First Depressed, Then Discriminated Against?

Stijn Baert, Sarah De Visschere, Koen Schoors,

Désirée Vandenberghe and Eddy Omey

Abstract

Each year a substantial share of the European population suffers from major depression. This mental illness may affect individuals' later life outcomes indirectly by the stigma it inflicts. The present study assesses hiring discrimination based on disclosed depression. To this end, between May 2015 and July 2015, we sent out 288 trios of job applications from fictitious candidates to real vacancies in Belgium. Within each trio, one candidate claimed to have become unemployed only recently, whereas the other two candidates revealed former depression or no reason at all for their unemployment during a full year. Disclosing a year of inactivity due to former depression decreases the probability of getting a job interview invitation by about 34% when compared with candidates who just became unemployed, but the stigma effect of a year of depression is not significantly higher than the stigma effect of a year of unexplained unemployment. In addition, we found that these stigmas of depression and unemployment were driven by our male trios of fictitious candidates. As a consequence, our results are in favour of further research on gender heterogeneity in the stigma of depression and other health impairments.

Keywords: Belgium; health; depression; hiring discrimination; field experiments.

1 Introduction

According to the 2011 meta-analysis on mental health in Europe by Wittchen et al. (2011), each year 6.9% of the European population suffers from major depression. In addition to the direct impact of this mental illness on people's lives, present and former depression affect individuals' (later) life outcomes indirectly by the stigma depression inflicts (McGinty et al., 2015; Schwenk, Davis, & Wimsatt, 2010; Whitley & Campbell, 2014). A frequently reported consequence of depression stigma is the discrimination the (formerly) depressed undergo in their search for suitable accommodation (Corrigan, Larson, Watson, Boyle, & Barr, 2006; Whitley & Campbell, 2014) and/or gainful employment (Brohan, Gauci, Sartorius, & Thornicroft, 2011; Corrigan et al., 2006). In the present study, we focus on the stigma of depression in the labour market.

Many studies have documented diminished labour market activity related to depression (see e.g. Frijters, Johnston, & Shields, 2010; Krause, 2013) and the consequent economic burden for both individuals and society (Kessler et al., 2008). Moreover, it has been shown that being distanced from the labour market makes depression more persistent (Gebel & Vossemer, 2014; Lloyd & Waghorn, 2007; Roy & Schurer, 2013). Therefore, not surprisingly, the reintegration into the labour market of employees inactive due to (former) depression is a key ambition of many OECD countries (OECD, 2013).

The development of adequate policy responses requires the assessment of the hurdles (formerly) depressed individuals face when attempting to reintegrate into the labour market. Next to supply side differences in human capital and preferences (Elinson, Houck, Marcus, & Pincus, 2004; Ettner, Frank, & Kessler, 1997), hiring discrimination based on (former) depression may be one of the key hurdles facing (formerly) depressed individuals. As predicted theoretically by Becker (1957) and Arrow (1973), employers may hesitate to hire employees with mental problems due to a distaste (of the employers, co-workers, or customers) to working with them, a fear of diminished productivity, or anticipated sick leave problems. Yet, the stigma effect of a depression-related sick leave period may be dominated by the well-documented stigma effect of a non-health-related unemployment period of comparable length (Vishwanath, 1989).

Some studies provide indicative *empirical* evidence of hiring discrimination based on disclosed depression (Ando et al., 2013; Brohan et al., 2011; Corrigan et al., 2006; Henderson, Little, Thornicroft, & Williams, 2013; Stuart, 2006). However, because these studies are based on survey data, their findings may reflect perceptions of discrimination and unobserved differences in human capital rather than causal evidence of unequal treatment. In addition, they might be biased due to reverse causality, i.e. due to an effect of economic attainment on mental health (Antonakakis & Collins, 2015; Barr, Kinderman, & Whitehead, 2015; Tøge & Blekesaune, 2015).

In this study, we assess hiring discrimination based on disclosed depression in a direct and causal way. To this end, we send out a total of 864 fictitious job applications to real vacancies in Belgium. These applications differ only in the labour market history of the candidates: one became unemployed (at most) a few weeks before the application, a second became unemployed one year earlier and does not provide the employer with a reason for her/his unemployment, and a third candidate became jobless at the same time as the second candidate but explains this break in employment by severe depression. By monitoring the subsequent reactions from the employer side, we are able to identify the effect on employment opportunities of disclosing a jobless year due to depression compared to two realistic counterfactual situations (i.e. no substantial break in employment and a comparable break without mentioning depression).

Drawing on the mentioned literature concerning the stigma of and discrimination based on depression, we propose the following hypotheses:

H1a: Individuals with a break in employment due to depression get less positive callback in response to their job applications compared to similar candidates with no substantial break in employment.

H1b: Individuals with a break in employment due to depression get less positive callback in response to their job applications compared to similar candidates with a comparable but unexplained break in employment.

The empirical literature documents lower epidemiology of and more negative attitudes towards depression among men (Berger, Addis, Reilly, Syzdek, & Green, 2012; Ogrodniczuk & Oliffe, 2011; Oliffe & Philips, 2008; Pattyn, Verhaeghe, & Bracke, 2015; Van de Velde, Bracke, & Levecque, 2010; Wittchen et al., 2011). In addition, other studies suggest that men disclosing depression may suffer more stigmatisation than their female peers (McCusker and Pez Galupo, 2011). We therefore proceed to inspect the candidate gender heterogeneity in unequal treatment due to disclosed depression and test the following hypothesis.

H2: Former depression hurts hiring chances more for male candidates.

2 The Experiment

We set up a correspondence experiment in the spirit of Bertrand and

Mullainathan (2004). Within such an experiment, fictitious job applications are sent to real vacancies. The applications that are sent to the same vacancy are equivalent, except for the characteristic of interest. By monitoring the subsequent callback, unequal treatment based on this single characteristic is identified. This correspondence testing framework is widely viewed as providing the most convincing evidence of hiring discrimination (Riach & Rich, 2002). Without such experimental data, researchers possess considerably less data than employers. For instance, data on general ability and work motivation are most of the time not observed in survey and administrative data. By consequence, applicants who appear similar to researchers on the basis of standard non-experimental data may in fact be very different in the eye of their prospective employers. As long as not all variables driving hiring, remuneration, and promotion decisions that may correlate with mental health are controlled by the researcher, analyses might suffer from selection bias. A correspondence experiment, in contrast, eliminates selection based on individual unobservable characteristics because the researcher fully controls the information available to the employer, allowing the researcher to disentangle discrimination from alternative explanations of heterogeneous hiring outcomes, such as differences in human capital or in employee preferences.

Our experiment was conducted between May 2015 and August 2015 in Flanders, the Dutch-speaking part of Belgium. Following the Belgian Health Interview Survey, gathered between 2001 and 2013, 10.1% (8.9%) of the population in Belgium (Flanders) suffered, at the moment of the survey, from symptoms of a depressive disorder (average over all age groups in, for each survey year, a random sample of 10,000 individuals residing in Belgium, regardless their place of birth, nationality or any other characteristic). These percentages were higher for women (12.9% in Belgium and 11.3% in Flanders) than for men (7.3% in Belgium and 6.6% in Flanders) (source: Scientific Institute of Public Health, Belgian Health Interview Survey). In addition, these percentages for Belgium and Flanders were somewhat higher than the European average presented by the aforementioned meta-analysis by Wittchen et al. (2011). Also following the European Core Health Indicators, Belgium was the country with the highest prevalence of depression in 2008 among 15 European countries (source: Eurostat, People reporting a chronic disease by disease, sex, age, and educational attainment level; based on national surveys using a common questionnaire; average over all ages). With respect to labour market performance, the unemployment rate in 2015 in Belgium (8.5%) and Flanders (5.2%) was lower than the average in the EU-27 (9.3%) (source: Eurostat, Unemployment rates by sex, age, and NUTS 2 regions). In addition, the labour market tightness was relatively high in Belgium in 2015: the job vacancy rate (i.e. the number of job vacancies as a percentage of the sum of the number of occupied posts and the number of job vacancies) was 2.1% in this country in the first quarter of 2015, while it was 1.7% in the EU-27 (source: Eurostat, Job vacancy rate).

Three applications of unemployed candidates were sent to 288 vacancies. From the database of the Public Employment Agency of Flanders — the region's major job search channel — we randomly selected 72 vacancies in the occupations of laboratory worker (ISCO-08 classification number 3111), representative (ISCO-08 3322), production worker (ISCO-08 81), and barkeeper (ISCO-08 5131). With respect to the broad occupation of production worker, in particular vacancies for the sub-occupations of chemical products plant and machine operators (ISCO-08 8131), plastic products machine operators (ISCO-08 8142), food and related products machine operators (ISCO-08 8160), and packing, bottling, and labelling machine operators (ISCO-08 8183) were tested.

These occupations were chosen for the expected variation in levels of skill and customer contact. In addition, the labour market tightness in these occupations differed. The median vacancy duration for all vacancies in the database of the Public Employment Agency of Flanders in 2015 was the highest in the occupation of barkeeper (78 days) and the lowest in the occupation of labelling machine operator (34 days). Given that the median vacancy duration had a mean of about 66 days and a standard deviation of about 59 over all 563 occupations observed in the database of the Public Employment Agency of Flanders in 2015, the occupations included in our experiment were all characterised by rather moderate labour market tightness.

We created three template types of resumes and cover letters, which we refer to as 'Type A', 'Type B', and 'Type C' applications, for each of the four aforementioned occupations, each matching the general requirements of these occupations in terms of schooling and experience. These application template types differed in inessential details and in layout but were, at the level of the occupation, identical in all job-relevant characteristics. To ensure that our resumes and cover letters were realistic, we used examples from the Public Employment Agency of Flanders as basic templates.

The Type A, Type B, and Type C resumes revealed the following information. All applicants were born, living, and had studied in Antwerp, the largest city in Flanders. They were 37 or 38 years old and married. The candidates applying for a position as a laboratory worker (representative) held a bachelor's degree in chemistry (commercial sciences). Those applying for the occupations of operator and barkeeper left education after high school with a technical degree in tourism (Type A), commerce (Type B), or secretary-languages (Type C). After leaving school, all candidates had been working in jobs similar to the one for which they applied until April 2014, a year before the start of our experiment. In addition, all fictitious individuals were unemployed at the start of our experiment, i.e. in May 2015. Finally, all applicants revealed the following characteristics: a Flemish-sounding name; a birthdate in 1977 or 1978; a Belgian nationality; a postal address in a middle-class neighbourhood (non-existing house number in an existing street); a mobile phone number and an email address (both from major providers); adequate Dutch, French, and English language skills; moderate ICT skills; one practised sport; and a driving license. To avoid employers' detection of the experiment, the resumes differed in layout, a variety of common wordings were used for the candidates' (bachelor's) degree and mentioned skills, and the candidates were given different sports activities (korfball, tennis, and fitness for the Type A, B, and C applicants, respectively).

All cover letters mentioned that the job applicant (i) found the vacancy in the database of the Public Employment Agency of Flanders, (ii) was an experienced candidate with the right qualifications, (iii) was motivated to start the job, and (iv) was looking forward to a job interview. Again, to avoid detection of the experiment, a variety of common wordings were used for the Type A, B, and C templates. The resumes and cover letters used are available on request.

We sent three applications — one of Type A, Type B, and Type C — to each selected vacancy. For each vacancy, we randomly assigned three situations faced by the applicants between May 2014 and April 2015 to the Type A, B, and C applications. Figure 1 schematises the resulting trajectories of the experimental identities.

A first (control) candidate was still employed between May 2014 and April 2015 in the same job as the one she/he held before May 2014. As a consequence, this identity was, at the moment of her/his application, unemployed for between less than one month (those who applied in May) and less than three months (those who applied in July).

A second member of the trio was out of work between May 2014 and April 2015. As a consequence, this candidate's unemployment lasted one year more than that of the first member of the trio. This experimental identity did not mention any reason for this unemployment in her/his cover letter. She/he mentioned in this cover letter the following: "As you can read in my resume, I have been unemployed during the last year. I am, however, very motivated to start in a new job." As a consequence, the "unemployed" candidates with whom we compare the (formerly) depressed candidates are actually candidates with an employment break which employers themselves can fill in with activities and reasons that are the most plausible in their perception. We come back to this point when discussing the limitations of this study in Section 4.

A last member of the trio indicated former severe depression as the reason for her/his year out of employment between May 2014 and April 2015. This was done by adding the clause: “In view of a trustful collaboration, I want to mention that during the last year I was inactive due to severe depression. Today, I have completely recovered and am ready for a new professional challenge.”

An alternative approach would have been to just compare two recently unemployed candidates of which one disclosed former depression. However, there is no reason why job candidates would reveal former depression if this depression had no labour market consequences for them. Therefore, such an experimental setting would not have been realistic, and the likeliness of detection of such an experiment by employers would have been too high. In contrast, longer unemployment periods may be perceived as a signal of lower ability and work motivation if they are not explained (see below). As a consequence, hardly any legal explanation (even former depression) for longer periods without work may be perceived as suspect by employers.

To eliminate any application type effects on callbacks, we alternately assigned the mentioned experimental identities to the Type A, Type B, and Type C applications. As a consequence of this random assignment, there was no correlation between experimental identity and application template type so that the aforementioned minimal differences between Type A, B, and C application templates could not bias our results.

To inspect the candidate gender heterogeneity in unequal treatment due to disclosed depression, we also alternated the gender of the three candidates between the trios. More concretely, to the odd vacancies we sent a trio of female candidates, whereas to the even vacancies we sent a trio of male candidates.

The resulting combinations were sent to the employers between 18 May 2015 and 11 July 2015, in an alternating order, each time with approximately 24 hours in between the trio members. Callbacks were received by telephone voicemail or email. In our analysis, we follow Baert et al. (2015) in distinguishing between two definitions of positive callback. Positive callback *sensu stricto* means the applicant was invited for an interview concerning the job for which she/he applied. Positive callback *sensu lato* also includes the request to contact the recruiter by telephone or to provide more information by email or the proposal of an alternative job. All callbacks received later than 30 days after sending out the application were ignored.

Between November 2013 and March 2014 we ran a pilot experiment in which only two applications were sent to 152 Flemish vacancies. These vacancies were unbalanced over the four occupations included in the final experiment. In this pilot, the experimental identity without a substantial break in employment was not included. However, with respect to the comparison of the identities included in this pilot experiment, the findings were very similar to those based on the experiment on which we report in the present study.

This research was reviewed and approved by the Ethical Committee of the Faculty of Economics and Business Administration of Ghent University at its meeting of 9 July 2013. The report of this meeting (in Dutch) is available on request.

3 Results

Table 1 describes the experimentally gathered data. In general, for 65 (115) vacancies, at least one of our three fictitious job applicants received a positive callback in strict (broad) sense. In 14 (49) of these vacancies, each of the three candidates received a positive callback. Next, in 19 (19), 14 (8), and 4 (7) of the situations, only the recently unemployed, long-term unemployed, and depressed identities, respectively, received a positive callback. Finally, in 5 (14) of the vacancies, there was positive callback *sensu stricto* (*sensu lato*) for only the recently unemployed and the long-term unemployed, in 3 (10) vacancies only for the recently unemployed and the depressed, and in 6 (8) vacancies only for the long-term unemployed and the depressed candidate.

TABLE 1 ABOUT HERE.

Based on these statistics, we can calculate the positive callback rates, i.e. the average probability of receiving a positive response, for our experimental identities. These rates are presented in column (2), column (3), and column (4) of Table 2. Panels A and C provide these statistics at the level of the total dataset. Overall, the recently unemployed candidate got a positive callback *sensu stricto* (*sensu lato*) in 14.2% (31.9%) of her/his applications. Her/his counterparts with a substantial employment break got a positive reaction in 10.1% (27.1%) of the cases when mentioning no reason for unemployment and in 9.4% (25.7%) of the cases when revealing former depression.

These rates suggest a preference for the recently unemployed over both the long-term unemployed and (formerly) depressed identities. However, we cannot assess the significance of their differences in callback chances based on these statistics. Therefore, we follow the literature by calculating two measures comparing callback outcomes identity-by-identity: the positive callback ratio, as outlined in the last three columns of Table 2, and the net discrimination rate, as presented in Table 3 (Baert et al., 2015, 2016; Bertrand & Mullainathan, 2004; McGinnity & Lunn, 2011; Riach & Rich, 2002).

The positive callback ratio is calculated by dividing the positive callback rate for a first group of candidates by the corresponding positive callback rate for a second group of candidates. Panel A and panel C of column (6) in Table 2 show that the probability of getting a positive callback is substantially higher for candidates without an employment break compared to candidates mentioning a year of depression. The positive callback ratio *sensu stricto* (*sensu lato*) is 1.519 (1.243) when comparing these experimental identities. This means that the former candidates receive 51.9% (24.3%) more job interview invitations (positive reactions in broad sense) than the latter candidates. In other words, the probability of getting invited to a job interview decreases by 34.2% (i.e. 1 – [1/1.519]) and the probability of any positive reaction decreases by 19.5% for candidates revealing a recent year out of work due to depression instead of a year of employment. Next to being substantial in magnitude, these ratios are also statistically significant at the 5% (1%) significance level. Thus, in line with our research hypothesis H1a, employers penalise job candidates for a recent career break due to depression *ceteris paribus*.

However, panel A and panel C of column (5) show that an employment break of comparable length, but without mentioning any reason, is subject to a similar penalisation. The latter result is in line with the literature indicating a negative relationship between (unexplained) unemployment duration and hiring chances (Eriksson & Rooth, 2014; Kroft, Lange, & Notowidigdo, 2013). In addition, this finding is complementary to the more general finding of “state dependence” in individual unemployment (or “unemployment scarring”), i.e. the pattern that individuals who are unemployed in the present are more likely to be unemployed in the future *ceteris paribus* (Arulampalam, Gregg, & Gregory, 2001; Cockx & Picchio, 2012; Daly & Delaney, 2013; Heckman & Borjas, 1980; Mooi-Reci & Ganzeboom, 2015). In particular, our result is related to (and empirically supports) one of the major explanations for state dependence in employment besides human capital depreciation and loss of self-esteem, i.e. unemployment as a signal of lower ability and motivation to employers.

Column (7) presents the positive callback ratios with respect to the comparison of the two experimental identities with a substantial employment break. It is found that the probability of positive callback *sensu stricto* (*sensu lato*) is 7.4% (6.8%) higher when not mentioning any reason for being out of work than when mentioning depression as a reason. However, these differences in positive callback are not statistically significant. So, we cannot reject that the stigma effect of a year of (inactivity due to) depression is equal in magnitude to the stigma effect of a year of unexplained unemployment. In other words, we do not find evidence for H1b.

TABLE 2 ABOUT HERE.

Next, we inspect the heterogeneity in unequal treatment by the gender of the candidate (based on panels B and D of Table 2). When splitting the data by the gender of the trios of candidates, significant measures of unequal treatment are only found for the subsample of male candidates. Male candidates without a break in employment are preferred above male candidates with such a break due to former depression. However, the same is true when comparing male candidates without a break and male candidates with a break due to (unexplained) unemployment. In other words, our general findings of comparable stigmas of a year of inactivity due to depression and a year of unemployment are driven by the male rather than the female trios. We come back to the significance of this dimension of heterogeneity in unequal treatment below.

The net discrimination rate comparing the callback for two experimental identities is calculated in two steps. First, we reduce the number of applications for which the first identity (e.g. the recently unemployed) received a positive callback and the second identity (e.g. the long-term unemployed) received none by the number of applications for which the reverse was true. Second, we divide the result of this calculation by the number of application pairs for which at least one of these two identities got a positive callback. So, in line with Baert et al. (2015), McGinnity and Lunn (2011), and Riach and Rich (2002), when, for a particular comparison, neither of the identities received a positive callback, we treat this as a non-observation in this analysis at the vacancy level. The final result is a net measure of the number of unfavourable unequal treatment acts that the latter applicant could expect to encounter per application for which at least one of the two identities under investigation received a positive callback. At the level of the total dataset, the net discrimination rates presented in Table 3 (panels A and C) lead to exactly the same conclusions as the aforementioned positive callback ratios. The net discrimination rates indicate a preference for the recently unemployed over both identities with a year out of work (with depression as the reason or no explanation at all). Also, the net discrimination rates by candidate gender (panels B and D of Table 3) lead to the same conclusions as those based on the aforementioned positive callback ratios.

TABLE 3 ABOUT HERE.

As the candidates' situations between May 2014 and April 2015 (employment, unemployment, or inactivity due to depression) were assigned randomly within our trios of applications, regressing positive callback (*sensu stricto* or *sensu lato*) on an indicator of these situations yields exactly the same conclusion as the one based on panels A and C of Table 2. In addition, as we randomly assigned the candidates' gender between trios, regressions including interactions between the candidates' treatment status and their gender should lead to the same empirical pattern as the one in panel B and panel D of Table 2, at least for a sample size approaching infinity. However, our sample size is finite. Thus, the gender of the trios might correlate with (un)observable vacancy characteristics. As these characteristics may affect the hiring outcomes of our candidate pairs, not controlling for them could yield biased measures of the heterogeneity of discrimination based on former depression by the candidates' gender. Therefore, in line with recent correspondence experiments (Baert et al., 2015, 2016; Drydakis, In press), we further explore the experimental data by a regression analysis including vacancy fixed effects. Another reason for conducting a regression analysis is that it allows us to investigate the heterogeneity of discrimination against formerly depressed job candidates by other dimensions of observed variation in the experimentally gathered data, i.e. variation in the occupation and contract type mentioned in the vacancy.

More specifically, we regress positive callback on (i) having been out of employment between May 2014 and April 2015, (ii) having been out of employment during this period due to (disclosed) depression, and (iii) various sets of variables interacted with (i) and (ii). The results of these regressions are presented in Table 4. In models (1) to (3) we adopt the *sensu stricto* definition of the dependent variable; in models (4) to (6) the *sensu lato* definition. In all regressions we control for vacancy fixed effects such that any impact of vacancy characteristics (without interaction with (i) and (ii)) is controlled. Random effect estimations yield equivalent results.

TABLE 4 ABOUT HERE.

In regression models (1) and (4), we only include the indicators of having been a year out of work and disclosing former depression as the reason for this break. By doing that, we find that a year out of employment decreases the probability of a job interview invitation (the probability of any positive reaction) by 4.5 (5.4) percentage points. These outcomes equal the differences between the overall positive callback rates *sensu stricto* (*sensu lato*) for the experimental identities discussed above based on panel A (panel C) of Table 2. For instance: -0.045 = (0.101 + 0.094)/2 - 0.142. In addition, disclosing depression as the reason for the one-year break has, compared with giving no reason other than unemployment, no significant effect on the probability of positive callback.

Secondly, in models (2) and (5), we include interactions between the treatment indicators adopted in models (1) and (4) and the gender of the candidate. In line with our discussion of panel B and panel D of Table 2, we find that male candidates are punished more for a jobless year than female candidates, irrespective of whether depression is the reason for this break in employment or not. In other words, we find a higher stigma of both former depression and unemployment for male candidates compared with female candidates. The higher stigma of depression for males is in line with the empirical evidence reviewed in the introduction (and, therefore, with our research hypothesis H2). The higher stigma of unemployment among them may be partly due to the fact that it is more common for women to have a period of inactivity around the age of 35 because of child-rearing activities (OECD, 2011). Consequently, as suggested by an anonymous reviewer of a former version of the present study, our findings may also have a cultural explanation. In addition, the general finding of less unequal treatment based on (former) depression and unemployment against female candidates might be explained by the lower overall number of applications from women for the tested occupations. Indeed, the fraction of women among the unemployed who indicated in 2015 to be interested in the occupations of representative, operator, and barkeeper in the database of the Public Employment Agency of Flanders was only about 32%, 26%, and 29%, respectively (for the occupation of laboratory worker, this was about 51%). Therefore, employers who pursued a gender balance in their firm might have been less picky with respect to female candidates for these occupations *ceteris paribus*.

Lastly, in models (3) and (6), we include additional interactions with indicators for jobs as a representative, jobs as a production worker, jobs as a barkeeper, vacancies offering temporary contracts, and vacancies offering part-time contracts. The effect of having been a year out of work and disclosing former depression as the reason for this break is somewhat less adverse for the occupations of representative, production worker, and barkeeper compared to the occupation of laboratory worker. Given that, as mentioned in Section 2, labour market tightness was the highest in the latter occupation, this finding contrasts to some extent to what Baert et al. (2015) found, i.e. that ethnic discrimination is lower in Flanders in occupations where labour market tightness is higher. However, none of the interactions added in models (3) and (6) have a statistically significant effect (at the 5% level) on the probability of positive callback.

4 Discussion

We investigated hiring discrimination based on former depression in a direct, empirical way. In contrast to former contributions on labour market discrimination against (formerly) depressed employees, we did not use survey data but gathered unique field experimental data. Thereby, we made sure that our results were not biased by inverse causality and did not reflect perceptions of discrimination and unobserved differences in human capital rather than causal evidence of unequal treatment. More concretely, we sent out trios of fictitious job applications to real vacancies in Belgium. We randomly assigned the treatment of a break in employment due to depression and two relevant counterfactual situations, i.e. no substantial break in employment and a comparable break without mentioning depression, within these trios. In addition, we alternated between male and female trios. This enabled us to contribute further to the literature by inspecting whether unequal treatment based on disclosed depression (and unexplained unemployment) was heterogeneous by candidates' gender.

We showed that, in the tested vacancies in the Belgian labour market, the probability of getting invited to a job interview decreased by about one third, and the probability of getting any positive reaction decreased by about one fifth for candidates revealing a recent year out of work due to depression (compared to the situation of no substantial employment break). In that respect, our study shows that the stigma of depression, as amply documented in the literature, causes discrimination in the labour market, at least in Belgium. However, the penalisation for disclosing inactivity due to former depression is not more severe than the penalisation for an unexplained unemployment period of comparable length. In addition, we found that these comparable stigmas of a year of inactivity due to depression and a year of unemployment were driven by the male rather than the female trios we sent out. As a consequence, our results are in favour of focusing further on gender heterogeneity in the stigma of depression and, by extension, the stigma of other health impairments.

We acknowledge four research limitations inherent to our experimental design. First, we gave no alternative explicit explanation for the unemployment of the experimental identity who did not mention (former) depression as a reason for her/his long-term unemployment. As a consequence, the “unemployed” applicant in our experiment could also be a formerly depressed applicant not disclosing her/his health problems. Stated otherwise, we did not observe the assumptions employers made about the health or other stigma conditions concerning the latter applicant. As mentioned in Section 2, the “depressed” candidates we compared to “unemployed” candidates in our framework were, therefore, actually “openly depressed” and the “unemployed” candidates were actually candidates with an employment break that employers themselves could fill in with activities and reasons that were the most plausible in their perception. As a result, we compared the costs associated with disclosing former depression to the costs associated with leaving employers uninformed about reasons for lack of employment. We believe that this is a relevant trade-off faced by formerly depressed job candidates in reality. In addition, when comparing the callback for the candidate without a substantial employment break and the (formerly) depressed candidate, we could not decompose the penalty for the latter candidate in a part inherent to her/his additional year outside the labour market and a part inherent to her/his depression. Nevertheless, we believe comparing the callback for these candidates is relevant, because also in real life, candidates who only differ in former depression status and *thereby* in the length of their non-employment spell have to compete.

A second limitation of this article is its focus on a particular labour market outcome, i.e. employers' first hiring decisions. We only measured differences in their callbacks so that our findings cannot be translated into divergences in final job offers, wages, or promotion opportunities. So, in comparison with former contributions to the literature on discrimination based on depression, we benefited from a research design guaranteeing causal discrimination measures at the cost of giving up on scope. However, Bendick, Brown, and Wall (1999), Bertrand and Mullainathan (2004), and Cédiey, Foroni, and Garner (2008) show that (i) reduced interview rates translate into reduced final job offers, wages, and promotion opportunities and that (ii) if discrimination occurs, it does so especially in the first phase of the hiring process.

Thirdly, we only measured the discrimination against (formerly) depressed (and long-term unemployed) candidates within jobs in the four tested occupations submitted to the database of the Public Employment Agency of Flanders. Although this limitation is less acute in our design compared to many former studies focusing on only one occupation (Riach & Rich, 2002), it is still possible that the stigma of depression (or unexplained unemployment) is more or less present in other occupations than those covered. For instance, in line with Baert et al. (2015), discrimination based on former depression could be expected to be lower (higher) in occupations with very high (low) labour market tightness. However, this limitation, if important, should cause a similar shift in the penalty for depression (and unemployment) for males and females. As a consequence, this finding should not bias our conclusions with respect to the heterogeneity in unequal treatment by the gender of the candidate. The same is true for the aforementioned limitations.

A final limitation has further repercussions for the generalisability of our findings. In this study, unequal treatment based on former depression and unemployment was measured for married people aged 37 or 38. As both treatments might have a different effect (as such and by gender) for different candidate characteristics, our results cannot be easily generalised to other age groups and people with a different civil status.

We end this article with two policy reflections. Firstly, in many OECD countries, a legal framework to punish labour market discrimination is available (Bassanini & Saint-Martin, 2008), so that in view of reducing discrimination against (formerly) depressed candidates, the main benefit seems to lie in a more vigorous detection of it. To implement this, one could consider a (systematic) application of the experimentation framework used in this study. Second, from an individual job candidate's perspective, our results do not yield a superior strategy with respect to (not) mentioning former depression as the reason for an employment break. However, they indicate that, compared to females, males have a greater interest in avoiding employment breaks, irrespective of whether they are related to health or other reasons.

References

Ando, S., Bacle, S.V., Bonetto, C., Cristofalo, D., Economou, M., Germanavicius, A., Lanfredi, M., Lasalvia, A., Lopez-Ibor, J.J., Reneses, B., Sartorius, N., Thornicroft, G., Van Audenhove, C., Van Bortel, T., Van Weeghel, J., Wahlbeck, K., Zoppei, S., & ASPEN/INDIGO Study Group. (2013). Global pattern of experienced and anticipated discrimination reported by people with major depressive disorder: A cross-sectional survey. *Lancet*, *381*, 55–62.

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., De Pauw, A.-S. & Deschacht, N. (2016). Do Employer Preferences Contribute to Sticky Floors? *ILR Review*, *69*, 714–736.

Antonakakis, N., & Collins, A. (2015). The impact of fiscal austerity on suicide mortality: Evidence across the ‘Eurozone periphery’. *Social Science & Medicine*, *145*, 63–78.

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

Arulampalam, W., Gregg, P., & Gregory, M. (2001). Unemployment scarring. *Economic Journal*, *111*, 577–584.

Barr, B., Kinderman, P., & Whitehead, M. (2015). Trends in mental health inequalities in England during a period of recession, austerity and welfare reform 2004 to 2013. *Social Science & Medicine*, *147*, 324–331.

Bassanini, A., & Saint-Martin, A. (2008). The price of prejudice: Labour market discrimination on the grounds of gender and ethnicity. In OECD. *Employment outlook*. Paris: OECD.

Becker, G.S. (1957). *The economics of discrimination*. Chicago: The University of Chicago Press.

Bendick, M., Brown, L.E., & Wall, K. (1999). No foot in the door: An experimental study of employment discrimination against older workers. *Journal of Aging & Social Policy*, *10*, 5–23.

Berger, J.L., Addis, M.E., Reilly, E.D., Syzdek, M.R., & Green, J.D. (2012).

Effects of gender, diagnostic labels and causal theories on willingness to report symptoms of depression. *Journal of Social and Clinical Psychology*, *31*, 439–457.

Bertrand, M., & Mullainathan, S. (2004). Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *American Economic Review*, *94*, 991–1013.

Brohan, E., Gauci, D., Sartorius, N., & Thornicroft, G. (2011). Self-stigma, empowerment and perceived discrimination among people with bipolar disorder or depression in 13 European countries: The GAMIAN-Europe study. *Journal of Affective Disorders*, *129*, 56–63.

Cédiey, E., Foroni, F., & Garner, H. (2008). Discrimination à l'embauche fondée sur l'origine à l'encontre des jeunes français(e)s peu qualifié(e)s. Dares Premières Infos Premières Synthèses 06.3.

Cockx, B., & Picchio, M. (2012). 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.

Corrigan, P.W., Larson, J.E., Watson, A.C., Boyle, M., & Barr, L. (2006). Solutions to discrimination in work and housing identified by people with mental illness. *Journal of Nervous and Mental Disease*, *194*, 716–718.

Daly, M., & Delaney, L. (2013). The scarring effect of unemployment throughout adulthood on psychological distress at age 50: Estimates controlling for early adulthood distress and childhood psychological factors. *Social Science & Medicine*, *80*, 19–23.

Drydakis, N. (In press). The effect of university attended on graduates' labour market prospects: A field study of Great Britain. *Economics of Education Review.* DOI: 10.1016/j.econedurev.2016.03.001.

Elinson, L., Houck, P., Marcus, S.C., & Pincus, H.A. (2004). Depression and the ability to work. *Psychiatric Services*, *55*, 29–34.

Eriksson, S., & Rooth, D.-O. (2014). Do employers use unemployment as a sorting criterion when hiring? Evidence from a field experiment. *American Economic Review*, *104*, 1014–1039.

Ettner, S., Frank, R., & Kessler, R. (1997). The impact of psychiatric disorders on labor market outcomes. *ILR Review*, *51*, 64–81.

Frijters, P., Johnston, D.W., & Shields, M.A. (2010). Mental health and labour market participation: Evidence from IV panel data models. IZA Discussion Paper 4883.

Gebel, M., & Vossemer, J. (2014). The impact of employment transitions on health in Germany. A difference-in-differences propensity score matching approach. *Social Science & Medicine*, *108*, 128–136.

Heckman, J.J., & Borjas, G.J. (1980). Does unemployment cause future unemployment? Definitions, questions and answers from a continuous time model of heterogeneity and state dependence. *Economica*, *47*, 247–283.

Henderson, C., Little, K., Thornicroft, G., & Williams, P. (2013). Mental health problems in the workplace: Changes in employers' knowledge, attitudes and practices in England 2006–2010. *British Journal of Psychiatry*, *202*, 70–76.

Kessler, R.C., Heeringa, S., Lakoma, M.D., Petukhova, M., Rupp, A.E., Schoenbaum, M., Wang, P.S., & Zaslavsky, A.M. (2008). Individual and societal effects of mental disorders on earnings in the United States: Results from the national comorbidity survey replication. *American Journal of Psychiatry*, *165*, 703–711.

Krause, A. (2013). Don't worry, be happy? Happiness and reemployment. *Journal of Economic Behavior & Organization*, *96*, 1–20.

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

Lloyd, C., & Waghorn, G. (2007). The importance of vocation in recovery for young people with psychiatric disabilities. *British Journal of Occupational Therapy*, *70*, 50–59.

McCusker, M.G., & Galupo, M.P. (2011). The impact of men seeking help for depression on perceptions of masculine and feminine characteristics. *Psychology of Men & Masculinity*, *12*, 275–284.

McGinnity, F., & Lunn, P.D. (2011). Measuring discrimination facing ethnic minority job applicants: An Irish experiment. *Work, Employment and Society*, *25*, 693–708.

McGinty, E.E., Goldman, H.H., Pescosolido, B., & Barry, B.L. (2015). Portraying mental illness and drug addiction as treatable health conditions: Effects of a randomized experiment on stigma and discrimination. *Social Science & Medicine*, *126*, 73–85.

Mooi-Reci, I., & 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.

OECD. (2011). *Babies and bosses - reconciling work and family life: A synthesis of findings for OECD countries*. Paris: OECD.

OECD. (2013). *Mental health and work*. Paris: OECD.

Ogrodniczuk, J.S., & Oliffe, J.L. (2011). Men and depression. *Canadian Family Physician*, *57*, 153–155.

Oliffe, J.L., & Philips, M.L. (2008). Men, depression and masculinities: A review and recommendations. *Journal of Men's Health*, *5*, 194–202.

Pattyn, E., Verhaeghe, M., & Bracke, P. (2015). The gender gap in mental health service use. *Social Psychiatry and Psychiatric Epidemiology*, *50*, 1089–1095.

Riach, P.A., & Rich, J. (2002). Field experiments of discrimination in the market place. *Economic Journal*, *112*, 480–518.

Roy, J., & Schurer, S. (2013). Getting stuck in the blues: Persistence of mental health problems in Australia. *Health Economics*, *22*, 1139–1157.

Schwenk, T.L., Davis, L., & Wimsatt, L.A. (2010). Depression, stigma, and suicidal ideation in medical students. *JAMA-Journal of the American Medical Association*, *304*, 1181–1190.

Stuart, H. (2006). Mental illness and employment discrimination. *Current Opinion in Psychiatry*, *19*, 522–526.

Tøge, A.G., & Blekesaune, M. (2015). Unemployment transitions and self-rated health in Europe: A longitudinal analysis of EU-SILC from 2008 to 2011. *Social Science & Medicine*, *143*, 171–178.

Van de Velde, S., Bracke, P., & Levecque, K. (2010). Gender differences in depression in 23 European countries: Cross-national variation in the gender gap in depression. *Social Science & Medicine*, *71*, 305–313.

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

Wittchen, H.U., Jacobi, F., Rehm, J., Gustavsson, A., Svensson, M., Jönsson, B., Olesen, J., Allgulander, C., Alonso, J., Faravelli, C., Fratiglioni, L., Jennum, P., Lieb, R., Maercker, A., van Os, J., Salvador-Carulla, L., Simon, R., & Steinhausen, H.C. (2011). The size and burden of mental disorders and other disorders of the brain in Europe 2010. *European Neuropsychopharmacology*, *21*, 655–679.

Whitley, R., & Campbell, R.D. (2014). Stigma, agency and recovery amongst people with severe mental illness. *Social Science & Medicine*, *107*, 1–8.

Figure 1: Experimental Identities.

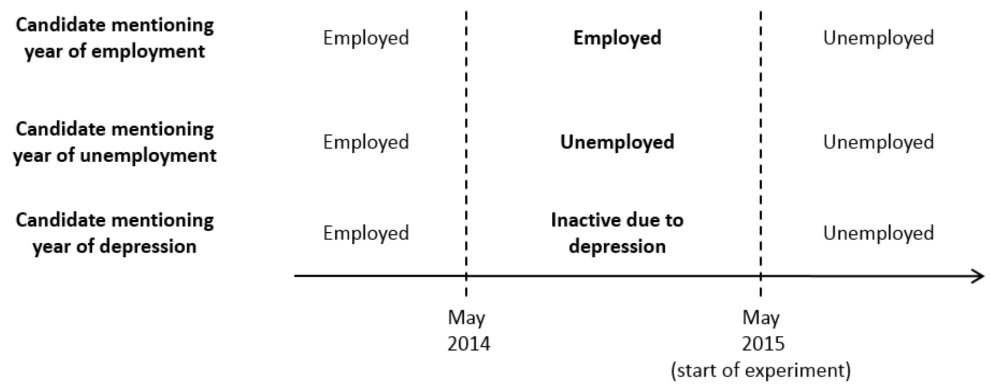

Table 1: The Probability of Positive Callback: Descriptive Statistics.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Observations	**Jobs**	**None of the three candidates positive callback**	**Each of the three candidates positive callback**	**Only candidate mentioning year of employment positive callback**	**Only candidate mentioning year of unemployment positive callback**	**Only candidate mentioning year of depression positive callback**	**Only candidate mentioning year of employment and candidate mentioning year of unemployment positive callback**	**Only candidate mentioning year of employment and candidate mentioning year of depression positive callback**	**Only candidate mentioning year of unemployment and candidate mentioning year of depression positive callback**
A. Positive callback* sensu stricto*: All observations									
All	288	233	14	19	4	4	5	3	6
B. Positive callback* sensu stricto*: Heterogeneity by gender of the candidate									
Males	144	113	9	13	1	3	3	1	1
Females	144	120	5	6	3	1	2	2	5
C. Positive callback* sensu lato*: All observations									
All	288	173	49	19	8	7	14	10	8
D. Positive callback* sensu lato*: Heterogeneity by gender of the candidate									
Males	144	84	27	13	1	3	8	5	3
Females	144	89	22	6	7	4	6	5	5

Table 2: The Probability of Positive Callback: Positive Callback Rates and Positive Callback Ratios.

Observations	(1) Jobs	(2) Positive callback rate candidate mentioning year of employment	(3) Positive callback rate candidate mentioning year of unemployment	(4) Positive callback rate candidate mentioning year of depression	(5) PCR employed versus unemployed: (2)/(3)	(6) PCR employed versus depressed: (2)/(4)	(7) PCR unemployed versus depressed: (3)/(4)
A. Positive callback **sensu stricto*: All observations***							
All	288	0.142	0.101	0.094	1.414** [2.133]	1.519** [2.419]	1.074 [0.499]
B. Positive callback **sensu stricto*: Heterogeneity by gender of the candidate***							
Males	144	0.181	0.097	0.097	1.857*** [3.082]	1.857*** [2.739]	1.000 [0.000]
Females	144	0.104	0.104	0.090	1.000 [0.000]	1.154 [0.532]	1.154 [0.705]
C. Positive callback **sensu lato*: All observations***							
All	288	0.319	0.274	0.257	1.165* [1.946]	1.243*** [2.622]	1.068 [0.799]
D. Positive callback **sensu lato*: Heterogeneity by gender of the candidate***							
Males	144	0.368	0.271	0.264	1.359*** [3.066]	1.395*** [2.959]	1.026 [0.810]
Females	144	0.271	0.278	0.250	0.975 [0.836]	1.083 [0.652]	1.111 [0.851]

Notes. The positive callback ratio (PCR) is calculated by dividing the positive callback rate for a first group of candidates by the corresponding positive callback rate for a second group of candidates. T-statistics, indicating whether the ratios are significantly different from 1 and based on standard errors corrected for clustering at the vacancy level, are between brackets. *** (**) ((*)) indicate significance at the 1% (5%) ((10%)) significance levels, respectively.

Table 3: The Probability of Positive Callback: Net Discrimination Rates.

Observations	(1) Jobs	(2) NDR employed versus unemployed	(3) NDR employed versus depressed	(3) NDR unemployed versus depressed
A. Positive callback* sensu stricto*: All observations				
All	288	0.235** [4.500]	0.275** [5.765]	0.056 [0.250]
B. Positive callback* sensu stricto*: Heterogeneity by gender of the candidate				
Males	144	0.429*** [9.000]	0.400*** [7.200]	0.000 [0.000]
Females	144	0.000 [0.000]	0.095 [0.286]	0.111 [0.500]
C. Positive callback* sensu lato*: All observations				
All	288	0.120* [3.756]	0.168*** [6.750]	0.052 [0.641]
D. Positive callback* sensu lato*: Heterogeneity by gender of the candidate				
Males	144	0.246*** [8.909]	0.254*** [8.333]	0.021 [0.059]
Females	144	-0.020 [0.043]	0.063 [0.429]	0.082 [0.727]

Notes. The net discrimination rate (NDR) is calculated by reducing the number of applications for which the former candidate was preferred by the number of applications for which the latter candidate was preferred, and this difference is then divided by the number of application pairs in which at least one received a positive callback. The chi-square test for the net discrimination rate tests the null hypothesis that both candidates are treated unfavourably with the same frequency. χ^2-statistics are between brackets. *** (**) ((*)) indicate significance at the 1% (5%) ((10%)) significance levels, respectively.

Table 4: The Probability of Positive Callback: Linear Probability Model Regression Estimates.

	(1)	(2)	(3)	(4)	(5)	(6)
Year out of work	-0.045** (0.019)	-0.045** (0.018)	-0.045** (0.018)	-0.054** (0.021)	-0.054*** (0.021)	-0.054*** (0.021)
Year out of work x Depression as reason (normalised)	-0.007 (0.014)	-0.007 (0.014)	-0.007 (0.014)	-0.017 (0.022)	-0.017 (0.022)	-0.017 (0.022)
Year out of work x Female candidate (normalised)		0.083** (0.039)	0.086** (0.039)		0.104** (0.046)	0.099** (0.046)
Year out of work x Depression as reason x Female candidate (normalised)		-0.014 (0.028)	-0.014 (0.029)		-0.021 (0.043)	-0.021 (0.042)
Year out of work x Representative (normalised)			0.015 (0.058)			0.023 (0.065)
Year out of work x Depression as reason x Representative (normalised)			0.042 (0.042)			0.028 (0.063)
Year out of work x Production worker (normalised)			0.029 (0.048)			0.014 (0.063)
Year out of work x Depression as reason x Production worker (normalised)			0.013 (0.037)			-0.015 (0.061)
Year out of work x Barkeeper (normalised)			0.046 (0.070)			0.062 (0.084)
Year out of work x Depression as reason x Barkeeper (normalised)			0.069 (0.052)			0.109* (0.065)
Year out of work x Temporary contract (normalised)			0.027 (0.064)			-0.116 (0.085)
Year out of work x Depression as reason x Temporary contract (normalised)			0.002 (0.053)			0.005 (0.085)
Year out of work x Part-time contract (normalised)			-0.071 (0.077)			-0.132 (0.089)
Year out of work x Depression as reason x Part-time contract (normalised)			0.032 (0.051)			0.064 (0.073)
Constant	0.142*** (0.012)	0.142*** (0.012)	0.142*** (0.012)	0.319*** (0.014)	0.319*** (0.014)	0.311*** (0.014)
Dependent variable: invitation to a job interview	Yes	Yes	Yes	No	No	No
Dependent variable: any positive reaction	No	No	No	Yes	Yes	Yes
Vacancy fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	864	864	864	864	864	864

Notes. Except for the variable 'Year out of work', all independent variables are normalised. The triple interactions are normalised by subtracting their mean among the subpopulation of formerly depressed candidates. The other variables are normalised by subtracting their mean among the subpopulation of candidates who were a year out of employment. Standard errors, corrected for clustering of the observations at the vacancy level, are in parentheses. *** (**) ((*)) indicate significance at the 1% (5%) ((10%)) levels, respectively.