



Gender, Parenthood, Ethnicity and Discrimination in the Labor Market

Experimental Studies on Discrimination in Recruitment in Sweden

Anni Erlandsson



Gender, Parenthood, Ethnicity and Discrimination in the Labor Market

Experimental Studies on Discrimination in Recruitment in Sweden

Anni Erlandsson

Academic dissertation for the Degree of Doctor of Philosophy in Sociology at Stockholm University to be publicly defended on Friday 25 March 2022 at 13.00 in hörsal 11, hus F, Universitetsvägen 10 F; Online via Zoom, public link is available at www.sociology.su.se.

Abstract

This dissertation uses experimental methods to study hiring discrimination based on gender, parenthood and ethnicity in the Swedish labor market. Also, the role of recruiter gender for gender and ethnic discrimination is studied. Three of the four empirical studies (Study I, Study II and Study IV), are based on field experiment data using a correspondence testing method. This involves fictitious job applications sent to announced jobs, and the employer responses to these. Signals of applicant characteristics such as gender, parenthood status, and ethnicity are randomly assigned to the job applications whereas qualifications are held constant (within occupations). Study III is based on a laboratory experiment in which (fictitious) job candidates are evaluated.

Study I does not show any evidence of discrimination based on gender or parenthood, or any combination of these, in the first step of the hiring process, neither in highly nor less qualified occupations. Study II shows that male job applicants are favored by male recruiters, especially in gender-balanced occupations.

Study III shows a statistically significant gender bias in job applicant ratings in favor of female applicants in a laboratory setting. This is particularly the case for female evaluators. Moreover, Study III shows no motherhood penalty in the applicant ratings.

Study IV presents evidence of ethnic discrimination against foreign-named job applicants by both male and female recruiters. Further, there is evidence of gendered ethnic discrimination, i.e., male applicants with foreign-sounding names receive considerably fewer positive responses than female applicants with foreign-sounding names. While female recruiters favor foreign-named female applicants over foreign-named male applicants, particularly in highly qualified occupations, male recruiters appear to prefer foreign-named females over foreign-named males in male-dominated occupations.

To summarize, the findings from this dissertation provide little support for the notion of discrimination in recruitment as an important mechanism behind gender inequalities in the Swedish labor market. However, the results indicate that discrimination in the recruitment process contributes to the labor market inequality of ethnic minorities, and of ethnic minority men in particular. Moreover, the findings suggest that recruiter gender matters for the success of male and female job candidates, and in particular for foreign-named men and women, at least in some occupational contexts.

Keywords: *discrimination, ethnicity, field experiment, gender, in-group bias, labor market, parenthood, recruitment.*

Stockholm 2022

<http://urn.kb.se/resolve?urn=urn:nbn:se:su:diva-201729>

ISBN 978-91-7911-792-4

ISBN 978-91-7911-793-1

ISSN 0283-8222



Stockholm
University

Department of Sociology

Stockholm University, 106 91 Stockholm

GENDER, PARENTHOOD, ETHNICITY AND DISCRIMINATION IN
THE LABOR MARKET

Anni Erlandsson



Gender, Parenthood, Ethnicity and Discrimination in the Labor Market

Experimental Studies on Discrimination in Recruitment in Sweden

Anni Erlandsson

©Anni Erlandsson, Stockholm University 2022

ISBN print 978-91-7911-792-4

ISBN PDF 978-91-7911-793-1

ISSN 0283-8222

Printed in Sweden by Universitetservice US-AB, Stockholm 2022

To the memory of my
grandparents

Contents

List of studies	i
Acknowledgements	ii
Abstract	iv
Sammanfattning	vi
Introduction	1
Discrimination and its grounds	2
Institutional context	4
Methodological context	5
Detecting discrimination empirically	5
Correspondence testing method	8
Critique of correspondence testing	9
Measuring discrimination	10
One application.....	10
Ethical considerations	11
Field experiment	11
Laboratory experiment.....	14
Summary of the four empirical studies	14
Concluding Discussion.....	16
References.....	19

List of studies

- I. 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(3), 337-348.
(Author contributions: the authors have contributed in equal shares, i.e., one third each.)
- II. Erlandsson, A. (2019). Do men favor men in recruitment? A field experiment in the Swedish labor market. *Work and Occupations*, 46(3), 239-264.
- III. Erlandsson, A., Bygren, M., and Gähler, M. She is a woman and women rule! A laboratory experiment on recruitment discrimination based on gender and parenthood. *Manuscript submitted*.
(Author contributions: Erlandsson 80%, Bygren 10%, Gähler 10%.)
- IV. Erlandsson, A. Gendered ethnic discrimination and the role of recruiter gender. A field experiment. *Manuscript submitted*.

Acknowledgements

This dissertation would not exist without two persons. Thus, I would like to thank my main supervisor Michael Gähler and my co-supervisor Magnus Bygren for including me in their research project from the very start. It has been a long journey, during which I have learned tremendously.

Michael, as my main supervisor, besides a detailed reading of my work and providing valuable feedback, you have repeatedly expressed that you believe in me, making me believe that I can do this. At the same time, you have pushed me to improve my texts endlessly, and you have patiently explained things when I did not understand, or when I argued against you. I am grateful for all this.

Magnus, I am most grateful for your genuine and insightful comments as well as the advice and expertise that you have shared with me throughout the years. While you have encouraged me to learn on my own, you have also helped me countless times to find solutions.

Last but not least, thank you, Karin Halldén, my other co-supervisor, for repeatedly reading my work, sharing your research knowledge and providing useful advice. You have always been encouraging and generous with your time to discuss numerous work-related issues and you have helped me to encounter them with a positive mindset. Again, thank you.

There are many colleagues at the Institute for Social Research (SOFI) and at the Department of Sociology who have helped me to improve this dissertation while also making my time as a PhD student more enjoyable. Thank you all, and in particular the members of the LNU group who have several times read and commented on my work. Special thanks to Lotta Magnusson and Magnus Nermo for valuable comments on earlier drafts of this dissertation and for encouragement throughout the years. To Erik Bihagen, thank you for useful feedback at several seminars. I also want to thank Moa Bursell for inspiration and much appreciated comments, especially at my half-time seminar. To my current and previous fellow PhD students and office mates, thank you for all the work-related and other conversations. These may not have been very frequent lately, yet much appreciated. I hope you know who you are. In addition, thank you, Eva-Lisa Palmtag, for peer support since the day I started at SOFI.

Finally, I am grateful for my family and friends, near and far, outside the academia. Thank you, mom and dad, Ritva and Olli, for supporting me in my

decisions. To my father-in-law, Ivar, thank you for a never-ending interest in my PhD studies. Most importantly, thank you, Henrik, for being there for me, always. At last, thank you, Jonathan and Emilia, for being you. You have given me a whole new perspective to life, work and research. There is nothing more challenging and rewarding than being your mom.

Stockholm, February 2022

Abstract

This dissertation uses experimental methods to study hiring discrimination based on gender, parenthood and ethnicity in the Swedish labor market. Also, the role of recruiter gender for gender and ethnic discrimination is studied. Three of the four empirical studies (Study I, Study II and Study IV), are based on field experiment data using a correspondence testing method. This involves fictitious job applications sent to announced jobs, and the employer responses to these. Signals of applicant characteristics such as gender, parenthood status, and ethnicity are randomly assigned to the job applications whereas qualifications are held constant (within occupations). Study III is based on a laboratory experiment in which (fictitious) job candidates are evaluated.

Study I does not show any evidence of discrimination based on gender or parenthood, or any combination of these, in the first step of the hiring process, neither in highly nor less qualified occupations. Study II shows that male job applicants are favored by male recruiters, especially in gender-balanced occupations.

Study III shows a statistically significant gender bias in job applicant ratings in favor of female applicants in a laboratory setting. This is particularly the case for female evaluators. Moreover, Study III shows no motherhood penalty in the applicant ratings.

Study IV presents evidence of ethnic discrimination against foreign-named job applicants by both male and female recruiters. Further, there is evidence of gendered ethnic discrimination, i.e., male applicants with foreign-sounding names receive considerably fewer positive responses than female applicants with foreign-sounding names. While female recruiters favor foreign-named female applicants over foreign-named male applicants, particularly in highly qualified occupations, male recruiters appear to prefer foreign-named females over foreign-named males in male-dominated occupations.

To summarize, the findings from this dissertation provide little support for the notion of discrimination in recruitment as an important mechanism behind gender inequalities in the Swedish labor market. However, the results indicate that discrimination in the recruitment process contributes to the labor market inequality of ethnic minorities, and of ethnic minority men in particular. Moreover, the findings suggest that recruiter gender matters for

the success of male and female job candidates, and in particular for foreign-named men and women, at least in some occupational contexts.

Sammanfattning

Denna avhandling använder experimentella metoder för att undersöka anställningsdiskriminering utifrån kön, föräldraskap och etnicitet på den svenska arbetsmarknaden. Vidare utreds huruvida rekryterarens kön spelar roll för köns- och etnisk diskriminering. Tre av de fyra empiriska studierna (Studie I, Studie II och Studie IV) är baserade på fältexperimentdata. Med så kallad *correspondence testing* besvaras jobbannonser med fiktiva ansökningar, och arbetsgivarnas svar på dessa observeras. Signaler om sökandes egenskaper såsom kön, föräldraskap och etnicitet tilldelas slumpmässigt till jobbansökningarna medan kvalifikationerna hålls konstanta (inom yrken). Studie III bygger på ett laboratorieexperiment där (fiktiva) jobbkandidater bedöms.

Studie I tyder inte på att systematisk diskriminering utifrån kön eller föräldraskap, eller en kombination av dessa, förekommer i första steget av anställningsprocessen vare sig i hög- eller lågkvalificerade yrken. Studie II visar att manliga arbetsökande favoriseras av manliga rekryterare, särskilt i könsbalanserade yrken.

Studie III visar en statistiskt signifikant könsbias till förmån för kvinnliga jobbkandidater i bedömningen av fiktiva arbetsökande. Detta gäller i synnerhet för kvinnliga bedömare. Vidare visar Studie III inte på någon diskriminering gentemot mödrar i bedömningen av jobbkandidater.

Studie IV visar att etnisk diskriminering av arbetsökande med utländskt klingande namn förekommer hos både manliga och kvinnliga rekryterare. Vidare tyder resultaten på etnisk könsdiskriminering i och med att manliga sökande med utländskt klingande namn får betydligt färre positiva svar än kvinnliga sökande med utländskt klingande namn. Medan kvinnliga rekryterare favoriserar kvinnliga sökande med utländskt namn framför manliga sökande med utländskt namn, särskilt i högkvalificerade yrken, föredrar manliga rekryterare kvinnor med utländskt namn framför män med utländskt namn i mansdominerade yrken.

Sammanfattningsvis ger resultaten från denna avhandling föga stöd för föreställningen om diskriminering vid rekrytering som en viktig mekanism bakom könsojämlikhet på den svenska arbetsmarknaden. Resultaten tyder dock på att diskriminering i rekryteringsprocessen bidrar till etnisk ojämlikhet på arbetsmarknaden, och i synnerhet för etniska minoritetsmän. Dessutom tyder resultaten på att rekryterarens kön spelar roll för manliga och

kvinnliga jobbkandidaters jobbchanser, i synnerhet för män och kvinnor med utländskt namn, åtminstone i vissa yrkessammanhang.

Introduction

Labor market inequality has many dimensions and different groups of individuals experience different types of obstacles in the labor market. In connection to this, labor markets tend to be segregated at two levels in most societies. There is vertical segregation, meaning that there is an unequal distribution of individuals from different groups, such as men and women, in positions based on the level of rewards, authority and/or status. Also, the labor market is horizontally segregated in the sense that there is an unequal distribution of individuals of certain groups, e.g., men and women as well as natives and immigrants, across different sectors and types of occupations. Moreover, female-dominated occupations generally have lower status and pay than male-dominated occupations. While this dissertation focuses on recruitment discrimination based on gender, parenthood and ethnicity together with recruiter gender, in a broader sense, the dissertation relates to themes such as social stratification, inequality, wellbeing and life chances of different groups of individuals in general, and in particular to labor market inequalities.

Research shows that, in general, men have an advantage over women, and fathers an advantage over mothers, in the labor market (e.g., Charles, 2011; Harkness and Waldfogel, 2003). Men tend to have higher wages and reach authority positions more often than women (Grönlund, Halldén and Magnusson, 2017). Researchers have suggested employer discrimination as one possible reason for the observed gender gaps in labor market outcomes, and not least so in the Swedish labor market with its generous rights to parental leave and reduced working hours following parenthood (Blau and Kahn, 2017; Bygren and Gähler, 2012; Gangl and Ziefle, 2009; Mandel and Semyonov, 2006). While there is some research on gender discrimination in recruitment, not many studies have examined discrimination based on parenthood, including motherhood and fatherhood, prior to this dissertation. There is an obvious reason for this void: discriminatory behaviors cannot be identified in quantitative research using register or survey data where a residual effect of gender, or gender interacted with parenthood may be explained by productivity-related characteristics observed by the recruiter, but not the researcher.

Moreover, little is known about the role of the recruiter for gender and ethnic discrimination. Because of data limitations, endogeneity and low diversity among evaluators (recruiters) it tends to be difficult to investigate

gender-based in-group bias in the real life, outside a laboratory setting (Sandberg, 2018). Thus, few field experiments focus on the role of recruiter gender for recruitment discrimination, and, to my knowledge, there is only one such field experiment studying (gendered) ethnic discrimination (Edo et al., 2019). Hence, it is unclear what role, if any, recruiter gender plays for labor market discrimination, and whether male and female recruiters differ in their tendency (if any) to favor job applicants of one gender over the other, and if such tendency differs by the ethnicity of the applicants. These are some of the main issues addressed by this dissertation.

This dissertation includes this introductory chapter and four empirical studies. The aim of the chapter is to place the dissertation in a broader sociological context and to describe its contribution to the field. Thus, while the chapter will shortly discuss the institutional (Swedish) and empirical contexts of the studies, the methodological context is presented in more detail. The chapter starts with a presentation of some central concepts. Next, the contexts of the studies are described. Here, the methods and data are presented and discussed, followed by an account of the ethical considerations related to the data collection methods. Finally, the results from the empirical studies are summarized, followed by a concluding discussion.

Discrimination and its grounds

In the studies included in this dissertation, *discrimination* refers to a negative *or* positive bias in the treatment of individuals (over and above any differences associated with objective indicators of productivity) based on their membership in certain demographic categories, such as the intersection of gender and parenthood, or gender and ethnicity. The field experiment data used here consist of non-authentic job applications (to which gender, parenthood and ethnicity were randomly assigned) that were submitted to real advertised jobs in the labor market, and the employer responses (callbacks) to these. Thus, in Study I, II and IV, any (significant) differences in positive employer responses between groups of equally qualified job candidates, i.e., men and women or Swedish-named and foreign-named, are considered as discrimination.

The studies included here use names as indicators of *gender* and *ethnicity*. Thus, distinct male and female first names are assumed to signal the *gender* of an individual, i.e., being a man or a woman. Also, foreign-sounding, versus typical Swedish-sounding, names are used as indicators of *ethnicity* and refer to *immigrant background*. Using names to indicate gender and ethnicity is a common practice in the field of labor market discrimination (cf., Gaddis, 2018). Although ethnicity and immigrant background are not the same, for the sake of simplicity, these concepts are used as synonyms here to refer to “foreignness”.

Different theoretical approaches are developed to explain discrimination, and a few of them are shortly described here. A theoretical branch in the sociological literature argues for the importance of status characteristics regarding, e.g., gender, ethnicity and parenthood, in the labor market. Predominant cultural status rankings concerning individuals of certain groups (based on e.g., gender or ethnicity) are frequently relied on in the evaluation of performance and characteristics, and this kind of status-related beliefs generally propose men as superior to women in many fields of social life (Ridgeway, 2011; Ridgeway and Smith-Lovin, 1999). Also, motherhood as a salient worker characteristic tends to be devalued (Ridgeway and Correll, 2004).

While discrimination based on beliefs related to status characteristics can stem from a cognitive bias that operates beyond group-related productivity (cf., Correll and Benard, 2006), a dominant economic theory in discrimination research is statistical discrimination, which is based on differences in the average productivity of groups (Arrow, 1973; Phelps, 1972). In taste-based discrimination theory, also common in economics, discrimination occurs because of an individual's prejudice against, or dislike of, a certain group of individuals (Becker, 1971). Whereas statistical discrimination is based on rationality in terms of estimated productivity while an individual's actual productivity remains unclear to the employer, taste-based discrimination basically ignores possible differences in productivity. Therefore, these theories suggest statistical discrimination against foreign-named job candidates to take place because of differences in average productivity between the foreign-named (ethnic minority) and Swedish-named (ethnic majority), e.g., because of differences in language skills, whereas taste-based discrimination against foreign-named candidates would arise from a recruiter's distaste of, or prejudice against, a group, i.e., an ethnic minority. Also, while the theory of status-based discrimination lies close to taste-based discrimination, their mechanisms differ slightly, i.e., a cognitive frame versus a taste respectively. Moreover, discrimination based on incorrect estimations of average group differences in productivity is sometimes labeled error discrimination (England, 1992). However, not all researchers make this distinction between correct and erroneous estimation of average group productivity, and it may also be considered as a subcategory of statistical discrimination (Bielby and Baron, 1986; Blau, 1984).

In addition, the characteristics of the recruiter, i.e., gender, may also be of theoretical importance. People tend to associate with people like themselves, which is conceptualized as homophily in sociology (e.g., Kandel, 1978; McPherson et al., 2001). Thus, discrimination can arise from in-group favoritism (e.g., Tajfel et al., 1971), or in-group bias (Hewstone et al., 2002), i.e., one discriminates in favor of individuals in one's own in-group, e.g., individuals of the same gender as oneself, and thereby against individuals from other groups, i.e., other gender (cf., Brewer and Kramer, 1985; Hewstone et al., 2002). Therefore, these approaches suggest that female recruiters would

discriminate against male job candidates and male recruiters would discriminate against female job candidates.

Institutional context

All studies in this dissertation are conducted in Sweden. Sweden provides an interesting, yet a special, cultural and institutional context to study discrimination based on gender, parenthood and ethnicity. On the one hand, in an international perspective, Sweden as well as the other Nordic countries tend to be seen as pioneers in gender equality, with fairly small gender differences in labor market outcomes, e.g., employment rates (OECD, 2017; United Nations Development Programme, 2016). Also, in a cross-country comparison, gender equal norms are widespread among Swedes, there is strong support for sharing work and care responsibilities equally (Edlund and Öun, 2016), and the gender gap in unpaid work (housework) is small (European Institute for Gender Equality, 2017). On the other hand, Sweden has received a rather high number of immigrants, at least in comparison to its Nordic neighbors. Consequently, foreign-born constitute about 20% of the Swedish population (Statistics Sweden, 2020). While many of the immigrants have come to Sweden because of humanitarian reasons, often without basic education or face trouble in transferring their qualifications, the employment gaps between natives and immigrants are among the largest in the OECD (Organisation for Economic Co-operation and Development), across the educational distribution (OECD, 2014:5).

Sweden is typically known for its family-friendly policies and dual-earner regulations, aimed at allowing both mothers and fathers to combine work and family life. Yet, considerable gender inequalities remain. Gender segregation in the labor market and the gender gap in wages appear high in Sweden in relation to other EU countries (European Institute for Gender Equality, 2017). Also, part-time employment remains more common among women than men in Sweden (OECD, 2017). Although fathers' take-up of parental leave has increased since early 2000s, (partly) as a response to leave earmarked for fathers (Duvander and Johansson, 2012), mothers still take up the vast majority of the total number of parental leave days and days paid caring for sick children (Swedish Social Insurance Agency, 2018).

While discrimination based on gender and ethnicity, among others, is prohibited by Swedish law, a preferential treatment of job applicants based on gender, e.g., applicants of the underrepresented gender, is sometimes allowed in hiring, for example in order to reach a more gender-balanced setting (Discrimination Ombudsman, 2021). The latter is not the case in many other countries, including the United States where it tends to be relatively easy for an employer to get rid of workers. Whereas in Sweden, permanent workers generally have a relatively high employment protection, i.e.,

stronger than the OECD average (OECD, 2021). This is likely to increase the pressure on recruiters in their recruitment decisions.

Methodological context

It is difficult for researchers to capture labor market discrimination, and especially to unravel its mechanisms empirically. Discrimination seems to be very subtle, and to take place more or less automatically (Cunningham et al., 2004; Rooth, 2010). This dissertation relies on experimental methods in studying labor market discrimination and exploits opportunities to combine methods in an attempt to give a nuanced view of this phenomenon. Thus, I do not only want to answer the “if question” of discrimination, but also (if so) to take a step in order to address the “why question” of discrimination. While three of the four studies, i.e., Study I, Study II and Study IV, use field experiment data based on a correspondence study design, Study III is based on a laboratory experiment.

First, this dissertation aims to answer the “if question”, i.e., if and to what extent discrimination (based on gender and parenthood as well as ethnicity) exists in the Swedish labor market, and whether recruiter gender plays a role here. For this purpose, a so-called correspondence study, i.e., sending fictitious job applications to advertised vacancies in the labor market, is conducted (Study I, II and IV). The method will be described in detail below.

Second, within the frame of this dissertation the aim is also to address the question of which mechanisms lead to discrimination (Study III). An attempt is done by conducting a laboratory experiment in which the participants (university students, i.e., potential future recruiters) are asked to rank a number of fictive job applications and to provide arguments for the rankings. Gender and parenthood, among others, in the fictive applications are varied while merits are held constant.

Detecting discrimination empirically

A range of different approaches have been applied in order to detect discrimination in the labor market. Naturally, all the approaches have certain weaknesses, though the suitability of the chosen method depends mostly on the aim of the study. I will now briefly describe some of the alternative methods to detect discrimination and discuss their limitations (cf., Bygren and Gähler, 2021).

Examining inequality in outcomes between different groups appears to be the most common approach to capture labor market discrimination (Blau and Kahn, 2017; Pager and Shepherd, 2008), in the past. Thus, statistical analyses, mainly the analysis of large-scale datasets such as register or survey data, provide one technique to investigate differences in outcomes between

the groups of interest, such as men and women. Discrimination is then measured as the remaining gap in the outcome between the groups, e.g., the gender difference in wages, after having controlled for observed characteristics. However, one should be careful with making causal inferences based on such measures and data as not all potential individual characteristics that are relevant to the recruiter can be observed and controlled by the researcher. While the approach is important in increasing knowledge of unexplained inequality in certain labor market outcomes, because of the nature of the information inherent in the data, i.e., omitted variable bias, it cannot adequately capture discrimination.

Analyzing actual interview evaluations or archived recruitment documents (e.g., Petersen and Togstad, 2006; Rivera and Owens, 2021) or anonymous recruitment processes (e.g., Åslund and Skans, 2012; Goldin and Rouse, 2000), represents another approach to detect discrimination. Yet, this method has selection issues that decrease the external validity of the studies: while one can only study workplaces and companies that agree to the research and maintain a detailed record of recruitment documents, the representativity of them is questionable. Therefore, while this approach can result in interesting knowledge about particular (types of) employers, it does not capture discrimination in the labor market at large.

Another technique to investigate discrimination is the use of surveys and qualitative interviews. One may pose questions about discriminatory behavior or attitudes to employers and recruiters, or one may ask workers and job applicants about their potential experiences related to discrimination. Yet, in a real recruitment situation, employers probably do not act in accordance to what they say (cf., Gaddis, 2018; Pager and Quillian, 2005). At the same time, job candidates often do not know whether they have actually been discriminated, and while studying their subjective experiences, or beliefs, related to potential recruitment discrimination can be of interest, it does not yield reliable evidence of discrimination acts. Earlier discrimination studies show that the actual extent of discrimination faced by the minority group often comes as a shock to the testers representing the group (Bovenkerk, 1992).

The limitations of the aforementioned approaches in measuring labor market discrimination are obvious, and the use of experimental methods to examine discrimination has become increasingly common in the past decade or so. A common technique to provide information on the causal relationships of discrimination is the use of laboratory experiments where researchers arrange different settings to study participant behavior, e.g., trust based on ethnicity (e.g., Ahmed, 2010) and wage-setting based on productivity and employer's gender (Dickinson and Oaxaca, 2009). The shortcoming of laboratory experiments is their typically low external validity, generalizability, due to the lack of real-life context. Thus, there is no guarantee that the results produced under controlled laboratory conditions would apply in the

empirical reality, i.e., the labor market outside the laboratory. This criticism about the external validity relates also to the laboratory experiment conducted in Study III in this dissertation.

The popularity of using field experiments in studying labor market discrimination has grown substantially (Gaddis, 2018), possibly because of the shortcomings of the other methods, as discussed above. Audit studies, including in-person audits (sometimes referred as situation tests, using real test persons) and correspondence audits (i.e., correspondence tests using written applications), represent the main type of field experiment used to study discrimination in the hiring process. An advantage of this technique is that through the real-world context, while still allowing a high level of control to the researcher, it provides more direct evidence of the extent of discrimination, in comparison to the other methods (Gaddis, 2018).

Audit testing was introduced in the 1960s in the United Kingdom and the United States to study discrimination based on race and ethnicity. Traditionally, an audit study entails two testers, one representing the majority group and the other the minority group, who are matched as closely as possible, in terms of qualifications, except for the majority and minority status. The testers then apply for real jobs and attend the interviews whenever invited to one. Discrimination is measured based on the success of each candidate in the hiring process, and any differential treatment of one group of candidates is considered as evidence of discrimination.

Audit studies, and especially in-person audits, have received their share of criticism. Heckman and Siegelman (1993) are highly skeptical to the possibility of the testers being perfectly matched in all aspects that may affect productivity. This is to say, that there may be unobserved characteristics unknown to the researcher but known to the employers, and it is also possible for the testers to unconsciously or consciously influence the outcomes as they are aware of the purpose of the study. Heckman (1998) argues that the extent of discrimination found in audits is determined by how these unobserved characteristics are distributed among the minority and majority groups and on the chosen standardization level, i.e., whether the level of qualifications assigned to the testers increases or decreases the probability of the tester being offered a job. He thereby argues against the careful matching of the testers because it can lead to a biased estimate of discrimination. In addition, Heckman finds the representativeness of audit studies questionable, mainly because the audits have focused on low skilled jobs applied by over-qualified undergraduate students. The audit studies also ignore the importance of social networks by only including advertised positions (Heckman 1998; Heckman and Siegelman 1993). In conclusion, audit studies, and in-person audits in particular, can detect discrimination when it does not in fact exist and they can fail to detect discrimination even when it actually exists. This relates also to correspondence testing, and I will return to it.

Correspondence testing method

To evaluate whether actual employers hiring employees for real jobs discriminate by gender and parenthood as well as by ethnicity, fictitious applications were sent to real jobs in the labor market, and the callbacks, i.e., employer responses, for these were observed. This approach, correspondence testing, is a scientifically well-established method in the field of discrimination. It has been successfully applied in a number of Swedish and international studies (cf., Baert, 2018; OECD, 2008). In many correspondence tests the researcher sends two equal applications that represent the groups of interest for each vacancy, though in some studies more than two applications have been sent to the same vacancy. Correspondence tests have mainly been conducted in the labor market, to investigate hiring discrimination, but the method has also been applied to other markets, such as housing (cf., Gaddis, 2018). Correspondence tests have been used to study discrimination based on ethnicity, race, gender, age, sexual orientation, social class, personality, and attractiveness, among others (cf., Baert, 2018). But with the exception of Correll, Benard and Paik (2007) and Petit (2007), the method had not been used to study potential discrimination by gender combined with parenthood, prior to conducting Study I. Yet, since then, some additional field experiments have studied discrimination based on parenthood (cf., Baert, 2018; Hipp, 2020) as well as potential fertility (Becker, Fernandes and Weichselbaumer, 2019; Li, Wen, Ye and Yu, 2022). Also, a few field experiments have examined the role of recruiter gender for gender or ethnic discrimination, prior to Study II (Booth and Leigh, 2010; Carlsson, 2011) and Study IV (Edo et al., 2019; Carlsson and Eriksson, 2019 [published after Study II]).

An obvious advantage of correspondence testing is that it allows for a real-world context – which provides external validity to the study – as real employers make real decisions based on “real” job applications that can be manipulated in such a way that characteristics other than the ones in focus, e.g., gender, parenthood and ethnicity, can be held constant. Another benefit with the method is the high degree of control achieved by using written applications rather than real persons. It minimizes the problems regarding the existence of unobservable variables, leading to a more objective measure of discrimination. Also, the high controllability allows for more variation in other applicant-related characteristics and helps to avoid the measured effect from being a very specific local average treatment effect – as could be the case if all characteristics were held constant – thereby contributing to a higher external validity of correspondence studies.

In comparison to audit studies, correspondence testing provides a much higher degree of representativeness and generalizability, partly because a range of different occupational categories with varying qualifications and education level can be studied simultaneously, but also because a greater number of jobs can easily be applied for in geographically different locations.

The extent of discrimination is measured directly based on employers' responses to the (fictive) applicants, which illustrates another advantage of the method, namely the simplicity of the experimental design and the easiness of interpreting the results (Jackson and Cox, 2013).

Critique of correspondence testing

Despite the obvious benefits of correspondence testing in studying discrimination, the method also has its limitations. One critique of the method is related to ethical considerations, namely the lack of informed consent from the participants. This will be discussed in more detail in a separate section below. Correspondence testing is also limited in its ability to produce an accurate measure of the *extent* of discrimination. The applicant characteristics that are signaled (and varied), number and types of occupations observed, and the geographical context, among others, influence the extent to which one measures a local effect. Also, correspondence testing only measures discrimination in the initial stage but not later in the hiring process, and thereby the final hiring decisions made by the employers remain unknown to the researcher. However, on the one hand, one could argue that this is not a major issue as the audit studies conducted by the International Labour Organization (ILO) show evidence of ethnic discrimination in Sweden being more common in the first stage, i.e., selecting applicants to an interview (Attström, 2007), than in later stages, and results from France (Cediey and Foroni, 2008) and Italy (Allasino et al. 2004) indicate that nearly 90 percent of the discrimination against ethnic minorities takes place in this stage. Yet, on the other hand, according to a recent meta-study, ethnic and racial discrimination appears to be stronger later in the process when job offers are made than early on when making callbacks (Quillian, Lee and Oliver, 2020). This hints that the correspondence testing method may underestimate discrimination.

Regarding the gender of the recruiter, the person responsible for the recruitment choices cannot with certainty be established, and several persons may be involved in the recruitment process. Yet, based on the callback contents, the person who is listed in the job announcement appears to be active in the recruitment decisions and is often the same person who responds the applicant (Erlandsson, 2019). Moreover, discrimination may also take place within an organization or a workplace, e.g., an employee might, because of gender, parenthood or ethnicity, be moved to another position, face layoffs or be treated negatively as regards wage setting, promotion and internal training opportunities (e.g., Petersen and Saporta, 2004). These potential types of discrimination are not observed by using correspondence testing.

One may also argue that another disadvantage of the method is that it only allows for applying for jobs that accept written applications, e.g., only announced vacancies can be applied for. Yet, written job applications are practically the norm in Sweden. Still, other routes to employment, such as social

networks and walk-ins, are ignored. Yet another critique of correspondence testing is that, while it answers the question of whether discrimination occurs, it does not adequately answer the question why discrimination takes place. The method, and design, adopted here does not provide much information about the fundamental causes of discrimination. In order to shed light on the causes of discrimination, a multi-method approach can be argued to be of importance (Bursell, 2012; Correll et al., 2007; Midtbøen and Rogstad, 2012). Thus, in an attempt to take a glimpse at the dynamics behind recruitment decisions, in this thesis the correspondence studies are complemented with a laboratory experiment.

Measuring discrimination

One should be aware of what is actually being measured in studies investigating discrimination based on characteristics that are concealable to some extent – such as parenthood – as correspondence testing does not allow for a clear distinction between discrimination against job applicants who are parents and job applicants who openly indicate parenthood (Tilcsik, 2011). As employers have only the information provided in the application documents, a correspondence test can solely detect discrimination against parent applicants who indicate their parental status in the application documents. Therefore, when speaking about discrimination based on parental status, i.e., motherhood or fatherhood, in connection to the correspondence test in Study I, what we actually measure is discrimination based on disclosed parenthood (of 31 years old job applicants) instead of parenthood generally. Thus, when signaling parental status and ethnicity in the application documents, there is a risk that the indicators used may add confounding factors to the study. In race discrimination studies, the use of distinct African-American names may be associated with lower social class, thereby making it difficult to distinguish between effects of race and socioeconomic status (Pager, 2007). Thus, one needs to be aware of the indicators of parental status and ethnicity, and preferably use as neutral indicators as possible if the aim is to avoid any additional (potentially negative) perceptions regarding any of the applicant groups. While the Swedish names used here can be considered neutral in terms of class, it is more difficult to evaluate the foreign-sounding names in this regard, and there may be a risk of confounding factors.

One application

Only one application is sent to each vacancy (for an example of a fictitious job application, see Appendix in Study I). The characteristics of the applicant, i.e., gender, parenthood status, ethnicity, are randomly assigned to each application. An alternative approach would be to send two or more applications (e.g., male/female*parent/nonparent) to each employer. A feature of

this paired, or matched, procedure is that it generates an intra-employer variation in the evaluation of the applicants. Still, the unpaired design, in which only one application is submitted, is preferred here. Some of the advantages of the unpaired procedure are that it minimizes employer costs in effort and time (which can be considered ethically more favorable), enables the use of identical applications, rather than using applications with equivalent merits, and reduces the risk of being exposed by the employers (cf., Vuolo, Uggen and Lageson, 2018). In addition, the procedure of using only one application avoids the risk of spillover effects (cf., Phillips, 2019).

There are only two possible outcomes when sending one application for each job, i.e., the applicant receives a callback or the applicant does not receive a callback. However, it should be noted that there are different types of employer responses (that are defined as callbacks), e.g., an invitation to an interview, a job offer, and an inquiry for more information, but the type of response does not seem to matter (cf., sensitivity analyses in Study I; Bygren and Gähler, 2021). If two applications were sent to each vacancy, the number of possible outcomes regarding parenthood would be increased to four: none of the applicants receives a callback, both applicants receive a callback, only the childless applicant receives a callback and only the applicant with children receives a callback. Consequently, contrary to sending two or more applications for each vacancy, where discrimination by individual employers may potentially be detected, discrimination by individual employers can never be detected when only sending one application. Regardless of only sending one application for each vacancy, it is still possible to calculate the relative callback rate the same way other studies (that send several applications) do, i.e., by dividing all callbacks for group A applicants (e.g., parent/ethnic majority) with all callbacks for group B applicants (non-parent/ethnic minority).

Ethical considerations

Field experiment

The method, i.e., correspondence testing, involves ethical concerns. There are three main ethical issues related to the use of the method. The first is the lack of informed consent. The employers, and their representatives, will not be given the opportunity to decide whether they want to participate in the study, and they will not be informed of the purpose of the study. Second, the method generates a cost in terms of time and resources that the employers use to handle a fictitious job application. The third problem is that employers are led to believe that the fictitious application documents are authentic, and the consequences that this may have. Thus, strong arguments are required in

order to justify such deviations from general principles regarding research ethics.

The first problem, i.e., the lack of informed consent, is inherent in the study design. If employers were informed about the purpose of the study, the employer behavior regarding discrimination that the researcher is interested in observing would not be informative. In order to be able to observe actual discrimination, it is not possible to inform employers that they are being observed, as those who may practice discrimination would then probably change their behavior. Thus, deviating from the ethical principle of obtaining informed consent should be weighed against the potential gain in terms of knowledge that is reached by conducting research using correspondence testing (cf., Swedish Research Council, 2002, for ethical principles concerning social sciences).

The second problem is the costs for employers, in terms of time and resources, due to dealing with fictitious applications. These costs are relatively small for individual employers, but the total cost for all employers must be weighed against the scientific and societal benefits of the research.

The third problem, the employers are being misled by a researcher, has two aspects. First, even if the risk of discovery for using the method is very small, there is still a risk that an individual employer will discover that the application (s)he received is fictitious. Second, when the study results based on the method are published, the method of data collection is revealed to a wider public. Employers' and the general public's trust in researchers can potentially be damaged when it becomes known that researchers mislead employers in the aforementioned way. Yet, to my knowledge, disseminating the results of studies presented here, and of other studies using similar methodology, in academic journals and mass media forums, has not led to any general harm for or reduced confidence in researchers.

Moreover, prior to beginning the data collection, and later to pool data with a previous data collection, ethical applications were submitted to The Regional Ethics Review Board in Stockholm (dnr. 2013/1237-31/5) and the Swedish Ethical Review Authority (dnr: 2019-01220) which decided not to review the applications. The board stated, in both cases, that the research project does not include the kind of processing of personal data that according to the Ethical Review Act is sensitive to the research persons, and that the research is therefore not covered by the Ethical Review Act. According to the accompanying statement of opinion, the committee's view was that there were no obstacles to carrying out the research. However, in the first case, one member (out of fourteen) objected to the decision, given the nature of the information collected and the way in which the information was obtained. These rulings are in line with a previous ruling made in 2006 regarding another research project where the correspondence testing method was used to study ethnic discrimination in the Swedish labor market (see Bursell, 2012).

The employers and their representatives will most likely not suffer any physical or mental damage or experience any pain, discomfort or invasion of privacy in connection with the research. Yet, the employers are likely to make a small loss of time and resources when handling the fictitious job applications. But for an individual employer the loss can be argued to be very small in absolute terms, and, in particular, in relation to the social and scientific benefit generated by the research in terms of knowledge gains about possible discrimination, if any, based, e.g., on gender, parenthood and ethnicity, in the labor market.

Several measures were taken to minimize the potential employer costs in time and resources, and the risk of damage to the employer. First, only one fictitious application is sent to each vacancy. This approach thereby differs from many earlier studies using correspondence testing method that send two or more applications per vacancy. Yet, recently, more and more studies appear to use the unpaired design, and send only one application per vacancy (cf., Larsen, 2020; Vuolo et al., 2018). Sending one job application, instead of two (or more) applications, to each employer means not only (at least) one application less for the employer to administer but it also implies that the employer does not need to make a comparison between two (or more) fictitious applications with equal merits, which obviously saves time. Second, the employers who contact the fictitious applicant via email or telephone will be promptly informed that the applicant is no longer available for the position. Third, employers remain anonymous, and the results are reported only in aggregate form. Individual employers can therefore not be identified when the results are published. Fourth, because only one application is sent to each vacancy, and information about other applicants is missing, any potential discrimination among individual employers cannot be identified. The purpose of the study is to detect general discrimination patterns, rather than to point out individual employers. Thus, the collected data can only be used to identify discrimination patterns in the Swedish labor market, and in certain occupations or occupational categories, in general. This approach further reduces the risk of damage to the employers who are included in the study.

While the research project brings no direct benefits for the employers, companies and organizations included in it, there is an indirect benefit of the project. If discrimination based on, e.g., applicant gender, parental status, ethnicity and/or recruiter gender takes place in the Swedish labor market, it indicates that the labor market does not function effectively, i.e., the right person may not end up in the right place and the most qualified candidate may not be selected for the job, which entails costs for both employees and employers. It should therefore be in the interest of all employers to detect and measure (potential) discrimination in recruitment. This type of research may also lead to employers becoming more aware of their behavior, and unintentional discrimination, which may lead to better recruitment decisions in the future.

Laboratory experiment

The laboratory experiment (Study III) – which here entails participants evaluating job applications in a controlled non-authentic recruitment setting, in contrast to field experiments studying job applications to announced vacancies in the real-world – is also related to ethical considerations. The experiment consists of student participants and the experiment took place in connection to two lectures at Stockholm University. The participants were informed about the experiment prior to deciding whether to participate in the experiment or not, they were notified that participation is voluntary, would not in any way affect their course grade (the researchers conducting the study have no direct connection to the course nor the participants), and they can at any point interrupt or withdraw their participation without providing an explanation. The students received an information sheet, and if they decided to participate after having read the information, it was considered as informed consent. In order to avoid documenting the names of the participants, the participants were not asked to sign an informed consent form. Age and gender are the only personal information that was asked about the participants. Thus, the identity of the participants is not revealed at any step of the study. Nonetheless, in order not to jeopardize the purpose of the experiment, the participants could not be informed that discrimination was the topic of interest of the study. This can be perceived as deception of the participants. Instead, the participants were told that recruitment decisions were the topic of the study. This ethical concern is, again, counterbalanced by the fact that the experiment could not be conducted if the participants were informed about discrimination, based on gender and parenthood in particular, being the main interest of the study. Nevertheless, it may be considered highly unlikely that the participants will experience any harm due to participation in the experiment. We submitted an application to the Regional Ethical Review Board in Stockholm in 2016 (dnr. 2016/1736-31/5) prior to conducting the laboratory experiment. The board decided not to review the application: it stated that the study was not covered by the Ethical Review Act and it did not see any obstacles to carrying out the research.

Summary of the four empirical studies

All the four empirical studies in the dissertation use experimental methods to examine labor market discrimination in Sweden. Study I, Study II and Study IV are based on a field experiment, i.e., correspondence testing, in which fictitious job applications are sent to real announced vacancies. Thus, three of the studies in this dissertation use large-scale field experiment data, comprising up to 5,641 job applications and vacancies in 20 occupations (Study IV)

that vary in terms of gender and immigrant composition, sector, and educational qualifications. While Study I focuses on discrimination on the basis of gender and parenthood, or any combination thereof, Study II examines the role of recruiter (employer) gender for gender discrimination in recruitment. Study IV studies the role of recruiter gender for ethnic and gender discrimination in recruitment. Study III examines recruitment discrimination based on gender and parenthood in a laboratory setting rather than by relying on field experiment data as the other studies.

Given the void in previous discrimination research regarding parenthood, motherhood and fatherhood, in Study I (Bygren, Erlandsson and Gähler, 2017), we ask whether mothers, or fathers, are discriminated against when applying for a job. Based on the callbacks (i.e., employer responses) received, no evidence of discrimination based on gender and parenthood, or any combination of these, is shown in the early stage of the hiring process, i.e., when job applications are submitted to vacant positions, in the Swedish labor market. This applies both to less and highly qualified occupations.

While few prior field experiments have focused on recruiter gender, Study II (Erlandsson, 2019) investigates whether recruiter gender plays a role for (potential) gender discrimination. The results show that male recruiters overall contact male applicants more often than female applicants. This is especially the case in gender-balanced occupations (occupations in which the proportion of women is 40% to 60%). However, while the callbacks by male recruiters are higher for male than female applicants in all categories based on the occupational gender composition, no statistically significant gender biases in callbacks by male or female employers are found in male- and female-dominated occupation. Although female employers in general are found to contact female job applicants slightly more often than male applicants, this result does not reach statistical significance. Therefore, the study shows no clear evidence of gender discrimination by female recruiters.

Study III aims at examining potential discrimination in recruitment evaluations based on gender and parenthood, and how recruitment decisions (by human relations and business students) are made. The study shows a statistically significant gender bias in job applicant ratings in favor of female applicants in a laboratory setting. In other words, female job applicants are more often than male applicants rated as the top candidate. Contrary to previous US-based laboratory experiments, the study shows no motherhood penalty in the applicant ratings. Moreover, there is a statistically significant pro-female bias in applicant ratings by female evaluators, both in the ratings made by the evaluators individually and collectively by all-female evaluation groups.

Given the scarcity of previous field experiments on the topic, Study IV is focused on the role of recruiter gender for ethnic and *gendered* ethnic discrimination in the Swedish labor market. While the study shows evidence of ethnic discrimination against foreign-named job applicants by both male and female recruiters, foreign-named male applicants are contacted significantly

less often than foreign-named female applicants. The results show gendered ethnic discrimination by female recruiters, i.e., female recruiters contact foreign-named female applicants more often than foreign-named male applicants, in particular in highly qualified occupations. Yet, male recruiters favor female candidates with foreign-sounding names over male candidates with foreign-sounding names in male-dominated occupations in particular. Thus, taken together, the main conclusion from Study II and Study IV about the role of recruiter gender for recruitment discrimination is that it appears to matter, at least in certain occupational contexts.

Although the results from Study I, II, III and IV may at first glance appear somewhat diverse, I argue that the studies complement each other. First, the overall callback rates between Study I, II and IV using field experiment data range from 36,7% (Study IV) to 49,0% (Study II), mostly because of differences in the study designs, i.e., differences regarding the proportion of foreign-named applicants, number of occupations included, and exclusion of observations for which recruiter gender was not documented. Yet, the overall gender difference in callbacks remains small and not statistically significant in Study I, II and IV.

Concluding Discussion

If discrimination against job applicants of the underrepresented gender occurred in gender-concentrated occupations, it could contribute to gender segregation in the labor market. Yet, the findings from this dissertation lend little support for gender discrimination in recruitment as an important contributing factor for gender segregation in the Swedish labor market. This suggests that alternative mechanisms, other than the demand side, are important in explaining gender inequality in the labor market, and that gender disparities in labor market outcomes are likely to originate largely in supply side factors. These include, e.g., horizontal segregation generated by the supply side and individuals' gender-specific choices related to work and family, and how to combine these. Yet, the evidence presented here support the notion of recruitment discrimination as *one* mechanism behind ethnic inequalities in labor market outcomes. Also, the findings suggest that recruitment discrimination contributes to the labor market inequality of ethnic minority men in particular, whereas other mechanisms may be of greater importance for the labor market inequality of ethnic minority women.

The results from the studies presented here indicate that recruiter gender, and thereby gender diversity among recruiters, is important for the success of male and female applicants, and in particular for the success of foreign-named male and female applicants in the recruitment process, at least in certain occupational contexts. The findings suggest that female recruiters may contribute to vertical segregation of ethnic minority men by disfavoring

them (over ethnic minority women) in highly qualified occupations. At the same time, recruiter gender does not appear to be a crucial factor in explaining gender and ethnic segregation or inequalities in labor market outcomes at large. However, other recruiter characteristics that are not observed here, i.e., recruiter ethnicity, may be important, as a small-scaled French field experiment of accounting jobs indicates (Edo et al., 2019). Thus, in future research, it would be interesting to study a potential in-group bias based on ethnicity, as well as the interaction of gender and ethnicity, in recruitment in the Swedish labor market.

The main method of study in this dissertation, i.e., a field experiment using correspondence testing, is a well-established method in discrimination research. Yet, the correspondence audits here (like audit studies in general) are restricted in terms of, e.g., occupations, applicant characteristics (i.e., 31-year-olds with certain names and backgrounds, among others), and geographic contexts (mainly jobs in and near the three largest Swedish cities) that were observed, and thereby the studies measure somewhat local effects. Thus, the studies here might produce biased estimates of labor market discrimination in Sweden in general. However, considering the large number and different types of occupations, in several geographic regions, the field experiment data here cover a considerable segment of the labor market and are, therefore, relatively representative of the (Swedish) labor market, especially in comparison to numerous other field experiments that are much more limited in scope. Therefore, the studies are, nevertheless, likely to provide a rather decent indicator of discrimination (based on gender, parenthood and ethnicity) in the Swedish labor market at large.

Study I does not show any evidence of discrimination based on gender or parenthood, or any combination of these, in general or in highly or less qualified occupations. Yet, Study II shows discrimination in favor of male applicants by male recruiters, especially in gender-balanced occupations. Thus, while the overall callback patterns for men and women, in general, are similar in Study I, II and IV, gender gaps in callbacks appear when different dimensions, i.e. recruiter gender, applicant ethnicity and separate occupational categories, are examined. Thus, although no general conclusions regarding *specific* occupations can be drawn here, one main conclusion to be made based on Study I, II and IV is that the occupational type matters, and thereby also the occupational context in general. Hence, the callback rates as well as the level and direction of discrimination are shown to differ by occupational category, and between specific occupations. There may be various occupation-specific reasons for this. It may have to do with the characteristics of the jobs, views on female- and male-typed jobs, gender composition in the workplace, supply and demand factors related to particular occupations, a combination of these factors or something else. Yet, the findings presented in this dissertation suggest that gender diversity among recruiters, or the lack of it, can be *one* factor in explaining some of the variation in discrimination

patterns across occupational types. Moreover, studying discrimination in and between specific occupations would be of great interest in future research. If occupation-specific characteristics could be identified, then these characteristics could potentially increase the understanding of why employer discrimination occurs.

Also, the result from Study III that female job candidates appear to be favored in a laboratory setting, and especially by female participants, can be interpreted in light of the occupational context, i.e., the specific position of an accounting manager. At first glance, one may think that this result contradicts the findings from Study I and Study II. Yet, if one looks at the callback rates for accountant/auditor positions in Study I, it appears that female applicants receive more callbacks than male applicants, although this difference is small and not statistically significant. Also, finding a pro-female bias in job applicant ratings by female recruiters in Study III can be seen to be partly in line with the callback patterns in Study II where female recruiters contact female applicants somewhat more often than male applicants, particularly in female-dominated and in gender-balanced occupations, although these differences are not statistically significant.

In addition, while one may consider the results from Study III in light of the age of the participants (being relatively young as they are undergraduate students), most of whom are lacking practical experience of recruiting, one may also interpret the results as representing the recruitment preferences of a somewhat younger generation than most recruiters currently in the labor market. Thus, it will be interesting to see if this pattern of recruitment preferences will remain in the future in the labor market outside the laboratory setting.

Besides a few concrete statements where participants consider parental status or gender, Study III shows no explicit evidence of gender or parental status being taken into consideration in the evaluation of job applicants. Thus, in the end, the study yields little information about the reasoning behind the applicant ratings, and potential mechanisms behind gender and ethnic discrimination. Therefore, in order to answer the “why question” of discrimination, this dissertation calls for more research about the mechanisms behind recruitment discrimination. For future research, this illustrates the importance of innovative experimental designs, combined with large-scaled field experiments, rather than field experiments that focus only on very limited occupational and geographical contexts.

Finally, this dissertation is important in raising both knowledge and awareness of recruitment discrimination, and in particular the gendered nature of ethnic discrimination, in the Swedish labor market as well as the importance of considering gender diversity among recruiters in the workplace and across the occupational distribution.

References

- Ahmed, A. (2010). What is in a surname? The role of ethnicity in economic decision making. *Applied Economics* 42(21):2715-2723.
- Allasino, E., Reyneri, E., Venturini, A., and Zincone, G. (2004). *Labour Market Discrimination against Migrant Workers in Italy*. International Migrations Papers 67. Geneva: International Labour Office.
- Åslund, O., and Skans, O. N. (2012). Do anonymous job application procedures level the playing field? *ILR Review*, 65(1), 82-107.
- Attström, K. (2007). *Discrimination against Native Swedes of Immigrant Origin in Access to Employment*. International Migrations Papers 86E. Geneva: International Labour Office.
- Baert, S. (2018). Hiring discrimination: An overview of (almost) all correspondence experiments since 2005. In Gaddis, S. M. (Ed.), *Audit studies: Behind the scenes with theory, method, and nuance* (pp. 63–77). Cham, Switzerland: Springer.
- Becker, G. (1971). *The Economics of Discrimination*. Chicago, IL: University of Chicago Press.
- 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.
- Bielby, W., and Baron, J. (1986). Men and women at work: Sex segregation and statistical discrimination. *American Journal of Sociology*, 91, 759-799.
- Blau, F.D. (1984). Occupational segregation and labor market discrimination. Pp. 117-143 in *Sex Segregation in the Workplace: Trends, Explanations, Remedies*, edited by B.F. Reskin. Washington, D.C: National Academy Press.
- Blau, F. D., and Kahn, L. M. (2017). The gender wage gap: Extent, trends, and explanations. *Journal of Economic Literature*, 55(3), 789-865.
- Booth, A., and Leigh, A., (2010). Do employers discriminate by gender? A field experiment in female-dominated occupations. *Economics Letters*, 107(2), 236-238.
- Bovenkerk, F. (1992). *Testing Discrimination in Natural Experiments: A Manual for International Comparative Research on Discrimination on the Grounds of Race and Ethnic Origin*. Geneva, International Labour Office.
- Brewer, M. B., and Kramer, R. M. (1985). The psychology of intergroup attitudes and behavior. *Annual Review of Psychology*, 36(1), 219–243.
- Bursell, M. 2012. *Ethnic Discrimination, Name Change and Labor Market Inequality: Mixed Approaches to Ethnic Exclusion in Sweden*. PhD Dissertation, Department of Sociology, Stockholm University.
- 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(3), 337-348.
- Bygren, M., and Gähler, M. (2012). Family formation and men's and women's attainment of workplace authority. *Social Forces*, 90, 795-816.
- Bygren, M., and Gähler, M. (2021). Are women discriminated against in countries with extensive family policies? A piece of the 'welfare state paradox' puzzle from Sweden. *Social Politics: International Studies in Gender, State & Society*, 28(4), 921–947.
- Carlsson, M. (2011). Does hiring discrimination cause gender segregation in the Swedish labor market? *Feminist Economics* 17, 71-102.

- Carlsson, M., and Eriksson, S. (2019). In-group gender bias in hiring: Real-world evidence. *Economics Letters*, 185, 108686.
- Cediey, E., and Foroni, F. (2008). Discrimination in access to employment on grounds of foreign origin in France. *International migration papers* 85E, Geneva, International Labour Office.
- Charles, M. (2011). A world of difference: International trends in women's economic status. *Annual Review of Sociology*, 37, 355-371.
- Correll, S. J., and Benard, S. (2006). Biased estimators? Comparing status and statistical theories of gender discrimination. *Advances in Group Processes*, 23, 89-116.
- Correll, S. J., Benard, S. and Paik, I. (2007). Getting a job: Is there a motherhood penalty? *American Journal of Sociology*, 112, 1297-1338.
- Cunningham, W. A., Johnson, M. K., Raye, C. L. Gatenby, J. C., Gore, J. C., and Banaji, M. R. (2004). Separable neural components in the processing of black and white faces. *Psychological Science*, 15, 806-813.
- Dickinson, D. L., and Oaxaca, R. L. (2009). Statistical discrimination in labor markets: An experimental analysis. *Southern Economic Journal*, 76(1), 16-31.
- Discrimination Ombudsman. (2021). Arbetslivet [Work life]. Retrieved from <https://www.do.se/diskriminering/diskriminering-olika-delar-samhallet/diskriminering-pa-jobbet>
- Duvander, A. Z., and Johansson, M. (2012). What are the effects of reforms promoting fathers' parental leave use? *Journal of European Social Policy*, 22(3), 319-330.
- Edlund, J., and Öun, I. (2016). Who should work and who should care? Attitudes towards the desirable division of labour between mothers and fathers in five European countries. *Acta Sociologica*, 59(2), 151-169.
- Edo, A., Jacquemet, N., and Yannelis, C. (2019). Language skills and homophilous hiring discrimination: Evidence from gender and racially differentiated applications. *Review of Economics of the Household*, 17(1), 349-376.
- England, P. (1992). *Comparable Worth: Theories and Evidence*. Hawthorne, NY: Aldine de Gruyter.
- Erlandsson, A. (2019). Do men favor men in recruitment? A field experiment in the Swedish labor market. *Work and Occupations*, 46(3), 239-264.
- European Institute for Gender Equality. 2017. *Gender Equality Index 2017. Measuring Gender Equality in the European Union 2005–2015*. Vilnius: European Institute for Gender Equality.
- Gaddis, S. M. (2018). An introduction to audit studies in the social sciences. In S. M. Gaddis (Ed.), *Audit studies: Behind the scenes with theory, method, and nuance* (pp. 3–44). Cham, Switzerland: Springer.
- 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.
- Goldin, C., and Rouse, C. (2000). Orchestrating impartiality: The impact of "blind" auditions on female musicians. *American Economic Review*, 90(4), 715–741.
- Grönlund, A., Halldén, K., and Magnusson, C. (2017). A Scandinavian success story? Women's labour market outcomes in Denmark, Finland, Norway and Sweden. *Acta Sociologica*, 60(2), 97-119.
- Harkness, S., and Waldfogel, J. (2003). The family gap in pay: Evidence from seven industrialized countries. *Research in Labor Economics*, 22, 369-413

- Heckman, J. (1998). Detecting discrimination. *The Journal of Economic Perspectives*, 12(2), 101-116.
- Heckman, J., and Siegelman P. (1993). The Urban Institute Audit Studies: Their Methods and Findings. In Fix, M. and Struyk R. J. (Eds.), *Clear and Convincing Evidence: Measurement of Discrimination in America*. Washington, DC: Urban Institute Press
- Hewstone, M., Rubin, M., and Willis, H. (2002). Intergroup bias. *Annual Review of Psychology*, 53(1), 575-604.
- Hipp, L. (2020). Do hiring practices penalize women and benefit men for having children? Experimental evidence from Germany. *European Sociological Review*, 36(2), 250-264.
- Jackson, M., and Cox, D. R. (2013). The principles of experimental design and their application in sociology. *Annual Review of Sociology*, 39(1), 27-49.
- Kandel, D. B. (1978). Homophily, selection, and socialization in adolescent friendships. *American Journal of Sociology*, 84(2), 427-436.
- Larsen, E. N. (2020). Induced competition in matched correspondence tests: conceptual and methodological considerations. *Research in Social Stratification and Mobility*, 65, 100475.
- Li, X., Wen, D., Ye, L., and Yu, J. (2022). Does China's fertility policy induce employment discrimination against women? Evidence from a nationwide correspondence experiment. Available at Research Square: <https://doi.org/10.21203/rs.3.rs-1161948/v1>
- 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, 1910-49.
- McPherson, M., Smith-Lovin, L., and Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, 27(1), 415-444.
- Midtbøen, A. H., and Rogstad J. (2012). Discrimination. Methodological controversies and sociological perspectives on future research. *Nordic Journal of Migration Research*, 2(3), 203-212.
- OECD. (2008). The Price of Prejudice: Labour Market Discrimination on the Grounds of Gender and Ethnicity. Legal and Institutional Framework in Force in 2007. Chapter 3 in *Employment Outlook 2008*. Paris: OECD.
- OECD. (2014). *Finding the way: A discussion of the Swedish migrant integration system*. Paris: OECD publishing.
- OECD. (2017). *OECD employment outlook 2017*. Paris, France: OECD Publishing. https://doi.org/10.1787/empl_outlook-2017-en
- OECD. (2021). *OECD Economic Surveys: Sweden 2021*. OECD Publishing, Paris. <https://doi.org/10.1787/f61d0a54-en>
- Tajfel, H., Billig, M. G., Bundy, R. P., et al. (1971). Social categorization and intergroup behaviour. *European Journal of Social Psychology*, 1(2), 149-178.
- Pager, D. (2007). The use of field experiments for studies of employment discrimination: Contributions, critiques, and directions for the future. *The Annals of the American Academy of Political and Social Science*, 609(1), 104-133.
- Pager, D., and Shepherd, H. (2008). The sociology of discrimination: Racial discrimination in employment, housing, credit, and consumer markets. *Annual Review of Sociology*, 34, 181-209.
- Pager, D., and Quillian, L. (2005). Walking the talk? What employers say versus what they do. *American Sociological Review*, 70(3), 355-380.

- Petersen, T., and Saporta, I. (2004). The opportunity structure for discrimination. *American Journal of Sociology*, 109(4), 852-901.
- Petersen, T., and Togstad T. (2006). Getting the Offer: Sex Discrimination in Hiring. *Research in Social Stratification and Mobility*, 24(3), 239–257.
- 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.
- Phillips, D. C. (2019). Do comparisons of fictional applicants measure discrimination when search externalities are present? Evidence from existing experiments. *The Economic Journal*, 129(621), 2240-2264.
- Ridgeway, C. (2011). *Framed by Gender: How Gender Inequality Persists in the Modern World*. New York: Oxford University Press.
- Ridgeway, C., and Correll S. J. (2004). Motherhood as a Status Characteristic. *Journal of Social Issues*, 60(4), 683–700.
- Ridgeway, C., and Smith-Lovin, L. (1999). The Gender System and Interaction. *Annual Review of Sociology* 25, 191-216.
- Rivera, L. A., and Owens, J. (2021). Glass floors and glass ceilings: Sex homophily and heterophily in job interviews. *Social Forces*, 99(4), 1363-1393.
- Rooth, D-O. (2010). Automatic Associations and Discrimination in Hiring: Real World Evidence. *Labour Economics*, 17, 523-534.
- Sandberg, A. (2018). Competing identities: a field study of in-group bias among professional evaluators. *The Economic Journal*, 128(613), 2131-2159.
- Statistics Sweden. (2020). Utrikes födda i Sverige [The foreign-born in Sweden]. Retrieved from <https://www.scb.se/hitta-statistik/sverige-i-siffror/manniskorna-i-sverige/utrikes-fodda/>
- Swedish Research Council. 2002. *Forskningsetiska principer inom humanistisk-samhällsvetenskaplig forskning*. Retrieved from <https://www.vr.se/analys/rapporter/vara-rapporter/2002-01-08-forskningsetiska-principer-inom-humanistisk-samhällsvetenskaplig-forskning.html>
- Swedish Social Insurance Agency. (2018). *Social Insurance in Figures 2018*. Stockholm: Swedish Social Insurance Agency.
- Tilcsik, A. (2011). Pride and Prejudice: Employment Discrimination against Openly Gay Men in the United States. *American Journal of Sociology*, 117(2), 586-626.
- United Nations Development Programme. (2016). Human development report 2016. Human development for everyone. New York, NY. Retrieved from http://hdr.undp.org/sites/default/files/2016_human_development_report.pdf
- Vuolo, M., Uggen, C., and Lageson, S. (2018). To match or not to match? Statistical and substantive considerations in audit design and analysis. In S. M. Gaddis (Ed.), *Audit studies: Behind the scenes with theory, method, and nuance* (pp. 119–140). Cham, Switzerland: Springer.