www.cesifo.org/wp

The Threat of Monitoring Job Search. A Discontinuity Design

Bart Cockx
Muriel Dejemeppe

CESIFO WORKING PAPER NO. 3267
CATEGORY 4: LABOUR MARKETS
NOVEMBER 2010

CESifo *Working Paper No. 3267*

The Threat of Monitoring Job Search. A Discontinuity Design

Abstract

Since July 2004 the job search effort of long-term unemployed benefit claimants is monitored in Belgium. We exploit the discontinuity in the treatment assignment at the age of 30 to evaluate the effect of a notification sent at least 8 months before job search is verified. The threat of monitoring increases transitions to employment, but of lower quality. In the less prosperous region, Wallonia, the impact is smaller, despite of the presence of specific counseling for the notified workers, and more heterogeneous. Moreover, in this region, the threat induces women to substitute sickness for unemployment benefits.

JEL-Code: J64, J65, J68, H43.

Keywords: evaluation, monitoring, job-search, threat effect, regression-discontinuity, grouped data.

Bart Cockx
Gent University
Tweekerkenstraat 2
9000 Gent
Belgium
Bart.Cockx@UGent.be

Muriel Dejemeppe
Catholic University of Louvain
Place Montesquieu 3
1348 Louvain-la-Neuve
Belgium
muriel.dejemeppe@uclouvain.be

18.11.2010
This paper is a thorough revision of an earlier version (Cockx and Dejemeppe, 2007). The major change concerns the outcome variables, since, in contrast to the previous paper, we can now identify employment and other destination states for which unemployment is left. In addition, this paper contains some methodological contributions regarding discontinuity design with a grouped continuous forcing variable. Conclusions also differ in major respects.

1 Introduction

The provision of Unemployment Insurance (UI) involves a trade-off between insurance and work incentives. Many economic researchers have studied how limiting the coverage of UI and the duration of benefit entitlement can restore work incentives (see e.g. Lalive et al., 2006). However, most UI schemes also provide work incentives by imposing job search requirements amongst others[1] on benefit claimants. Boone et al. (2007) argue that such requirements, enforced by monitoring and sanctions, may deliver the right incentives by imposing less costs than the aforementioned alternatives. This paper is about the impact of the introduction of such a monitoring scheme in Belgian UI in 2004.

In most countries[2] job search requirements are implemented as a sequence of regular face-to-face interviews in which caseworkers monitor job search methods and employer contacts reported by the benefit claimant. The claimant who is detected not to comply with the rules or with an individual agreement drawn up at an earlier meeting may be sanctioned, usually in the form of a temporary benefit reduction. The new scheme that was introduced in Belgium shares these general features, but was particular in that the monitoring interviews did not start early in the unemployment spell, as in many countries, but were targeted on the long-term unemployed claiming benefits for more than 21 months.[3] A consequence of this timing is that the benefit claimant may anticipate the monitoring interview and change her (job search) behavior beforehand. In this study we focus on this "threat effect" (Black et al., 2003) of job search monitoring by analyzing the impact of a written notification sent to benefit claimants several months before the interview.

This focus is relevant, since in the presence of threat effects a monitoring scheme does not only affect the population attending the interviews. Consequently, if one neglects these effects, the impact of such a scheme may be underestimated. Moreover, if threat effects are important, this also has consequences for the design of the monitoring scheme. For it suggests that one can save on caseworker personnel costs by not starting with monitoring interviews immediately after the intake in unemployment, but later on, so that there is time for the threat effect to realize.

The reform in Belgium is implemented differently according to the region of living. In Wal-

[1] Most common other requirements are to be registered as job seeker, to be available for work and to participate in training or work experience programs.

[2] See OECD (2007) for an international comparison of monitoring schemes.

[3] For youth aged less than 25 years this threshold was set at 15 months. Since this study focuses on benefit claimants aged 30, we will describe in the sequel only the regulations that apply to this age group.

lonia, the French speaking region in the South of the country, the notified unemployed are invited to counseling interviews within two months after dispatch of the notification.[4] This means that we identify in this region the *combined* effect of notification and counseling. By contrast, in Flanders, the Flemish speaking region in the North, the notified group is not systematically counseled, so that we identify in this region the *pure* threat effect induced by the notification.

The analysis is based on rich administrative data, which do not only allow to identify the impact on the job finding rate, but also on the quality of employment in terms of its starting wage and duration, and on exits to training and out of the labor force, including sickness benefits. A discontinuity design (DD), resulting from the gradual phasing in of the new monitoring scheme by age group, identifies these effects under weak assumptions (Thistlethwaite and Campbell, 1960; Hahn et al., 2001; Lee and Lemieux, 2009). Between July 2004 and June 2005 the job search requirements were only imposed on benefit claimants younger than 30 years on July 1, 2004. In the following years the older age groups were gradually integrated. This study exploits the discontinuity in the treatment assignment at the age of 30. Since the discontinuity disappears after a year, we can only identify the threat effect of notification and not the ex post effects of the monitoring interviews.

The DD inference is slightly complicated by the measurement of age in monthly intervals. By this grouping of the forcing variable no data are available in a close neighborhood around the discontinuity which makes non-parametric estimation of the treatment effect infeasible. The chosen functional form for the relationship between the outcome of interest and age is thus subject to a specification error. Lee and Card (2008) provide a solution to this problem for the case of a *discrete* forcing variable. In this research we adjust this solution for a *grouped continuous* forcing variable, such as age, and discuss the interest of such an adjustment. In addition, we propose a goodness-of-fit test for the chosen functional form that explicitly takes the aforementioned specification error into account.

The outline of the article is as follows. In the next section we describe the institutional setting and the features of the new monitoring scheme. In Section 3 we review the literature and the theoretical predictions of job search theory. Section 4 describes the data. The estimation method, focusing on the methodological contribution of this research, is presented in Section 5. Section 6 reports the treatment effects on various outcomes and contains a number of validity checks. A final section concludes.

[4]A similar procedure applies to Brussels, but this region is not retained in the analysis, since the available sample is to small.

2 Institutional setting

2.1 Before the 2004 reform

Belgium is a federal state that has decentralized certain competencies to the regional level. UI is organized at the federal level. It pays out unemployment benefits (UB) and issues sanctions if the unemployed don't comply with the rules. The Public Employment Services (PES) are organized at the regional level. They are in charge of job search assistance, intermediation services and training of unemployed and employed workers.

In Belgium a worker is entitled to UI in two instances:[5] (i) after graduation from school conditional on a waiting period of 9 months;[6] (ii) after involuntary dismissal from a sufficiently long-lasting job.[7] In contrast to many other countries the benefit entitlement does not exhaust after a certain duration.[8] School-leavers are entitled to flat rate benefits while dismissed workers earn a gross replacement rate ranging between 40% and 60% of past earnings, which is bracketed by a floor and a cap. The benefit level depends on household type (head of household, cohabitant or single) and on unemployment duration for dismissed singles and cohabitants.

Sanctions could be imposed for not complying to administrative rules: (i) making a false declaration (e.g. with regards the household type or an undeclared employment relationship) or (ii) being unavailable for the labor market (not registered as a job searcher at the regional PES, not turning up at an appointment in the PES or UI agency, refusing a 'suitable' job offer or refusing job search assistance or participation in a training program, etc.). Before the reform in 2004, roughly 80% of the monitoring reports regarding availability concerned not turning up at an appointment (RVA, 2006, p. 72). Job search was not monitored at that time.

Job seekers may on their own initiative make free of charge use of the services provided by the regional PES. Aside of this, before the 2004 reform only the Flemish PES was following-up all registered job seekers systematically. Inspired by the first European guidelines for employment, job seekers were invited to a meeting with a counselor a few months[9] after the initial registration. Depending on the outcome of this meeting, an action plan was drawn up and followed up. In Brussels and Wallonia such systematic counseling was targeted to

[5]See www.onem.be for more information.

[6]This waiting period lasts only 6 months for those aged less than 18 years and 12 months for youth between 26 and 30.

[7]The length of the contribution period depends on age.

[8]There existed an exception for cohabitants living with a partner earning a sufficiently high income (Cockx and Ries, 2004). However, this scheme was abolished together with the 2004 reform.

[9]In 2004 this occurred after 3 to 4 months for unemployed younger than 25 and for the older after 9 months.

low-skilled youth aged less than 25 only.

2.2 The 2004 reform

The most essential feature of the 2004 reform was the introduction of a job search monitoring scheme. The monitoring procedure consists in several steps:[10] a notification and up to three face-to-face interviews. The notification is sent by mail and states that one is required to actively search for a job and to participate in any action proposed by the regional PES. Some examples of search methods are provided and it is clearly stated that one should collect written proofs of the undertaken search actions. The letter announces that one will be invited at the UI office to evaluate the undertaken actions and that this interview will take place at the moment that the unemployment duration attains 21 months, i.e. not earlier than eight months later (since the letter is sent after 13 months). The letter also mentions that one should contact the regional PES if one has not yet been individually counseled. Finally, a folder explaining the different steps of the monitoring procedure is enclosed.

Interviews last approximately half an hour. At each interview the undertaken search actions are evaluated. If search effort at the first meeting is deemed insufficient an action plan is drawn up. If at the next meeting four months later it is established that the worker does not fulfill the plan, a second, stricter action plan is imposed and benefits are temporarily (possibly partially) withdrawn. If, again four months later, at the third meeting the worker does not comply, benefits are completely withdrawn and the worker can regain entitlement only after being full-time employed during at least one year. After a positive evaluation at the first (second) meeting, a new sequence of meetings is scheduled 12 (16) months later. This contrasts quite starkly with the frequency of monitoring in many other countries: half of OECD countries require reporting of job search (in most cases) every two weeks or at least monthly (OECD, 2007).

Job search effort is evaluated on the basis of proofs delivered by the unemployed worker (copies of letters of application, registration in temporary help agencies, proofs of participation in selection procedures, etc.). Regulations don't specify, however, a minimum number of employer contacts to submit. Consequently, caseworkers have quite some discretion in the evaluation process. This discretion together with the fact that in the beginning phase of the reform no information was available on the evaluation practice could explain why such a relatively loose monitoring scheme from an international perspective could cause the important threat effects reported in the empirical analysis.

The 2004 reform did not only concern monitoring of job search. The supply of the services

[10] See Cockx et al. (2007) and Cockx and Dejemeppe (2007) for more details.

provided by the regional PES was significantly enhanced in all three regions. In Flanders this increase was essentially allocated to expand existing services. By contrast, in Brussels and Wallonia the PES gradually introduced a systematic follow-up scheme as the one that was already operating in Flanders. In addition, specific counseling was targeted on the population concerned by the monitoring scheme. Within two months of receipt of the aforementioned notification, unemployed benefit claimants are invited to a collective information meeting at the PES. At this meeting the monitoring procedure and the services of the PES are elucidated. Soon after this meeting an individual face-to-face interview takes place and, if necessary, an action plan is drawn up, which could result in job search assistance or training.

3 Literature review and theoretical predictions

According to standard job search theory[11] the threat of being sanctioned if job search effort is insufficient results in more search effort and thereby speeds up the transition to employment. At the same time, monitoring will make the worker less selective so that she may accept jobs of lower quality. These effects realize as soon as the worker is informed about the monitoring scheme in place.

Monitoring need not always enhance the job finding rate, however. First, van den Berg and van der Klaauw (2006) argue that only *formal* proofs of job search can be verified. In an extended job search model they demonstrate that monitoring may then substitute informal by formal job search, so that the total job search intensity may hardly change or even decline. This effect would, however, be less relevant for the long-term unemployed, on whom we focus in this research, since the informal search channels are likely to have "dried up" after a long spell of inactivity. Second, Manning (2009) shows that the imposition of stricter job search requirements does not enhance search incentives for all workers, since, if behavior is followed-up too closely, workers may find it too onerous to continue claiming benefits and may withdraw from the labor force instead. However, this argument must be qualified here. Since it's costly to forgo the UB, it is unlikely that such withdrawals occur before the moment at which the monitoring interview takes place, unless the unemployed can qualify for other benefits of comparable level. In Belgium, there is an incentive to transit to sickness insurance (and to disability insurance subsequently, after 6 months), since the benefit level corresponds to that of UI. Entitlement to sickness benefits requires a doctor's certificate and is, in principle, verified by an independent physician.

If monitoring is combined with counseling and if the latter is effective, the job arrival rate will increase even further and the impact on the job finding rate will be even stronger.[12]

[11] See e.g. Mortensen (1986).

[12] In principle the job finding rate could decrease if the indirect effect via the effect on the reservation wage

In contrast, unless workers dislike it (Black et al., 2003), counseling raises the reservation wage, so that the net effect of the combined program on the quality of employment is indeterminate. In addition, counseling may enhance transitions into training and discourage withdrawals from the labor force.

The U.K. Restart unemployment program introduced in April 1987 shares a number of the features of the monitoring scheme introduced in Belgium: (i) it is targeted to long-term unemployed workers although of shorter duration (6 months); (ii) it starts with a dispatch of a letter; (iii) it continues with regular interviews (every 6 months); (iv) as in Wallonia (but not in Flanders), it consists of a combination of counseling and tighter enforcement of eligibility conditions. Dolton and O'Neill (1996, 2002) evaluate this program. On the basis of experimental evidence they conclude the following: (i) the program reduces the time spent in unemployment in the short run for both men and women, but in the long run only for men; (ii) the threat of monitoring is responsible for the short run effect, whereas counseling for the long-run effect; (iii) the program is more likely to induce men to leave unemployment for a job, while for women it encourages retreat from the labor force. Some of these conclusions are confirmed in our empirical analysis.

Other researchers[13] report that schemes that combine tighter enforcement with counseling stimulate the transition from unemployment to work, but could not identify which of the two components contributed to this success. Researchers who have studied the impact of job search monitoring in isolation mostly find that the duration of benefit claim is reduced,[14] but not necessarily that transitions to employment increase, as predicted by basic job search theory. Borland and Tseng (2007) find that enhanced monitoring of job search in Australia significantly increases exits from paid unemployment, but their study could not identify the exit destination. Studies of Johnson and Klepinger (1994) in the U.S. and McVicar (2008) in Northern Ireland demonstrate that monitoring increases transitions to work. According to Johnson and Klepinger this effect is not long-lasting, however, since after one year the probability of employment is no longer significantly different between treated and controls. In contrast, Klepinger et al. (2002) report that tightening work-search requirements slightly decreases employment and earnings. This is in line with the evidence of Petrongolo (2008) that enhanced requirements in the UK reduce transitions from unemployment to work, the number of weeks worked and earnings during about 3 years after job loss. Finally, only two

dominates. Under weak restrictions on the wage offer distribution this possibility is not relevant (van den Berg and van der Klaauw, 2006).

[13]See Meyer (1995) for a summary of experimental evidence in the US and Gorter and Kalb (1996) and van den Berg and van der Klaauw (2006) in the Netherlands.

[14]Ashenfelter et al. (2005) find insignificant effects, but Klepinger et al. (2002, p.19) claim that the discrepancy with their conclusions is essentially caused by the smaller sample size of Ashenfelter et al.'s study, resulting in less precision.

studies document the effect on withdrawals from the labor market. Petrongolo reports that the UK reform increased transitions to incapacity benefits. McVicar also finds evidence that the suspension of monitoring reduces transitions to 'other benefits'. However, at the same time workers are more likely to enter training and education.

The delay in the monitoring procedure implies that the imposed job search requirements may affect the behavior of the benefit claimants before the first interview takes place. Such anticipation effects have been observed and studied before, but not, to our knowledge, the threat of *pure* monitoring of job search. On the one hand there is an extensive literature on the anticipation of the expiration of UI entitlement.[15] On the other hand there is a growing literature that studies the threat effects of (potential) mandatory participation to active labor market programs (ALMP) or of a pending sanction.[16] Both strands of this literature have shown that such threat effects may form a major share of the total impact of these policies on the return to employment. Effects on the quality of employment is mixed. Card et al. (2007), van Ours and Vodopivec (2006) and Lalive et al. (2005) studying the threat effect of benefit expiration don't find any significant effect on wages or subsequent employment duration. In contrast, a recent study of Caliendo et al. (2009) reveals that workers who leave unemployment close to the moment of benefit expiration earn significantly lower wages and experience shorter employment spells, and Arni et al. (2009) report that a warning of a pending sanction reduces earnings, but not employment duration. Finally, the threat effect on withdrawals from the labor force has hardly been studied. Arni et al. (2009) report that announcement of a sanction nearly doubles the exit out of the labor force, but this exit is often only temporary.

4 The Data

We exploit administrative data from several sources: (i) the federal UI agency for monthly information on UB claims, the new monitoring procedure and the return to regular education; (ii) the regional PES for participation in training and job search assistance provided to the unemployed; (iii) the Crossroads Bank for Social Security[17] which matches the aforementioned information to records of all federal Social Security institutions. These matched data allowed us to construct monthly indicators of employment (including self-employment) and starting wage for salaried employment, sickness insurance claims and a residual state (i.e. neither being employed or UI claimant). This information is available for all sampled indi-

[15] See e.g. Fredriksson and Holmlund (2006) and Boone and van Ours (2009).

[16] See Black et al. (2003), Geerdsen (2006), Geerdsen and Holm (2007), Rosholm and Svarer (2008), Graversen and van Ours (2008) and Rosholm (2008). Arni et al. (2009) study both the ex post effect of a warning of a pending sanction and the ex ante effect of sanctions.

[17] See www.ksz-bcss.fgov.be/

viduals from January 2001 until the end of 2006.

4.1 Sample selection criteria and descriptive statistics

The sample contains *all* claimants of UB, who on the 1^{st} of July 2004 were between 25 and 34 years old and who became exactly 13 months entitled to UI between May 1 and August 31 of the same year. For administrative reasons, the duration criterion (13 months) is determined on the basis of payments made two months before the notification is (theoretically) sent.[18] To estimate the threat effect of monitoring, we retain only those workers claiming UI for at least one day during the second month after which unemployment duration attained 13 months, since only to these workers (aged below 30) a notification was sent. This dispatch occurs between July and October 2004.[19]

In principle all sampled individuals younger than 30 should receive the notification. However, as a consequence of the discrepancy between the moment of data selection (April 2006) and the actual dispatch of the notification (starting in June 2004), a small fraction[20] of these workers is actually not notified in the month in which they should have been according to regulations.[21] As a consequence, the DD design is "fuzzy". In section 5 we briefly discuss the methodological implications.

Table 1 reports descriptive statistics of the sampled population theoretically eligible for the dispatch of the notification letter. In accordance with the subsequent analysis the statistics are reported separately for Flanders and Wallonia.[22] For each of these regions the first column refers to the sample of unemployed workers between 25 and 29 years old (the "treated") and the second to those aged between 30 and 34 (the "controls"). The data inform about the following characteristics of the unemployed workers: the starting date of the observation window corresponding to (potential) dispatch of the notification letter (July, August, September or October 2004), the age reported in years (but measured in months) on the 1^{st} of July 2004, the gender, the nationality, the level of education, the household-type determining the UB level (head of household, single or cohabitant), the type of entitlement (school-leaver or dismissed), an indicator of recent participation in training (including a return to regular education and job search assistance) and of recent employment experience, both during the year before sample selection, and the unemployment rate by district of living. Time-varying

[18]This sample allows therefore to check whether claimants anticipate the notification: in that case the exit rate during these two months would differ between the treatment and control group (see Section 6.1.2).

[19]To estimate a placebo treatment effect (see Section 6.1.2), we also selected a sample according to exactly the same sample selection criteria one year earlier, in 2003.

[20]This fraction is 15%, but in most cases the letter is dispatched a few months later.

[21]This is essentially a consequence of *ex post* rectifications in the administrative data (Cockx and Dejemeppe, 2007).

[22]Recall that for reasons of too small a sample size, we did not include Brussels in the analysis.

Table 1: Descriptive Statistics by Region of Living and Age Group

	Flanders		Wallonia	
	25-29 years	**30-34 years**	**25-29 years**	**30-34 years**
Number of individuals	1 311	1 165	1 310	1 069
Month of (potential) notification				
July	28.4%	28.8%	27.2%	28.9%
August	25.3%	24.0%	20.8%	22.5%
October	22.7%	24.5%	23.7%	26.0%
November	23.7%	22.8%	28.2%	22.6%
Age				
mean age in years on July 1, 2004	26.9	32.0	26.8	32.0
(standard deviation)	(1.4)	(1.4)	(1.4)	(1.4)
Sex				
women	51.9%	53.1%	45.9%	46.3%
Nationality				
Belgian	90.1%	86.8%	90.4%	85.5%
EU15 (excluding Belgian)	2.8%	5.1%	7.3%	10.1%
others	7.1%	8.2%	2.3%	4.4%
Schooling level				
primary	14.8%	22.3%	14.0%	20.4%
lower secondary	19.6%	17.7%	19.5%	21.6%
upper secondary	45.0%	42.0%	41.2%	35.2%
higher education	20.5%	17.8%	22.8%	15.1%
other studies	0.1%	0.3%	2.6%	7.8%
Category of insured unemployment[(a)]				
head of household	14.8%	23.4%	19.5%	25.9%
single	24.3%	22.8%	27.5%	24.0%
cohabitant	61.0%	53.7%	53.0%	50.1%
Type of entitlement to benefits[(a)]				
work experience	83.0%	98.9%	74.2%	97.9%
school-leaver	17.0%	1.1%	25.8%	2.1%
Recent participation in training[(a) (b)]	18.0%	19.1%	13.4%	11.8%
Recent work experience[(a) (b)]	53.6%	42.8%	42.6%	31.2%
mean number of days in employment	87.3	80.7	93.1	78.2
(standard deviation)	(71.6)	(70.9)	(76.4)	(71.2)
Mean unemployment rate[(c)] by district	8.3%	8.2%	21.7%	21.0%
of living (standard deviation)	(1.8)	(1.8)	(5.2)	(5.4)

[(a)] At the sampling date, i.e. 2 months prior to the (potential) dispatch of the notification letter.
[(b)] During the year before the sampling date.
[(c)] The ratio of the number of benefit claimants available for the labor market and the number of workers insured against unemployment, by district of living on May 31 (ONEM, 2004, p.65).

variables are evaluated at the sampling date, i.e. two months prior to the (potential) dispatch of the notification letter.

The composition of the population varies across both, regions and age groups. This matches expectations. Younger workers e.g. are generally more educated and obviously more entitled to benefits as school-leaver. Note that this relation between observed characteristics and age does not compromise the evaluation as long as it does not display a discontinuity at the age of 30 (see Section 6.1.2).

Between 31% and 54% of the long-term unemployed workers has been employed during the year prior to the selection date. This apparent contradiction is related to the administrative definition of unemployment duration within the monitoring procedure. In this definition the duration counter[23] is reset to zero only if the worker has been 12 months full time employed within the preceding 15 months. Consequently, the sample does not only contain genuinely long-term unemployed workers, but also workers with unstable labor market careers. In the empirical analysis below it will be revealed that in Wallonia the treatment effect differs importantly between these two groups.

Finally, the last line of Table 1 reveals the major divide in the labor market conditions between the northern and southern regions in Belgium. Even if the unemployment rates reported in Table 1 are biased upwards,[24] statistics based on the standard ILO definition confirm this divide. According to the ILO definition 5.5% of the Flemish and 12.1% of the Walloon labor force was unemployed in 2004.[25]

4.2 The outcome variables

The outcome variables are measured at the end of each month following the (potential) dispatch of notification up to the eighth month. The benchmark outcome is a discrete indicator measuring whether any job has been found since the notification. More precisely, the indicator is set to one as soon as, within a month, one does not claim any UB as full-time unemployed worker and is officially registered as a salaried or self-employed worker. In order to measure the quality of employment, we use two indicators: the gross full-time equivalent (FTE) monthly wage and the employment duration in months. Since the wage is only known for *salaried* employment, we measure the quality on this subset only. Note that we measure employment duration and not job tenure.

Three additional outcome variables are analyzed. The first is an indicator of participation in training. This indicator comprises not only training *strictu sensu*, but also job search assistance and re-enrollment in regular education. The second outcome variable is an indicator of withdrawal from the labor force, defined as the residual state. Finally, among withdrawals from the labor force, we distinguish those to sickness insurance.

[23]The duration in months is obtained by dividing the number of claimed UB days by 26 and rounding down to the nearest integer.

[24]The unemployment figures reported in Table 1 are based on an administrative definition of unemployment in which the denominator is underestimated, since it excludes employment in public administrations. These are, however, the only available unemployment rate statisitics at the district level.

[25]Source: www.steunpuntwse.be/view/nl/18767

5 The Econometric Model

The empirical analysis aims at identifying the effect on the various outcome variables described in the previous section of a notification announcing that job search effort will be evaluated at least 8 months later. Identification is based on the discontinuity of the treatment at 30 years during the first year of the reform starting on July 1, 2004. DD analysis is by now well established in the economics literature (see e.g. Imbens and Lemieux, 2008; Lee and Lemieux, 2009). Nevertheless, we propose two contributions to this literature.

First, we consider a method that deals with a continuous forcing variable, age in this application, that is grouped into intervals. This is closely related to the method of Lee and Card (2008) for a discrete forcing variable. Lee and Card explain that this makes the method essentially parametric and that standard errors must be adjusted for the specification error that follows on. The specification error at each discrete value of the forcing variable induces correlation between observations with the same discrete value. Cluster-consistent standard errors tackle this problem. However, if the specification error differs between the actual and counterfactual treatment status at the (discrete) discontinuity threshold, then the standard errors must be further inflated for the additional uncertainty. We show that if the forcing variable is a grouped *continuous* variable (and the within-group distribution of the forcing variable is known) this additional correction is not required. Intuitively, this is because the underlying *continuous* forcing variable contains more information than a discrete one in that it identifies the exact location of the discontinuity point.[26] The treatment effect being exactly defined at this point is no longer affected by specification error.

Second, Lee and Card (2008) present a goodness-of-fit statistic to verify whether conventional least squares inference is appropriate. However, this statistic does not take the aforementioned specification error into account. We propose an adjustment that takes this into account. In addition, we propose a grouping procedure that aims at avoiding the poor small sample properties of this statistic. The test is related to the modified logit estimator proposed by Amemiya and Nold (1975) for grouped discrete outcomes.[27] Their estimation procedure could also be adapted to DD. This would yield a more efficient estimator. However, to avoid small sample bias, more observations within each grouping of the forcing variable than available in our data would be required.

In Section 4.1 we mentioned that the DD design is "fuzzy".[28] Nevertheless, throughout most

[26] The discontinuity point must be at the boundary of a grouping interval.

[27] Cockx and Ridder (2001), Cockx and Dejemeppe (2005) and Dejemeppe (2005) applied both, the estimation method and test, to grouped duration models.

[28] Note that the fuzzy design is particular in that the treatment is only available at one side of the discontinuity threshold: no individual older than 30 is assigned to the treatment. Battistin and Rettore (2008) show that in this

of this section we assume that the design is sharp. There are two reasons for this choice. First, model selection is simpler in the sharp case and, since the fraction of treated is never lower than 85%, ignoring the fuzzy nature of the design will only lead to a slight under-estimation of the treatment effect. Second, we don't present any methodological contibution related to the fuzziness of the design. In the presentation of the empirical results we will just report the two-stage least squares (TSLS) estimator for the preferred specifications.

We now proceed by deriving our econometric model in the presence of a grouped continuous forcing variable. In a second step, we propose the goodness-of-fit test with a specification error.

5.1 DD with a grouped continuous forcing variable

Consider a continuous forcing variable X (age in the empirical application) which is measured in deviation from x_0, the point at which the treatment status changes discontinuously: D(X)=1[X < 0].[29] In the data this forcing variable is grouped in $2J$ equally spaced (monthly) intervals: $[-J, -J+1), ..., [-1, 0), [0, 1), ...[j, j+1), ...[J-1, J)$. Let Y_k^1 and Y_k^0 denote the potential outcomes of being treated or not as measured in the k^{th} month after the assignment to the (counterfactual) treatment.[30] Assume that the expectations of these potential outcomes, conditional on the continuous forcing variable, take the following form: $E[Y_k^0|X=x] = \alpha_k + h_k^0(x)$ and $E[Y_k^1|X=x] = \alpha_k + \beta_k + h_k^1(x)$, where $h_k^0(.)$ and $h_k^1(.)$ are continuous functions, each determined by at most $J-1$ parameters,[31] and where we normalize $h_k^0(0) = h_k^1(0) = 0$. From these we define the average treatment effects (ATE) at the threshold of interest[32] as follows

$$E[Y_k^1 - Y_k^0|X=0] = \beta_k \tag{1}$$

It's important to see that, in contrast to Lee and Card (2008), the ATE does not depend on a specification error. The reason is that the underlying forcing variable X is continuous and not discrete, so that the exact location of the discontinuity point is identified and can be conditioned upon.

case identification assumptions of the sharp design are sufficient to identify the average treatment effect of the treated (ATT) at the discontinuity point.

[29] In the empirical application $x_0 = 360$ months.

[30] In the empirical analysis this evaluation occurs in the k^{th} month after the notification is (theoretically) dispatched.

[31] As a consequence of the grouping of the data, non-parametric identification of the treatment effect is not feasible (Lee and Card, 2008).

[32] As mentioned, we abstract from the fact that the DD is fuzzy. If this is taken into account, we can only identify the ATT, i.e. $E[Y_k^1 - Y_k^0|X=0, D=1]$.

Now assume that we approximate $h_k^0(.)$ and $h_k^1(.)$ by polynomials of order $P < 2(J-1)$:

$$h_k^d(x) = \sum_{p=1}^{P} x^p(\tilde{\gamma}_{kp}^0 + d\tilde{\gamma}_{kp}^1) + a_k^0(x) + d\left[a_k^1(x) - a_k^0(x)\right] \tag{2}$$

for $d \in \{0,1\}$ and where $a_k^0(x)$ and $a_k^1(x)$ are specification errors that indicate the degree to which the true functions $h_k^0(.)$ and $h_k^1(.)$ deviate from the polynomial function. Note $a_k^0(0) = a_k^1(0) = 0$ by the aforementioned normalization.

Denoting the observed outcome by $Y_k \equiv Y_k^0 + D(X)(Y_k^1 - Y_k^0)$, then, using (2), the regression equation, as a function of the continuous forcing variable, can be expressed as

$$E[Y_k|X=x] = \alpha_k + D(x)\beta_k + \sum_{p=1}^{P} x^p\left(\tilde{\gamma}_{kp}^0 + D(x)\tilde{\gamma}_{kp}^1\right) + a_k(x) \tag{3}$$

where $a_k(x) \equiv a_k^0(x) + D(x)[a_k^1(x) - a_k^0(x)]$.

In order to relate the grouped regression function to the continuous one as expressed in (3) we rely on the law of iterated expectations:

$$E[Y_k|j \leq X < j+1] = \int_j^{j+1} \left[E[Y_k|X=x]f(X|j \leq X < j+1)\right]dx \tag{4}$$

where $f(X|j \leq X < j+1)$ is the conditional density of the forcing variable, conditional on being in the interval $[j, j+1)$. Since the data are grouped by intervals, this density cannot be identified from the data. However, it is reasonable to assume that the within month distribution of birthdays of the sampled population is uniform.[33] If we denote individual observations by subscript i ($i = 1, 2, ..., N$), this assumption combined with (3) and (4) yields the following specification of the regression model for the micro data:

$$Y_{ijk} = \alpha_k + D(x)\beta_k + \sum_{p=1}^{P}\left[(j+1)^{(p+1)} - j^{(p+1)}\right]\left(\gamma_{kp}^0 + D_j\gamma_{kp}^1\right) + a_{jk} + \epsilon_{ijk} \tag{5}$$

where $\epsilon_{ijk} \equiv Y_{ijk} - E[Y_k|j \leq X_i < j+1]$, $\gamma_{kp}^d \equiv \tilde{\gamma}_{kp}^d/(p+1)$ for $d = 0,1$, $D_j \equiv \int_j^{j+1} D(x)dx = D(j)$, and $a_{jk} \equiv \int_j^{j+1} a_k(x)dx$. To ensure consistency of the least squares (LS) estimator, we

[33] Note that a misspecified density ($f(X|j-1 \leq X < j) = 1+u(X)$, where $u(X)$ is the specification error) may induce correlation between the residual, which includes the specification errors, and the included regressors and thereby affect the consistency of the least squares estimator. Gans and Leigh (2009) cite studies reporting that birth timing is not completely random: there are fewer births during week-ends, public holidays and during weeks at which obstetrics yearly conferences take place; more births on auspicious days and 7-10 days prior to Christmas. However, these days are not systematically related to particular days within a randomly selected month and therefore no cause for concern.

assume in addition that $E[a_{jk}|j \leq X_i < j+1] = 0, \forall j$.

In principle we can estimate this regression model separately for each k. However, since the sample size is relatively small, this yields imprecise estimates. In the empirical application, we therefore impose the restriction that the polynomials in age do not vary over k: $\gamma^0_{kp} = \gamma^0_p$ and $\gamma^1_{kp} = \gamma^1_p$, $\forall k$ and $\forall p$. As a consequence, the residuals are not only correlated between individuals i within age groups j through the specification error a_{jk}, but also across time k, both through a_{jk} and ϵ_{ijk}. Cluster-consistent standard errors at the group level j can take this correlation into account.

5.2 A goodness-of-fit test in the presence of specification error

In this section we propose a goodness-of-fit test in the presence of specification error. The idea is essentially that the average residual of a correctly specified model should be close to zero. This can be implemented as a m-test (White, 1994). Assume that the specification error a_{jk} is Normally distributed (with zero mean and constant variance) and independent between age groups.[34] If groups are sufficiently large ($n_{j.} \to \infty$, where $n_{j.}$ denotes the number of observations in group jk[35]), the sum of squared grouped residuals weighted by the inverse of the variance matrix can be shown to converge, under the null hypothesis of correct specification, to a χ^2 distribution with $2JK - M$ degrees of freedom. K is the number of evaluation months after treatment, so that $2JK$ is the number of groups. M is the number of estimated parameters in the regression model. This χ^2-statistic forms the basis of our goodness-of-fit test.

More formally, let us denote the grouped residual by $\bar{e}_{jk} \equiv a_{jk} + \bar{\epsilon}_{ijk}$, where $\bar{\epsilon}_{jk} \equiv \frac{1}{n_{j.}} \sum_{i=1}^{n_{j.}} \epsilon_{ijk}$, and $\bar{e}_j \equiv [\bar{e}_{j1} ... \bar{e}_{jk} ... \bar{e}_{jK}]'$ then we have that

$$\sum_{j=1}^{J} \bar{e}_j' \hat{V}_j^{-1} \bar{e}_j \quad \sim \quad \chi^2(2JK - M) \tag{6}$$

where $\hat{V}_j$ denotes the estimated variance-covariance matrix of $\bar{e}_j$. To calculate this statistic we need to construct $\hat{V}_j$. How this can be done is shown in the Appendix.

A problem with this goodness-of-fit statistic is that it is known to behave poorly in small samples (in terms of $n_{j.}$). Given that in our data the average group size is only just more than 20 units, this problem is pertinent in this empirical application. To avoid this problem we therefore group the data over wider age intervals, but only for testing, not for estimation.

[34] The specification error can be correlated between k's and between i's within j's.

[35] Note that $n_{j.} = n_{j1} = ... = n_{jk} = ... = n_{jK}$: in age group j the sample contains the same number of individuals in each month k after dispatch of the notification.

Since we consider only discrete binary outcomes, we use rules of thumb in the literature (Box et al., 1978; Johnson et al., 2005) that ensure that the Normal approximation of the binomial distribution is adequate. We enlarge the grouping until the three following conditions are all satisfied:[36] (i) $n_{j.}\bar{Y}_{jk}(1-\bar{Y}_{jk}) > 9$; $\left|1/\sqrt{n_{j.}}\left[\sqrt{(1-\bar{Y}_{jk})/\bar{Y}_{jk}} - \sqrt{\bar{Y}_{jk}/(1-\bar{Y}_{jk})}\right]\right| < 0.3$; (iii) $n_{j.}\bar{Y}_{jk} \pm 3\sqrt{n_{j.}\bar{Y}_{jk}(1-\bar{Y}_{jk})} \in [0, n_{j.}]$. Note that these conditions require a sufficiently large number of "successes"/"failures" ($n_{j.}\bar{Y}_{jk}/n_{j.}(1-\bar{Y}_{jk})$) in the outcome variable. Since right after the (theoretical) dispatch of the letter only few individuals have a successful outcome (found a job for the benchmark outcome), we consider only the effects on the outcome for $k \in [2, 8]$: we leave out $k = 1$, so that $K = 7$.

We use this goodness-of-fit test to select the appropriate (i) degree of the polynomial and (ii) window width. For each window width, we estimate 3 spline polynomials of degree zero up to two. We then choose within each window the polynomial that fits best according to the P-value of the goodness-of-fit tests. Subsequently, the trade-off between bias and precision is realized by retaining the widest window for which the model with the preferred polynomial is not rejected at the conventional level of 5%.

The aforementioned testing procedure is related to one proposed by Lee and Lemieux (2009, pp. 44-45; 54) and Lee and Card (2008, p. 658). They suggest to add $J-2$ (in our case $2K(J-1)$) bin dummies to the polynomial in the forcing variable and test whether the coefficients of these bin dummies are jointly different from zero. In the case of a grouped forcing variable, the model augmented with bin dummies is saturated at the grouped level. The test then boils down to a goodness-of-fit test of the polynomial against the saturated model. This corresponds exactly to the test we propose, except that we explicitly allow for a specification error and that we group the data further to ensure that the distributional assumptions of the test are satisfied.

Lee and Lemieux (2009) argue that this test can be interpreted as a falsification test of the DD, since a rejection of the test is evidence for discontinuities at other thresholds of the forcing variable. Here the model may be rejected for another reason: the restriction that the polynomials in age are constant over k may not be satisfied. The falsification test then consists in testing the polynomial regression model against one in which the polynomial in age is replaced by a set of indicators for each age group j, but common for each k: the difference of the goodness-of-fit statistics between these models is χ^2 distributed with $2(J-P)$ degrees of freedom. In fact we never reject any of the models on the basis of this statistic.

[36] In the few cases we could not fulfill this condition, we explicitly report this.

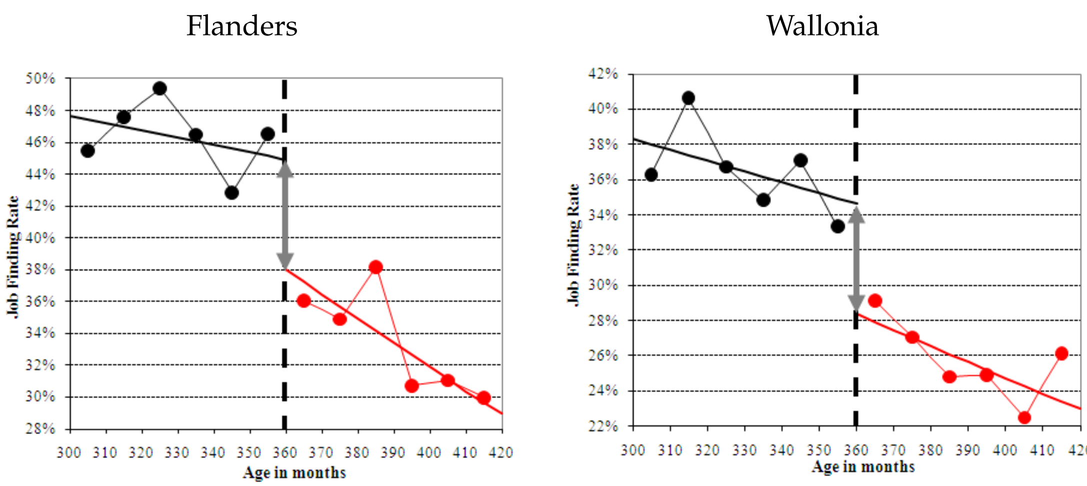
6 The Empirical Findings

6.1 The effect of notification on the job finding rate

6.1.1 The average treatment effects and sensitivity analyses

Figure 1 displays descriptive evidence of the effect of notification eight months after dispatch. The average transition rate in bins with a bandwidth of 10 months is plotted in a window of 5 years on either side of the cutoff point of 360 months. This bandwidth corresponds to the finest grouping for which the Normal approximation to the binomial distribution is adequate (see Section 5.2). A narrower bandwidth would make graphical evidence too noisy.

Figure 1: Discontinuity of the Job Finding Rate 8 Months after Notification



Both in Flanders and in Wallonia the job finding rate is declining with age. At 360 months the job finding rate declines more sharply, somewhat more clearly in Flanders than in Wallonia. This is evidence that the notification sent to the unemployed less than 30 years old speeds up the transition to employment.

Table 2 reports the estimation results of linear regressions of the outcome variable on various spline polynomials up to a second degree ($P = 2$) for sub-samples selected within an increasingly narrow age window around the cutoff point. Although all models estimate simultaneously the impact between $k = 2$ and $k = 8$ months after notification, we only report the results at $k = 8$.[37] Since the goodness-of-fit test requires a grouping of 10-monthly intervals, we consider 6 discontinuity samples: $DS \pm 1 * 10$-$DS \pm 6 * 10$. The discontinuity sample $DS \pm 1 * 10$ consists of the treated within the age bracket [350, 359] months and the control

[37] In both regions the treatment effect increases gradually over time. The time profile is available on request.

group in the bracket [360, 369]. The other discontinuity samples are defined accordingly.

Table 2: The Effect on the Job Finding Rate 8 Months after Notification (β_8) for Various Windows and Polynomials

Window width	$DS \pm 6 * 10$	$DS \pm 5 * 10$	$DS \pm 4 * 10$	$DS \pm 3 * 10$	$DS \pm 2 * 10$	$DS \pm 1 * 10$
Region			**Flanders**			
Polynom. order						
Zero	0.130** (0.022)	0.124** (0.023)	**0.099**** (0.024)	0.073** (0.028)	0.084** (0.038)	0.100** (0.043)
	[0.000]	[0.000]	[0.418]	[0.165]	[0.898]	[na]
One	0.069** (0.035)	0.059 (0.037)	0.041 (0.042)	0.073 (0.049)	0.116** (0.054)	0.193** (0.083)
	[0.000]	[0.000]	[0.080]	[0.129]	[0.902]	[na]
Two	0.085* (0.048)	0.098* (0.053)	0.134** (0.054)	0.185** (0.057)	0.214** (0.077)	0.361** (0.122)
	[0.000]	[0.000]	[0.090]	[0.017]	[0.788]	[na]
Optimal order[(a)]	na	na	0	0	1	na
# indiv. = N	2.476	2.045	1.592	1.176	769	391
Region			**Wallonia**			
Polynom. order						
Zero	0.109** (0.018)	0.105** (0.021)	0.081** (0.023)	0.068** (0.026)	0.051* (0.028)	0.049 (0.048)
	[0.881]	[0.969]	[0.987]	[0.949]	[0.752]	[na]
One	**0.062**** (0.031)	0.032 (0.034)	0.040 (0.038)	0.029 (0.049)	0.004 (0.057)	0.019 (0.071)
	[0.898]	[0.993]	[0.983]	[0.949]	[0.724]	[na]
Two	0.015 (0.047)	0.036 (0.053)	0.013 (0.062)	-0.004 (0.067)	-0.013 (0.067)	0.123* (0.064)
	[0.867]	[0.985]	[0.968]	[0.751]	[0.552]	[na]
Optimal order[(a)]	1	1	0	0	0	na
# indiv. = N	2.379	1.953	1.513	1.127	752	372

[(a)] Order of the model with the highest P-value according to the goodness-of-fit test.
Standard errors in parentheses. P-values of the goodness-of-fit test in square brackets. na = "not available". In bold the preferred estimate: model with optimal polynomial in the widest window not rejected at a P-value of 5%.
* Significant at the 10% level; ** Significant at the 5% level.

The goodness-of-fit test rejects for Flanders all models within the two widest windows ($DS \pm 5 * 10$ and $DS \pm 6 * 10$). Since the falsification tests described at the end of the previous section reject none of these models,[38] this rejection is no proof of discontinuity at other thresholds of the forcing variable, but rather of non-constant polynomials in age. This is confirmed if we re-estimate the models for these age windows separately for each k. In that case we can always find a spline polynomial with $P \leq 2$ for which the model is not rejected. However, the treatment effects of the preferred specifications are very imprecisely estimated and never significantly different from zero.[39] We therefore prefer the estimates for narrower windows in Flanders.

For the narrower windows we always find a polynomial for which the goodness-of-fit test is not rejected in Flanders.[40] The optimal polynomial degree (in terms of the P-value of the

[38]The lowest P-value of the X^2-tests of the 6 considered models is 0.23.

[39]For $DS \pm 6 * 10$ and $DS \pm 5 * 10$ the retained polynomial degree, P, is respectively 1 and 2, and the treatment effect (standard error) at $k = 8$ months 0.042 (0.042) and 0.103 (0.065).

[40]Note that for $DS \pm 1 * 10$ we cannot calculate the P-value of the goodness-of-fit statistic, since, by aggregation, the number of estimated parameters exceeds the number of grouped data points. In this case it is reasonable to choose $P = 0$.

goodness-of-fit) ranges between zero and one and the corresponding treatment effects between 7.3 and 11.6 percentage points and is always significant at the 5% level. Our preferred estimate on which we will perform a number of additional sensitivity analyses is the 9.9 percentage points effect of $DS \pm 4 * 10$. This corresponds to a proportional increase of the transition to employment of 28%.

For Wallonia the goodness-of-fit test rejects none of the estimated models. The effects corresponding to the models with optimal polynomial range between 2.9 and 8.1 percentage points and are not always significantly different from zero. So, the findings for Wallonia tend to be somewhat lower and less robust than for Flanders. The preferred estimate for $DS \pm 6 * 10$ is 6.2 percentage points or 22% in proportional terms.

Table 3: The Effect on the Job Finding Rate 8 Months after Notification (β_8): Controlling for Covariates and TSLS

	Flanders			**Wallonia**		
	Benchmark	**with X**	**TSLS**	**Benchmark**	**with X**	**TSLS**
$\alpha_8{}^{(a)}$	0.351**	0.351**	0.351**	0.284**	0.284**	0.284**
(standard error)	(0.018)	(0.018)	(0.018)	(0.024)	(0.024)	(0.024)
$\beta_8{}^{(b)}$	0.099**	0.096**	0.110**	0.062**	0.069**	0.067**
(standard error)	(0.024)	(0.022)	(0.027)	(0.031)	(0.031)	(0.033)
Order of the polynomial		0			1	
Window width		$DS \pm 4 * 10$			$DS \pm 6 * 10$	

(a) The job finding rate without notification for 30 year olds.
(b) The ATE for 30 year olds.
* Significant at the 10% level; ** Significant at the 5% level.

Table 3 reports two further sensitivity analyses for the preferred estimates: (i) adding in the control variables reported in Table 1; (ii) explicitly accounting for the "fuzzy" nature of the DD by implementing a TSLS estimator. In line with expectations, the treatment effects are hardly affected.

6.1.2 Testing the validity of the DD

In the previous section we already discussed that there is no evidence for discontinuities at other thresholds of the forcing variable. In this section an additional falsification test is presented. We check for the presence of a placebo treatment effect in the year prior to the reform. Furthermore, we will test for *precise* manipulation of the DD.

Table 4 reports, for the preferred estimates, the effect of a placebo treatment on a sample of unemployed workers aged between 25 and 34 years old to whom a notification would have been sent if the reform was in place in 2003. The estimated impact in both regions is only 2 percentage points and not significantly different from zero. In the sensitivity analysis (not reported) only one of the effects, not rejected on the basis of the goodness-of-fit test, was

Table 4: Placebo Treatment Effect in 2003

	Flanders		**Wallonia**	
	Benchmark	**2003**	**Benchmark**	**2003**
α_8 [(a)]	0.351**	0.336**	0.284**	0.253**
(standard error)	(0.018)	(0.015)	(0.024)	(0.022)
β_8 [(b)]	0.099**	0.018	0.062**	0.020
(standard error)	(0.024)	(0.021)	(0.031)	(0.028)
Order of the polynomial	0		1	
Window width	$DS \pm 4 * 10$		$DS \pm 6 * 10$	

[(a)] The job finding rate without notification for 30 year olds.
[(b)] The ATE for 30 year olds.
* Significant at the 10% level; ** Significant at the 5% level.

significant.[41]

The DD approach is only valid to the extent that individuals cannot alter their behavior - "manipulate the forcing variable" - to avoid (or to benefit from) the treatment, i.e. “to *precisely* “sort” around the discontinuity threshold” (Lee, 2008; Lee and Lemieux, 2009). In this study the forcing variable, age, can not be manipulated directly, but indirectly, since assignment to the treatment is not solely based on age, but also on being a benefit claimant for 13 months or more. This unemployment duration can be manipulated if the unemployed workers anticipate the dispatch of the notification and leave unemployment even before the notification. Note, however, as stressed by Lee (2008), that this manipulation invalidates the DD only to the extent that workers have a *precise* control over exit from unemployment. For instance, if anticipation implies increased job search, this need not invalidate the DD, since the outcome of the job search process is random, so that the worker cannot avoid the notification with certainty. It would be more problematic if the worker would, in anticipation of the notification, decide to decline her entitlement to UB. However, it’s unlikely that a written notification raises the perceived marginal cost to that extent. Nevertheless, we test for the presence of this kind of precise manipulation.

Table 5: The Exit Rate from Unemployment One Month Prior to Notification

	Flanders	**Wallonia**
α_8 [(a)]	0.085**	0.060**
(standard error)	(0.008)	(0.011)
β_8 [(b)]	-0.001	-0.024
(standard error)	(0.012)	(0.016)
Order of the polynomial	0	1
Window width	$DS \pm 4 * 10$	$DS \pm 6 * 10$

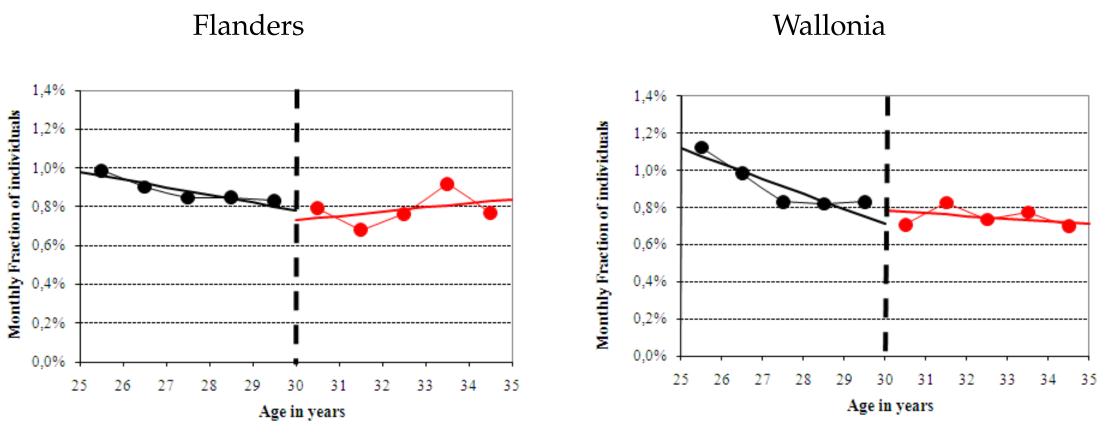
[(a)] The exit rate without notification for 30 year olds.
[(b)] The ATE for 30 year olds.
* Significant at the 10% level; ** Significant at the 5% level.

We implement three tests. First, since the sample is selected two months before notification (cf. Section 4.1), we can verify by DD whether workers slightly below the age of 30 years are more likely to leave unemployment (we include exit to training and from the labor force

[41]This occurs for $DS \pm 4 * 10$ and $P = 0$, but for $P > 0$ the effects are virtually zero.

in this definition) one month before notification than workers slightly older than 30. As can be deduced from Table 5, the exit rate from UB in the month prior to notification does not exhibit any significant discontinuity at 360 months. Second, we test whether the 18 predetermined characteristics (reported in Table 1) are smooth around the cutoff age. To do so, we follow the suggestion of Lee and Lemieux (2009) to estimate for the widest window size all 18 regression equations with a linear spline jointly in a Seemingly Unrelated Regression (SUR) and test for the joint significance of discontinuities at the cutoff point. The absence of a discontinuity cannot be rejected at a P-value of 18% for Flanders and 44% for Wallonia.[42] Finally, we test whether the density of the forcing variable is continuous around the threshold (McCrary, 2008). This density test is graphically represented in the two panels of Figure 2. From this figure, but also from the formal tests (not reported), there is no evidence for any discontinuity.

Figure 2: Graphical Density Test of the Forcing Variable



6.2 Heterogeneity of the effect on the job finding rate

In this section we briefly investigate whether the effect on the transition to employment differs importantly between sub-populations. In view of the relatively small sample size, the analysis of treatment heterogeneity is limited. Moreover, the results are less precise and robust than for the main analysis. They should therefore be interpreted with some caution.[43]

To avoid the multiple comparisons problem, the choice of the heterogeneity dimensions is based on a priori grounds. First, since for most households it is still the case that the respon-

[42]For Flanders 3 variables were individually significant at the 5% level: lower secondary school, school-leaver and numbers of days employed during the year before the sampling date. For Wallonia none of the variables were individually significant.

[43]The sensitivity analysis can be obtained from the authors on request. Note, to avoid small sample problems in the goodness-of-test statistic, the data had to be grouped over wider age intervals: the width of grouping intervals varies between 12 and 20 months, depending on region and sub-population. In one case (the low educated $DS \pm 3 * 20$ in Wallonia) the grouping criteria of Section 5.2 were not satisfied for one cell.

sibility for child and elderly care falls disproportionately on the female partner, we expect, as reported by Dolton and O'Neill (2002), that women are less likely to leave unemployment for a job than men. Second, we expect the threat effect of monitoring, combined with the effect of counseling in Wallonia, to be larger for workers with more favorable re-employability chances, such as the higher educated and those living in districts with relatively low unemployment rates: changes in job search behavior are more likely to result in a successful transition to employment. Finally, we expect smaller effects for those with a recent employment experience,[44] since these workers can more easily prove that they have searched and continue searching for jobs.

Table 6: Heterogeneity of the Treatment Effect

Region	Flanders							
Sub-populations	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
α_8 (a)	0.354**	0.403**	0.424**	0.307**	0.308**	0.440**	0.439**	0.255**
(standard error)	(0.027)	(0.046)	(0.024)	(0.031)	(0.023)	(0.036)	(0.024)	(0.019)
β_8 (b)	0.125**	0.004	0.062*	0.079*	0.124**	0.021	0.115**	0.104**
(standard error)	(0.035)	(0.058)	(0.037)	(0.040)	(0.033)	(0.052)	(0.030)	(0.027)
Optimal polynomial order (c)	0	1	0	0	0	1	0	0
P-value goodness-of-fit	0.166	0.095	0.367	0.096	0.323	0.807	0.548	0.578
P-value F equality test	0.091		0.749		0.103		0.777	
# indiv. = N	927	1 028	595	580	1 043	912	1 201	1 275
Window width (d)	$DS \pm 4 * 12$		$DS \pm 2 * 15$		$DS \pm 4 * 12$		$DS \pm 4 * 15$	

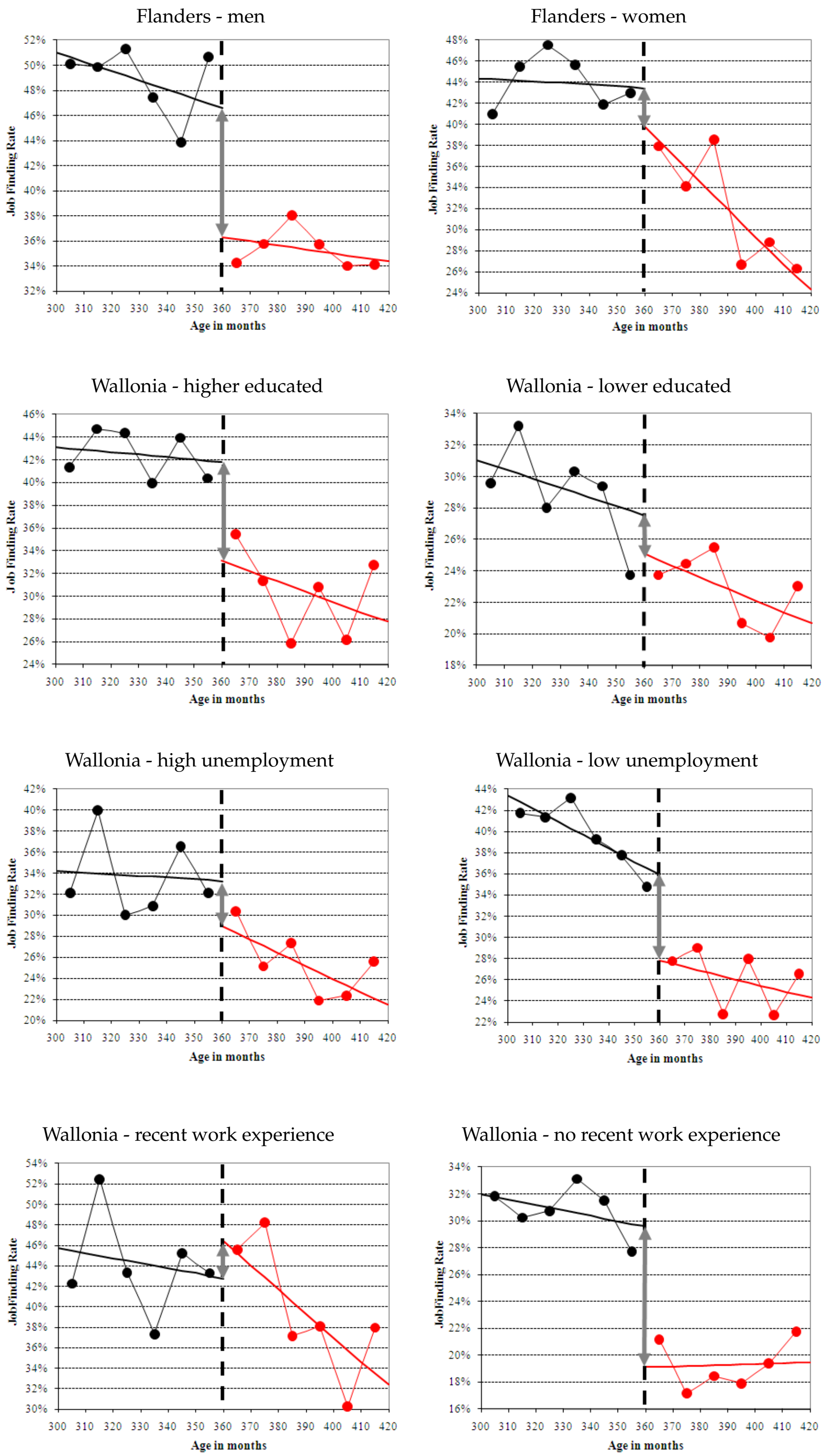
Region	Wallonia							
Sub-populations	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
α_8 (a)	0.265**	0.271**	0.306**	0.251**	0.265**	0.270**	0.466**	0.193**
(standard error)	(0.023)	(0.020)	(0.019)	(0.033)	(0.022)	(0.022)	(0.056)	(0.012)
β_8 (b)	0.101**	0.055*	0.119**	0.024	0.041	0.124**	-0.037	0.116**
(standard error)	(0.030)	(0.030)	(0.025)	(0.043)	(0.032)	(0.032)	(0.063)	(0.021)
Optimal polynomial order (c)	0	0	0	1	0	0	1	0
P-value goodness-of-fit	0.493	0.304	0.362	0.068	0.312	0.380	0.647	0.458
P-value F equality test	0.240		0.051		0.072		0.025	
# indiv. = N	833	680	1 177	1 085	764	749	891	1 488
Window width (d)	$DS \pm 2 * 20$		$DS \pm 3 * 20$		$DS \pm 2 * 20$		$DS \pm 3 * 20$	

(1) Men; (2) Women; (3) Higher education; (4) Lower education; (5) High (above regional median) unemployment rate; (6) Low (below median) unemployment rate; (7) Recent work experience; (8) No recent work experience.
* Significant at the 10% level; ** Significant at the 5% level.
(a) The job finding rate without notification for 30 year olds.
(b) The ATE for 30 year olds.
(c) Order of the model with the highest P-value according to the goodness-of-fit test.
(d) Widest window for which no sub-model is rejected at a P-value of 5%.

Table 6 displays the treatment effects for the various sub-populations in the two regions. We report the estimates of the models with "optimal" polynomial and with the widest age grouping for which the goodness-of-fit test does not reject the model for any of the two considered subpopulations (e.g. men and women) at a 5% level. The table includes the P-value of an F-test of the equality of the subgroup effects.

[44] "Recent" is within one year before the sample selection date.

Figure 3: Significant Heterogeneity of the Treatment Effect



If we set the P-value of this F-test at 10%, there is no evidence of treatment heterogeneity in Flanders, except for gender. In Wallonia, by contrast, the F-test is significant for all the variables, except for gender. Figure 3 illustrates the DD for the cases in which the heterogeneity is significant. The lack of heterogeneity in Flanders may be related to the following factors. First, in a tighter labor market, such as in Flanders, the impact of enhanced job search translates more easily in employment, irrespective of the employability of the workers. Second, in a tight labor market, the sorting process weeds out the most employable workers more rapidly than when the labor market is slack. This means that in a tight labor market there is less heterogenity in a population of long-term unemployed workers. Thirdly, since notified benefit claimants were only counseled in Wallonia, it could be that heterogeneity is induced by counseling. Counselors elucidate the monitoring procedure and, as a consequence, some groups might better understand that they are less at risk of a negative evaluation, such as workers with recent employment experience, but also low educated workers and workers living in high unemployment districts, since in the evaluation of job search effort it is taken into account that these factors make it more difficult to find job offers.

In both regions, the threat of monitoring has, as expected, more impact on men than on women. In Wallonia this difference is not significant, but this may be related to the imprecision of the estimates. In Wallonia, the differential impacts of all the other variables are significant and have the sign that was expected on a priori grounds. In Flanders, as mentioned and justified, no treatment heterogeneity for the other variables is significant.[45]

6.3 The effect on the quality of employment

Imposing job search requirements may speed up the transition to work, but at the same time there is a risk that the quality of employment is affected. Lower quality employment is defined as employment with a lower FTE starting wage (employment duration) than the median for a 30 year old worker.[46] Note, however, that the DD cannot identify the causal impact on quality. If the treatment induces more transitions to lower quality jobs, this can be either because the treatment causes employment to be of lower quality or because the treatment causes transitions to employment that would have been of lower quality anyway.

Table 7 summarizes the results of the preferred models according to the goodness-of-fit tests.[47] Since wages and employment durations are only observed for salaried employment,

[45] In Flanders the ATE increases (at the margin of significance) with the unemployment rate, but this finding is not robust (not reported).

[46] These medians are estimated on the control group (the 30-34 year olds) from a quantile regression with linear polynomial in age.

[47] A sensitivity analysis with respect to window widths and polynomial degrees is reported in Tables A-1-A-4 in Appendix A-2.

Table 7: The Effect on the Quality of Employment

Region	**Flanders**			
Outcome variable	Wage <50%[a]	Wage >=50%[a]	Duration < 50%[a]	Duration >=50%[a]
Outcome without treatment (α_8)	0.172	0.170	0.148	0.162
(standard error)	(0.013)	(0.018)	(0.013)	(0.012)
Treatment effect (β_8)	0.059**	0.000	0.083**	0.044**
(standard error)	(0.019)	(0.026)	(0.017)	(0.017)
Polynomial order[b]	0	1	0	0
P-value for goodness-of -fit	0.408	0.069	0.741	0.074
P-value for F equality test	0.101		0.145	
# indiv. = N	1 592	2 476	2 476	2 476
Window width[c]	$DS \pm 2 * 20$	$DS \pm 3 * 20$	$DS \pm 3 * 20$	$DS \pm 3 * 20$
Region	**Wallonia**			
Outcome without treatment (α_8)	0.115	0.130	0.110	0.137
(standard error)	(0.009)	(0.012)	(0.009)	(0.011)
Treatment effect (β_8)	0.051**	0.040**	0.071**	0.002
(standard error)	(0.013)	(0.017)	(0.014)	(0.015)
Polynomial order[b]	0	0	0	0
P-value for goodness-of-fit	0.252	0.277	0.861	0.328
P-value for F equality test	0.629		0.002	
# indiv. = N	2 379	1 513	2 379	1 513
Window width[c]	$DS \pm 3 * 20$	$DS \pm 2 * 20$	$DS \pm 3 * 20$	$DS \pm 2 * 20$

* Significant at the 10% level; ** Significant at the 5% level.
[a] Median FTE gross starting wage and median employment duration for 30 year olds in the absence of notification in Flanders (Wallonia) = 1964€ (1911€) and 10 (9.4) months.
[b] Order of the model with the highest P-value according to the goodness-of-fit test.
[c] Widest window in which the goodness-of-fit test is not rejected at a P-value of 5%.

we restricted the outcomes accordingly. The second and third column contain the policy impact on transitions below and above the median of the outcomes, wage and employment duration.[48] The P-value of the reported F-test allows to check for each outcome whether the effects on the transitions below and above the median are significantly different from each other.

First, it strikes that in Flanders only the effect on transitions to low-wage jobs is significant whereas in Wallonia the effect on both, transitions to low and high wage jobs is significant. This is consistent with the theoretical prediction (cf. Section 3) that the pure threat of monitoring reduces the reservation wage, but that this effect can be undone if it is combined with counseling, since this raises the reservation wage.

Second, in both regions transitions to short-term employment increase significantly and more than to long-term employment. In addition, in Flanders the effect on transitions to long-term employment differs significantly from zero. However, the latter effect is not at all robust if the window width is narrowed: the effect for other window widths is virtually

[48]The sum of the impacts for each outcome does not correspond to the total impact reported in Table 2 for two reasons: (i) Table 2 includes self-employment; (ii) the sum equals the total only if each sub-impact is estimated with the same window size and polynomial degree as for the total effect.

zero.[49] We therefore conclude that the evidence for Flanders suggests that a pure threat of monitoring lowers the quality of employment, both in terms of wages and employment duration.

Despite the presence of counseling in Wallonia no significant increase to long-term employment is observed. This may be related to the fact that counseling often results in participation in training (as documented in the next section): as a consequence of locking-in, transitions to more long-lasting employment may occur after the end of the observation period.

6.4 The effect on transitions into training, out of the labor force and into sickness insurance

Table 8 shows the results of the analysis on three other outcome variables: the transition into training, out of the labor force and into sickness insurance. It reports the treatment effects (slightly underestimated, since the fuzzy design is not taken into account) for the optimal polynomial and the window for which the model is not rejected on the basis of the goodness-of-fit statistic. For each of the outcome variables both sensitivity analyses and validity checks were performed as for the benchmark outcome (see Section 6.1).[50]

Since participation in training may positively affect the evaluation at the monitoring interview, we expect an increase in the participation rate. In Flanders, eight months after notification the participation rate is indeed 3.6 percentage points higher than without a notification. This effect is only significant at the 10% level, however. Systematic counseling for the notified group, as present in Wallonia, increases training participation dramatically by 7.7 percentage points. This is illustrated in Figure 4.

As argued in Section 3, we do not expect that the threat of monitoring should affect much the transitions out of the labor force except for transitions to sickness insurance, since workers are entitled to the same benefit level as in UI. In Flanders we don't observe any significant increase in the exit from the labor force. In Wallonia, by contrast, this exit rate increases significantly by 4.3 percentage points. If we decompose this effect, we find that this effect is virtually completely caused by a 6.4 percentage point increase in the transition rate of women to sickness insurance (see also Figure 5).

Why is this effect only present for women in Wallonia? As already mentioned, since women take disproportionately more responsibility for child and elderly care, they have more incen-

[49] See Table A-4 in the Appendix.

[50] The sensitivity analysis with respect to window widths and polynomial degrees is reported in Tables A-5-A-7 in Appendix A-2.

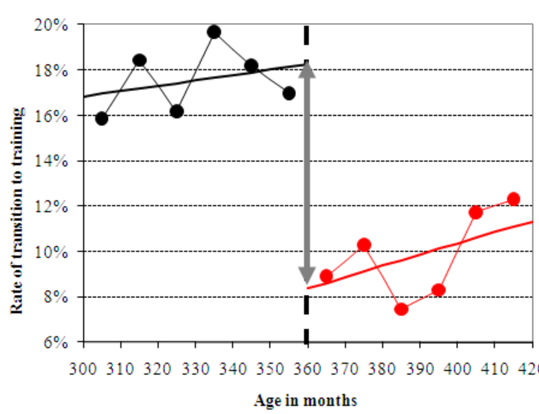
Table 8: The Effect on Transitions into Training, out of the Labor Force and into Sickness Insurance

Region	**Flanders**						
Outcome	Training	Out of Labor Force			Sickness Insurance		
		All	Men	Women	All	Men	Women
α_8 [(a)]	0.192	0.175	0.133	0.216	0.074	0.037	0.107
(standard error)	(0.015)	(0.014)	(0.020)	(0.026)	(0.009)	(0.009)	(0.014)
β_8 [(b)]	0.036*	0.010	0.013	0.001	0.012	-0.010	0.033
(standard error)	(0.021)	(0.025)	(0.030)	(0.038)	(0.012)	(0.010)	(0.021)
Pol. order[(c)]	0	1	1	1	0	0	0
P-value fit	0.854	0.750	-[(d)]	-[(d)]	0.467[(e)]	-[(d)]	-[(d)]
# indiv. = N	1 176	2 476	1 177	1 299	2 476	1 177	1 299
Window width[(f)]	$DS \pm 3 * 10$	$DS \pm 3 * 20$			$DS \pm 3 * 20$		

Region	**Wallonia**						
Outcome	Training	Out of Labor Force			Sickness Insurance		
		All	Men	Women	All	Men	Women
α_8 [(a)]	0.098	0.087	0.088	0.086	0.041	0.031	0.053
(standard error)	(0.009)	(0.012)	(0.016)	(0.019)	(0.006)	(0.007)	(0.010)
β_8 [(b)]	0.077**	0.043**	-0.012	0.114**	0.023**	-0.012	0.064**
(standard error)	(0.013)	(0.017)	(0.022)	(0.028)	(0.009)	(0.009)	(0.017)
Pol. order[(c)]	0	0	0	0	0	0	0
P-value fit	0.598	0.190[(e)]	-[(d)]	-[(d)]	0.314[(e)]	-[(d)]	-[(d)]
# indiv. = N	2 379	1 513	833	680	2 379	1 283	1 096
Window width[(f)]	$DS \pm 3 * 20$	$DS \pm 2 * 20$			$DS \pm 3 * 20$		

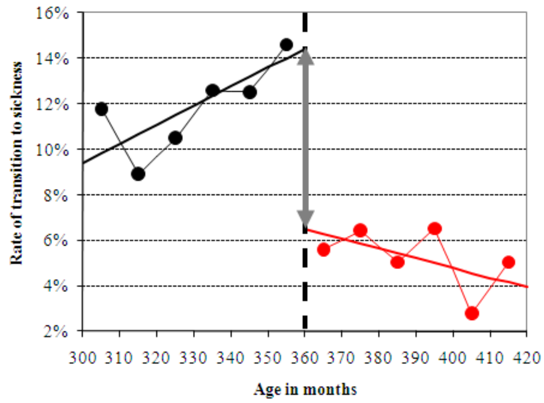
* Significant at the 10% level; ** Significant at the 5% level.
[(a)] The job finding rate without notification for 30 year olds.
[(b)] The ATE for 30 year olds.
[(c)] Order of the model with the highest P-value according to the goodness-of-fit test.
[(d)] Goodness-of-fit statistic completely unreliable because of too small sample size.
[(e)] Rules of thumb not satisfied for all groups: goodness-of-fit statistics are not reliable.
[(f)] Widest window in which the goodness-of-fit test is not rejected at a P-value of 5%.

Figure 4: The Effect on the Transition into Training in Wallonia



tives to find alternatives to working. Moreover, since this reason for not searching for jobs is socially more accepted for women than for men, counselors of PES may suggest this option to women and physicians might be more willing to prescribe sickness or disability, even if it is not strictly justified. Alternatively, long-term dependency on unemployment may affect the mental health of women. However, there is evidence that unemployment causes mental health problems only for men and not for women, at least shortly after job loss (Kuhn et

Figure 5: The Effect on the Transition into Sickness Insurance for Women in Wallonia



al., 2009).

Why is this not observed in Flanders? To answer this question, note first that in Flanders the transition rate to sickness insurance for 30 year olds is in the absence of notification nearly double as high as in Wallonia: 7.4% versus 4.1%. The notification therefore merely shifts the transition level in Wallonia to the Flemish level. This may be related to the fact that, irrespective of the notification, the unemployed in Flanders have more contacts with the regional PES than in Wallonia.[51] Thus, both Flemish and Walloon counselors might suggest to women the option of a transition to sickness benefits. Since the contact rate with the regional PES is in Wallonia much lower for those aged more than 30 than for the younger group, we observe this discontinuity in Wallonia, but not in Flanders.

7 Conclusion

This research studied the impact of an important reform in Belgium which introduced for the first time job search requirements into UI. More specifically, it investigated whether a notification sent to long-term unemployed benefit claimants had any impact on the job finding rate, on the quality of employment and on the rate of labor force withdrawal. The analysis was performed separately in two regions. In Wallonia the regional PES systematically starts counseling the unemployed two months after notification. In Flanders a *pure* threat effect of the notification is measured. The estimation is based on a DD at the age of 30 years that resulted from a gradual introduction of the reform by age group.

Eight months after the notification the transition rate to employment was roughly 10 and 6

[51] In the year preceding sample selection nearly 60% of the Flemish unemployed aged 25-29 had a contact with the regional PES (collectitive/individual meeting, training, etc.) against only 18% of the Walloon unemployed.

percentage points higher, respectively in Flanders and Wallonia. This corresponds to a proportional increase of 28% and 22%. The larger impact in Flanders is probably related to the more favorable labor market conditions in this region than in Wallonia. However, the presence of systematic counseling of the notified group in Wallonia may also play a role. On the one hand counseling may elucidate for some unemployed workers, such as those with a recent employment experience, living in high unemployment districts and the lower educated, that the threat of monitoring is less acute. On the other hand, it stimulates participation in training which, through a locking-in effect, may delay the transition to employment beyond the observation period.

However, consistent with theoretical predictions, in Flanders the benefit of the pure threat effect in terms of transitions to employment comes at a cost of the quality of employment, both in terms of the wage level and the duration of the employment relationship. In Wallonia counseling undoes the negative effect on the wage, but not on employment duration. The latter finding may, however, be a consequence of the too short observation period: workers may be locked into training programs and transit to more long-lasting employment afterwards.

Finally, in both regions the reform stimulates more the re-employment of men than of women. In addition, in Wallonia the policy induces women, but not men, to withdraw from the labor force. These women transit to sickness insurance which entitles them to equally large benefits as UI. This is consistent with the social norm that women are more responsible for caring and less obliged to search for jobs. Counselors and physicians of the sickness insurance may therefore be more lenient for women than for men. In Flanders no increase in the withdrawal rate from the labor force is observed. However, in Flanders the rate of transition to sickness insurance is already at the high post-reform level of Wallonia in the absence of notification. We argue that this is a consequence of the higher contact rate with the regional PES in Flanders irrespective of notification: counselors may suggest women the option of sickness insurance as a way to fulfill their caring duties without income loss.

Acknowledgements

We acknowledge funding by the Belgian federal administration Science Policy within the program "Society and Future" (contract no. TA/00/11) and the Interuniversity Poles of Attraction [Contract No. P6/P7]. We thank Bruno Crépon, Anne Defourny, Marc Gurgand, Michael Lechner, Thierry Magnac, William Parienté, Matteo Picchio, Jeffrey Smith, Sébastien Van Bellegem and Bruno van der Linden for discussions and comments.

Appendix

A-1 Construction of the variance matrix of the residuals

The variance matrix of the residuals $\hat{V}_j$ is constructed in three steps. First, if we assume that the "approximation errors", ϵ_{ijk}, are i.i.d. within groups jk, then a consistent (for $n_{j.} \to \infty$) estimate of the variance ($k = l$) and covariance ($k \neq l$) is given by[52]

$$\widehat{var}(\epsilon_{ijk}\epsilon_{ijl}) \equiv \hat{\sigma}_{\epsilon jkl} = \frac{1}{(n_{j.}-1)} \sum_{i=1}^{n_{j.}} (Y_{ijk} - \bar{Y}_{jk})(Y_{ijl} - \bar{Y}_{jl}) \equiv \frac{1}{(n_{j.}-1)} \sum_{i=1}^{n_{j.}} \hat{\epsilon}_{ijk}\hat{\epsilon}_{ijl} \tag{A-1}$$

where $\bar{Y}_{jk} \equiv \frac{1}{(n_{j.}-1)} \sum_{i=1}^{n_{j.}} Y_{ijk}$.

Second, let $\hat{\bar{e}}_{jk}$ denote the (unweighted) residuals of a weighted LS regression of model (5) in which the data are grouped by age group j and by month k and in which the weights are set to $n_{j.}/(N/J)$.[53] It is not difficult to show that these residuals are equivalent to the average residuals of the micro-regression: $\hat{\bar{e}}_{jk} = (\bar{Y}_{jk} - W_{jk}\hat{\theta}_k)$, where W_{jk} is the vector of explanatory variables in (5) and where $\hat{\theta}_k$ is the vector of parameter estimates of the unweighted OLS regression on micro-data. Since the residual is a sum of a specification and a grouped approximation error ($\bar{e}_{jk} = a_{jk} + \bar{\epsilon}_{jk}$), a consistent (for $J \to \infty$) estimator of variance and covariance of the specification errors a_{jk} can be found by subtracting from the weighted (co-)variance[54] of the grouped regression residuals $\hat{\bar{e}}_{jk}$ the weighted (co-)variance of the grouped approximation errors $\bar{\epsilon}_{jk}$:

$$\widehat{var}(a_{jk}a_{jl}) \equiv \hat{\sigma}_{akl} = \frac{1}{J} \sum_{j=1}^{J} \frac{n_{j.}}{(N/J)} \left[\hat{\bar{e}}_{jk}\hat{\bar{e}}_{jl} - \frac{\hat{\sigma}_{\epsilon jkl}}{n_{j.}} \right] = \frac{1}{N} \sum_{j=1}^{J} \left[n_{j.}\hat{\bar{e}}_{jk}\hat{\bar{e}}_{jl} - \hat{\sigma}_{\epsilon jkl} \right] \tag{A-2}$$

where we set $\hat{\sigma}_{akk} = 0$ if the right-hand side becomes negative and $\hat{\sigma}_{akl} = 0$ if $\hat{\sigma}_{akk} = 0$ or $\hat{\sigma}_{all} = 0$.

Finally, if $\hat{v}_{jkl}$ denotes the k^{th} row and l^{th} column of $\hat{V}_j$, we have

$$\hat{v}_{jkl} = \hat{\sigma}_{akl} + \frac{\hat{\sigma}_{\epsilon jkl}}{n_{j.}} \tag{A-3}$$

[52]If, as in the empirical application, the outcome variable is discrete indicator of a past transition to another labor market state (implying that the indicator cannot revert to zero once it has changed to one) it is not difficult to show that $\hat{\sigma}_{\epsilon jkl} = \bar{Y}_{jk}(1 - \bar{Y}_{jl})n_j/(n_j - 1)$, reflecting the binomial distribution of the discrete indicator.

[53]N/J is equal to the average group size.

[54]Note that, in contrast to $\hat{V}_j$ this (co-)variance is *un*conditional on j.

A-2 Sensitivity Analysis for Various Outcomes

Table A-1: The Effect (β_8) on Transitions to Low Wage Jobs

Window width	$DS \pm 3 * 20$	$DS \pm 2 * 20$	$DS \pm 1 * 20$
Region		**Flanders**	
Polynomial order			
Zero	0.077** (0.017) [0.001]	**0.059**** (0.019) [0.408]	0.069** (0.031) [na]
One	0.050* (0.028) [0.000]	0.046 (0.034) [0.201]	0.070 (0.046) [na]
Two	0.076* (0.039) [0.000]	0.092** (0.044) [0.085]	0.134** (0.060) [na]
Optimal order[(a)]	0	0	na
# indiv. = N	2 476	1 592	769
		Wallonia	
Zero	**0.051**** (0.013) [0.252]	0.031* (0.017) [0.218]	0.027 (0.022) [na]
One	0.029 (0.022) [0.222]	0.027 (0.027) [0.426]	0.005 (0.035) [na]
Two	0.021 (0.031) [0.223]	0.003 (0.037) [0.260]	-0.043 (0.036) [na]
Optimal order[(a)]	0	1	na
# indiv. = N	2 379	1 513	752

* Significant at the 10% level; ** Significant at the 5% level. Standard errors in parentheses. P-values of the goodness-of-fit test in square brackets. na = "not available". In bold the preferred estimate: model with optimal polynomial in the widest window not rejected at a P-value of 5%. A low wage is defined as a FTE gross starting wage below the median for 30 year olds in the absence of notification: in Flanders (Wallonia) this wage is 1964€ (1911€).

[(a)] Order of the model with the highest P-value according to the goodness-of-fit test.

Table A-2: The Effect (β_8) on Transitions to High Wage Jobs

Window width	$DS \pm 3 * 20$	$DS \pm 2 * 20$	$DS \pm 1 * 20$
Region		**Flanders**	
Polynomial order			
Zero	0.050** (0.016) [0.042]	0.032 (0.021) [0.015]	0.001 (0.030) [na]
One	**0.000** (0.026) [0.069]	-0.018 (0.033) [0.018]	0.032 (0.044) [na]
Two	-0.002 (0.037) [0.047]	0.031 (0.045) [0.001]	0.050 (0.066) [na]
Optimal order[(a)]	1	1	na
# indiv. = N	2 476	1 592	769
		Wallonia	
Zero	0.050** (0.013) [0.002]	**0.040**** (0.017) [0.277]	0.022 (0.024) [na]
One	0.031 (0.025) [0.002]	0.019 (0.031) [0.210]	0.003 (0.044) [na]
Two	0.004 (0.037) [0.001]	0.011 (0.047) [0.112]	0.031 (0.058) [na]
Optimal order[(a)]	0	0	na
# indiv. = N	2 379	1 513	752

* Significant at the 10% level; ** Significant at the 5% level. Standard errors in parentheses. P-values of the goodness-of-fit test in square brackets. na = "not available". In bold the preferred estimate: model with optimal polynomial in the widest window not rejected at a P-value of 5%. A high wage is defined as a FTE gross starting wage below the median for 30 year olds in the absence of notification: in Flanders (Wallonia) this wage is 1964€ (1911€).

[(a)] Order of the model with the highest P-value according to the goodness-of-fit test.

Table A-3: The Effect (β_8) on Transitions to Short Employment Spells

Window width	$DS \pm 3 * 20$	$DS \pm 2 * 20$	$DS \pm 1 * 20$
Region		**Flanders**	
Polynomial order			
Zero	**0.083**** (0.017) [0.741]	0.077** (0.021) [0.130]	0.077** (0.031) [na]
One	0.055** (0.028) [0.658]	0.051 (0.033) [0.132]	0.103** (0.043) [na]
Two	0.086** (0.037) [0.385]	0.108** (0.045) [0.033]	0.117** (0.055) [na]
Optimal order[(a)]	0	1	na
# indiv. = N	2 476	1 592	769
		Wallonia	
Zero	**0.071**** (0.014) [0.861]	0.069** (0.017) [0.000]	0.040* (0.023) [na]
One	0.053** (0.025) [0.798]	0.026 (0.031) [0.212]	-0.008 (0.047) [na]
Two	-0.006 (0.039) [0.680]	-0.009 (0.050) [0.854]	-0.053 (0.068) [na]
Optimal order[(a)]	0	2	na
# indiv. = N	2 379	1 513	752

* Significant at the 10% level; ** Significant at the 5% level. Standard errors in parentheses. P-values of the goodness-of-fit test in square brackets. na = "not available". In bold the preferred estimate: model with optimal polynomial in the widest window not rejected at a P-value of 5%. A short employment spell lasts less than the median employment duration for 30 year olds in the absence of notification: in Flanders (Wallonia) this is 10 (9.4) months.
[(a)] Order of the model with the highest P-value according to the goodness-of-fit test.

Table A-4: The Effect (β_8) on Transitions to Long Employment Spells

Window width	$DS \pm 3 * 20$	$DS \pm 2 * 20$	$DS \pm 1 * 20$
Region		**Flanders**	
Polynomial order			
Zero	**0.044**** (0.017) [0.074]	0.013 (0.020) [0.348]	-0.007 (0.028) [na]
One	-0.005 (0.026) [0.071]	-0.023 (0.032) [0.241]	-0.002 (0.042) [na]
Two	-0.013 (0.037) [0.045]	0.015 (0.041) [0.132]	0.067 (0.049) [na]
Optimal order[(a)]	0	0	na
# indiv. = N	2 476	1 592	769
		Wallonia	
Zero	0.030** (0.014) [0.049]	**0.002** (0.015) [0.328]	0.009 (0.023) [na]
One	0.008 (0.022) [0.028]	0.020 (0.025) [0.180]	0.016 (0.035) [na]
Two	0.032 (0.030) [0.021]	0.023 (0.034) [0.094]	0.041 (0.046) [na]
Optimal order[(a)]	0	0	na
# indiv. = N	2 379	1 513	752

* Significant at the 10% level; ** Significant at the 5% level. Standard errors in parentheses. P-values of the goodness-of-fit test in square brackets. na ="not available". In bold the preferred estimate: model with optimal polynomial in the widest window not rejected at a P-value of 5%. A long employment spell lasts less than the median employment duration for 30 year olds in the absence of notification: in Flanders (Wallonia) this is 10 (9.4) months.
[(a)] Order of the model with the highest P-value according to the goodness-of-fit test.

Table A-5: The Effect (β_8) on Transitions into Training

Window width	$DS \pm 6 * 10$	$DS \pm 5 * 10$	$DS \pm 4 * 10$	$DS \pm 3 * 10$	$DS \pm 2 * 10$	$DS \pm 1 * 10$
Region			**Flanders**			
Polynomial order						
Zero	0.034** (0.015) [0.025]	0.036** (0.017) [0.018]	0.027 (0.019) [0.003]	**0.036*** (0.021) [0.854]	0.034 (0.025) [0.704]	0.052 (0.038) [na]
One	0.040 (0.026) [0.021]	0.040 (0.028) [0.011]	0.074** (0.032) [0.003]	0.065* (0.037) [0.818]	0.060 (0.049) [0.802]	-0.023 (0.061) [na]
Two	0.063 (0.039) [0.022]	0.089** (0.044) [0.015]	0.063 (0.049) [0.001]	0.037 (0.054) [0.692]	0.011 (0.068) [0.651]	-0.053 (0.093) [na]
Optimal order[(a)]	0	0	1	0	1	na
# indiv. = N	2 476	2 045	1 592	1 176	769	391
Window width	$DS \pm 3 * 20$	$DS \pm 2 * 20$	$DS \pm 1 * 20$			
		Wallonia				
Zero	**0.077**** (0.013) [0.598]	0.091** (0.016) [0.676]	0.070** (0.022) [na]			
One	0.099** (0.023) [0.563]	0.087** (0.029) [0.469]	0.077* (0.046) [na]			
Two	0.081** (0.035) [0.505]	0.063 (0.044) [0.307]	0.026 (0.062) [na]			
Optimal order[(a)]	0	0	na			
# indiv. = N	2 379	1 513	752			

* Significant at the 10% level; ** Significant at the 5% level. Standard errors in parentheses. P-values of the goodness-of-fit test in square brackets. na = "not available". In bold the preferred estimate: model with optimal polynomial in the widest window not rejected at a P-value of 5%.

[(a)] Order of the model with the highest P-value according to the goodness-of-fit test.

Table A-6: The Effect (β_8) on Transitions out of the Labor Force

Window size	$DS \pm 3 * 20$	$DS \pm 2 * 20$	$DS \pm 1 * 20$
Region		**Flanders**	
Polynomial order			
Zero	0.042** (0.015) [0.728]	0.032* (0.019) [0.720]	0.038 (0.027) [na]
One	**0.010** (0.025) [0.750]	0.008 (0.031) [0.893]	-0.019 (0.050) [na]
Two	0.028 (0.037) [0.607]	0.016 (0.050) [0.779]	0.023 (0.060) [na]
Optimal order$^{(a)}$	1	1	na
# indiv. = N	2 476	1 592	769
		Wallonia	
Zero	0.050** (0.013) [0.033]$^{(b)}$	**0.043**** (0.017) [0.190]$^{(b)}$	0.034 (0.025) [na]
One	0.034 (0.023) [0.015]$^{(b)}$	0.051* (0.029) [0.173]$^{(b)}$	0.060 (0.040) [na]
Two	0.053 (0.034) [0.005]$^{(b)}$	0.050 (0.042) [0.088]$^{(b)}$	0.135** (0.049) [na]
Optimal order$^{(a)}$	$0^{(b)}$	$0^{(b)}$	na
# indiv. = N	2 379	1 513	752

* Significant at the 10% level; ** Significant at the 5% level. Standard errors in parentheses. P-values of the goodness-of-fit test in square brackets. na = "not available". In bold the preferred estimate: model with optimal polynomial in the widest window not rejected at a P-value of 5%.

(a) Order of the model with the highest P-value according to the goodness-of-fit test.

(b) Rules of thumb not satisfied for all groups: goodness-of-fit statistics are not reliable.

Table A-7: The Effect (β_8) on Transitions to Sickness Insurance

Window size	$DS \pm 3 * 20$	$DS \pm 2 * 20$	$DS \pm 1 * 20$
Region		**Flanders**	
Polynomial order			
Zero	**0.012** (0.012) [0.467]$^{(b)}$	0.009 (0.015) [0.413]$^{(b)}$	0.029 (0.020) [na]
One	0.021 (0.019) [0,433]$^{(b)}$	0.037 (0.022) [0.209]$^{(b)}$	0.045 (0.033) [na]
Two	0.058** (0.026) [0.267]$^{(b)}$	0.054 (0.033) [0.119]$^{(b)}$	0.088** (0.041) [na]
Optimal order$^{(a)}$	$0^{(b)}$	$0^{(b)}$	na
# indiv. = N	2.476	1.592	769
		Wallonia	
Zero	**0.023**** (0.009) [0.100]$^{(b)}$	0.021* (0.012) [0.314]$^{(b)}$	0.019 (0.017) [na]
One	0.021 (0.016) [0.043]$^{(b)}$	0.035* (0.019) [0.000]$^{(b)}$	0.047* (0.027) [na]
Two	0.036 (0.022) [0.019]$^{(b)}$	0.041 (0.029) [0.000]$^{(b)}$	0.093** (0.034) [na]
Optimal order$^{(a)}$	$0^{(b)}$	$0^{(b)}$	na
# indiv. = N	2.379	1.513	752

* Significant at the 10% level; ** Significant at the 5% level. Standard errors in parentheses. P-values of the goodness-of-fit test in square brackets. na = "not available". In bold the preferred estimate: model with optimal polynomial in the widest window not rejected at a P-value of 5%.

(a) Order of the model with the highest P-value according to the goodness-of-fit test.

(b) Rules of thumb not satisfied for all groups: goodness-of-fit statistics are not reliable.

References

Amemiya, T., and F. Nold (1975) 'A Modified Logit Model.' *The Review of Economics and Statistics* 57(2), 255–257

Arni, P., R. Lalive, and J. C. van Ours (2009) 'How effective are unemployment benefit sanctions? Looking beyond unemployment exit.' IFAU working paper 2009:22

Ashenfelter, O., D. Ashmore, and O. Deschênes (2005) 'Do Unemployment Insurance Recipients Actively Seek Work? Evidence from Randomized Trials in Four U.S. States.' *Journal of Econometrics* 125(1-2), 53–75

Battistin, E., and E. Rettore (2008) 'Ineligibles and Eligible Non-Participants as a Double Comparison Group in Regression Discontinuity Designs.' *Journal of Econometrics* 142, 715–730

Black, D. A., J. A. Smith, M.C. Berger, and B.J. Noel (2003) 'Is the Threat of Reemployment Services More Effective than the Services Themselves? Evidence from Random Assignment in the UI System.' *American Economic Review* 93, 1313–1327

Boone, J., and J.C. van Ours (2009) 'Why Is There a Spike in the Job Finding Rate at Benefit Exhaustion?' IZA Discusion Paper 4523

Boone, J., P. Fredriksson, B. Holmlund, and J. van Ours (2007) 'Optimal Unemployment Insurance with Monitoring and Sanctions.' *The Economic Journal* 117(518), 399–421

Borland, J., and Y.-P. Tseng (2007) 'Does a Minimum Job Search Requirement Reduce the Time Spent on Unemployment? Evidenc from the Job Seekers Diary in Australia.' *Industrial and Labor Relations Review* 60(3), 357–378

Box, G. E. P., W. G. Hunter, and J. S. Hunter (1978) *Statistics for Experimenters. An Introduction to Design, Data Analysis and Model Building* (New York:Wiley)

Caliendo, M., K. Tatsiramos, and A. Uhlendorff (2009) 'Benefit Duration, Unemployment Duration and Job Match Quality.' IZA Discussion Paper 4670

Card, D.E., R. Chetty, and A. Weber (2007) 'The Spike at Benefit Exhaustion: Leaving the Unemployment System or Starting a New Job?' *American Economic Review Papers and Proceedings* 97(2), 113–118

Cockx, B., A. Defourny, M. Dejemeppe, and B. Van der Linden (2007) *Le nouveau système de suivi des chômeurs* (Louvain-la-Neuve)

Cockx, B., and G. Ridder (2001) 'Social Employment of Welfare Recipients in Belgium: An Evaluation.' *The Economic Journal* 111, 322–352

Cockx, B., and J. Ries (2004) 'The Exhaustion of Unemployment Benefits in Belgium. Does it Enhance the Probability of Employment?' Discussion Paper No. 1177, Bonn: IZA

Cockx, B., and M. Dejemeppe (2005) 'Duration Dependence in the Exit Rate Out of Unemployment in Belgium. Is it True or Spurious?' *Journal of Applied Econometrics* 20, 1–23

—— (2007) 'Is the Notification of Monitoring a Threat to the Unemployed? A Regression Discontinuity Approach.' IZA discussion paper 2854

Dejemeppe, M. (2005) 'A complete decomposition of unemployment dynamics using longitudinal grouped duration data.' *Oxford Bulletin of Economics and Statistics* 67(1), 47–70

Dolton, P., and D. O'Neill (1996) 'Unemployment Duration and the Restart Effect: Some Experimental Evidence.' *The Economic Journal* 106(435), 387–400

—— (2002) 'The Long-Run Effects of Unemployment Monitoring and Work-Search Programs: Experimental Evidence from the United Kingdom.' *Journal of Labor Economics* 20(2), 381–403

Fredriksson, P., and B. Holmlund (2006) 'Improving Incentives in Unemployment Insurance: A Review of Recent Research.' *Journal of Economic Surveys* 20(3), 357–386

Gans, J. S., and A. Leigh (2009) 'Born on the first of july: An (un)natural experiment in birth timing.' *Journal of Public Economics* 93, 246–263

Geerdsen, L. P. (2006) 'Is there a Threat Effect of Labour Market Programmes? A Study of ALMP in the Danish UI System.' *The Economic Journal* 116, 738–750

Geerdsen, L. P., and A. Holm (2007) 'Duration of UI Periods and the Perceived Threat Effect from Labour Market Programmes.' *Labour Economics* 14(3), 639–652

Gorter, C., and G.R.J. Kalb (1996) 'Estimating the Effect of Counseling and Monitoring the Unemployed Using a Job Search Model.' *Journal of Human Resources* 31, 590–610

Graversen, B. K., and J. C. van Ours (2008) 'How to Help Unemployed to Find Jobs Quickly: Experimental Evidence from a Mandatory Activation Program.' *Journal of Public economics* 92, 2020–2035

Hahn, J., P. Todd, and W. van der Klaauw (2001) 'Identification and Estimation of Treatment Effects by Regression Discontinuity Design.' *Econometrica* 63(3), 201–209

Imbens, G. W., and T. Lemieux (2008) 'Regression discontinuity designs: A guide to practice.' *Journal of Econometrics* 142, 615–635

Johnson, N. L., A.W. Kemp, and S. Kotz (2005) *Univariate Discrete Distributions*, third ed. (Hoboken, New Jersey: Wiley)

Johnson, T. R., and D. H. Klepinger (1994) 'Experimental Evidence on Unemployment Insurance Work-Search Policies.' *The Journal of Human Resources* 29(3), 695–717

Klepinger, D. H., T.R. Johnson, and J.M. Joesch (2002) 'Effects of Unemployment Insurance Work Search Requirements: The Maryland Experiment.' *Industrial and Labor Relations Review* 56(1), 3–22

Kuhn, A., R. Lalive, and J. Zweimüller (2009) 'The public health costs of job loss.' *Journal of Health Economics* 28, 1099–1115

Lalive, R., J. C. van Ours, and J. Zweimueller (2005) 'The Effect of Benefit Sanctions on the Duration of Unemployment.' *Journal of the European Economic Association* 3(6), 1386–1417

Lalive, R., J. van Ours, and J. Zweimuller (2006) 'How Changes in Financial Incentives Affect the Duration of Unemployment.' *The Review of Economic Studies* 73, 1009–1038

Lee, D. S. (2008) 'Randomized Experiments from Non-random Selection in U.S. House Elections.' *Journal of Econometrics* 142(2), 675–697

Lee, D. S., and D. Card (2008) 'Regression Discontinuity Inference with Specification Error.' *Journal of Econometrics* 142, 655–674

Lee, D. S., and T. Lemieux (2009) 'Regression Discontinuity Designs in Economics.' NBER Working Paper 14723

Manning, A. (2009) 'You Can't Always Get What You Want: The Impact of the UK Jobseeker's Allowance.' *Labour Economics* 16(3), 239–250

McCrary, J. (2008) 'Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test.' *Journal of Econometrics* 142, 698–714

McVicar, D. (2008) 'Job search monitoring intensity, unemployment exit and job entry: Quasi-experimental evidence from the UK.' *Labour Economics* 15, 1451–1468

Meyer, B. D. (1995) 'Lessons from the U.S. Unemployment Insurance Experiments.' *Journal of Economic Literature* 33, 91–131

Mortensen, D. (1986) 'Job Search and Labor Market Analysis.' In *Handbook of Labor Economics, Vol. II*, ed. O. Ashenfelter and R. Layard (Amsterdam: North-Holland) pp. 849–919

OECD (2007) *Employment Outlook* (Paris: OECD)

ONEM (2004) *Bulletin mensuel Juin* (Brussels: ONEM)

Petrongolo, B. (2008) 'The Long-Term Effects of Job Search Requirements: Evidence from the UK JSA Reform.' IZA Discussion Paper 3856

Rosholm, M. (2008) 'Experimental Evidence on the Nature of the Danish Employment Miracle.' IZA Discussion Paper 3620

Rosholm, M., and M. Svarer (2008) 'The Threat Effect of Active Labour Market Programmes.' *The Scandinavian Journal of Economics* 110(2), 385–401. Discussion Paper No. 1300, Bonn: IZA

RVA (2006) *Jaarverslag 2005* (Brussels: RVA)

Thistlethwaite, D., and D. Campbell (1960) 'Regression-discontinuity analysis: an alternative to the ex post facto experiment.' *Journal of Educational Psychology* 51, 309–317

van den Berg, G. J., and B. van der Klaauw (2006) 'Counseling and Monitoring of Unemployed Workers: Theory and Evidence from a Controlled Social Experiment.' *International Economic Review* 47(3), 895–936

van Ours, J. C., and M. Vodopivec (2006) 'Does Reducing Unemployment Insurance Generosity Reduce Job Match Quality?' *Journal of Public Economics* 92, 684–695

White, H. (1994) *Estimation, Inference, and Specification Analysis* (Cambridge University Press, Cambridge, UK)