



ELSEVIER

Contents lists available at ScienceDirect

Research in Social Stratification and Mobility

journal homepage: www.elsevier.com/locate/rssm

Gender, parenthood, and hiring intentions in sex-typical jobs: Insights from a survey experiment

Gabriele Mari^{a,*}, Ruud Luijkx^b^a Erasmus University Rotterdam, the Netherlands^b Tilburg University, the Netherlands

ARTICLE INFO

Keywords:

Hiring
Gender
Parenthood
Survey experiment
Statistical discrimination
Status-characteristic theory

ABSTRACT

We ran a survey experiment with Dutch employers to investigate hiring discrimination in sex-typical jobs. We ask if women are especially discriminated against when they have children, whether discrimination applies similarly in different occupations, and whether statistical discrimination or status-characteristic theories best account for discriminatory practices (if any).

Employers rate fictitious candidates for either a female-typical job (primary-school teacher) or a male-typical job (software engineer). Employers are found to display a slight preference for female candidates when filling a teacher post, although such bias is less strong for female applicants with children. No such ranking is found for a software engineer vacancy, nor do we find different salary offers across candidates and across vacancies. Employers do not appear to favour men over women for positions likely to be on the career track, as predicted by statistical discrimination theories, nor do they expect women to be less capable than men, as posited by status-characteristic theory. Female candidates with children, however, are expected to be less committed to their job and work fewer hours, especially in the teacher experiment. Such expectations seem to have small consequences for the hiring decisions and salary offers Dutch employers make in our study.

1. Introduction

Employers may shape gender disparities in the workplace when they treat differently otherwise equally-productive men and women. Favouring male candidates over comparable female ones at the hiring stage is one instance of such discriminatory practices. Hiring is increasingly understood as a key stage in the (re)production of labour market disparities (e.g. Bills, Di Stasio, & Gërkhani, 2017; Di Stasio & van de Werfhorst, 2016; Protsch & Solga, 2015) and gender discrimination in hiring is well-documented (Azmat & Petrongolo, 2014; Neumark, 2016). Questions still loom, however, on whether women are especially discriminated against when they have children, whether discrimination applies similarly in different occupations, and which mechanisms drive discrimination. We address here this threefold gap by means of a survey experiment on the hiring intentions of Dutch employers.

First, previous research has yielded mixed support for gender-by-parental-status discrimination resulting in a motherhood penalty at the hiring stage. A seminal study combining a lab and a field experiment in the US found mothers to be the least-preferred candidate for hire in marketing and business jobs (Correll, Benard, & Paik, 2007). Field

experiments in the financial sector in France (Petit, 2007) and across a wide array of jobs in Sweden (Bygren, Erlandsson, & Gähler, 2017) have not replicated this finding. Differently, a comparably large field experiment in Spain found lower callback probabilities for women especially if with children (José González, Cortina, & Rodríguez, 2019). A vignette study in Switzerland found evidence of a motherhood penalty for women applying for a HR assistant position (Oesch, Lipps, & McDonald, 2017), and a field experiment in Germany found a motherhood penalty in hiring for an event manager position (Hipp, 2019). We add to this debate by examining how employers evaluate job candidates at the intersection of gender and parental status in the Netherlands. The Dutch context features a “residual” gender wage gap of around 8% that studies attribute in part to discrimination (Fransen, Plantenga, & Vlasblom, 2012), as well as persistent gender segregation by field of study and, consequently, occupation (OECD, 2017b: 144; van de Werfhorst, 2017). At the same time, the evidence for a motherhood penalty is mixed in the Dutch labour market (cf. Davies & Pierre, 2005; De Hoon, Keizer, & Dykstra, 2017) making the Netherlands a compelling *litmus test* for the analysis of gender-by-parental-status discrimination.

Our second contribution consists in running one experiment for a

* Corresponding author.

E-mail addresses: mari@essb.eur.nl (G. Mari), r.luijkx@uvt.nl (R. Luijkx).

<https://doi.org/10.1016/j.rssm.2019.100464>

Received 27 February 2019; Received in revised form 19 December 2019; Accepted 19 December 2019

0276-5624/ © 2020 Elsevier Ltd. All rights reserved.

male-typical job and one for a female-typical job. We picked a software engineer job vacancy and a primary-school teacher vacancy, respectively, for they both require similar levels of tertiary education in the Netherlands, but widely differ with respect to opportunities for career progression and task content. We leverage these latter features to put to test distinct theories of labour market discrimination. Also, research on the motherhood penalty in hiring has so far focused on mixed occupations (for an exception, [Bygren et al., 2017](#)). Yet experimental studies on gender discrimination at large have found differential treatment *in favour* of women in female-typical jobs and, albeit to a lesser extent, against women in male-typical jobs ([Neumark, 2016](#); [Riach & Rich, 2002](#); [Rich, 2014](#)). We complement these findings by assessing if and how parental status modifies gender discrimination in sex-typical occupations.

Finally, we develop and test the predictions of two theories, statistical discrimination theory and status-characteristic theory (e.g. [Correll & Benard, 2006](#)). While the former framework would predict particularly pronounced gender differentials in male-typical jobs and holding irrespective of parental status, the latter would suggest that mothers may be better off in female-typical jobs. Further, employers' propensity to discriminate statistically may vary depending on the length of the prospective employment relationship, with permanent contracts being conducive to higher risk-aversion on the part of employers and thereby lower hiring chances for the groups employers deem more 'risky' to employ. Alternatively, we measure the expectations employers form with respect to the ability and commitment of different job candidates, and assess the role of these constructs in accounting for hiring patterns as predicted by status-characteristic theory.

In short, we ask if there is a motherhood penalty in hiring chances, whether it varies along occupational and type-of-contract lines, and if such variation fits with the predictions of either economic or socio-psychological models of discrimination. Rather than relying on undergraduate students as in most experimental simulations of the hiring process ([Koch, D'Mello, & Sackett, 2015](#)), participants in our study are real-world employers selected among the respondents of LISS, a web panel carried out since 2007 on a true probability sample of the Dutch population ([Scherpenzeel, 2011](#)). While field experiments involve real-world employers too, what is typically measured is only the outcome of employers' decision over a candidate, that is, whether or not a given job applicant gets a callback. We are able here not only to survey employers' decision over a candidate, but also to inquire about employers' motives.

2. Background

2.1. Statistical discrimination: risk aversion against women

The design of our study is informed by economic and socio-psychological models of discrimination, namely statistical discrimination theory and status-characteristic theory. Both are commonly featured in the literature on motherhood penalties in labour markets (e.g. [Bygren et al., 2017](#); [Correll et al., 2007](#); [Gangl & Ziefle, 2009](#); [Oesch et al., 2017](#)). Both theories also pin down employer discrimination to stereotypical beliefs about groups. We distinguish them for analytical purposes, emphasising, at the outset, what tells them apart; we will return to their commonalities in our concluding remarks.

Under statistical discrimination, employers are rational actors aiming at maximizing expected profits in labour markets characterized by imperfect information (for a review, [Fang & Moro, 2011](#)). Specifically, employers may find it difficult or too expensive to access precise information on the individual productivity of job applicants. Group markers such as a candidate's sex are instead easily accessible during the hiring process and, for example, already when employers are screening CVs. Uncertain about individual productivity, statistically-discriminating employers may make inferences based on such group markers.

Even if confronted with the same signal (say, candidates with equal educational credentials), employers might find the signal given by women less reliable than that of men. Female employees, in the eyes of employers, might be more likely than men to take career breaks, reduce their working hours, or leave their job altogether for family-related reasons. And thus, conditional on the same signal (e.g. the same educational credentials), employers might assign more variance¹ to the expected productivity of women than that of men ([Aigner & Cain, 1977](#); [Charles & Guryan, 2011](#)). If risk averse, further, employers might thus prefer the low-variance group ([Fang & Moro, 2011](#): 138), in this case men. Women will be hired less often or paid lower wages than equally-qualified men, as employers insure themselves against the risk of future family-related productivity losses².

In this framework, employers are forward-looking and their concern lies with women *potentially* becoming parents and altering their labour supply in ways harmful to productivity. Women of childbearing age, irrespective of current parental status, could be discriminated against with respect to men ([Baert, 2014](#); [Becker, Fernandes, & Weichselbaumer, 2019](#); [Biewen & Seifert, 2016](#); [Gupta & Smith, 2002](#); [Jessen, Jessen, & Kluge, 2019](#); [Petit, 2007](#); [Yip & Wong, 2014](#)): we name this our *family-risk hypothesis*. Under statistical discrimination, therefore, we do not expect a motherhood penalty. In fact, a strong formulation of the *family-risk hypothesis* would suggest that mothers may be even preferred to childless women. Employers might assume mothers completed their maternity leave period(s) at another firm and, conversely, fear childless women will have their maternity break at some point if hired (see [Jessen et al., 2019](#); [Petit, 2007](#)).

What can be done then to modify employers' propensity to discriminate statistically? Two variables can be manipulated, namely the information available for each job applicant and the risk employers associate with hiring him or her. Supplying precise and cheap information on individual productivity could help employers correct the signal-to-noise ratio in their productivity estimates (e.g. [Altonji & Pierret, 2001](#); [Pinkston, 2006](#)). Although information on job candidates has indeed been manipulated to detect statistical discrimination in experimental settings, doing so is problematic for two reasons ([Neumark, 2016](#)). First, researchers may not know which information employers need to reduce the expected variance around group productivity.

¹ We test this 'variance' model of statistical discrimination, rather than the classic 'mean' model ([Phelps, 1972](#)) for two main reasons. First, the Phelps model would posit that employers rely on their estimate for *group-average* (rather than group variance) productivity more for women than for men, being conversely more capable to extract a signal of *individual* productivity among men rather than women. In this model, discrimination arises out of error ([Aigner & Cain, 1977](#)). Women whose individual productivity in fact exceeds the group-average estimate made by employers will be discriminated against (e.g. underpaid). In our setting, though, we cannot assess this discrepancy between individual productivity and the group-average productivity estimated by the employer. One cannot simply measure the actual individual productivity of a fictitious job candidate. Further, the Phelps model is not guaranteed to produce unequal group outcomes (e.g. [Aigner & Cain, 1977](#); [Correll & Benard, 2006](#): 96; [Fang & Moro, 2011](#): 137-138), e.g. lower wages for women with respect to men. Women whose productivity is higher than the group average estimated by employer will be underpaid, but women whose productivity is lower than the group average estimated by employers will be overpaid. To the extent that the two balance each other out in the long run, the pay of women as a group might accurately reflect their actual average productivity. Women as a group will thus be paid less than men only if there are gender differences in actual average productivity. Yet this latter scenario is irrelevant for a theory of labour market discrimination in so far as we are willing to assume that actual gender differences in average productivity are "pre-market", they pre-exist any (experienced or expected) discriminatory behaviour by employers (for a model building on the opposite assumption, see, e.g., [Arrow, 1973](#)).

² Symmetrically, this risk aversion also involves discounting the (possibly) higher gains of employing a woman v. a man when both are actually in the upper tail of their respective productivity distribution.

Besides, even if such information is known to the experimenters, withholding it in a “low-information” experimental condition may prove to be unrealistic with respect to real-world job applications. For example, Dutch employers particularly value a candidate’s field of study and occupation-specific degrees (Di Stasio & van de Werfhorst, 2016). We felt that omitting such a commonplace information in order to simulate a low-information setting would have further compromised the external validity of our study.

Instead of information then, we turn to the manipulation of risk as a function of a) the length of the work contract being offered and b) the type of occupation. For the statistically-discriminating employer, investing in female job applicants may be more risky in permanent employment relationships as compared to temporary ones. While a temporary contract may be used to screen employees and get rid of bad matches at a low or no cost when the contract itself expires, permanent contracts are typically harder to terminate. This especially holds in contexts such as the Dutch one, where recent labour market reforms have relaxed the level of employment protection attached to temporary contracts but left largely untouched that of permanent contracts (e.g. Mooi-Reci & Dekker, 2015) – and the latter remains above the European average in the Netherlands (OECD, 2014). Other than with a costly dismissal, permanent contracts are also more likely to be associated with employer-funded training than are temporary contracts (Akgündüz & van Huizen, 2015; Fouarge, de Grip, Smits, & de Vries, 2012). Therefore, given that hiring an employee on a permanent contract entails higher economic costs, risk-averse Dutch employers will discriminate statistically against female applicants particularly when filling a permanent post, more so than when confronted with a temporary job vacancy (*contract-risk hypothesis*).

Finally, we might expect both our *family-risk* hypothesis and our *contract-risk* hypothesis to find more support in male-typical rather than in female-typical lines of work (*occupation risk hypothesis*). Occupations that are held predominantly by men often have steeper career ladders than female-dominated ones. Foreseeing further investments in terms of training and promotions, employers in male-typical jobs may thus be reluctant to hire female applicants, especially if permanently. The evidence for gender discrimination in hiring in male-dominated jobs is mixed though, with some field experiments finding lower callback rates for women vis-à-vis men and others finding such rates to be substantially and statistically indistinguishable (e.g. Bygren et al., 2017; Carlsson, 2011; Riach & Rich, 2006; for reviews, Neumark, 2016; Riach & Rich, 2002; Rich, 2014). To the best of our knowledge, only one previous study incorporated the type of work contract as an additional experimental condition, finding hiring discrimination against young childless women for permanent vacancies in the financial sector in France (Petit, 2007). More to this point, Baert, De Pauw, and Deschacht (2016) find substantially lower callback rates for young women for jobs that imply promotion prospects in a large field experiment in Belgium. Keeping this in mind, we argue that the long-term investment of hiring an applicant for a ‘career’ rather than simply for a ‘job’ makes filling permanent posts in male-typical occupations more prone to statistical discrimination against women of childbearing age.

2.2. Status-characteristic theory: status beliefs penalize mothers

According to status-characteristic theory, no matter risk aversion and the properties of expected productivity distributions, status beliefs may bias employers’ evaluations against women and, particularly, mothers (Correll & Benard, 2006; Correll et al., 2007; Ridgeway & Correll, 2004). Status beliefs are conceived as a particular class of stereotypes (e.g. Fiske, Cuddy, Glick, & Xu, 2002) by virtue of which individuals categorize members of social groups on the basis of perceived competence. Any nominal characteristic that groups together individuals in a social setting – sex, ethnicity, sexual orientation, and so forth – may become a status characteristic if actors share beliefs regarding that group’s competence, further conceptualized as the sum of

ability and commitment (Berger, Berger, Fisek, & Norman, 1977; Mark, Smith-Lovin, & Ridgeway, 2009; Ridgeway & Correll, 2006; Ridgeway, Backor, Li, Tinkler, & Erickson, 2009). Seeking coordination with one another, individuals may activate performance expectations regarding how capable and committed others will be with respect to the task at hand, depending on the salient social memberships. Performance expectations may then drive the distribution of rewards such that, in the hiring setting, low-status actors end up penalized in terms of hiring chances or salary offers (Correll et al., 2007; Pedulla, 2016, 2018; for a review, Ridgeway, 2011).

Motherhood is such a status characteristic insofar as it amplifies status beliefs morphed along gender lines (Correll et al., 2007; Ridgeway & Correll, 2004), following a pattern of “amplified congruence” between status characteristics (Pedulla, 2018). Specifically, women, especially if mothers, are perceived to be less capable than men in workplace settings (Correll et al., 2007; Cuddy, Fiske, & Glick, 2004; Thébaud, 2015). Superior ability may be granted to mothers only for tasks that involve nurturance and care (Ridgeway & Correll, 2004), in line with stereotypes broadly associating women with communion, i.e. being selfless and concerned with others, rather than with agency, i.e. being assertive and motivated to master a task (for a review, Ellemers, 2018). As for commitment, women and mothers in particular are not expected to prioritize work over family obligations nor to make sacrifices to build a career as much as men would (Correll et al., 2007; Ridgeway & Correll, 2004; Rivera & Tilcsik, 2016). Differently from our predictions under statistical discrimination, we now expect mothers to be the least-preferred candidate for hire and to receive the lowest salary offers (*motherhood penalty hypothesis*). Such penalty will be due to the fact that employers attribute the lowest levels of expected competence, i.e. of ability and commitment, to female job applicants who are also mothers (*status belief hypothesis*).

Yet, in female-typical jobs where nurturing and caring skills are crucial (teaching, nursing, social work etc.), proponents of status-characteristic theory suggest women and mothers might actually be highly regarded in terms of competence (Ridgeway & Correll, 2004). This proposition may help explain the large body of evidence on gender discrimination in favour of women in female-typical lines of work (e.g. Booth & Leigh, 2010; Carlsson, 2011; Carlsson & Eriksson, 2017; Rich, 2014). Statistical discrimination is somewhat at odds with this finding (see, e.g., Neumark, 2016: 77). Employers in female-typical jobs might still expect that young childless women will have children in the future. While arguably lower than in male-typical jobs, there will be costs (e.g. substituting those who will go on maternity leave) to hiring female candidates vis-à-vis men. Hence, risk-averse employers should still prefer men over (childless) women in female-typical jobs, under statistical discrimination.

Status-characteristic theory, on the contrary, predicts that employers will act upon their beliefs regarding what women and men “are good at”, rather than make decisions influenced by risk aversion. Women, and mothers in particular, may top employers’ ratings of competence in female-dominated jobs and be therefore more likely to be hired and receive higher salary offers than men (see also Koch et al., 2015). We name this our *status reversal hypothesis*.

Finally, differently from statistical discrimination, we remain agnostic on whether status-based discrimination may differ depending on the length of the work contract being offered.

3. Empirical approach

3.1. Data, sample, and setup

Participants in our study are sample members of the Longitudinal Internet Study for the Social Sciences (LISS), a web household panel carried out since October 2007 on a true probability sample of the Dutch population (e.g. Scherpenzeel, 2011). The core modules of the LISS questionnaire allow to track the lifecycle changes and living

conditions of panel members on a yearly basis, much like in traditional household panel surveys. *Ad hoc* modules can be submitted for consideration to CentERdata, a survey research institute affiliated with Tilburg University, and appended to one of the monthly rounds of interviews. From December 2016 to August 2017, we were thus able to run our survey experiments thanks to the LISS infrastructure. All experiment materials are available at https://www.dataarchive.lissdata.nl/study_units/view/781.

To select our “employers”, we administered four filter questions targeting respondents aged 25–65 and in paid work at the time of the most recent interview of LISS Work & Schooling module (Wave 8, April–May 2015). Of the 2985 individuals on target, 2252 (75.4%) were re-interviewed in December 2016 and 2207 (73.9%) answered to all of our filter questions that month.

Four yes/no filter questions covered respondents’ HR responsibilities in their current job. Specifically, we asked whether respondents 1) had the power to hire/fire employees, 2) took part in any phase of the recruitment process (screening of CVs, job interviews, etc.), 3) could set or influence the rate of pay received by other employees, and 4) could have influence on or decide over the promotion of other employees. 749 respondents, who answered “Yes” to at least one of these four questions, form the pool of potential participants in our study. We refer to them as employers in the study jargon.

Employers were later involved in a simulation of the hiring process (e.g. Correll et al., 2007; Di Stasio & G’erxhani, 2015; Di Stasio & van de Werfhorst, 2016). Employers were solely instructed that they would be taking part in such a simulation, with no explanation regarding the goals of the study. We ran such survey experiment across two different job vacancies (software engineer or primary-school teacher). After being randomly allocated to one of the two job vacancies, employers were presented with a CV for a first job applicant and asked to fill in a survey to evaluate him or her for the position. Employers were then presented with a second CV and repeated the questionnaire for a second candidate for the same position.

In March 2017, 150 employers were contacted for the pre-test phase and 132 (88%) completed the task, 66 for the software engineer vacancy and 66 for the teacher vacancy. In June and August 2017, a total of 480 respondents were targeted for the test phase, so that each survey experiment would have a sample size of 240 employers³. Restricting our analysis to complete cases⁴, we work with 239 participants – and 478 observations, since each employer evaluates a pair of CVs – for the teacher experiment and 237 (474 observations) for the software engineer experiment. As reported in Table 1, employers in both experiments are more likely to be men, highly-educated, and to have a partner (either married or cohabiting). Looking at their real-world HR responsibilities, the overwhelming majority of employers is involved in the screening of job applicants, matching well with our focus on employers’ ratings of job candidates based on their CVs.

3.2. Choice of jobs and job ads

We picked *software engineer* and *primary-school teacher* for our male-typical and female-typical job, respectively. Both are examples of gender-segregated occupations in top-income countries. In the Netherlands, women made up to 86% of primary-school teachers as recently as 2015 (OECD, 2017a: 400), while only around 12% of Dutch

³ Both for the pre-test and test phase sample size was determined with reference to previous studies with a similar design (Correll et al., 2007) and drawing from the total pool of 749 potential participants previously identified in LISS through our filter questions. Prior to data collection, we defined clear stopping rules for both the pre-test and the test keeping into account that the response rate in LISS usually exceeds 80%. Employers involved in the pre-test phase were not re-contacted for the test.

⁴ We dropped one respondent who expressed suspicions over the scope of the experimental task.

Table 1

Summary of sample features for each survey experiment.

Source: LISS 2016.

	Primary-school teacher	Software engineer
Female	0.44	0.37
Age	46.57	47.08
Degree (HBO or WO)	0.60	0.62
Have a partner	0.81	0.78
Have children	0.49	0.49
Have the power to hire/fire	0.47	0.43
Screen candidates (CVs, interviews)	0.90	0.87
Set or influence pay	0.44	0.35
Influence or decide on promotions	0.48	0.45
Number of employers	239	237

women figured among computing professionals in 2007 (broadly defined, ISCO-88 code 213; see Bettio & Verashchagina, 2009)⁵.

Beyond being on opposite sides of the spectrum in terms of gender segregation, our two jobs require similar educational qualifications (namely HBOs, the degrees obtained in the vocational track of the Dutch university system) yet widely differ in task content (e.g. Gathmann & Schönberg, 2010). Teaching primarily entails interactive tasks, with nurturance and care of small children being key at the primary-school level. Status beliefs on women’s and mothers’ higher competence might thus be activated in these realms, as predicted by status-characteristic theory. Further, even if in the public sector, primary-school teachers are hired directly by schools in the Netherlands, a feature that makes the hiring process comparable to that of private-sector jobs⁶. Software engineering jobs, on the other hand, not only primarily involve analytical tasks, but are also training-intensive and typically installed on well-structured career ladders, making them ideal to test our statistical discrimination argument.

As for previous studies on gender discrimination in these lines of work, only Bygren et al. (2017) have investigated the intersection of sex of the applicant and parental status in similar occupations. As part of a larger correspondence study in Sweden, they have found similar odds of a callback by sex and parental status for “computer specialists” and for “elementary school teachers”. More generally, evidence is often nil (Carlsson, 2011; Di Stasio & van de Werfhorst, 2016; Fernandez & Campero, 2017; Petersen, Saporta, & Seidel, 2000; but see Riach & Rich, 2006) on gender discrimination in hiring in tech jobs. Our choice of jobs therefore provides a stringent test for our theoretical predictions.

Job ads in our study consisted of paragraphs of around seven lines of text (see Supplementary material). The wording and information provided was derived from real ads accessible through online databases⁷ and from fake ads used in previous studies (e.g. Di Stasio & G’erxhani, 2015). Ads included a range of possible salary offers for applicants with different levels of experience, as well as a range of possible weekly working hours (32–40). Employers were referred to such ranges when asked about their salary offer to a given candidate and about the weekly working hours they would have expected the candidate to put in if hired. Ads also comprised our first manipulation, posting either a

⁵ These numbers largely reflect persistent gender segregation in the respective fields of study (Barone, 2011; van de Werfhorst, 2017), yet we ask here if employers practice hiring discrimination even if confronted with candidates with equal educational qualifications (work experience) in the relevant field (occupation).

⁶ For Dutch teachers, salaries are set with reference to a national standard nonetheless (*Salarisschaal*).

⁷ We made inquiries into online jobs postings of websites such as indeed.nl, nationalevacaturebank.nl, and meesterbaan.nl.

permanent or a temporary job vacancy to manipulate the risk associated with the prospective employment relationship.

3.3. CVs, job candidates, and manipulations

The CVs we presented to employers are stylized syntheses of real CVs accessed through free online databases (e.g. indeed.nl). CVs included the following information regarding the candidate: name, date of birth, marital status, previous work experience, maximum level and field of education attained, language skills, training courses for teachers/programming languages for engineers, and hobbies. Following previous studies (Correll et al., 2007; Petit, 2007), we sought to create a pair of similar but not identical CVs (CV A and CV B) for each job type, to minimize participants' suspicion about manipulations and about the purpose of the experiment. Specifically, we tweaked the precise date of birth, the dates, locations, and content of previous work experiences, where and when candidates got their degree, the content and dates of training courses for teachers, which programming languages engineers knew and with which proficiency, and the hobbies listed.

In the pre-test phase, we made sure that such heterogeneity across the CV pair did not result in systematic differences in the evaluation of candidate A vis-à-vis candidate B. Employers evaluated each candidate separately on a set of items for ability, commitment, and expressed the likelihood with which they would hire each candidate on a scale from 0 to 100 (see Section 3.4 for details on our dependent variables). Similar to later in the test, we looked for differences between candidates A and B in terms of each outcome by means of multilevel linear models. None of these differences were statistically significant at conventional levels nor sizeable (for all outcomes, Cohen's $d < .2$). Nevertheless, to further ensure that CV type (A or B) did not influence our results, treatments were assigned orthogonally to CV type in the test phase.

In line with common practices in the literature (for a review, Neumark, 2016), we manipulated our second binary treatment variable, sex of the job applicant, by varying the names of job candidates. Four pairs of Dutch names and surnames were randomly extracted from a list compiled in a previous study similarly manipulating names, albeit to gauge the presence of ethnic discrimination in the Dutch labour market (Blommaert, Coenders, & Van Tubergen, 2014). The four resulting pairs were Roos Kosters and Eline Vos for women, Thijs Blom and Bram Ouwehand for men. We stuck to Dutch-native names and surnames, as the analysis of the intersection of sex and ethnic discrimination is beyond the scope of our experiment. Considering that the average age at first childbirth for Dutch women is around 29 years old (OECD, 2016), we set the age of candidates at around 30 years old. This choice makes it credible for a job candidate to have one child or, if not, to be of childbearing age anyway. Indeed, under our model of statistical discrimination, we expect employers to display risk aversion against women of childbearing age, regardless of actual parental status. As for the exact manipulation of parental status, our third binary treatment, the candidate is either referred to as "married, with one child" for parents or simply "married" otherwise (e.g. Bygren et al., 2017; Petit, 2007).

To recap, our survey experiment implements a 2 (sex) by 2 (parental status) by 2 (type of contract) factorial design. To be clear, our manipulations enable us to recover the effects of signalling or giving evidence of being a man/woman, parent/non-parent (see e.g. Wenz and Hoenig 2020). Sex of the applicant is a between-subject factor, to avoid excessive suspicion on the purpose of the survey on the part of participants. The type of contract is also a between-subject factor to ease participants' task so that they would have to read only one job ad. Finally, parental status is a within-subject factor coherently with previous studies (Correll et al., 2007). This means that, for instance, a given participant could start from a job ad for a primary-school teacher, with a temporary contract being offered, and evaluate first a CV (type B) for the candidate Roos Kosters without kids and secondly evaluate a

CV (type A) for the candidate Eline Vos with kids⁸. The order in which CV types A and B, names of the candidates, and mentions of parental status appeared was counterbalanced across participants in such a way to avoid any systematic order effect.

3.4. Outcomes and models: candidate evaluations and hiring intentions

Once presented with a CV, employers evaluated the candidate first in terms of expected *competence*, further decomposed into the two constructs of ability and commitment (Correll et al., 2007; Ridgeway, 2011; Ridgeway & Correll, 2004). To capture the ability component we adopted seven items from previous studies (e.g. Correll et al., 2007; Fiske et al., 2002) and asked how "competent", "confident", "independent", "competitive", "intelligent", "skilled", and "well-trained" employers found a given candidate. For commitment (e.g. Correll et al., 2007; Heilman & Okimoto, 2008), we asked to what extent employers expected the candidate "to be very committed to the company (school)", "to make sacrifices for the job", and "to make work a top priority" if hired. Each of these items was measured on a 7-point scale, with responses ranging from "not at all" (1) to "extremely" (7). Both in the pre-test and test phases we achieved high reliability for the resulting composite measures of ability and commitment, with Cronbach's α ranging from a minimum of 0.77 to a maximum of 0.93.

We also included a more factual measure of work commitment by asking employers how many hours they would expect a given candidate to work on a weekly basis. Similar to the online job ads we surveyed, our fictitious job ads featured a 32- to 40-h working week. Given that part-time work (under 35 h) is notoriously widespread in the Netherlands, especially among women and not limited to those among them with small children (e.g. Bosch, Deelen, & Euwals, 2010), we are cautious nonetheless on the extent to which this variable may tap expected commitment.

Finally, employers were asked to estimate with what likelihood they would recommend a candidate for hire on a scale from 0 to 100 (e.g. Di Stasio & Gërkhani, 2015) and what monthly salary offer they would make to the candidate (e.g. Correll et al., 2007). Both outcomes are transformed, dividing by 100 the likelihood of hire and taking the logarithm of the salary offer, to enable the interpretation of regression coefficients in terms of percentage-point changes⁹. Hiring intentions (likelihood of hire, salary) and candidate ratings (ability, commitment, and expected hours) are our dependent variables in multilevel linear regression models. Multilevel models are chosen to account for the nested nature of our data (CVs nested within employers/subjects). Alternatively, one could opt to include employer fixed effects in simple linear regression and thereby retain the possibility of estimating coefficients for within-subject factors and for the interaction(s) of within- and between-subject factors. Such fixed-effects estimates are indistinguishable from their counterparts in our multilevel models of choice.

For each of our dependent variables, our main specification includes the main effects of job candidates' sex and parental status. The interaction between the two is our main focus as it may capture a motherhood penalty, if any (e.g. Correll et al., 2007). For both job types, we ran our models first on the full sample and secondly separate by type of work contract. Full estimation results, including a simple bivariate model comprising only the main effect for sex of the job candidate, are available in the Appendix.

⁸ Throughout the test phase, employers give individual ratings to job applicants, meaning that there are no instructions that would explicitly encourage any comparison or choice between the two candidates forming a given pair. Previous research on gender discrimination in hiring has found no evidence of heterogeneous effects depending on the type of rating (individual v. comparative) participants have been asked to give (Koch et al., 2015).

⁹ Such transformations do not alter, substantively or statistically, our results.

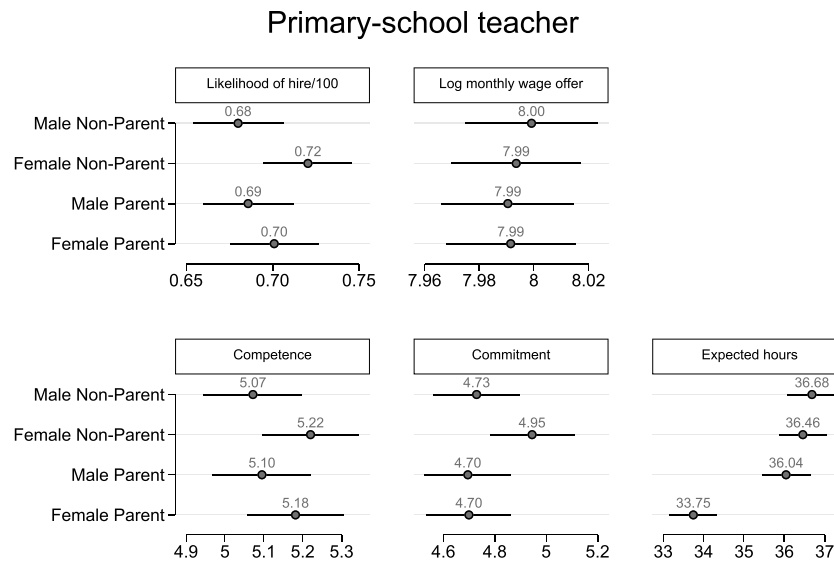


Fig. 1. Linear predictions and 95% confidence intervals. Estimate refer to multilevel linear models for hiring intentions and candidate ratings for a primary-school teacher vacancy.

4. Findings

We present our results in the form of linear predictions, focusing first on a specification including sex of the candidate, parental status, and the interaction thereof. Evidence from our primary-school teacher experiment is summarized in Fig. 1 (columns 2 in Tables A1–A5). The ranking of candidates for hire in the teacher vacancy sees childless women come first (.72, 95% CI = .69, .75), followed by mothers (.70, 95% CI = .69, .73), fathers (.69, 95% CI = .66, .71), and childless men (.68, 95% CI = .65, .71). Differences between groups are extremely small, with the most distance between childless women and childless men (≈ 4 percentage points, $p = .030$). Further, no ranking is discernible in the second panel of Fig. 1 when looking at salary offers. Hence, while we do find some support for a pro-women bias in female-typical jobs, candidates giving evidence of being a mother share smaller advantages over men than childless women do. Our status reversal hypothesis is thus only partially supported in that employer prefer childless women, and not mothers, for hire in a female-typical vacancy. This ranking with respect to hiring intentions is also inconsequential for salary offers.

To make sense of hiring intentions we then look at ability and commitment ratings, as well as at expected working hours in the third to fifth panel of Fig. 1. While we reach no conclusion on ability rankings, we find a clearer pattern similar to the one highlighted for the likelihood of hire for commitment ratings. These are slightly skewed in favour of childless women, but again differences are very small. As for expected working hours, mothers are expected to work the least amount of hours (33.7, 95% CI = 33.1, 34.3) compared to both childless women and men. Overall this clear-cut expectation has little consequences for employers' hiring intentions in our female-typical job vacancy.

Turning to the software engineer job vacancy in Fig. 2 (see column 2 in Tables A6–A10), we surprisingly find the same ranking of candidates in terms of likelihood of hire, with childless women topping employers' preferences, followed by mothers and men. Differences between candidates are even smaller in size, as compared to the teacher experiment, and none of them reaches statistical significance at conventional levels. Also similar to our female-typical vacancy, we cannot detect any gradient by sex of the candidate and parental status in terms of salary offers or ability.

For our measures of commitment, mothers receive the lowest ratings (4.53, 95% CI = 4.36, 4.68) albeit difference across candidates are

once again very small. Somewhat differently from the teacher experiment, employers rank childless men at the top of expected hours, followed by childless women and fathers, and finally mothers (34, 95% CI = 33.5, 34.6). The distance between childless men and mothers in terms of expected working hours is of around 3.4 h (per Table A10, $-1.1 + (-1.6) + (-.7) = -3.4$, $p < .001$). Yet such sizeable difference in terms of expected working hours does not result in differential hiring intentions. Taken together, findings from both experiments fit with neither economic nor sociopsychological models of discrimination. Contrary to our *family-risk* hypothesis, core to the statistical discrimination scenario, women are not disadvantaged irrespective of parental status. Importantly, this lack of pro-male bias is more evident in male-typical jobs where our *occupation-risk* hypothesis would have predicted harsher statistical discrimination against women. As for status-characteristic theory, we find little support for a *motherhood penalty*, if not in terms of expected working hours. Also, if *status beliefs* are at work, they do so similarly across our male- and female-typical jobs. We cannot discern a status ranking in terms of ability, but only in terms of expected commitment, and differences are overall negligible in size.

In line with our last, *contract-risk* hypothesis, we investigate heterogeneity by work contract. This may unveil patterns coherent with statistical discrimination, which we expected to be more prominent when permanent contracts are offered in male-typical jobs. We examine this prediction in Figs. 3 and 4 (corresponding to columns 3 and 4 of Tables A1–A10). Comparing coefficients across models separate by type of contract, we find little evidence of heterogeneous effects. If anything, for our male-typical job vacancy in particular (Fig. 4 and Table A6), female candidates actually score higher on the likelihood of hire when a permanent post is offered, the opposite of our expectation under statistical discrimination.

5. Discussion and conclusions

Confronting fictitious and identical CVs, Dutch employers in our study seem to prefer childless women to fill a primary-school teacher vacancy, but differences are very small. We cannot detect discrimination in average salary offers, both for a teacher post and a software engineer post. Female candidates that give evidence of being a mother are expected to be less committed to their job than their childless counterparts, more than fathers as compared to childless men, but again this disadvantage is small in size. Clear-cut and sizeable effects are only

Software engineer

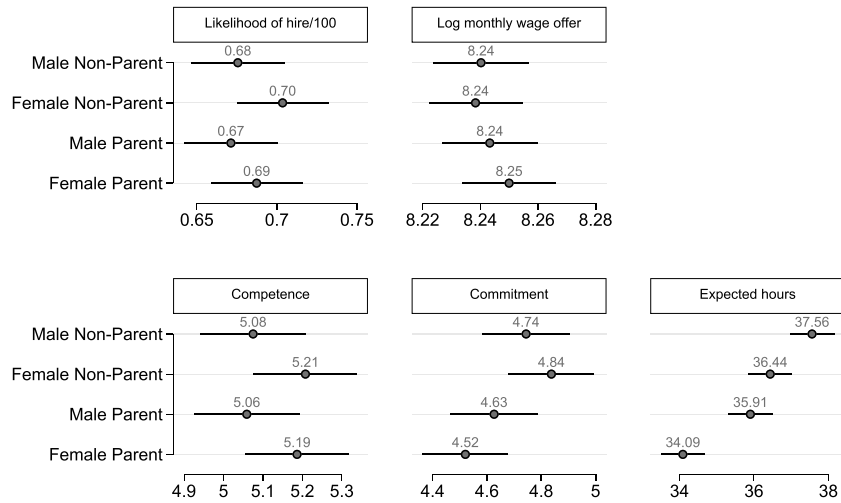


Fig. 2. Linear predictions and 95% confidence intervals. Estimate refer to multilevel linear models for hiring intentions and candidate ratings for a software engineer vacancy.

detected for expected working hours, with mothers expected to work the lowest amount of weekly hours. Notably then, this expectation is inconsequential for hiring intentions in our experimental setting.

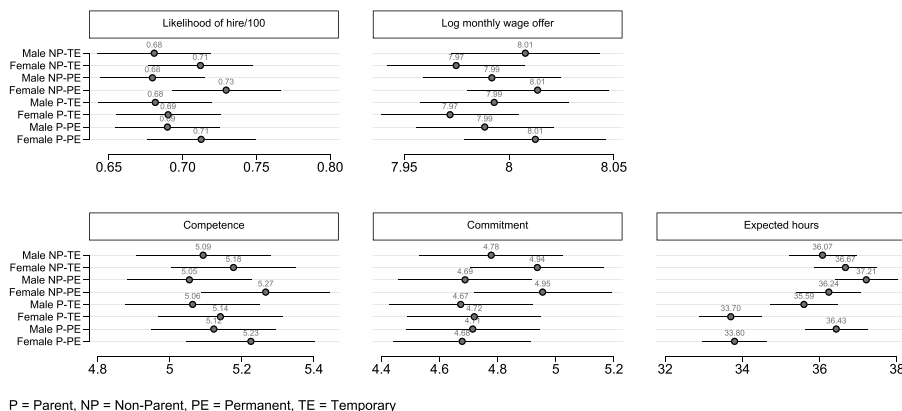
Our conclusions are, first of all, sample-dependent. We were able to draw on a sample of Dutch individuals, participating in a pre-existing online panel survey and with HR responsibilities in their current job. The latter feature in particular enhances the external validity of our study, improving on previous research typically relying on undergraduates in controlled experimental settings (Correll et al., 2007; Koch et al., 2015). Since our employers came from all sectors of the economy, however, they may not have had specialised knowledge of the two occupation-specific job markets of primary school teacher and software engineer. As such, caution should be applied when generalising our findings to Dutch employers in these specific lines of work. Yet, previous studies could not detect a preference for male over female candidates for the position of software engineer, even when involving Dutch employers in the IT sector (Di Stasio & van de Werfhorst, 2016). For the teacher vacancy, our findings are in line with those of two field experiments ran in Sweden for similar job titles (Bygren et al., 2017; Carlsson, 2011). Effect sizes are also in line with those found investigating gender-by-parental-status interactions in previous studies for other European countries (e.g. Bygren et al., 2017; Oesch et al., 2017).

The small magnitude of these effects stands out, particularly in comparison with that highlighted for the US context (Correll et al., 2007).

External validity concerns may also be raised, more specifically, with our manipulation of risk. Being a simulation of the hiring process, employers do not bear any actual hiring cost and ours could be considered a litmus test for what one could expect then in the field. Null findings concerning the manipulation of work contract, nonetheless, may reflect our focus on hiring intentions rather than actual hiring decisions. Also, being part of a three-way interaction in our factorial design, we might have not been able to detect differential treatment by type of contract (if any) in our setting due to the lack of statistical power. Large-sample replications and field replications in particular (e.g. Baert et al., 2016; Petit, 2007) could further probe heterogeneity in employer discrimination with respect to work contract (see also Becker et al., 2019).

We also acknowledge that our design grants, in principle, more statistical power to the detection of the effects of parental status, given that it is manipulated within subjects and the correlation of repeated measurements within subjects is positive and substantial (see the Intraclass Correlations reported in the full models; e.g. Maxwell & Delaney, 2004: 561). High-powered replications might thus prove particularly useful to further investigate our between-subject factor, i.e.

Primary-school teacher



P = Parent, NP = Non-Parent, PE = Permanent, TE = Temporary

Fig. 3. Linear predictions and 95% confidence intervals. Estimate refer to multilevel linear models, separate by work contract, for hiring intentions and candidate ratings for a primary-school teacher vacancy.

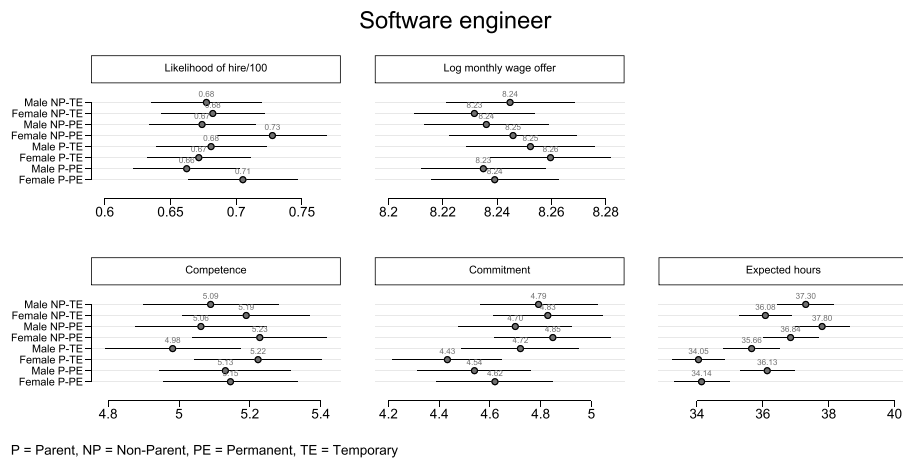


Fig. 4. Linear predictions and 95% confidence intervals. Estimate refer to multilevel linear models, separate by work contract, for hiring intentions and candidate ratings for a software engineer vacancy.

differential treatment of women and men regardless of parental status. In this latter regard, it is also the case that, by letting sex of the candidate vary only between subjects, respondents could not explicitly compare male and female candidates. We cannot assess therefore if respondents have different standards or scales with which they evaluate men and women, as they only evaluate one or the other. While between-subjects manipulation may shield from social desirability bias, it also limits the external validity of our experimental setting with respect to real-world hiring. In the market, employers typically compare male and female candidates – at least if, and to the extent that, both apply for a given job vacancy.

Finally, the (internal) validity of our findings is also bounded to the successful manipulation of our treatments. Once employers finished rating CVs, we asked them *a*) whether the candidates they just evaluated were both female or both male, *b*) whether one of the candidates had a kid and the other did not, and *c*) whether the vacancy they were required to fill was permanent or temporary. While more than 70% of employers correctly recalled features *a*) and *b*) of job candidates in both experiments, only around 47% correctly recalled the type of work contract being offered in the job ad. While many factors can bias recall, our estimates should nonetheless be regarded as intention-to-treat.

Overall we find little evidence of a motherhood penalty in hiring in two sex-typical jobs in the Netherlands, in line with what field experiments and vignette studies have shown in some

European countries (Bygren et al., 2017; Petit, 2007), but not others (Hipp, 2019; José González et al., 2019; Oesch et al., 2017). In our experiments, this holds even if Dutch employers expect mothers to be less committed to their job and work fewer hours a week than men and women without kids. This may speak to the relevance of context for hiring intentions (e.g. Bills et al., 2017). In the Netherlands, women in part-time jobs have similar chances to receive firm-sponsored training (Picchio & van Ours, 2016) and experience comparatively small wage penalties (Fouarge & Muffels, 2009) vis-a-vis their full-time counterparts. The spread of part-time work arrangements (Bosch et al., 2010) and the weakening of the ‘work obligation’ norm (Wielers & Raven, 2013), one giving work the priority over all other spheres of life, may also explain why Dutch employers’ expectations regarding a candidate’s commitment and working hours have small or nil spill-over effects on hiring intentions. Future research could shed light on whether employers hold similar expectations, but to much greater detriment for the hiring of female job candidates, in contexts where women (mothers) are expected to work part-time yet part-time jobs are of much poorer quality career-wise.

On balance, results in this study fail to meet our predictions under both statistical and status-based models of discrimination. Admittedly, we followed the predictions of one theory of statistical discrimination, namely

that based on expected variance, while others (for a review, Fang & Moro, 2011) were beyond the remit of this experiment. Nonetheless, differently from previous studies, we consider both economic and sociopsychological theories while examining gender-by-parental-status hiring discrimination (cfr. Correll et al., 2007). We particularly focused on the screening of CVs: most of our employers carried out real-world screenings in their jobs and such stage of recruitment arguably presents a favourable opportunity structure for discrimination (e.g. Petersen & Saporta, 2004). Yet, we cannot assess here whether discrimination, either of the statistical or of the status-based kind, may occur later in the hiring process or whether and how it may affect career progression. The current debate could thus benefit from accumulating tests of these two theories in other contexts, or for different job titles, but also considering more fully the stages in which employers reward or penalise job candidates and current employees. Our findings are also mute on whether motives coherent with both theories might be at work and, perhaps, cancel each other out. After all, both theories commonly predict that employers are driven by stereotypical beliefs about groups. Future research could aim at integrating and/or reformulating these two theories to better grasp hiring discrimination, or lack thereof, in sex-typical jobs.

Although not the main focus in our survey experiments, investigating gender discrimination in hiring in sex-typical jobs may also highlight whether or not employers contribute to occupational segregation in labour markets. In this respect, Dutch employers’ preference for female candidates *all else equal* for a primary-school teacher vacancy is consistent with previous experimental research on hiring discrimination in female-typical jobs (Booth & Leigh, 2010; Carlsson, 2011; Rich, 2014). Yet, the presence of, if anything, very small effects may rather strengthen the stance of those pointing to persistent gender segregation by field of study (Barone, 2011; van de Werfhorst, 2017) as the main source of occupational segregation.

Acknowledgements

In this paper we make use of data of the LISS (Longitudinal Internet Studies for the Social sciences) panel administered by CentERdata (Tilburg University, The Netherlands). The LISS panel data were collected by CentERdata (Tilburg University, The Netherlands) through its MESS project funded by the Netherlands Organization for Scientific Research. Funding for this research project was provided by Tilburg University, School of Social and Behavioral Sciences.

The authors wish to thank Marije Oudejans for managing the data collection process. We also wish to thank participants at the 2017 ESRA Conference in Lisbon, as well as in seminars at Tilburg University and at the University of Trento for their comments and suggestions. All errors or omissions remain the authors’ own.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:<https://doi.org/10.1016/j.rssm.2019.100464>.

References

- Aigner, D. J., & Cain, G. G. (1977). Statistical theories of discrimination in labor markets. *Industrial & Labor Relations Review*, 30(2), 175–187.
- Akgündüz, Y. E., & van Huizen, T. (2015). Training in two-tier labor markets: The role of job match quality. *Social Science Research*, 52, 508–521.
- Altonji, J. G., & Pierret, C. R. (2001). Employer learning and statistical discrimination. *The Quarterly Journal of Economics*, 116(1), 313–350.
- Arrow, K. J. (1973). The theory of discrimination. In O. Ashenfelter, & A. Rees (Eds.), *Discrimination in labor markets* (pp. 3–33). Princeton, NJ: Princeton University Press.
- Azmat, G., & Petrongolo, B. (2014). Gender and the labor market: What have we learned from field and lab experiments? *Labour Economics*, 30, 32–40.
- Baert, S. (2014). Career lesbians. Getting hired for not having kids? *Industrial Relations Journal*, 45(6), 543–561.
- Baert, S., De Pauw, A.-S., & Deschacht, N. (2016). Do employer preferences contribute to sticky floors? *ILR Review*, 69(3), 714–736.
- Barone, C. (2011). Some things never change: Gender segregation in higher education across eight nations and three decades. *Sociology of Education*, 84(2), 157–176.
- Becker, S. O., Fernandes, A., & Weichselbaumer, D. (2019). Discrimination in hiring based on potential and realized fertility: Evidence from a large-scale field experiment. *Labour Economics*. <https://doi.org/10.1016/j.labeco.2019.04.009>.
- Berger, J., Berger, J., Fiske, M. H., & Norman, R. Z. (1977). *Status characteristics and social interaction: An expectation-states approach*. New York: Elsevier.
- Bettio, F., & Verashchagina, A. (2009). *Gender segregation in the labour market: Root causes, implications and policy responses in the EU*. Luxembourg: Publications Office of the European Union.
- Biewen, M., & Seifert, S. (2016). *Potential parenthood and career progression of men and women: A simultaneous hazards approach*. IZA Discussion Paper No.10050.
- Bills, D. B., Di Stasio, V., & Gërkhani, K. (2017). The demand side of hiring: Employers in the labor market. *Annual Review of Sociology*, 43, 291–310.
- Blommaert, L., Coenders, M., & Van Tubergen, F. (2014). Discrimination of Arabic-named applicants in the Netherlands: An internet-based field experiment examining different phases in online recruitment procedures. *Social Forces*, 92(3), 957–982.
- Booth, A., & Leigh, A. (2010). Do employers discriminate by gender? A field experiment in female-dominated occupations. *Economics Letters*, 107(2), 236–238.
- Bosch, N., Deelen, A., & Euwals, R. (2010). S part-time employment Here to stay? Working hours of dutch women over successive generations. *Labour*, 24(1), 35–54.
- Bygren, M., Erlandsson, A., & Gähler, M. (2017). Do employers prefer fathers? Evidence from a field experiment testing the gender by parenthood interaction effect on call-backs to job applications. *European Sociological Review*, 33(3), 337–348.
- Carlsson, M. (2011). Does hiring discrimination cause gender segregation in the Swedish labor market? *Feminist Economics*, 17(3), 71–102.
- Carlsson, M., & Eriksson, S. (2017). *The effect of age and gender on labor demand – Evidence from a field experiment*. IFAU Working Paper (2017:8).
- Charles, K. K., & Guryan, J. (2011). Studying discrimination: Fundamental challenges and recent progress. *Annual Review of Economics*, 3(1), 479–511.
- Correll, S. J., & Benard, S. (2006). Biased estimators? Comparing status and statistical theories of gender discrimination. *Social Psychology of the Workplace: Advances in Group Processes*, 23, 89–116.
- Correll, S. J., Benard, S., & Paik, I. (2007). Getting a Job: Is There a Motherhood Penalty? *The American Journal of Sociology*, 112(5), 1297–1339.
- Cuddy, A. J., Fiske, S. T., & Glick, P. (2004). When professionals become mothers, warmth doesn't cut the ice. *The Journal of Social Issues*, 60(4), 701–718.
- Davies, R., & Pierre, G. (2005). The family gap in pay in Europe: A cross-country study. *Labour Economics*, 12(4), 469–486.
- De Hoon, S., Keizer, R., & Dykstra, P. (2017). The influence of motherhood on income: Do partner characteristics and parity matter? *Community, Work & Family*, 20(2), 211–225.
- Di Stasio, V., & G'èrxhani, K. (2015). Employers' social contacts and their hiring behavior in a factorial survey. *Social Science Research*, 51, 93–107.
- Di Stasio, V., & van de Werfhorst, H. G. (2016). Why does education matter to employers in different institutional contexts? A vignette study in England and the Netherlands. *Social Forces*, 95(1), 77–106.
- Ellemers, N. (2018). Gender stereotypes. *Annual Review of Psychology*, 69, 275–298.
- Fang, H., & Moro, A. (2011). Theories of statistical discrimination and affirmative action: a survey. In J. Benhabib, M. O. Jackson, & A. Bisin (Vol. Eds.), *Handbook of social economics: Volume 1A*. North Holland: Elsevier.
- Fernandez, R. M., & Campero, S. (2017). Gender sorting and the glass ceiling in high-tech firms. *ILR Review*, 70(1), 73–104.
- Fiske, S. T., Cuddy, A. J., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, 82(6), 878.
- Fouarge, D., & Muffels, R. (2009). Working part-time in the British, German and Dutch labour market: Scarring for the wage career? *Schmollers Jahrbuch*, 129(2), 217–226.
- Fouarge, D., de Grip, A., Smits, W., & de Vries, R. (2012). Flexible contracts and human capital investments. *De Economist*, 160(2), 177–195.
- Fransen, E., Plantenga, J., & Vlasblom, J. D. (2012). Why do women still earn less than men? Decomposing the Dutch gender pay gap, 1996–2006. *Applied Economics*, 44(33), 4343–4354.
- Gangl, M., & Ziefle, A. (2009). Motherhood, labor force behavior, and women's careers: An empirical assessment of the wage penalty for motherhood in Britain, Germany, and the United States. *Demography*, 46(2), 341–369.
- Gathmann, C., & Schönberg, U. (2010). How general is human capital? A task-based approach. *Journal of Labor Economics*, 28(1), 1–49.
- Gupta, N. D., & Smith, N. (2002). Children and career interruptions: The family gap in Denmark. *Economica*, 69(276), 609–629.
- Heilman, M. E., & Okimoto, T. G. (2008). Motherhood: A potential source of bias in employment decisions. *The Journal of Applied Psychology*, 93(1), 189.
- Hipp, L. (2019). Do hiring practices penalize women and benefit men for having children? Experimental evidence from Germany. *European Sociological Review*. <https://doi.org/10.1093/esr/jcz056>.
- Jessen, J., Jessen, R., & Kluge, J. (2019). Punishing potential mothers? Evidence for statistical employer discrimination from a natural experiment. *Labour Economics*. <https://doi.org/10.1016/j.labeco.2019.04.002>.
- José González, M., Cortina, C., & Rodríguez, J. (2019). The role of gender stereotypes in hiring: A field experiment. *European Sociological Review*, 35(2), 187–204.
- Koch, A. J., D'Mello, S. D., & Sackett, P. R. (2015). A meta-analysis of gender stereotypes and bias in experimental simulations of employment decision making. *The Journal of Applied Psychology*, 100(1), 128.
- Mark, N. P., Smith-Lovin, L., & Ridgeway, C. L. (2009). Why do nominal characteristics acquire status value? A minimal explanation for status construction. *The American Journal of Sociology*, 115(3), 832–862.
- Maxwell, S. E., & Delaney, H. D. (2004). *Designing experiments and analyzing data: A model comparison perspective* (2nd ed.). Routledge.
- Mooi-Reci, I., & Dekker, R. (2015). Fixed-term contracts: Short-term blessings or long-term scars? Empirical findings from the Netherlands 1980–2000. *British Journal of Industrial Relations*, 53(1), 112–135.
- Neumark, D. (2016). *Experimental research on labor market discrimination*. NBER Working Paper 22022.
- OECD (2014). *OECD employment outlook*. Paris: Oecd.
- OECD (2016). *Society at a glance 2016: OECD social indicators*. Paris: OECD.
- OECD (2017a). *Education at a glance 2017: OECD social indicators*. Paris: OECD.
- OECD (2017b). *The pursuit of gender equality: An uphill battle*. Paris: OECD Publishing.
- Oesch, D., Lipps, O., & McDonald, P. (2017). The wage penalty for motherhood: Evidence on discrimination from panel data and a survey experiment for Switzerland. *Demographic Research*, 37, 1793–1824.
- Pedulla, D. S. (2016). Penalized or protected? Gender and the consequences of non-standard and mismatched employment histories. *American Sociological Review*, 81(2), 262–289.
- Pedulla, D. S. (2018). How Race and Unemployment Shape Labor Market Opportunities: Additive, Amplified, or Muted Effects? *Social Forces*. <https://doi.org/10.1093/sf/soy002>.
- Petersen, T., & Saporta, I. (2004). The opportunity structure for discrimination. *The American Journal of Sociology*, 109(4), 852–901.
- Petersen, T., Saporta, I., & Seidel, M.-D. L. (2000). Offering a job: Meritocracy and social networks. *The American Journal of Sociology*, 106(3), 763–816.
- 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(3), 371–391.
- Phelps, E. S. (1972). The statistical theory of racism and sexism. *The American Economic Review*, 62(4), 659–661.
- Picchio, M., & van Ours, J. C. (2016). Gender and the effect of working hours on firm-sponsored training. *Journal of Economic Behavior & Organization*, 125, 192–211.
- Pinkston, J. C. (2006). A test of screening discrimination with employer learning. *ILR Review*, 59(2), 267–284.
- Protsch, P., & Solga, H. (2015). How employers use signals of cognitive and noncognitive skills at labour market entry. Insights from field experiments. *European Sociological Review*, 31(5), 521–532.
- Riach, P. A., & Rich, J. (2002). Field experiments of discrimination in the market place. *The Economic Journal*, 112(483), F480–F518.
- 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(2).
- Rich, J. (2014). *What do field experiments of discrimination in markets tell us? A meta analysis of studies conducted since 2000*. IZA Discussion paper No.8584.
- Ridgeway, C. L. (2011). *Framed by gender: How gender inequality persists in the modern world*. Oxford: Oxford University Press.
- Ridgeway, C. L., & Correll, S. J. (2004). Motherhood as a status characteristic. *The Journal of Social Issues*, 60(4), 683–700.
- Ridgeway, C. L., & Correll, S. J. (2006). Consensus and the creation of status beliefs. *Social Forces*, 85(1), 431–453.
- Ridgeway, C. L., Backer, K., Li, Y. E., Tinkler, J. E., & Erickson, K. G. (2009). How easily does a social difference become a status distinction? Gender matters. *American Sociological Review*, 74(1), 44–62.
- Rivera, L. A., & Tilesik, A. (2016). Class advantage, commitment penalty: The gendered effect of social class signals in an elite labor market. *American Sociological Review*, 81(6), 1097–1131.
- Scherpenzeel, A. (2011). Data collection in a probability-based internet panel: How the LISS panel was built and how it can be used. *Bulletin of Sociological Methodology/Bulletin de Méthodologie Sociologique*, 109(1), 56–61.
- Thébaud, S. (2015). Status beliefs and the spirit of capitalism: Accounting for gender biases in entrepreneurship and innovation. *Social Forces*, 94(1), 61–86.
- van de Werfhorst, H. G. (2017). Gender segregation across fields of study in post-secondary education: Trends and social differentials. *European Sociological Review*, 33(3), 449–464.
- Wielers, R., & Raven, D. (2013). Part-time work and work norms in the Netherlands. *European Sociological Review*, 29(1), 105–113.
- Yip, C. M., & Wong, R. S.-K. (2014). Gender-oriented statistical discrimination theory: Empirical evidence from the Hong Kong labor market. *Research in Social Stratification and Mobility*, 37, 43–59.