Why Are Employers Put Off by Long Spells of Unemployment?

Abstract

Recent evidence from large-scale field experiments has shown that employers use job candidates' unemployment duration as a sorting criterion. In the present study, we investigate what underlies this pattern. To this end, we conduct a survey experiment in which employers make hiring decisions concerning fictitious job candidates who have experienced spells of unemployment of different length. In addition, candidates are rated on several statements that are central to four signals often associated with unemployment: (i) a signal of trainability, (ii) a signal of other fixed skills, (iii) a signal of skill loss, and (iv) a signal of negative evaluation by other employers. We use these ratings to estimate a multiple mediation model, in which the effect of the duration of unemployment on hiring intentions is mediated by the four signals. Our findings indicate that longer unemployment spells are mainly perceived by employers as a signal of lower motivation and, as a result, the long-term unemployed have lower chances to be hired or even be invited to a job interview. Understanding the reasons why employers are reluctant to hire the long-term unemployed is crucial to devise proper activation measures to facilitate their re-employment. Our study is a contribution in this direction.

Keywords: unemployment scarring; signalling theory; employers; factorial surveys; vignettes.

1. Introduction

In social stratification research, the experience of unemployment has been described as a *trigger event* (DiPrete, 2002; Gangl, 2004; 2006), that is, a critical, stressful and potentially disruptive life course event often taking a severe economic and psychological toll on those affected (for a review: Brand, 2015). With the economic downturn of recent years, the number of people going through a spell of unemployment as well as the average length of unemployment spells have been on the rise (OECD, 2013), drawing renewed attention to the potential scarring effect of unemployment on future re-employment chances. As employers are particularly wary of lengthy gaps in the résumé that are unaccounted for (Bills, 1990), unemployment tends to be self-reinforcing, possibly stigmatising the long-term unemployed (hereafter: LTU) in employers' perceptions. Indeed, a number of studies in both sociology and economics have pointed to the negative duration dependence of unemployment—the observation that an individual's probability of exiting unemployment decreases the longer she/he is unemployed (e.g. Cockx and Picchio, 2013; Luijkx and Wolbers, 2009; Mooi-Reci and Ganzeboom, 2015).

Recently, large-scale field experiments conducted in Sweden and the United States have shown that at least part of the negative duration dependence of unemployment has a demand-side explanation: employers are reluctant to hire LTU (Eriksson and Rooth, 2014; Kroft et al., 2013). In these résumé-based audit studies, fictitious job applicants with a longer unemployment spell received significantly fewer job interview invitations than identical applicants with a shorter spell. However, while field experiments of this kind are convincing for the clean measurement of unemployment scarring, they do not allow disentangling the

reasons for this pattern: long-term unemployment is shown to be used as a negative signal by employers, but it remains unclear what exactly is signalled by longer unemployment spells.

In this study, we explore the empirical importance of four perceptions potentially underlying employers' reluctance to hire LTU, namely, the perception that LTU: (i) possess skills or characteristics that are not directly observed but considered less than optimal for the job, (ii) have experienced a deterioration of skills during the unemployment spell, (iii) are less trainable than candidates without long unemployment spells, and (iv) have been negatively evaluated by other employers and therefore deemed undesirable employees. To this end, we propose a state-of-the-art vignette experiment conducted in Flanders, Belgium,[i] in which professionals involved in real-life hiring processes reveal their hiring intentions with respect to job candidates with different unemployment durations. In addition, the survey module in which the vignette experiment is embedded provides us with rich information about the reasons underlying employers' preferences. This allows us to examine the empirical power of the four signals by estimating a multiple mediation model. Thereby, our study complements (and is consistent with) the evidence obtained from employer surveys (Atkinson et al., 1996; Bonoli, 2014) which, however, are more likely to be biased by socially desirable response patterns. In comparison, vignettes are a powerful method to analyse socially sensitive questions (Auspurg et al., 2014) and the possibility they afford to present employers with detailed scenarios is an important methodological advantage as employers are more likely to report negative views of specific unemployed applicants than when questioned in very general terms (Bonoli, 2014).

This study contributes to the literature on unemployment scarring by looking more closely at the demand-side mechanisms that can trap unemployed job seekers in long-lasting periods

of joblessness. Our findings show that employers' reluctance to hire LTU is to a large extent mediated by their perception of unemployment as signalling lower motivation. A smaller fraction of the total effect of unemployment duration on hiring intentions is associated with rational herding, that is, the belief that other employers found the candidate's productivity to be low. Understanding why employers refrain from hiring LTU is crucial to design activation policies that are effective in re-inserting them into the labor market. Our study is a contribution in this direction. For example, if the unemployed (and caseworkers) are made aware of the (mis)perceptions standing in the way of their employment opportunities, they may attempt to compensate for these perceptions, for instance, by underlining relevant personal characteristics and attainments in their résumé.

The remainder of this article is structured as follows. Section 2 gives a brief overview of the four theoretical explanations for employers' reluctance to hire LTU, and the associated signals, as found in the multidisciplinary literature on this topic. Section 3 describes the experiment we conducted. The experimental data is then analysed in Section 4. Section 5 concludes with some take-away messages for scholars as well as for interested policy makers. In addition, in this last section, we discuss the limitations of our experimental design.

2. Theoretical Framework

Theories explaining the phenomenon of negative duration dependence of unemployment are abundant in both the fields of sociology of work and occupations and labour economics. The observed reluctance to hire LTU can have many possible sources, both on the demand- and

supply-side of the labour market. While the demand-side explanations reviewed in this study influence the unemployment duration through the perceptions of employers, supply-side explanations attribute to the negative duration dependence by actual changes in the behaviour or productivity of workers over the course of the unemployment spell.[ii] However, in our vignette experiment, explanations for the negative duration dependence of unemployment that are situated on the supply-side are ruled out by design.

Under the umbrella of signalling theory, we can bracket various models in the social and behavioural sciences, arguing that when people are confronted with asymmetric information, they use the limited available information as a signal for other, unobserved factors related to one's productivity (Arrow, 1973; Eriksson and Rooth, 2014; Kroft et al., 2013; Spence, 1973; Vishwanath, 1989). Accordingly, employers could rely on candidates' employment history as a screening device to filter out job candidates. What remains unclear however, is what exactly is signalled by a long unemployment spell. In this study, we focus on four signals that are related in the literature to long-term unemployment: (i) a signal of (lower) fixed skills and characteristics, (ii) a signal of skill loss, (iii) a signal of (lower) trainability and (iv) a signal of rejection by other employers.

In the most direct interpretation of signalling theory, employers could see a long unemployment duration as a signal of unobserved skills or characteristics that are innate or fixed over time. In this sense, a long unemployment spell can be a signal of lower motivation (Luijkx and Wolbers, 2009) or lower intellectual and social capabilities (Vishwanath, 1989), both of which are negatively associated with productivity. As these characteristics are unobserved by employers at point of hire, unemployment spells may be used as proxies instead.

On the other hand, employers could also believe that a worker's productivity is dynamic and deteriorates over the course of an unemployment spell. Put differently, employers could believe in skill loss or skill depreciation. This mechanism is related to human capital theory, as first described by Becker (1962; 1994). Crucial is that it is costly for the unemployed to maintain their skill level during the stretch of unemployment (Acemoglu, 1995; Mincer and Ofek, 1982). Moreover, employers cannot detect the genuine level of skill depreciation of a (long-term) unemployed applicant. As shown by Acemoglu (1995), these two observations may result in an inefficient equilibrium in which employers discriminate against LTU due to the perceived skill loss (and, as a result, the unemployed do not invest to maintain their skill level).

Two more specific applications of signalling theory are also widely cited in this context. A first particular application relates long-term unemployment to (a signal of) lower trainability. Following queuing theory (Thurow, 1975), employers may rank all job candidates by their (perceived) trainability, with the person they believe will be easiest to train holding the first position in the queue and the person they perceive as the least trainable holding the last. Subsequently, these employers decide on a cut-off and only the individuals above the cut-off are invited for a job interview. Because employers, again, do not possess full information, they have to use the limited information available to assess a job applicant's trainability (Di Stasio, 2014). If employers believe unemployment has a negative effect on trainability, people with a longer unemployment spell will be ranked lower in the labour queue and, as a consequence, have a lower chance of getting invited for a job interview.

The final application of signalling theory we consider stipulates that, when making the

decision to invite someone for a job interview, employers follow the behaviour of other employers —a behaviour also known as rational herding (Banerjee, 1992; Bonoli and Hinrichs, 2012; Oberholzer-Gee, 2008). One such factor from which employers might infer the screening behaviour of their colleagues is job candidates' unemployment durations. Qualitative studies have indicated that employers assume the time out of work is spent looking for a job, but, since the candidate is still unemployed, this search must have been unsuccessful (Bonoli, 2014). If the unemployment spell is relatively long, employers might conclude that other employers have repeatedly found the candidate's productivity to be low and decide that it is unprofitable to hire the candidate.

In what follows, we will explore how these key perceptions mediate the effect of unemployment duration on hiring intentions. We should note two things. First, apart from a person's unemployment history, these signals could vary with a number of different factors, including gender, work experience, social participation and education level. We will take this into account when designing the experiment. Second, we do not intend to demonstrate that, for example, LTU actually lose specific skills or become less motivated while out of work (i.e. to test supply-side explanations), but only that employers believe they do. In other words, when looking at unemployment scarring from a demand-side perspective, employers' perceptions are both crucial and sufficient for scarring effects to materialise.

Correspondence tests have provided evidence for negative signalling effects related to long-term unemployment. In this kind of experiment, sets of fictitious résumés, differing only in the characteristic of interest that is randomly assigned, are sent to real job openings. By measuring the subsequent invitations received from employers (i.e. callbacks) unequal treatment can be

identified in a causal manner (Baert, 2018b). Using this methodology, it has been shown that a wide range of factors constitute a signal in the hiring process, including ethnicity (Baert et al., 2015; Kaas and Manger, 2012; Oreopoulos, 2011), gender (Baert et al., 2016; Petit, 2007; Riach and Rich; 2006), and age (Ahmed et al. 2012; Baert et al., 2016; Lahey, 2008). Studies using this methodology have also looked at the signal of long unemployment durations. While Farber et al. (2016) found no significant scarring effect of long unemployment spells on callbacks, the majority of studies reported, indeed, lower callback probabilities for LTU (Eriksson and Rooth, 2014; Kroft et al., 2013; Oberholzer-Gee, 2008).

Having established that a long unemployment spell is a negative signal towards employers, the question remains what is signalled by this long unemployment spell. This has been the topic of a number of qualitative studies. Atkinson et al. (1996) administered a telephone survey with 800 representative employers in the United Kingdom. They concluded that employers believe LTU do possess the necessary skills but they are nevertheless less attractive due to a recent deterioration in these skills—pointing towards a negative signal of skill loss—and, most importantly, a lower motivation. A perceived lower motivation was also the main reason why 722 Swiss employers surveyed by Bonoli (2014) were reluctant to hire LTU. Bonoli and Hinrichs (2012) reached similar conclusions based on 41 semi-structured interviews with employers in six European countries. In addition, they found evidence for rational herding, i.e. the employers stated that LTU must have been deemed unproductive by previous employers. Lastly, Oberholzer-Gee (2008) carried out 766 telephone surveys with Swiss employers and found evidence for a signal of skill loss and a signal of negative evaluation by other employers. To the best of our knowledge, we are the first to approach this question using experimental methods (and to tease out the signals' relative importance).

3. Experiment

In order to not only determine whether job candidates' unemployment duration affects their hiring chances, but also gain an insight into the thought process leading to this pattern, we conducted a vignette study. Vignette studies are based on the factorial survey method (Auspurg and Hinz, 2014; Rossi and Nock, 1982) and are commonly used to study human judgements (Jasso, 2006; Wallander, 2009). In recent years, this method has been increasingly used to study dynamics in hiring decisions (e.g. Di Stasio, 2014; Liechti et al., 2017).

Each participant in a vignette experiment is asked to judge several short hypothetical descriptions of situations or individuals described on vignettes, whose characteristics (*factors*) vary randomly or systematically over a defined number of categories (*levels*). As a consequence, correlations between the vignette factors are minimised to a value close to 0. This orthogonal design allows a causal interpretation of the effects of the vignette factors on participants' judgements. When employed to study hiring intentions, vignettes typically list various characteristics of fictitious job applicants who are evaluated by the participants of the experiment. The simultaneous manipulation of different applicant characteristics closely resembles the multidimensional nature of selection decisions in the field, as in practice employers also compare candidates who vary on a number of characteristics, such as gender, level of education, and employment history.

3.1. Vignette Design

We asked a sample of professionals familiar with real-life hiring processes (referred to as

employers from here on) to evaluate a set of five vignettes describing each a fictitious job applicant. The job applicants varied in five factors, presented in Table 1.[iii] The vignette factor of main interest for our study is the unemployment duration, operationalised as the number of months a candidate reported to have been unemployed prior to the job application. In line with Kroft et al. (2013), this number could take on any integer from 1 to 36 (resulting in 36 vignette levels for this factor). By means of this flexible approach, we did not have to make any prior judgement on the time-pattern of unemployment scarring. As can be seen from Table 1, the fictitious candidates also differed in gender (male or female), highest degree obtained (secondary education or bachelor's degree), work experience (two or five years), and participation in social activities (none or volunteering activities). These factors were chosen on the basis of our literature review and tested over the course of explorative interviews with three HR professionals. We also ran a pilot study with 30 master's students in economics to assess whether our vignettes were perceived as credible, which reassured us that no crucial information was omitted. We should make two important notes. Firstly, our choice to include a continuous unemployment duration, resulting in one vignette factor with 36-levels (as opposed to two levels for the other factors), can cause a 'number of levels' effect (Auspurg and Hinz, 2014). However, as the aim of our study is not to compare the relative importance of different vignette factors, we do not think this is a major issue. Moreover, including these 36 levels in our models allows us to exploit a larger variance in this variable and avoids a choice for arbitrary vignette levels capturing short- and long-term unemployment. Secondly, it could be the case that some combinations of vignette factors are implausible. Indeed, even though long-term unemployment is high in Belgium (see endnote i) one could imagine that employers are unlikely to have been confronted with, for instance, candidates with a bachelor degree

and/or five years of experience who have been unemployed for the full 36 months. Therefore, we will report on a robustness check in which implausible vignettes were excluded.

< Table 1 >

After fully crossing all the vignette levels for the five factors, we obtained a vignette universe of 576 (i.e. 36 x 2 x 2 x 2 x 2) vignettes. We sampled 300 vignettes out of this universe using a D-efficient randomisation following the Kuhfeld (2010) algorithm as explained in Auspurg and Hinz (2014). This resulted in a very high D-efficiency of 99.820. In a second step, we grouped these vignettes (again following Kuhfeld (2010)) to create 60 decks with five vignettes each. These decks were distributed at random to the participants. It is important to note that one of these decks was not effectively evaluated, while the other (59) decks were evaluated at least once. This could result in a low efficiency of the post-survey sample. The ensuing post-survey correlations among the vignette factors are shown in Table A1 (in the Supplementary Material). While this is no test of post-survey efficiency, it is nevertheless comforting that all of these correlations are sufficiently small and not statistically different from 0.

3.2. Data Collection

Our vignette experiment was integrated into a large-scale web-based survey sent to individuals living in Flanders, in January 2017. More concretely, the survey was sent to 89,847 individuals who selected themselves into a database of people interested in participating in research on human resource management (in response to calls via e-mail and social media). In the first question, each individual was asked whether she/he had been involved in evaluating job

candidates for a minimum of five vacancies over the last year. In order to closely mimic real-life hiring decisions, we wanted to conduct our experiment exclusively with professionals familiar with the hiring process. Therefore, the answer to this first question determined whether a person was eligible to take part in our experiment. If this first question was answered positively, she/he was assigned with a chance of 0.50 to our experiment (and with a similar chance to another one). Otherwise, she/he was referred to a regular, policy-oriented survey on burnout. A total of 10,488 individuals answered this first question, giving us an overall response rate of about 12%. Out of these respondents, 475 indicated being actively involved in the hiring process a minimum of five times over the last year, of which 242 were assigned to our experiment. Twenty-three among them left one or more questions unanswered, leaving us with a final sample of 219 participants with complete responses. These 219 participants were comparable to the initial 242 participants in terms of the participant characteristics that are discussed below and reported in Table A2 in the supplementary material.[iv] As each participant rated five vignettes, the number of (participant x vignette) observations is 1,095.

At the beginning of the web-based survey, participants were introduced to their role as employer at a fictitious company selling building materials. This company was in search of a counter assistant, which corresponds to ISCO-08 category 4200 (customer services clerks). We selected this occupation because it is transversal to a number of industries, thus increasing the chance that respondents would be familiar with it (we discuss the research limitations related to this choice in Section 5). Participants were explicitly informed that this counter assistant should be (i) customer-oriented, (ii) service-minded and (iii) commercially oriented. The assistant was also expected to be efficient and reliable in managing administrative tasks. These instructions were presented to all participants in the same way at the beginning of the survey.

Subsequently, participants were shown the vignettes describing five fictitious candidates. It was stressed that these candidates were formally qualified for the job. Information about the candidates was presented in a tabulated way. We chose this format because 'tabular vignettes might be better suited to decision tasks (i.e. resumes or many consumer product descriptions), which frequently involve lists of decision criteria[, compared to text vignettes]' (Auspurg and Hinz 2014: p. 70). Participants were not informed about the goal of the experiment.

After this, participants were asked to indicate, for each vignette, their intention to hire the candidate by rating the statements 'The probability that I will invite this candidate for a job interview is high' and 'The probability that I will hire this candidate for the position is high' on a 7-point Likert scale (with 1 'completely disagree' and 7 'completely agree'). We will refer to these items as the 'interview scale' and the 'hiring scale', respectively, and consider both outcomes separately.

In view of investigating the signals associated with the unemployment duration, participants were additionally prompted to rate eight statements for each candidate, linked to the four signals described in Section 2, on a 7-point Likert scale. These statements are reported, signal by signal, in Table 2.[v]

< Table 2 >

To make sure that our selection of signals was exhaustive, we complemented our literature review with three exploratory interviews with HR professionals (as described in Section 3.1). Here we asked whether they would hire a person with a long unemployment spell and, if not, which reasons they voiced for this decision.[vi] Independently, all HR professionals linked long-term unemployment to lower motivation and/or fewer hard or soft skills. Related to skill loss,

the fact that the workplace goes through quick technological changes over the course of an unemployment spell was also cited multiple times. Next, we discussed the four signals we selected and whether any of these perceptions had ever driven their hiring decisions in practice. The HR professionals evaluated all four signals as relevant.

Firstly, we included three statements to test for the possibility that long-term unemployment may signal (a lower level of) fixed skills and characteristics. Participants were asked whether they thought the candidate was sufficiently motivated (statement 1) and had a high enough level of intellectual ability (statement 2) and social ability (statement 3) for the job. Secondly, three statements tested for perceived skill loss of the candidate. Inspired by the interviews with HR professionals, the candidate was scored with respect to being up to date with technologies (statement 4). In addition, perceived deterioration in general skills (statement 5) and social skills (statement 6) were scored. Thirdly, closely linked to queuing theory, participants were asked to rate the candidate's trainability (statement 7). Fourthly, participants judged whether the candidate had been rejected often by other employers (statement 8), which is the explanation for the negative duration dependence of unemployment put forward by rational herding.[vii]

A definition of all variables collected by means of this vignette experiment and used in our analyses is given in Table A2 of the Supplementary Material. An English translation of the experimental instructions and an example of a vignette (and the related items) can be found in Section B of this Supplementary Material.

In the mediation model presented in Section 4, we include four mediators, one for each signal, based on the eight statements reported in Table 2. The first mediator, the fixed skills

scale, groups statements 1 to 3 (Cronbach's alpha for internal consistency: $\alpha = 0.763$). Its value is, for each observation, computed as the average over these three statements. The second mediator, the skill loss scale, is based on the scores of statements 4 to 6 ($\alpha = 0.716$). The scores of statement 4 were reverse-coded (so that a higher score became consistent with higher perceived skill loss). The third mediator, the trainability scale, reflects the score of statement 7. The fourth and final mediator, the rational herding scale, corresponds to the score of statement 8.

Our choice to group statements together as we did is, to some extent, arbitrary. Therefore, we tested the sensibility of our results with respect to other strategies. For instance, an approach in which the scores of the statements were first standardised (by subtracting their sample mean and dividing the result by these scores' sample standard deviation) before grouping them did not substantially affect the results presented in Section 4. In addition, factor analysis yielded the same number (i.e. four) of scales, with a comparable composition. Note that we also present the mediating role of the eight separate statements (i.e. without grouping them) in an alternative mediation model.

After judging the five job candidates, participants were asked to provide some personal information, including their gender, level of education, frequency of taking hiring decisions and experience with the hiring process (Table A2, Supplementary Material). Overall, about 57% of our participants were female. They were mainly highly educated (almost 90% had completed some form of tertiary education), with an average age of about 42 and an average of around 10 years of experience as an HR professional. Table A3 (in the Supplementary Material) reports the distribution of our participants according to the unemployment duration

of the candidates they judged to check whether our randomisation was successful. For instance, as shown in Panel A, the subsample of vignettes disclosing 3 months of unemployment or fewer and the subsample of vignettes disclosing more than 3 months of unemployment were scored by participants with comparable characteristics.

It should be noted that our sample is not representative of the population of Belgian employers, for which a sampling frame is unfortunately not readily available. We do not consider this a substantial shortcoming. Samples gathered by field experiments are similarly non-representative (they only target employers who post their job ads online in specific job banks) but still widely employed to causally test the scarring effects of unemployment. Moreover, our sample is very comparable in age and gender distribution with Belgian HR professionals in the European Social Survey, even though our sample seems slightly higher educated—the formal comparison is included as Table A4 in the Supplementary Material. We come back to this and other issues related to our experimental design in the conclusion.

4. Results

We estimate a multiple mediation model (Hayes, 2013) to analyse the total effect of unemployment duration on hiring intentions as well as the part of this effect passing through the four mediators. A simplified version of the estimated model is depicted in Figure 1.

< Figure 1 >

In a first step (Section 4.1), we estimate the total effect of the unemployment duration of our fictitious job candidates on the employers' hiring intentions. Subsequently, we explore the

mediation effects related to the fixed skills, skill loss, trainability, and rational herding scales. Each mediation effect is calculated as the product of the effect of unemployment duration on the respective mediation scale and the association of this scale on the outcome scale (i.e. $\delta_i\theta_i$, with i ranging from 1 to 4, in Figure 1). In Section 4.2 we explore the mediation effects separately and in Section 4.3 we estimate the complete mediation model, in which the mediation scales are included jointly. The latter model allows us to decompose the total effect of unemployment duration into four 'indirect' effects via the mediators and a remaining 'direct' effect δ' (so that the total effect δ equates $\delta' + \sum_{i=1}^{4} \delta_i\theta_i$).

We stress that we follow the literature when labelling $\delta_i\theta_i$ as mediation *effects* but refrain from giving them a causal interpretation. The unemployment duration of our fictitious job candidates is experimentally manipulated and, as a consequence, δ and δ_i are causal effects. However, our mediators are not exogenous. Although we attempt to capture, based on our literature review, the most relevant signals potentially explaining the lower hiring chances of LTU, it is still possible that our mediators correlate with other, unobserved, employer perceptions related to candidates' unemployment. For this reason, θ_i should be seen as associations rather than as causal effects. We return to this point in Section 5.

4.1. Bivariate Analysis

To get a first impression of the (total) effect of the candidates' unemployment duration on their hiring intentions, we plot the average scores on the interview scale of the 1,095 evaluated vignettes, by unemployment duration. As is clear from Figure 2, the likelihood of getting invited for an interview exhibits a clear downward trend as the unemployment duration

increases. A similar pattern emerges for the hiring scale.

< Figure 2 >

However, due to the relatively low number of observations for each potential unemployment duration (between 23 and 40 observations), Figure 2 captures some noise. A clearer picture of the total effect is presented in Table 3, where we compare the outcome scales for candidates with an unemployment spell of 3 months or fewer to the outcome scales for candidates with an unemployment spell of more than 3 months (Panel A), and repeat this with 12 months (Panel B) and 24 months (Panel C) as cut-off points. A t-test is used to determine whether the difference in invitation and hiring probability between these subsamples are significantly different from 0.[viii]

< Table 3 >

As shown in Table 3, the probability of getting invited for a job interview is always significantly higher for candidates belonging to a subsample with a shorter unemployment spell compared to candidates belonging to a subsample with a longer unemployment spell, regardless of the chosen cut-off. For instance, the average score on the interview scale for those with an unemployment duration of 3 months or fewer is 5.515 (i.e. just between an evaluation of 'somewhat agree' and 'agree' with respect to the statement 'The probability that I will invite this candidate for a job interview is high') while it is 4.050 (i.e. close to 'neither agree or disagree') for those with an unemployment duration of more than 3 months. A similar pattern is found for the probability that a candidate is hired for the position.

Due to the orthogonal design, candidates with a longer unemployment spell are (on average) equal to candidates with a shorter unemployment spell on all vignette factors, other than their

unemployment duration. In other words, the measured differences in interview invitations presented in Table 3 can only be driven by differences in unemployment duration. A regression-based approach yields exactly the same conclusion: a clear scarring effect of long-term unemployment.

4.2. Exploration of the Mediation Effects

A significant role for the mediation scales in explaining the negative relationship between unemployment duration and hiring intentions is conditional on two things. Firstly, candidates' unemployment duration should affect the mediation scales (left part of Figure 1). Secondly, these mediation scales should affect participants' hiring intentions (right part of Figure 1). In this subsection, we explore both conditions separately.

To get a first idea of the effect of unemployment duration on the four mediation scales, we examine the candidates' scores for these scales by their unemployment duration. In addition to the scores at the aggregate level, we present the scores on the level of the individual statements.

< Table 4 >

As Table 4 shows, the unemployment duration has a significant effect on all four mediators. Candidates with a longer unemployment spell score significantly lower on the 'positive' mediators (fixed skills and trainability), while they score significantly higher on the 'negative' mediators (skill loss and rational herding). When we look at the individual statements, it is apparent that the subsample means differ highly significantly for all statements and in the expected direction.

We have also checked that a positive evaluation with respect to the mediation scales is correlated with higher hiring intentions. To this end, we calculated correlations between the mediation scales (and their underlying statements) and the interview and hiring scales. A correlation matrix is presented in Table A5 (in the Supplementary Material): all correlations are significantly different from 0 and have the expected sign.

4.3. Multiple Mediation Regression Model

In the multiple mediation regression model all four mediators are included jointly, following a system of linear regression equations (by analogy with Hayes, 2013):

$$M_1 = \alpha_{M_1} + \beta_{M_1} CC + \gamma_{M_1} PC + \delta_1 UD + \varepsilon_{M_1}; \quad (1)$$

$$M_2 = \alpha_{M_2} + \beta_{M_2} CC + \gamma_{M_2} PC + \delta_2 UD + \varepsilon_{M_2}; \quad (2)$$

$$M_3 = \alpha_{M_3} + \beta_{M_3} CC + \gamma_{M_3} PC + \delta_3 UD + \varepsilon_{M_3}; \quad (3)$$

$$M_4 = \alpha_{M_4} + \beta_{M_4} CC + \gamma_{M_4} PC + \delta_4 UD + \varepsilon_{M_4}; \quad (4)$$

$$Y = \alpha_Y + \beta_Y CC + \gamma_Y PC + \delta' UD + \theta_1 M_1 + \theta_2 M_2 + \theta_3 M_3 + \theta_4 M_4 + \varepsilon_Y. \quad (5)$$

M_1, M_2, M_3, and M_4 are fixed skills, skill loss, trainability, and rational herding mediation scales, respectively; UD is the candidate's unemployment duration; CC is a vector of other vignette factors; PC is a vector of participant characteristics; and Y is the interview or hiring scale. β_{M_i}, γ_{M_i}, and δ_i are the (vectors of) parameters associated with CC, PC, and UD in the equations with M_i as dependent variable, with α_{M_i} being the intercept. β_Y, γ_Y, δ', and α_Y are the corresponding parameters in the equation with Y as dependent variable. Finally, θ_1, θ_2, θ_3,

and θ_4 are the parameters associated with the mediator scales in the latter equation. As a consequence, δ' is the remaining direct effect of the unemployment duration after controlling for the mediators. Our main interest lies in the products $\delta_i\theta_i$, namely the indirect effects of the unemployment duration on Y through each mediator M_i. In line with Hayes (2013), we estimate equations (1) to (5) simultaneously and correct the standard errors ε_{M_1}, ε_{M_2}, ε_{M_3}, ε_{M_4}, and ε_Y for clustering of the observations at the participant level.

In order to capture hiring intentions, we look at two outcomes: the interview and the hiring scale. The main results of our mediation analysis with the interview scale (hiring scale) as the Y-variable are depicted in Figure 1 (Figure A1 in the Supplementary Material). The corresponding full estimation results are reported in Table 5 and Table A6.

The total effect of unemployment duration on the interview scale (δ = -0.062; $p \leq 0.001$) is in line with what was reported in Section 4.1. One additional month of unemployment decreases the interview scale by 0.062 (i.e. about one sixteenth of a unit decrease on this scale ranging from 1 to 7). This total effect can be broken down into one direct effect and four indirect effects (one for each mediator). The direct effect, which can be interpreted as the part of the total effect that does not pass through any of the four mediators, is substantial (δ' = -0.026; $p \leq 0.001$). It accounts for 41.9% (i.e. 0.026 divided by 0.062) of the total effect, while all mediation effects together account for the remaining 58.1%—we will come back to this in Section 5.

Next, we investigate the relative importance of the four mediators. On the one hand, unemployment duration significantly affects all four mediation scales in the expected direction. On the other hand, only three of the mediation scales—the fixed skills scale (θ_1 = 0.851; $p \leq$

0.001), the trainability scale (θ_3 = 0.106; p = 0.039), and the rational herding scale (θ_4 = -0.117; p = 0.003)—appear to significantly influence the interview probability. Multiplying the first set of coefficients by the second set yields the mediation effects. In line with Hayes (2013), the confidence intervals for these mediation effects are based on 10,000 bootstrap samples. We find three significant mediation effects. Firstly, the effect of the unemployment duration on the interview outcome is highly significantly mediated by the fixed skills scale ($\delta_1\theta_1$ = -0.025, i.e. the product of -0.029 and 0.851; $p \leq 0.001$). This mediation effect accounts for 38.7% of the total effect. In addition, we find a smaller—but still highly significant—mediation via rational herding ($\delta_4\theta_4$ = -0.005; p = 0.005) and a small mediation via perceived trainability ($\delta_3\theta_3$ = -0.004; p = 0.049). No significant mediation via perceived skill loss is found. In other words, employers seem to believe that unemployment duration correlates with fixed (unobservable) employee characteristics rather than that the unemployment spell causes skills to deteriorate.

< Table 5 >

The total, direct, and indirect effects of unemployment duration on the hiring scale are similar to what is found with respect to the interview scale. Other secondary results, pertaining to the role of employers' characteristics, are reported in Panel B and Panel C of both Table 5 and Table A6 in the Supplementary Material. We do not discuss them any further as they fall outside the scope of this paper.

As stated in Section 3.1 we perform a robustness analysis where we exclude candidates with a bachelor degree and/or five years of experience in combination with an unemployment duration of two years or more, as these combinations of vignette levels could be perceived as implausible. The results of this analysis (in which 108 of the 300 sampled vignettes are

excluded) are reported in Figure A2 of the Supplementary Material. [ix] It is clear that our results are robust to the exclusion of these potentially implausible vignettes.

To get a picture of the relative weights of the individual statements, we re-estimate our mediation model using eight separate mediators instead of the four mediation scales. Estimation results are given in Table A8 and Table A9 (Supplementary Material). These results indicate that the dominant mediation through the fixed skills scale is mainly driven by a long unemployment spell being viewed as a signal of lower motivation. Moreover, there is some evidence for an indirect effect through the 'not up to date with technologies' statement. This did not translate into a significant effect of the overall skill loss scale in our benchmark mediation model because of the (insignificant) effect of the statements capturing general skill loss and/or social skill loss.

5. Discussion and conclusion

This study contributed to the multidisciplinary literature on the negative duration dependence of unemployment. It complemented recent large-scale field experiments showing that at least part of this negative duration dependence can be given a demand-side explanation: employers are reluctant to hire long-term unemployed job candidates. Using vignettes, we took the logical next step in this literature and empirically explored four theoretical explanations for unemployment scarring. Our analyses provided evidence that employers' reluctance to hire LTU is to a large extent mediated by their perception of unemployment as a signal of lower motivation. This is very much in line with findings from the qualitative study of Bonoli and

Hinrichs (2012) as well as with results obtained by Atkinson et al. (1996) and Bonoli (2014) on the basis of employer surveys. We also found that a smaller fraction of the total effect of unemployment duration on hiring intentions was associated with rational herding, that is, the belief that other employers found the candidate's productivity to be low (in line with Oberholzer-Gee, 2008).

From a policy point of view, our findings show that LTU might benefit from including in their job applications a detailed statement about their motivation to find work as well as a credible justification for their time out of work. We believe that the focus in this respect should be on work motivation and not on general motivation because an additional mediation analysis with interaction variables showed that the effect of unemployment duration on hiring intentions was not moderated by applicants' engagement in volunteer work.[x] Furthermore, labour market policies should also take into account potential asymmetric information between employers and job candidates. Indeed, policies aiming to increase productivity of LTU might be ineffective if this increased productivity is not properly signalled to employers when applying to their vacancies.

We end this article by acknowledging limitations inherent to our experiment and briefly highlighting related directions for further research. Most importantly, while the estimated total effect of unemployment duration on hiring intentions (i.e. the δ of our mediation model) and its effect on the tested candidate perceptions (i.e. our δ_i) can be given a causal interpretation, this is not the case for the estimated association of these perceptions with hiring intentions (i.e. our θ_i). Given that the aim of our study is to explore all potential signals related to a long unemployment duration, we would have to experimentally manipulate these perceptions

separately to be able to measure their causal impact. However, we do not see a setting in which jointly manipulating these perceptions would be feasible. Indeed, it would be very difficult to signal, for example, skill loss in a vignette in a realistic way. Nevertheless, it would make an interesting follow-up study to experimentally manipulate some of the different signals. Another interesting avenue for future research into the mechanisms behind signalling would be to experimentally manipulate the timing and continuity of the unemployment spell(s). In this way one could causally test whether these factors serve as independent signals or whether they substitute or reinforce one another.

While we found a number of interesting and significant mediation effects, we nevertheless also reported a large and significant direct effect, indicating that a considerable portion of the scarring effect of unemployment still remained unexplained (Shrout and Bolger, 2002; Zhao et al., 2010). This suggests the need for further theoretical development going beyond the four signals included. Our experiment does not allow us to identify the direction this future theory development should take, so we can only speculate. One interesting avenue could be to look into a signal of overqualification. It could indeed be the case that when a person remains unemployed for a longer period, she/he will cast a wider net during the job search and apply for positions for which she/he is overqualified. If employers assume this to be the case, this could be a potential negative signal associated with a long unemployment spell (as overqualified candidates may not fit their low-status vacancy). The negative effect of a bachelor degree on hiring intentions is consistent with this explanation. On the other hand, the significant direct effect can also result from our statements imprecisely measuring the four signals. Indeed, measurement errors in our mediators may have resulted in downward-biased estimates for the mediation effects and an upward-biased estimate for the direct effect (Judd

and Kenny, 1981, Vanderweelde et al., 2013).

Contrary to field experiments, the data collection within a vignette experiment does not take place under real-life circumstances and participants are aware to take part in an experiment. Although this is an advantage from a research-ethical point of view (Charness et al., 2013; Riach and Rich, 2004) and necessary to get an insight into thought processes (Baert and De Pauw, 2014; Van Hoye and Lievens, 2003), participants may answer in a socially desirable way when not exposed to the urgency of real-life decision-making. While this is considered a serious issue for direct question-based surveys (Auspurg and Hinz, 2014), we believe this to be less of a concern in vignette experiments in general, and in our design in particular, for two main reasons. Firstly, the widespread use of vignette studies in the social and behavioural sciences is related to the fact that self-reported measures of perceptions have been shown to correlate highly with actual behaviour and that changes in intentions clearly result in actual behavioural changes (Hainmueller et al., 2015). Secondly, in a vignette experiment each participant is only shown a small number of vignettes that vary with regard to multiple factors and therefore it is almost impossible for the participant to know what the socially desirable answer is (Auspurg and Hinz, 2014; Liechti et al., 2017; Mutz, 2011). In this respect, the reader should also note that the factor of interest in our study (unemployment duration) is a generally socially acceptable screen (Bills, 1990)—much less sensitive than, for example, race—and, as a consequence, socially desirable answers are expected to be negligible.

With respect to the generalisability of our findings, our approach is subject to the same limitations as those found in the field experiments we mimicked. We only measured unequal treatment based on a single recent unemployment spell towards individuals with a specific

profile (i.e. two or five years of experience, with a secondary education degree or a bachelor's degree) applying for a specific position in a specific context (i.e. Flanders). As a consequence, our findings cannot be easily generalised to settings with jobs and candidate profiles different from those used in this study, or to other geographical regions. Indeed, it is possible that the stigma of unemployment is more or less present in other settings. In particular, there may be systematic variation across countries, as unemployment is differently regulated across institutional contexts (Gangl, 2004). Similarly, the relative value of some signals related to unemployment may differ across occupations. For instance, the value of social capabilities could be lower in occupations without (much) contact with customers or co-workers. Alternatively, the reported lack of significance for the skill loss scale may be due to the fact that the occupation of counter assistant requires mainly general skills that are less subject to depreciation. More generally, Mosthaf (2014) argues that as the incidence of unemployment is more typical for low-skilled workers, the negative signals related to long-term unemployment may be weaker for them (compared with high-skilled workers).

This being said, the consistency of our results with findings from earlier studies conducted in very different contexts, namely Switzerland (Bonoli, 2014) and the United Kingdom (Atkinson et al., 1996), and different populations, including low-educated LTU in six European countries (Bonoli and Hinrichs, 2012), suggests—at the very least—that the belief that LTU are particularly lacking in motivation is widespread across employers. Nevertheless, further research is necessary to ensure the robustness of our results in other settings. With the recent economic downturn, many people have suffered a spell of unemployment: we welcome a program of research that looks more closely at the scars they carry from a demand-side perspective. For instance, semi-structured interviews with employers (e.g. Bonoli and Hinrichs,

2012) and/or employees could deepen the insights from our study. In addition, research that combines testing in the field with psychological tests in the manner of Rooth (2010) or that integrates vignettes in large-scale and possibly representative employer surveys could be very fruitful.

References

Acemoglu, D. (1995). Public policy in a model of long-term unemployment, *Economica*, **246**, 161-178.

Ahmed, A. M., Andersson, L., & Hammarstedt, M. (2012). Does age matter for employability? A field experiment on ageism in the Swedish labour market, *Applied Economics Letters*, **19**, 403-406.

Arrow, K. J. (1973). The theory of discrimination. In O. Ashenfelter and A. Rees (Eds.), *Discrimination in labour markets*. Princeton: Princeton University Press.

Atkinson, J., Giles, L. and Meager, N. (1996). *Employers, Recruitment and the Unemployed*. Brighton: The Institute for Employment Studies.

Auspurg, K., and Hinz, T. (2014). *Factorial survey experiments*. Thousand Oaks: Sage.

Auspurg, K., Hinz T., and Liebig, S. (2009). *Complexity, Learning Effects, and Plausibility of Vignettes in Factorial Surveys*. Working Paper 4. Research Project: The Factorial Survey as a Method for Measuring Attitudes in Population Surveys

Auspurg, K., Hinz, T., Liebig, S. and Sauer, C. (2014). The factorial survey as a method for

measuring sensitive issues. In U. Engel, B. Jann, P. Lynn, A. Scherpenzeel, and P. Sturgis (Eds.), *Improving survey methods: Lessons from recent research*. New York: Routledge.

Baert, S. (2018a). Facebook profile picture appearance affects recruiters' first hiring decisions. *new media & society*, **20**, 1220-1239.

Baert, S. (2018b). Hiring discrimination: an overview of (almost) all correspondence experiments since 2005. In *Audit Studies: Behind the Scenes with Theory, Method, and Nuance* (pp. 63-77). Springer, Cham.Baert, S., Cockx, B., Gheyle, N., & Vandamme, C. (2015). Is there less discrimination in occupations where recruitment is difficult?, *ILR Review*, **68**, 467-500.

Baert, S. and De Pauw, A. S. (2014). Is ethnic discrimination due to distaste or statistics?, *Economics Letters*, **125**, 270-273.

Baert, S., De Pauw, A. S., & Deschacht, N. (2016). Do employer preferences contribute to sticky floors?, *ILR Review*, **69**, 714-736.

Baert, S., Norga, J., Thuy, Y., & Van Hecke, M. (2016). Getting grey hairs in the labour market. An alternative experiment on age discrimination. *Journal of Economic Psychology*, **57**, 86-101.

Banerjee, A. (1992). A simple model of herd behaviour, *Quarterly Journal of Economics*, **107**, 797-817.

Becker, G. S. (1962). Investment in human capital: A theoretical analysis, *Journal of Political Economy*, **70**, 9-49.

Becker, G. S. (1994). *Human capital: A theoretical and empirical analysis with special*

reference to education. Chicago: University of Chicago Press.

Bills, D. B. (1990). Employers' use of job history data for making hiring decisions: A fuller specification of job assignment and status attainment. *The Sociological Quarterly*, **31**, 23-35.

Bonoli, G. (2014). Employers' attitudes towards long-term unemployed people and the role of activation in Switzerland, *International Journal of Social Welfare*, **23**, 421-430.

Bonoli, G. and Hinrichs, K. (2012) Statistical discrimination and employers' recruitment: Practices for low-skilled workers, *European Societies*, **14**, 338-361.

Brand, J. E. (2015). The far-reaching impact of job loss and unemployment, *Annual Review of Sociology*, **41**, 359-375.

Calvó-Armengol, A. and Jackson, M. O. (2004). The effects of social networks on employment and inequality, *American Economic Review*, **94**, 426-454.

Charness, G., Gneezy, U. and Kuhn, M. A. (2013). Experimental methods: Extra-laboratory experiments-extending the reach of experimental economics, *Journal of Economic Behavior and Organization*, **91**, 93-100.

Clark, A., Georgellis, Y. and Sanfey, P. (2001). Scarring: The psychological impact of past unemployment, *Economica*, **68**, 221-241.

Cockx, B. and Picchio, M. (2013). Scarring effects of remaining unemployed for long-term unemployed school-leavers, *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, **176**, 951-980.

DiPrete, T. A. (2002). Life course risks, mobility regimes, and mobility consequences: A

comparison of Sweden, Germany, and the United States, *American Journal of Sociology*, **108**, 267-309.

Di Stasio, V. (2014). Education as a signal of trainability: Results from a vignette study with Italian employers, *European Sociological Review*, **30**, 796-809.

Eriksson, S., and Rooth, D. O. (2014). Do employers use unemployment as a sorting criterion when hiring? Evidence from a field experiment, *American Economic Review*, **10**, 1014-1039.

Farber, H. S., Silverman, D. and Von Wachter, T. (2016). Determinants of callbacks to job applications: An audit study, *American Economic Review*, **106**, 314-318.

Gangl, M. (2004). Welfare states and the scar effects of unemployment: A comparative analysis of the United States and West Germany, *American Journal of Sociology*, **109**, 1319-1364.

Gangl, M. (2006). Scar effects of unemployment: An assessment of institutional complementarities, *American Sociological Review*, **71**, 986-1013.

Gerard M and Valsamis D (2015) Arbeidstekorten op de Europese arbeidsmarkt. *Over.Werk*, **25**, 8–14

Hainmueller, J., Hangartner, D. and Yamamoto, T. (2015). Validating vignette and conjoint survey experiments against real-world behaviour, *Proceedings of the National Academy of Sciences*, **112**, 2395-2400.

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. New York: Guilford Press.

International Labour Organization (ILO). International Labour Organization Database

(ILOSTAT) – Unemployment by sex, age and duration. International Labour Organization (ILO).

Jasso, G. (2006). Factorial survey methods for studying beliefs and judgments, *Sociological Methods and Research*, **34**, 334-423.

Judd, C. M., & Kenny, D. A. (1981). Process analysis: Estimating mediation in treatment evaluations. *Evaluation review*, **5**, 602-619.

Kaas, L., & Manger, C. (2012). Ethnic discrimination in Germany's labour market: a field experiment. *German economic review*, **13**, 1-20.

Kroft, K., Lange, F. and Notowidigdo, M. J. (2013). Duration dependence and labour market conditions: Evidence from a field experiment, *Quarterly Journal of Economics*, **128**, 1123-1167.

Kuhfeld, W. F. (2010). *Statistical graphics in SAS: an introduction to the graph template language and the statistical graphics procedures*. SAS Publishing.

Lahey, J. N. (2008). Age, women, and hiring an experimental study. *Journal of Human resources*, **43**, 30-56.

Liechti, F., Fossati, F., Bonoli, G. and Auer, D. (2017). The signaling value of labor market programs, *European Sociological Review*, **33**, 257-*274.*

Luijkx, R. and Wolbers, M. H. (2009). The effects of non-employment in early work-life on subsequent employment chances of individuals in the Netherlands, *European Sociological Review*, **25**, 647-660.

McFarland, S. G. (1981). Effects of question order on survey responses. *Public Opinion*

Quarterly, **45**, 208-215.

Mincer, J. and Ofek, H. (1982). Interrupted work careers: Depreciation and restoration of human capital, *Journal of Human Resources*, **17**, 3-24.

Mooi-Reci, I. and Ganzeboom, H. B. (2015). Unemployment scarring by gender: Human capital depreciation or stigmatization? Longitudinal evidence from the Netherlands, 1980–2000, *Social Science Research*, **52**, 642-658.

Mosthaf, A. (2014). Do Scarring Effects of Low-Wage Employment and Non-Employment Differ BETWEEN Levels of Qualification?. *Scottish Journal of Political Economy*, **61**, 154-177.

Mutz, D. C. (2011). *Population-based survey experiments*. Princeton: Princeton University Press.

Oberholzer-Gee, F. (2008). Nonemployment stigma as rational herding: A field experiment, *Journal of Economic Behavior & Organization*, **65**, 30-40.

OECD. (2013). *Employment outlook 2013*. Paris: OECD.

Oreopoulos, P. (2011). Why do skilled immigrants struggle in the labor market? A field experiment with thirteen thousand resumes. *American Economic Journal: Economic Policy*, **3**, 148-71.

Petit, P. (2007). The effects of age and family constraints on gender hiring discrimination: A field experiment in the French financial sector. *Labour Economics*, **14**, 371-391.

Riach, P. and Rich, J. (2004). Fishing for discrimination, *Review of Social Economy*, **62**, 465-486.

Riach, P. A., & Rich, J. (2006). An experimental investigation of sexual discrimination in hiring in the English labor market. *Advances in Economic Analysis & Policy*, **5**.

Rooth, D. A. (2010). Automatic associations and discrimination in hiring: Real world evidence, *Labour Economics*, **17**, 523-534.

Rossi, P. H. and Nock, S. L. (1982). *Measuring social judgments: The factorial survey approach*. Thousand Oaks: Sage.

Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: new procedures and recommendations. *Psychological methods*, **7**.

Spence, M. (1973). Job market signalling, *Quarterly Journal of Economics*, **87**, 355-374.

Thurow, L. C. (1975). *Generating inequality: Mechanisms of distribution in the U.S. economy*. New York: Basic Books.

VanderWeele, T. J., Valeri, L., & Ogburn, E. L. (2012). The role of measurement error and misclassification in mediation analysis. *Epidemiology (Cambridge, Mass.)*, **23**.

Van Hoye, G. and Lievens, F. (2003). The effects of sexual orientation on hirability ratings: An experimental study, *Journal of Business and Psychology*, **18**, 15-30.

Vishwanath, T. (1989). Job search, stigma effect, and escape rate from unemployment, *Journal of Labor Economics*, **7**, 487-502.

Wallander, L. (2009). 25 years of factorial surveys in sociology: A review, *Social Science Research*, **38**, 505-520.

Zhao, X., Lynch Jr, J. G., & Chen, Q. (2010). Reconsidering Baron and Kenny: Myths and truths about mediation analysis. *Journal of consumer research*, **37**, 197-206.

Endnotes

[i] Belgium is a federal state with three regions. Flanders is the largest region, situated in the North. The Flemish hiring landscape is an interesting setting for this study in at least two ways. First, while unemployment rates in Belgium are comparable to the average of the Eurozone, the share of long-term unemployment (i.e. one year or more) is more than 50% (ILOSTAT), which is fairly high in international comparison. In particular, in Flanders, the share of long-term unemployment was 50.3% in 2018 (source: Public Employment Agency of Flanders). Second, overall, the competition for human capital is relatively high in comparison to other European countries (Baert, 2108a; Gerard and Valsamis, 2015). Indeed, in the first quarter of 2018, the job vacancy rate in Flanders was 3.37% as opposed to 2.2% for EU-28 (source: Eurostat).

[ii] We note three such supply-side explanations. First, a long unemployment spell might reduce one's search intensity when looking for a job. Clark et al. (2001) showed that the unemployed can become indifferent to the prospect of becoming employed after a lengthy unemployment spell. A second explanation is the lack of a network experienced by LTU (Calvó-Armengol and Jackson, 2004). Finally, human capital theory (Becker, 1962; 1994) predicts that LTU will experience skill loss over the course of the unemployment spell. It is important to note that these supply-side explanations could have a demand-side effect through the associated perceptions of employers. Indeed, the important difference between both groups of explanations is the mechanism behind them. While the demand-side explanations assume that the hiring process is characterised by asymmetric information and that, as a result, employers make assumptions based on group differences, the supply-side explanations on the other hand assume that employers adequately evaluate changes in productivity due to the long unemployment spell.

[iii] In the methodological literature on vignette experiments (Auspurg and Hinz 2014), five is the lower bound suggested for the number of vignette factors. We decided to stick to this minimum to limit respondents' fatigue, taking into account the relatively large number of judgements we asked them to make for each vignette (see Section 3.2).

[iv] We assessed the difference in means between the initial 242 participants and the 219 participants with complete responses using t-tests. The results of these tests are available upon request.

[v] One should note that the order of these statements did not vary between vignettes, therefore we cannot exclude an order effect (McFarland, 1981).

[vi] The HR professionals were first shown a résumé of a candidate with an unemployment spell of four years and were asked whether they would consider hiring this candidate, and why (not). In the second part of the interview, we talked about 'long unemployment spells' in more general terms, allowing it up to the discretion of the HR professional to determine how she/he interpreted this.

[vii] Oberholzer-Gee (2008) also prompts participants to rate statements to test for different signals. The statement related to skill loss ('I prefer the candidate with a job because the unemployed applicant has lost some skills and she is not familiar with recent developments in the profession') is very close to our three statements capturing this signal. Additionally, he also includes a statement for rational herding: 'I prefer the candidate with a job because the unemployed applicant is probably not very productive. If she were productive, she would have been hired by another firm.'

[viii] With respect to the calculation of these t-statistics, it is important to account for the nested structure of data collected through a vignette experiment, with multiple vignettes judged by the same participant (Jasso, 2006). To this end, we take into account the dependence of the error term within participants by clustering all estimated t-values at the participant level.

[ix] The corresponding correlation matrix is reported in Table A7.

[x] The results of this analysis are available on request.