European Sociological Review, 2021, 1–18
doi: 10.1093/esr/jcab043
Original Article

Gender Discrimination in Hiring: Evidence from a Cross-National Harmonized Field Experiment

Gunn Elisabeth Birkelund [1,*], Bram Lancee[2], Edvard Nergård Larsen[1], Javier G. Polavieja[3], Jonas Radl[3,4] and Ruta Yemane[4]

[1]Department of Sociology and Human Geography, University of Oslo, Oslo, Norway, [2]Department of Sociology', University of Amsterdam, Amsterdam, The Netherlands, [3]Department of Social Sciences, Universidad Carlos III de Madrid, Getafe, Spain and [4]Research Unit Migration, Integration, Transnationalization, WZB Social Science Center Berlin, Berlin, Germany

*Corresponding author: Email: g.e.birkelund@sosgeo.uio.no

Submitted October 2020; revised August 2021; accepted September 2021

Abstract

Gender discrimination is often regarded as an important driver of women's disadvantage in the labour market, yet earlier studies show mixed results. However, because different studies employ different research designs, the estimates of discrimination cannot be compared across countries. By utilizing data from the first harmonized comparative field experiment on gender discrimination in hiring in six countries, we can directly compare employers' callbacks to fictitious male and female applicants. The countries included vary in a number of key institutional, economic, and cultural dimensions, yet we found no sign of discrimination against women. This cross-national finding constitutes an important and robust piece of evidence. Second, we found discrimination against men in Germany, the Netherlands, Spain, and the UK, and no discrimination against men in Norway and the United States. However, in the pooled data the gender gradient hardly differs across countries. Our findings suggest that although employers operate in quite different institutional contexts, they regard female applicants as more suitable for jobs in female-dominated occupations, *ceteris paribus*, while we find no evidence that they regard male applicants as more suitable anywhere.

Introduction

Women have traditionally been disadvantaged in the labour market, and much scholarship has documented patterns of and trends in gender inequalities (e.g. Weichselbaumer and Winter-Ebmer, 2005; Carlsson, 2011). However, women's and men's working lives have changed considerably since the mid-20th century (Goldin, 2014). In nearly all OECD countries, women now have higher educational attainment than men (OECD, 2015). In many countries, women comprise more than 40 per cent of the labour force (Pew Research Center, 2017), and, although the process is slow, there is some evidence that the gender gap in earnings is converging (Jacobsen, Khamis and Yuksel, 2015; Blau and Kahn, 2017; Neumark, 2018). People's attitudes have also changed; in particular, we have seen decreasing support for traditional gender norms and increasing support for women's employment (Fernández, 2013).

All trends towards equalization notwithstanding, gender inequalities in the labour market still exist.

Broadly construed, there are two explanations for why this is the case. First, women are treated differently from men within the same jobs, and second, women and men are sorted into different jobs, with lower earnings and fewer promotion prospects in typically female-dominated jobs. Studies have, however, shown that when men and women work in the same jobs in the same firms, gender differences in earnings are significantly diminished or even eradicated (e.g. Petersen and Morgan, 1995). This gives more credibility to the sorting explanation. Indeed, we know that occupational sex segregation is widespread (Chang, 2004), and that men and women work in jobs with unequal compensation (Levanon and Grusky, 2016). Scholars have therefore argued for the exigency to better understand the sorting process of men and women into different jobs (Petersen and Saporta, 2004). We can think of two competing explanations. First, the supply side argument addresses educational and occupational choices: men and women choose different occupations and therefore apply for different jobs. Alternatively, men and women apply for the same jobs, but women are discriminated against when they apply for jobs with higher earnings, more responsibilities, etc. This demand side argument is related to employers' hiring decisions, and this study aims to make a contribution to the literature by testing the discrimination explanation.

Hiring processes are contingent on employers' decision-making, and crucial elements of their decisions usually remain opaque to researchers. Thus, measuring discrimination is difficult. Supply-side data can reveal gender gaps in labour market outcomes, but we can never rule out the possibility that observed gender gaps are driven by unobserved factors pertaining to the supply side rather than by employers' discriminatory practices on the demand side. Therefore, experimental designs are more suitable for detecting discrimination (Azmat and Petrongolo, 2014; Gaddis, 2018). While a weakness of laboratory experiments is external validity, field experiments can, through manipulation of one (or more) treatment variable(s), e.g. the applicant's gender, provide real-world causal estimates of treatment effects on employers' hiring decisions.

Previous Research

Social scientists have conducted randomized field experiments to detect hiring discrimination since the 1970s (Riach and Rich, 2002). Perhaps surprisingly, previous studies on hiring discrimination of male and female job applications show very mixed findings. Table 1 gives an overview of the most relevant field experiments on gender discrimination in hiring, and we comment on the most important findings below.

Some experiments found advantages for men over women (Neumark, Bank and Van Nort, 1996; Petit, 2007; Zhou, Zhang and Song, 2013; Duguet, Loïc and Petit, 2017; González, Cortina and Rodríguez, 2019), whereas other experiments found advantages for women over men (Jackson, 2009; Carlsson, 2011; Carlsson and Eriksson, 2017). Some studies found hiring discrimination against both men and women, depending on parental status (Correll, Benard and Paik, 2007) or gender composition and type of job (Weichselbaumer, 2004; Yavorsky, 2019), while other studies found no gender discrimination at all (Albert, Escot and Fernández-Cornejo, 2011; Capéan *et al.*, 2012; Carlsson *et al.*, 2014; Carlsson and Erikson, 2017; Bygren, Erlandsson and Gähler, 2017). Some studies found evidence of hiring discrimination against women in high-level jobs (Riach and Rich, 2002; Baert, De Pauw and Deschacht, 2016), while others did not (Williams and Ceci, 2015). These inconsistencies in findings might reflect true cross-national differences in gender discrimination. If institutional contexts, such as labour market policies, affect employers' hiring decisions, they might, all else equal, behave differently in different national contexts (Gangl and Ziefle, 2009). However, as these experiments are adapted to national contexts, and the included occupations vary considerably, inconsistencies in findings might also be an artefact of heterogeneity of research designs.

More consistently across contexts, field experiments on gender discrimination show that men are discriminated when they apply for female occupations, and women when they apply for male occupations (Riach and Rich, 2002, 2006; Booth and Leigh, 2010; Carlsson, 2011; Rich, 2014). 'However, discrimination against men in "female" occupations was always much higher than that against women in "male" occupations' (Riach and Rich, 2002: pp. F504–505). One study also found discrimination of men in female-dominated occupations, and no gender differences in hiring in mixed or male-dominated occupations (Ahmed, Granberg and Khanna, 2021). Thus, despite the obvious temptation, we cannot directly compare field-experimental evidence on gender discrimination across countries, due to heterogeneity in research designs across countries and time-periods.

To address this limitation, we make use of a harmonized cross-national field experiment in six countries: Germany, the Netherlands, Norway, Spain, the United Kingdom, and the United States [The Growth, Equal Opportunities, Migration and Markets (GEMM) study, conducted by Lancee *et al.*, 2019b].[1] To our knowledge,

Table 1. Previous field experiments on gender discrimination in hiring

Authors	Applicant ages	Country	No. of occupations	Blue/white collar	Qualifications	Occupations
Ahmed, Granberg and Khanna (2021)	28	Sweden	15	BW	Lo-Med-Hi	Store clerk, vehicle mechanic, cleaner, enrolled nurse, waitstaff, chef, truck/delivery driver, warehouse worker, preschool teacher, IT developer, B2B sales, accounting clerk, customer service, telemarketing, childcare
Albert, Escot and Fernández-Cornejo (2011)	24; 28; 38	Spain	6	W	Med-Hi	Sales representatives, marketing technicians, accountant's assistants, accountants, administrative assistants/receptionists, executive secretaries
Baert (2015); Baert, De Pauw and Deschacht (2016)	NA	Belgium	2	W	High	Business administration for BA and business economics for MA
Berson (2012)	20	France	1	W	Low	Cashier works in retail stores
Booth and Leigh (2010)	NA	Australia	4	W	Low	Waitstaff, data-entry, customer service, sales
Bygren, Erlandsson and Gähler (2017); Brandén, Bygren and Gähler (2018)	31	Sweden	18	W	Med-Hi	Accountant/auditor, assistant nurse, chef, cleaner, elementary school teacher, computer specialist, engineer, financial assistant, high school teacher, nurse, preschool teacher, receptionist, salesperson, store personnel, or cashier
Capéau *et al.* (2012)	23; 35; 47; 53	Belgium	12	BW	Lo-Med-Hi	Industry and manufacturing; commerce, transport, and catering; communication, administration, and financial services; public sector, health care, non-profit, and other services
Carlsson *et al.* (2014)	NA	Sweden	11	BW	Lo-Med-Hi	Cleaners, restaurant workers, accountants, nurses, primary school teachers, shop sales assistants, high school teachers, business sales assistants, construction workers, motor-vehicle drivers, and computer professionals
Carlsson and Eriksson (2017)	35–70	Sweden	7	BW	Low/Medium	Administrative assistants, chefs, cleaners, food serving and waitstaff, retail sales persons and cashiers, sales representatives, truck drivers
Carlsson (2011)	24–29	Sweden	13	BW	Lo-Med-Hi	Construction, motor-vehicle drivers, nurses, secondary school teachers (math, science, language), shop sales assistants, computer professionals, preschool teachers, business sales assistants, cleaners, accountants, restaurant workers
Duguet *et al.* (2012)	25	France	1	W	High	Software developers
Duguet, Loïc and Petit (2017)	23–24	France	3	B	Medium	Construction (masonry, plumbing, and electricity)
González, Cortina and Rodríguez (2019)	37–39	Spain	18	BW	Lo-Med-Hi	Delivery, waitstaff, sales clerks, computer technician, estate agents, office clerks, industrial engineers, tax advisors, physiotherapists, foremen/women, head chefs, store managers, heads of logistics, warehouse managers, supervising clerks, marketing directors, senior lawyers, senior nurses

(continued)

Table 1. (Continued)

Authors	Applicant ages	Country	No. of occupations	Blue/white collar	Qualifications	Occupations
Jackson (2009)	NA	UK	NA	W	High	Professional and managerial positions
Neumark, Bank and Van Nort (1996)	NA	US	1	W	Low	Waitstaff
Petit (2007)	25; 37	France	12	W	Lo-Med-Hi	Administrative technician, administrative clerk, accounting clerk, executive manager, portfolio manager, recovery manager, accounting manager; receptionist, counter clerk, customer consultant, sales manager, customer assistant
Riach and Rich (1987)	NA	Australia	7	BW	Med-Hi	Computer analyst programmer, computer operator, computer programmer, gardener, industrial relations officer, management accountant, payroll clerk
Riach and Rich (2006)	NA	UK	4	W	Med-Hi	Computer analyst, electrical and mechanical engineer, secretary, trainee chartered accountant
Rivera and Tilcsik (2016)	NA	US	1	W	High	Summer associate positions of large law firms (interpreted as quasi full-time job offer due to sectoral characteristics of summer associate positions as job entry into the law sector)
Weichselbaumer (2004)	NA	Austria	4	W	Med-Hi	Network technicians, computer programmers, accountants, secretaries
Williams and Ceci (2015)	NA	US	1	W	High	Tenure-track assistant professorships
Yavorsky (2019)	NA	US	8	BW	Lo-Med-Hi	Administrative support, human resource associate, financial analyst, sales representative; housekeeping, customer service, manufacturing, maintenance/janitor
Zhou, Zhang and Song (2013)	25; 28	China	4	W	Med-Hi	Engineers, accountants, secretaries, and marketing professionals

Note: B = blue collar; W = white collar.
Source: own elaboration.

the GEMM study is the first randomized field experiment with a deliberate cross-national comparative design (Di Stasio and Lancee, 2019). These data allow us to provide new and rigorous evidence on gender discrimination in the first phase of the hiring process in six occupations in six countries. We contribute to the literature by analysing hiring discrimination within and across countries with different institutional characteristics.

Gender Discrimination: Theoretical Considerations

Hiring new employees always involves an element of risk-taking, as employers cannot know beforehand how an individual will perform. Employers rely on the information available in the cover letter and CV but may still be uncertain about the applicants' skills. If employers believe members of a particular group are more productive than others, they might regard group membership as an informative cue. Obviously, employers' expectations might be wrong, as they may rely on unfounded stereotypes about certain groups. In addition, even if employers' beliefs are correct in terms of average group-level characteristics, individual job applicants may deviate substantially from a given group characteristic.[2]

Discrimination against Women

Several perspectives explain why employers discriminate against women. We have grouped the relevant theoretical approaches into two broader categories: (i) cultural perspectives focusing on social norms and gender stereotypes, and (ii) the economic-rational perspective addressing statistical discrimination.

According to cultural perspectives, employers rely on gender stereotypes and gender-differentiated work expectations. In Joan Acker's seminal work on gendered organizations, gender inequality is an inbuilt characteristic of work organizations (Acker, 1990; Rudman and Phelan, 2008; Williams, Muller and Kilanski, 2012). Of particular importance is the norm of the 'ideal worker', working full-time without family obligations. As women's work traditionally has been confined to the domestic sphere, this norm would disadvantage women in hiring situations (Acker, 1990). Even in large, modern organizations, there is evidence that women are held to other standards than men, which might explain the persistence of the glass ceiling in career promotion. The so-called 'paradox of meritocracy' (Castilla and Benard, 2010) implies that top-down directives oriented towards fairness and efficiency seem incapable of neutralizing discriminatory gender attitudes and may even reinforce the adverse effects of unconscious bias. Thus, despite societal trends towards gender convergence, theories about gendered organizations lead us to expect that men have an advantage over women in virtually all hiring processes.

The theory of statistical discrimination builds on the assumption that employers engage in cost-benefit calculations (Arrow, 1972; Phelps, 1972). This economic-rational perspective leads us to expect that employers assess the potential productivity of job applicants by their observable characteristics, such as human capital, and attribute average group characteristics to them to assess their unobservable characteristics (Fang and Moro, 2011). Due to productivity gains and because hiring in itself is costly, employers can be expected to be looking for stable workers. Given that women are more likely to be absent due to family responsibilities, employers would assess men's productivity higher and discriminate against women, all else equal.

To summarize, both cultural and economic-rational perspectives lead us to expect discrimination of female applicants, primarily due to employers' beliefs about women's higher level of absence associated with childcare.

Discrimination against Men and Women

As noted above, previous experiments show differential gender discrimination across male- and female-dominated occupations. The cultural perspectives might explain why. Psychologists have developed the stereotype content model, which proposes that people tend to perceive men as competent but not warm, and women as warm but not competent (Glick and Fiske, 1996). People also perceive male-dominated jobs as requiring more competence and female-dominated jobs as requiring more warmth (Cuddy, Fiske and Glick, 2008). As these stereotypes are associated both with individuals and jobs, it is highly plausible that employers discriminate applicants with the 'wrong' gender (Bobbitt-Zeher, 2011). Thus, 'if a caregiving job is thought to require warmth and men are thought to not possess much warmth, individuals may expect that a man will not be successful at a caregiving job' (Halper, Cowgill and Rios, 2019: p. 2). By the same logic, employers would form negative performance expectations of women in—for instance—technical jobs. Thus, employers' gender stereotypes might steer the process of matching jobs and job applicants. Theoretically, this argument is captured by the concept of sex typing of jobs (Bielby and Baron, 1986; Glick, Zion and Nelson, 1988; Reskin and Roos, 1990), the role congruency model (Cejka and Eagly, 1999), and the theory of gender categorization within work organizations (Ridgeway, 1997).

The theory on statistical discrimination can also explain differential gender discrimination across male- and female-dominated occupations. As noted, most employers are looking for stable employees, and studies have documented that workers' employment duration is sensitive to the sex typing of the job, so that women who enter a male-dominated occupation and men who enter a female-dominated occupation have disproportionately higher exit risks (Torre, 2014, 2018). Employers might be aware of this association and act accordingly. On closer inspection therefore, the differences between the cultural and the economic-rational perspectives are rather subtle, as both perspectives are compatible with the assumption that gender stereotypes are exogenously given and that employers are looking for the best match between an applicant and a job.[3] Both perspectives, therefore, lead us to expect discrimination against the minority sex in sex-typed jobs and to expect to find no prevalence of discrimination in gender-balanced jobs, *ceteris paribus*. The norm of the 'ideal worker', however, leads us to the generic expectation that women are discriminated against, independently of the sex typing of the job.

Theories on discrimination are primarily concerned with individual-level explanations, largely ignoring the role of country-level institutional contexts (Reskin, 2000). However, the 'opportunity structure for discrimination' (Petersen and Saporta, 2004) is likely to differ by macro-level factors, which we explain below.

Selection of Countries

The GEMM study is a fully harmonized field experiment on job hiring across six advanced economies that differ in a number of relevant macro-level characteristics. Because the number of policy and institutional characteristics varying across these countries is larger than the number of countries analysed and because these characteristics are highly endogenous, it is not possible to identify the effect of a single policy or institutional dimension. Our goal is therefore more modest: we want to test whether estimates of hiring discrimination of male and female applicants are robust across different policy and institutional contexts. If they are, we conclude that, despite their institutional differences, there is a common trend across these societies. If they are not, we interpret cross-national variation by considering country-specific characteristics that may affect employers' propensity to discriminate. We consider three macro dimensions: (i) general labour market regulations and conditions, (ii) family policies, and (iii) cultural norms.

First, labour market regulations can influence employers' hiring decisions by affecting the costs of job mismatch. When these costs are high, employers are likely to be more risk averse and to draw on statistical discrimination to reduce contractual hazards. If employment contracts with low termination costs are available to employers and if such contracts can be used for long time-periods, the match-or-miss pressure for employers will wane, thus reducing the impact of risk aversion on hiring decisions. The included countries differ markedly in the extent of labour market regulation (see Table 2); and we expect more gender discrimination related to the sex typing of jobs in countries with higher dismissal costs, such as Germany and the Netherlands. Another potential factor affecting the costs of discriminating is labour market tightness. If employers have a large pool of potential candidates, they are more prone to discriminate, even if only as a heuristic strategy to simplify the screening procedure (Birkelund, 2016), than when they have a restricted supply of workers (Baert, De Pauw and Deschacht, 2016). Spain is an outlier, with a high unemployment rate, which could fuel hiring discrimination.

Family policies can potentially influence employers' hiring decisions by affecting the costs associated with childbirth. Although often considered mutually complementary interventions, public support for childcare (through direct provision or subsidies) and parental leave policies actually have very different implications. Childcare support policies likely reduce the duration of post-birth work interruptions, and, because they are funded through general taxes, their costs are not borne by employers in particular. In contrast, generous maternity leave policies that establish mandatory job retention over a specified period around childbirth impose significant nonwage costs to employers, which will be greater for tasks where interruptions provoke severe human capital depreciation (Stier, Lewin-Epstein and Braun, 2001; Mandel and Semyonov, 2006; Gangl and Ziefle, 2009). The probability that employers discriminate against women should thus be greater in contexts where maternity leave arrangements are generous, such as Norway, and in contexts with less public provision of childcare, such as the United States (see Table 2).

Our countries of study also differ with respect to gender norms, which are associated with labour market and family policies (see Table 2). There is a close association between female employment rates and support for gender stereotypes (Fortin, 2005; Polavieja, 2015) and we expect more hiring discrimination of women in

Table 2. Societal factors potentially associated with gender discrimination propensities

	Dis-missal costs[a]	Duration of paid maternity leaveb	Public spending on early childhood (as per cent of GDP)[c]	Unemployment rate (2016–2018)[d]	Female-to-male employment ratio 2017[e]	Part-time employment (women)[f]	Mean age at first birth[g]	Gender egalitarian attitudes[h]	Hofstede's Masculinity Dimension[i]	Gender equality Index[j]
Germany	2.5	58	0.6	3.75%	0.91	36.6%	29.6	53.6%	66	0.778
Netherlands	2.8	42	0.6	4.89%	0.89	58.%	29.9	70.5%	14	0.737
Spain	2.0	16	0.5	17.37%	0.84	21.6%	30.9	61%	42	0.746
Norway	2.2	87	1.3	4.21%	0.96	27.7%	29.3	90.2%	8	0.83
United Kingdom	1.2	39	0.6	4.38%	0.89	36.4%	28.9	61.9%	66	0.77
United States	0.5	0	0.3	4.37%	0.86	17.2%	26.8	61.4%	62	0.718

Sources:

[a]OECD Index of regulation on individual dismissal of workers with regular contracts. 0 = very loose, 5 = very strict. The index refers to the year 2013 (OECD, 2020a).

[b]Data from the OECD for 2013. Total duration for which mothers can be on paid leave (OECD, 2020 b).

[c]Includes public spending on early childhood education and care, OECD Family Database for 2015 or latest available year (OECD, 2020c).

[d]Data from OECD for 2019 (OECD, 2019).

[e]OECD Short-Term Labor Market Statistics 2017 (OECD, 2017).

[f]Data from the OECD, referring to 2018 (OECD, 2020d).

[g]Data from OECD Family Data Base for 2015 or latest available year (OECD, 2020c).

[h]*Source*: Own calculations. 'When jobs are scarce, men should have more right to a job than women', per cent (strongly) disagree minus per cent (strongly) agree. Averages based on available data, European Values Survey 2008, 2017, as well as World Value Survey Waves 5 (2005–2009) and 6 (2011–2015).

[i]Numbers provided by Hofstede Insights, comparing countries' scores on the Masculinity Index (see Hofstede Insights, 2020).

[j]The World Economic Forum: The Global Gender Gap Report 2017. Global Gender Gap Index (The World Economic Forum, 2017).

countries with higher support for traditional gender attitudes, such as Germany. Notably, such norms go beyond mere attitudinal indicators and include sex-typical behaviours that can shape expectations (Polavieja, 2012). Relevant behaviours with a normative dimension include fertility behaviour (e.g. average age at first birth) and gender differences in employment rates and working hours that can 'inform' employers about the 'risks' of employing women (Bygren, Erlandsson and Gähler, 2017; Becker, Fernandes and Weichselbaumer, 2019). The selected countries differ in both gender attitudes and behaviours potentially affecting employers' hiring decisions.

Table 2 summarizes the indicators that characterize the countries included in the study. The list of indicators is not exhaustive, but the table illustrates the degree of variation across these countries. In accordance with the above theories, we expect the probability of observing gender discrimination in hiring to be higher in macro-level contexts where the costs of job mismatch are high due to labour-market regulation or—conditions and where traditional gender norms prevail, as expressed through attitudes and values or through gendered behaviours. These arguments, based on a small selection of the contextual measures that could have been included, are tentative. Moreover, contextual factors are only relevant if employers know about them or act upon related beliefs. Both assumptions are disputable (Birkelund *et al.*, 2019). Hence, our aim is not to identify the effect of any single dimension, which would be impossible given the small sample of countries, but to determine if our findings hold across different country contexts, and, in the event they do not, whether we can meaningfully interpret national variation by accounting for these institutional, cultural, and economic dimensions.

A Harmonized Cross-National Field Experiment

From 2016 to 2018, we sent fictitious cover letters and CVs sent to 21,318 vacant jobs advertized on national online platforms, and gathered and coded all responses from the employers (for an overview of the data, see Lancee *et al.*, 2019a,b). The experiment was primarily designed to measure hiring discrimination against immigrants and their descendants.[4] To compare their callbacks with those received by the majority population, 25 per cent of the applications in each country included a majority identity, 4,279 in total, which are the data that are used here. The fictitious job applicants, hereafter applicants, were given education levels that matched the (average) job requirements, which varied between a high school diploma to a bachelor's degree. All applicants had CVs with four years occupation-specific work experience at two different employers,[5] and we varied their age between 22 and 26 years.[6] The design is unmatched, which means that one application was sent to each vacancy. Some field experiments send two—or more—applications per vacancy, allowing the researchers to measure individual employer behaviour in addition to average employer behaviour within occupations and countries, which we measure here. Although both matched and unmatched designs have distinct advantages, the strength of the unmatched design is that one can easily implement multiple treatments. Furthermore, the risk of detection is minimal. There is also evidence that unmatched designs provide the most comparable and externally valid estimates of hiring discrimination, by avoiding potential issues of induced competition (see Vuolo, Uggen and Lageson, 2018; Lancee, 2019; Larsen, 2020 for discussions) and they minimize harm to employers by reducing their time spent in reading fictitious applications. Applications were sent to nationally advertized job vacancies within each country, which means that, although limited by occupational constraints (six occupations), the study covers national labour markets.

Occupations

The occupations included are as comparable across the six countries as possible. The selected occupations have different levels of customer contact and different educational requirements. We were looking for occupations that were available on job search platforms within each country, and for which there were sufficient numbers of vacant jobs within a time limit of maximum 2 years. To decide which occupations we should chose, we discussed a range of occupational covariates that one might not need to worry about in national studies, but which could be highly relevant in a cross-national design. We decided to exclude jobs in the public sector, which often have their own recruitment organizations. This implies that many female dominated occupations, such as nurses and teachers, are not included in our data, since they are mostly found in the public sector. We also decided to avoid occupations that often rely on informal recruitment of workers. This implies that many male-dominated occupations, such as mechanics or plumbers, are not included in our data, since they seem to rely on informal networks when they recruit new workers. Since we need the same occupations across all countries, we only need one country in which some of these considerations matter, to influence the data collection.

After these market discussions, we carefully considered the comparability of job tasks and content, and we decided to include four occupations with low or middle qualifications (cook, receptionist, store assistant, and payroll clerk), and two occupations which require education up to a bachelor's degree (software developer and sales representative). Three of these occupations have relatively little customer contact (software developer, payroll clerk, and cook), whereas the other three imply higher customer contact (sales representative, receptionist, and store assistant). The following occupations are included (ISCO codes in parentheses): Cook (512), payroll clerk (2411, 3313, 411, 412), receptionist (422), sales representative (3322), software developer (252), and store assistant (522). These occupations cover approximately 15–20 per cent of the work force within each country.

Many occupations are likely to comprise different sex-typed jobs, and the occupations included here vary in their gender profiles.[7] Supplementary Table S1 provides an overview of the gender distribution in each country within each occupational category based on national statistics the year before the field experiment took place (Lancee *et al.*, 2019b). We note that receptionists and payroll clerks are female dominated, in particular in Netherlands, Norway, and the United States, whereas software developers are clearly male dominated in all countries.

The size of the labour market differs across these countries, and as the data collection took place within a limited time, the availability of job vacancies varied. This implies that in the data, for some countries, some occupations are under-represented. For instance, Norway has a low share of receptionists (4 per cent), whereas Spain has a low share of software developers (6 per cent) and sales representatives (7 per cent). We therefore add occupational controls in all our analyses.

Treatment Variable

Gender, our main treatment variable, randomly assigned the job applications, is coded '1' for females and '0' for males.[8] The experiment also included other treatments (see Lancee *et al.*, 2019a). As these treatments are orthogonal to gender, there is no need to control for them.

Dependent Variable: Employer Response

Our main dependent variable is employer callback, which includes an invitation to an interview, an invitation to a pre-interview, and/or a request for more information. In Supplementary Information, we include analyses using only 'invitation to an interview', a stricter measurement of callback. As there are cross-national differences in the likelihood that employers ask job applicants for an interview (see Lancee *et al.*, 2019a), we prefer the broader definition of callbacks that includes an invitation for a pre-interview and/or a request for more information. A callback rate of 0.49 means that 49 per cent of the applicants received a callback. We also calculate gender ratios, dividing female by male callback rates. A gender ratio above 1 means that male applicants are discriminated, whereas a gender ratio below 1 means that female applicants are discriminated.

Estimation Strategy

To examine cross-country variation in hiring discrimination, we start by documenting callback ratios for each occupation in each country; see Table 3. We then estimate country-specific linear probability regression models; regressing callbacks on gender (see Supplementary Table S2 and Figure 1).[9] The gender coefficient provides an estimate of gender discrimination in hiring within each country, with associated standard error.

Findings

Table 3 shows the callback rates and related gender ratios by country and occupation. We first note that out of 36 possible outcomes, 23 favour *females*, as indicated by callback gender ratios > 1. This is interesting, but due to the small sample for each occupation within each country, most of these outcomes are not significant by conventional standards (see right-hand column). In Germany, we find statistically significant hiring discrimination against male applicants for receptionist and store assistant jobs, with callback ratios of 1.4 and 1.9, respectively. In the Netherlands, we find evidence of hiring discrimination against male applicants for store assistant jobs, with a callback ratio of 2.2. In Spain, we find clear evidence of hiring discrimination of males in two occupations, with callback ratios of 1.9 (payroll clerk) and 4.5 (receptionist). In the United Kingdom, we find strong evidence of hiring discrimination against males in payroll clerk jobs (callback ratio of 4.8, the highest of all). Interestingly, in the data, we find no evidence of gender discrimination in hiring in Norway or the United States. Thus, the evidence shows hiring discrimination against male, not female, job applicants in 1–3 occupations within four of the six countries.

Based on country-specific regression models, Figure 1 (and Supplementary Table S2) shows the probability of receiving a callback separately for each

Table 3. Callback ratios by country, occupation, and gender

Country	Occupation	*N* Male/Female	Callback rate Male	Callback rate Female	Callback gender ratio	*P*
Germany	Cook	66/55	0.77	0.67	0.87	0.36
Germany	Payroll clerk	61/62	0.16	0.29	1.77	0.13
Germany	Receptionist	61/66	0.57	0.79	1.37	0.01
Germany	Sales representative	49/72	0.47	0.42	0.89	0.79
Germany	Software developer	58/54	0.67	0.81	1.21	0.16
Germany	Store assistant	51/62	0.25	0.48	1.90	0.01
Netherlands	Cook	113/133	0.80	0.76	0.95	0.71
Netherlands	Payroll clerk	97/89	0.26	0.35	1.35	0.29
Netherlands	Receptionist	62/50	0.27	0.46	1.68	0.06
Netherlands	Sales representative	83/68	0.37	0.47	1.26	0.39
Netherlands	Software developer	82/72	0.83	0.78	0.94	0.65
Netherlands	Store assistant	65/68	0.20	0.44	2.21	0.00
Norway	Cook	36/41	0.33	0.34	1.02	1.00
Norway	Payroll clerk	46/43	0.33	0.26	0.78	0.71
Norway	Receptionist	9/11	0.44	0.18	0.41	0.35
Norway	Sales representative	91/84	0.25	0.32	1.27	0.51
Norway	Software developer	59/53	0.46	0.51	1.11	0.82
Norway	Store assistant	35/39	0.09	0.21	2.39	0.20
Spain	Cook	175/189	0.22	0.23	1.05	0.96
Spain	Payroll clerk	86/81	0.14	0.26	1.86	0.07
Spain	Receptionist	76/51	0.05	0.24	4.47	0.00
Spain	Sales representative	34/35	0.38	0.31	0.82	0.79
Spain	Software developer	28/23	0.57	0.52	0.91	0.92
Spain	Store assistant	105/76	0.10	0.17	1.80	0.21
United Kingdom	Cook	61/49	0.41	0.45	1.10	0.90
United Kingdom	Payroll clerk	115/93	0.06	0.29	4.77	0.00
United Kingdom	Receptionist	53/51	0.19	0.12	0.62	0.53
United Kingdom	Sales representative	67/71	0.18	0.21	1.18	0.86
United Kingdom	Software developer	64/50	0.30	0.38	1.28	0.57
United Kingdom	Store assistant	49/63	0.33	0.17	0.53	0.10
United States	Cook	37/40	0.54	0.45	0.83	0.65
United States	Payroll clerk	55/34	0.13	0.15	1.16	0.96
United States	Receptionist	46/38	0.15	0.21	1.38	0.72
United States	Sales representative	37/39	0.38	0.28	0.75	0.59
United States	Software developer	36/46	0.36	0.35	0.96	0.99
United States	Store assistant	43/51	0.26	0.33	1.30	0.62

country. According to these estimates, we find evidence of hiring discrimination against male applicants in United Kingdom, Spain, Germany, and the Netherlands. The gender differences range from 0 per cent in the US to 9 percentage points in Germany. Thus, we observe gender discrimination in hiring against men in four out of six countries.[10]

Taken separately, the country-specific estimates therefore add to the heterogeneity of findings reported in existing single-country studies. We also note fairly large standard errors associated with the country-specific coefficients, implying overlapping confidence intervals. However, visually overlapping confidence intervals in independent samples do not necessarily imply that the coefficients are not significantly different from each other. As this is a harmonized design, we can compare the countries directly. To test directly for country differences in the gender coefficients, we pooled the data using the following linear probability regression model:

$$Y = a + b1 * Female + b2 * Country + b3 * Female * Country + e$$

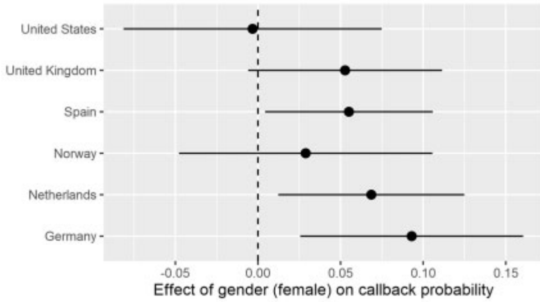

Figure 1. Effect of gender on callback probability. *Note:* Coefficients with 95 per cent confidence intervals from linear probability models estimated for each country, including occupation controls (Supplementary Table S2, models 1–6)

As shown in Supplementary Table S3, only one of the contrasts is significant, namely, that between the United States and Germany, the countries with the lowest and highest gender coefficients, respectively. However, given that there are 30 contrasts in this equation, we would expect to observe 1–2 significant outcomes (5 per cent) by chance.

Thus far, the field experiment has revealed that employers discriminate against male but not female applicants. Second, although the gender coefficients are statistically significant in four out of six countries (United Kingdom, Germany, the Netherlands, and Spain), we find no convincing evidence of cross-national differences in gender discrimination.[11] Given the widespread evidence of female labour market disadvantage and the large cross-national variation in structural, institutional, and cultural dimensions documented in Table 2, our finding of no cross-national differences in hiring discrimination is surprising. However, no previous study has examined this topic in a rigorous comparative way.

When using invitation for an interview, a stricter definition of callbacks, as the dependent variable, we find smaller country differences in gender discrimination in hiring (compare Figure 1 with Supplementary Figure S1). As the stricter version of callback (invitation for an interview) are less frequent than the wider version, the standard errors for these estimates are slightly larger, which can be seen by comparing Figure 1 with Supplementary Figure S1. This means that for the interview variable, the 95 per cent confidence intervals are slightly wider, and that it is only for Spain where the estimate is statistically significant.

Summary and Conclusion

Despite recent changes, on average, women still have lower earnings and worse career prospects. These well-known facts are true according to reliable and national representative data, such as labour force surveys and register data. The key question is why. Broadly speaking, two explanations have been provided. First, women and men might sort into different jobs because of their different educational and occupational choices, and their different work–life balance preferences and constraints, all of which accumulate to different employment trajectories and outcomes. This is the supply-side story. Second, men and women might sort into different jobs because employers discriminate women, particularly in the best-paid jobs. According to this demand-side explanation, hiring discrimination against women would be an important explanation for women's labour-market disadvantage. Because studies based on observational data cannot empirically adjudicate between supply and demand side explanations, there is a need for field experiments to provide reliable and valid estimates of employers' hiring discrimination.

Interestingly, the story jointly told by previous field experiments clashes with the conventional account of female disadvantage. It is often the fictitious male applicants, not the females, who are discriminated in hiring processes. In particular, there is evidence that women are favoured in female-dominated occupations. However, the heterogeneity of previous studies, in terms of occupations included, timing of the studies, and at what geographical level (local or national) they took place, makes comparisons difficult. Against this background, we made use of a harmonized field experiment in six countries to provide comparable, reliable, and balanced cross-national documentation of hiring discrimination against men and women.

The field experimental data show *no evidence of hiring discrimination against women* in any of the occupations in any of the countries included. The countries vary in a number of institutional, economic, and cultural dimensions potentially affecting employers' likelihood of discriminating against women. We also included occupations varying in skill requirements and customer contact. And, as documented in footnote 7, the manual job content of our occupations vary from high (cooks) to low (payroll clerks). The findings reported in this study therefore constitute an important and robust piece of evidence that young women are not discriminated in the first phase of the hiring process in any of the occupations studied in any of the countries studied.

Second, we found hiring discrimination against men in Germany, the Netherlands, Spain, and the United Kingdom, where male applicants were less likely to receive a callback when they applied for jobs as store assistants (Germany and the Netherlands), receptionists

(Spain and Germany), and payroll clerks (Spain and the United Kingdom). We found no hiring discrimination against men in Norway and in the United States. However, when pooling the data, we found no statistically significant differences across countries, perhaps with the exception of the contrast between Germany and the United States.

Understanding Gender Discrimination

With these findings in mind, how can we better understand gender discrimination in hiring? We did not find any support for the generic belief that women are disadvantaged in hiring processes, as implied both in models of cultural stereotypes and statistical discrimination, where employers are assumed to believe that women are potentially unstable workers, more likely to quit their jobs to attend their families and/or generally less committed to their firms. Gender stereotypes where women are seen as mothers and housewives seem less important in hiring processes today than in the past. According to our findings, these stereotypes seem not to operate at all. We suggest a few tentative interpretations of why this is the case. First, most women today are not primarily homemakers. Second, females are more likely to be hiring agents, in particular in female-dominated occupations, and we cannot rule out the possibility of in-group (same gender) favouritism benefiting female candidates. Third, in female occupations, hiring agents might find women more stable employees than men, who might be more likely to pursue a career, thereby leaving the job they were hired for. We should also remember that the job candidates we constructed are young workers with only 4 years of working experience. This means the presented evidence does not preclude the possibility of discrimination against women in hiring, earnings, or promotion opportunities later in the career.

Interestingly, the evidence on hiring discrimination against men would seem compatible with existing theories about gender stereotypes that were formulated to account for women's disadvantage. Perspectives emphasizing the sex typing of jobs, gender categorization within work organizations, role congruency, and stereotype contents, all seem relevant for explaining discrimination against men in the matching process. Theoretically, these cultural perspectives are also compatible with the economic model of employers as (limited) rational actors who try to find the best match between job tasks and job applicants. If employers perceive certain jobs as more appropriate for women, male applicants, even if formally qualified, may be devaluated because employers believe that they are poor matches for the sex-typed job tasks. For jobs that are not sex-typed, gender stereotypes do not seem to matter in the matching process.

The above-mentioned theories should lead to *symmetrical expectations* of hiring discrimination against applicants with the 'wrong sex' in sex-typed jobs. Thus, they cannot help us understand why women were not discriminated in the male-dominated occupation we included: software developers, an occupation which requires continuous training and where job disruptions are particularly hazardous for employers. To understand this, we can only speculate. It could be that the IT sector is more tolerant, pioneering a new work–life gender-egalitarian culture (Faulkner, 2009, but see Bertogg *et al.*, 2020). Alternatively, given the low proportion of women who enter STEM fields, IT employers might believe female applicants are positively selected in unobserved characteristics. Another possibility is that employers might be nervous that they have implicit or hidden bias against women. As a result, they may overreact and give women advantages in hiring. Whatever the reason is, finding no hiring discrimination against women in IT jobs constitutes an important challenge to both cultural and economic theories of 'gender' discrimination.

However surprising, the presented evidence is not at odds with previous research on hiring discrimination. The key to explaining divergent results likely lies in the occupations studied. For balanced studies, including both female- and male-dominated occupations, and gender-neutral occupations, the aggregate outcome would be close to zero gender discrimination in hiring. For more unbalanced studies, like the GEMM study, which includes two clearly female-typed occupations, and only one strongly male-dominated occupation, we might expect an aggregated pattern showing hiring discrimination against men. In principle, the same logic should apply for unbalanced studies including a higher proportion of male dominated occupations, but then we would expect an aggregated pattern of hiring discrimination of females. Yet the findings regarding the male-dominated occupation we included cast doubts on the symmetrical nature of hiring discrimination by gender. Interestingly, when scholars plan to study gender differences in hiring discrimination, we tend to think about discrimination of women, not men, yet previous experiments seem to include more female- than male-dominated occupations. More research including more occupations is needed.

Lack of Cross-National Variation

Despite differences in labour market conditions, family policies, and cultural norms, we found no clear evidence of cross-national variation in hiring discrimination. An explanation might be that the associations of gender stereotypes and jobs, while culturally embedded, are fairly universal across advanced Western economies (but see Supplementary Table S1 for national variations in occupational gender distributions), and hiring agents across these societies are similarly influenced by these views. Given the embeddedness of job-specific gender stereotypes, one might be pessimistic with regard to the possibilities of policy reforms to encourage gender balance. In addition, the implications of our study appear even more serious given that male-dominated occupations related to the industrial society are gradually vanishing. On the other hand, if gender-neutral occupations are growing in size, gender stereotypes will become less important over time. Thus, we have a cultural and a structural argument, and future research would benefit from addressing both arguments.

Naturally, this study has limitations. Field experiments investigate discrimination in the initial stages of the hiring process and do not give information about who gets the jobs, at what wages, and with what career opportunities. Second, the field experiment provides information about the outcomes of job applications for young applicants 22–26 years of age, and we cannot know what the situation would have looked like if we had included older fictitious applicants. Similarly, we have not tested employers' reactions to applicants with family obligations. It should be noted though, that a Swedish study including older applicants, found no difference in employers' reactions to mothers and fathers (Bygren, Erlandsson and Gähler, 2017).

Field experiments cannot cover the whole labour market, and the outcomes of these experiments are only representative for the included occupations. The GEMM study includes six occupations, requiring an educational level varying from a high school diploma to a bachelor's degree. With a limited number of male and female applications within each occupation, we are abstained from analysing in more detail the variation in types of jobs within occupations (e.g. managerial jobs).

We believe that the implications of our findings are important. In particular, we need to update our knowledge of gender discrimination and the belief that women are always the disadvantaged group. This belief might have been correct earlier, but today, at least for the occupations we examined, we found no evidence of hiring discrimination against female job applicants in any of the six countries included. Rather, we observed hiring discrimination against males in female-dominated jobs, whereas female applicants were favoured in female-dominated occupations and not discriminated in the other occupations we included. Future research should explore more in-depth the mechanisms associated with this (reversed) gender gap in hiring discrimination and delineate its boundary conditions.

Supplementary Data

Supplementary data are available at *ESR* online.

Acknowledgements

This project received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 649255; the Research Council of Norway, grant number 287016; The Netherlands Organization for Scientific Research (NWO), (016.Vidi.185.041). We thank Laura García Llamas and Louis Klobes for valuable research assistance.

Conflict of interest

We are aware of no potential conflict of interest that might raise questions of bias in our work.

Endnotes

1 For information on 'Growth, Equal Opportunities, Migration and Markets' (GEMM) project, financed by Horizon2020, see http://gemm2020.eu/.

2 If employers act upon a perceived group difference in *the variance* of unobserved expected productivity, field experimental evidence of discrimination may not be very informative (Heckman and Siegelman, 1993). Using the method proposed by Neumark (2012), Baert (2015) found no evidence of this bias related to gender heterogeneity.

3 Several concepts have been introduced to differentiate so-called error discrimination (England, 1994) and stereotype-based discrimination (Bobbitt-Zeher, 2011) from the economic-rational model, but the theory of statistical discrimination (albeit with bounded rationality) can easily accommodate the notion of stereotypes affecting employers' hiring decisions.

4 See Di Stasio and Larsen (2020) for a study of the combined effects of ethnicity and gender on employers callbacks, based on the GEMM occupations.

5 To find suitable names for the applicants, an online name search was conducted on the websites of national name registers and the most frequent names in the applicants' birth year were listed. Names were then carefully chosen to avoid connotations to

religion or class. Finally, we used official register data to identify the most common surnames in each country. For the United States, we used census data (U.S. Census Bureau, 2010) to ensure that employers would identify the names as typical white names.

6 The age used for fictitious job applicants in field experiments of gender discrimination in hiring varies. See Table 1.

7 The O*NET dataset (previously called the Dictionary of Occupational Titles) provides very detailed information of the task-content of occupations in the United States. It covers 449 detailed occupations and provides 277 descriptors for each occupation. Using these data, we performed a factor analysis to measure the manual skill content of the jobs. We converted 2,000 US Census occupations into their ISCO-88 four-digit equivalents by means of a crosswalk provided by the Centre for Longitudinal Studies, Institute of Education, University of London. We found that the GEMM occupations vary between having a manual job content score of 0.76 (cooks) to 0.23 (payroll clerks). See also Ortega and Polavieja (2012).

8 We would have needed a much larger sample if we were to include more than a binary gender variable.

9 Due to the well-known problems with logistic regression (Mood, 2010), especially concerning comparisons across samples and interaction effects, we do not present logit models here. The results are generally similar and are available upon request.

10 Using a narrower definition of callbacks, see Supplementary Information, we find significantly higher callbacks to women (0.07 and 0.06) in Spain and the Netherlands, whereas the gender coefficient, albeit positive in favour of females, is not significant in the other countries.

11 The constant terms in Supplementary Table S2 indicate the probability of receiving a callback for male applicants. They vary from low (Spain: 0.19) via moderately low in the United Kingdom, Norway, and the United States (with intervals between 0.32 and 0.50), to high in Germany and the Netherlands (0.70–0.74). These cross-national differences in baseline callbacks reflect country-level differences in demand for labour and/or a better fit of the applications.

References

Acker, J. (1990). Hierarchies, jobs, bodies: a theory of gendered organizations. *Gender & Society*, **4**, 139–158.

Ahmed, A., Granberg, M. and Khanna, S. (2021). Gender discrimination in hiring: an experimental reexamination of the Swedish case. *Plos One*, **16**, 0245513.

Albert, R., Escot, L. and Fernández-Cornejo, J. A. (2011). A field experiment to study sex and age discrimination in the Madrid labour market. *International Journal of Human Resource Management*, **22**, 351–375.

Arrow, K. J. 1972. Models of job discrimination. In Pascal A. H. (Eds.), *Racial Discrimination in Economic Life*. New York: Lexington Books, pp. 83–102.

Azmat, G. and Petrongolo, B. (2014). Gender and the labor market: what have we learned from field and lab experiments? *Labour Economics*, **30**, 32–40.

Baert, S. (2015). Field experimental evidence on gender discrimination in hiring: biased as Heckman and Siegelman predicted?. *Economics*, **9**, 1–11.

Baert, S., Pauw, A.-S. D. and Deschacht, N. (2016). Do employer preferences contribute to sticky floors? *Industrial and Labor Relations Review*, **63**, 714–736.

Becker, S. O., Fernandes, A. and Weichselbaumer, D. (2019). Discrimination in hiring based on potential and realized fertility: evidence from a large-scale field experiment. *Labour Economics*, **59**, 139–152.

Berson, C. 2012. *Does Competition Induce Hiring Equity?*, available from: https://halshs.archives-ouvertes.fr/halshs-00718627/document [accessed 24 September 2021].

Bertogg, A. *et al.* (2020). Gender discrimination in the hiring of skilled professionals in two male-dominated occupational fields: a factorial survey experiment with real-world vacancies and recruiters in four European countries. *Köln Z Soziol* **(Suppl 1) 72**, 261–289.

Bielby, W. T. and Baron, J. N. (1986). Men and women at work: sex segregation and statistical discrimination. *American Journal of Sociology*, **91**, 759–799.

Birkelund, G. E. (2016). Rational laziness – when time is limited, supply abundant, and decisions have to be made. *Analyse & Kritik. Zeitschrift Für Sozialtheorie*, **38**, 203–226.

Birkelund, G. E. *et al.* (2019). Do terrorist attacks affect ethnic discrimination in the labour market? Evidence from two randomized field experiments. *British Journal of Sociology*, **70**, 241–260.

Blau, F. D. and Kahn, L. M. (2017). The gender wage gap: extent, trends, and explanations. *Journal of Economic Literature*, **55**, 789–865.

Bobbitt-Zeher, D. (2011). Gender discrimination at work: connecting gender stereotypes, institutional policies, and gender composition of workplace. *Gender & Society*, **25**, 764–786.

Booth, A. and Leigh, A. (2010). Do employers discriminate by gender? A field experiment in female-dominated occupations. *Economics Letters*, **107**, 236–238.

Brandén, M., Bygren, M. and Gähler, M. (2018). Can the trailing spouse phenomenon be explained by employer recruitment choices? *Population, Space and Place*, **24**, e2141.

Bygren, M., Erlandsson, A. and Gähler, M. (2017). Do employers prefer fathers? Evidence from a field experiment testing the gender by parenthood interaction effect on callbacks to job applications. *European Sociological Review*, **33**, 337–348.

Capéau, B. *et al.* (2012). *Two Concepts of Discriminaiton: Inequality of Opportunity versus Unequal Treatment of Equals*. ECARES Working Paper No. 2012/58.

Carlsson, M. (2011). Does hiring discrimination cause gender segregation in the Swedish labor market? *Feminist Economics*, **17**, 71–102.

Carlsson, M. and Eriksson, S. (2017). *The Effect of Age and Gender on Labor Demand Evidence from a Field Experiment*. Working Paper No. 2017:4. Sweden: Linnaeus University.

Carlsson, R. *et al.* (2014). *Testing for Backlash in Hiring: A Field Experiment on Agency, Communion, and Gender*. Working paper. Sweden: Linnaeus University.

Castilla, E. J. and Benard, S. (2010). The paradox of meritocracy in organizations. *Administrative Science Quarterly*, **55**, 543–676.

Cejka, M. A. and Eagly, A. H. (1999). Gender-stereotypic images of occupations correspond to the sex segregation of employment. *Personality and Social Psychology Bulletin*, **25**, 413–423.

Chang, M. L. (2004). Growing pains: cross-national variation in sex segregation in sixteen developing countries. *American Sociological Review*, **69**, 114–137.

Charles, M. (2011). A world of difference: international trends in women's economic status. *Annual Review of Sociology*, **37**, 355–371.

Correll, S. J., Benard, S. and Paik, I. (2007). Getting a job: is there a motherhood penalty? *American Journal of Sociology*, **112**, 1297–1338.

Cuddy, A. J. C., Fiske, S. T. and Glick, P. (2008). Warmth and competence as universal dimensions of social perception: the stereotype content model and the BIAS map. *Advances in Experimental Social Psychology*, **40**, 61–149.

Di Stasio, V. and Lancee, B. (2019). Understanding why employers discriminate, where and against whom: the potential of cross-national, factorial and multi-group field experiments. *Research in Stratification and Mobility*, available from: 10.1016/j.rssm.2019.100463

Di Stasio, V. and Larsen, E. N. (2020). The racialized and gendered workplace: applying an intersectional lens to a field experiment on hiring discrimination in five European labor markets. *Social Psychology Quarterly*, **83**, 229–250.

Duguet, E. *et al.* (2012). First order stochastic dominance and the measurement of hiring discriminaiton: a ranking extension of correspondence testing with an application to gender and origin, available from: https://halshs.archives-ouvertes.fr/halshs-00731005/

Duguet, E., Loïc, D. and P. and Petit, P. (2017). *Hiring discrimination against women: distinguishing taste based discrimination from statistical discrimination*. Available at SSRN: https://ssrn.com/abstract=3083957 or 10.2139/ssrn.3083957.

England, P. (1994). Neoclassical economists' theories of discrimination. In Burstein, P. (Ed.), *Equal Employment Opportunity: Labor Market Discrimination and Public Policy*. New York: Aldine De Gruyter, pp. 59–70.

Fang, H. and Moro, A. (2011). Theories of statistical discrimination and affirmative action: a survey. In Benhabib, J., Bisin, A. and Jackson, M. O. (Eds.), *Handbook of Social Economics*. San Diego: Elsevier, Chapter 5, pp. 133–200.

Faulkner, W. (2009). Doing gender in engineering workplace cultures. I. Observations from the Field. *Engineering Studies*, **1**, 3–18.

Fernández, R. (2013). Cultural change as learning: the evolution of female labor force participation over a century. *American Economic Review*, **103**, 472–500.

Fortin, N. M. (2005). Gender role attitudes and the labour-market outcomes of women across OECD countries. *Oxford Review of Economic Policy*, **21**, 416–438.

Gaddis, S. M. (Ed.). (2018). An introduction to audit studies in the social sciences. In *Audit Studies: Behind the Scenes with Theory, Method, and Nuance*. Cham: Springer, pp. 3–44.

Gangl, M. and 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**, 341–369.

Glick, P. and Fiske, S. T. (1996). The ambivalent sexism inventory: differentiating hostile from benevolent sexism. *Journal of Personality and Social Psychology*, **70**, 491–512.

Glick, P., Zion, C. and Nelson, C. (1988). What mediates sex discrimination in hiring decisions?. *Journal of Personality and Social Psychology*, **55**, 178.

Goldin, C. (2014). A grand gender convergence: its last chapter. *American Economic Review*, **104**, 1091–1119.

González, M. J., Cortina, C. and Rodríguez, J. (2019). The role of gender stereotypes in hiring: a field experiment. *European Sociological Review*, **35**, 187–204.

Halper, L. R., Cowgill, C. M. and Rios, K. (2019). Gender bias in caregiving professions: the role of perceived warmth. *Journal of Applied Social Psychology*, **49**, 1–14.

Heckman, J. J. and Siegelman, P. (1993). The urban institute audit studies: their methods and findings. In Fix, M. and Struyk, R. (Eds.), *Clear and Convincing Evidence: Measurement of Discrimination in America*. Washington, DC: Urban Institute Press.

Hofstede Insights (2020). *Compare Countries*, available from: https://www.hofstede-insights.com/product/compare-countries/ [accessed 25 June 2020].

Jackson, M. (2009). Disadvantaged through discrimination? The role of employers in social stratification. *The British Journal of Sociology*, **60**, 669–692.

Jacobsen, J., Khamis, M. and Yuksel M. (2015). Convergence in men's and women's life patterns: lifetime work, lifetime earnings, and human capital investment. *Research in Labor Economics*, **41**, 1–33.

Lancee, B. (2019). Ethnic discrimination in hiring: comparing groups across contexts. Results from a cross-national field experiment. *Journal of Ethnic and Migration Studies*, **47**, 1181–1200.

Lancee, B. *et al.* (2019a). *The GEMM Study: A Cross-National Harmonized Field Experiment on Labour Market Discrimination: Codebook*.http: //dx.doi.org/10.2139/ssrn.3398273

Lancee, B. *et al.* (2019b). *The GEMM Study: A Cross-National Harmonized Field Experiment on Labour Market Discrimination: Technical Report*. 10.2139/ssrn.3398191

Larsen, E. N. (2020). Induced competition in matched correspondence tests: conceptual and methodological considerations. *Research in Social Stratification and Mobility*, **65**, 100475.

Levanon, A. and Grusky, D. B. (2016). The persistence of extreme gender segregation in the twenty-first century. *American Journal of Sociology*, **22**, 573–619.

Mandel, H. and Semyonov, M. (2006). A welfare state paradox: state interventions and women's employment opportunities in 22 countries. *American Journal of Sociology*, **111**, 1910–1949.

Mood, C. (2010). Logistic regression: why we cannot do what we think we can do, and what we can do about it. *European Sociological Review*, **26**, 67–82.

Neumark, D. (2012). Detecting discrimination in audit and correspondence studies. *Journal of Human Resources*, **47**, 1128–1157.

Neumark, D. (2018). Experimental research on labor market discrimination. *Journal of Economic Literature*, **56**, 799–866.

Neumark, D., Bank, R. J. and Van Nort, K. D. (1996). Sex discrimination in restaurant hiring: an audit study. *The Quarterly Journal of Economics*, **111**, 915–941.

OECD. (2015). *Education at a Glance 2015*, available from: https://www.oecd.org/gender/data/gender-gap-in-education.htm [accessed 5 January 2020].

OECD. (2017). *Short-Term Labour Market Statistics*, available from: https://stats.oecd.org/OECDStat_Metadata/ShowMetadata.ashx?Dataset=STLABOUR&ShowOnWeb=true&Lang=en) [accessed 24 June 2020].

OECD. (2019). *"Unemployment rate"*. OECD *Employment Outlook*, available from: https://data.oecd.org/unemp/unemployment-rate.htm

OECD. (2020a). *Index of Regulation on Individual Dismissal of Workers with Regular Contracts*, available from: https://www1.compareyourcountry.org/employment-protection-legislation/en/0/178/ranking/ [accessed 23 June 2020].

OECD. (2020b). *Length of Maternity Leave, Parental Leave and Paid Father-Specific Leave*, available from: https://www.oecd.org/gender/data/length-of-maternity-leave-parental-leave-and-paid-father-specific-leave.htm [accessed 25 June 2020].

OECD. (2020c). *OECD Family Database*, available from: http://www.oecd.org/els/family/database.htm [accessed 25 June 2020].

OECD. (2020d). *Exployment: Share of Employed In Part-Time Employment, by Sex and Age Group*, available from: https://stats.oecd.org/index.aspx?queryid=54746 [accessed 25 June 2020].

Petersen, T. and Morgan, L. A. (1995). Separate and unequal: occupation establishment sex-segregation and the gender wage-gap. *American Journal of Sociology*, **101**, 329–365.

Petersen, T. and Saporta, I. (2004). The opportunity structure for discrimination. *American Journal of Sociology*, **109**, 852–901.

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

Pew Research Center. (2017). *In Many Countries at Least Four-in-Ten in the Labor Force are Women*, available from: https://www.pewresearch.org/fact-tank/2017/03/07/in-many-countries-at-least-four-in-ten-in-the-labor-force-are-women/ [accessed 25 June 2020].

Phelps, E. S. (1972). The statistical theory of racism and sexism. *American Economic Review*, **62**, 659–661.

Polavieja, J. G. (2012). Socially embedded investments: explaining gender differences in job-specific skills. *American Journal of Sociology*, **118**, 592–634.

Polavieja, J. G. (2015). Capturing culture: a new method to estimate exogenous cultural effects using migrant populations. *American Sociological Review*, **80**, 166–191.

Reskin: Plenum, B. F. (2000). Employment discrimination and its remedies. In Berg, I. and Kalleberg, A. (Eds.), *Handbook on Labor Market Research*. New York.

Reskin, B. F. and Roos, P. A. (1990). *Job Queues, Gender Queues: Explaining Women's Inroads into Male Occupations*. Philadelphia: Temple University Press.

Riach, P. A. and Rich, J. (1987). Testing for Sexual Discrimination in the Labour Market. *Australian Economic Papers*, **26**, 165–178.

Riach, P. A. and Rich, J. (2002). Field experiments of discrimination in the market place. *The Economic Journal*, **112**, F480–F518.

Riach, P. A. and Rich, J. (2006). An experimental investigation of sexual discrimination in hiring in the English labor market. *The B.E. Journal of Economic Analysis & Policy*, **6**, available from: http://www.bepress.com/bejeap/advances/vol6/iss2/art1

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. Available at SSRN: https://ssrn.com/abstract=2517887.

Ridgeway, C. L. (1997). Interaction and the conservation of gender inequality: considering employment. *American Sociological Review*, **62**, 218–235.

Rivera, L. A. and Tilcsik, A. (2016). Class advantage, commitment penalty: the gendered effect of social class signals in an elite labor market. *American Sociological Review*, **81**, 1097–1131.

Rudman, L. and Phelan, J. E. (2008). Backlash effects for disconfirming gender stereotypes in organizations. *Research in Organizational Behavior*, **28**, 61–79.

Stier, H., Lewin-Epstein, N. and Braun, M. (2001). Welfare regimes, family-supportive policies, and women's employment along the life-course. *American Journal of Sociology*, **106**, 1731–1760.

The World Economic Forum. (2017). The *Global Gender Gap Report 2017*, available from: https://www.weforum.org/reports/the-global-gender-gap-report-2017

Torre, M. (2014). The scarring effect of “women’s work”: the determinants of women’s attrition from male-dominated occupations. *Social Forces*, **93**, 1–29.

Torre, M. (2018). Stopgappers? The occupational trajectories of men in female-dominated occupations. *Work and Occupations*, **45**, 283–312.

U.S. Census Bureau. (2010). *Frequently Occurring Surnames from the 2010 Census*, available from: https://www.census.gov/topics/population/genealogy/data/2010_surnames.html [accessed 18 January 2019].

Vuolo, M., Uggen, C. and Lageson, S. (2018). To match or not to match? Statistical and substantive considerations in audit design and analysis. In Gaddis, S. M. (Ed.), *Audit Studies: Behind the Scenes with Theory, Method, and Nuance*. Cham: Springer, pp. 119–140.

Weichselbaumer, D. (2004). Is it sex or personality? The impact of sex stereotypes on discrimination in applicant selection. *Eastern Economic Journal*, **30**, 159–186.

Weichselbaumer, D. and Winter-Ebmer, R. (2005). A meta-analysis of the international gender wage gap. *Journal of Economic Surveys*, **19**, 479–511.

Williams, C. L., Muller, C. and Kilanski, K. (2012). Gendered organizations in the new economy. *Gender & Society*, **26**, 549–573.

Williams, W. M. and Ceci, S. J. (2015). National hiring experiments reveal 2:1 faculty preference for women on STEM tenure track. *PNAS*, **112**, 5360–5365.

Yavorsky, J. E. (2019). Uneven patterns of inequality: an audit analysis of hiring-related practicies by gendered and classed contexts. *Social Forces*, **98**, 461–492.

Zhou, X., Zhang, J. and Song, X. (2013). *Gender Discrimination in Hiring: Evidence from 19,130 Resumes in China*. MPRA paper No. 43543. Available at SSRN: https://ssrn.com/abstract=2195840 or 10.2139/ssrn.2195840.

Gunn Elisabeth Birkelund is a Professor of Sociology at University of Oslo. Her main research interests include analytical sociology, labor market studies, social inequalities, and population dynamics. She is a Fellow at The European Academy of Sociology, and Secretary General at the Norwegian Academy of Science and Letters. Her articles have appeared in *European Sociological Review, Social Forces, International Migration Review,* European Societies, and, earlier, in *American Journal of Sociology* and *American Sociological Review*.

Bram Lancee is an Associate Professor of Sociology at the University of Amsterdam. Current research interests include social capital, ethnic minorities and the labour market, inequality, attitudes towards immigration, and ethnic discrimination. His work has been published in journals, such as *Social Forces, European Sociological Review, International Migration Review, Journal of Ethnic and Migration Studies,* and *Social Science Research*.

Edvard N. Larsen is a postdoctoral researcher in Sociology at the University of Oslo, and Researcher II at the KIFO Institute of Church, Religion, and Worldview research. His main research interests are social inequality, migration, labor market discrimination, and religion. His work has been published in the journals *Journal of Ethnic and Migration Studies*, *Social Psychology Quarterly*, and *Research on Social Stratification and Mobility*.

Javier Polavieja (Oxford University PhD in Sociology, 2001) is Banco Santander Professor of Sociology and Director of the D-Lab at the Department of Social Sciences, University Carlos III of Madrid, as well as Research Fellow at the Institute of Economics and the Carlos III-Juan March Institute. His main fields of research are social stratification, political sociology, and migration research. His work has been published in *American Journal of Sociology*, *American Sociological Review*, *European Sociological Review*, *Social Forces*, *Socio-Economic Review*, *Labour Economics*, *Political Behavior*, *Electoral Studies*, *International Migration*, and *Social Indicators Research*.

Jonas Radl is an Associate Professor of Sociology at Universidad Carlos III de Madrid and Head of the Research Group ‘Effort and Social Inequality’ at WZB Berlin Social Science Center. Current research interests comprise social stratification and the life course. His work has been published in journals such as *European Sociological Review*, *Social Forces*, and *Socio-economic Review*.

Ruta Yemane is a Research Fellow at WZB Berlin Social Science Center in the migration, integration, and transnationalization research unit. Her research focuses on labor market discrimination, racism, and stereotypes. Her work has been published in the *British Journal of Social Psychology* and the *Journal of Ethnic and Migration Studies*.