The Gender Employment Gap among Refugees and the Role of Employer Discrimination: Experimental Evidence from the German, Swedish and Austrian Labor Markets

Note: published as: Fossati, F., Knotz, C., Liechti, F., & Otmani, I. (2022, online first). The Gender Employment Gap among Refugees and the Role of Employer Discrimination: Experimental Evidence from the German, Swedish and Austrian Labor Markets. *International Migration Review*, 019791832211342. https://doi.org/10.1177/01979183221134274

Open access version is available on the journal's website.

Flavia Fossati[1], Carlo Knotz[2], Fabienne Liechti[3] and Ihssane Otmani[4]

[1] IDHEAP, University of Lausanne, nccr-on the move and NCCR LIVES, ORCID: 0000-0002-9218-5422

[2] University of Stavanger, ORCID: 0000-0003-0042-3065

[3] IDHEAP, University of Lausanne, ORCID: 0000-0001-5509-3353

[4] IDHEAP, University of Lausanne

Abstract

Compared to their male counterparts, refugee women exhibit low employment rates in many countries. Discrimination by recruiters is a possible explanation for this phenomenon, but there is so far little direct evidence on this. This study addresses this gap. We develop a set of hypotheses about the effects of gender and family status on refugees' labor market integration and then test these hypotheses using data from an original survey experiment administered in 2019 to online panels of recruiters in three major refugee-receiving countries (Germany, Austria, and Sweden). We find that recruiters in fact prefer female over male refugees across different job types, all else equal. However, we also find evidence of a disadvantage connected with motherhood among refugees. Overall, our findings raise doubts about the relevance of discrimination as an explanation for the employment gap between male and female refugees.

Keywords: Refugees; Labor Market Integration; Discrimination; Survey Experiment

Acknowledgments

This research was financed by the nccr-on the move IP7 (Grant number: 51NF40-182897), which is funded by the Swiss National Science Foundation. We are grateful to Anna Wilson for the translation of the questionnaires into Swedish. We extend our thanks to Jeremias Stadlmair and the INEX group at the Political Science Department of the University of Vienna for their feedback on the research questionnaire.

Introduction

Gainful employment is one of the central dimensions of immigrant integration into their host countries (Harder et al. 2018). Of course, simply having *a* job is not equivalent to successful integration (e.g., Ballarino and Panichella 2015), but *not* having a job will very likely impede other forms of social integration (e.g., Gallie 1999). Unfortunately, immigrants in many Western societies struggle to find employment, and this is particularly the case for refugees and other humanitarian immigrants (Hooijer and Picot 2015; de Vroome and van Tubergen 2010). In addition, within this already disadvantaged group, female refugees' employment rates tend to be significantly lower as those of their male peers (Cheung and Phillimore 2017; Bloch 2007). In this paper, we focus on this gender employment gap among refugees.

There are different explanations for why female refugees are less likely to be employed than their male peers. One potential explanation is that female refugees are being discriminated against by virtue of being refugees and immigrants (Zschirnt and Ruedin 2016), but also because they are women (González, Cortina, and Rodríguez 2019). This added disadvantage in recruitment experienced by refugee women might explain at least a part of the gap between refugee men and women. But the case for the "double burden" hypothesis is by far not open-and-shut, because it is conceivable that refugee women might actually have an *advantage* in recruitment over their male peers. This is because any negative attributes applied to refugees ("unreliable, untrustworthy"; e.g., Kotzur et al. 2019) are applied primarily to refugee *men* (Eagly and Kite 1987) but less so to women. In addition, stereotypes commonly attached to women (e.g., warmth, communality; see Ellemers 2018) could counteract negative stereotypes often applied to refugees. If that were the

case, and if refugee women accordingly hold an advantage over male refugees in recruitment, this would imply that the employment gap among refugees would result from other factors than discrimination in recruitment. Such other factors could be traditional beliefs about proper gender roles held in immigrant and refugee communities, which lead women in these communities to adopt the role of homemaker instead of seeking to join the labor force (Fernández and Fogli 2009; Koopmans 2016; but see also Breidahl and Larsen 2016), or institutional factors, such as a lack of affordable childcare and other supportive family policies in their destination countries (Dumper 2002).

Empirically, the question of whether discrimination or other factors are more important determinants of refugee gender employment gap is not yet settled. Existing research has largely relied on indirect evidence such as testimonies from refugees themselves (e.g., Bloch 2007; Dumper 2002), or survey or registry data, where employment gaps that cannot be attributed to differences in productivity can serve as a sign of discrimination (Bevelander 2011). An arguably superior approach is to study the perceptions and hiring behavior of recruiters directly, but this has only rarely been done (but see Vernby and Dancygier 2019; Lundborg and Skedinger 2016).

In this contribution, we provide new evidence on the role of discrimination for the gender employment gap among refugees. To do so, we first develop a set of hypotheses about the likely patterns of recruiter discrimination that refugees of different genders and with different family backgrounds might face, building on theories of discrimination and stereotypes from economics

and psychology. We then test these hypotheses using data from an original survey experiment administered to an online panel of recruiters from Austria, Germany, and Sweden in 2019.[1]

Our main finding is that—all else equal—female refugees are indeed *preferred* to their male peers. We also find that this female advantage disappears once children enter the picture. Without children, female refugees are preferred over their male peers (again all else equal); with children, female refugees are on par with their male peers. This finding is robust across two distinct types of jobs and the three countries.

Our findings have implications for our understanding of the mechanisms of refugee integration as well as public policy. For one, our findings suggest that recruiter discrimination against women might not actually be the main mechanism behind the lower employment rates of female refugees compared to male refugees. Our findings also suggest that any anti-discrimination initiatives to support refugees should not bracket out the situation of refugee *men*.

The remainder of this article proceeds as follows: the next section develops hypotheses about hiring discrimination against refugees based on a discussion of theories about stereotyping in recruitment.

[1] The recruiters who participated in this study were members of an incentivized online panel that is run by an international market and public opinion research firm (Qualtrics). More information on how the sample was selected is provided below.

The third section then presents our experiment, data collection, and estimation method. The fourth section presents the results and the final section concludes.

Hiring discrimination and the role of stereotypes

There is a wealth of evidence for the presence of discrimination against immigrants, women, and other groups in job recruitment as well as for the relevance of stereotypes as a driver of such discrimination (e.g., Zschirnt and Ruedin 2016; Ellemers 2018; Midtbøen 2014). It is therefore plausible—but not self-evident—that stereotyping is also the mechanism behind refugee women's low employment rates. In this section, we discuss different scenarios of how stereotypes could impact male and female refugees in recruitment procedures. We define stereotypes as oversimplified (and frequently false) expectations or beliefs about social groups and the attributes that are characteristic for them. Stereotypes can be positive ("Germans are punctual") or negative ("Immigrants are criminals") (see also Ellemers 2018; Cuddy et al. 2009; Bordalo et al. 2016).

The "double burden" hypothesis

Whatever labor market discrimination refugees in general are facing (Zschirnt and Ruedin 2016), it is possible that refugee women suffer an additional disadvantage connected to their female gender and the stereotypes attached to it (Ellemers 2018; Cejka and Eagly 1999). Such a "double burden" for female refugees could result from two related forms of stereotyping. In the first case, refugee women—because they are women—might be seen by recruiters as more likely to be less committed

to their jobs and less reliable than men due to current or potential future caring responsibilities. In addition, prescriptive stereotypes about the "proper role" of women in society might also lead recruiters to discriminate against female applicants in general (González, Cortina, and Rodríguez 2019; see also Phelps 1972). In both cases, stereotyping based on gender would create an additional disadvantage for female refugees compared to male refugees.

This leads us to expect that the employment gap between male and female refugees might result from the fact that female refugees are affected by negative stereotypes attached to both their gender and their immigration status, creating the type of "double burden" or "double jeopardy" that has also been observed for women from other disadvantaged groups (King 1988; Taylor, Charlton, and Ranyard 2012). Specifically, if the "double burden" hypothesis holds then we would expect to find the following:

H1: Female refugees will be considered as less employable than male refugees, all other factors held equal.

The "female advantage" hypothesis

An alternative scenario is that female refugees are in fact perceived as *more* employable than their male peers because negative stereotypes about refugees in general are applied more strongly to males than to females.

Stereotypes about refugees are indeed often negative. For example, in their study of perceptions of refugees and asylum seekers conducted in Germany, Kotzur et al. (2019; see also Kotzur, Forsbach, and Wagner 2017; Cuddy et al. 2009) found that "generic" refugees are seen as low in both likability and trustworthiness (*warmth*) and in ability and independence (*competence*). Similar results have been reported by Froehlich and Schulte (2019), who compared stereotypes toward different immigrant groups in Germany and found that immigrants from countries associated with recent refugee outflows such as Syria or Afghanistan are rated as low in both warmth and competence compared to Germans, but the same holds true for immigrant groups with a longer presence in Germany such as Turks.

Yet, it has also been found that any negative stereotypes about foreigners are unequally applied to men and women. Notably, Eagly and Kite (1987) found that negative traits associated with a given foreign nationality are mostly attributed to men with that nationality rather than to women. They suggest that this is because stereotypes about foreigners are mostly informed by "indirect contact," for example portrayals in news media. And it is here where women—due to their generally lower profile in public life—are less likely to be shown. For example, women are less likely to be seen in news reporting of events that can be seen as negative or problematic (e.g., irregular border crossings that involve scuffling with security forces).[2]

[2] Systematic differences in how men and women from the same country of origin are depicted have indeed been observed in the media coverage of the recent (but also past) refugee inflows into Western countries (e.g., Wilmott 2017; Wigger 2019; Berry, Garcia-Blanco, and Moore 2015). Both newspaper articles and photographs have been

In addition, it is possible that the stereotypes attached to refugee women due to their gender (warmth, compassion; Ellemers 2018) can counteract the stereotypes attached to them due to their refugee status, and make them appear less threatening than their male peers. This reduced threat would be particularly advantageous in recruitment processes for jobs that are relatively undesirable and low in status and that involve boring, repetitive, or dirty tasks. In these types of jobs, "tractability" is an important attribute—and it is also these types of jobs that are most likely to be open to applicants with few or no formal credentials or networks such as refugees (e.g., Bonoli and Hinrichs 2012; Waldinger and Lichter 2003; Zamudio and Lichter 2008). In other words, because female refugees are less likely to be perceived as being "threatening" or having "the wrong attitude", they might have an advantage compared to their male peers.

H2: Refugee women are perceived as more employable compared to male refugees, all other factors held equal.

It is important to point out the further implications of this last hypothesis: if female refugees are indeed seen as no less or even more employable than male refugees, then this would imply that discrimination by recruiters might not be the explanation for the observed employment gap between male and female refugees. Other potential alternative explanation could be an adherence

shown to use different frames to portray refugees (Wilmott 2017), and negative stereotypes are more likely to be elicited by media portrayals of (young) (Muslim) refugee men.

to traditional gender roles within refugee communities that leads female refugees to drop out or stay out of the labor force, as mentioned above (Koopmans 2016).

However, we also consider the possibility that the dynamics of stereotyping described above work differently once children enter the picture, and that women might then suffer from a motherhood penalty whereas men can potentially benefit from a fatherhood advantage.[3] If present, such effects might change the overall picture again, especially if refugees come from origin countries with high fertility rates and retain these after migrating (e.g., Scott and Stanfors 2011).

The motherhood penalty hypothesis

Existing evidence points strongly to a significant penalty for mothers in the labor market in many (but not all) advanced economies (Budig, Misra, and Boeckmann 2016). This motherhood penalty can be explained, in part, by objective factors such as the loss of work experience suffered by women that leave the labor force after having children or differences in human capital (Anderson, Binder, and Krause 2002; Budig, Misra, and Boeckmann 2016). However, studies also find a remaining penalty that cannot be explained by observable characteristics, and this remaining penalty may indicate discrimination against mothers by recruiters and employers (Budig and England 2001; Anderson, Binder, and Krause 2002; Correll, Benard, and Paik 2007)—including

[3] In Switzerland it is still customary to disclose personal details such as family status, number and age of children, or nationality in an application. Oftentimes, applicants also include pictures of themselves, which can provide further information about the applicant's gender or ethnicity.

because mothers that seek to work are perceived as warmer, but also less competent (Cuddy, Fiske, and Glick 2004).

If this also applies to refugee women, then this would imply the following:

H3: All else equal (including human capital), refugee mothers are seen as less employable compared to childless refugee women.

The fatherhood advantage hypothesis

Finally, we also consider that any potential disadvantage experienced by male refugees might be reduced if they are married and/or have children, because being a husband and father can reduce their perception as "threatening".

Equivalent to the evidence for a motherhood penalty in many labor markets, there is also considerable evidence for a fatherhood advantage or bonus, meaning that men benefit professionally from having children. More importantly, here again research suggests that at least a part of this effect works through factors other than objective ones such as behavioral changes (Killewald 2013; Bygren and Gähler 2012; but see also Bygren, Erlandsson, and Gähler 2017). The research by Cuddy et al. (2004) is again illustrative as they find that men with children are seen as warmer than men without children, but equally competent.

If this also applies to refugee men, then it is possible that they can compensate for otherwise their lower perceived warmth by becoming fathers.

H4: Refugee fathers are seen as more employable as refugee men without children, all other factors held equal.

Methods

Research design

To evaluate our hypotheses, we use data from an original survey experiment in which we studied how a sample of recruiters evaluates refugee job applicants who differ in gender, family status, and several other attributes.

Specifically, we conducted a so-called factorial or vignette survey experiment. In this type of experiment, participants are asked to evaluate brief descriptions of fictional persons or objects ("vignettes") that vary on a set of attributes (Auspurg and Hinz 2015; Wallander 2009; Jasso 2006). In our experiment, participants were shown brief vignettes of refugee jobseekers (described in more detail below) and were asked to evaluate the employability of each fictional vignette person.

In our vignette experiment, treatment assignment was randomized in two respects: first, each vignette displayed a random combination of jobseeker attributes—in other words, the fictional persons shown were equally likely to be male or female, to have children or be childless, to come from different origin countries, etc. This random composition of vignettes ensures that the unique and unbiased effect of each individual attribute on the participants' evaluations can be identified. In addition, the randomized composition of vignettes and the fact that each vignette included information about several different attributes at a time made it more difficult for participants to

identify which particular attribute is of central interest in the experiment, which helps to reduce social desirability bias (Auspurg and Hinz 2015, 11). Especially when studying discrimination, this is an important precondition to gathering valid measurements. Second, vignettes were also randomly assigned to participants, which means that the experiment controls for potential confounding effects of respondent-level variables.

The vignettes varied along the following set of attributes: first, we included as attributes our core variables of interest, namely gender (male, female) and family status (single, married without children, married with one child aged five). In addition, each vignette description also included information about the fictional refugee's age (24, 37, or 48 years), country of origin (Afghanistan, Syria, Turkey), year of arrival (2015 or 2018), language proficiency (local language A2, local language B2, local language A2 and English B1), former occupation in the country of origin (elementary school teacher, medical doctor, administrative assistant, cleaner, or temporary worker) and integration measure participation (integration course, work practice private sector, work practice public sector, wage subsidy or volunteering) (for more details, see Table S3, Supplementary Material).[4]

The refugees were presented as applicants for two distinct types of jobs. The first position was a job as an administrative assistant, which involved basic office duties such as delivering internal

[4] The Common European Framework of Reference for Languages, also referred to as CEFR or CEF or CEFRL, is a guideline used to describe achievements of learners of foreign languages and validate their language ability. There are six reference levels going from lower to higher (A1, A2, B1, B2, C1, C2).

mail, sorting office material, and copying documents. The second position was as a janitor/caretaker and consisted of tasks such as cleaning offices and taking care of the establishment's outside area. The job descriptions (see experimental protocol, Tables S2a and S2b and Figure S2) were limited to lower-skilled occupations because refugees, even if they are highly qualified, often only have access to lower-skilled jobs due to a lack of language skills, formal credentials, or networks (e.g., Bloch 2007). We presented each respondent with four vignettes of applicants for each of the two positions (i.e., eight vignettes in total), one vignette at a time. Each time, participants evaluated how likely they were to invite the candidate for a job interview on a scale from 1 to 10 (from not at all likely to very likely). Both the position described in the survey and the order of the four vignettes per job were randomized to avoid order effects. Participants were instructed that all applicants had been officially recognized as refugees and had valid work permits to prevent diverging assumptions regarding the bureaucratic hurdles to hiring these workers and/or uncertainty about their ability to stay in the country.

By declaring that all fictional applicants were refugees, we focus specifically on the recruitment experience of the refugee population—and thus, our experiment does not allow us to compare refugees to other immigrant groups or to members of the majority population. This is for two reasons. First, and most importantly, we are interested in the gender employment gap *within* the group of refugees, hence the central comparison is between refugees of different genders. Second, while it would, in principle, be interesting to include natives as a comparison group, there is a risk that the likely strong preference among recruiters for native applicants (e.g., Zschirnt and Ruedin 2016) and the tendency to engage in "attention discrimination" against less appealing groups of

candidates in recruitment (Bartoš et al. 2016) might drown out the finer gender differences within the group of interest: refugees.

We include Syria and Afghanistan as countries of origin because these are the two largest groups of recent refugees in all three countries in which we conducted the experiment (see Table S1, Supplementary Material). However, by having Turkey as a third country of origin, we also include one other group that has a longer immigration history and larger diaspora in Western Europe—but still a valid cause of refugee emigration in the form of the failed coup d'état in 2016 and the generally eroding civil liberties in the years before and after this event (e.g., Esen and Gumuscu 2016). This allows us to account for the possibility that stereotypes might differ between immigrant groups depending on the size of their diaspora and their historical presence in a destination country (Froehlich and Schulte 2019).

We also want to point out an important potential drawback of our experimental setup, which is that this type of experiment can only approximate a real hiring situation and can only capture the hiring *intent*, but not actual hiring behavior. However, as Hainmueller et al. (2015) have recently shown, intentions stated in vignette experiments correspond well to observed real-world behavior, which testifies to the external validity of this type of experiment. An added advantage is that vignette experiments do not have the same ethical issues as alternative approaches such as audit studies or correspondence tests. Finally, vignette experiments are particularly appropriate when studying recruitment for low-skilled positions, where recruitment practices are less structured and application processes do not require highly formalized, individualized, and extensive application documents—which is precisely what our experiment captures.

Survey setup and country case selection

The experiment was embedded within a survey questionnaire that was distributed via the internet and self-administered. Participants were first asked a set of screening questions (see below) before they proceeded to the experiments.

The survey was fielded in three European countries: Austria, Germany, and Sweden. The rationale behind selecting these three countries was as follows. First, these countries were severely confronted with above-average numbers of refugees during the 2015-2017 period, making refugee socioeconomic integration an important and salient topic (see Figure S1, Supplementary Material). Moreover, the composition of the refugee population was similar across the three countries: most refugees were of Syrian, Iraqi or Afghan nationality, and most of them were young, i.e., between 18 and 34 years of age, and male (e.g., Martin et al. 2016).

However, these three countries also differ significantly in their labor market and welfare state institutions, which are factors that have been found to influence the labor market integration of migrants in general, and migrant women in particular (e.g., Ballarino and Panichella 2018; Kogan 2007; Reyneri and Fullin 2011). Germany and Austria have prototypical conservative welfare systems with a strong focus on status maintenance and a bias toward the 'traditional' family model, while Sweden has a Nordic or social democratic welfare state with a more egalitarian orientation—including with regard to the gender-dimension (Esping-Andersen 1990). In addition, these countries also differ in their immigration history. Continental countries such as Germany or Austria are traditional guest-worker countries that relied heavily on foreign workers to meet the demands of their booming labor markets after WWII. Conversely, Sweden took a different route and opted

to expand female labor force participation to reduce the labor market reliance on foreign workers (Afonso 2019). By conducting our survey and experiment in different countries with different labor market and migration traditions, we can account for the possibility that recruiter evaluations of refugees might be affected by different policy regimes—e.g., that recruiters are less likely to penalize refugee women where policies allow for an easier reconciliation of work and family life (e.g., Budig, Misra, and Boeckmann 2012).

Participant recruitment

We sought to select participants with substantial recruitment experience who could therefore deliver a realistic assessment of the fictional refugee job candidates. Obviously, obtaining such a specific participant sample is not a straightforward task. Our approach was to select participants from an online panel operated by Qualtrics, using a set of requirements and quotas.

Upon starting the survey, all initially recruited participants were asked if they had any experience with job recruitment processes. Those participants who indicated that they had been involved in hiring processes during the 12 months prior to the survey were selected; all other participants were screened out. After those without recruitment experience were screened out, we applied quotas based on the participant's age (50% of participants had to be older than 35), gender (50% had to be female), and firm size (60% had to work in firms with up to 250 employees) to obtain a diverse sample. Overall, we had 368 participants from Germany, 228 from Austria, and 363 from Sweden.

These participants rated a total of ~7,600 vignettes. We present more detailed descriptive statistics for our respondent sample in Table S5 in the Supplementary Material.[5]

Empirical model

Given that our dependent variable is metric and our data have a hierarchical structure, with multiple vignette evaluations pooled within many participants, the recommended empirical model is a linear random effects multilevel model (see e.g., Auspurg and Hinz 2015, Chapter 5), and we choose this as our main specification. Our models include dummy variables for all the various vignette attributes listed above as the main predictors; interaction terms are added to some models, as described in detail below. Unless otherwise indicated, our models were estimated on the pooled data that includes observations from all three countries in which we conducted our survey. These models include country dummies and, in some specifications, interactions, to capture between-country differences. We also estimated the main model separately for each country (reported in Table A1), and we re-estimated both the pooled and the country-specific models using a fixed-effects specification and cluster-robust standard errors (see Table S9 in the Supplement) to check if our results are stable across estimation techniques (Bryan and Jenkins 2016). The results do not change substantively across specifications.

[5] And robustness tests for the vignette variables in Tables S4a, S4b and S4c.

Results

The main effects of gender and family status

We start the analysis by inspecting the main coefficient estimates for all vignette attributes from the main pooled model, which are shown in Figure 1 (the detailed estimation results are reported in Table A1, under "Model 1").

The first result that stands out is the overall small magnitude of the effects of the vignette attributes. The largest effect estimate (the dummy for "Janitor" as the position applied for) is around 0.5 points on the 10-point response scale, and other coefficients are half of that or less. We suspect that this might result from the focus on refugees alone in this experiment, and that recruiters may find it difficult to differentiate more strongly between candidates within this narrow group. Had the experiment pitted refugee candidates against other immigrant groups or the native-born, stronger differences might have emerged. Nevertheless, we do find statistically significant effects of several vignette attributes, which indicates that recruiters do make some systematic distinctions within the group of refugees.

Among these, we first point to the significant and negative effect of being male compared to being female on recruiters' evaluations of employability (holding constant other characteristics, including the type of job applied for). This finding is consistent with our second hypothesis, which predicts that men with a refugee background are perceived as less employable because they are more strongly connected to their respective culture's negative stereotypes. By implication, the "double burden" hypothesis is rejected.

<Figure 1 about here>

The results show also that married individuals with a child are, on average, evaluated less positively than married applicants without children or single applicants. Single persons are evaluated better than married ones without children, but this difference is not statistically significant.

When looking at the other vignette attributes, we do not find significant age-related discrimination or country-of-origin effects. The explanation for the latter might be due to the fact that all fictional candidates were presented as having been awarded refugee status, and our participants might in this case trust that their residence in the destination country was legitimate, regardless of their country of origin. Humanitarian concerns might then play into recruiters' evaluations (Bansak, Hainmueller, and Hangartner 2016), leading them to treat Afghans, Syrians, and Turks equally. We stress, however, that this does not mean that such individuals would not be discriminated against if compared to other low-skilled individuals on the labor market (Bloch 2007; Bevelander 2011).

Regarding the other migration-related variables, the results show that, on average, speaking the local language (Swedish or German) at a B2 level or a combination of knowledge of the local language at the A2 level and English at the B2 level does not lead to a statistically significant improvement in recruiter evaluations compared to having only A2 skills in the local language. Refugees who held lower-skilled positions in their country of origin (i.e., were cleaners or had only temporary jobs) were evaluated significantly worse than both former teachers and medical doctors,

and these differences are statistically significant.[6] Refugees who were assigned to a second integration measure in addition to a mandatory integration course were also evaluated significantly better than those who only participated in a mandatory course (the reference category). However, there are no discernible differences between the effect estimates of the different integration programs. This suggests that it does not matter what kind of measure or volunteering activity refugees take part in (see also Author, 2020 forthcoming), as long as they do something that goes beyond a simple integration program.

Finally, applicants for a janitor position are evaluated significantly more favorably than those applying for a position as an administrative assistant. As mentioned above, this is also the effect with the largest magnitude. We interpret this as an additional indication for the relevance of stereotypes in hiring processes. Stereotypes, or more specifically what Fiske and Taylor (2013) describe as "jobholder schemas", provide a plausible explanation for why refugees are consistently sorted into jobs that involve less prestigious work: recruiters have expectations about which profile is most likely to work in such a position, and they find refugees to be more suited for positions involving 'dirtier' and less prestigious work (see also Auer et al. 2019; Pager, Bonikowski, and Western 2009).

[6] However, if we run separate models (not shown) for the janitor and the administrative assistant position, we see that the professional background matters in different ways. For assistant positions, having worked as a cleaner or at temporary jobs decreases employers' evaluation of the profiles. Conversely, when hiring cleaners, they prefer individuals who worked as cleaners in their country of origin. This points to preferences for relevant skills by employers.

<Figure 2 about here>

The conditioning effect of family status

We found so far that refugee women are, overall, preferred in recruitment to refugee men. However, we also expected that this effect might differ depending on whether refugees have children or not. To evaluate our third and fourth hypotheses about how a refugee's family status moderates the effects of gender, we add an interaction between gender and family status to our main specification and compute predicted ratings for each combination of family status and gender from the estimation results. These ratings are shown in Figure 2.[7]

We find, firstly, that the estimates for men with different family statuses are not significantly different from each other. Being male has a consistent negative effect, and this is the same irrespective of whether an applicant is married or has children. This absence of a conditioning effect of fatherhood runs counter to our fourth hypothesis, i.e., the idea that refugee men can compensate for their disadvantage by becoming fathers.

On the other hand, we do find support for the hypothesis that being married with a child makes a difference for women – and here the effect is negative, as expected. Thus, we do find that refugee

[7] The detailed estimation results are shown in Table A2.

women experience a setback in recruiter evaluations once they have children. Interestingly, the predicted rating for refugee mothers is essentially the same as the ratings for refugee men across the different family statuses. In other words, it seems recruiters cluster applicants into two groups: in one group are childless women, who are preferred; in the other group are mothers and all male applicants.

Further analyses and robustness checks

To further probe the stability of the negative main effect of being male compared to female, we also test if women may be considered more suitable for the administrative assistant position whereas men may be seen as better janitors (Cejka and Eagly 1999). We do not find such an effect (see Table S6 and Figure S3 in the Supplementary Material). In addition, we consider whether the gender effect differs across countries as it is conceivable that recruiters in societies with strong gender-equality norms such as Sweden are less prone to discriminate than those elsewhere. Again, we fail to find such differences (see Table S7 and Figure S4 in the Supplementary Material).

Finally, we also investigate if the interaction effects between gender and family status we uncovered above differ between the three countries. The reasoning here is that the better availability of employment-friendly childcare and family policies in Sweden might reduce in particular the penalty experienced by mothers (e.g., Budig, Misra, and Boeckmann 2016). To see if this is the case, we estimate a further model using the pooled data that includes a three-way interaction between gender, family status, and country (Austria, Germany, or Sweden, see Table S8,

Supplementary Material) and test for the joint significance of all interaction terms in the model using a Wald test.[8] All interaction terms are jointly insignificant ($\chi^2 = 9.16$, df = 10, $p = 0.512$). We also test for the joint significance of only the three-way interaction terms between gender, family status, and country and can here as well not reject the null hypothesis ($\chi^2 = 2.13$, df = 4, $p = 0.712$). This indicates that the motherhood penalty does not differ between the three countries we study.

Conclusion

We investigated whether discrimination during recruitment could be an explanation for the widely observed employment gap between male and female refugees. Our findings suggest that this is not the case. Female refugees are, on average, preferred over males in the job recruitment process. This finding is robust across occupations and countries. We suggest that this is because negative stereotypes that are often attributed to refugees (lacking reliability and trustworthiness) are primarily attributed to *male* refugees. We also found that children change this pattern as refugee mothers are no longer evaluated better—but also not worse—than their male counterparts.

Our findings contribute to the literature on the economic integration of immigrants and refugees in western countries (e.g., van Tubergen, Maas, and Flap 2004; de Vroome and van Tubergen 2010;

[8] To be specific, the model includes three-way interaction terms between family status, gender, and country, in addition to two-way interactions between all pairs of these three variables.

Koopmans 2016), and we think that in particular our finding that male refugees never enjoy an advantage over their female peers in recruiting has an interesting implication for this literature. Specifically, it suggests that it is unlikely that gender-based discrimination in recruitment is behind the employment gap between refugee men and women, at least in the three countries we studied. Hiring discrimination could only be a contributing factor if female refugee applicants would at some point be actively disadvantaged compared to their male counterparts—but that is not what we found. We found a female advantage (for non-parents) or the absence of differences (when children are present). This suggests that the factors that produce the employment gap between male and female refugees should be located on the *supply*- rather than the demand-side of the labor market.

One such supply-side factor that could keep female refugees from joining the labor force could be insufficient access to childcare and generally weak policies to support the reconciliation of work and family life (Budig, Misra, and Boeckmann 2016; Bonoli 2013), and another could be an adherence to more traditional gender roles in refugee communities (e.g., Koopmans, 2016). Additional research is needed to investigate this further.

Our study has of course some limitations, and these produce some avenues for future research. A first limitation of our study is that we considered only a small set of refugee origin countries, and it would be worthwhile to see if the discrimination against male refugees we found occurs against refugees from different backgrounds and from earlier refugee waves (e.g., refugees from the Balkan wars of the 1990s, or from Iran in the 1970s). Indeed, although the origin countries we consider (Afghanistan, Syria, and Turkey) are different in many cultural, geographic, and economic

respects, they are all majority-Muslim countries that experienced refugee emigration recently. It is in this context important to point to the study by Eagly and Kite (1987), who found that although stereotypes about nationalities are more strongly applied to males than females, this does not apply equally to all nationalities but mostly to those associated with less gender equality and the absence of democracy and liberal values. Therefore, it would be relevant to study whether there is less between-gender discrimination in recruitment when it comes to refugees from countries that are perceived as more gender-equal and liberal than Afghanistan, Syria, or Turkey.

A related limitation of our study is the fact that we covered only recruiter perceptions in three West European countries, two of which belong to the Continental European or Conservative welfare state regime (Esping-Andersen 1990). A natural way to extend our analysis would be to include further countries, in particular non-European countries with welfare state models that are classified as liberal (in the sense of a strong free-market orientation, low taxes, little regulation or redistribution) such as Canada, the United States, Australia, or New Zealand but potentially also countries in the Mediterranean area with less developed welfare states such as Italy or Greece (Esping-Andersen 1990; Ferrera 1996).

Furthermore, an obvious limitation of our study is that we focused on differences within the group of refugees and did not also include a comparison to natives or other groups of immigrants. As already mentioned above, including natives or other groups of immigrants in the experimental design would create a risk that any differences within the group of refugees get drowned out due to "attention discrimination" by recruiters (Bartoš et al. 2016). Still, conducting this type of experiment-—ideally with a large sample of participants to ensure sufficient statistical power to

detect even small effects—would be worthwhile since this would allow researchers to test the "double burden" hypothesis more explicitly and directly than we have been able to do here.

Last but not least, we want to point out that the focus of our study–labor market integration–is only one of several core dimensions of host country integration, next to social, political, or psychological integration (Harder et al. 2018). In addition, and as mentioned already at the start, simply having a job is not equivalent to being successfully integrated because jobs vary in quality and the extent to which they are appropriate given workers' skills and abilities. It is in this area where immigrants, including immigrant women, have been found to be facing significant disadvantages (Ballarino and Panichella 2018; 2015). Although this particular form of disadvantage has not been our focus here, our finding that refugee applicants are strongly preferred if they apply for the position as janitor instead of the (more prestigious) position as administrative assistant is in line with the finding from other studies that recruiters tend to channel applicants with foreign or minority backgrounds into less prestigious positions (Auer et al. 2019; Pager, Bonikowski, and Western 2009). Of course, this pattern might look differently if one were to study recruitment patterns in countries other than those we have studied. Based on the results of previous research, we would expect this to be less pronounced in liberal countries such as the United Kingdom but even more pronounced in Southern Europe (Ballarino and Panichella 2015). This again highlights the value of extending our analysis to these contexts.

Our results also have implications for policymaking, the main one being that when it comes to addressing labor market discrimination against immigrants and refugees in general, a major focus should be on countering stereotypes held by employers (see also Vernby and Dancygier 2019).

Particularly stereotypes about *male* refugees should be countered. Existing research suggests a number of potential interventions such as de-biasing training, fostering positive contact, or institutional reforms such as mandatory quotas or diversity offices (see Paluck and Green 2009; Kalev, Dobbin, and Kelly 2006).

De-biasing interventions involve exposing decisionmakers such as recruiters to stimuli and exercises that reduce any biases they may have, including biases that are subconscious or "implicit" (Greenwald and Banaji 1995). Especially "perspective taking" and exposing individuals to counter stereotypes, namely images of famous and admired exponents of the minority group, have proven particularly effective at reducing biases (e.g., Lai et al. 2014; Dasgupta and Greenwald 2001; Finnegan, Oakhill, and Garnham 2015).

A second type of intervention draws on the logic of the contact hypothesis (Pettigrew and Tropp 2006), which predicts that promoting contact between different groups can help reduce any biases and prejudices that group members might hold about each other, provided that the experiences are meaningful, positive, and constructive. In a labor market context, subsidized temporary employment or internship programs could be used to foster such contact between employers and refugees (Hirst et al. 2019; but see also Ortlieb et al. 2020). Similarly, mentoring programs in which refugees could get advice and are introduced into professional networks might be relevant here (Liechti 2020).

Finally, changing rules and institutions can help the inclusion of disadvantaged groups such as (male) refugees. One measure in this category would be to create structures within companies that assign direct responsibility for countering bias and increasing diversity such as diversity managers

or affirmative action plans (Kalev, Dobbin, and Kelly 2006). A second institutional measure would be mandatory quotas or affirmative action policies. Albeit controversial, these types of policies have proven effective at helping different minority groups across different context because they trigger changes in recruitment procedures more generally, which in turn results in fairer and more formalized processes (Miller 2017).

References

Afonso, Alexandre. 2019. “Migrant Workers or Working Women? Comparing Labour Supply Policies in Post-War Europe.” *Journal of Comparative Policy Analysis: Research and Practice* 21 (3): 251–69.

Anderson, Deborah J., Melissa Binder, and Kate Krause. 2002. “The Motherhood Wage Penalty: Which Mothers Pay It and Why?” *American Economic Review* 92 (2): 354–58.

Auer, Daniel, Giuliano Bonoli, Flavia Fossati, and Fabienne Liechti. 2019. “The Matching Hierarchies Model: Evidence from a Survey Experiment on Employers’ Hiring Intent Regarding Immigrant Applicants.” *International Migration Review* 53 (1): 90–121.

Auspurg, Katrin, and Thomas Hinz. 2015. *Factorial Survey Experiments*. Quantitative Applications in the Social Sciences 175. Los Angeles and others: Sage.

Ballarino, Gabriele, and Nazareno Panichella. 2015. “The Occupational Integration of Male Migrants in Western European Countries: Assimilation or Persistent Disadvantage?” *International Migration* 53 (2): 338–52.

———. 2018. “The Occupational Integration of Migrant Women in Western European Labour Markets.” *Acta Sociologica* 61 (2): 126–42.

Bansak, K., J. Hainmueller, and D. Hangartner. 2016. “How Economic, Humanitarian, and Religious Concerns Shape European Attitudes toward Asylum Seekers.” *Science* 354 (6309): 217–22.

Bartoš, Vojtěch, Michal Bauer, Julie Chytilová, and Filip Matějka. 2016. “Attention Discrimination: Theory and Field Experiments with Monitoring Information Acquisition.” *American Economic Review* 106 (6): 1437–75.

Berry, Mike, Inaki Garcia-Blanco, and Kerry Moore. 2015. *Press Coverage of the Refugee and Migrant Crisis in the EU: A Content Analysis of Five European Countries*. Geneva: United Nations High Commissioner for Refugees.

Bevelander, Pieter. 2011. “The Employment Integration of Resettled Refugees, Asylum Claimants, and Family Reunion Migrants in Sweden.” *Refugee Survey Quarterly* 30 (1): 22–43.

Bloch, Alice. 2007. “Refugees in the UK Labour Market: The Conflict between Economic Integration and Policy-Led Labour Market Restriction.” *Journal of Social Policy* 37 (1): 21–36.

Bonoli, Giuliano. 2013. *The Origins of Active Social Policy: Labour Market and Childcare Policies in a Comparative Perspective*. Oxford University Press.

Bonoli, Giuliano, and Karl Hinrichs. 2012. “Statistical Discrimination and Employers’ Recruitment.” *European Societies* 14 (3): 338–61.

Bordalo, Pedro, Katherine Coffman, Nicola Gennaioli, and Andrei Shleifer. 2016. “Stereotypes.” *The Quarterly Journal of Economics* 131 (4): 1753–94.

Breidahl, Karen N., and Christian Albrekt Larsen. 2016. “The Myth of Unadaptable Gender Roles: Attitudes towards Women’s Paid Work among Immigrants across 30 European Countries.” *Journal of European Social Policy* 26 (5): 387–401.

Bryan, Mark L., and Stephen P. Jenkins. 2016. “Multilevel Modelling of Country Effects: A Cautionary Tale.” *European Sociological Review* 32 (1): 3–22. https://doi.org/10.1093/esr/jcv059.

Budig, Michelle J., and Paula England. 2001. “The Wage Penalty for Motherhood.” *American Sociological Review* 66 (2): 204–25.

Budig, Michelle J., Joya Misra, and Irene Boeckmann. 2012. “The Motherhood Penalty in Cross-National Perspective: The Importance of Work–Family Policies and Cultural Attitudes.” *Social Politics: International Studies in Gender, State & Society* 19 (2): 163–93.

———. 2016. “Work–Family Policy Trade-Offs for Mothers? Unpacking the Cross-National Variation in Motherhood Earnings Penalties.” *Work and Occupations* 43 (2): 119–77.
Bygren, Magnus, Anni Erlandsson, and Michael Gähler. 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): 337–48.
Bygren, Magnus, and Michael Gähler. 2012. “Family Formation and Men’s and Women’s Attainment of Workplace Authority.” *Social Forces* 90 (3): 795–816.
Cejka, Mary Ann, and Alice H. Eagly. 1999. “Gender-Stereotypic Images of Occupations Correspond to the Sex Segregation of Employment.” *Personality and Social Psychology Bulletin* 25 (4): 413–23.
Cheung, Sin Yi, and Jenny Phillimore. 2017. “Gender and Refugee Integration: A Quantitative Analysis of Integration and Social Policy Outcomes.” *Journal of Social Policy* 46 (2): 211–30.
Correll, Shelley J., Stephen Benard, and In Paik. 2007. “Getting a Job: Is There a Motherhood Penalty?” *American Journal of Sociology* 112 (5): 1297–1338.
Cuddy, Amy J. C., Susan T. Fiske, and Peter Glick. 2004. “When Professionals Become Mothers, Warmth Doesn’t Cut the Ice.” *Journal of Social Issues* 60 (4): 701–18.
Cuddy, Amy J. C., Susan T. Fiske, Virginia S. Y. Kwan, Peter Glick, Stéphanie Demoulin, Jacques-Philippe Leyens, Michael Harris Bond, et al. 2009. “Stereotype Content Model across Cultures: Towards Universal Similarities and Some Differences.” *British Journal of Social Psychology* 48 (1): 1–33.
Dasgupta, Nilanjana, and Anthony G. Greenwald. 2001. “On the Malleability of Automatic Attitudes: Combating Automatic Prejudice with Images of Admired and Disliked Individuals.” *Journal of Personality and Social Psychology* 81 (5): 800–814.
Dumper, Hildegard. 2002. *Missed Opportunities: A Skills Audit of Refugee Women in London from the Teaching, Nursing and Medical Professions*. London: Mayor of London & Refugee Women’s Association.
Eagly, Alice H., and Mary E. Kite. 1987. “Are Stereotypes of Nationalities Applied to Both Women and Men?” *Journal of Personality and Social Psychology* 53 (3): 451.
Ellemers, Naomi. 2018. “Gender Stereotypes.” *Annual Review of Psychology* 69 (1): 275–98.
Esen, Berk, and Sebnem Gumuscu. 2016. “Rising Competitive Authoritarianism in Turkey.” *Third World Quarterly* 37 (9): 1581–1606.
Esping-Andersen, Gøsta. 1990. *The Three Worlds of Welfare Capitalism*. Polity Press.
Fernández, Raquel, and Alessandra Fogli. 2009. “Culture: An Empirical Investigation of Beliefs, Work, and Fertility.” *American Economic Journal: Macroeconomics* 1 (1): 146–77.
Ferrera, Maurizio. 1996. “The ‘Southern Model’ of Welfare in Social Europe.” *Journal of European Social Policy* 6 (1): 17–37.
Finnegan, Eimear, Jane Oakhill, and Alan Garnham. 2015. “Counter-Stereotypical Pictures as a Strategy for Overcoming Spontaneous Gender Stereotypes.” *Frontiers in Psychology* 6.
Fiske, Susan T., and Shelley E. Taylor. 2013. *Social Cognition: From Brains to Culture*. Los Angeles and others: SAGE.
Froehlich, Laura, and Isabel Schulte. 2019. “Warmth and Competence Stereotypes about Immigrant Groups in Germany.” *PLOS ONE* 14 (9): e0223103.
Gallie, Duncan. 1999. “Unemployment and Social Exclusion in the European Union.” *European Societies* 1 (2): 139–67.
González, M José, Clara Cortina, and Jorge Rodríguez. 2019. “The Role of Gender Stereotypes in Hiring: A Field Experiment.” *European Sociological Review* 35 (2): 187–204.
Greenwald, Anthony G., and Mahzarin R. Banaji. 1995. “Implicit Social Cognition: Attitudes, Self-Esteem, and Stereotypes.” *Psychological Review* 102 (1): 4–27.

Hainmueller, Jens, Dominik Hangartner, and Teppei Yamamoto. 2015. “Validating Vignette and Conjoint Survey Experiments against Real-World Behavior.” *Proceedings of the National Academy of Sciences* 112 (8): 2395–2400.

Harder, Niklas, Lucila Figueroa, Rachel M. Gillum, Dominik Hangartner, David D. Laitin, and Jens Hainmueller. 2018. “Multidimensional Measure of Immigrant Integration.” *Proceedings of the National Academy of Sciences* 115 (45): 11483–88.

Hirst, Giles, Sally Curtis, Ingrid Nielsen, Russell Smyth, Alexander Newman, and Ning Xiao. 2019. “Refugee Recruitment and Workplace Integration: An Opportunity for Human Resource Management Scholarship and Impact.” *Human Resource Management Journal* forth.

Hooijer, Gerda, and Georg Picot. 2015. “European Welfare States and Migrant Poverty: The Institutional Determinants of Disadvantage.” *Comparative Political Studies* 48 (14): 1879–1904.

Jasso, Guillermina. 2006. “Factorial Survey Methods for Studying Beliefs and Judgments.” *Sociological Methods and Research* 34 (3): 334–423.

Kalev, Alexandra, Frank Dobbin, and Erin Kelly. 2006. “Best Practices or Best Guesses? Assessing the Efficacy of Corporate Affirmative Action and Diversity Policies.” *American Sociological Review* 71 (4): 589–617.

Killewald, Alexandra. 2013. “A Reconsideration of the Fatherhood Premium: Marriage, Coresidence, Biology, and Fathers’ Wages.” *American Sociological Review* 78 (1): 96–116.

King, Deborah K. 1988. “Multiple Jeopardy, Multiple Consciousness: The Context of a Black Feminist Ideology.” *Signs: Journal of Women in Culture and Society* 14 (1): 42–72.

Kogan, Irena. 2007. *Working through Barriers: Host Country Institutions and Immigrant Labour Market Performance in Europe*. Springer.

Koopmans, Ruud. 2016. “Does Assimilation Work? Sociocultural Determinants of Labour Market Participation of European Muslims.” *Journal of Ethnic and Migration Studies* 42 (2): 197–216.

Kotzur, Patrick F., Nora Forsbach, and Ulrich Wagner. 2017. “Choose Your Words Wisely: Stereotypes, Emotions, and Action Tendencies toward Fled People as a Function of the Group Label.” *Social Psychology* 48 (4): 226–41.

Kotzur, Patrick F., Maria-Therese Friehs, Frank Asbrock, and Maarten H. W. van Zalk. 2019. “Stereotype Content of Refugee Subgroups in Germany.” *European Journal of Social Psychology* 49 (7): 1344–58.

Lai, Calvin K., Maddalena Marini, Steven A. Lehr, Carlo Cerruti, Jiyun-Elizabeth L. Shin, Jennifer A. Joy-Gaba, Arnold K. Ho, Bethany A. Teachman, Sean P. Wojcik, and Spassena P. Koleva. 2014. “Reducing Implicit Racial Preferences: I. A Comparative Investigation of 17 Interventions.” *Journal of Experimental Psychology: General* 143 (4): 1765.

Liechti, Fabienne. 2020. “Connecting Employers and Workers: Can Recommendations from the Public Employment Service Act as a Substitute for Social Contacts?” *Work, Employment and Society* 34 (4): 587–604.

Lundborg, Per, and Per Skedinger. 2016. “Employer Attitudes towards Refugee Immigrants: Findings from a Swedish Survey.” *International Labour Review* 155 (2): 315–37. https://doi.org/10.1111/ilr.12026.

Martin, Iván, Albert Arcarons, Jutta Aumüller, Pieter Bevelander, Henrik Emilsson, Sona Kalantaryan, Alastair Maciver, et al. 2016. *From Refugees to Workers : Mapping Labour Market Integration Support Measures for Asylum-Seekers and Refugees in EU Member States. Volume II: Literature Review and Country Case Studies*. Florence: Migration Policy Centre (MPC), European University Institute.

Midtbøen, Arnfinn H. 2014. “The Invisible Second Generation? Statistical Discrimination and Immigrant Stereotypes in Employment Processes in Norway.” *Journal of Ethnic and Migration Studies* 40 (10): 1657–75.
Miller, Conrad. 2017. “The Persistent Effect of Temporary Affirmative Action.” *American Economic Journal: Applied Economics* 9 (3): 152–90.
Ortlieb, Renate, Petra Eggenhofer-Rehart, Sandra Leitner, Roland Hosner, and Michael Landesmann. 2020. “Do Austrian Programmes Facilitate Labour Market Integration of Refugees?” *International Migration* forth.
Pager, Devah, Bart Bonikowski, and Bruce Western. 2009. “Discrimination in a Low-Wage Labor Market: A Field Experiment.” *American Sociological Review* 74 (5): 777–99.
Paluck, Elizabeth L., and Donald P. Green. 2009. “Prejudice Reduction: What Works? A Review and Assessment of Research and Practice.” *Annual Review of Psychology* 60: 339–67.
Pettigrew, Thomas F., and Linda R. Tropp. 2006. “A Meta-Analytic Test of Intergroup Contact Theory.” *Journal of Personality and Social Psychology* 90 (5): 751.
Phelps, Edmund S. 1972. “The Statistical Theory of Racism and Sexism.” *American Economic Review* 62 (4): 659–61.
Reyneri, Emilio, and Giovanna Fullin. 2011. “Labour Market Penalties of New Immigrants in New and Old Receiving West European Countries.” *International Migration* 49 (1): 31–57.
Scott, Kirk, and Maria Stanfors. 2011. “The Transition to Parenthood among the Second Generation: Evidence from Sweden, 1990–2005.” *Advances in Life Course Research*, Family Life Transitions of the Second Generation, 16 (4): 190–204.
Taylor, Susan D., John P. Charlton, and Rob Ranyard. 2012. “Ethnic and Gender Differences in the Labour Market Perceptions of Post-Higher Education Job Seekers: ‘Double Jeopardy’ or ‘Ethnic Prominence’?” *Journal of Occupational and Organizational Psychology* 85 (2): 353–69.
Tubergen, Frank van, Ineke Maas, and Henk Flap. 2004. “The Economic Incorporation of Immigrants in 18 Western Societies: Origin, Destination, and Community Effects.” *American Sociological Review* 69 (5): 704–27.
Vernby, Kåre, and Rafaela Dancygier. 2019. “Can Immigrants Counteract Employer Discrimination? A Factorial Field Experiment Reveals the Immutability of Ethnic Hierarchies.” *PLOS ONE* 14 (7): e0218044.
Vroome, Thomas de, and Frank van Tubergen. 2010. “The Employment Experience of Refugees in the Netherlands.” *International Migration Review* 44 (2): 376–403.
Waldinger, Roger David, and Michael Ira Lichter. 2003. *How the Other Half Works: Immigration and the Social Organization of Labor*. Berkeley, Calif: University of California Press.
Wallander, Lisa. 2009. “25 Years of Factorial Surveys in Sociology: A Review.” *Social Science Research* 38 (3): 505–20.
Wigger, Iris. 2019. “Anti-Muslim Racism and the Racialisation of Sexual Violence: ‘Intersectional Stereotyping’ in Mass Media Representations of Male Muslim Migrants in Germany.” *Culture and Religion* 20 (3): 248–71.
Wilmott, Annabelle Cathryn. 2017. “The Politics of Photography: Visual Depictions of Syrian Refugees in U.K. Online Media.” *Visual Communication Quarterly* 24 (2): 67–82.
Zamudio, Margaret M., and Michael I. Lichter. 2008. “Bad Attitudes and Good Soldiers: Soft Skills as a Code for Tractability in the Hiring of Immigrant Latina/Os over Native Blacks in the Hotel Industry.” *Social Problems* 55 (4): 573–89.
Zschirnt, Eva, and Didier Ruedin. 2016. “Ethnic Discrimination in Hiring Decisions: A Meta-Analysis of Correspondence Tests 1990–2015.” *Journal of Ethnic and Migration Studies* 42 (7): 1115–34.

Figures

Figure 1: The influence of vignette variables on refugee applicants' ratings by recruiters (both occupations and all countries)

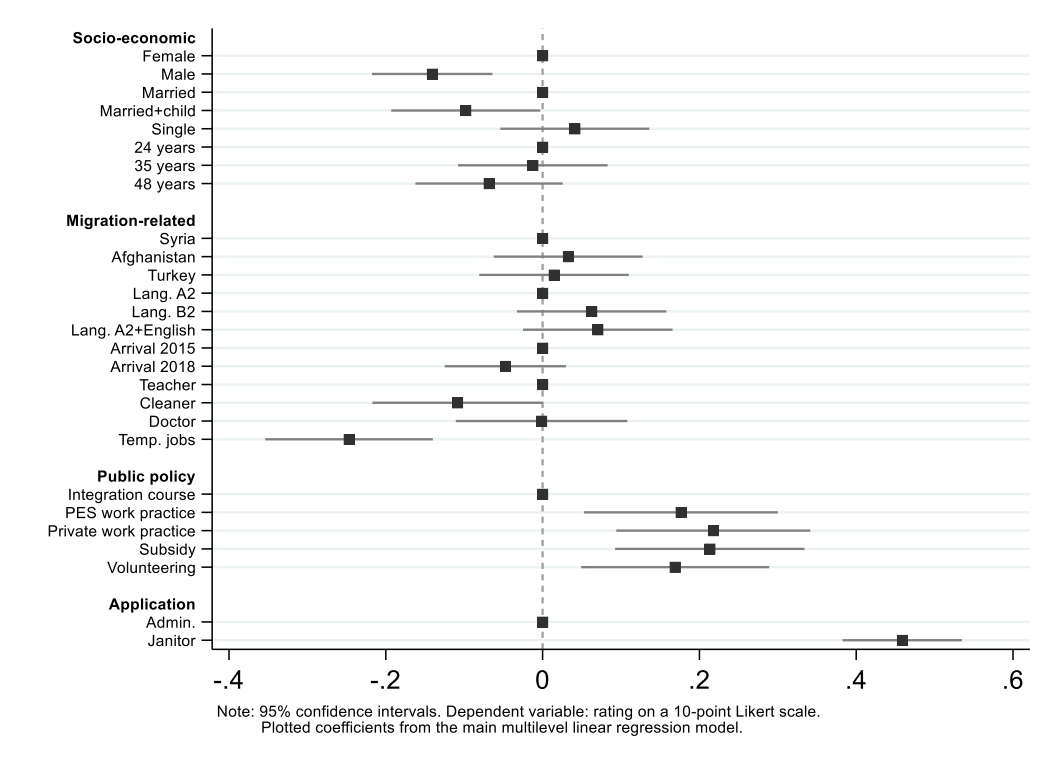
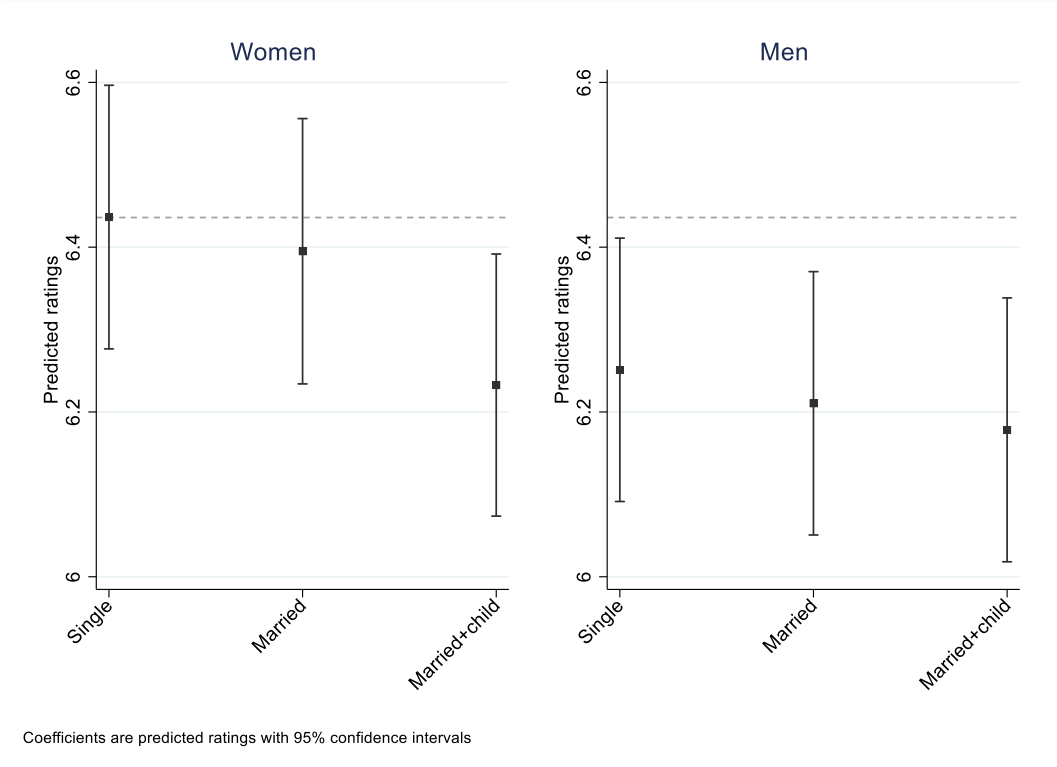


Figure 2: Predicted ratings



Appendix: Tables and Figures

Table A1: Effects of vignette variables on recruiters' ratings of refugee candidates

	Model 1 All countries		**Model 2 Germany**		**Model 3 Austria**		**Model 4 Sweden**	
Female (ref.)								
Male	-0.141***	(0.039)	-0.102	(0.064)	-0.158°	(0.089)	-0.160**	(0.057)
Married (ref.)								
Married+child	-0.098*	(0.048)	-0.099	(0.080)	-0.082	(0.110)	-0.106	(0.071)
Single	0.041	(0.048)	0.069	(0.080)	0.176	(0.110)	-0.075	(0.071)
24 years (ref.)								
35 years	-0.012	(0.049)	-0.011	(0.080)	-0.118	(0.111)	0.066	(0.071)
48 years	-0.068	(0.048)	-0.007	(0.079)	-0.245*	(0.109)	-0.001	(0.070)
Syria (ref.)								
Afghanistan	0.033	(0.048)	-0.030	(0.080)	0.127	(0.109)	0.011	(0.071)
Turkey	0.015	(0.049)	-0.061	(0.080)	-0.009	(0.110)	0.095	(0.071)
Lang. A2 (ref.)								
Lang. B2	0.063	(0.049)	0.129	(0.080)	0.192°	(0.111)	-0.096	(0.071)
Lang. A2+English	0.070	(0.049)	0.003	(0.080)	0.218*	(0.110)	0.046	(0.071)
Arrival 2015 (ref.)								
Arrival 2018	-0.047	(0.039)	-0.041	(0.065)	-0.105	(0.090)	-0.029	(0.058)
Teacher (ref.)								
Cleaner	-0.109°	(0.055)	-0.002	(0.091)	-0.247°	(0.126)	-0.134°	(0.081)
Doctor	-0.001	(0.056)	0.057	(0.092)	-0.125	(0.126)	-0.002	(0.082)

Temporary	-0.247***	(0.055)	-0.237**	(0.090)	-0.348**	(0.124)	-0.206*	(0.080)
Integration course (ref.)								
PES work practice	0.176**	(0.063)	0.110	(0.104)	0.157	(0.144)	0.250**	(0.092)
Private work practice	0.218***	(0.063)	0.179°	(0.105)	0.135	(0.141)	0.319***	(0.093)
Subsidy	0.213***	(0.062)	0.147	(0.102)	0.240°	(0.137)	0.270**	(0.091)
Volunteering	0.169**	(0.061)	0.243*	(0.101)	0.155	(0.139)	0.104	(0.090)
Admin (ref.)								
Clean	0.459***	(0.039)	0.503***	(0.064)	0.425***	(0.088)	0.432***	(0.057)
Germany (ref.)								
Austria	-0.389*	(0.177)						
Sweden	0.027	(0.156)						
Constant	6.152***	(0.137)	6.098***	(0.175)	5.849***	(0.238)	6.196***	(0.162)
Var. vignettes	4.067	(0.202)	3.977	(0.321)	4.157	(0.431)	4.102	(0.326)
Var. respondents	2.873	(0.050)	2.979	(0.083)	3.496	(0.124)	2.332	(0.065)
N vignettes	7634		2928		1814		2892	
N respondent	959		368		228		363	
AIC	32171.404		12450.248		7995.470		11681.623	
BIC	32331.033		12575.872		8111.039		11806.987	
Ll	-16062.702		-6204.124		-3976.735		-5819.812	

Table A2: Interaction gender and marital status, all countries and both jobs

	Model 1	
Female (ref.)		
Male	-0.185**	(0.072)
Married (ref.)		
Married+child	-0.162*	(0.070)
Single	0.041	(0.071)
Interaction		
Male # married+child	0.130	(0.102)
Male # single	-0.001	(0.103)
24 years (ref.)		
35 years	-0.010	(0.049)
48 years	-0.070	(0.048)
Syria (ref.)		
Afghanistan	0.032	(0.048)
Turkey	0.015	(0.049)
Lang. A2 (ref.)		
Lang. B2	0.059	(0.049)
Lang. A2+English	0.068	(0.049)
Arrival 2015 (ref.)		
Arrival 2018	-0.045	(0.040)
Teacher (ref.)		
Cleaner	-0.105°	(0.056)
Doctor	-0.006	(0.056)
Temporary	-0.249***	(0.055)
Integration course (ref.)		
PES work practice	0.172**	(0.063)
Private work practice	0.214***	(0.063)
Subsidy	0.210***	(0.062)
Volunteering	0.163**	(0.061)
Admin (ref.)		
Clean	0.458***	(0.039)
Germany (ref.)		
Austria	-0.390*	(0.177)
Sweden	0.027	(0.156)
Constant	6.179***	(0.141)
Var. vignettes	4.066	(0.202)

Var. respondents	2.872	(0.050)
N	7634.000	
AIC	32173.238	
BIC	32346.748	
ll	-16061.619	

Supplementary Material:

The Gender Employment Gap among Refugees and the Role of Employer Discrimination: Experimental Evidence from the German, Swedish and Austrian Labor Markets

---

Flavia Fossati, Carlo Knotz, Fabienne Liechti and Ihssane Otmani

IDHEAP, University of Lausanne, Rue de la Mouline 28, 1015 Lausanne

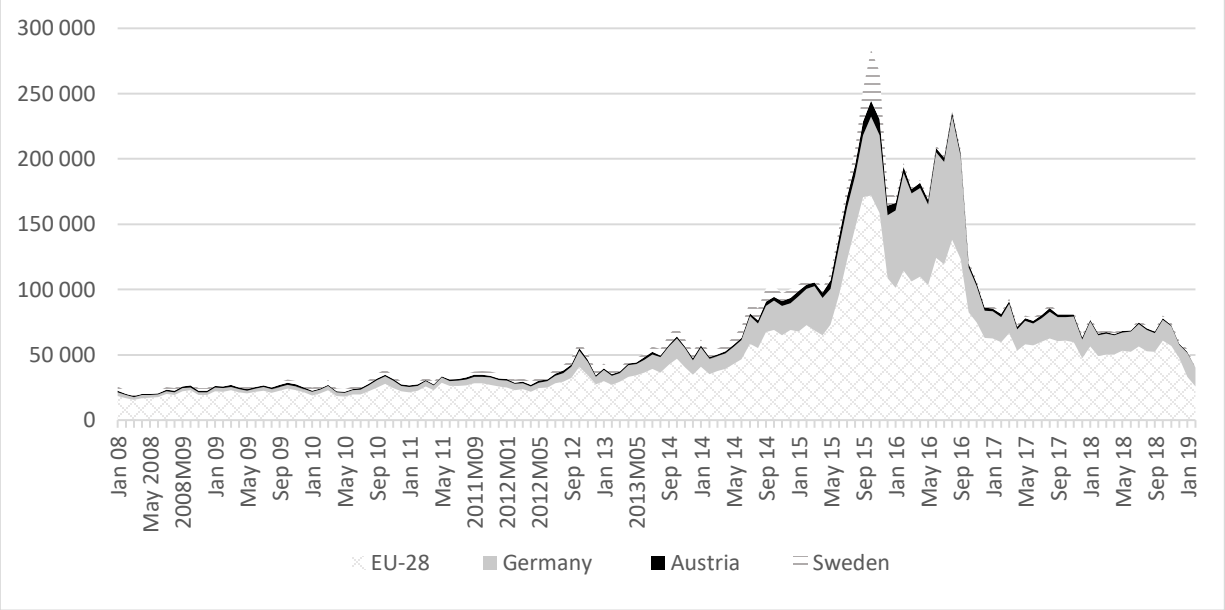
Corresponding author: flavia.fossati@unil.ch

Table S1: Citizenships of first-time non-EU asylum applicants (2015)

Citizenship	Germany	Austria	Sweden
Syria	158655	24720	50890
Albania	53805	135	2565
Kosovo	33425	1605	1560
Afghanistan	31380	24840	41190
Iraq	29785	13225	20190
Iran	5395	3380	4270
Pakistan	8200	2890	515
Stateless	3885	2000	7445
Eritrea	10875	85	6515
Turkey	1500	190	220
Total	**441900**	**85520**	**156195**

Note: in grey are the five main citizenships of non-EU asylum applicants per country

Figure S1: Number of asylum seekers (2008 -2019)



Source: Eurostat (2020)

Experimental protocol

D-efficient sample

From the whole vignette universe of 8,100 possible combinations, we draw a d-efficient sub-sample of 220 vignettes per job that minimizes the correlation between the different dimensions in the vignette universe (Auspurg and Hinz, 2015).[9] The 220 vignettes were divided into 55 blocks of 4 vignettes each that were randomly distributed to respondents. We chose to have 4 vignettes per block because this resulted in eight vignettes per respondent (four for each job), and this is the number of vignettes respondents are usually able to evaluate without fatigue effects (Auspurg and Hinz, 2015).

Implementation in Qualtrics

Table S2a: Job description administrative assistant

Rating of candidates

Imagine that in your company, you have an open position for an unskilled worker in the area of administration and you are involved in the recruitment process.

The tasks are the following:

- Internal mailing
- Put away office material
- Copying and arrange documents

All candidates are recognized refugees and have the permission to work.

Please indicate how likely you are to invited to following candidates for a job interview (1_very unlikely, 10 very likely).

[9] A d-efficient design draws a subset of vignettes to be presented to the respondents from the vignette universe, it is a technique appropriate for small samples of respondents. We used the SAS algorithm *mktex* to identify a sub-sample that maximizes the orthogonality of the profiles, thereby also maximizing the statistical power one can obtain from a given number of observations (Auspurg and Hinz, 2015). Drawing a deficient sample (in contrast to a random sample) allows us to specify which effects can be estimated (we specified all main effects and all two-way interactions). Our vignette sample has a d-efficiency of 90.1, which allows to reduce correlations between dimensions to below 0.05.

Table S2b: Job description caretaker

Rating of candidates

Imagine that in your company, you have an open position for an unskilled worker in the area of cleaning and maintenance and you are involved in the recruitment process.

The tasks are the following:

- Cleaning the office space
- Taking care of the outside space and the green area

All candidates are recognized refugees and have the permission to work.

Please indicate how likely you are to invited to following candidates for a job interview (1_very unlikely, 10 very likely).

Figure S2: Vignette example translation in English

Applicant Administration

- Ms Bakhtari fled Syria because of political prosecution and has been living in Germany since beginning of 2015 and is looking for a job.
- She is 24 years old, married and has one child of five years of age.
- She speaks German at level A2 and English at level B2 and worked as a medical doctor in her country of origin.
- Ms Bakhtari is doing community work twice a week in the framework of a voluntary project of the Red Cross in an elderly home.
-

Please indicate how likely it is that you would invite the candidate for a job interview. (1=very unlikely, 10=very likely).

	1	2	3	4	5	6	7	8	9	10
Invite for interview	○	○	○	○	○	○	○	○	○	○

Note: The vignette for the position as caretaker are the same, only the job description changes.

Table S3: Vignette Dimensions

	Dimension	Level
1	Gender	Male
		Female
2	Refugee from	Syria
		Afghanistan
		Turkey
3	Year of arrival in host country	2015
		2018
4	Age	24 years
		37 years
		48 years
5	Family status	Married, no children
		Married, one child of 5 years of age
		Single
6	Language	Local language A2
		Local language B2
		local language A2 + English B1
7	Training in home country	Elementary school teacher
		Medical doctor
		Administrative assistant
		Cleaner
		Different temporary jobs
8	ALMP	Basic integration course
		1-month work practice organized by the jobcentre
		1-month work practice in a private firm
		Wage subsidy 40% for six months
		Volunteering work twice a week for a local

Robustness

Table S4a: Correlations for vignette dimensions, both jobs

Administrative assistant	**Age**	**Gender**	**Nationality**	**Public policy**	**Language**	**Occupation**	**Year arrival**	**Children**
Vignette dimensions								
Age	1.00							
Gender	-0.01	1.00						
Nationality	-0.01	0.00	1.00					
Public policy	0.02	-0.00	-0.01	1.00				
Language	0.01	-0.00	0.00	0.02	1.00			
Occupation	-0.00	-0.03	-0.00	0.01	0.00	1.00		
Year arrival	-0.01	0.02	0.00	0.00	-0.01	0.01	1.00	
Children	0.01	-0.00	0.01	0.02	-0.01	0.02	-0.00	1.00

Table S4b: Correlations for vignette dimensions, administrative assistant

Administrative assistant	**Age**	**Gender**	**Nationality**	**Public policy**	**Language**	**Occupation**	**Year arrival**	**Children**
Vignette dimensions								
Age	1.00							
Gender	-0.02	1.00						
Nationality	-0.00	0.01	1.00					
Public policy	0.02	-0.00	-0.01	1.00				
Language	0.01	0.01	-0.00	0.02	1.00			
Occupation	-0.01	-0.03	-0.01	0.01	0.01	1.00		
Year arrival	-0.01	0.02	-0.00	0.01	-0.01	-0.00	1.00	
Children	0.01	-0.01	0.00	0.01	-0.01	0.02	0.00	1.00

Table S4c: Correlations for vignette dimensions, caretaker

Cleaner	Age	Gender	Nationality	Public policy	Language	Occupation	Year arrival	Children
Vignette dimensions								
Age	1.00							
Gender	-0.00	1.00						
Nationality	-0.02	-0.00	1.00					
Public policy	0.02	0.00	-0.00	1.00				
Language	0.01	-0.01	0.01	0.02	1.00			
Occupation	-0.00	-0.03	-0.00	0.01	0.00	1.00		
Year arrival	-0.01	0.02	0.00	-0.01	-0.01	0.01	1.00	
Children	0.01	0.01	0.02	0.02	-0.01	0.02	-0.01	1.00

Table S5: Respondent sample

Variable	
Size	
1-9 employees	11.78
10-49 employees	21.27
50-249 employees[1]	26.69
250-499 employees	12.62
More than 500 employees	27.63
Sector	
Agriculture	2.29
Mining/Energy/Waste	2.92
Production	10.68
Construction	5.94
Wholesale	10.01
Transport	3.44
Information	7.61
Hospitality	4.8
Finance	3.65
Education	5.32
Health and social services	12.30
Other services	5.53
Public administration	3.96
Urban	
Urban	36.91
Suburban	22.52
Middle town	21.38
Rural	12.20

Table S6: Interaction gender and occupation

	Model 1	
Female (ref.)		
Male	-0.146**	(0.055)
Admin. Assistant (ref.)		
Janitor	0.454***	(0.055)
Interaction		
Janitor # male	0.010	(0.078)
Married (ref.)		
Married+child	-0.098*	(0.048)
Single	0.041	(0.048)
24 years old (ref.)		
35 years old	-0.013	(0.049)
48 years old	-0.068	(0.048)
Syria (ref.)		
Afghanistan	0.033	(0.048)
Turkey	0.015	(0.049)
A2		
B2	0.063	(0.049)
A2 + English B2	0.070	(0.049)
2015 (ref.)		
2018	-0.047	(0.039)
Teacher (ref.)		
Cleaner	-0.109°	(0.055)
Medical doctor	-0.001	(0.056)
Temporary work	-0.247***	(0.055)
Integration course (ref.)		
Public work	0.177**	(0.063)
Private work	0.218***	(0.063)
Subsidy	0.213***	(0.062)
Volunteering	0.169**	(0.061)
Austria	-0.389*	(0.177)
Sweden	0.027	(0.156)
Constant	6.155***	(0.139)
Var respondent	4.067	(0.202)
Var. vignette	2.873	(0.050)
N respondents	959	
N vignettes	7634	
AIC	32173.389	
BIC	32339.958	
ll	-16062.694	

Standard errors in parentheses
° p<0.1, * p<0.05, ** p<0.01, *** p<0.001

Figure S3: Interaction occupation and gender

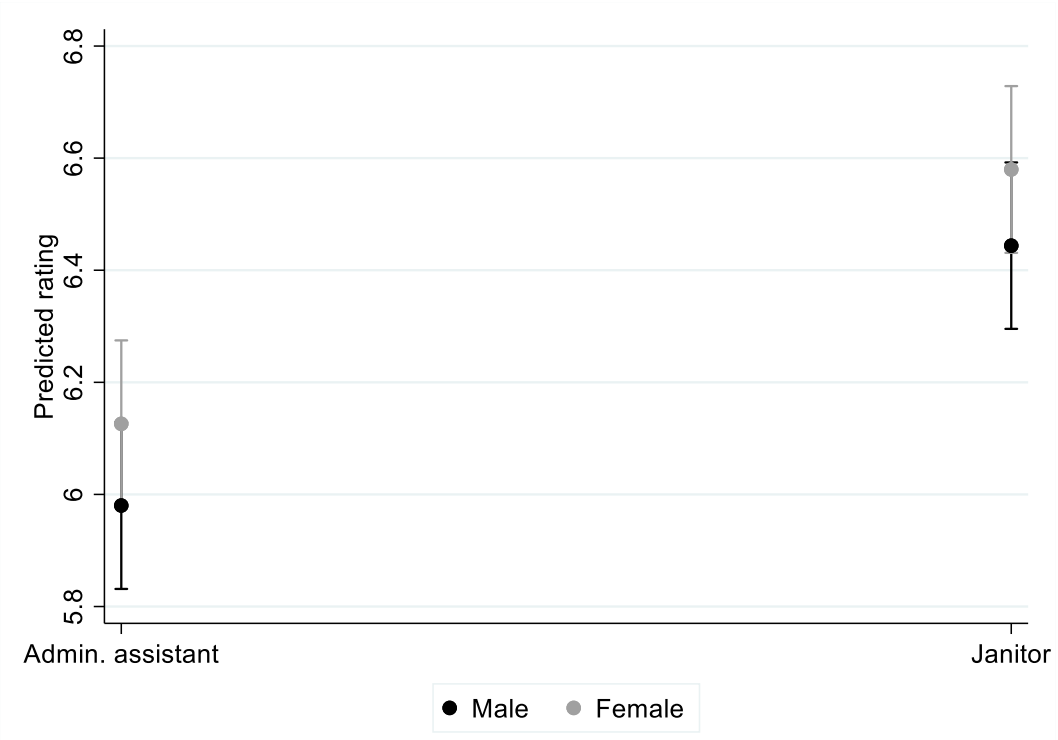

Predicted rating
5.8
6
6.2
6.4
6.6
6.8
Admin. assistant
Janitor
Male
Female

Table S7: Interaction gender and country

	Model 1	
Female (ref.)		
Male	-0.107°	(0.063)
Germany (ref.)		
Austria	-0.366*	(0.185)
Sweden	0.057	(0.162)
Interaction		
Male # Austria	-0.046	(0.102)
Male # Sweden	-0.060	(0.089)
Admin. Assistant (ref.)		
Married (ref.)		
Married+child	-0.098*	(0.048)
Single	0.041	(0.048)
24 years old (ref.)		
35 years old	-0.012	(0.049)
48 years old	-0.068	(0.048)
Syria		
Afghanistan	0.032	(0.048)
Turkey	0.015	(0.049)
A2 (ref.)		
B2	0.063	(0.049)
A2&EnglishB2	0.070	(0.049)
Arrival 2015		
Arrival 2018	-0.047	(0.040)
Teacher (ref.)		
Cleaner	-0.109°	(0.055)
Doctor	-0.002	(0.056)
Temporary work	-0.247***	(0.055)
Integration course (ref.)		
Public work	0.176**	(0.063)
Private work	0.218***	(0.063)
Subsidy	0.214***	(0.062)
Volunteering	0.170**	(0.061)
Janitor	0.459***	(0.039)
Constant	6.135***	(0.140)
Var. respondent	4.067	(0.202)
Var. vignettes	2.873	(0.050)
N	7634.000	
AIC	32174.917	
bic	32348.427	
ll	-16062.459	

Standard errors in parentheses
° p<0.1, * p<0.05, ** p<0.01, *** p<0.001

Figure S4: Interaction gender and country

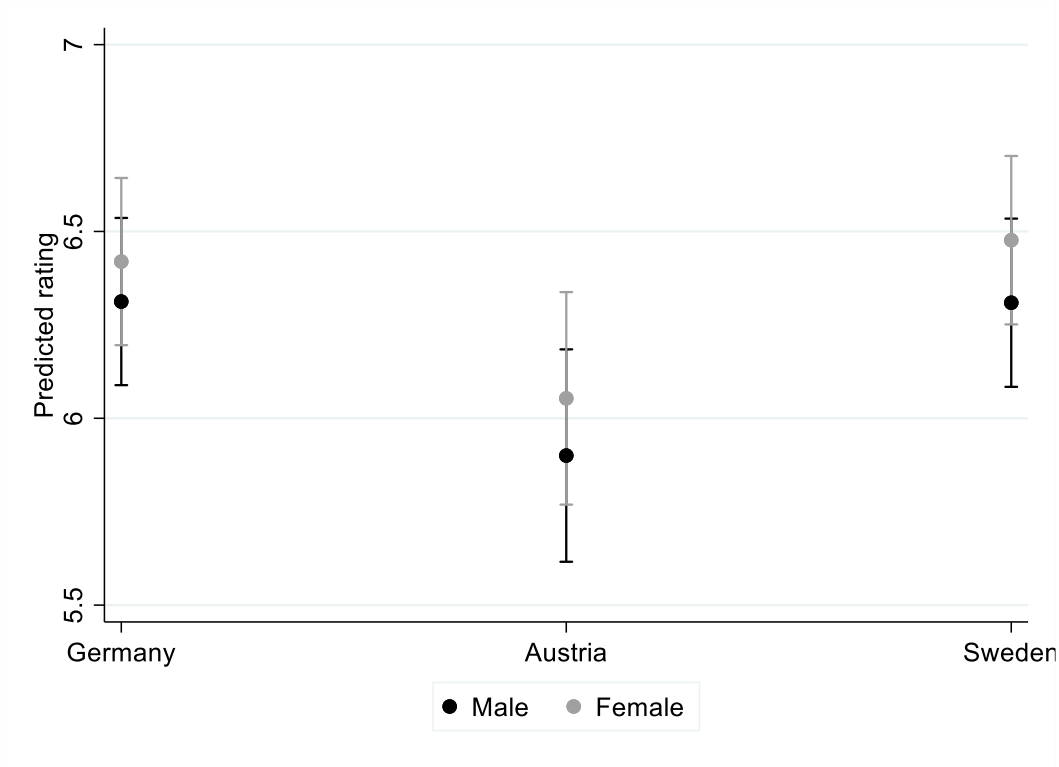


Table S8: Interaction gender, occupation and marital status

	Model 1	
Female (ref.)		
Male	-0.184	(0.115)
Married (ref.)		
Married+child	-0.148	(0.113)
Single	0.007	(0.115)
Germany (ref.)		
Austria	-0.406°	(0.214)
Sweden	0.063	(0.187)
Interaction		
Male # Austria	-0.049	(0.184)
Male # Sweden	0.032	(0.162)
Male # married+child	0.097	(0.163)
Male # single	0.131	(0.166)
Austria # married+child	-0.053	(0.182)
Austria # single	0.170	(0.185)
Sweden # married+child	-0.004	(0.159)
Sweden # single	-0.015	(0.162)
Ma e # Austria # married child	0.152	(0.261)
Male # Austria # single	-0.137	(0.267)
Male # Sweden # married+child	-0.011	(0.231)
Male # Sweden # single	-0.256	(0.234)
24 years old (ref.)		
35 years old	-0.009	(0.049)
48 years old	-0.067	(0.048)
Syria (ref.)		
Afghanistan	0.031	(0.048)
Turkey	0.015	(0.049)
A2 (ref.)		
B2	0.059	(0.049)
A2&English B2	0.069	(0.049)
Arrival 2015		
Arrival 2018	-0.046	(0.040)
Teacher (ref.)		
Cleaner	-0.107°	(0.056)
Doctor	-0.008	(0.056)
Temporary work	-0.250***	(0.055)
Integration course (ref.)		
Public work	0.171**	(0.063)
Private work	0.216***	(0.063)
Subsidy	0.211***	(0.062)
Volunteering	0.166**	(0.061)
Admin. Assistant		
Clean	0.457***	(0.039)
Constant	6.169***	(0.153)

Var. respondent	
Var. vignettes	
N respondents	
N	7634
AIC	32185.768
BIC	32428.681
Ll	-16057.884

Standard errors in parentheses
° p<0.1, * p<0.05, ** p<0.01, *** p<0.001

Fixed effects estimation

Table S9: Fixed effects estimation with clustered standard errors

	All countries Model 1		Germany Model 2		Austria Model 3		Sweden Model 4	
Female (ref.)								
Male	-0.137***	(0.041)	-0.120°	(0.065)	-0.159°	(0.093)	-0.124*	(0.061)
Married (ref.)								
Married+child	-0.114*	(0.054)	-0.114	(0.090)	-0.086	(0.117)	-0.129	(0.085)
Single	0.017	(0.054)	-0.042	(0.087)	0.212°	(0.120)	-0.041	(0.084)
24 years (ref.)								
35 years	0.010	(0.053)	0.037	(0.083)	-0.114	(0.121)	0.073	(0.083)
48 years	-0.046	(0.058)	0.055	(0.094)	-0.248°	(0.127)	-0.005	(0.088)
Syria (ref.)								
Afghanistan	0.035	(0.054)	0.001	(0.084)	0.145	(0.132)	-0.030	(0.080)
Turkey	0.025	(0.058)	-0.108	(0.100)	0.061	(0.126)	0.127	(0.085)
Lang. A2 (ref.)								
Lang. B2	0.063	(0.057)	0.124	(0.088)	0.175	(0.130)	-0.085	(0.089)
Lang. A2+English	0.061	(0.053)	0.037	(0.090)	0.199°	(0.117)	-0.001	(0.080)
Arrival 2015 (ref.)								
Arrival 2018	-0.063	(0.045)	-0.071	(0.071)	-0.158	(0.104)	-0.008	(0.068)
Teacher (ref.)								
Cleaner	-0.117*	(0.054)	0.003	(0.092)	-0.294*	(0.121)	-0.129°	(0.076)
Doctor	-0.005	(0.050)	0.066	(0.085)	-0.150	(0.109)	-0.005	(0.077)
Temporary	-0.253***	(0.051)	-0.262**	(0.084)	-0.321**	(0.115)	-0.212**	(0.075)
Integration course (ref.)								
PES work practice	0.128°	(0.073)	0.046	(0.123)	0.150	(0.168)	0.192°	(0.103)

Private work practice	0.231**	(0.073)	0.218°	(0.124)	0.102	(0.161)	0.329**	(0.104)
Subsidy	0.182*	(0.073)	0.119	(0.121)	0.225	(0.157)	0.221°	(0.115)
Volunteering	0.120°	(0.070)	0.240*	(0.116)	0.151	(0.161)	-0.025	(0.103)
Admin (ref.)								
Clean	0.457***	(0.050)	0.509***	(0.084)	0.426***	(0.115)	0.422***	(0.072)
Germany								
Austria	-0.389*	(0.179)						
Sweden	0.033	(0.155)						
Constant	6.183***	(0.144)	6.133***	(0.194)	5.870***	(0.255)	6.232***	(0.175)
N vignettes	7634		2928		1814		2892	
N respondent	959		368		228		363	
AIC	36488.918		14019.913		8878.390		13622.641	
BIC	36634.665		14133.572		8982.952		13736.065	
ll	-18223.459		-6990.956		-4420.195		-6792.320	

Standard errors in parentheses, fixed effects estimates with clustered standard errors.
° p<0.1, * p<0.05, ** p<0.01, *** p<0.001