Kent Academic Repository

Downloaded from

The version of record is available from

This document version

Author's Accepted Manuscript

DOI for this version

Licence for this version

Additional information

Versions of research works

Versions of Record

Author Accepted Manuscripts

Enquiries

Imperfect Monitoring of Job Search:
Structural Estimation and Policy Design[*]

Bart Cockx[†], Muriel Dejemeppe[‡], Andrey Launov[♭] and Bruno Van der Linden[⋆]

[†] Sherppa, Ghent U., IRES, IZA and CESIfo, [‡] IRES, Université catholique de Louvain,
[♭] U. of Kent, IZA and CESIfo, [⋆]IRES, Université catholique de Louvain, FNRS, IZA and CESIfo

Date: December 2, 2016

Abstract

We build and estimate a non-stationary structural job search model that incorporates the main stylized features of a typical job search monitoring scheme in unemployment insurance (UI) and acknowledges that search effort and requirements are measured imperfectly. Based on Belgian data, monitoring is found to affect search behavior only weakly, because (i) assessments were scheduled late and infrequently; (ii) the monitoring technology was not sufficiently precise, (iii) lenient Belgian UI results in caseloads that are less responsive to incentives than elsewhere. Simulations show how changing the aforementioned design features can enhance effectiveness and that precise monitoring is key in this.

Keywords: Monitoring, sanctions, non-stationary job search, unemployment benefits, structural estimation.

JEL Classification: J64, J68, C41.

---

[*]Corresponding author: Bart Cockx, Ghent University (SHERPPA and Faculty of Economics and Business Administration), Sint-Pietersplein 6, B-9000 Gent, Belgium. Email: bart.cockx@ugent.be, phone +32 9 264 78 77, fax: +32 9 264 89 96. Muriel Dejemeppe: muriel.dejemeppe@uclouvain.be, Andrey Launov: a.launov@kent.ac.uk, Bruno Van der Linden: bruno.vanderlinden@uclouvain.be.

1 Introduction

By now, the way limiting the coverage of UI and the duration of benefit entitlement can restore work incentives has been well studied. At the same time, apart from limiting entitlement and coverage, most UI schemes also provide incentives by imposing job search requirements on benefit claimants, monitoring the resulting search activity, and issuing sanctions in case of non-compliance. Recently, there has been growing interest among researchers in the evaluation of such job search monitoring (see below). However, the behavioral channels through which these policies work remain largely unknown, in particular when the unemployed are informed about the timing of the monitoring interviews and sanctions. By integrating a monitoring scheme and the announced timing of associated interviews and sanctions within a structural job search model, we shed light on the dynamic behavioral adjustments induced by such a scheme.

In many OECD countries, monitoring of job search effort is organized along fairly standardized lines (OECD, 2007). It starts off with a notification (often at initial registration) informing the unemployed worker of the search requirements and the proofs thereof which are to be provided. Simultaneously, the worker is informed of the timing of the evaluations of search effort and of the associated sanctions in case of noncompliance. At the prescribed dates, past job search effort is evaluated on the basis of transmitted paper proofs of job applications or in face-to-face interviews. If the outcome of the evaluation is negative, a sanction in the form of a temporary and partial reduction of UI benefits usually follows. In addition, another monitoring interview is planned shortly thereafter. In the case of a second negative evaluation, the penalty is reinforced and can even lead to a permanent withdrawal of benefits. If the outcome of the evaluation is positive, no sanction is imposed and a subsequent monitoring interview is scheduled, but usually not so quickly as in the case of a negative evaluation.

In this paper, we develop and estimate a *non-stationary* structural job search model which integrates the aforementioned key ingredients of a typical job search monitoring scheme[1], as well as an imperfect monitoring technology that acknowledges that job search effort and requirements are not perfectly measurable. Taking this imperfection into account is crucial, because ignoring it would result in too favorable an evaluation of the monitoring scheme (Cockx et al., 2014). By integrating the key ingredients of a typical scheme, our model can also be used to study the impact of alternative designs implemented in other countries, rather than only that of the design selected for the estimation.

In this research, we tailor our model to capture the main features of the scheme that was introduced in the Belgian UI in 2004. As the scheme was gradually implemented by age, starting with the youth, the estimation of our model is restricted to the population of 25- to 30-years-olds. From an international perspective, the Belgian UI is very lenient and the monitoring scheme introduced in 2004 is no exception. The scheme was targeted at the long-term unemployed only, and characterized by a low monitoring frequency and precision. Consequently, we find that the scheme hardly enhanced the job finding rate. Based on simulations, we find that designing a "stricter" scheme, more in line with those existing in other countries, reinforces the effect of monitoring on unemployment duration by nearly a factor three. We also show that enhancing monitoring precision further boosts the impact on the job finding rate. Nevertheless, these factors alone are not sufficient to explain the discrepancy with the stronger effects found by other researchers, based on structural models for the U.S. and the Netherlands (Paserman, 2008; van den Berg and van der Klaauw, 2015). We conjecture that the remaining differences can be explained by the particular composition of Belgian UI caseloads and by the restriction to the youth, for whom active labor market policies are typically found to be less effective than for the prime aged (Card, Kluve and Weber, 2015, e.g.).

Various papers have estimated non-stationary job search models under the assumption of exogenous search effort.[2] Some papers have estimated job search models with endogenous job search intensity,

---

[1]Since the timing is announced, it is crucial to embed this in a *non-stationary* model, since the unemployed will anticipate this timing and adjust their behavior as they approach the time when their job search is evaluated.

[2]See in particular Wolpin (1987),van den Berg (1990), Garcia-Perez (2006), Frijters and van der Klaauw (2006), and

though assuming a stationary environment.[3] DellaVigna and Paserman (2005) and Paserman (2008) are, to our knowledge, the first to have allowed for endogenous search in a non-stationary setting, where non-stationarity stems from the finite entitlement to UI.[4] The model of Paserman (2008) does not consider monitoring of job search effort, but simulations based on this model investigate the implications of a simplified monitoring scheme in which benefits are withdrawn if search effort falls below a particular threshold. We add to this literature by combining non-stationarity, endogeneity of search effort, and sanctions in a unified empirical framework which explicitly takes unobserved heterogeneity into account.

To our knowledge, van den Berg and van der Klaauw (2015) are the only ones to have developed and estimated a stationary structural job search model which integrates job search monitoring. Their approach of modeling monitoring imperfection is different from ours. We assume that the evaluation of search effort is a random process in which the probability of a negative evaluation is a decreasing function of average past job search effort (Boone et al., 2007). Following van den Berg and van der Klaauw (2006), they assume that the unemployed search through two imperfectly substitutable channels: a formal one and an informal one. Search effort is perfectly observed and monitored in the formal channel, while it cannot be verified at all and is not monitored in the informal channel. They consider an intensive monitoring policy targeted at short-term, relatively skilled unemployed individuals in the Netherlands. Their model reveals that job search channel substitution reduces the effectiveness of monitoring, but at the same time, together with on-the-job search, it also mitigates the adverse effect of monitoring on job quality, as measured by accepted wages and job duration.

Other researchers have developed partial and equilibrium job search models that comprise job search monitoring, but they do not estimate these models structurally. Moreover, either they assume that the monitoring technology is perfect or that other behaviors than search are controlled.[5] An imperfect monitoring technology in which the probability of a negative evaluation depends on search effort would, however, square better with existing schemes (OECD, 2007, p.218). To our knowledge, only two studies have considered such imperfect monitoring technology. Setty (2015) builds a dynamic principal-agent model with two levels of search effort to address a question of optimal UI with a probabilistic monitoring technology in which the probability of a negative evaluation decreases with the effort level. Boone et al. (2007) consider a stationary equilibrium search model with sanctions. They assume, as we do, that effort is imperfectly observable to the planner, but, unlike us, that the unemployed are perfectly informed about the search requirements (see their Appendix C). Here, we do not develop an equilibrium job search model. The assumption that the wage offer distribution and the supply of vacancies are unaffected is reasonable for the Belgian monitoring scheme, since it only targets the long-term unemployed who represent a small fraction of potential recruits. However, when considering policy reforms in which monitoring would target the short-term unemployed, our partial equilibrium framework should be treated as an approximation.

There exists a more abundant empirical literature evaluating the effectiveness of monitoring schemes within reduced-form models. Early studies[6] found that a combination of monitoring and counseling had a positive effect on the job finding rate. A number of later contributions have successfully isolated the pure effects of monitoring. The evidence about its impact on transitions to employment remains

---

Lollivier and Rioux (2010).

[3]Bloemen (2005), Fougère et al. (2009) and van der Klaauw and van Vuuren (2010) among others.

[4]Launov and Wälde (2013) formulate and estimate a non-stationary matching model with endogenous effort and time-dependent benefits, but focus rather on equilibrium effects of UI benefit reduction in a Mortensen-Pissarides setting.

[5]Some authors assume that job search effort can be perfectly monitored (Pavoni and Violante, 2007; Manning, 2009; Petrongolo, 2009; Wunsch, 2013); for others, the probability of a negative evaluation is independent of search effort (Abbring et al., 2005), or it depends on the acceptance of suitable job offers (Ljungqvist and Sargent, 1995; Boone et al., 2009), or on voluntary quits (van den Berg et al., 2015).

[6]See Meyer (1995) for a review of U.S. studies, and Gorter and Kalb (1996) and Dolton and O'Neill (1996, 2002) for a review of European studies.

rather mixed.[7] Increases in withdrawals from the labor force or decreases in employment stability and post-unemployment earnings are emphasized by Manning (2009), Petrongolo (2009) and Arni et al. (2013) among others.

Using the same data as in this paper, in an application of a regression discontinuity design at the age threshold of 30, Cockx and Dejemeppe (2012) found that job search monitoring did increase the transition rate to employment by nearly nine percentage points before the first evaluation of job search effort. Given that at this first evaluation the unemployed do not risk a benefit sanction (only in the case of recidivism), this high treatment effect is puzzling. One explanation is that, despite being robust, the treatment effect was imprecisely estimated. By imposing the restrictions implied by our structural model, we indeed obtain a much smaller treatment effect.

Our paper is organized as follows. Section 2 provides information on the institutional setting. Section 3 describes the sample selection and the data. In Section 4, we present the job search model incorporating the main features of the Belgian monitoring scheme which we estimate subsequently. Section 5 elaborates on the econometric model and discusses identification. Section 6 reports the estimation results. In Section 7, on the basis of simulations, we evaluate the impact of the existing scheme, as well as investigate whether and in which direction the design of the scheme in place (in terms of timing of assessments, strength of sanctions and precision of the monitoring technology) could be improved. Section 8 summarizes our key results and sets avenues for future research. The Online Appendix contains further technical details regarding the theoretical and econometric model and its solution.[8]

2 The Belgian Job Search Monitoring Scheme

In Belgium, UI, monitoring and sanctions are organized at the federal level. The Public Employment Services (PES) are under regional authority. They are in charge of counseling, job search assistance, intermediation services and training. A worker in Belgium is entitled to UI in two instances: (i) after graduation from school conditional on a waiting period of 9 months; (ii) after being laid off from a job which was held for a qualifying period of time. In contrast to many other countries, there is in principle no time limit to UI. However, sanctions may imply losing one's entitlement to UI, in particular in the context of the policy evaluated in this paper (see below). School-leavers are entitled to flat rate benefits, while laid-off 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)[9] and on unemployment duration for those who are single and for cohabitants. These principles are valid for the period covered by our empirical analysis (July 2004 - December 2006).

In July 2004, the UI agency began monitoring job search effort. Since then, the agency has been able to withdraw benefits in case of insufficient search effort. During the period we cover, there were no other labor market reforms affecting the target group. The analysis is restricted to the Flemish region, where the scheme was implemented without the provision of more counseling. Between July 2004 and June 2005, only unemployed workers younger than 30 were concerned. Afterwards, the monitoring scheme was gradually extended to older age groups.

The monitoring process consists of a notification and a sequence of face-to-face interviews. Figure 1 summarizes the timing of the notification, the first interview and the subsequent interviews in case

---

[7]Positive effects were found by Klepinger et al. (1997), Lalive et al. (2005), McVicar (2008) and Bolhaar et al. (2016) among others. No effect or negative effects were found by Ashenfelter et al. (2005), Klepinger et al. (2002) and van den Berg and van der Klaauw (2006) among others.

[8]This appendix can be downloaded together with the data, the code and the instructions for using the code from https://www.sites.google.com/site/andreylaunov/research/CDLV.zip.

[9]An unemployed individual is regarded as the head of household if she lives together with a partner or relatives (children or other) whose labor earnings or allowances do not exceed a threshold set by regulations; otherwise, she is a cohabitant, or single, if living alone.

Figure 1: Timing of the Monitoring Process in Case of a Negative Evaluation
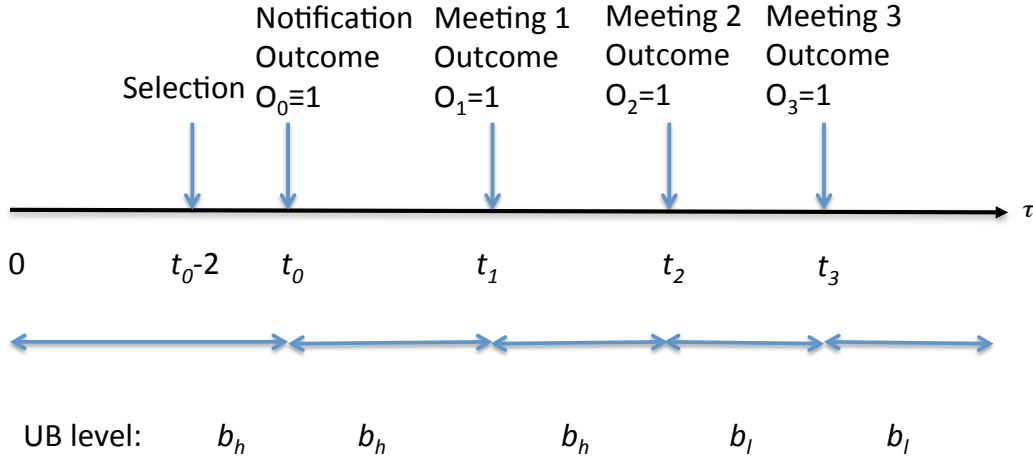
of a negative evaluation. If the outcome of the evaluation is positive at any of the interviews, a new sequence of interviews will start one year later at the earliest.[10]

First, at $t_s = t_0 - 2$, the administration selects individuals who have been entitled to UI for more than 12 but less than 13 months. In the second month after selection, i.e. at $t_0 = 14$, a notification letter is sent by mail. It states that entitlement to UI benefits requires actively searching for a job and participating 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 search actions undertaken. The letter does not, however, specify what evidence of search is sufficient. It announces that one will be invited to the UI office to evaluate the actions undertaken and that these evaluations will start 8 months after the notification has been dispatched ($t_1 = t_0 + 8 = 22$).

These monitoring interviews last approximately half an hour. If search effort at the first interview is deemed insufficient, an action plan is drawn up, but the worker is not sanctioned yet. A first sanction is only imposed in case of non-compliance with the action plan 4 months later ($t_2 = t_1 + 4 = 26$). This sanction consists in a second, stricter, action plan and a temporary reduction of the benefits b_h to the means-tested social assistance benefits b_ℓ until the third interview four months later ($t_3 = t_2 + 4 = 30$). In case of a negative assessment at this third interview, the benefits are permanently withdrawn. The individual can regain entitlement only after having been uninterruptedly employed full-time for at least one year.

An action plan is an administrative form containing a list of types of job search activities (renewing contact with the regional PES, sending application letters, registering with a temporary work agency, and the like) imposed by the caseworker, together with an indication of the number of times each activity type should be chosen to comply with the search requirements at the next meeting. As this is the prerogative of the regional PES, caseworkers are neither allowed to offer job vacancies, nor may they propose participation in training programs. Furthermore, sanctioning the refusal of suitable offers is the responsibility of another service within the federal UI agency. An action plan is therefore an individually tailored guideline for the job seeker. However, these guidelines were not very strictly followed and considerable discretion remained. Interviews with caseworkers in charge revealed that trade-offs between activity types were possible. Moreover, depending on the caseworker, action plans that were not completely satisfied did not automatically lead to sanctions, while, conversely, action plans that were followed were not always exempted from negative evaluations. The content of action plans is not available in our data. In order to capture the aforementioned fuzziness and discretion in the assessments, we model the relationship between job search effort and the outcome of the evaluations

---

[10]16 months later for the first interview and 12 months later otherwise.

as stochastic (see Subsection 4.1).

The low frequency of monitoring contrasts quite starkly with that in many other countries: half of OECD countries require reports of job search (in most cases) every two weeks or at least monthly (OECD, 2007). On the other hand, sanctions in case of non-compliance with the action plan seem generally tougher in Belgium than in other OECD countries. For instance, in the Netherlands, a typical punishment for insufficient job search is a 10% reduction of unemployment benefits for a period of 2 months (van den Berg and van der Klaauw, 2006). In the 2010-2012 period in Switzerland, Arni and Schiprowski (2015) report that "the median amount of a sanction is the monetary equivalent of 7 days of UI benefits" (p. 10). In Section 7, we will study the implications of these design features of the monitoring scheme.

During the period covered by our empirical analysis, the probability of a negative evaluation in the population of Flemish UI recipients increased with the order of the meetings: 33.5% at the first meeting, 45.8% at the second one, and 57.8% at the third. This may reflect both more strictness and the sorting process by which less intensive job seekers are more likely to be subjected to later interviews.

Only in some districts were unemployed individuals assigned to the same caseworker in principle, though even in these districts the assignment was often not followed through for practical reasons. There is therefore little scope for learning about evaluation standards across interviews. Regardless, we cannot identify caseworkers in our data.

3 Data

3.1 Sample

The data originate from several administrative sources available from 2001 to 2006. We have access to monthly information on UI benefit claims, the timing and outcomes of the monitoring scheme, employment (including self-employment) and earnings (for salaried workers). Information about the job search actions reported at the evaluation meetings is not available. Since our model does not explain the choice of working hours, we just retain full-time occupations.[11]

Our sample consists of individuals aged 25 to 30 who were unemployed for 12 to 13 months between May and August 2004 and to whom a notification was sent between July and October 2004 if they were still receiving UI benefits at the time. To avoid modeling non-stationary behavior induced by a declining benefit level after notification, we discarded cohabitants with sufficient past work experience to be at risk of such a decline from our sample (see Section 2). This results in a sample of 903 individuals.

Since the sampling occurs two months before notification, we can check whether claimants anticipated the notification by leaving the unemployment register beforehand. Cockx and Dejemeppe (2012) could not find any evidence of such anticipation (see their Section 6.1.2). This means that we can safely assume that the moment of notification came as a surprise. This can be rationalized by the complexity of the duration counter used in the monitoring scheme.

3.2 Descriptive Statistics

Table 1 reports summary statistics for the sample selected in 2004. Time-varying variables are evaluated at the sampling date. Monthly earnings are measured at the start of a salaried employment spell. All monetary variables are measured in 2004 euros. The table presents the individual characteristics conditioned upon in the empirical analysis (gender, level of education, household type

---

[11]We assume that someone hired for two-thirds or more of a full-time position is classified as a "full-time worker". This represents nearly 80% of exits to employment. The actual post-tax remuneration of the part-time workers retained has been scaled to a full-time job.

Table 1: Descriptive Statistics of the Individual Characteristics in the Sample Retained

Number of individuals	903
Gender	
Women	45.2%
Schooling level[a]	
Primary or lower secondary (low)	34.8%
Upper secondary (middle)	40.0%
Higher education (high)	25.2%
Type of entitlement[a] **(monthly UI benefit level in 2004 €**[a]**)**	
Entitled by work experience	69.2%
Head of household ([865 -1005])	22.1%
Single ([725 -835])	32.7%
Cohabitant (385)	14.4%
Entitled by schooling	30.8%
Head of household (835)	1.8%
Single (595)	7.2%
Cohabitant (325)	21.8%
Unemployment benefits[a]	
Mean (2004 €)	646
Standard deviation	(242)
25%	385
Median	725
75%	835
Observed net monthly earnings (1st spell)	
Number of individuals	427
Mean (2004 €)	1,199
Standard deviation	(279)
25%	1,066
Median	1,214
75%	1,358

[a] At the sample selection date.

Table 2: Assistance Benefit Levels b_ℓ in Case of a Negative Evaluation at the Second or Third Meeting and Size of the Sanction (Monthly Level in 2004 €)

Type of entitlement[†]	Range of b_h	b_ℓ	Lowest sanction[*]	Highest sanction[*]
Entitled by work experience				
Head of household	865 -1005	802	63	203
Single	725 -835	601	124	234
Cohabitant	385	0	385	385
Entitled by schooling				
Head of household	835	802	33	33
Single	595	595	0	0
Cohabitant	325	0	325	325

[†]At the sample selection date.

[*] I.e. benefit loss $b_h - b_\ell$.

determining the benefit level) and the type of entitlement (school-leaver or work experience). The monthly benefit level b_h varies between 325 € and 1,005 €, and is on average 646 €. Those who find a job earn 1,200 € (net of taxes) on average.

Table 3: Sampled Population at Each Step of the Monitoring Process[a]

Number of individuals	903
Steps of the monitoring process	
Notification letter	723
	(80.1%)
First interview	162
	(17.9%)
Positive evaluation	112
	(69.1%)
Negative evaluation	50
	(30.9%)
Second interview	18
	(36.0%)
Positive evaluation	16
Negative evaluation	2
Third interview	1
	(50.0%)
Positive evaluation	1
Negative evaluation	0

[a]% in the population at risk.

8

Table 2 describes benefit levels, b_ℓ, in case of a sanction. These are means-tested social assistance benefits. Recall that no sanction is imposed after the first negative evaluation. After subsequent negative evaluations, the same benefit amount is withdrawn, but the sanction is temporary after the second evaluation, while it is permanent after the third. The benefit loss $b_h - b_\ell$ ranges between 0 and 385 €/month. Table 2 reports the minimum and maximum sanctions by category. School-leavers living alone (i.e. *singles*) are effectively not subject to a benefit sanction, since they remain entitled to equivalent social assistance benefits. This concerns about 7% of the sample.

Table 3 displays the number of sampled benefit claimants at the various steps of the process and the outcomes of each evaluation. Since these individuals may have found a job between being selected and being notified, only 723 out of the 903 were notified. Among those notified, 162 attended the first interview, of which about one third were evaluated negatively. Since the evaluation process is long, delays in the scheduled timing are important, and many job seekers left the unemployment registers before the interviews took place. Consequently, only very few sampled individuals were observed to be evaluated for a second and third time. Subsection 5.2 will explain how we deal with these low frequencies.

4   Job Search in the Presence of Monitoring

In this section, we formalize the job search behavior of infinitely-lived unemployed workers whose job search is monitored similarly to the Belgian scheme. As shown in Figure 1, the unemployment spell, and hence, the job search behavior, can be divided into five sub-periods: $[0, t_0)$, $[t_0, t_1)$, $[t_1, t_2)$, $[t_2, t_3)$ and $[t_3, \infty)$.[12] Since the unemployed in the retained sample are unaware until notifcation (at t_0) that they will be monitored subsequently, the scheme does not affect job search within the first sub-period $[0, t_0)$. Since before the reform the unemployed were entitled to a constant unemployment benefit b_h without any time limit, their behavior within this first sub-period can be described by a standard stationary job search model and is not explicitly formulated here.[13] However, once the new scheme has been implemented and everyone knows about it, monitoring is anticipated from the start of the unemployment spell. When we simulate policy reforms in Section 7, we will allow for this anticipatory behavior.

In case of compliance, the next evaluation will not take place within a year. We therefore assume that a standard stationary job search model in which the benefit level is set to b_h without any time limit can approximate the resulting behavior. Similarly, we ignore the fact that, after a third negative evaluation, individuals may be entitled again to UI if they remain employed full-time for at least one year. As such, the standard stationary job search model with the benefit set to b_ℓ can describe the behavior after t_3.

We focus our discussion on how we model behavior within the three remaining sub-periods $[t_{k-1}, t_k)$, i.e. from the notification to the first meeting ($k = 1$) and after a negative assessment of search effort at the start of the two subsequent sub-periods ($k = 2$ or $k = 3$). At the beginning of each sub-period, it is assumed that the unemployed are well informed about the remaining stages of the monitoring scheme. Below, we look at a generic sub-period k. Hence, the endogenous (control) variables for such a generic problem depend on k. However, to avoid notational clutter, we do not explicitly denote this dependence.

A feature that complicates the analysis is that, in reality, evaluations do not take place at the scheduled times t_1, t_2 and t_3, but are delayed for various reasons. These delays are important and cannot be ignored in the empirical analysis. However, since accounting for these delays complicates

---

[12]Calendar time starts at entry in unemployment so that (calendar) time and unemployment duration are synonyms.

[13]For *laid-off* cohabitants and singles, the benefit level could be step-wise decreasing in this first period, but since the abatement is not very large, we ignore it. In Subsection 6.3, we demonstrate that this approximation does not invalidate our analysis.

the analysis substantially without affecting the main insights, we will first ignore them. Subsection 4.4 discusses how these delays are taken into account.

4.1 Monitoring Technology

Since job search requirements need not be satisfied at every moment of the evaluation period, a caseworker bases her evaluation on the *average* job search effort $\bar{S}(t_k, t_{k-1})$ exerted between t_{k-1} and t_k:[14]

$$\bar{S}(t_k, t_{k-1}) = \frac{\int_{t_{k-1}}^{t_k} s(\tau) d\tau}{t_k - t_{k-1}}, \tag{1}$$

where $s(\tau)$ denotes the instantaneous job search effort at time $\tau \in [t_{k-1}, t_k)$, and here and in all subsequent equations, $k \in \{1, 2, 3\}$.

Researchers have generally assumed that the precision of the monitoring technology is perfect in that job seekers are evaluated negatively if job search effort falls short of some predetermined target and positively if effort exceeds this threshold (Paserman, 2008; van den Berg and van der Klaauw, 2015, e.g.). At this threshold, the first derivative of the probability of a negative evaluation is infinite and the probability of a negative evaluation drops discretely from one to zero as the threshold is crossed. Other authors have instead stressed that it is very difficult for caseworkers to measure the search intensity of the unemployed without error (see e.g. Boone et al., 2007). As explained in Section 2, in addition caseworkers have some discretion in determining whether search effort is sufficient. These two features are captured through a probability of a negative evaluation at t_k[15]

$$\pi_k \left[\bar{S}(t_k, t_{k-1}) \right] \quad \text{with} \quad \pi_k' \left[\bar{S}(t_k, t_{k-1}) \right] \leq 0. \tag{2}$$

In other words, the unemployed cannot increase the probability of a negative evaluation by raising their job search effort.

4.2 Workers' Problem

Workers discount the future at rate $\rho > 0$ and consume their current income entirely. Assuming risk-neutrality, non-labor income other than unemployment benefits does not affect behavior and can thus be normalized to zero.[16] Workers can either be employed full-time or unemployed. If employed, workers earn a constant net wage $w > 0$. There are no job-to-job transitions, and jobs dissolve at an exogenous constant Poisson rate $\delta > 0$. According to the legislation, if a temporary job interrupts unemployment, workers remain entitled to benefits and the counter $\tau \in [t_{k-1}, t_k)$ determining the timing of the next evaluation remains fixed to its value before the interruption.

With these assumptions, the expected lifetime utility of a worker transiting to employment at τ is:

$$W_k(w; \tau) = \frac{w + \delta U_k(\tau)}{\rho + \delta}, \tag{3}$$

where $U_k(\tau)$ denotes the expected lifetime utility at reentry in unemployment after a spell of temporary employment that started at τ.

As long as the unemployed worker is not sanctioned, she is entitled to a flat benefit level b_h. A sanction is imposed only from the second negative assessment onwards. At this point, the benefit level

---

[14]Alternatively, the caseworker might place higher weight on more recent job search effort. Introducing such weighting into eq. (1) is easy, but will not qualitatively affect the results.

[15]For a derivation with explicit micro foundations, see Cockx, Dejemeppe, Launov and Van der Linden (2011).

[16]Risk-neutrality is a strong assumption. Yet, since we lack complete information on non-labor income, this is the best that we can do.

falls temporarily to $b_\ell < b_h$. A third negative evaluation lowers the benefit permanently. Formally, for $k \in \{1, 2, 3\}$, we denote $O_k = 1$ if the outcome of the evaluation is negative and $O_k = 0$ if it is positive, and at notification $O_0 \equiv 1$. If $t_4 \equiv +\infty$ and $b(\tau)$ is the benefit level at time τ, then $\forall k \in \{1, 2\}, \tau \in [t_{k-1}, t_k) : b(\tau) = b_k \equiv b_h$ and $\forall k \in \{3, 4\}, \tau \in [t_{k-1}, t_k) : b(\tau) = b_k \equiv b_h - O_{k-1}(b_h - b_\ell)$. Apart from unemployment benefits, the unemployed enjoy a value of leisure (net of stigma costs) equal to ν. Search effort $s(\tau)$ is costly. We denote the cost of search function by $c[s(\tau)]$ and assume it possesses standard properties, namely $c(0) = 0$, $c'[s(\tau)] > 0$ and $c''[s(\tau)] > 0$. In sum, the net instantaneous utility of an unemployed worker at $\tau \in [t_{k-1}, t_k)$ is

$$y_k(\tau) \equiv b_k + \nu - c[s(\tau)]. \tag{4}$$

Since $W_k(w; \tau)$ is strictly increasing in w, the worker's optimal strategy is equivalent to accepting any offer that exceeds a reservation wage $w_r(\tau)$. Therefore, if $F(\cdot)$ denotes the wage offer distribution and $\bar{F}(\cdot) \equiv 1 - F(\cdot)$, the transition rate from unemployment to employment at time τ is

$$p(\tau) \equiv p[s(\tau), w_r(\tau)] = \lambda s(\tau) \bar{F}[w_r(\tau)] \geq 0, \tag{5}$$

where λ denotes the job arrival rate per unit of search effort. The survivor function at τ, conditional on being unemployed at $t_{k-1} \leq \tau$, is

$$P(\tau, t_{k-1}) = \exp\left\{-\int_{t_{k-1}}^{\tau} p(x) \, dx\right\}. \tag{6}$$

The optimal behavior over the interval $[t_{k-1}, t_k)$ is found by choosing the paths of control variables $\{s(\tau), w_r(\tau)\}_{\tau \in [t_{k-1}, t_k]}$ that maximize the expected lifetime utility of an unemployed worker at t_{k-1}, subject to the laws of motions of the two state variables: the survival probability $P(\tau, t_{k-1})$ and the average search effort $\bar{S}(\tau, t_{k-1})$. Agents are assumed to be perfectly forward looking. This means that they anticipate within each stage of the monitoring regime that the expected lifetime utility decreases in case of a negative evaluation at the end of each stage. This anticipation induces non-stationary behavior similar to the one that prevails when approaching the exhaustion of entitlement to UI benefits. The non-stationary problem can be solved by backward induction, starting with the stationary problem in case of a positive evaluation at any meeting, or with a negative evaluation at the third meeting. One proceeds further backwards, considering each time an earlier meeting, up to the moment of notification.

Formally, the expected lifetime utility of an unemployed worker at t_{k-1}, for $k \in \{1, 2, 3, 4\}$, is:[17]

$$U_k(t_{k-1}) = \int_{t_{k-1}}^{t_k} \left[y_k(\tau) + p(\tau)\bar{W}_k(\tau)\right] P(\tau, t_{k-1}) e^{-\rho(\tau - t_{k-1})} d\tau$$

$$+ \mathbb{U}_k(t_k) P(t_k, t_{k-1}) e^{-\rho(t_k - t_{k-1})}, \tag{7}$$

$$\mathbb{U}_k(t_k) = \pi_k[\bar{S}(t_k, t_{k-1})] U_{k+1}(t_k) + (1 - \pi_k[\bar{S}(t_k, t_{k-1})]) U^+. \tag{8}$$

The right-hand side of (7) is the discounted sum of three terms: (i) the "sum" from t_{k-1} to t_k of the instantaneous utility in unemployment $y_k(\tau)$ weighted by the probability $P(\tau, t_{k-1})$ of still being unemployed at each moment $\tau \in [t_{k-1}, t_k)$; (ii) the "sum" from t_{k-1} to t_k of the expected utility of employment conditional on acceptance $\bar{W}_k(\tau) \equiv E[W_k(w; \tau)|w > w_r(\tau)]$ weighted by the density of unemployment duration at τ, $p(\tau) P(\tau, t_{k-1})$; (iii) the expected lifetime utility right before the monitoring interview, denoted $\mathbb{U}_k(t_k)$, weighted by the probability $P(t_k, t_{k-1})$ of surviving in unemployment up to t_k. Right before the job search assessment, the unemployed individual expects a positive evaluation with probability $(1 - \pi_k[\bar{S}(t_k, t_{k-1})])$. In this case, she is no longer monitored and

---

[17]In Online Appendix A, we show how $U_k(t_{k-1})$ can be derived from the limit of its recursive definition in discrete time.

is permanently entitled to the high benefit level b_h. The corresponding stationary expected lifetime utility level is denoted U^+. The unemployed individual expects a negative evaluation with probability $\pi_k\left[\bar{S}(t_k, t_{k-1})\right]$. In this case, she enters the next stage of the monitoring regime, with the associated lifetime utility $U_{k+1}(t_k)$. The latter declines over the different stages k.

After the first negative evaluation ($k = 1$), the individual has not been sanctioned yet, i.e. $y_2 = y_h$, but the expected lifetime utility drops relative to its value $\mathbb{U}_1(t_1)$ just before the first interview, because she has missed the opportunity of being permanently entitled to the high benefit level b_h, and because she anticipates the risk of the temporary sanction $(y_h - y_\ell)$ at the second interview, which may turn into a permanent one at the third: $U_2(t_1) < \mathbb{U}_1(t_1)$. Right after the second negative evaluation ($k = 2$), the lifetime utility drops further relative to its value just before this assessment for similar reasons and, in particular, because of the enhanced likelihood of being permanently sanctioned: $U_3(t_2) < \mathbb{U}_2(t_2)$. Finally, right after the third negative evaluation, the lifetime utility drops to its stationary lower bound, i.e. the lifetime utility U^- after a permanent withdrawal of the UI benefit, when the individual is just entitled to the social assistance benefit $b_\ell < b_h$: $U_4(t_3) \equiv U^- < \mathbb{U}_3(t_3)$. These discrete drops in lifetime utility across the different stages of the monitoring regimes provide valuable variation in the behavior of agents that will be exploited in the identification of the structural parameters of the model (see Subsection 5.2). The derivation of the optimality conditions described in the next section can be found in Online Appendix C.[18]

4.3 Optimality Conditions

The pair of optimal paths $\{w_r(\tau), s(\tau)\}$ obey two first-order conditions (FOC). The first one is:

$$U_k(\tau) = W_k(w_r(\tau); \tau) = \frac{w_r(\tau) + \delta U_k(\tau)}{\rho + \delta} \quad \Leftrightarrow \quad w_r(\tau) = \rho U_k(\tau) \tag{9}$$

This states that the reservation wage is chosen such that the expected lifetime utility in unemployment and in employment at the reservation wage should be equal at every instant of time τ. In the previous subsection, we explained that expected lifetime utility falls discontinuously after each negative evaluation and smoothly in anticipation of these drops. From the last expression in (9), it follows that the evolution of the reservation wage is just a scaled-down version of the evolution of lifetime utility.

After some manipulations, we obtain a generalization of the condition reported by van den Berg (1990, Eq. (3), p.259):[19]

$$w_r(\tau) + c[s(\tau)] = b_k + \nu + p(\tau)\frac{E\left[w - w_r(\tau)|w > w_r(\tau)\right]}{\rho + \delta} + \dot{U}_k(\tau). \tag{10}$$

The interpretation is as follows. The right-hand side represents the benefits of continuing search if one is offered a job that pays the reservation wage. It consists of three components: (i) the continuing flow of income b_k together with the net value of leisure; (ii) the probability of finding a job times the expected discounted lifetime wage gains relative to the reservation wage, given that no job paying below the reservation wage is accepted; and (iii) the rate of appreciation of the asset value of unemployment. At the optimum, these benefits should be equal to the costs of continuing search, as expressed on the left-hand side of Equation (10). They correspond to the opportunity cost of not accepting the job and the cost of search effort. The benefit reduction from b_h to b_ℓ induced by the sanctions in case of negative assessments in the second and third interviews and the depreciation of the asset value of unemployment in anticipation of these sanctions decrease the benefits of search both *within* and *between* the different stages of the monitoring scheme.

---

[18]In Online Appendix C, the problem is generalized along the lines indicated in Subsection 4.4.

[19]We use the facts that $\dot{U}_k(\tau) = \rho U_k(\tau) - y_{j(k)}(\tau) - p(\tau)\left[\bar{W}_k(\tau) - U_k(\tau)\right]$ and, by (3) and (9), that $p(\tau)\left[\bar{W}_k(\tau) - U_k(\tau)\right] = p(\tau)E[W_k(w; \tau) - U_k(\tau)|w > w_r(\tau)] = p(\tau)\frac{E[w - w_r(\tau)|w > w_r(\tau)]}{\rho + \delta}$. van den Berg (1990) assumes an exogenous job arrival rate and no job destruction ($\delta = 0$).

The second FOC is:

$$c'[s(\tau)] = \frac{\lambda \bar{F}[w_r(\tau)]E\left[w - w_r(\tau)|w > w_r(\tau)\right]}{\rho + \delta} + \frac{\pi'_k\left[\bar{S}(t_k, t_{k-1})\right]}{t_k - t_{k-1}}\left[U_{k+1}(t_k) - U^+\right]P(t_k, \tau)e^{-\rho(t_k - \tau)} \quad (11)$$

This generalizes the familiar condition that the marginal cost of search should equal its marginal return (Mortensen, 1986, p. 871). The monitoring of job search enhances the marginal return by the second term on the right-hand side of (11). Increasing job search marginally at τ decreases the probability of negative evaluation by $-\pi'_k\left[\bar{S}(t_k, t_{k-1})\right]/(t_k - t_{k-1})$. The division by $(t_k - t_{k-1})$ reflects that the evaluation occurs on the basis of average rather than instantaneous search effort. Since this return realizes only to the extent that the worker is unemployed at t_k, we need to weigh the value of avoiding a sanction, $[U^+ - U_{k+1}(t_k)]$, by the survivor probability between τ and t_k. In addition, since the evaluation occurs in the future ($t_k \geq \tau$), the return is discounted by $e^{-\rho(t_k - \tau)}$. Therefore, if $\pi'_k(.) < 0$, whatever the value of $w_r(\tau)$, the incentive to search harder increases between the successive stages of the monitoring scheme since the loss $[U^+ - U_{k+1}(t_k)]$ cannot decline with k. However, even if $\pi'_k(.) = 0$, which later turns out to be our estimation result, the marginal return to search increases over the successive stages. This is because then the marginal return to search effort increases as more wage opportunities open up when wage aspirations are gradually set at a lower level,[20] which follows from the declining lifetime utility over the successive stages (cf. supra).

4.4 Extensions

In this subsection, we discuss an aspect which has been neglected until now: the role of a temporary occupation on the behavior of caseworkers and the issue raised by delays in the scheduled timing of the job search assessments. The description of the monitoring technology in subsection 4.1 applies to workers who have been continuously unemployed. For them, only the evidence about job search effort matters. The guidelines for assessment instruct caseworkers to take job acceptance as sufficient evidence of compliance with the search requirements. In the empirical analysis, we demonstrate that work experience significantly increases the probability of positive evaluation, but it does not *guarantee* such an outcome, suggesting again caseworkers' discretion in the assessment. In the model, we take this feature into account, but for purposes of tractability, we simplify by assuming that, in case of an interruption, the probability of a negative evaluation no longer depends on past job search effort. Let superscript i denote whether a worker has interrupted unemployment ($i = 1$) or not ($i = 0$) between two interviews. Then, $\pi^i_k\left[\bar{S}^i(t_k, t_{k-1})\right]$ denotes the probability of a negative evaluation for $i \in \{0, 1\}$, and $\pi^1_k\left[\bar{S}^1(t_k, t_{k-1})\right] = \pi^1_k$, where π^1_k is a fixed number. Notice that by allowing this dependence on $i \in \{0, 1\}$, the (optimal) behavior of individuals depends on i. Online Appendix B explains how FOC (10) and (11) must be adapted. Intuitively, for an unemployed individual who has not yet been temporarily employed during the current interval (t_{k-1}, t_k), accepting a temporary job creates an "entitlement effect" if, once the person is unemployed again, caseworkers are more inclined to evaluate her effort positively.

Evaluations do not to take place at the scheduled moments t_1, t_2 and t_3, but are rather delayed for various reasons. The average delay is 5.75 months. A delay implies that each period $[t_{k-1}, t_k)$ is split into two sub-periods: a "scheduled interval"[21] followed by a "delay interval". According to the regulations, no evaluation can take place within a scheduled interval. During the delay interval, the meeting is assumed to arrive at a rate q.[22] In Online Appendix C, we derive the FOC of the optimization problem on the corresponding intervals. It turns out that the FOC of the reservation wage, (9) and (10), is not affected when we replace the expected lifetime utility of the unemployed

---

[20]The derivative of the first term on the right-hand side of (11) with respect to the reservation wage is strictly negative.

[21]Lasting for $k = 1$ and $k \in \{2, 3\}$, respectively 8 and 4 months.

[22]Delays are assumed to be distributed exponentially with rate q, truncated at a maximum observed delay. Truncation from above and rate q are identical for all meetings.

by the corresponding ones in the scheduled and delay periods. By contrast, the FOC of search effort does change. To understand why, notice that at the end of the scheduled interval, Expression (8) is replaced by the lifetime value of entering the delay interval. The incentives of the monitoring scheme are intuitively weakened, because delays introduce some probability that the meeting will not take place until the maximum delay. The interested reader can find the precise formula in Online Appendix C.

5 The Econometric Model

In this section, we discuss the specification and identification of the econometric model. The derivation of the likelihood contributions can be found in Online Appendix F. Online Appendix G describes in detail how the model is solved and estimated.

5.1 Specification

The estimation of the structural model requires the specification of the unknown functions $c(.)$, $F(.)$, $\pi_k^i(.)$ (for $i = 0, 1$ and $k = 1, 2, 3$) and choosing the way in which these functions and unknown parameters of the model (ρ, ν, δ and q) depend on individual characteristics. We allow the cost of search, the separation rate, the value of leisure, and the mean of the wage distribution to depend on gender and three levels of education (low, medium and high, defined in Table 1), which we denote by $\mathbf{x_1}$. Furthermore, the cost of search and the value of leisure are functions of the household type $\mathbf{x_2}$ (head of household, cohabitant or single), since this affects the preferences for work and is informative of other income in the household. Following the traditional labor supply literature (e.g. Mroz, 1987), we exclude the household type as determinant of the wage distribution and the job separation rate. Let $\mathbf{x} \equiv (\mathbf{x_1} \; \mathbf{x_2})$. Finally, in view of the limited set of observable individual characteristics, the cost of search is also allowed to depend on unobservables u, independently distributed of $\mathbf{x}$. Because of computational limitations, we exclude this dependence from the other functions and impose that u follow a discrete distribution with two points of support: $u \in \{v_1, v_2\}$, where Q_1 and $Q_2 = (1 - Q_1)$ respectively denote the probability that $u = v_1$ and $u = v_2$ at entry in unemployment.

More specifically, the value of leisure and the job separation rate are specified as $\nu(\mathbf{x}) = \mathbf{x}'\zeta_\nu$ and $\delta(\mathbf{x_1}) = \exp\{\mathbf{x_1}'\zeta_\delta\}$, respectively. The functional form for the cost of search effort is

$$c(s; \mathbf{x}, u) = e^{\mathbf{x}'\zeta_\varepsilon + u} [e^{\varepsilon s} - 1], \tag{12}$$

where $\varepsilon > 0$. Since $\mathbf{x}$ includes a constant term, by normalization $v_1 = 0$ to represent the reference category: a male, highly educated head of household with $u = v_1$.

The heterogeneity u is unobserved by the econometrician, but not necessarily by the caseworker. Therefore, the realization of u can affect not only search effort, but also the probability of a negative evaluation. At the kth interview and for someone who has not left unemployment since notification, $i = 0$ (respectively, someone who has returned to unemployment after a temporary job, $i = 1$), we assume that this probability takes the following functional form:

$$\pi_{k,u}^0 \left[\bar{S}^0 (\tau, t_{k-1}; \mathbf{x}, u) \right] = \exp \left\{ -(\alpha_{k,u} + \beta_k \, \bar{S}^0 (\tau, t_{k-1}; \mathbf{x}, u)) \right\}, \; \alpha_{k,u}, \beta_k \geq 0, \; u \in \{v_1, v_2\} \tag{13}$$

$$\pi_k^1 = \exp\{-\gamma_k\}, \; \gamma_k \geq 0. \tag{14}$$

This specification depends on the observed characteristics $\mathbf{x}$ only through the average search effort. We also experimented with a specification in which direct dependence on $\mathbf{x}$ was allowed for. However, since it did not have any implications on the dependence of this probability on average search effort, we did not retain it in the final estimations to avoid further increasing the computational complexity. In Subsection 6.1, we will justify why we maintain the dependence on unobserved characteristics u.

The net wage offer density $f(w)$ is assumed to be log-normal: $w \sim \mathcal{LN}(\mu, \sigma)$, with $\mu(\mathbf{x_1}) = e^{\mathbf{x_1}'\zeta_\mu}$. Observed net wages w^o are measured with a multiplicative error m: $w^o = w \cdot m$, and the density

function of the measurement error $h(m)$ is a unit-mean log-normal: $m \sim \mathcal{LN}\left(-\omega^2/2, \omega\right)$. Following Christensen and Kiefer (1994), it can be shown that the density function of observed *accepted* wages $f_o(w^o; \tau)$ if unemployment is left at τ is given by

$$f_o(w^o; \tau) = \int_0^{w^o/w_r^i(\tau)} \frac{f(w^o/m)}{\bar{F}[w_r^i(\tau)]} \frac{1}{m} h(m) dm. \tag{15}$$

5.2 Identification

Throughout this subsection, except for the parameters of unobserved heterogeneity, we discuss the identification of the unknown parameters for a subpopulation of individuals with a given value of the discrete observed characteristics $\mathbf{x}$. The data contain monthly information on the transition rate from unemployment to employment ($p(\tau)$) and non-employment, the net accepted wage (w^o), and the hazard rate from employment back to unemployment (δ). Since transitions to non-employment are not modeled, they are treated as right-censored observations. The identification of the reservation wage ($w_r(\tau)$), the wage offer distribution (μ and σ) and the variance of the measurement error (ω^2) is standard (Flinn and Heckman, 1982; Eckstein and van den Berg, 2007; Keane et al., 2011). By the assumption that the wage offer distribution is log-normal, the complete distribution can be recovered from the truncated distribution of accepted wages. Moreover, the parametric assumption on the measurement error in observed wages is sufficient to identify the complete wage offer distribution (μ and σ) and the variance of the measurement error (ω^2). Given that the reservation wage and the wage offer distribution are identified, the job arrival rate, i.e. the *effective* search effort $\tilde{s}(\tau) \equiv \lambda s(\tau)$, can be recovered from the transition rate to employment: $\tilde{s}(\tau) = p(\tau)/\bar{F}(w_r(\tau))$.

A distinguishing feature of our study is that we follow individuals across different job search monitoring regimes in which the expectations about future benefit levels differ. The variation in expectations *across* monitoring regime stages ($k \in \{1, 2, 3, 4\}$) generates additional variation in the job search and acceptance behavior, which helps identify more unknown parameters than when only one regime is considered. Formally, if the behavior is stationary within each of these regimes, each regime generates two independent FOC, allowing each time to identify two unknown parameters. In principle, this could identify ten unknown parameters: two in the regimes prior to notification or after a positive evaluation, two after notification, and six after the three negative evaluations. In Online Appendix D, we show how we can thus identify the value of leisure $\nu(\mathbf{x})$, the scale parameter of the cost of search function ζ_ε and the ratio ε/λ. The separate identification of parameters ε and λ from the ratio ε/λ would require information on job search effort, such as e.g. the number of job applications sent out per time period. The absence of this information makes us normalize λ to unity. In principle, we should also be able to identify the discount rate ρ by exploiting this variation across monitoring regimes. However, since trial estimates converge to unreasonably high values, following van den Berg and van der Klaauw (2015), we decided to fix the discount rate to 5% per year. This means that in the end we exploit the across-regime variation to identify only one parameter in addition to the two that are already determined by the FOC in a single regime. Our model is therefore overidentified. Yet, in practice, the overdentifying restrictions in this study are weak, since we hardly observe any individuals who reach the second and third meetings.

The structural model implies that job search effort should be increasing and the reservation wage decreasing over the different stages of the monitoring scheme. As a result, the hazard rate to employment must be an increasing function of unemployment duration. However, the observed hazard rate exhibits negative duration dependence (Subsection 6.2 below). This negative duration dependence is assumed to be induced by the dynamic sorting over the unemployment spell, which causes the more employable individuals to leave earlier than the less employable ones. By matching the observed transitions (conditional on the observables $\mathbf{x}$) to those implied by the theoretical model, we can identify the distribution of unobserved heterogeneity (see e.g. Paserman, 2008), i.e. Q_1 and v_2.

The data provide information on the timing of the notification (t_0) and of the monitoring interviews ($\{t_k\}_{k=1}^3$), as well as on the outcomes of these interviews ($\{O_k\}_{k=1}^3$). Since we observe few individuals who are monitored for a second and third time, the identification strategy of $\alpha_{k,u}, \beta_k$ and γ_k depends on whether $k = 1$ or $k \in \{2, 3\}$. We first explain the identification procedure for $k \in \{2, 3\}$ under the assumption that the parameters at the first meeting are identified. Subsequently, we discuss the identification of the latter parameters.

We adopt the following identifying assumption. The sample average of the expected sanction probability at meeting 2 (respectively, 3) relative to that at the first meeting in the sample is equal to the corresponding relative probability in the population:

$$\forall k \in \{2, 3\} : \frac{\bar{\pi}_k}{\bar{\pi}_1} = \frac{\pi_k^a}{\pi_1^a} \tag{16}$$

The right-hand side of this equality is the ratio of the aggregate fractions π_k^a in the population reported at the end of Section 2. On the left-hand side, the denominator can be estimated from the data, and the numerator is a function of the unknown parameters of the sanction probabilities at meeting 2 (respectively, 3). We can thus infer the latter if we assume that *all* parameters at the kth meeting (for $k \in \{2, 3\}$) are tied to the estimated parameters of the first meeting by the same factor of proportionality κ_k: $\forall k \in \{2, 3\}, u \in \{v_1, v_2\} : \hat{\alpha}_{k,u} = \kappa_k \hat{\alpha}_{1,u}, \hat{\beta}_k = \kappa_k \hat{\beta}_1$, and $\hat{\gamma}_k = \kappa_k \hat{\gamma}_1$. We choose κ_k to satisfy the equality in (16). Online Appendix E provides more details.[23]

Consider now the identification of $\alpha_{1,u}, \beta_1$ and γ_1 determining the probability of a negative evaluation at the first meeting. As for γ_1, this parameter is directly identified from the fraction of individuals evaluated negatively at the first interview among those whose unemployment spell had been interrupted by temporary employment prior to the assessment. In Online Appendix D, we demonstrate that, with the available data, we identify the ratio β_1/λ. Consequently, by the aforementioned normalization $\lambda = 1$, we identify β_1.

As for α_{1,v_1} and α_{1,v_2}, the identification argument is as follows. The observed fraction of negatively evaluated individuals in the sample is a weighted average of the probabilities of negative evaluations for individuals of each unobserved type. Since the dynamic sorting implies that the fraction of unobserved types evolves with unemployment duration according to each individual's search effort and reservation wage and because these fractions are identified, the weights are known and vary across individuals. The individual variation in these weights then ensures that α_{1,v_1} and α_{1,v_2} can be separately identified. In the specification of the likelihood function in (F-5) in Online Appendix F, we show how the observed fraction of negatively evaluated individuals can be written as a weighted average.

Summing up, we can identify all the parameters of our model once we have normalized λ to unity. Henceforth, given this normalization, the arrival rate of job offers is simply equal to the level of search effort.

6 Results

6.1 Estimated Parameters

Table 4 presents the parameter estimates. During the maximization of the likelihood function, the mass point v_2 was systematically converging to very high values, meaning that the cost of search of type v_2 individuals tends to infinity and hence their effective search effort (i.e. their exit probability) converges to zero. Therefore, the ultimate specification of the mixture removes v_2 and places a positive probability mass at infinite duration. The share of the unemployed who do not search for a job is equal to 20.4% at entry (see the estimated value of Q_2 in Table 4). This high share can be explained by the leniency of the UI in Belgium where, before the reform, non-searchers could claim UI

---

[23]See Ridder and van den Berg (2003) for another example in which the parameters of interest can be identified from aggregate data if individual micro data are unavailable.

Table 4: Estimated Parameters [a]

		Coeff.	(S.E.)			Coeff.	(S.E.)
ζ_ε	constant	1.092	(1.225)	ζ_δ	constant	−4.082	(0.174)
	female	−5.418	(3.175)		female	0.475	(0.150)
	skill-low	−1.722	(0.882)		skill-low	1.225	(0.203)
	skill-med	−1.568	(0.803)		skill-med	0.791	(0.185)
	single	4.268	(1.033)				
	cohabitant	4.401	(1.015)				
ζ_ν	constant	−0.795	(0.203)	ζ_μ	constant	7.040	(0.090)
	female	0.615	(0.389)		female	−0.170	(0.055)
	skill-low	1.931	(0.614)		skill-low	−0.036	(0.037)
	skill-med	1.018	(0.599)		skill-med	−0.055	(0.032)
	single	1.375	(0.498)				
	cohabitant	4.431	(0.535)				
$\ln(\varepsilon)$		1.735	(0.539)	α_{1,v_1}		1.563	(1.286)
σ		0.175	(0.030)	$\alpha_{1,\infty}$		0.852	(0.206)
ω		0.067	(0.008)	β_1		0.000	−[c]
q		0.159	(0.032)	γ_1		1.705	(0.320)
Q_2		0.204	(0.018)				

Log-likelihood   −3969.86

Probability π of negative evaluation	1st meeting	2nd meeting[b]	3rd meeting[b]
Non-job seekers	0.43	0.49	0.57
Job seekers	0.21	0.27	0.36
Job seekers after job interruption	0.18	0.24	0.32

[a] Symbols: ζ_ε: Coefficient vector of covariates in the cost of search function;

ζ_ν: Coefficient vector of covariates in the value of leisure, i.e. $\nu(\mathbf{x})$ measured in hundreds of €;

$\ln(\varepsilon)$: ln of the multiplier of search in the cost of effort;

σ: Standard deviation of the log-wage offer distribution;

ω: Standard deviation of the log-measurement error distribution;

q: Arrival rate of the meeting in the delay period;

Q_2: Share of people at entry with infinite cost of search, i.e. non-job seekers;

ζ_δ: Coefficient vectors of covariates in the separation rate;

ζ_μ: Coefficient vector of covariates in the mean of the wage distribution;

α_{1,v_1}: Constant term in the probability of negative evaluation at the 1st meeting when the cost of search is finite and $i = 0$;

$\alpha_{1,\infty}$: Same parameter but when the search cost tends to ∞;

β_1: Coefficient of $\bar{S}(t_1, t_0)$ in the probability of negative evaluation at the 1st meeting;

γ_1: Constant term in the probability of negative evaluation at the 1st meeting if $i = 1$.

[b] The probability of a negative evaluation at the 2nd and 3rd meetings is obtained by multiplying the parameters determining this probability at the 1st meeting by $\kappa_2 = 0.829$ and $\kappa_3 = 0.653$, respectively. The latter are solved such that the sample averages of the individual probabilities at the second and third interviews are compatible with the aggregate observed frequencies (see Subsection 5.2 and Online Appendix E).

[c] Since the estimate of β_1 lies on the boundary of the parameter space, the standard error is not computed.

Figure 2: The Share of the Unemployed Searching for a Job
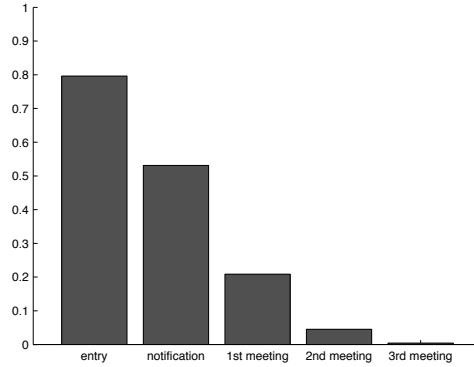
indefinitely without their job search being monitored throughout the spell. The introduction of job search monitoring in 2004 cannot have affected the composition of the UI benefit claimants retained in the sample analyzed, since these individuals were informed by surprise only after 14 months that their job search effort would subsequently be monitored. However, even when the scheme is fully instituted and UI benefit claimants are informed from the start of unemployment that their job search will be monitored, there is no reason initially for these non-job seekers not to continue claiming UI rather than the lower means-tested welfare benefits, since monitoring starts so late in the unemployment spell. This is very different from the institutional environment in several other OECD countries, where job search monitoring starts early in the unemployment spell and is much more intensive, so that from the start non-job seekers do not claim UI. We return to this point in Subsection 7.4 where we make an international comparison and explain that the effectiveness of the Belgian monitoring scheme is not only much reduced by the drawn-out timing of the monitoring interviews, but also by this compositional effect. Figure 2 illustrates that dynamic sorting leads to negligible shares of the unemployed searching for a job upon reaching the second and third meetings. Consequently, as the scheme that we evaluate to a large extent monitors individuals with extremely high search costs, the behavioral impact and the implied budgetary savings that the scheme generates cannot be important.

In the specification, we allow the probability of a negative evaluation $\pi^0_{k,u}(.)$ to depend on the unobserved characteristics u. We allow for this dependence, because the aforementioned estimation result revealed that a non-negligible fraction of UI claimants are not searching for a job: we would not have introduced this dependence if the point of support v_2 of the unobserved heterogeneity had converged to a finite value. Benefit claimants are aware that some caseworkers have the technology to identify who is not searching for a job and that caseworkers are more likely to evaluate the behavior of such individuals negatively. However, since this technology is imperfect and since caseworkers have the discretion to consider these non-job seekers, typically people with health or psychological problems, or singles caring for young children, as "deserving", not all of them will receive a negative assessment. The probability of a negative evaluation is estimated to be 0.43 at the first meeting, 0.49 at the second one and 0.57 at the third.

Benefit recipients who do search for a job are about half as likely to be negatively evaluated at the first meeting as non-seekers: 0.21. At the two subsequent evaluations, a sanction is imposed respectively with probability 0.27 and 0.36. While raising job search effort at the *extensive* margin pays, it does not pay at the *intensive* margin: $\hat{\beta}_1$, and, hence, the corresponding parameters at later meetings converge to zero. This finding is robust to the inclusion of the observed characteristics $\mathbf{x}$ as control variables in the specification of $\pi^0_{k,u}(.)$ (not reported). The convergence of $\hat{\beta}_1$ to the boundary of the parameter space is not the consequence of non-identification due to lack of variation in $\bar{S}^0(\tau, t_{k-1}; \mathbf{x}, \hat{v}_1)$. The coefficient of variation of this variable in the sample of individuals present at

the first meeting and used in the estimation of $\hat{\beta}_1$ is 0.67. Thus, the issue is rather that this variation is unrelated to the probability of a negative evaluation. Apparently, search effort is measured with too much error and/or caseworkers have too much discretion to ensure that the probability of a negative evaluation decreases marginally with job search intensity. Consequently, the additional search incentive identified by the second term on the right-hand side of the FOC for job search effort in Equation (11) is zero and eventually does not play a role. In Section 7, we will nevertheless illustrate its impact based on simulations. Interrupting unemployment to take a temporary job ($i = 1$) further reduces the likelihood of a negative assessment to 0.18 at the first meeting, and to 0.24 and 0.33 at the two subsequent interviews.

According to our estimates, the marginal costs of search are lowest for women and the low educated. These findings are unexpected, since these two groups are usually found to be disadvantaged on the labor market. A possible explanation is that employers recognize that these groups have more difficulties on the labor market, so that being unemployed does not convey as negative a signal for them as for more advantaged groups. The latter are expected to directly transit to employment when they enter the labor market and to transit to other jobs through on-the-job search. Hence, employers may be less likely to offer jobs[24] to these advantaged groups when they are unemployed (Eriksson and Gottfries, 2005; Eriksson and Lagerström, 2006; Longhi, 2015). By contrast, heads of households may perceive lower marginal costs, because their family responsibility may push them to find a job more quickly than cohabitants or singles.

In line with expectations, the value of leisure is higher for women, the low educated, cohabitants, and to a lesser extent, for singles. The high estimated value of leisure for cohabitants (443 €/month) suggests that household activities within this group are highly valued. By contrast, the reference individual (highly educated male head of household) seems to perceive stigma costs to being unemployed of about 80 €/month.

Disadvantaged workers on the labor market, such as women and the low educated, are more likely to separate from jobs. Women enter lower paying jobs more often than men do, but the educational level does not seem to play a significant role in the determination of entry wages. This finding is in line with the aforementioned interpretation that the highly educated are less likely to be offered jobs, because of the negative signal their unemployment conveys. Hence, if they are offered a job, it does not pay more than for the low educated. Note that the standard deviation of the log wage distribution is estimated to be about 6.7% of the mean offered log wages. It accounts for just 13% ($= \hat{\omega}^2 / (\hat{\omega}^2 + \hat{\sigma}^2)$) of the variance of the observed accepted log wages. This relatively small measurement error suggests that the log-normal wage offer distribution fits the observed wages quite well.

Finally, q is estimated to be equal to 0.159. Since the delay is capped by the maximum observed one (see Subsection 4.4), the average delay is 5.75 months, which is somewhat shorter than the average of the corresponding untruncated exponential distribution, i.e. $1/0.159 = 6.3$, but still remains substantial. This points to an additional source of inefficiency in the scheme.

6.2 Internal Validation: Goodness of Fit

This subsection reports the within-sample fit of the model. First, in the left panel of Figure 3, the solid line represents the hazard function of job finding since notification as predicted by the structural model, and the dashed line, the smoothed non-parametric counterpart (see Tanner and Wong, 1983). The optimal bandwidth for this non-parametric estimator is chosen by cross-validation, as suggested by Tanner and Wong (1984). The 95% confidence interval around the smoothed non-parametric hazard function (dotted lines) is based on the bootstrap with 1,000 replications. The right panel of Figure 3 displays the density function of observed net monthly earnings predicted by the structural model (solid line) and the kernel density of observed wages from the data (dashed line). The Kernel function

---

[24]Recall that our model cannot discriminate between the marginal effect of search on costs for the unemployed and on the job arrival rate (see Subsection 5.2).

Figure 3: Goodness of Fit of the Model
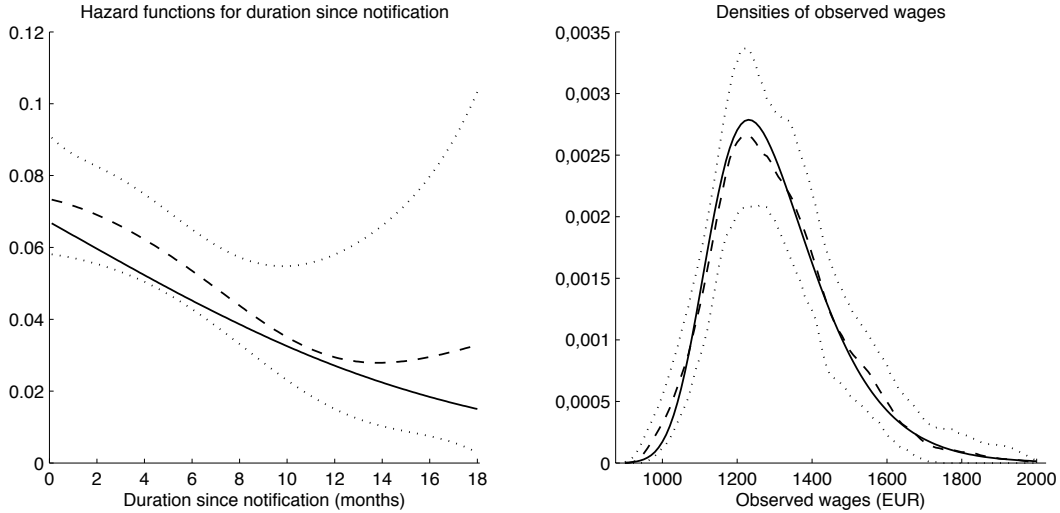
The solid lines display the predictions of the hazard and the wage density based on the simulated structural model. The middle dashed lines represent the smoothed non-parametric hazard to employment based on the notified individuals in the sample (left panel) and the density of observed accepted wages (right panel). The dotted lines around the dashed lines delineate the 95% confidence intervals around the aforementioned non-parametric estimates.

is Epanechnikov and the optimal bandwidth is Silverman. The 95% confidence interval around the kernel density estimate (dotted lines) is likewise based on the bootstrap with 1,000 replications.

The sorting induced by unobserved heterogeneity generates negative duration dependence in the hazard rate to employment. The fitted hazard of job finding and density of observed wages are almost everywhere within the confidence bounds of the corresponding nonparametric estimates. This underlines a very good fit of our structural model to the data.

6.3 External Validation

We propose an out-of-sample validation that is somewhat different from the standard approach (see e.g. Todd and Wolpin, 2006). In the standard approach, researchers would typically use pre-reform data to check whether the predictions of the model about the impact of the program track the (non-)experimental post-reform impact estimates. However, we could not follow such a strategy,[25] since it turns out that economic conditions were notably worse in the pre-program period than during the time when the program was implemented. GDP real growth reached only 0.8% in 2003 against 2.7% on average between 2004 and 2006, which adversely affected pre-reform exit rates to employment. This would therefore require that our structural model not only predict the impact of the program, but also the impact of the improved economic conditions, which is clearly too ambitious.

We therefore propose a less ambitious validation exercise. We aim at checking whether, in the absence of the policy reform, a stationary version of our structural model can reproduce the dynamic sorting between entry in unemployment and the date of sample selection (i.e. 14 months later), accounting for observed and unobserved heterogeneity as estimated. If the model reproduces the pre-reform sorting correctly, we can safely perform counterfactual policy experiments with notification set at any point in time. Of special interest will be the notification at the moment of entry to

---

[25]Notice that this strategy would in any case be partial, since the parameters of the probability of a negative evaluation can only be identified on the basis of post-reform data.

20

unemployment, which corresponds to the monitoring scheme not coming as a surprise anymore (see Section 7).

In order to implement this validation exercise, we have obtained from the UI agency an additional random sample of individuals born and starting their unemployment spell at the same time period as those retained for the estimation. On this sample, we simulate our model until the counterfactual sample selection date, i.e. at an elapsed unemployment duration of 14 months. We then check whether the distribution of observed personal characteristics in this simulated external sample is statistically indistinguishable from the corresponding distribution in the sample that was selected for estimation. This is not obvious, since the unobserved heterogeneity induces sorting over time which changes the composition of observables over the course of the unemployment spell at different speeds: among those with observed characteristics that enhance employability, this sorting process operates more rapidly than among those who are less employable (see e.g. Ridder, 1984).[26]

The simulation is repeated 1,000 times to account for the classification error of an individual searching or not for a job, as well as to account for sampling error in both the inflow sample and the sample retained for estimation. The first column of Table 5 reports sample fractions, $\hat{r}$, of the mutually exclusive and exhaustive partition of individual characteristics in the sample that we use to estimate our structural model. The second column of this table reports sample fractions, $\hat{r}_{sim}$, in the very same partition of individual characteristics at the counterfactual date of selection simulated using the aforementioned inflow sample into unemployment. The last two columns show the t-statistic and the p-value of the test of $H_0 : r = r_{sim}$, variable by variable, respectively.

Table 5: Distribution of Personal Characteristics at the (Counterfactual) Selection

Variable	$\hat{r}$ (s.e.)	$\hat{r}_{sim}$ (s.e.)	t-stat.	p-Value
"female" × "skill-low" × "single"	0.021 (0.005)	0.020 (0.006)	0.116	0.908
"female" × "skill-low" × "cohabitant"	0.030 (0.006)	0.065 (0.011)	−2.850	0.004
"female" × "skill-low" × "hh-head"	0.081 (0.009)	0.059 (0.010)	1.560	0.119
"female" × "skill-med" × "single"	0.062 (0.008)	0.063 (0.011)	−0.085	0.933
"female" × "skill-med" × "cohabitant"	0.086 (0.010)	0.096 (0.013)	−0.601	0.548
"female" × "skill-med" × "hh-head"	0.054 (0.008)	0.061 (0.010)	−0.589	0.556
"female" × "skill-high" × "single"	0.046 (0.007)	0.059 (0.010)	−1.034	0.302
"female" × "skill-high" × "cohabitant"	0.059 (0.008)	0.063 (0.011)	−0.369	0.712
"female" × "skill-high" × "hh-head"	0.013 (0.004)	0.011 (0.004)	0.405	0.686
"male" × "skill-low" × "single"	0.112 (0.011)	0.119 (0.014)	−0.400	0.690
"male" × "skill-low" × "cohabitant"	0.044 (0.007)	0.036 (0.008)	0.730	0.466
"male" × "skill-low" × "hh-head"	0.062 (0.008)	0.051 (0.010)	0.841	0.400
"male" × "skill-med" × "single"	0.105 (0.010)	0.095 (0.013)	0.637	0.524
"male" × "skill-med" × "cohabitant"	0.063 (0.008)	0.048 (0.009)	1.249	0.212
"male" × "skill-med" × "hh-head"	0.025 (0.005)	0.033 (0.008)	−0.788	0.431
"male" × "skill-high" × "single"	0.055 (0.008)	0.045 (0.009)	0.910	0.363
"male" × "skill-high" × "cohabitant"	0.079 (0.009)	0.067 (0.011)	0.826	0.409
"male" × "skill-high" × "hh-head"	0.005 (0.002)	0.011 (0.005)	−1.276	0.202

We see that, on the variable-by-variable basis, we do not reject the hypothesis that the relative sizes of all cells in the partition of individual characteristics, except for the "female" × "skill-

---

[26]Furthermore, if the neglect of declining benefits for *laid-off* cohabitants and singles at the beginning of the unemployment spell (see footnote 13) mattered, then this should also lead to a rejection of the validation test.

low" × "cohabitant" one, are equal at any standard level of significance. The test statistic of joint equality of the actual distribution of covariates and the distribution of covariates generated by the model, $H_0 : \mathbf{r} = \mathbf{r}_{sim}$, is given by

$$(\hat{\mathbf{r}} - \hat{\mathbf{r}}_{sim}) \, \hat{\mathbf{\Sigma}}^- \, (\hat{\mathbf{r}} - \hat{\mathbf{r}}_{sim})' \sim \chi^2_{rk[\mathbf{\Sigma}]}$$

where $\hat{\mathbf{\Sigma}}^-$ is the Moore-Penrose generalized inverse of the estimated covariance matrix of the difference between actual and model-based cell frequencies[27] and $rk[\mathbf{\Sigma}]$ denotes its rank (see Andrews (1988a, b) for a general theory and applications). Since these 18 cell frequencies add up to one, we have that $rk[\mathbf{\Sigma}] = 17$ and the value of this test statistic is 18.95 with a p-value of 0.33. Hence, we cannot reject the hypothesis that the actual and model-based distributions of covariates at the moment of selection are equal.

6.4 Implied Behavioral Responses

Since the parameter estimates are not informative about the behavioral responses to the monitoring scheme over time in terms of search effort and reservation wage, we simulate the optimal behavior of a job seeker with average observed characteristics. We assume that she does not leave unemployment before the third meeting, she does not experience any temporary employment after notification ($i = 0$), the outcome of the evaluations are always negative and there are no delays in the timing of the meetings.

Figure 4 displays the time paths of the simulated behavioral responses. It comprises four panels. The upper-left panel plots the evolution of effective search effort, i.e. the monthly arrival rate of job offers. The upper-right panel traces the monthly net reservation wage (in euros) and the lower-left one, the corresponding acceptance rate. Finally, the lower-right panel shows how the product of the job arrival rate and the acceptance probability translates into the monthly transition rate towards full-time employment. On the horizontal axis, duration, measured in months, is normalized to zero at notification. The solid lines display the behavior predicted by the model evaluated at the parameter estimates reported in Table 4, apart from $q = 0$ (no delays). The dashed lines will be discussed in Subsection 7.5.

Before notification, nearly 0.70 job offers arrive on average per month and about 20% of these jobs are accepted, so that the average monthly exit rate is about 14% ($0.7 * 0.2 = 0.14$). Since the environment is stationary before notification, this corresponds to an average unemployment duration of about 7 months. After the third negative evaluation, the unemployed are permanently sanctioned and entitled to the lower welfare benefits. This raises the job arrival rate to about 0.77 and the job acceptance probability to about 27%, resulting in an exit rate to employment of about 21% and an average residual unemployment duration of slightly less than 5 months. The search effort, acceptance probability and job finding rate rise, as a consequence of anticipatory behavior, gradually between meetings, and abruptly just after each negative assessment, from the levels before notification to those after the third negative evaluation. The job finding rate is eventually raised by about 50%, which is substantial. However, notice that the rate of increase in all three variables is very slight between notification and the first interview. This is because the first negative evaluation does not involve an immediate sanction, but only the threat of a temporary withdrawal of benefits at the second assessment. The transition rate to employment is actually not substantially affected before the second meeting. It surges most after the third negative evaluation when unemployment benefits are permanently withdrawn. The spread-out timing of the monitoring interviews substantially reduces the effectiveness of the scheme in raising the exit rate to employment of the notified unemployed. This is further reinforced by the delays in the scheduled timing of the meetings, which are ignored

---

[27]By independence between these samples, the covariance matrix of the difference is equal to the sum of the estimated covariance matrices.

Figure 4: Predicted Optimal Behavioral Responses to the Monitoring Scheme at the Sample Average [a]
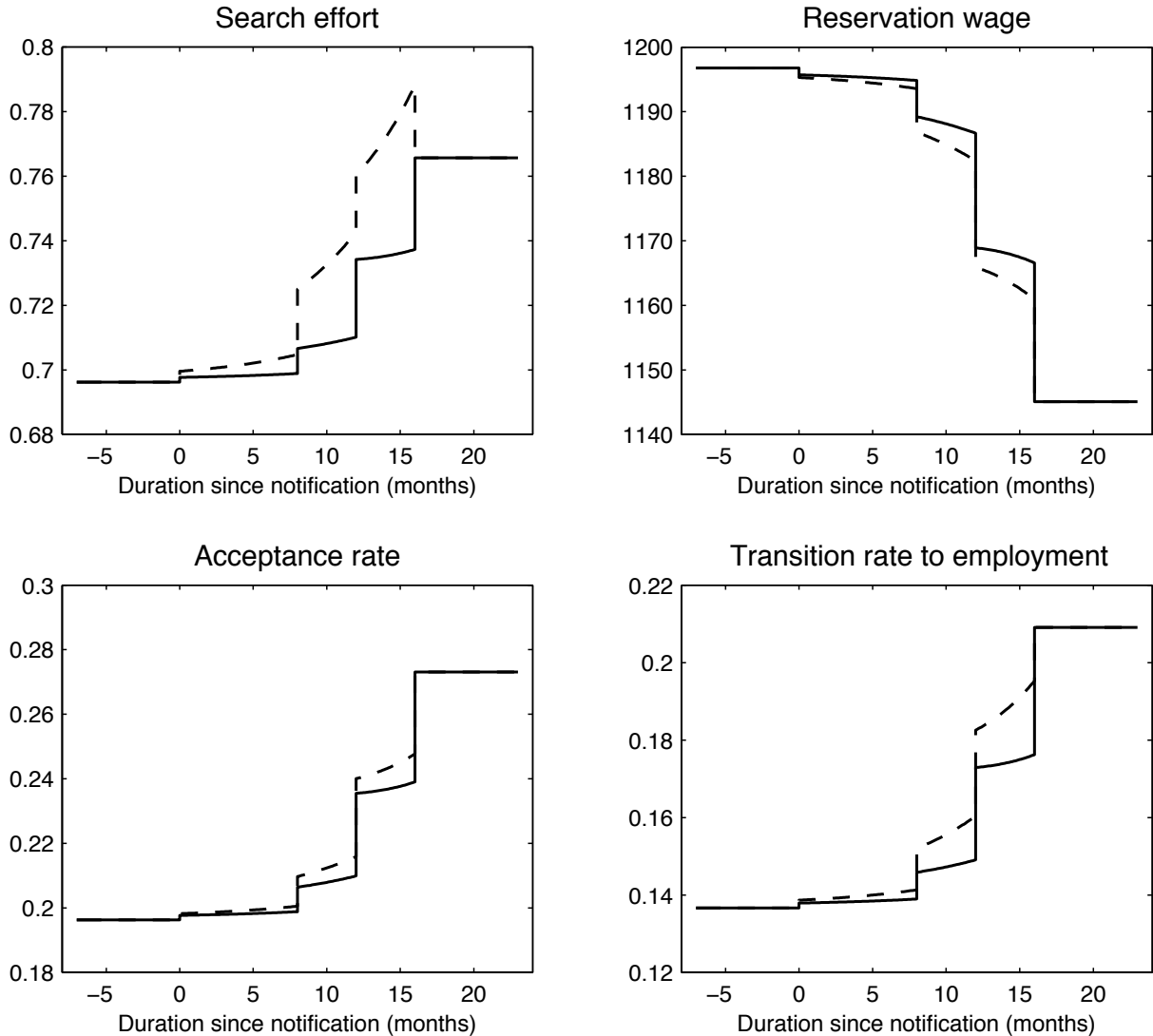
[a] For the parameter estimates reported in Table 4 and a job seeker with average characteristics, the solid lines display the time profile of the behavioral responses to the monitoring scheme in the absence of delays (setting $q = 0$), and conditional on remaining unemployed and being negatively evaluated at each meeting. The dashed lines display these profiles for a monitoring technology that generates "front-loading", i.e. for $\alpha_{1,v_1} = \hat{\alpha}_{1,\infty} = 0.85$ and $\beta_1 = 0.82$ (See Subsection 7.5). Without delays the 1st, 2nd and 3rd meetings take place respectively 8, 12 and 16 months after notification.

in Figure 4 for illustrative purposes. Beyond the second interview - the point from which the impact of monitoring starts to matter - about 90% of notified individuals are actually not seeking a job (cf. Figure 2) and, hence, are not influenced at all by the job search assessments. This means that the monitoring scheme did not manage to raise the average job finding rate of notified individuals in any significant way. In the next subsection, we investigate whether a change in the design of the scheme could affect this conclusion.

7 Policy Evaluation

An acknowledged advantage of structural estimation is that it allows simulation and the subsequent evaluation of alternative policy designs in a behaviorally consistent framework. In this section, we

evaluate the current monitoring policy, as well as its sensitivity to some alternative design features. In doing so, we consider the average treatment effects on the treated (ATT). Throughout this section, the treatment effect is defined relative to an environment without monitoring of search effort where a flat benefit level b_h is paid out indefinitely. We consider five indicators: (i) the expected unemployment duration, (ii) the expected lifetime utility of an unemployed individual, $U_k^0(t_0)$, (iii) the expected lifetime labor earnings, which differs from the previous indicator in that it excludes unemployment and assistance benefits, the value of leisure, and the cost of search, (iv) the expected expenditures of public authorities on unemployment and assistance benefits, and on implementing job search assessments,[28] (v) the "net social gain", which measures the difference between $U_k^0(t_0)$ (indicator ii) and the expected public expenditures (indicator iv). We must stress that this is *not* a measure of "social welfare" as the data required that, like most researchers, we impose risk-neutrality.[29] The last four indicators are computed as the average net present values discounted from the start of the treatment.[30] Their formal definitions can be found in Section H of the Online Appendix.

We start by evaluating the existing 2004 policy design. First, we simulate the ATT of introducing the monitoring scheme on the population of 25- to 30-year-olds with and elapsed unemployment duration of 14 months ($t_0 = 14$) notified by surprise that their search effort will be monitored for the first time eight months later ($t_1 = 14 + 8 = 22$). Once the monitoring scheme was fully operational, the notification no longer comes as a surprise after 14 months. Thus, next, we repeat the evaluation for the sample of entrants in unemployment that we used in the external validation reported in Subsection 6.3, assuming that the notification arrives as of entry ($t_0 = 0$) and that the first interview is scheduled after 22 months. We then simulate the ATT under two major sets of alternative policy reforms in which we alter some design features of the monitoring scheme. In these simulations, we maintain the assumption that benefit recipients are informed about the reform as of entry into unemployment. In the first set of reforms, we keep the monitoring technology unaffected, but change the scheduled timing of the monitoring interviews and study the effect of introducing a temporary sanction in case of a negative evaluation at the first assessment. In the second set of reforms, we consider the same reforms as in the first set, but increase the precision of the monitoring technology by allowing the probability of a negative evaluation $\pi_{k,u}^0[.]$ at all meetings to depend on the average search effort since the previous meeting, i.e. we allow $\beta_k > 0$. A brief subsection puts our results in international perspective by comparing them to those obtained from other structural models estimated on U.S. and Dutch data. In a final subsection, we discuss an interesting property, which we label "front-loading" of search effort. We demonstrate that when the monitoring technology is sufficiently sensitive to search effort (i.e. β_k sufficiently large), the latter may even exceed the post permanent sanction level.

7.1 The Effect of the 2004 Reform

In Panel A of Table 6, we report the effects of introducing the monitoring scheme on the long-term unemployed workers notified by surprise in their 14th month of unemployment. We disentangle the total effects into those for non-job seekers ($u = +\infty$) and job seekers ($u = v_1$).

Monitoring does not affect the unemployment duration of non-job seekers. The loss in UI payments when there is a sanction reduces the expected discounted lifetime utility on average by 6,022 €. This loss in UI payments is a direct transfer to the UI agency, which must, however, also finance the operating costs possibly of multiple job search assessments. Net of financing operating costs, the expected discounted gain is 5,871 €.

Among job seekers notified at $t_0 = 14$, monitoring reduces the average unemployment duration by about 6 days (0.18 months). Monitoring imposes a much lower utility loss on job seekers than on

---

[28]Based on accounting information from the UI agency, each assessment costs 100 € on average.

[29]See footnote 16.

[30]Since our framework assumes rational expectations, an individual is "treated" from the moment she is informed about the monitoring process, because she can anticipate the monitoring as from that moment, i.e. at notification when the program was introduced, but once the scheme was fully operational, at the start of unemployment.

Table 6: ATT for the Benchmark/Enhanced Monitoring Schemes and for Policy Reforms

	Expected Unemployment Duration (months)	Lifetime Utility (€)	Lifetime Earnings (€)	Government Discounted Expenditures (€)	Net social Gain (€)
A. Implemented Monitoring Scheme ($t_0 = 14$)					
○ Non-job Seekers	0.00	-6022	0	-5871	-151
○ Job Seekers	-0.18	-184	325	-368	184
○ All UI Recipients	-0.10	-2789	176	-2817	28
B. Current Monitoring Technology (All UI Recipients)					
B.1. Benchmark Design					
○ $\alpha_{1,\infty}, \alpha_{1,v_1}$ and β_1 as in Table 4					
○ Notification at $t_0 = 0$	-0.09	-1291	215	-1364	73
B.2. Design minimizing unemployment duration:*					
Meeting Month(s) since Meeting $k-1$[†]					
(a) $k=1$ 1	-0.15	-1463	268	-1548	85
(b) $k=2$ 1	-0.09	-1311	218	-1388	77
(c) $k=3$ 4	-0.09	-1291	215	-1364	73
Sanction at Meeting 1[‡]:					
(d) Full sanction (100%)	-0.17	-1624	266	-1619	-5
(e) Combining (a) to (d)	-0.26	-1783	318	-1783	0
C. Enhanced Monitoring Technology (All UI Recipients)					
C.1. Benchmark Design					
○ $\alpha_{1,\infty} = \alpha_{1,v_1} = 0.85, \beta_1 = 0.82$					
○ Notification at $t_0 = 0$	-0.13	-1295	247	-1427	132
C.2. Design minimizing unemployment duration:*					
Meeting Month(s) since Meeting $k-1$[†]					
(a) $k=1$ 1	-0.31	-1501	412	-1739	238
(b) $k=2$ 1	-0.14	-1316	254	-1458	142
(c) $k=3$ 4	-0.13	-1295	247	-1427	132
Sanction at Meeting 1[‡]:					
(d) Full sanction (100%)	-0.24	-1630	314	-1704	74
(e) Combining (a) to (d)	-0.52	-1851	515	-2018	167

Notes:
* Minimizing unemployment duration by varying each dimension (a)-(d) separately within the ranges displayed in Figure 5, while keeping the other dimensions fixed to the benchmark.

[†] Since notification (at entry in unemployment) if $k = 1$.

[‡] Ranging between 0% and 100% of the (temporary) sanction at the second meeting.

25

non-job seekers: 184 vs 6,022 €. As already mentioned in Subsection 6.4, this is because monitoring occurs very late in the unemployment spell and most notified workers, even without modifying their search behavior, have already left unemployment before the second interview takes place and the first sanction is potentially imposed. The additional working time more than compensates for the negative effect on wages, i.e. -2 €/month (not reported in Table 6). Hence, lifetime earnings increase by 325 €. The government's discounted expenditures decrease by 368 €.

Overall, the program reduces the duration of unemployment by about 3 days, and although lifetime earnings increase by 176 €, the expected loss in lifetime utility for the unemployed (2,787 €) is of the same order of magnitude as the government's savings (2,817 €). Thus, the net social gain of the implemented reform is positive, but only marginally so.

Figure 5: The Partial Effects of Alternative Policy Design Features: The Benchmark Monitoring Technology
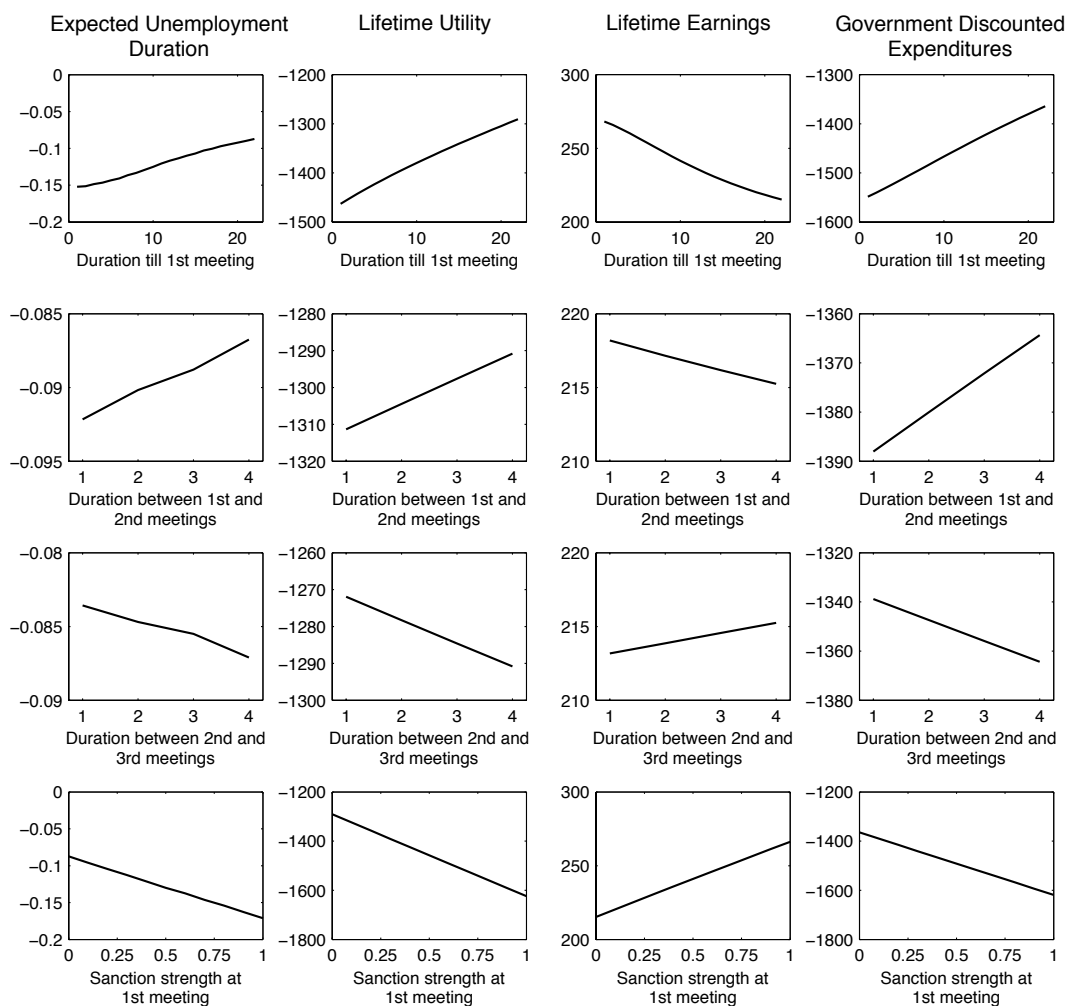
7.2 Policy Reforms without Affecting the Monitoring Technology

The negligible impact on unemployment duration and post-unemployment earnings reported in the previous subsection were obtained because (i) the current monitoring scheme cannot have a behavioral impact on the notified non-job seekers, (ii) the timing of the monitoring is very much spread out over the unemployment spell, and (iii) sanctions are imposed very late. In this subsection, we investigate whether changing the design of the monitoring scheme would lead to more positive conclusions. We measure the average impact on the whole population since caseworkers do not have access to a *costless* technology allowing them to distinguish job seekers from non-job seekers. Before dealing with reforms, first we consider the effects of the scheme that was implemented in 2004 as a benchmark, but instead of assuming that the unemployed were notified by surprise about the monitoring after 14 months (as in Panel A of Table 6), henceforth we consider a fully operational scheme in which workers have already been informed about all the features from the start of the unemployment spell. This benchmark is reported in Panel B.1 of Table 6. Compared to Panel A, the effect on unemployment duration is slightly lower, but the net social gain is somewhat higher.

Next, we study the impact of gradually changing the design features of the monitoring scheme, namely: (a) reducing the scheduled duration from entry in unemployment to the first interview from 22 months to 1 month; (b) reducing the scheduled timing of the second interview from 4 months to 1 month; (c) doing the same for the third assessment; and (d) increasing the sanction intensity at the first interview from 0% to 100% of the temporary sanction imposed in the benchmark scheme at the second interview. To this purpose, we first evaluate the *partial* effects of each of these four design features separately. These effects are displayed in Figure 5. Subsequently, for each design feature, we choose the value that reduces unemployment duration the most over the range of values considered in Figure 5. In lines (a) to (d) of Panel B.2 in Table 6, we report the ATT of these values one by one. In a third step, we combine design features (a)-(d) and impose them simultaneously. The corresponding outcomes are reported in line (e).

Figure 5 illustrates that a stricter monitoring scheme along any of the dimensions (a), (b) and (d) always reduces unemployment duration and lifetime utility, and increases lifetime earnings and government savings. However, a stricter scheme also always entails a higher loss in lifetime utility. Shortening the duration between the second and third meetings, i.e. (c), has the opposite effect on all indicators. Why is this? Being negatively evaluated at the second meeting involves a temporary sanction. From the third meeting onwards, the benefit level can only increase: in case of a negative assessment, the benefit level does not change, while otherwise the higher benefit level, b_h, is paid out indefinitely. Hence, after the third meeting, the incentive to search falls. As shortening the duration between the second and third interviews shifts the moment from which incentives are reduced forward in time, the monitoring becomes more lenient rather than stricter. Hence, unemployment duration is not minimized when the duration between the last two meetings is set to its current value (four months), but we impose the status quo in Table 6.

With the three scheduled intervals set respectively at 1, 1 and 4 months and with the full sanction $b_h - b_l$ imposed already at the first meeting (line (e)), the ATT on the various indicators differ notably from those in the benchmark (Panel B.1): the decline in average duration is almost three times higher, the impact on lifetime earnings (50% higher) and public savings (30% higher) is much more favorable, yet at the cost of larger losses in lifetime utility (40% higher). Whether introducing a more precise monitoring technology (i.e. $\hat{\beta}_k \neq 0$ for all meetings k) can reduce government expenditures even more without substantially harming lifetime utility is the object of the next subsection.

7.3 Policy Reforms in Case of Enhanced Monitoring Precision

The precision of the monitoring technology crucially depends on the sensitivity of the probability of a negative evaluation to average search effort. In this subsection, we study what happens if we increase the precision of this monitoring technology, without making it perfect, however. In order

to focus on the effect of enhancing the precision and not on increasing the level, we first set the probability of a negative evaluation of job seekers with a zero effort level at the first meeting to that of non-job seekers, i.e. $\alpha_{1,v_1} = \alpha_{1,\infty} = 0.85$ and the probability of a negative evaluation is, hence, 43%. Second, for those who search, we raise β_1 to a strictly positive level such that the probability of a negative evaluation at the first meeting remains at its value in Table 4, namely 21%, assuming that average search effort remains unchanged as well. This corresponds to setting $\beta_1 = 0.82$. At the second (resp., third) meeting, β_k is also adjusted and is equal to $\kappa_k \cdot 0.82$ for $k = 2$ (resp., 3).[31] We assume throughout that enhancing the precision of the monitoring technology is costless. This might not be so unrealistic in the Belgian context though, since presumably precision could simply be much improved by decreasing the discretionary power of the counselors in charge of the assessments and replacing it with some strict uniform guidelines.

In Panel C of Table 6, we report how this enhanced precision in the monitoring technology affects the ATT by proceeding in the same way as in Subsection 7.2. To start with, let us compare the results of Panel C.1 to the corresponding ones in Panel B.1. The enhanced precision decreases average unemployment duration by 44% (from -0.09 to -0.13 months). At the same time, lifetime utility falls by only 4 € (less than 0.5% of the loss in Panel B.1) and lifetime earnings rise by 32 € (15% of the gain in Panel B.1), while 63 additional euros are saved on public outlays (about 5% of the savings in Panel B.1). The net social gain is 80% larger than in Panel B.1. Given the share of the unemployed whose behavior is unaffected, this points to a non-negligible reduction in public spending without inflicting too much utility loss (assuming risk neutrality). Observe that this is very different from introducing more "strictness", i.e. contrasting each line, apart from line (c), in Panel B.2 to the benchmark in Panel B.1 (cf. Subsection 7.2). The latter always generates a utility loss of the same order of magnitude or higher than the savings in government expenditures. This contrasting finding partly reflects our assumption that enhancing precision comes at no cost. We have argued that this is not unrealistic for Belgium, but more research on this dimension, in particular for other countries, is required to evaluate whether this justifies reforms of monitoring schemes along these lines.

Next, we can see again that, also with enhanced precision, a stricter monitoring scheme along any of the dimensions (a), (b) and (d) always reduces unemployment duration and lifetime utility, and increases lifetime earnings and the savings made by the government.[32] The combination of features (a) to (d) in line (e) of Panel C.2 leads to an ATT on unemployment duration of half a month. This is twice as large as the effect without enhanced precision (Panel B.2.(e)). Furthermore, a strict monitoring scheme (Panel C.2.(e)) increases government savings more than losses in lifetime utility, which was not the case when the monitoring technology was blunt (Panel B.2.(e)): the net social gain increases from zero to 167 €. This confirms that the precision of the monitoring technology is key.

7.4 International Comparison

When comparing these overall findings to those of two other studies that evaluate job search monitoring on the basis of a similar approach, the impact on unemployment duration is much smaller in our paper. Paserman (2008) reports that monitoring in the U.S. decreases unemployment duration by 4.7 weeks (6.3 weeks) assuming a (Log-) Normal wage offer distribution.[33] In the Dutch case, van den Berg and van der Klaauw (2015) report that compared to the counterfactual of no monitoring, unemployment duration decreases by 3.6 weeks.[34] In both papers, the impact is much greater than that of the monitoring scheme introduced in Belgium: if fully operational, it would reduce expected unemployment duration on average only by 0.09 months, i.e. 0.4 weeks. Part of the explanation is that

---

[31]See note b of Table 5 for more information.

[32]Qualitatively, the pattern is similar to the one displayed in Figure 5. The corresponding figure can be obtained from the authors upon request.

[33]Paserman (2008) focuses his analysis on the unemployed with hyperbolic time preferences, but also considers those with time consistent preferences. We restrict our comparison to the latter case presented in his Table 5.

[34]We deduce this from their Table 5.

the monitoring in Belgium starts late in the unemployment spell and, furthermore, is much spread out over time. The findings in Panel B. 2 line (e) reveal that imposing a much "stricter" monitoring scheme that would be more in line with the U.S. and Dutch schemes makes up part of the difference: average unemployment duration can then be expected to fall by 1.1 weeks. Increasing the monitoring precision further reduces the gap. With such enhanced monitoring precision, the strict monitoring regime shortens the unemployment spell by 2.3 weeks (Panel C.2 line (e)).

We did not check whether by further increasing the precision of the monitoring, the impact would be comparable to what is achieved abroad, but we doubt it for the two following reasons. First, the institutional context of UI is very different in Belgium than in the two aforementioned countries. Through the eligibility based on educational qualification, individuals with little or no work experience can also be entitled to UI, job search requirements are only imposed late in the unemployment spell (affecting the composition at entry) and UI benefits are not exhausted after a particular duration. Consequently, Belgian UI caseloads will include individuals who in most other countries would only be eligible for social assistance. Our estimation finding that job search costs are extremely high for about 20% of UI entrants is in line with this hypothesis. For instance, if we assumed that this group of non-searchers did not claim UI, monitoring would decrease unemployment duration by nearly 3 weeks instead of the aforementioned 2.3 weeks (Panel C.2 line (e)). This further closes the gap with the two other countries, but we have to keep in mind that this probably still underestimates the impact of the different institutional design of UI, because the composition of job seekers is most likely also affected by this design as well.

A second reason for the difference is that our evaluation was limited to a population of 25- to 30-year-olds, while the two other studies did not restrict their analysis to a particular age class. Active labor market policies are generally found to be less effective for the youth (Card et al., 2015, e.g.).

7.5   "Front-Loading" of Search Effort

In this subsection, we illustrate that if β_1 is sufficiently high, the monitoring of search effort could induce "front-loading" of search effort. If the opportunity to avoid a sanction is enhanced by raising search effort, one may raise job search effort before the meeting even above the level attained after a permanent sanction. This can be shown to be quite a general property. In their Proposition 2, Cockx et al. (2011) provide a sufficient condition for search effort to increase above the post-sanction level. The value that is chosen in Subsection 7.3 satisfies this condition.

In Figure 4, the dashed lines illustrate that front-loading indeed occurs if we set the parameters of the probability of a negative evaluation to the values that generate the enhanced monitoring precision discussed in the Subsection 7.3. Search effort prior to the third meeting is higher than after the permanent sanction is imposed. Moreover, the figure clearly illustrates that more precision leads to raising search effort and reducing the reservation wage well before the actual meetings.

8   Conclusion

This paper sets up a structural job search model that formalizes key elements of monitoring job search intensity in insured unemployment (UI). We innovate relative to the existing literature by explicitly taking into account the fact that the measurement of search effort is imperfect and that the assessments of search effort take place at prescribed moments in time, so that the unemployed can anticipate them and gradually adjust their behavior nearer to these interviews, i.e. by allowing for non-stationary behavior.

This model is estimated with administrative data collected around the introduction of a scheme designed to monitor the search effort of young, long-term unemployed individuals in Belgium. The estimation results point to a very weak behavioral impact of the monitoring scheme. According to these results, once the scheme was fully operational, a population of 25- to 30-year-old entrants into

unemployment is expected to find a job on average 3 days earlier than in the absence of monitoring. This effect is much lower than the impact of the order of 3 to 6 weeks found in the U.S. and the Netherlands (Paserman, 2008; van den Berg and van der Klaauw, 2015). Essentially this is because (i) the timing of the monitoring is very much spread out over time, with the first assessment of search effort (out of three at most) does not occur earlier than the 22nd month of UI; (ii) the monitoring technology is not sufficiently precise; (iii) in contrast to the other studies, the evaluation is targeted at the youth, for whom active labor market policies are generally found to be less effective than for prime aged workers; and (iv) UI in Belgium is also more lenient than in many other countries in other respects, which makes UI caseloads more comparable to social assistance claimants and less responsive to incentives. For instance, in Belgium, based on their educational qualifications, individuals with little or no work experience are entitled to benefits and no time limit is imposed on eligibility. Our finding that as much as 20% of UI entrants are estimated to have such high search costs that they cannot be induced to search for a job supports this interpretation and underlines the fact that the effectiveness of a job search monitoring scheme depends on the overall institutional design of UI.

The advantage of estimating this structural model is that it allows us to study how modifying the design changes the impact of the monitoring scheme. We therefore simulated the effects of changing some features of the monitoring scheme on a range of outcomes. These simulations show that reinforcing the strictness of the monitoring, i.e. placing the meetings much earlier in the spell and introducing the threat of sanctions at the first meeting, already has sizable implications: the duration of benefit receipt is reduced by nearly a factor three, i.e. to 8 instead of 3 days. Since we found that the Belgian monitoring scheme was very blunt in that the probability of a negative assessment was insensitive to marginal changes in job search effort, we also investigated the implications of enhancing the precision of the monitoring technology. From this analysis, we have concluded that introducing more precision not only increases the effectiveness of monitoring in that it enhances job finding (e.g. the effect of the aforementioned reinforcement of strictness further reduces unemployment duration by a factor two, i.e. to 16 days), but also in that it generally increases savings in public expenditures more than it reduces the lifetime utility of the unemployed. The latter may, however, partly reflect our assumption that enhancing precision is costless. We argued this to be realistic for the Belgian scheme, but this may not be the case elsewhere. Anyhow, this suggests that the precision of the monitoring technology is key.[35] Currently, we know very little about this precision and we therefore invite scholars to include this in their research agenda.[36]

Finally, as any study, this research is not exempt from limitations. Let us list here a couple of these without claiming to be exhaustive. A first weakness of this study is that we were not able to distinguish between the job arrival rate and the job search effort. One way to solve this is to collect information on indicators of job search effort (such as indicators of the types of search channels used or the number of application letters sent), as van den Berg and van der Klaauw (2015) do. Second, since we lacked complete information on non-labor income, we had to assume risk-neutral agents. Introducing risk aversion into the analysis is crucial not only to study the impact of job search monitoring on the optimal trade-off between insurance and work incentives, but also to better understand how monitoring should be designed to minimize the distortions that the pursuit of these conflicting objectives generate. Third, Lalive, Landais and Zweimüller (2015) have shown that UI reforms can have market externalities. Thus, the impact of large monitoring schemes would be better studied in an equilibrium search model, such as the one proposed by Boone et al. (2007). All these exiting topics are left as avenues for further research.

---

[35]This was also the conclusion of Cockx et al. (2014), who studied the theoretical welfare implications of job search monitoring of the unemployed with time inconsistent preferences.

[36]Recently, Arni and Schiprowski (2015) found that setting job search requirements above the job seeker's unconstrained effort choice can enhance the job finding rate substantially. These authors study the Swiss system in which job search requirements are quite sharply defined. This supports our finding that a sufficiently accurate monitoring scheme can indeed be effective in raising job search effort. However, more research is required to obtain insights into the sensitivity of this effectiveness with respect to measurement precision and the implications on the design of job search monitoring.

Acknowledgements

References

Abbring, J. H., G. J. van den Berg, and J. C. van Ours (2005) 'The Effect of Unemployment Insurance Sanctions on the Transition Rate from Unemployment to Employment.' *The Economic Journal* 115(505), 602–630

Andrews, D. (1988a) 'Chi-square diagnostic tests for econometric models: Introduction and applications.' *Journal of Econometrics* 37(1), 135–156

―― (1988b) 'Chi-square diagnostic tests for econometric models: Theory.' *Econometrica* 56(6), 1419–1453

Arni, P., and A. Schiprowski (2015) 'The effects of Binding and Non-Binding Job Search Requirements.' IZA Discussion Paper 8951, Bonn: IZA

Arni, Patrick, Rafael Lalive, and Jan van Ours (2013) 'How Effective Are Unemployment Benefit Sanctions? Looking Beyong Unemployment Exit.' *Journal of Applied Econometrics* 28, 1153–1178

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

Bloemen, H. G. (2005) 'Job Search, Search Intensity, and Labor Market Transitions an Empirical Analysis.' *The Journal of Human Resources* 40(1), 231–269

Bolhaar, Jonneke, Nadine Ketel, and Bas van der Klaauw (2016) 'Job-Search Periods for Welfare Applicants: Evidence from a Randomized Experiment.' IZA Discussion Paper 9786, Bonn: IZA

Boone, J., A. Sadrieh, and J. van Ours (2009) 'Experiments on Unemployment Benefit Sanctions and Job Search Behavior.' *European Economic Review* 53(8), 937–951

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

Card, D., J. Kluve, and A. Weber (2015) 'What Works? A Meta-Analysis of Recent Active Labor Market Program Evaluations.' IZA Discussion Paper 9236, Bonn: IZA

Christensen, B., and N. Kiefer (1994) 'Measurement Error in the Prototypal Job Search Model.' *Journal of Labor Economics* 12(4), 618–639

Cockx, B., and M. Dejemeppe (2012) 'Monitoring Job Search Effort. An Evaluation Based on a Regression Discontinuity Design.' *Labour Economics* 19(5), 729–737

Cockx, B., C. Ghirelli, and B. Van der Linden (2014) 'Is it socially efficient to impose job search requirements on unemployed benefit claimants with hyperbolic preferences?' *Journal of Public Economics* 113, 80–95

Cockx, Bart, Muriel Dejemeppe, Andrey Launov, and Bruno Van der Linden (2011) 'Monitoring, sanctions and front-loading of job-search in a non-stationary model.' IZA Discussion Paper 6181, Bonn: IZA

DellaVigna, S., and D. Paserman (2005) 'Job Search and Impatience.' *Journal of Labor Economics* 23(3), 527–588

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

Eckstein, Z., and G. J. van den Berg (2007) 'Empirical Labor Search: A Survey.' *Journal of Econometrics* 136(2), 531–564

Eriksson, S., and J. Lagerström (2006) 'Competition between Employed and Unemployed Job Applicants; Swedish Evidence.' *Scandinavian Journal of Economics* 108(3), 373–396

Eriksson, S., and N. Gottfries (2005) 'Ranking of job applicants, on-the-job search and persistent unemployment.' *Labour Economics* 12, 407–428

Flinn, C., and J. Heckman (1982) 'New Methods for Analyzing Structural Models of Labor Market Dynamics.' *Journal of Econometrics* 18(1), 115–168

Fougère, D., J. Pradel, and M. Roger (2009) 'Does the Public Employment Service Affect Search Effort and Outcomes?' *European Economic Review* 53(7), 846–869

Frijters, P., and B. van der Klaauw (2006) 'Job search with nonparticipation.' *The Economic Journal* 116(508), 45–83

Garcia-Perez, J. I. (2006) 'Job separation in a non-stationary search model: A structural estimation to evaluate alternative unemployment insurance systems.' *Journal of Applied Econometrics* 21(2), 245–272

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(3), 590–610

Keane, M. P., P. E. Todd, and K. I. Wolpin (2011) 'The Structural Estimation of Behavioral Models: Discrete Choice Dynamic Programming Methods and Applications.' In *Handbook of Labor Economics,* ed. Orley Ashenfelter and David Card, O. Ashenfelter and D. Card ed., vol. IV.A (Elsevier) pp. 331–461

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

Klepinger, D. H., T. R. Johnson, J. M. Joesch, and J. M. Benus (1997) *Evaluation of the Maryland Unemployment Insurance Work Search Demonstration* (Seattle and Bethesda: Batelle Memorial Instiute and Abt Associates Inc.)

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

Lalive, Rafael, Camille Landais, and Josef Zweimüller (2015) 'Market externalities of large unemployment insurance extension programs.' *American Economic Review* 105(12), 3564–3596

Launov, A., and K. Wälde (2013) 'Estimating incentive and welfare effects of non-stationary unemployment benefits.' *International Economic Review* 54(4), 1159–1198

Ljungqvist, L., and T. Sargent (1995) 'The Swedish Unemployment Experience.' *European Economic Review* 39(5), 1043–1070

Lollivier, S., and L. Rioux (2010) 'An Empirical Examination of the Sources of Changes over Time in the Job Finding Rate Using Reservation Wages and Rejected Wage Offers.' *International Economic Review* 51(4), 1039–1069

Longhi, S. (2015) 'Do the Unemployed Accept Jobs Too Quickly? A Comparison with Employed Job Seekers.' IZA Discussion Paper No. 9112, Bonn: IZA

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

McVicar, D. (2008) 'Job Search Monitoring Intensity, Unemployment Exit and Job Entry: Quasi-experimental Evidence from the UK.' *Labour Economics* 15(6), 1451–1468

Meyer, B. D. (1995) 'Lessons from the U.S. Unemployment Insurance Experiments.' *Journal of Economic Literature* 33(1), 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

Mroz, T. A. (1987) 'The Sensitivity of an Empirical Model of Married Women's Hours of Work to Economic and Statistical Assumptions.' *Econometrica* 55(4), 765–799

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

Paserman, D. (2008) 'Job Search and Hyperbolic Discounting: Structural Estimation and Policy Evaluation.' *The Economic Journal* 118(531), 1418–1452

Pavoni, N., and G. Violante (2007) 'Optimal Welfare-to-Work Programs.' *Review of Economic Studies* 74(1), 283–318

Petrongolo, B. (2009) 'The Long-Term Effects of Job Search Requirements: Evidence from the UK JSA Reform.' *Journal of Public Economics* 93(11-12), 1234–1253

Ridder, G. (1984) 'The distribution of single-spell duration data.' In *Studies in Labor Market Dynamics,* ed. G.R. Neumann and N.C. Westergård (Darmstadt: Springer-Verlag) chapter 3, pp. 45–73

Ridder, G., and G. van den Berg (2003) 'Measuring Labor Market Frictions: A Cross-Country Comparison.' *Journal of the European Economic Association* 1(1), 224–244

Setty, Ofer (2015) 'Optimal Unemployment Insurance with Monitoring.' Working Paper, Tel Aviv University

Tanner, M., and W. Wong (1983) 'The estimation of the hazard function from randomly censored data by the kernel method.' *Annals of Statistics* 11(3), 989–993

_____ (1984) 'Data-based nonparametric estimation of the hazard function with applications to model diagnostics and exploratory analysis.' *Journal of the American Statistical Association* 79(385), 174–182

Todd, P. E., and K. I. Wolpin (2006) 'Assessing the Impact of a School Subsidy Program in Mexico: Using a Social Experiment to Validate a Dynamic Behavioral Model of Child Schooling and Fertility.' *American Economic Review* 96(5), 1384–1417

van den Berg, G. J. (1990) 'Nonstationarity in Job Search Theory.' *Review of Economic Studies* 57(2), 255–277

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

_____ (2015) 'Structural Empirical Evaluation of Job Search Monitoring.' IFAU Discussion Paper No. 2015:16

van den Berg, G. J., H. Förster, B. Hoffman, and A. Uhlendorff (2015) 'Job search, sanctions, and sickness absence.' *mimeo*, University of Mannheim

van der Klaauw, B., and A. van Vuuren (2010) 'Job Search and Academic Achievement.' *European Economic Review* 54(2), 298–320

Wolpin, K. I. (1987) 'Estimating a Structural Job Search Model: The Transition from School to Work.' *Econometrica* 55(4), 801–818

Wunsch, C. (2013) 'Optimal Use of Labor Market Policies: The Role of Job Search Assistance.' *Review of Economics and Statistics* 95(3), 1030–1045