the segment below the threshold would experience no incentive to change labour supply. For those with earnings bunching just below the threshold and on the horizontal segment h*-h**, increasing the threshold is equivalent to a reduced tax on labour income. Because their after tax wage increases, the substitution effect will stimulate them to increase their number of hours worked. On the other hand, for those with earnings above the threshold, to the left of h**, given their amount of hours worked, an increased threshold yields an increase in net income. If leisure is a normal good, they will supply fewer hours of work while earning the same net income as before. Hence the aggregate labour supply effect depends on the relative strength of the income (measured by virtual income) and substitution (measured by wage) effects and on the distribution of individuals across these segments. Theoretically, the impact of the earnings test on labour supply is hence inconclusive.

[11] It also depends on whether the employee has dependent children this is taken into account in the analysis. For brevity, we only display the amounts of the typical case without dependent children.

Figure 1



Note that the Belgian old-age pension scheme still has a statutory retirement age of 65. Early claiming of old-age benefits is possible for private wage-earners at the age of 60 conditional on a career of 35 years. It is calculated as *N/45**average lifecycle earnings**k*, where N represents the number of years worked or spent on replacement income and *k* is a replacement rate, which takes on the value of 0.60 or 0.75 depending on whether the pension recipient claims benefits with or without a working partner. Hence pension rights accrue by merely 1/45 (=+/-2%) per additional year working. For those having a career longer than 45 years, a dropout-year provision replaces low-earnings years with higher ones. There is no actuarial adjustment of benefits in case of early claiming. Neither is there a deferral rate in case of delayed claiming post the statutory retirement age of 65. Similarly, civil servants may claim old-age benefits early at 60 conditional on a career

of 5 years. Again the deferral rate after 65 is zero and there is no actuarial adjustment in case of claiming before 65. Civil servant pensions are a fraction of average earnings during the last five years before retirement. This fraction is calculated as the number of years the person worked in the public service divided by a benefit accrual factor. The latter equals 60 for the majority of civil servants but it is 50 for policemen, 55 for teachers in secondary schools,. In any case, the fraction cannot exceed 75%(=45/60).

An obvious weakness of the theory of labour supply behind Figure 1, is that it supposes a one-period framework which intuitively conflicts with the fact that the decision to claim pension benefits supposes at least a two-period framework. It does not incorporate actuarial adjustment of benefits in case of deferral. However given the zero deferral rates in the Belgian old-age pension scheme, we argue that this is not a strong limitation while for instance in the US it is 8% and in the UK 10.4%. If the earnings threshold increase were accompanied by the introduction of an actuarially fair rate of deferral, that would have neutralised in a sense the negative income effect that affects the labour supply of workers in the upper segment. This is probably why Disney-Smith(2002) found that in the UK the removal of the earnings test on pensioners led to a large labour supply increase.

As one can notice from table 1, several OECD countries increased their earnings test threshold on pension benefits post the statutory retirement age, often combined with an increase of the statutory age itself and the introduction of a deferral rate. Notwithstanding common perceptions that the earnings test should be eliminated for giving strong work disincentives, according to the empirical literature the aggregate labour supply impact of an earnings test removal is rather small, with only a sizeable positive impact for workers located around the convex kink. We summarize in table 3 the main research results on the labour supply impact of the earning test. In contrast to difference-in-difference models, a structural labour supply model designates a model whose parameters can be directly used to “feed” a corresponding utility function

Table 3: Overview empirical literature on the relationship earnings test- labour supply						
author	**Country (data)**	**What is simulated?**	**Simulated labour supply impact**	**Who is affected?**	**Deferral rate applicable to sample**	**Type model**
Burtless-Moffitt (1985)	US (RHS 1969-1977)	Removal for all ages	1.3 (6.2) hours per week for men (men at or above the kink)	12% of 62-69 is working	1%	Structural
Honig-Reimers (1989)	US (RHS and CPS 1986)	Removal for all ages	13-20% hours increase at kink and 1% hours decrease in upper segment	Of 62-69: 22% at kink, middle or upper segment	3%	Structural
Friedberg (2000)	US (CPS 1976-1987)	Removal above 65	5.3% increase hours at or above kink	Of 62-69: 4.1%kink 5.3% middle segment 15.2% upper segment	1-3%	Structural
		Doubling threshold[12] above 65	0.2% decrease hours at or above kink			
Haider-Lochran(2008)	US (CPS)	Removal above 65 in 2000 (with increase NRA 65-67)	Increase 1.55 hours per week among working men	25% of 65-75 is working	3-8%	Difdif
Song-Manchester (2007)	US (CWHS)	Removal above 65 in 2000(with increase NRA 65-67)[13]	Increase employment rate with 0.8-2%point	29% of 65-69 is working	3-8%	Difdif
Gustman-Steinmeier(2004)	US (HRS 1992-2002)	Removal between 62-64	Increase employment rate with 1.3-1.6(1.3) %point 62-64 (65-69) for each age	Of 65-69: 68% retired 15% part time 17% fulltime	3-8%	structural
Disney-Smith(2002)/ Disney-Tanner(2000)	UK (family expenditure survey 1984-1994)	Removal above 65	Increase 3-4 (2) hours per week for men (women)	7.58 (5.2)% of men (women) of 65-69 is working	7.4%	Difdif
Zabalza et al(1980)	UK (population census 1977)	Removal for all ages	Increase 10.24 (7.04) hours per year for men (women)	Of men 65-69 (women 60-64): 9(12)%fulltime work 16(26)%part time 75(62)%retired[14]	6.5%	Structural
Benjamin-	Canada	Removal	Increase 5-6	25-30% of 65-69	0%	Difdif

[12] Increase from $14500 in 1998 to $30000 in 2002.

[13] Note that the removal of the earnings test in the US in 2000 above 65 was simultaneous to a gradual increase of the NRA from 65 to 67 which makes the earnings test more stringent again for those *younger than* the NRA

[14] For men 60-64(women 55-59): 79(39)% full-time, 3(40)% part-time, 18(21)% retired.

Baker (1999)	(SCF 1972-1980)	above 65	weeks worked per year (from part-year to full-year) for working men	is working		
Hernaes-Zhiyang(2012)	Norway (-2008)	Removal above 67	Increase 8.7%point of individuals at or above kink	Those working at 67-70	0%	Difdif

2. Description of the sample

Our sample is based on a random selection out of the Belgian National Register of 80,000 national identification numbers of Belgian men between 57 and 72 years old on 31-12-2003[15]. These numbers are matched for the years 1999-2009 with variables registered by different social security institutions to generate a panel with variables like year and month of birth, age of other household members, number of children at charge, marital status, region (Flanders, Wallonia, Brussels), sector of activity, replacement income (early retirement, unemployment, disability, old-age,..), earnings, earnings of the spouse, means-tested benefits, professional code (like agriculture, banking-insurance,…), full/part-time contract (albeit defined differently before/after 2005 and not very reliable), number of years of the career, occupational pension benefits and for each quarter of the year we dispose as well of the effective number of hours worked during a given quarter, as a percentage of a benchmark (40 hours per week). This is the labour supply variable of interest. The sample size decreases slightly over time due to increasing number of deaths of older people over time. The red lines in table 4 show when and for whom an earnings test threshold increase took place: the 1st January 2002, 2004, 2006 and 2007 the threshold was raised for those 65 and older; for those between 60-65 it was only raised once on 1st January 2002. We believe that the fact of having individuals subject to different taxation rules will enhance the identification of the estimates in the model.

[15] It is only since 2003 that the Datawarehouse registers data from all social security institutions and hence covers 100% of the Belgian citizens. Because we want a random sample of the population the selection year had to be 2003.

Table 4: Sample size by year and age											
age	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
53	6945										
54	5752	6945									
55	5693	5752	6945								
56	5421	5693	5752	6945							
57	4681	5421	5693	5752	6945						
58	4347	4681	5421	5693	5752	6945					
59	4582	4347	4681	5421	5693	5752	6898				
60	5084	4582	4347	4681	*5421*	*5693*	*5713*	*6838*			
61	5237	5084	4582	4347	*4681*	*5421*	*5644*	*5658*	*6768*		
62	5018	5237	5084	4582	*4347*	*4681*	*5373*	*5583*	*5609*	*6691*	
63	4694	5018	5237	5084	*4582*	*4347*	*4640*	*5317*	*5519*	*5565*	*6626*
64	4543	4694	5018	5237	*5084*	*4583*	*4290*	*4595*	*5269*	*5463*	*5491*
65	4496	4543	4694	5018	5237	5083	4509	4225	4517	5197	5389
66	4552	4496	4543	4694	5018	5237	5021	4456	4171	4452	5113
67	4572	4552	4496	4543	4694	5018	5147	4943	4381	4106	4382
68	4383	4572	4552	4496	4543	4694	4936	5059	4862	4293	4023
69		4383	4572	4552	4496	4543	4609	4847	4962	4778	4223
70			4383	4572	4552	4496	4441	4527	4763	4863	4687
71				4383	4572	4552	4388	4335	4426	4649	4759
72					4383	4572	4424	4280	4228	4331	4543
73						4383	4432	4276	4169	4128	4218
74							4247	4272	4141	4061	4024
75								4087	4116	4008	3940
76									3929	3967	3858
77										3769	3819
78											3611
total	80000	80000	80000	80000	80000	80000	78712	77298	75830	74321	72706

In our sample, few old-age pensioners combine pension benefits with labour income: only 8.4%(=5795/68628) of all pension beneficiaries in ages 60-64 and 6.2% (=12411/200341) of the pension beneficiaries in ages 65-69. So the large majority of pensioners have zero hours of paid work.

Note that for the subsequent analysis several groups are deleted from the sample. Firstly, as the numbers of hours worked are unavailable for the self-employed we delete them.[16] Consequently, we assume that there is no problem of tax evasion or measurement error as the model is confined to employees whose earnings are declared by third-party employers. Secondly, note that we only keep individuals eligible for old-age pension benefits and older than 59. However next to the

[16] In the future, we hope however to incorporate the self-employed in a labour supply model. The bunching behaviour of self-employed as emphasized by Saez(2010) turns out to be very relevant also in Belgian tax returns data.

official old-age pension scheme, several forms of early retirement-programs have developed since the 1970's, based on collective agreements between employees and employers: they allow retirement at 58, 56, 55, 54 or even at 52 conditional on career requirements and sector of activity. Those who are early retired stay in the early retirement-scheme until the age of 65 at which they are transferred to the old-age pension scheme. All observations concerning these early retirement beneficiaries will be deleted from our sample. We assume that people who have been retired from the labour market for 5-10 years do not come back after 65: the % of early retired that re-enters the labour market is 0.07%. Thirdly, military personnel and teachers whose NRA itself is lower than 65 (e.g. for military personnel 54, for policemen 56,..) face specific old-age eligibility rules requiring specific data that we miss. Also civil servants often have the alternative to quit work early through disability pensions where the screening seems to be weak. However as disability pensions of civil servants are calculated exactly as old-age pensions and subject to the same earnings test, these observations are treated just like old-age pensions. Finally, as the earnings coming from the exercise of political mandates can be cumulated without any limit[17] with pension benefits we delete observations related to these persons from the sample.

We summarize some relevant descriptive statistics of the remaining sample in table 5. This shows that about 72% of our sample have zero hours, 1.26% are bunching and 20% are working hours above the threshold. It also informs about the vector of individual characteristics X_i that includes a linear age variable and a dummy for age over/below 65,17 dummies for sector of activity, number of household members, a dummy for having been unemployed previously, 2 regional dummies.

[17] This is allowed if the start of the political mandate is anterior to the claiming of pension benefits and inspection of the data reveals that this is a common practice.

Table 5: summary statistics		
	All (hours of work>=0)	Working (hours of work>0)
Number of observations	83205	23246
Age 60-64	57.6%	77.09%
Age 65-69	36.12%	16.74%
Age 70+	6.28%	6.17%
Having been unemployed previously	10.07%	12.88%
Walloon	30.06%	29.92%
Brussels	6.87%	8.55%
Flemish	63,07%	61.53%
Private sector	54.69%	71.79%
Public sector	45.31%	28.21%
Single	12.39%	13.37%
Number household members=1	78.25%	72.89%
Number household members=2	15.58%	18.40%
Number household members>2	6.17%	8.71%
Mean hours per week	8,07	28.8
Mean wage per hour	12.27	12.20
Mean virtual income per year(/1000)	17.86	6.10
Location on budget constraint:		
Zero hours	72.07%	/
Segment between zero and convex kink	6.71%	24.02%
Convex kink	1.26%	4.52%
Middle horizontal segment	0.13%	0.48%
Upper segment	19.81%	70.95%
We only kept 23246 observations of individuals eligible for pension benefits. We deleted: observations of self-employed, with wage<1 or >63, of individuals who were fully retired with zero hours in first year of observation, of individuals without pension benefits while having earnings less than the threshold (because the reason for not claiming is not related to the earnings test), of individuals in the year they claim benefits because they have in that year also some labour income but, as we have only yearly data, we are unable to identify whether this labour income is earned before the exact data of claiming or afterwards.		

3. Some descriptive bunching evidence around the convex kink

The next figures visualize the bunching phenomenon just below the earnings test threshold and how working pension beneficiaries respond to a change in the threshold. This gives us a feel of the substitution effect on labour supply. We will compare the distribution of labour incomes of a treatment (65-69) and control (60-64) group in the years before any increase (1999-2002) to the years 2005-2009 when the main threshold increases took effect. In particular, we group labour incomes, in nominal[18] terms, in intervals of 1000euros relative to the pre-reform threshold of 1999-2001. Somebody in the '0' interval has earnings exactly equal to the threshold.

Figure 2 starts with employees older than 64: there appears to be a clear concentration of labour incomes in the intervals just below the pre-reform threshold: see the yellow bars. Over time, as the threshold increases, earnings are spread more and more equally: see the dark grey and black bars. This figure has to be compared with figure 3 that shows the same data for the age group younger than 65. For them bunching remains strongly in place over time: both the yellow and dark grey bars represent permanently about 12% of working pension beneficiaries. That is because there was a once-only threshold increase from 6857 to 7421euro in 2002 that shifted the bunching to the +1000interval. The fact that bunching remains strongly in place strengthens our hypothesis that the shift in the earnings distribution for those above 65 is not driven by some macro-economic trend affecting all age groups but rather by the earnings test reform.

[18] Notice that the threshold is never indexed to price increases while earnings are in Belgium always indexed to price increases. This implies that somebody who is bunching below the threshold should decrease his number of hours with about 2% a year, as his earnings are price-indexed while the threshold isn't. We chacked whether Figures 2 and 3 being in real or nominal terms wouldn't change the picture at all. To avoid a distorted picture, we also eliminated all observations of those in the first year of pension receipt as this is nearly always a year where the individual reveives earnings until the month in which the starts to receive pension benefits.

Figure 2: earnings of pension beneficiaries, age>64, relative to the pre-reform threshold

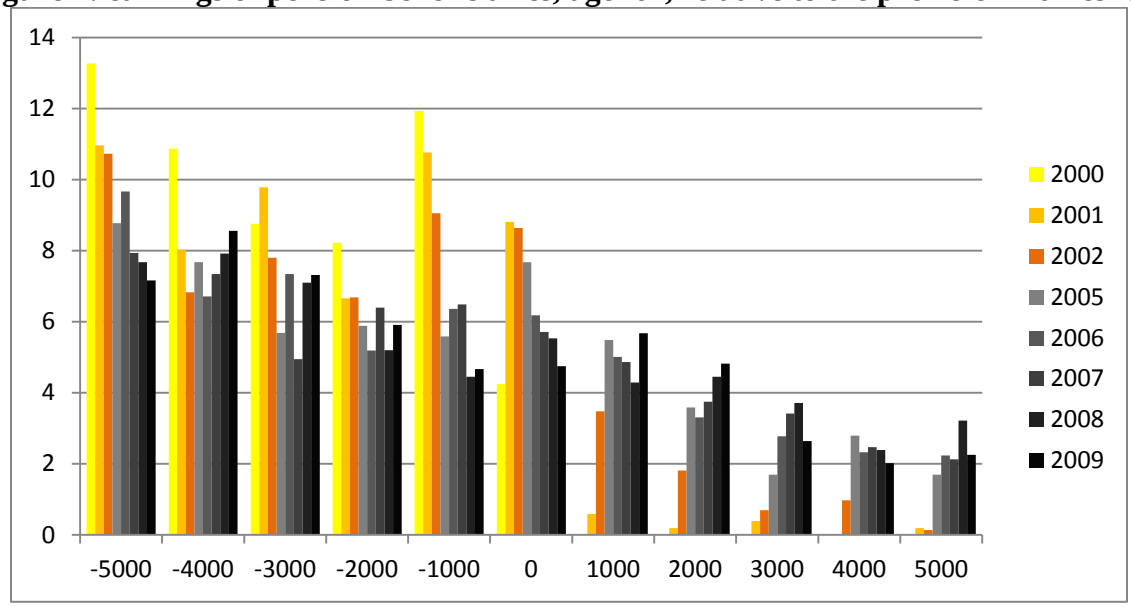
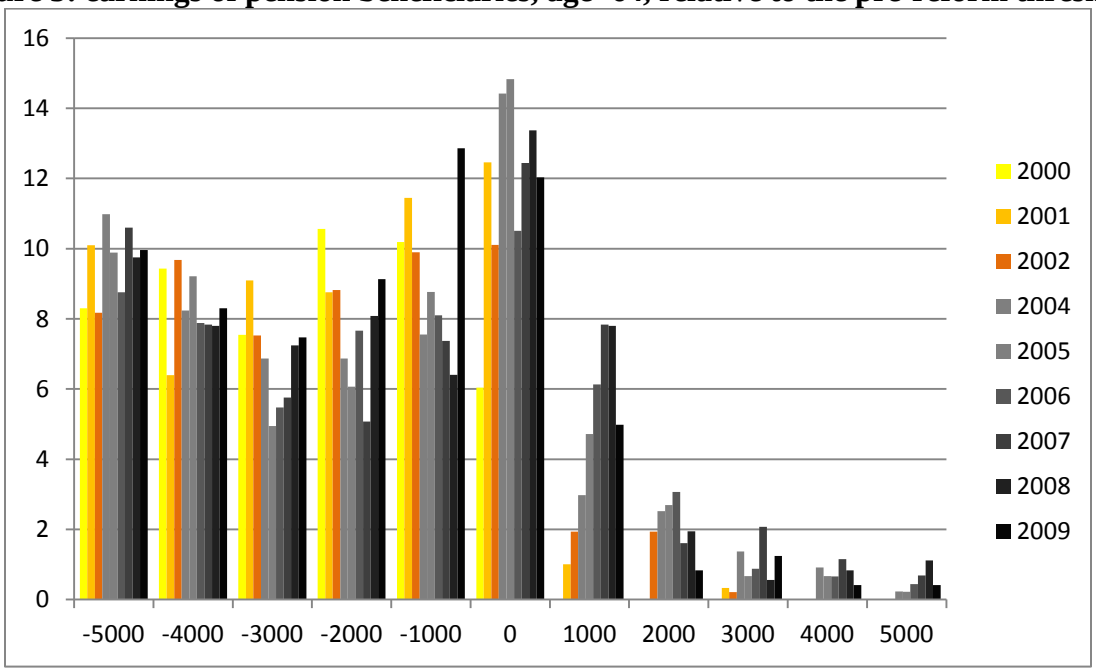


Figure 3: earnings of pension beneficiaries, age<64, relative to the pre-reform threshold



These figures visualized the bunching phenomenon just below the threshold and the visibly strong adjustment of labour supply behaviour to a newer and higher threshold. However, only a small part of the population is bunching and it remains to be seen how people located above the threshold will react in order to conclude about the aggregate impact on labour supply.

4. Estimation of a structural labour supply model

4.1. Methodology

The bunching evidence mainly reports what happens at the convex kink of the budget constraint and feeds the substitution effect on labour supply. Income effects induced along other parts of the nonlinear budget constraint depicted in figure 1 remained out of the picture till now. To take the whole piecewise-linear budget constraint into account, the main econometric complication arises from its nonlinearity. Following Friedberg(2000), we estimate the labour supply equation that corresponds to the piece-wise linear budget constraint of figure 1:

$$H_i = S_{1i}\left[Z_{1i}\theta + \alpha_i\right] + S_{2i}\left[Z_{2i}\theta + \alpha_i\right] + S_{3i}\left[Z_{3i}\theta + \alpha_i\right] + (1 - S_{1i} - S_{2i} - S_{3i})\left[H_i *\right] \qquad (1)$$

where H stands for weekly[19] hours of work, S_j is a dummy indicating whether utility is maximized respectively on segment j=1, 2, 3 or at the kink:

S_1= 1 if $H* - Z_1\theta > \alpha$, =0 otherwise

S_2=1 if $H** - Z_2\theta < \alpha$, =0 otherwise

S_3=1 if $H** - Z_3\theta > \alpha > H* - Z_3\theta$, =0 otherwise

1-S_1-S_2-S_3=1 if $H* - Z_3\theta > \alpha > H* - Z_1\theta$, =0 otherwise

with $Z_j\theta = \gamma X + \delta W_j + \beta Y_{Vj}$, W for net wage per hour, Y_V for virtual income, X for a vector of covariates and α for unobserved preferences. This way of writing down the labour supply equation shows each segment in terms of its wage rate and virtual income. We will be particularly interested in the signs and magnitudes of δ and β representing the substitution and income effect[20].

[19] We converted FTE in a measure of hours of work per week: given that there are 52 weeks in a year and 40hours of work per week for a full-time contract: FTE*40/52.

[20] We did not impose any a priori constraints on the sign of these coefficients, during the estimation procedure. It is after the estimation that we observe that these estimates have the sign one would intuitively expect.

The obvious question is how to obtain consistent estimates of θ. One could try to estimate equation (1) by OLS for the individuals on each segment separately. However if the estimation takes place on the individuals per separate segment, this leads to a truncated distribution of the error term and bias. It would not account either for the fact that due to policy change individuals may jump from one segment to another. Alternatively if all the individuals are pooled in one sample and one estimates by OLS an equation where the wage and virtual income of an individual are the ones of the segment where he is actually situated, the inconsistency may become even worse. The inconsistency with OLS arises because wage and virtual incomes are determined by the number of hours that an individual works. So the error term will be systematically correlated with the wage and virtual income variables. To overcome this endogeneity problem, one could consider IV estimation. The adequate instrumental variable should be correlated with the endogenous variable but uncorrelated with the error term; otherwise it leads to misspecification error which also results in inconsistent estimates. Suppose that the problem of endogeneity could be solved with good instrumental variables. This still leaves us with the problem that OLS is inconvenient to model individuals located at a convex kink: what is the wage rate and virtual income that should be imputed to individuals at the kink? They are simply undefined. As we showed before, bunching at the convex kink is a real issue for our study and should be taken into account in the estimation, especially the behaviour of these bunching individuals after policy changes. Burtless-Hausman (1978) showed that the coefficients θ of equation (4) can be consistently estimated with maximum likelihood. As Blundell et al. (1999) explained, "The essential idea underlying the development of likelihood functions in the presence of nonlinear constraints involves defining a set of "states of the world". Each state designates a particular segment of the budget set, with states being *mutually exclusive* and states *jointly covering* all parts of budget constraints'(p.4704). By taking into account the entire exogenous budget constraint, one removes the endogeneity that is associated with the joint choice of hours and net wage in a particular segment of the constraint. This involves defining for each individual a set of virtual incomes and net wages depending on the segment of the budget constraint where he may locate.

Evidence in our sample shows that when faced with an identical budget constraint, individuals choose nonetheless to spread themselves differently over the budget set. This is where the

parameter α comes into the picture. In particular, a random distribution of preferences accounts for the fact that the convex kink in the budget constraint is compatible with an entire range of utility functions. As explained by Moffitt (1986), "heterogeneity of preferences tends to generate clusters of observations at the kink of a convex budget constraint and tends to disperse observations away from the kink of a nonconvex budget constraint" (p. 321).

Applying maximum likelihood requires specifying the joint probability of the value of H_i along a segment or kink and the probability of choosing that segment or kink. This leads to the individual likelihood:

$$L(H_i)= pr\left[\alpha_i = H_i - Z_{1i}\theta\right]^{S_1}\ pr\left[H_i^* - Z_{1i}\theta < \alpha_i < H_i^* - Z_{3i}\theta\right]^{K}$$
$$pr\left[\alpha_i = H_i - Z_{3i}\theta; \alpha_i < \alpha_i'(Z_2, Z_3)\right]^{S_3}\ pr\left[\alpha_i = H_i - Z_{2i}\theta; \alpha_i > \alpha_i'(Z_2, Z_3)\right]^{S_2} /$$
$$(1 - pr\left[\alpha = -Z\theta\right]) \qquad (2)$$

Where the first term denotes the probability of the value of H_i along segment 1, the second term the probability of choosing the kink (where the probability of the value of H_i is one) and the third term the joint probability of choosing a value of H_i along the segment and the probability of choosing that segment, given the nonconvex kink. We condition this on the probability of having positive hours of work. If we assume a normal distribution for $\alpha\ N(0, \sigma^2)$, we estimate the parameters θ (and σ^2) so as to maximize $\sum_{i=1}^{N} \log L(H_i)$ with:

$$\log L(H_i) = S_{1i} \log\left[\frac{1}{\sigma_\alpha}\phi\left(\frac{\alpha_i = H_i - Z_{1i}\theta}{\sigma_\alpha}\right)\right] + K_i \log\left[\int_{Hi - Z_{1i}\theta}^{Hi - Z_{2i}\theta} \frac{1}{\sigma_\alpha}\phi\left(\frac{\alpha_i}{\sigma_\alpha}\right)\right] +$$

$$S_{2i} \log\left[\frac{1}{\sigma_\alpha}\phi\left(\frac{\alpha_i = H_i - Z_{2i}\theta}{\sigma_\alpha}\right)\Phi(\alpha')\right] + S_{3i} \log\left[\frac{1}{\sigma_\alpha}\phi\left(\frac{\alpha_i = H_i - Z_{3i}\theta}{\sigma_\alpha}\right)(1 - \Phi(\alpha')\right]$$
$$- \log\left[1 - \int_{-\infty}^{-Z_1\theta} \frac{1}{\sigma_\alpha}\phi\left(\frac{\alpha_i}{\sigma_\alpha}\right) d\alpha_i\right] \qquad (3)$$

With ϕ the standard normal pdf and Φ the cdf, as in Friedberg (2000)[21]. The second term for the convex kink expresses that the individual would desire to work more than H* if he had the wage and virtual income of segment 1 and less than H* if he would have the wage and virtual income of segment 3, neither of which is feasible. Individuals are attributed to the kink if they have hours of work in an interval between h* and h*-3.84hours per week[22]. Notice that there is no term for the nonconvex kink as it will never be chosen. However a special problem arises in the case of a nonconvex kink and normal indifference curves. The nonconvex kink generates the possibility of two tangencies between budget constraint and indifference curves. To solve for the global maximum one must compare the utility of being in the horizontal segment with utility of being in the upper segment (see Moffitt (1986)). We define α' as the cut-off value of α that makes the individual indifferent between the horizontal segment and the segment above the threshold. Finally, the last term expresses that we condition on people with positive number of hours, as we deleted individuals with zero hours of work from our sample.

[21] We use SAS optmodel and SAS/IML to perform the estimation with Newton Raphson methods.
[22] we would like to integrate besides the heterogeneity in preferences a measurement error in the estimation such that we can avoid this arbitrary intervention. We experimented with several intervals.

4.2. Estimation results

Table 6 displays the estimation results of maximizing the sum of individual loglikelihoods of equation (3).

Table 6: Estimation results; dependent variable=weekly hours of work	
Intercept	37.11 ()
wage rate per hour	0.049 (0.004)
Virtual income per year/1000	-0.592 (0.005)
Dummy for age >65	-6.75 (0.13)
Age-60	-0.764 (0.013)
Wallonia	0.590 (0.114)
Brussels	2.356 (0.214)
Number household members	0.364 (0.038)
Having been unemployed once	-3.761 (0.174)
σ	9.45
Uncompensated wage elasticity with respect to hours (at sample means):0.0212 Income elasticity with respect to hours (at sample means):-0.13	
Log likelihood:-73065	
Number observations: 23246; Standard errors in parenthesis. We also controlled for 14 sectors of activity.	

The fact that the sample includes individuals who were and who were not affected by policy changes enhances the identification of the coefficients of interest, that is the wage and virtual income variables. We found at the sample means (only composed of men) an uncompensated wage elasticity of 0.0212 which is, as usual, very small and which suggests a forward bending labour supply curve. The income elasticity is much higher in magnitude, in line with results of the literature, and equals -0.13.

It may be useful to compare these magnitudes to those of other structural labour supply models that explicitly integrated the earnings test and made use of the Burtless-Hausman technique.

Friedberg (2000) estimated exactly the same model on US data, hence with continuous hours distribution, and found at the sample means an uncompensated wage elasticity of 0.225 and an income elasticity of -0.297. A different model is that of Zabalza et al. (1980) on UK data who distinguished between 3 discrete states: full-time work, part-time work and full-time retirement. They obtained for men a wage elasticity of -0.020 ("familiar backward bending shape with a very small elasticity", p.261)) and an income elasticity of -0.023. As they suspected the statutory retirement age of 65 to affect their results, they estimated the regression separately on men younger than 65 and older than 64. The wage elasticity is -0.0078 at age 59 but increases to 0.012 at age 68: the backward bending curve was caused by the influence of men under the age of 65. Burtless-Moffitt (1985) found an uncompensated wage elasticity of 0.028 and an income elasticity of -0.428. They estimated jointly a retirement age equation and a postretirement hours of work equation. Interestingly they allow the impact of virtual wage and virtual income on postretirement hours of work to depend explicitly on the age and found that a higher age increases the impact of virtual income and decreases the impact of wages on hours of work. This Burtless-Moffitt (1985) model extends the Friedberg-model by estimating simultaneously hours of work after retirement with the retirement decision itself. Their sample includes retirees as well as non-retirees.

To check the goodness of fit of the estimation we look at the actual and predicted distributions of observations across the different segments of the budget constraint, as Zabalza et al. (1980). About 86% of all observations are correctly predicted on the segment they are actually located. Of those individuals located in the lower segment, 72% were correctly predicted to be below the threshold. Of those individuals in the upper segment, 96% were correctly predicted to be in that segment. However of the few individuals located at the kink or at the horizontal segment, the percentage correctly predicted is small. The following figures allow a visual comparison of the distribution of actual and of predicted[23] hours of work per week. We show these distributions for each segment of the budget constraint. Our model reproduces fairly well the spikes at full-time work. Although few people at the kink and in the horizontal segment were predicted to be exactly at that kink and exactly in that small horizontal segment, the predicted distribution of hours corresponds relatively well to the actual distribution. It over predicts however the number

[23] We simulated a random distribution of alfa with N(0,sigma).

of hours for those at the kink: this is due to the fact that 8% of those at the kink work full-time and hence have extremely low wage rates.

Figure 5: upper segment

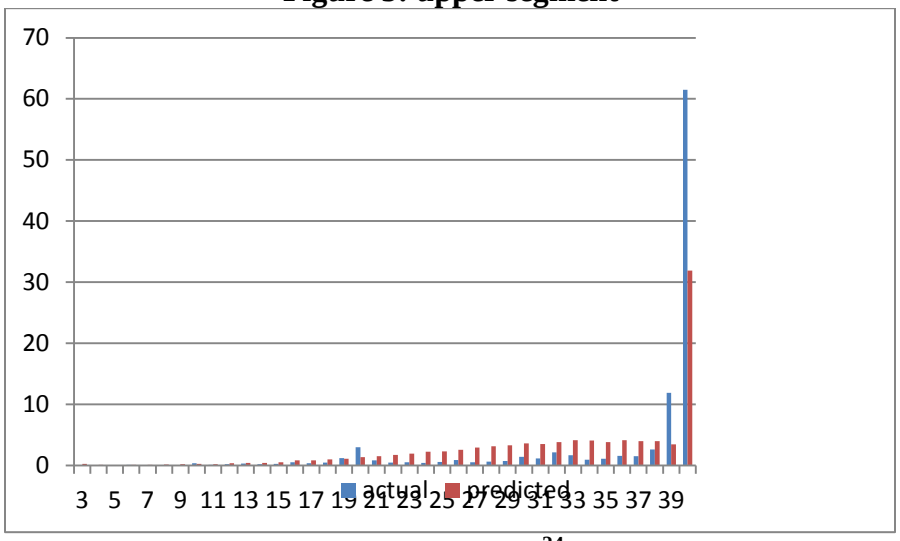
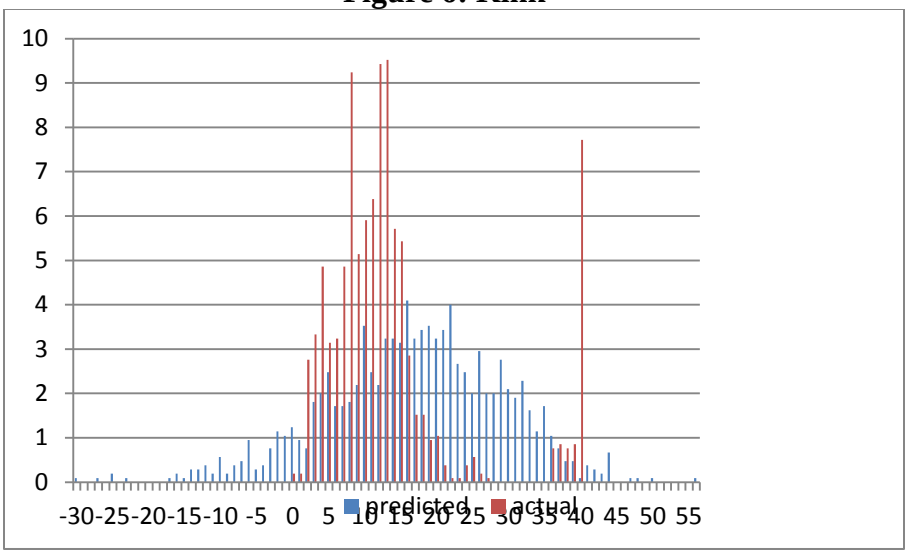


Figure 6: Kink[24]



[24] Notice the peak at 40 hours a week in the distribution of actual hours: it means that some people receive a yearly labour income exactly equal to the kink and are registered in the Datawarehouse Social Seucirty as working full-time. This implies that they have a very low wage. We think that this is measurement error which could be better integrated in the model.

Figure 6: Lower segment

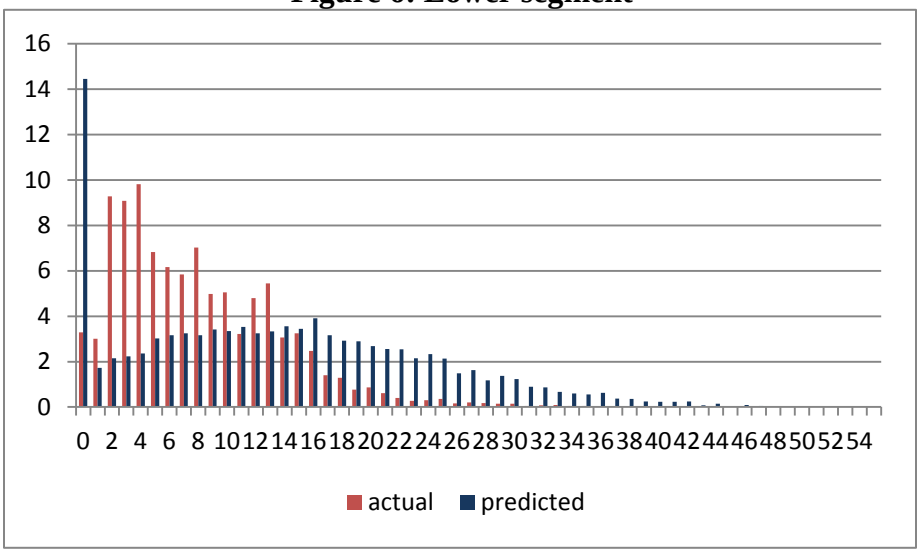
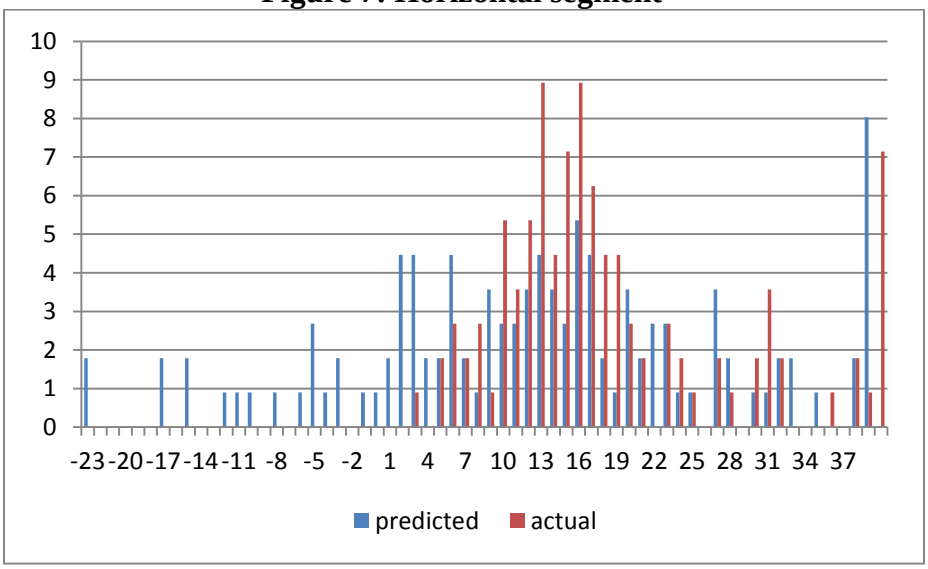


Figure 7: Horizontal segment



5. Policy simulations of a removal of the earnings test

5.1. Simulated impact on hours of work per week

The estimates of the previous section can be used to simulate policy changes. In particular, we simulate **the impact of a complete removal of the earnings test**. This corresponds to fully linearizing the piece-wise linear budget constraint of figure 1. We simulate the impact on hours worked per week, for the whole sample and for some subsamples (depending on whether the individual is younger or older than 65 at the time of reform. The results are in table 7.

Table 7: simulated and actual distribution of hours of work per week						
Percentile in the hours distribution	Simulated	Actual	Age 65 and over		Age 60-64	
			Simulated	actual	simulated	Actual
90%	35.1	40	25.1	40	36.53	40
75%	28.2	40	18.98	15.7	30.07	40
50%	20.5	39.6	12.04	9.6	22.9	40
25%	12.5	14.6	4.75	4.8	15.6	32
10%	4.7	5.5	-1.77	2.7	8.5	12.9
Mean	20.1	28.8	11.8	13.16	22.5	33.5

The results seem quite interesting and according to what we intuitively expect. A removal of the earnings test threshold for those aged 65 or older increases median hours worked per week with 2.44hours per week(from 9.6 to 12.04) and in the 75% percentile even with 3.28hours (from 15.7 to 18.98)[25]. This suggests that for the majority of individuals in this age category the substitution effect dominates the income effect. That is also what one could expect because the bunching individuals are concentrated in this age group. This is fully in line with e.g. Friedberg (2000). However a removal of the earnings test in the age group below 65 would lead to a strong decrease of hours worked: this suggests that the negative income effect plays an important role in this age group, as can be expected because the majority of these individuals are working full-time and situated in the upper segment above the threshold.

5.2. Simulated impact on the governmental budget

The **budgetary impact** of the removal of the earnings test results on the one hand from the change in social contribution revenues for the government: the removal of the earnings test may change hours worked and hence earnings. On the other hand it may change the flow of pension benefits to be paid out by the government as people may decide to advance claim their pension benefits due to the removal of the earnings test. Our results suggest unambiguously that the removal of the earnings test leads to government deficits.

According to our simulations, removing the earnings test **over 60** would decrease social contributions revenues by 38%. This is due to the strong decrease in hours worked in case the

[25]Average hours of work decrease because some individuals are predicted to have negative hours of work while the actual distribution does not allow negative values of hours of work: if we set the minimum of predicted hours equal to zero then average hours would increase after the earnings test removal.

earnings test is removed. This "average" of 38% masks interesting opposing results depending on the segment (or kink) in which individuals are located. Among those who were bunching below the threshold hours worked would increase a lot (due to a substitution effect) resulting in a 196% increase of social contribution revenues coming from those in the lower segment. In contrast, as those in the upper segment would reduce hours worked (due to an income effect) they would cause a loss of social contribution revenues for the government of 40%. What about removing the earnings test for the age group that reached the statutory age **of 65**? In that age group the majority of workers is in the lower segment or bunching below the threshold and as this group will supply more labour contributions revenues for the government go up by 35%. This group is probably the reason why governments want to remove a threshold in order to "make work pay". However, the increased social contribution revenues coming from this group, is more than counterbalanced by the income effect that hits those who are on the upper segment. We notice a 68% decrease of social contribution revenues due to reduced labour supply in the upper segment. In the end, the net effect on social contribution revenues is a loss of 22% revenues for the government. This is definitely a disappointing result for policymakers who are naively hoping that removing the earnings test threshold would raise more revenues for the government.

Table 8: percentage change on social contribution revenues for the government due to the full removal of the earnings test					
Aged 60-64			Aged 65 and older		
All (N=17926)	Hours<h** (N=2026)	Hours>=h** (N=15905)	All (N=5327)	Hours<h** (N=4639)	Hours>=h** (N=689)
-0.38	1.96	-0.4	-0.22	0.33	-0.68

The reason why the income effect weighs so heavily in social contribution revenues compared to the substitution effect is that in Belgium social contributions are a fixed percentage of earnings: the substitution effect is concentrated among low earners working a few hours a week and whose increased contributions are completely outweighed by the loss of contributions made by high earners / fulltime workers who are mainly hit by an income effect.

On top of this, the budgetary impact depends not only on the change in social contribution revenues but also upon the change in the flux of pension payments to be made by the government to early claimants. Our model is about the impact on hours worked, but having

figure 1 in mind, one sees that, following the earnings test removal, all those in the upper and horizontal segment will claim pension benefits. The budgetary cost due to advanced benefit claiming adds hence to the already significant budgetary cost.

The budgetary cost of the removal of the earnings test could be weakened by making the Belgian pension system actuarially fair at the margin, in particular by implementing a deferral rate on delayed benefit claiming. In that case, as in the UK and US system, the negative income effect would have only a marginal impact on the simulations. This explains why studies in the US estimate that, although in the short term the budgetary cost is high, in the long term, as the US pension system is becoming actuarially fair, the budgetary cost of an earnings test removal will decrease to zero. To take the impact of the degree of actuarially fairness of the pension system into account, our hours of work equation has to be extended with a "benefit claiming age"-equation in the vein of Burtless-Moffitt (1985).

Conclusion

In Belgium cumulating pension benefits with labour income is allowed up to a certain earnings threshold above which pension benefits are heavily taxed or even withheld. This earnings test is commonly perceived as a powerful work disincentive and a reason for policymakers to raise the earnings test threshold. Nonetheless, according to previous research the earnings test and its removal have an ambiguous and very modest impact on labour supply. The aim of our study was to examine the impact of the steady increase of the earnings test threshold (and even its complete removal) in recent years on the labour supply of older people in Belgium and furthermore on the budget of the government.

On the basis of longitudinal Belgian administrative data covering years before and after an earnings test reform, applying to people of a specific age group and not to people in another age group, we provide evidence of working pension beneficiaries with earnings concentrated just below the threshold who increase their labour supply as soon as the threshold increases. This reflects a substitution effect on labour supply. However those bunching individuals concern a minority and the aggregate employment impact depends also on the negative income effect that affects those who, before the reform, were supplying hours of work well above the threshold and

who, following a relaxation of the threshold, may decide to reduce hours of work. A structural labour supply model was estimated that accounts for a piecewise-linear budget constraint. We found significant income elasticity of labour supply as well as significant wage elasticity.

Our estimates were used to simulate a policy change that was in fact very recently implemented by the Belgian government on 1st January 2013: a complete removal of the earnings test above the age of 65. Our model suggests that this would raise median hours of work per week by 2.44. However, it would make the pension system, which is already financially unsustainable, even more unsustainable: not only will this policy advance claiming of pension benefits, it will also decrease social contribution revenues for the government by 22%. This is because the substitution effect that makes people work more is concentrated among low earners while the income effect mainly affects high earners to work less and claim pension benefits than they would otherwise do. We suggest that the magnitude of the governmental deficit is strongly related to the lack of actuarial fairness of the Belgian pension scheme.

References

Benjamin, M. - Baker, D. (1999), "How do retirement tests affect the labour supply of older men?", *Journal of Public Economics* 71(1), 27-51.

Blomquist, N. - Hansson-Brusewitz, U. (1990), "The effect of taxes on male and female labor supply in Sweden", *Journal of Human Resources*, 317-357.

Blomquist, N. (1995), "Restrictions in labor supply estimation: is the MaCurdy critique correct?", *Economics Letters*, 47, 229-235.

Blöndal, S. - Scarpetta, S. (1998), The retirement decision in OECD countries, OECD WP 202, Paris.

Brinch, C. - Hernaes, E. – Zhiyang, J. (2012), "Labor supply on the eve of retirement", DP Statistics Norway.

Burtless, G. - Hausman, J. (1978), "The effect of taxation on labor supply: evaluating the Gary negative income tax experiment", *Journal of Political Economy* 86(6), 1103-1130.

Burtless, G. - Moffitt, R. (1985), The joint choice of retirement age and postretirement hours of work", *Journal of Labor Economics*, 3(2), 209-236.

Cremer,H.-Lozachmeur, J.- Pestieau, P. (2009), "Social desirability of the earnings test", *German Economic Review* 9(2), 114-134.

Creedy, J. - Kalb, G. (2005), "Discrete labour supply modelling: specification, estimation and simulation", *Journal of Economic Surveys* 19 (5), 697-734.

Dagsvik, J.- Zhiyang, J. -Kornstad,T. - Thoresen, T. (2014), "Theoretical and practical arguments for modelling labour supply as a choice among latent jobs", *Journal of economic surveys* 28(1), 134-151.

Disney, R. - Smith, S. (2002), “The labor supply effect of the abolition of the earnings rule for older workers in the United Kingdom”, *Economic Journal* 112(478), C136-C152.
Ecklof, M. – Sacklén, H. (2000), “The Hausman-MaCurdy controversy: why do the results differ across studies?”, *Journal of Human Resources* 35(1), 204-220.
Gruber, J. - Orszag, M. (2003), “Does the social security earnings test affect labor supply and benefits receipt?”, *National Tax Journal* 56(4), 755-773 .
Friedberg, L. (2000), “The labor supply effects of the social security earnings test”, *Review of Economics and Statistics* 82(1), 48-63.
Gelbers, A. - Jones, J. - Sachs, D. (2013), “Earnings adjustment frictions: Evidence from the social security earnings test”, NBER 19491.
Gruber, J. – Saez, E.(2002), “The elasticity of taxable income: evidence and implications“, *Journal of Public Economics* 84, 1–32 .
Gustman, A. - Steinmeier, T.(2004), “The social security retirement earnings test, retirement and benefit claiming”, *NBER WP 10905*.
Haider, S. - Lochran, D. (2008), “The effect of the social security earnings test on male labor supply: new evidence from survey and administrative data”, *Journal of Human Resources* 43, 57-87.
Honig, M. - Reimers, C.(1989), “Is it worth eliminating the retirement test?, Current Issues in social security”, *AEA Papers and proceedings*, 79(2), 103-107.
Kuismanen, M. (2009), “Piece-wise or differentiable budget constraint? Estimating labour supply function for Finnish females”, *Applied Economics* 41, 1461-1472.
Maes,M.(2014), “The earnings test in Belgium: a difference-in-difference analysis”, forthcoming Discussion Paper.
Moffitt, R. (1986), “The econometrics of piecewise-linear budget constraints: a survey and exposition of the maximum likelihood method”, *Journal of Business & Economic Statistics* 4(3), 317-328.
Saez, E. (2010), “Do Taxpayers Bunch at Kink Points?”, *American Economic Journal: Economic Policy*, 2(3): 180-212.
Shimizutani, S. (2013), “social security earnings test and the labour supply of the elderly: new evidence from unique survey responses in Japan”, *Japanese Economic Review*, 64(3): 399-413.
Song, J. - Manchester, J. (2007), “New evidence on earnings and benefit claims following changes in the retirement earnings test in 2000”, *Journal of Public Economics* 91(3-4), 669-700 .
Michaud, P.-C. - Van Soest, A. (2007), “How Did the Elimination of the Earnings Test above the Normal Retirement Age affect retirement expectations“, *Fiscal Studies*, 29(2), 197-231.
van Soest, A.(1995), “Structural Models of Family Labour Supply”, *Journal of Human Resources* 30: 63-28.
van Soest, A - Das, M. - Gong, X. (2002), “A structural labour supply model with flexible preferences” *Journal of Econometrics*, 107(1-2), 345-374.
Zabalza, A. - Pissarides, C. - Barton, M. (1980), “Social security and the choice between full-time work, part-time work and retirement”, *Journal of Public Eeconomics*, 14, 245-276.