Differential Discrimination against Mobile EU Citizens: Experimental Evidence from Bureaucratic Choice Settings

Christian Adam,
Xavier Fernández-i-Marín,
Oliver James,
Anita Manatschal,
Carolin Rapp,
Eva Thomann

Abstract

EU citizens have rights when living in a member state other than their own. Bureaucratic discrimination undermines the operation of these rights. We go beyond extant research on bureaucratic discrimination in two ways. First, we move beyond considering mobile EU citizens as homogenous immigrant minority to assess whether EU citizens from certain countries face greater discrimination than others. Second, we analyse whether discrimination patterns vary between the general population and public administrators regarding attributes triggering discrimination and whether accountability prevents discrimination. In a pre-registered design, we conduct a population-based conjoint experiment in Germany including a sub-sample of public administrators. We find that (1) Dutch and fluent German speakers are preferred, i.e. positively discriminated, over Romanians and EU citizens with broken language skills, that (2) our way of holding people accountable was ineffective, and that (3) in all these regards discriminatory behaviour of public administrators is similar to the general population.

Keywords: Accountability; Conjoint experiment; Discrimination; EU citizenship; Street-level bureaucracy

-   Author Accepted Manuscript -

Introduction

Citizens of the European Union (EU) are not only free to move to other member states to work and study. EU citizenship also confers access to public services as well as certain political rights. These rights, however, are not self-enforcing. To a large extent, the ability of mobile EU citizens to exercise their rights in other member states depends on the behaviour and routines inside their host country's public administration (Dörrenbächer, 2018; Blauberger and Schmidt, 2014). Since administrative procedures reflect a restricted interpretation of the rights associated with EU citizenship, the ability of mobile EU citizens to fully exercise their rights is limited (Heindlmaier and Blauberger, 2017; Martinsen et al., 2019; Sampson Thierry and Martinsen, 2018).

While research on the evolution and practice of EU citizenship has uncovered the struggles of mobile EU citizens to profit from EU law in practice, it tends to discuss EU citizens as a homogeneous group. Yet, in today's EU, with 27 culturally and socio-economically different member states, some EU citizens might be treated more equally to national citizens than other EU citizens. After all, frontline bureaucrats' attitudes and behaviour can affect administrative burdens, rendering access to social assistance, education, child care, and employment easier or more cumbersome (Moynihan et al. 2015). Research has documented bureaucratic discrimination by officials at the frontline of bureaucracy based on race, gender, or political beliefs (e.g., Adam et al. 2020; Grohs et al., 2016; Hemker and Rink, 2017; Schütze and Johansson, 2020; White et al., 2015).

We integrate both research perspectives by assessing potential patterns of differentiated discrimination against mobile EU citizens. Discrimination refers to unequal treatment that includes not only actions that directly harm or disadvantage certain groups, but also actions that unfairly favour other groups (Dovidio et al. 2010: 9). We examine the preferential treatment, or positive discrimination, of some mobile EU citizens relative to others by conducting a conjoint experiment in Germany with a general population sample and a sub-sample of

2

employees in the public administration. Both groups are confronted with hypothetical bureaucratic choice settings in areas where EU citizenship ensures important social and political rights: (1) access to social benefits and (2) the ability to vote in local elections. We ask experimental participants to prioritize the handling of identical requests by various mobile EU citizens. This way, we bring a behavioural perspective to the study of EU citizenship and free movement that connects the literatures on policy implementation, bureaucratic discrimination, and public accountability.

The study reported in this paper goes beyond extant research on bureaucratic discrimination in two main regards. First, studies on bureaucratic discrimination typically focus on systematic differences in the treatment of members of a minority as compared to members of the national majority (e.g., Grohs et al., 2016; Hemker and Rink, 2017; White et al., 2015). In contrast, we focus on whether within the non-German minority group certain EU citizens face greater risk of discrimination than other EU citizens. In particular, we ask whether signalling cultural and socio-economic similarity to the host country leads to preferential treatment within bureaucratic choice settings. To answer this question within the German context, we assess whether Dutch EU citizens are systematically prioritised over Romanian EU citizens and whether mobile EU citizens with good knowledge of German are systematically prioritized over EU citizens with no such language skills. Moreover, to dissect multiple dimensions of discrimination, we not only vary EU citizens' nationality and language skills but also their professional skills, age, and gender.

Secondly, our participants include both bureaucrats and people from the broader population. By asking whether bureaucratic discrimination is merely representative of societal discrimination or better treated as a distinct phenomenon, the paper addresses an important gap within research on bureaucratic discrimination and joins the few existing comparative assessments aimed at analysing whether discrimination in the public sector is more or less pronounced than in other contexts (e.g., Jilke et al. 2018).

3

Our findings challenge optimistic expectations about public administration's ability to suppress or prevent discrimination. Despite a strong rule-of-law orientation in the German public sector, we hardly find behavioural differences between the general population and respondents working within public administration: we observe negative discrimination against Romanian citizens as Dutch citizens are prioritized in bureaucratic choice settings. We also observe negative discrimination against mobile EU citizens with poor language skills as mobile EU citizens fluent in German are more likely to be prioritized. In both regards, discriminatory behaviour by public administrators was representative of discriminatory behaviour within the general population. This also holds for our attempt to reduce discrimination through an accountability treatment, which proved ineffective across both groups of respondents.

Theory and Hypotheses

Bureaucratic discrimination is discrimination by bureaucrats against clients and constitutes an essential obstacle to equality before the law with far-reaching societal consequences (Baumgartner et al. 2018; Epp et al. 2017; Thomann and Rapp 2018). We focus on differentiated discrimination against different groups within the category of EU citizens.

Generally, discrimination can result from prejudice or stereotypes. Stereotypes are specific beliefs about a group, for example, descriptions of what members of a particular group look like (Vescio and Weaver 2013), while prejudice is a consequence or result of positive, negative, conscious, or unconscious attitudes and feelings that people have about members of other groups (Allport 1954). Typically, one distinguishes different forms and sources of discrimination: e.g., statistical discrimination (stereotyping based on aggregated experiences; Arrow 1998; Phelps 1972), taste-based discrimination (Allport 1954; Becker 1971), or stereotyping based on socialization (Jilke and Tummers 2018; Schneider and Ingram 1993). We focus on the extent to which different characteristics of migrants that display their degree of integration, conformity with the prevailing culture, as well as their general phenotypical

'otherness' can trigger discrimination (Kleinpenning and Hagendoorn 1991).

Research on ethnic hierarchies (Kleinpenning and Hagendoorn 1991), social identity theory (Tajfel and Turner 2004), and on welfare chauvinism that is inspired by the idea of inter-ethnic competition theory (Hjorth 2016) all make similar suggestions: if resources are scarce, such as welfare benefits or administrative resources, individuals tend to adhere to in-group favouritism and out-group discrimination. In this sense, the likelihood of discriminatory behaviour and attitudes may thus depend on perceived similarities and differences between the outgroup and one's in-group (Valentino et al. 2019). Consequently, we expect preferential treatment, meaning positive discrimination against mobile EU citizens who come from a country that is culturally and economically close to the host country when compared to EU citizens coming from more – culturally, socially, and economically – distant member states. We assume that in terms of cultural and economic similarity, the Netherlands are more similar to Germany than Romania and hypothesize that bureaucratic choices will reflect this.

H1: Mobile EU citizens of countries that are more similar to the in-taking country are treated more favourably than EU citizens of countries that are different from the country of destination.

Migrants' nationality is not the only characteristic that may trigger discrimination. Language skills, are considered a significant marker of integration (Ager and Strang 2008; Hainmüller and Hopkins 2015). The language criterion also plays an important political role. Some civic integration policies even tie access to social benefits or legal rights to language proficiency (Goodman 2010). We thus expect that language proficiency in host country's primary language influences discrimination against mobile EU citizens.

H2: Treatment of mobile EU citizens will be more favourable when they are more proficient in their host country's language.

While extant research implicitly assumes that bureaucratic discrimination is a distinct phenomenon, it remains an open question whether public bureaucrats really differ in their degree of discrimination from the general public. Surely, bureaucratic discrimination takes place within a very distinct empirical context – i.e., during interactions between public administrators and citizens – but should we assume that this context brings forward substantially different patterns and extents of discrimination? The optimistic expectation about bureaucratic discrimination is that it is rare. This assumes that, on the one hand, legal rules confine administrative behaviour, and that people who have self-selected to work in the public sector display a common-good oriented public service motivation is opposed to discrimination (Albrow 1992; Frederickson 1990; Romzek and Dubnick 1998). On the other hand, this expectation takes into account that the general public remains highly critical of free movement (Vasilopoulou and Talving 2019) and that competition between nationals and EU immigrants over public services and rights can trigger anti-immigrant attitudes in the form of 'welfare chauvinism' (Garand et al. 2017; Hjorth 2016; Thomann and Rapp 2018). We thus hypothesize that:

H3: Street-level bureaucrats are less likely to discriminate against mobile EU citizens compared to the general population.

Finally, behavioural assessments of discrimination should be concerned with the question of how discrimination can be prevented. This relates to the general attempt to hold bureaucrats accountable and thereby reduce discrimination (Andersen and Guul, 2018; Cantarelli et al. 2020; Tetlock and Mitchell 2009). We focus on accountability mechanisms that refer to arrangements in which an actor is held accountable by a forum for discriminatory behaviour (Bovens et al. 2014; Brodkin 2008; Lieberherr and Thomann 2019). Various accountability

mechanisms have been suggested as a means to counter discriminatory practices as they decrease bias in individual decisions, encourage extensive and conscientious information processing and self-critical awareness of one's judgment processes (Cohen and Gershgoren 2016; Dobbs and Crano 2001; Epp et al. 2017; Paolini et al. 2009; Silitonga et al. 2019). Bodenhausen et al. (1994: 621) find that individuals avoid 'the influence of stereotypes in their judgments when situational factors provide a motivational impetus for such effort'. Such an effect 'is most likely to be activated when decision-makers learn prior to forming any opinions that they will be accountable to an audience a) whose views are unknown, b) who is interested in accuracy, c) who is interested in processes rather than specific outcomes, d) who is reasonably well-informed, and e) who has a legitimate reason for inquiring into the reasons behind participants' judgements (Lerner and Tetlock 1999: 259). In this vein, we focus on a simple intervention: managerial accountability, where street-level bureaucrats need to justify their behaviour toward their managers as the relevant accountability forum. Managers' authority is given by their formal position in the public administration, which formally does not allow for discrimination. We thus assume that managers are seen to disapprove of discrimination and hypothesize that knowing about the need to account for one's decisions vis-à-vis a manager will reduce discriminatory behaviour.

H4: The presence of a managerial accountability mechanism reduces discrimination.

Data and Methods

We conduct a conjoint survey experiment in Germany to test our hypotheses. The empirical focus on Germany seems justified because it represents a founding EU member state as well as the member state with the highest influx of migrants (from within and outside of the EU) over

the last years.[1] This has turned migration and mobility into highly salient and politicized topics (Grande et al. 2019).

The design was preregistered (available at: https://osf.io/fgbxu).[2] The extent to which survey experiments are able to capture discrimination remains contested as, particularly, respondents striving for socially desirable answers can undermine their external validity (see e.g., research by Wulff and Villadsen 2020; but also see the discussion in Baekgaard and George 2018). In our view, this method is the best available alternative to using field-based audit experiments, which have come to be criticised as ethically problematic since they typically lack informed consent by participants, deceive participants through fake interactions, and use up scarce public resources (James et al. 2017: 107-110; Teele 2014). The conjoint experiment allows us to consider the potentially multidimensional nature of discrimination. The experiment is conducted with a general population sample as well as with a sub-sample of respondents working within the public administration. This allows us to compare these groups' responses. While this design exploits the advantages of experimental approaches, it complies with current ethical research standards as it works based on participants' informed consent.

Our data stems from a representative survey of the general population conducted in March 2020 by YouGov in Germany.[3] The sample includes 2,974 respondents and is representative of the general population. In addition, we include 779 oversampled respondents working within the public administration. Respondents were recruited into the oversample based on a two-stage procedure: first, they were asked in which sector they work. Those working in the public sector then had to further specify the area within the public sector, i.e., police, schools, public administration. We only focus on those working within the core part of the public administration (*öffentliche Verwaltung*). This top-up sample of public administrators is not representative of the general population of people working within the public administration in Germany and does not only strictly include "frontline bureaucrats". As it is impossible to obtain a representative sample of street-level bureaucrats, our approach resembles the best available

solution to approximate such a frontline bureaucrat sample. This approximation seems legitimate as it can be plausibly assumed that frontline bureaucrats and public administrators more generally share a considerable public service motivation and rule-of-law orientation as key principles endorsed by the educational training for and on-the job in Germany.

Both samples (general public and public administration) are equally distributed in terms of gender, political ideology, and migration background. However, the public administration sample reveals a slightly higher level of both education and personal income as well as a significantly lower average age (44 vs 49).[4] These differences are unsurprising as the sub-sample only includes the working age population. Moreover, public administration employees typically go through an education process that includes at least three years of vocational training whereas the general population also includes individuals without vocational training or school diplomas. We thus assume that behavioural differences between the general public and public bureaucrats will be plausibly attributable to the latter being able to self-select into the specific context of the public administration as hypothesized above.

Tasks, Set-up, and Procedure

Respondents perform two tasks: in one task they handle a hypothetical application for the welfare programme colloquially dubbed 'Hartz IV'; a means-tested benefit for (long-term) unemployed job-seekers. In a second task, respondents handled a question about missing voting documents needed to participate in municipal elections (see Figure 1). The right of EU citizens to vote in local elections in their host member state is defined in Council Directive 94/80/EC. We asked respondents to put themselves into the position of frontline workers who, due to time constraints, have to prioritize the processing of the request of one EU citizen over the request of another from a pair of EU citizens presented to them (see figure 1). In both contexts, we describe that preferential treatment of one individual will have negative consequences for the other individual. Per task (welfare benefits vs. political rights), each respondent was

sequentially confronted with six pairs of EU citizens with randomized characteristics. To avoid order effects, we randomized the order of the tasks.

We simultaneously randomized five characteristics of the applicant to investigate their causal effect on the response. These characteristics could plausibly be part of an application procedure in the two policy contexts for the situation of a resident newly arriving in a local area in Germany. (1) nationality – Dutch vs. Romanian, with the latter being more culturally distant and potentially subject to more discrimination compared to the former. In migration research, gender, language proficiency, profession, and age are further essential aspects of migrants. Consequently, we included following characteristics of the applicants: (2) gender – male vs. female (3) language proficiency – broken vs. fluent German (4) profession – nurse vs. medical doctor (to capture the difference in the professional status of the applicant) and (5) age – 25 vs. 40 vs. 55 (to examine differences across three main categories of working age). The list of attributes represents information typically required for identification in real world applications for welfare benefits (i.e., nationality, age, gender, profession). Overall this renders 48 possible combinations of personal traits.[5] Figure 1 displays the five characteristics as well as their attributes. The order of the attributes was not randomised as research has shown that there are no order effects in conjoint analyses, except in complex settings with eight or more attributes (Auspurg and Jäckle 2017: 525). Research using eye-tracking methodology further suggests that participants do not merely look at attributes listed first but visually search across the list of attributes; particularly when this list is as short as just five attributes (Jenke et al. 2021).

Figure 1: Conjoint design

Social benefits example
'Please consider a scenario in which you have a job working for the local government in your municipality. You receive two applications for a social benefit, the so-called Hartz IV. As EU-citizens, both applicants are eligible to apply for this social benefit. Due to personal time constraints, you are required to make a choice of which applicant you process first, which will lead to a delay in the payment to the other applicant.'

Voting rights example
'Please consider a scenario in which you have a job working for the local government in your municipality. You receive two requests concerning missing vote documents for the upcoming local election. As EU-citizens, both persons are eligible to vote in this election. However, without the necessary documents, they will not be able to cast their vote. Due to personal time constraints, you are required to make a choice of which request you process first, which will lead to the other person cannot vote in the upcoming election.'

	Applicant (Request) 1	Applicant (Request) 2
Nationality	Dutch	Romanian
Gender	female	male
Language proficiency	broken German	fluent German
Profession	nurse	medical doctor
Age	25 years old	40 years old

Which of the two applications (requests) will you process first?
<1>   Applicant (Request) 1
<2>   Applicant (Request) 2

Additionally, we introduced an accountability treatment. Specifically, we inform a treatment group that they will have to afterwards justify their decisions toward their putative manager. The control group receives no such information. We randomly assigned respondents to either treatment or control group: 1,875 (49.96%) respondents received the treatment, and 1,878 (50.04%) did not receive the treatment. The median time taken to read the information with the treatment is approximately one second longer for both tasks, which by itself is unlikely to affect participants' interaction with the experimental materials.

Analysis

For each task, we compare respondents' choices between successively presented pairs of EU citizens. These choices are made under three different conditions: (1) they are made by different kinds or respondents (public administrators vs. general population), (2) with and without an accountability treatment (treatment shown vs. not shown), and (3) in two different policy

contexts presented by the two different tasks (welfare benefits vs. political rights). Across both samples, we have a total of 3,753 individuals who have prioritized a total of 12 pairs of profiles (6 for each task). We thus obtain a total of (3,753 individuals * 12 decisions * 2 EU citizen profiles =) 90,072 binary responses. These responses are not independent, but rather depend on the three conditions aforementioned. Therefore, we employ a hierarchical structure (population, treatment and outcome) to partially pool the responses of the logistic regression, allowing all groups to borrow strength from the main effects.

We follow Hainmueller et al.'s (2014) approach to estimate the Average Marginal Component Effects (AMCEs) for the effects of interest. This allows us to estimate effects without having to rely on a specific decision-making process and functional form assumptions. In other words, we do not have to assume that individuals follow a particular behavioural model (e.g., the utility maximization or bounded rationality). We make the assumption of stability and no carryover effects (this assumption implies that in the course of the experiment previous decisions do not alter following ones), as the profiles are randomized. We do, however, control for profile-order effects and find that, on average, by outcome, population, and treatment situation, profiles that are presented first in the pair are 36% more likely to be prioritized. We employ Bayesian inference to get the posterior distributions of the parameters of interest (Fernández-i-Marín 2016).

Results

Figure 2 shows the posterior distributions of the main θ parameters of interest. The dot represents the median expected effect, and the band covers 95 percent of the credible interval containing that percentage of probable values (also called the highest posterior density, HPD). The parameter estimates are analogous to the Average Component Marginal Effects (AMCEs) in Hainmueller et al. (2014), defined as the 'causal quantity of interest', and are effects of each of the individual treatment components. To present our results graphically, we use the treatment

12

of a reference profile, which is a 40-year-old Romanian female with broken German and working as a doctor, as the intercept. As long as the credible intervals for an attribute (e.g., nationality) overlaps with the intercept, there is no differential treatment between the reference profile and a person who is different in terms of this attribute (e.g., has a different nationality). We present odds ratios. Values >1 indicate higher odds of positive discrimination (preference for prioritization) over the reference profile, whereas values <1 point towards higher odds of negative discrimination (preference against prioritization).

Figure 2 presents the results concerning preferences in the policy field of welfare distribution (i.e., welfare benefits) on the left side and the preferences in the policy field of political rights on the right-hand side. Above this, the upper part of Figure 2 refers to the results under the non-treatment condition, and the lower part shows the results for respondents who received the managerial accountability treatment. Overall, our model correctly predicts 62% of the hypothetical decisions made in the context of voter registration (political rights) and 60% of the decisions on processing unemployment benefits (social policy).

Figure 2: Posterior distributions of the main θ parameters of interest.
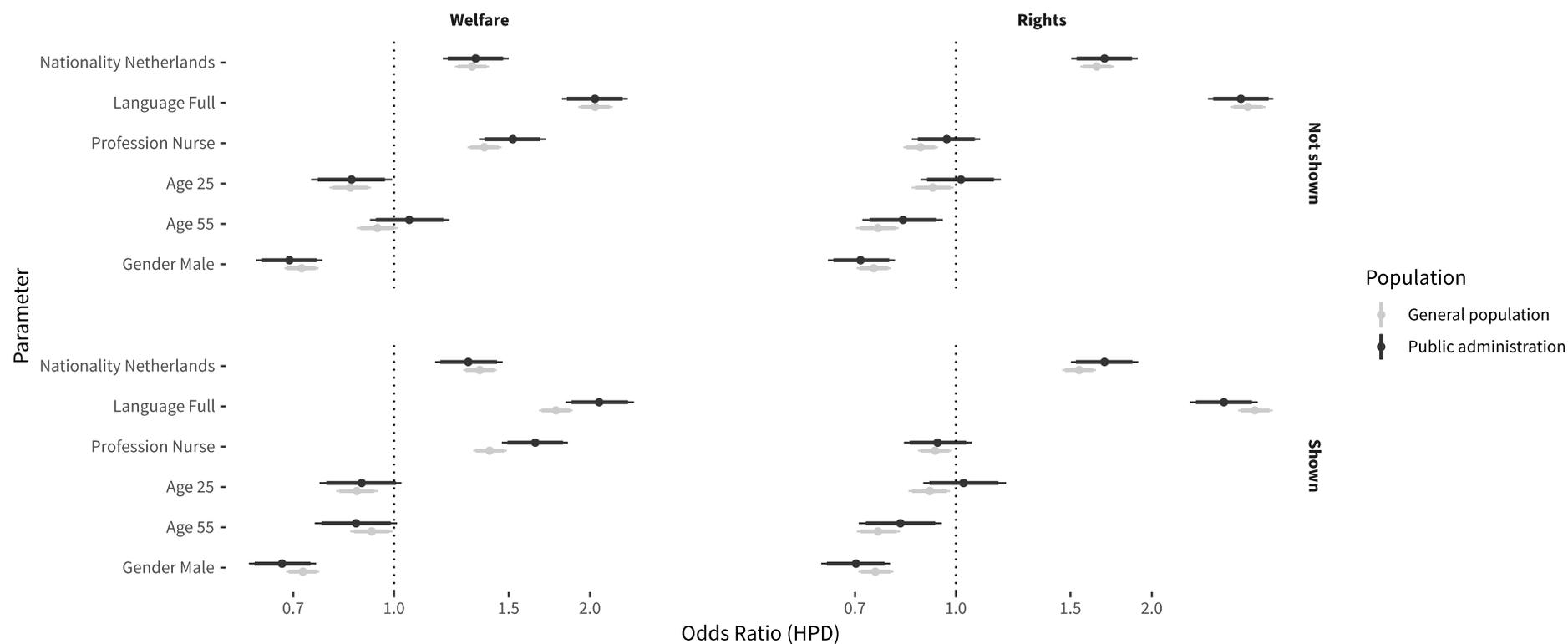
Notes: Dots represent the expected odds, and thick and thin lines, the 90 and 95 percent Bayesian credible intervals, respectively. Colours represent different population groups.

The upper row corresponds to effects without treatment and the lower row to the treatment group. The columns separate the different policies (welfare/rights).

As hypothesized (H1) and independent of the responding group – general public vs. public administration – the results support our first hypothesis: there is a clear tendency to prefer Dutch applicants over Romanian applicants. This is shown in Figure 2 by the larger than 1 odds ratio of the nationality variable in the four conditions (both policy fields and treatment/non-treatment condition). Table 1 presents the combined expected effect magnitude of all attributes; again across all four conditions: Dutch applicants are 44% more likely to be prioritized than Romanian applicants.

The attribute with the largest effect on the odds of being prioritized is language proficiency (see table 1). Profiles with full German proficiency are 2.1 times (117%) more likely to be prioritized than profiles with little knowledge of German. This supports hypothesis 2 and its underlying assumption that the degree of integration, as demonstrated by language proficiency, is a crucial trigger for discrimination. This is also reflected in figure 2: Irrespective of the policy sector, the responding group, or treatment condition, language proficiency triggers discrimination (odds ratio >1).

Table 1: Odds ratios of expected grand-effects

Feature	Odds Ratio	Expected effect
Language – fluent	2.17407	▲ 117%
Nationality – Dutch	1.44029	▲ 44%
Gender – Male	0.73550	▼ 26%
Profession – Nurse	1.14477	▲ 14%
Age 55	0.86526	▼ 13%
Age 22	0.92439	▼ 7.6%
First shown	1.35513	▲ 36%

While the patterns of discrimination are quite similar within both policy sectors, there remain interesting differences between them. E.g., the prioritization of Dutch applicants as well as the prioritization of applicants fluent in German, is more pronounced when processing voter registration requests (political rights) than when processing welfare benefits (welfare).

Moreover, while men are generally less often prioritized than women in both policy sectors, the unemployment requests of nurses are generally prioritized over those of doctors. While results for the effect of age are less consistent, applicants in the middle age group tend to be prioritized over older and younger applicants (age 55 and 25).

In contrast to hypothesis 3, we find no behavioural differences between the general population and people working within the public administration. Figure 2 displays distinct credible interval bands for these two groups in different shades to allow for easy comparison between two groups. Unsurprisingly, public administration effects have wider uncertainty bands, as their sample is smaller compared to our general population sample. Generally, effects for the two populations overlap, indicating similar distributions of the estimated parameters (e.g., no differential effect).

Finally, we also find no support for hypothesis 4, which dealt with the potential impact of holding people accountable for their decisions vis-à-vis a manager or supervisor. We rearrange the findings displayed in figure 2 with in figure 3 for better visual assessment of this question. Our accountability treatment does not influence discriminatory behaviour: the parameter distributions stay almost identical across both conditions. The only observable difference occurs when respondents of the general population are asked to process welfare benefit requests. Here, the accountability treatment seems to reduce the extent of discrimination against non-fluent speakers to some extent. However, the accountability treatment does not result in an overall behavioural change within either population. In some regards, bureaucrats and the general population react differently, with bureaucrats reacting more to the treatment stimulus but in a non-intended way, by making their preferences slightly more marked in the case of age.

More specifically, when the accountability treatment is shown, discrimination in voter registration contexts increases on nationality (with 92% certainty, against our expectations) but decreases for the professional characteristics (88.5% certainty) and language (72% certainty). On the other hand, discrimination when processing welfare benefit requests decreases on nationality (74% certain) when the accountability treatment is shown. For the rest of the cases, including public administration, the evidence of differential effects is fragile.

Figure 3. Posterior distributions of the main θ parameters of interest
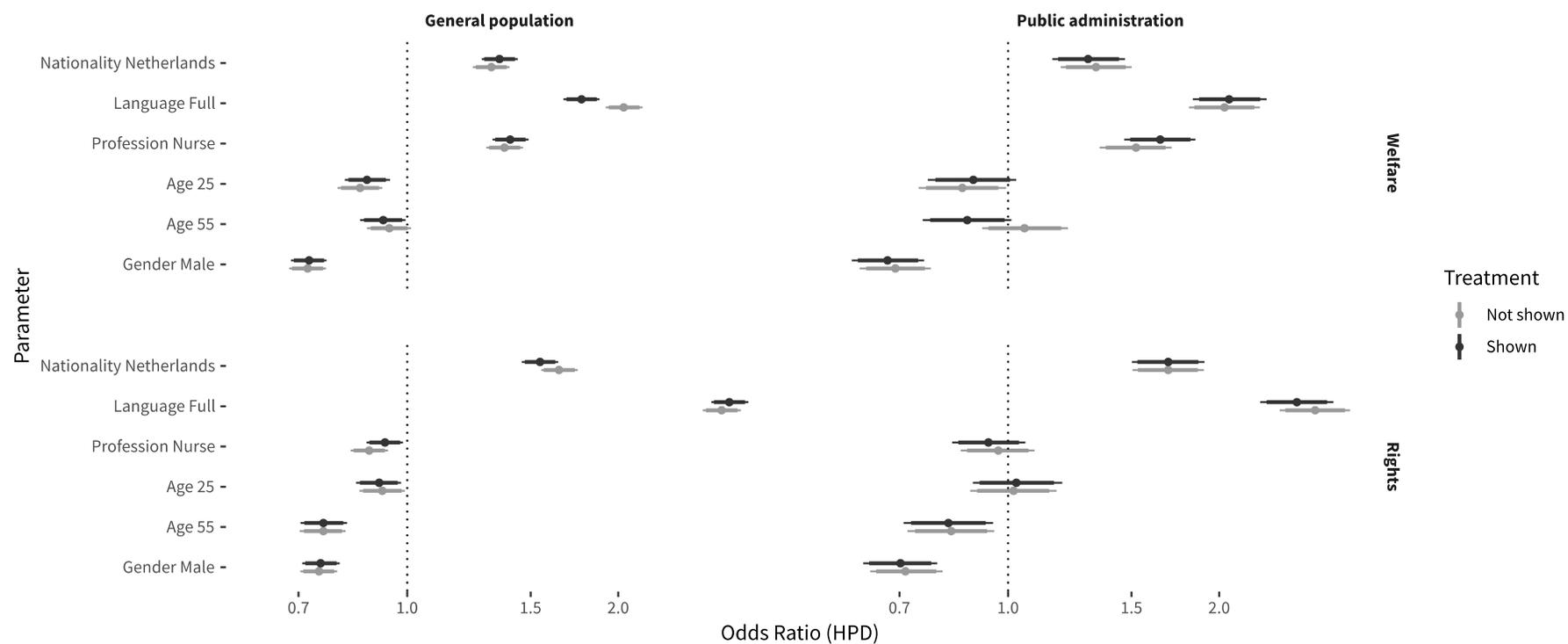
Notes: Dots represent the expected odds, and thick and thin lines, the 90 and 95 percent Bayesian credible intervals, respectively. Shades of colour represent different treatment situations. The upper row corresponds to effects on welfare and the lower row on rights. The columns separate the two populations (general population and public administration).

Conclusion

This paper pursued two main objectives: (1) to assess whether some mobile EU citizens face a greater risk of discrimination than others and (2) to analyse whether discrimination patterns vary between the general population and public administrators. Our findings show that the strongest triggers of discrimination are migrants' nationality and language skills. Dutch migrants are more likely to receive preferential treatment than Romanian migrants. This supports the argument that cultural, social, and economic similarity gives rise to ethnic hierarchies that result in differential patterns of discrimination. Also, migrants fluent in the host country's language are more likely to receive preferential treatment than less proficient migrants. Interestingly, discrimination is slightly more pronounced when processing voter registration than unemployment requests. This might suggest that political rights cut deeper into the fear that immigrants from more culturally and economically different countries may get more rights and, as a consequence thereof, more influence in one's country (Vernby 2013). Moreover, a simple accountability treatment administered during the experiment proves unable to suppress discrimination. In all of these regards, respondents working within public administration show strikingly similar patterns of discriminatory behaviour as respondents representing the general public. This finding challenges the implicit assumption underlying existing research that bureaucratic discrimination is a distinct phenomenon. While bureaucratic discrimination takes place within a distinct setting – i.e., within the public administration – and therefore warrants explicit analysis, we find no evidence for a distinct pattern of magnitude of discrimination within the public sector compared to the general public. Overall, this provides no support for an optimistic view, which expected that a rule-of-law orientation and public service motivation would reduce or even prevent discrimination.

Moreover, our findings support extant research on discrimination against migrants, which highlights that not all migrants face the same challenges but face differential treatment based on arbitrary characteristics such as provenance (Dancygier et al. 2015; Zschirnt and Ruedin

2016) or religious orientation (Di Stasio et al. 2019, Valentino et al. 2019). Our results thus seem to align more with a pessimistic view, which observes a transformation of the unique idea of free movement of persons in the EU to just another form of "immigration" which can be subject to selectivity and exclusion (Barbuescu and Favell 2020).

Our study yields certain policy implications: if nationality is a crucial trigger of discrimination, anonymity in bureaucratic encounters could be of value. This is, however, very difficult to implement in bureaucratic encounters. After all, eligibility for voter registration and welfare benefits is established based on people's identity. While changes towards more anonymity in the process might not be impossible, directing efforts at quick and proficient language acquisition is likely a more fruitful approach. Our results thus support calls promoting quick language acquisition (Goodman 2010). Since it is mainly in the early period after their arrival that immigrants have to interact with the administration, assistance from specialized international offices might help to reduce the burdens for newly arrived immigrants and mobile EU citizens by providing information in English, or sometimes even in their native languages. Multilingual information has been shown to be an important instrument to facilitate equal access, e.g., to voting rights (Filindra and Manatschal 2019).

Limitations of this study create the need for future research. Whilst our findings caution against accountability as a panacea, we need to assess the effectiveness of different accountability treatments within field experiments and within survey experimental settings before discarding accountability as not important; particularly, as accountability is shown to be an important treatment in other contexts (e.g., Christensen and Moynihan 2020). Second, although we were unable to conduct our experiment in the field with actual frontline bureaucrats who process voter registration and welfare benefit requests, we were able to use an oversample of people working within the public administration. This is a major advance on representativeness compared to using student samples or exclusively general population samples. However, if non-trivial problems of access can be overcome, future research's external validity would benefit

from assessing bureaucratic discrimination in the field at the administrative frontline without recurring to ethically problematic audit experiments. This requires, however, significant political efforts to gain access to these organizations. Moreover, the pattern that Dutch enjoy preferential treatment over Romanians in the German context might have other or additional causes than cultural and socio-economic similarity. Future research could further support the relevance of the similarity-mechanism by including several countries with varying degree of similarity. Finally, the study has not looked at potentially heterogeneous effects based on individual and contextual characteristics. Future research should explore whether regional differences and place-based stereotyping affect discrimination patterns.

Biographical Notes

Christian Adam is Professor of Comparative Politics at Zeppelin University, Germany, ORCID: https://orcid.org/0000-0002-1755-5243

Xavier Fernández-i-Marín is Lecturer at the LMU Munich, Germany, ORCID: https://orcid.org/0000-0002-9522-8870

Oliver James is Professor of Political Science at Exeter University, UK, ORCID: https://orcid.org/0000-0001-6031-2581

Anita Manatschal is Assistant Professor of Migration Policy Analysis at the University of Neuchâtel, Switzerland,

Carolin Rapp is Assistant Professor of Political Science at the University of Copenhagen, Denmark, ORCID: https://orcid.org/0000-0001-5152-6799

Eva Thomann is Professor of Public Administration at the University of Konstanz, Germany, ORCID: https://orcid.org/0000-0001-6837-4787

Contact Details

Prof. Dr. Christian Adam, Am Seemooser Horn 20, DE-88045 Friedrichshafen, Germany,

christian.adam@zu.de

Notes

1. See https://data.oecd.org/migration/permanent-immigrant-inflows.htm, https://migrationdataportal.org/data?i=inflow_freemovement&t=2017 (last accessed: April 30, 2020).

2. We commit to our preregistered plan. However, we branched out one hypothesis to be addressed within its own paper dealing with potentially heterogeneous effects for different kinds of respondents. This decision was not made based on the results obtained but due to space constraints and conference feedback.

3. Data collection took place between February 21 and March 16. The survey was fielded at the same time as the COVID-19 crisis accelerated in Germany but finished before Germany issued strict social distance measures on March 22.

4. Average age of public sector workers in Germany is 44.5 years: http://bit.ly/3sFKaNF; last accessed 15 Oct 2020.

5. We conducted a pre-test based on 101 respondents to test the understanding of the wording of the questions as well as the quality of responses.

Supplementary Material

Accessible at [link].

References

Adam, C., Grohs. S. and Knill, C. (2020) 'Discrimination based on political beliefs: A field experiment on the freedom of assembly', *Public Policy and Administration,* 1-22.

Ager, A. and Strang, A. (2008) 'Understanding Integration: A Conceptual Framework', *Journal of Refugee Studies* 21(2): 166-91.

Albrow, M. (1992) 'Sine ira et studio — or do organizations have feelings?', *Organization Studies*, 13(3): 313–29.

Allport, G.W. (1954) *The nature of prejudice,* Cambridge: Addison Wesley.

Andersen, S.C. and Guul, T.S. (2018) 'Reducing minority discrimination at the front line—combined survey and field experimental evidence', *Journal of Public Administration Research and Theory,* 29(3): 429-44.

Arrow, K.J. (1998) 'What Has Economics to Say about Racial Discrimination?', *The Journal of Economic Perspectives,* 12(2): 91–00.

Auspurg, K. and Jäckle, A. (2017) 'First Equals Most Important? Order Effects in Vignette-Based Measurement'. *Sociological Methods & Research,* 46(3):490-539.

Baekgaard, M. and George, B. (2018) 'Equal Access to the Top? Representative Bureaucracy and Politicians' Recruitment Preferences for Top Administrative Staff', *Journal of Public Administration Research and Theory,* 28(4): 535-550.

Baumgartner, F.R., Epp, D.A. and Shoub, K. (2018) *Suspect citizens: What 20 million traffic stops tell us about policing and race*. Cambridge: Cambridge University Press.

Barbuescu, R. and Favell, A. (2020) 'Commentary: A citizenship without social rights? EU freedom of movement and changing access to welfare rights, *International Migration,* 58(1): 151-65.

Becker, S.G. (1971) *The economics of discrimination,* Chicago: University of Chicago Press.

Blauberger, M. and Schmidt, S.K. (2014) 'Welfare migration? Free movement of EU citizens

and access to social benefits', *Research & Politics,* 1(3): 1-7.

Bodenhausen, G.V., Kramer, G.P. and Süsser, K. (1994) 'Happiness and stereotypic thinking in social judgment', *Journal of Personality and Social Psychology,* 66(4): 621-32.

Bovens, M., Schillemans, T. and Goodin, R.E. (2014) 'Public Accountability', in Bovens, M., Goodin, R.E. and T. Schillemans (Eds.), *The Oxford Handbook of Public Accountability,* Oxford: Oxford University Press, pp. 1-20.

Brodkin, E.Z. (2008) 'Accountability in street-level organizations', *International Journal of Public Administration,* 31(3): 317-36.

Cantarelli, P., Bellé, N. and Belardinelli, P. (2020) 'Behavioral Public HR: Experimental Evidence on Cognitive Biases and Debiasing Interventions, *Review of Public Personnel Administration,* 40(1): 56-81.

Christensen, J. and Moynihan, D. P. (2020) 'Motivated reasoning and policy information: politicians are more resistant to debiasing interventions than the general public', *Behavioural Public Policy*: 1-22.

Cohen, N. and Gershgoren, S. (2016) 'The incentives of street-level bureaucrats and inequality in tax assessments', *Administration & Society,* 48(3): 267-89.

Dancygier, R., Lindgren, K., Oskarsson, S. and Vernby. K. (2015) 'Why are immigrants underrepresented in politics? Evidence from Sweden', *American Political Science Review,* 109 (4): 703-24.

Decker, O. and Brähler E. (2018) *Flucht ins Autoritäre - Rechtsextreme Dynamiken in der Mitte der Gesellschaft*, Giessen: Psychosozial-Verlag.

Di Stasio, V., Lancee, B., Veit, S. and Yemane, R. (2019) Muslim by default or religious discrimination? Results from a cross-national field experiment on hiring discrimination. *Journal of Ethnic and Migration Studies,* 1-22.

Dobbs, M. and Crano, W.D. (2001) 'Outgroup accountability in the minimal group paradigm: Implications for aversive discrimination and social identity theory', *Personality and Social*

Psychology Bulletin, 27(3): 355-64.

Dörrenbächer N. (2018) 'Frontline uses of European Union (EU) law: a parallel legal order?, *Journal of Public Policy,* 38(4): 455-79.

Dovidio, J.F., Hewstone, M., Glick, P. and Esses, V.M. (Eds.) (2010). *The SAGE Handbook of Prejudice, Stereotyping and Discrimination*. Thousand Oaks: SAGE.

Epp, C.R., Maynard-Moody, S. and Haider-Markel, D. (2017). 'Beyond profiling: The institutional sources of racial disparities in policing, *Public Administration Review,* 77(2): 168-78.

Fernández-i-Marín, X. (2016) 'ggmcmc: Analysis of MCMC samples and Bayesian inference', *Journal of Statistical Software,* 70(9): 1-20.

Filindra, A. and Manatschal, A. (2019) 'Coping with a changing integration policy context: American state policies and their effects on immigrant political engagement', *Regional Studies,* 1-12.

Frederickson, H.G. (1990) 'Public Administration and Social Equity', *Public Administration Review,* 50(2): 228–37.

Garand, J.C., Xu, P. and Davis, B.C. (2017) 'Immigration Attitudes and Support for the Welfare State in the American Mass Public', *American Journal of Political Science,* 61(1): 146–62.

Goodman, S. (2010) 'Integration requirements for integration's sake? Identifying, categorizing and comparing civic integration policies', *Journal of Ethnic and Migration Studies,* 36(5): 753-72.

Grande, E., Schwarzbözl, T. and Fatke. M. (2019) 'Politicizing immigration in Western Europe', *Journal of European Public Policy,* 26(10): 1444-63.

Grohs, S., Adam, C. and Knill, C. (2016) 'Are some citizens more equal than others? Evidence from a field experiment, *Public Administration Review,* 76(1): 155-64.

Hainmueller, J., Hopkins, D.J. and Yamamoto, T. (2014) 'Causal inference in conjoint analysis: Understanding multidimensional choices via stated preference experiments', *Political Analysis,*

22(1): 1-30.

Heindlmaier, A. and Blauberger, M. (2017) 'Enter at Your Own Risk: Free Movement of EU Citizens in Practice' *West European Politics,* 40(6): 1198-1217.

Hemker, J. and Rink, A. (2017) 'Multiple dimensions of bureaucratic discrimination: Evidence from German welfare offices, *American Journal of Political Science,* 61(4): 786-803.

Hjorth, F. (2016) 'Who benefits? Welfare chauvinism and national stereotypes', *European Union Politics,* 17(1): 3-24.

Jenke, L., Bansak, K., Hainmueller, J. and Hangartner, D. (2021) 'Using Eye-Tracking to Understand Decision-Making in Conjoint Experiments', *Political Analysis*, 29(1): 75-101

Jilke, S. and Tummers, L. (2018) 'Which clients are deserving of help? A theoretical model and experimental test', *Journal of Public Administration Research and Theory,* 28(2): 226-38.

Jilke, S., Van Dooren, W. and Rys, S. (2018) 'Discrimination and Administrative Burden in Public Service Markets: Does a Public-Private Difference Exist?', *Journal of Public Administration Research and Theory,* 28(3): 423-39.

James, O., John, P. and Moseley, A. (2017) 'Field Experiments in Public Management', in James, O., Jilke, S., and G. Van Ryzin (Eds.), *Experiments in Public Management Research: Challenges and Contributions*, Cambridge: Cambridge University Press, pp. 89-116.

Kleinpenning, G. and Hagendoorn, L. (1991) 'Contextual aspects of ethnic stereotypes and interethnic evaluations', *European Journal of Social Psychology,* 21(4): 331-48.

Lerner, J.S. and Tetlock, P.E. (1999) 'Accounting for the effects of accountability', *Psychological Bulletin,* 125(2): 255-75.

Lieberherr, E. and Thomann, E. (2019) 'Street-level bureaucracy research and accountability beyond hierarchy', in Hupe, P. (Ed), *Research Handbook on Street-Level Bureaucracy: The Ground Floor of Government in Context,* Cheltenham: Edward Elgar.

Martinsen D.S., Blauberger M., Heindlmaier A. and Sampson Thierry, J. (2019) 'Implementing European case law at the bureaucratic frontline: How domestic signalling influences the

outcomes of EU law', *Public Administration,* 97(4): 814-28.

Moynihan, D., Herd, P. and Harvey, H. (2015) 'Administrative burden: Learning, psychological and compliance costs in citizen-state interactions', *Journal of Public Administration Research and Theory,* 25(1): 43–69.

Paolini, S., Crisp, R.J. and McIntyre, K. (2009) 'Accountability moderates member-to-group generalization: Testing a dual process model of stereotype change, *Journal of Experimental Social Psychology,* 45(4), 676-85.

Phelps, E.S. (1972) 'The statistical theory of racism and sexism', *The American Economic Review,* 62(4): 659–61.

Romzek, B.S. and Dubnick, M.J. (1998) 'Accountability' in Shafritz, J.M. (Ed.), *International Encyclopedia of Public Policy and Administration*, Boulder: Westview Press, pp. 6–11.

Sampson Thierry, J. and Martinsen D.S. (2018) 'Lost in translation: how street-level bureaucrats condition Union solidarity', *Journal of European Integration,* 40(6): 819-34.

Schneider, A.L. and Ingram, H.M. (1993) 'Social construction of target populations: Implications for politics and policy', *The American Political Science Review,* 87(2): 334–47.

Schütze, C. and Johansson, H. (2020) 'The importance of discretion for welfare services to minorities: Examining workload and anti-immigration attitudes', *Australian Journal of Public Administration,* 79: 426-443.

Silitonga, M.S., van Duijn, M., Heyse, L. and Wittek, R. (2019) 'Setting a good example? The effect of leader and peer behaviour on corruption among Indonesian senior civil servants', *Public Administration Review,* 79(4): 565–79.

Tajfel, H. and Turner, J.C. (2004) 'The social identity theory of intergroup behavior', in Jost, J.T. and J. Sidanius (Eds.), *Political Psychology: Key Readings*, New York: Psychology Press, pp. 276-93.

Thomann, E. and Rapp, C. (2018) 'Who deserves solidarity? Unequal treatment of immigrants in Swiss welfare policy delivery', *Policy Studies Journal,* 46(3): 531-52.

Teele, D.L. (2014) 'Reflections on the ethics of field experiments' in D.L. Teele (Ed.), *Field Experiments and Their Critics: Essays on the Uses and Abuses of Experimentation in the Social Sciences*, New Haven: Yale University Press, pp. 115-140.

Tetlock, P.E. and Mitchell, G. (2009) 'Implicit bias and accountability systems: What must organizations do to prevent discrimination?' *Research in Organizational Behavior* 29: 3-38.

Valentino, Nicholas A., Stuart N. Soroka, Shanto Iyengar, Toril Aalberg, Raymond Duch, Marta Fraile, Kyu S. Hahn, Kasper M. Hansen, Allison Harell, Marc Helbling, Simon D. Jackman, and Tetsuro Kobayashi. (2019) 'Economic and cultural drivers of immigrant support worldwide', *British Journal of Political Science,* 49(4): 1201-26.

Vasilopoulou, S. and Talving, L. (2019) 'Opportunity or threat? Public attitudes towards EU freedom of movement', *Journal of European Public Policy,* 26(6): 805-23.

Vernby, K. (2013) 'Inclusion and Public Policy: Evidence from Sweden's Introduction of Noncitizen Suffrage', *American Journal of Political Science,* 57(1): 15-29.

Vescio, T.K. and Weaver, K. (2013). *Prejudice and Stereotyping*, Oxford: Oxford University Press.

Vogel D. (2011) *Dem Gemeinwohl verpflichtet? - Was motiviert die Beschäftigten des öffentlichen Dienstes?,* Potsdam: Universitätsverlag Potsdam.

White A.R, Nathan, N.L. and Faller, J.K. (2015) 'What do I need to vote? Bureaucratic discretion and discrimination by local election officials, *American Political Science Review,* 109(1): 129-42.

Wulff, J. N. and Villadsen, A. R. (2020) 'Are Survey Experiments as Valid as Field Experiments in Management Research?', *European Management Review, 17*(1), 347-356.

Zschirnt, E. and Ruedin, D. (2016) 'Ethnic discrimination in hiring decisions: a meta-analysis of correspondence tests 1990–2015', *Journal of Ethnic and Migration Studies,* 42(7): 1115-34.

Word count:

6.860 words (title, abstract, notes, references) + (2.4 pages of figures/tables * 475 words=)

1.140 = **8.000 words**.