Jnl Soc. Pol. (2018), 47, 1, 39–56 © Cambridge University Press 2017
doi:10.1017/S0047279417000113

Discrimination Against Ethnic Minorities in Activation Programme? Evidence from a Vignette Experiment

LARS INGE TERUM*, GAUTE TORSVIK AND EINAR ØVERBYE*****

*Centre for the Study of Professions, Oslo and Akershus University College, Oslo, Norway
email: larsinge.terum@hioa.no
**Department of Economics, University of Oslo, and Centre for the Study of Professions, Oslo and Akershus University College, Oslo, Norway
email: gaute.torsvik@econ.uio.no
***Department of Social Work, Child Welfare and Social Policy, Oslo and Akershus University College, Oslo, Norway
email: einar.overbye@hioa.no

Abstract

We conducted a vignette experiment involving 470 Norwegian front-line workers to investigate whether their decisions to sanction non-compliance of activation requirements varied with the ethnicity of the welfare claimant. This is the first vignette experiment on ethnic discrimination in the administration of activation programmes in Europe. The study shows that front-line workers did not sanction claimants with a North African name more often than claimants with a native Norwegian name. However, among front-line workers who had experience with the relevant activation programme, a male claimant with a North African name was sanctioned less often than a male claimant with a native Norwegian name. Thus, we find some degree of reverse discrimination on the part of experienced front-line workers. This finding is contrary to a similar US vignette experiment that detected discrimination (not reverse discrimination) with regard to claimants with an ethnic minority name. The most likely explanation for the difference concerns the different institutional-cultural contexts within which Norwegian and US social policy programmes operate.

Introduction

Discretion, which is an inescapable feature of most administrative decisions, may be performed in ways that discriminate against specific categories of claimants. Two social policy trends make the risk of ethnic discrimination particularly salient. The first trend, known as *activation turn*, imposes activation requirements on welfare recipients and sanctions on those who fail to comply (Watts *et al.*, 2014). Activation increases the discretionary power of front-line workers (Brodkin, 2013). The second trend is increased net migration into affluent democracies. Migration increases ethnic heterogeneity and changes the composition of welfare recipients; hence, front-line workers are increasingly in a position where they may

https://doi.org/10.1017/S0047279417000113 Published online by Cambridge University Press

discriminate against or in favour of ethnic minorities. Discrimination occurs when two similar applicants are treated differently *because* of characteristics that should not be relevant, such as ethnicity. The aim of this study is to examine whether front-line workers discriminate when they impose sanctions on participants in an activation programme.

Despite its importance, few studies have investigated discrimination in the implementation of social policies. A reason may be the methodological challenges involved in studying discrimination. In a recent study, Campbell (2015) found that teachers in England systematically evaluated the abilities of ethnic minority pupils below their actual scores on standardised tests of reading and maths compared with other pupils in the sample. Campbell interpreted the findings as indicating negative stereotyping of such pupils by teachers. Comparing the actual abilities of a group of people with how others evaluate their abilities is an effective way to detect negative stereotyping. This method requires independent assessments of actual abilities and a measure of the evaluations that others make about their abilities (be they teachers, employers or other individuals). When such information is unavailable, an alternative strategy involves using an experimental design. An experimental design measures discrimination directly by investigating whether the assessment or treatment of a person varies with the identity markers (e.g. ethnicity and gender) of that individual. To investigate ethnic discrimination, researchers must choose a research design that aims to hold everything constant, apart from the ethnicity of the applicant.

In this study, we use a vignette experiment to investigate whether front-line workers[1] make different decisions regarding how to treat clients depending on the clients' ethnicity and gender. In a survey sent to front-line workers in the Norwegian Labour and Welfare Administration (NAV), we included a vignette of a claimant who did not comply with the mandatory activation requirements for the Qualification Programme, a Norwegian welfare programme. Front-line workers who affirmed that they had experience with the programme read the vignette and were asked to decide whether to sanction the claimant. To investigate whether their decision varied with the ethnicity of the claimant, we randomly varied the name of the claimant in four equal samples.

Experimental studies of discrimination

Numerous studies have used experimental designs to study discrimination in labour and housing markets but almost none have been undertaken with regard to discrimination in social welfare programmes. We identified only two earlier studies that used an experimental design to study discrimination in welfare programmes. Before we discuss these studies, we will review the findings with regard to discrimination in markets.

https://doi.org/10.1017/S0047279417000113 Published online by Cambridge University Press

Discrimination in markets

According to Becker (1957), discrimination in labour markets occurs if employers treat workers (or job applicants), with identical individual productivity characteristics, differently (less favourably) *because* they belong to a specific group (such as ethnicity).

Economists differentiate between taste-based discrimination and so-called statistical discrimination, initially described by Phelps (1972) and Arrow (1973).[2] First, employers may simply dislike members of a certain group; if so, discrimination is based on taste. Second, employers with no bias against a group may nevertheless find it 'rational' – in a profit-maximising sense – to discriminate against members of the group if (1) they have limited information about the skills of the individual job applicants in the group and (2) the group indicator carries information about the value of the unobserved productivity parameters. In the latter case, employers use observable group characteristics such as ethnicity to infer an expected (in a statistical sense) value of the productivity parameters in situations where they lack individual information.

Two types of experiments are used to study discrimination: audit experiments and correspondence experiments. In audit experiments, the experimenter instructs real persons (actors) with different ethnicities to present the same information and behave in the same manner towards an employer. In correspondence experiments, written résumés that are identical except for the ethnicity of the jobseekers are sent to employers.

Several audit and correspondence experiments have examined discrimination in labour markets and housing markets.[3] Most have been conducted in the US, but an increasing number are being conducted in Europe, including Scandinavia. Norwegian and Swedish studies have shown that an ethnic minority name reduces the likelihood that a landlord contacts a prospective tenant who has responded to a rent advertisement. The negative effect is greater for men than for women (Andersson *et al.*, 2012; Bursell, 2014). An ethnic minority name similarly reduces the likelihood of a job applicant being granted a job interview. Here the negative effect is also greater for men than for women (Midtbøen and Rogstad, 2012). The pattern is the same in US studies (Bertrand and Mullainathan, 2004).

The discovery of greater negative effects for ethnic minority men than for ethnic minority women suggests that social categories such as gender and ethnicity should not be studied in isolation. The study of English teachers, referred to in the Introduction, found that identity markers such as ethnicity and gender intersected in creating stereotypes (Campbell, 2015). Stereotypes appear to be intersectional, implying that identity markers should be studied in different combinations (see Bursell, 2014). Consequently, we differentiated not only between native and ethnic minority names in our experiment but also between male and female names.

https://doi.org/10.1017/S0047279417000113 Published online by Cambridge University Press

Discrimination in social programmes

Few previous studies have used vignette experiments to study possible discrimination in the implementation of social programmes. One study used a survey experiment to investigate gender bias in the allocation of public long-term care for elderly persons in Norway (Jakobsson *et al.*, 2015). It found that a female applicant with an adult daughter was offered less help than a female applicant with an adult son. Apparently, Norwegian front-line workers who allocate long-term care expect female applicants to receive more help from a daughter than from a son. Hence, they allocate less public care to female applicants with daughters, although they are not formally allowed to consider the gender of adult children. This can be interpreted as a form of statistical discrimination.

Another study used a vignette experiment to investigate sanctions in a welfare programme targeting low-income mothers among 137 caseworkers in Florida (Schram *et al.*, 2009). It found that a claimant with a Spanish or African-American name was sanctioned more often than a claimant with a Caucasian name if the claimant violated the activation requirements in the programme. Discrimination was more evident if ethnicity (name) was coupled with stereotypical information aimed at identifying the ethnic minority to which the participant belonged.

Assumptions and social policy context

Unlike the US, Norway has no large African-American or Spanish-speaking minority. However, minorities are over-represented in welfare programmes. North Africans (particularly Somalis) constitute the largest minority group on social assistance. Norwegian ethnographic studies indicate that front-line workers experience Somalian men to be particularly demanding (Engebrigtsen, 2005; Fangen, 2008). Somalian recipients have also reported being treated arbitrarily and disrespectfully and experiencing personally invasive behaviour by front-line workers (Friberg and Elgvin, 2014). This suggests that we should find similar discrimination in Norway towards claimants with North African names, as Schram *et al.* (2009) found with regard to Spanish and (in particular) African-American names in the US.

However, substantial institutional and cultural differences exist between the welfare systems of the US and Norway, which may dampen or reduce an eventual similar tendency in Norway. First, Norway redistributes income on a larger scale, and social programmes are more generous and reach a larger proportion of citizens. Norwegian 'active' welfare programmes are to a lesser extent workfare strategies directed mainly towards ethnic minorities (Alesina and Glaeser, 2004; Scruggs, 2008; Scruggs and Allan, 2006). Our initial expectation was that ethnic discrimination could also be identified in Norway but to a lesser extent than in the US. We further assumed that Norwegian front-line workers would treat North African male claimants more harshly than they would treat North African female

https://doi.org/10.1017/S0047279417000113 Published online by Cambridge University Press

claimants because previous research on ethnic discrimination in the Norwegian housing and labour markets had revealed greater discrimination towards ethnic minority men than towards ethnic minority women.

The Qualification Programme

With the aforementioned hypotheses in mind, we believe the Qualification Programme (QP) in Norway to be very well suited for studying the enforcement of sanctions. The QP was established in 2007 as the main policy instrument to fight poverty and social exclusion (Gubrium *et al.*, 2014; Norwegian Ministry of Labour and Inclusion, 2006). The aim of the programme is to improve the labour market attachment of claimants who have complex problems and cannot immediately be integrated into the labour market but, nonetheless, are deemed capable of working. The programme targets individuals who are long-term recipients of social assistance. Unlike social assistance, the QP is not only a benefit scheme but also a full-time activation programme, where claimants are referred to as *participants*. Each participant in the programme has a right to an individually designed weekly plan that involves 37.5 hours of extensive training, counselling and related activities geared towards increasing their opportunities of finding ordinary work. Whilst they are in the programme, which usually lasts for two years, they are entitled to a special Qualification Allowance (QA). Unlike social assistance, the QA is not a means-tested benefit. Instead, it provides a regular income equal to the minimum state pension. In addition to the QA, participants may be entitled to child allowances and supplementary social assistance benefits on a means-tested basis.

The QP is administered by the NAV. The NAV administers all types of social security benefits in Norway, with recipients comprising approximately 20 per cent of the working-age population. A majority of front-line workers in the NAV have higher education, and more than 40 per cent are trained as social workers (Hagelund and Terum, 2015). The QP has the highest caseworker/client ratio of all programmes in the NAV.

In principle, signing up for the QP is voluntary. However, in practice, turning down an offer from the NAV to join the programme can be difficult. To be eligible, recipients must have a permanent address in Norway and undergo a work capability assessment (Gjersøe, 2016). The average age of participants is 34 years, 44 per cent are women, and just over half are non-native Norwegians. The largest group comprises immigrants from Somalia (Lima and Naper, 2013).

In the QP, participants are placed in a typical 'throffer' situation, where a threat is accompanied by an offer (Goodin, 1998; Molander and Torsvik, 2015). The assumption is that participants perceive the QA to be more attractive than social assistance, owing to the predictability of income and the lack of a means test. In addition, participants receive a great deal of extra follow-up and help that are not offered to those on social assistance. This is crucial for the construction of

https://doi.org/10.1017/S0047279417000113 Published online by Cambridge University Press

sanctions. In the carrot and stick approach for motivation, the greater generosity and predictability of the scheme represent 'carrots' for those who sign up for the programme. The scheme is also the basis for the threat of sanctions ('stick'). Participants who do not comply with the activation requirements can lose part or all of the benefit. *Weak sanctions* imply a reduction in the QA. *Strong sanctions* imply being ejected from the programme and losing the QA. Claimants who are ejected may still be eligible for regular social assistance.

Data and methods

As part of a larger survey sent to front-line workers in the NAV, we included an experimental vignette. Front-line workers read nearly identical versions of a summary of a person who had violated the rules in the QP. One group of front-line workers received a vignette with a typical Norwegian male name, a second with a typical Norwegian female name, a third with a North African male name and a fourth with a North African female name. They were then asked whether the violations should lead to ejection from the programme.

The vignette

The participant was married, and the name of the partner was chosen to indicate the same ethnic group as the QP participant. The name of the participant was mentioned nine times in the vignette, and the name of the partner was mentioned once. We repeated the name to make the gender and ethnicity of the participant salient. This is the vignette (translated from Norwegian):

X [Anja, Jens, Amina, Jamir] is a 37-year-old [man/woman] who has received various social assistance benefits during the last four years, primarily economic support. [He/she] applied to attend the Qualification Programme (QP) after some pressure from the social assistance office and initially received the right to an eight-month stay in the programme. There are now four months left.

At the first meeting with [his/her] caseworker after having entered the programme, [he/she] seemed active and motivated to obtain paid work. During the following weeks, this changed. The caseworker detected several signs that X did not much care about adhering to the programme rules. X sometimes did not respond to text messages and did not always show up for scheduled talks. [He/she] is signed up for a course that includes on-the-job training, but the organiser of the course has reported that [he/she] has not turned up at the workplace. The course organiser has therefore cut part of the QA money, because of the participant's failure to show up.

The caseworker has informed X several times that the programme can be stopped if [he/she] does not follow the agreed plan. X has then given varying and diffuse reasons for not showing up at the workplace. In the contact with X, it has been revealed that the [husband/wife], X, feels depressed. [He/she] works in the canteen at a hospital and is often on short-term sickness leave. The couple have a two-year-old boy and a four-year-old girl. Both children attend kindergarten. X says [he/she] feels torn between the expectations of the family and the QP (NAV).

https://doi.org/10.1017/S0047279417000113 Published online by Cambridge University Press

A while ago, a letter was sent to X, informing [him/her] that the programme will be stopped if [he/she] does not follow up according to plan. Since then, NAV has not heard anything from [him/her]. Nor has X replied to a text message from the caseworker.

After reading this summary, the front-line workers were asked two questions (translated from Norwegian): (1) 'Given this information, do you think a decision should be made to cease participation in the QP? Yes or No'. (2) 'Does X [Anja, Jens, Amina, Jamir] resemble cases you often encounter? Yes or No'

We developed the vignette in consultation with experienced NAV front-line workers and conducted a pilot test with staff who had experience with the QP.

The Survey

The survey was distributed online to 2,747 front-line workers working in 108 municipalities. The municipalities were a random sample of Norway's 428 municipalities. The response rate was 63 per cent (1,735 front-line workers). Only those NAV front-line workers who confirmed they had experience working with the QP received the vignette, and 470 responded to the vignette.

We used a random number generator to allocate the participant identity (gender and ethnicity) to the vignettes the front-line workers received. The randomisation procedure was successful. None of the variables influencing the decision to sanction is correlated with the identity of the person in the vignette. This implies a balanced sample in each of the 'treatments' (i.e. identities of the vignette person).[4]

Results

Because the identity of the vignette participant was randomly allocated to the front-line workers, we can identify discrimination simply by comparing mean sanction rates across the types of participants. Table 1 reports the fraction of front-line workers sanctioning each participant.

Table 1 shows that most front-line workers (on average, 81 per cent) would remove the participant from the programme. No discrimination based on the gender or ethnicity of the participant was detected. The fraction of front-line workers who would sanction the non-native vignette participants (Jamir and Amina) is slightly *lower* than the fraction that would sanction the native vignette participants (Jens and Anja), but the difference is not statistically significant.

We also performed a regression analysis where we controlled for several caseworker and workplace characteristics (gender, age and number of participants; education of the caseworker; size of the office; and fraction of the

https://doi.org/10.1017/S0047279417000113 Published online by Cambridge University Press

TABLE 1. Fraction of caseworkers who decided to sanction the participant

	Jamir	Amina	Jens	Anja	Non-Native	Native	Total
Fraction who sanction (*SD*)	0.76 (0.43)	0.83 (0.38)	0.81 (0.39)	0.83 (0.38)	0.80 (0.40)	0.82 (0.38)	0.81 (0.39)
n =	105	125	129	111	230	240	470

workforce receiving health-related social benefits in the municipality where the office is located). We constructed an indicator variable based on whether the caseworker agreed with the following statement: 'The generous social security benefits in Norway encourage migrants to come to the country'[5]. The regression analysis confirms the finding in Table 1; that is, front-line workers do not discriminate based on the ethnicity and gender of the participant, not even those front-line workers who believe generous social security benefits attract migrants.

All NAV front-line workers who received the vignette had experience with the QP but not everyone agreed that the vignette represented a familiar case. On average, 66 per cent of the front-line workers found the vignette to be familiar. The degree of familiarity did not vary across the gender and ethnicity of the vignette client. Next, we examined whether caseworkers, who were acquainted with participants who violate activation requirements, sanction differently from those who were *not* familiar with this type of behaviour. Here we find differences between the front-line workers. Those who were familiar with the case described in the vignette sanctioned far less often than those who did not recognise this as a familiar case. Table 2 (last row) shows that 90 per cent of the caseworkers who were unfamiliar with such participants decided to eject the participant from the programme compared with only 76 per cent of those who were familiar with such participants. The difference (14 percentage points) in the overall propensity to sanction is statistically significant. (The Fisher exact test statistic for equality in the propensity to sanction across these groups gives a p value of .0019.)

This finding corresponds with a finding reported by Schram *et al.* (2009), who found that experienced caseworkers sanction participants less often than inexperienced caseworkers do. Schram *et al.* offered two explanations from a US context. First, caseworkers with greater experience may have witnessed the negative effects of sanctioning on the future lives of participants – they have a stronger premonition that ejecting a participant will result in a difficult or hopeless life situation. Second, caseworkers with greater experience may be more motivated to support participants. Although our question concerns whether the caseworkers were familiar with participants who violate activation requirements, rather than whether the caseworkers were more 'experienced', these hypotheses

https://doi.org/10.1017/S0047279417000113 Published online by Cambridge University Press

TABLE 2. Fraction of caseworkers who decided to sanction, in terms of familiarity with the vignette

	Familiar		Unfamiliar	
	Sanction (*SD*)	*n*	Sanction (*SD*)	*n*
Jamir	0.67 (0.47)	70	0.94 (0.34)	35
Amina	0.81 (0.39)	79	0.86 (0.34)	46
Jens	0.78 (0.42)	88	0.90 (0.30)	41
Anja	0.79 (0.41)	71	0.90 (0.30)	41
Total	0.76 (0.42)	308	0.90 (0.29)	163

may also explain our findings. Caseworkers who are familiar with participants who violate activation requirements may also be more aware that sanctioning involves conflicts with clients and extra paperwork. In addition, a selection effect is possible: front-line workers who continue to work with participants who break the rules (rather than apply for other work in NAV or leave NAV altogether) may be a select group of front-line workers who find working with 'difficult' participants to be rewarding.

Schram *et al.* (2009) found that experienced caseworkers discriminated against Spanish and African-American participants compared with Caucasian participants.[6] We, however, find no such pattern. Table 2 shows that among caseworkers who were not acquainted with participants similar to those described in the vignette, there is no statistical difference in sanctioning. A large majority (between 86 and 94 per cent) want to eject the participant, regardless of gender or ethnicity. The only exception concerns front-line workers who are acquainted with such cases. They are actually less inclined to sanction Jamir (the North African male) than the other participants. If we aggregate Amina, Jens and Anja and compare them with Jamir, the difference in the propensity to sanction between 'familiar and unfamiliar' is 13 percentage points, with a p value of .035 (Fisher exact test). Thus, contrary to our initial hypothesis – that is, the North African male is sanctioned more often – we find either no discrimination (among caseworkers unfamiliar with such cases) or a tendency to be *less* strict towards North African male claimants (among caseworkers familiar with such cases).[7]

In summary, ethnic minority participants are not sanctioned more harshly in Norway if they do not comply with the activation requirements in the QP; nor are men sanctioned more than women. Front-line workers are strict regardless of the gender or ethnicity of the participant. However, front-line workers who are unfamiliar with participants who violate activation requirements are much harsher than front-line workers who have previous experience with such participants. Front-line workers who are familiar with such participants are less

https://doi.org/10.1017/S0047279417000113 Published online by Cambridge University Press

likely to sanction, and they are likely to treat a male participant with a non-native name more leniently than other participants. The latter can be perceived as reverse discrimination. Individuals who do not meet the requirements of the programme are less likely to be sanctioned if they are males from an ethnic minority group, provided that their caseworker has previous experience with rule-breaking participants.[8]

Discussion

Our findings raise several questions. Why do vignette experiments reveal discrimination in the labour and housing markets in Norway but not in social programmes? Why did the surveyed Norwegian front-line workers who reported familiarity with rule-breaking participants sanction Jamir *less* often than they did Jens, Anja and Amina? Finally, why does Norway differ from the US, where Schram *et al.* (2009) found negative discrimination against ethnic minorities, in particular African-Americans? In the following subsections, we suggest some answers to these questions.

Why does discrimination differ between markets and public welfare?

As aforementioned, vignette experiments – in Norway and elsewhere – have revealed that landlords and employers discriminate against tenants or jobseekers with a non-native name, and that such discrimination is particularly prevalent if the person is a male. Why do we find no such pattern in the QP?

First, perhaps the risk-reward calculus is different for a front-line worker than for an employer or a landlord. By *risk-reward calculus*, we mean the balance between the perceived risks and rewards associated with making a decision. If it differs between actors, they are likely to choose different courses of action. Front-line workers who administer welfare programmes have fewer incentives to discriminate, because they are not at risk of incurring any type of loss by selecting the 'wrong' recipient for sanctioning. Landlords, by contrast, risk something if they rent an apartment to someone who later turns out to be a troublesome tenant, because a tenant – once chosen – can be difficult to evict. Similarly, employers risk something if they employ a jobseeker who subsequently turns out to be an inefficient worker, because it can be costly and difficult to lay off an employee later. Therefore, landlords and employers have greater reasons to discriminate, in the form of statistical discrimination, compared with front-line workers. In addition, front-line workers may be punished harder than landlords or employers if they are found to discriminate based on ethnicity and/or gender. Front-line workers who exhibit such behaviour may destroy their career prospects, and even risk being fired. Landlords and employers may also face negative reactions if they discriminate against a minority but these reactions are less likely to result in loss of work or livelihood.

https://doi.org/10.1017/S0047279417000113 Published online by Cambridge University Press

Why some reverse discrimination?

If Norwegian caseworkers believe that accusations of discrimination may jeopardise their career prospects, they may be sensitive to a mere *suspicion* that they discriminate against ethnic minorities. Working within the Norwegian welfare system, they may 'overcompensate' by being especially lenient towards ethnic minority claimants – in particular male claimants, since male immigrants receiving welfare benefits are arguably the group most widely suspected by the media of being subjected to public discrimination. This hypothesis turns the initial suspicion of discrimination on its head: the judgement of caseworkers is possibly coloured by general stereotypes regarding which groups are discriminated against, but this colouration may work in a way that *reduces* sanctioning of such claimants.

In addition, Norwegian social workers, who constituted the majority of caseworkers, are sensitised to the plight of ethnic minorities during their training. They are taught a cognitive frame that more often portrays welfare claimants as victims of circumstances beyond their control, rather than as strategic 'takers' who cannot be trusted and must be disciplined (Halvorsen *et al.*, 2013). In short: Norwegian caseworkers are possibly concerned with avoiding accusations of ethnic discrimination, and with preventing exclusion of ethnic minorities, to such an extent that some degree of reverse discrimination may even take place.

Why does discrimination in public welfare systems differ from the US?

The aforementioned study by Schram *et al.* (2009) found negative discrimination of ethnic minority claimants in the US welfare system. Why do results in Norway differ?

We do not have the data to do a comparative analysis of the diverging results between the US and the Norwegian experimental studies. However, we can offer some hypotheses as suggestions for further research.

First, the research design is not strictly the same. The Schram *et al.* study included two vignettes depicting single mothers: one Latina/Caucasian single mother and one African American/Caucasian single mother. Both were contrasted to a Caucasian (white) single mother. Schram *et al.* detected discrimination against the African-American claimant in particular (relative to the Caucasian claimant), and discrimination was more pronounced when the vignette contained information that the claimant had previously violated activation requirements. Based on the findings, Schram *et al.* (2009: 398) claimed the following: 'Overall, our study clarifies how racial minorities, especially African Americans, are more likely to be punished for deviant behaviour in the new world of disciplinary welfare provision'.

In contrast to this study, we did not present the respondents with different versions of the vignette; we did not vary the extent to which the claimant had previously violated activation requirements. However, our vignette depicts a case

https://doi.org/10.1017/S0047279417000113 Published online by Cambridge University Press

where the claimant has previously violated the activation requirements several times. Hence, our vignette corresponds to the worst-case vignette described by Schram *et al.* In addition, the sample in our study is larger (N = 470) than the sample used by Schram *et al.* (N = 137), thereby allowing us to obtain a more precise estimate of discrimination. Thus, although the research designs differ somewhat, we find it unlikely that this is the main reason why the results differ.

Alternative explanations must necessarily be speculative. One possibility concerns how ethnic minority welfare claimants are cognitively 'framed' by mass media and in the public discourse. In the US, 'welfare' is strongly associated with ethnic/racial characteristics (Alesina and Glaeser, 2004; Schram *et al.*, 2003). Most recipients of Temporary Assistance for Needy Families are poor single mothers and a large proportion are Hispanics and African Americans. Mass media, and the public discourse, often portray claimants as 'strategic takers'. There can be an ethnic component to this portrayal, in particular with regard to African-Americans (Schneider and Ingram, 1993; Gilens, 1999; Schram *et al.*, 2009). Norwegian media, on the other hand, tend to portray claimants more as victims of circumstances beyond their control, and the public debate focuses on the desirability of social integration of ethnic minorities (Bay and Stang, 2009). If such general stereotyping processes influence caseworkers, this may explain the diverging results. In this context, it can be noted that a larger percentage (>80 per cent) of Norwegian front-line workers have tertiary education, compared with approximately 60 per cent of US caseworkers in the study by Schram *et al.* (2009).[9] Previous studies indicate that people with higher education tend to be less ethnocentric – although these studies have not looked specifically at caseworkers (Van Der Waal *et al.*, 2010). However, and drawing in the opposite direction, concerns about civil rights and non-discrimination has a longer history, and therefore probably a stronger and more salient role, in the US than Norway.

Another possible explanation could be the different design of the welfare system itself, rather than the broader social culture. Decisions made by front-line workers are likely to depend on the historical and administrative contexts of the particular street-level organisation they work within (Watkins-Hayes, 2013). US welfare schemes are separated from social security programmes serving the majority. This is less the case in Norway. QP participants in Norway are served by the same welfare agency and have the same caseworkers as the majority. The QP is administered by NAV, which is responsible for all social security benefits and most social services. Everyone older than 66 years and approximately 20 per cent of the working-age population (16–66 years) are served by the same NAV caseworkers (Terum and Hatland, 2014). Different institutional contexts may impact on the selection of caseworkers, as well as on how they perceive their clients. Perhaps the broader client base in the Norwegian system somehow reduces the 'otherness' of ethnic minority claimants.

https://doi.org/10.1017/S0047279417000113 Published online by Cambridge University Press

A third set of possible explanations concerns how work is organised *within* welfare agencies. Despite the similar role as programmes of 'last resort', the role caseworkers play in administering such welfare programmes, through tailoring support to the actual needs of recipients, varies considerably cross-nationally (Jewell, 2007, 181). In the US, welfare programmes represent bureaucratic organisations, staffed by non-professionals who make eligibility assessments based on extensive regulations from above. Jerwell (op.cit.) contrasts the US with Sweden, where social assistance resembles a professional organisation, staffed by social workers with a large amount of discretionary power. Norway is similar to Sweden in this regard. US caseworkers may have less autonomy and be held to a stricter targets-and-performance system than Norwegian caseworkers (Brodkin, 1997, 2013). Following Lipsky (1980), larger professional discretion on behalf of Norwegian caseworkers may imply they are more autonomous in reducing the 'hassle' of sanctioning, including the risk of exposing themselves to complaints and criticism from clients that are sanctioned. Perhaps they are particularly sensitive about not exposing themselves to complaints from ethnic minority claimants in this regard. If so, the larger degree of autonomy among Norwegian caseworkers compared to US caseworkers may contribute to an explanation of the different empirical findings.

As already mentioned, we do not have sufficient empirical material to test these (partly competing) hypotheses; they are suggested avenues for further cross-national research with regard to how caseworkers treat ethnic minority claimants.

Concluding comment

Schram *et al.* (2009) used their findings to develop the so-called Racial Classification Model (RCM), as a theoretical tool to understand caseworker behaviour related to ethnic minorities. The RCM model is based on an underlying premise that salient social classifications and group representations are necessary for individuals (here: caseworkers), in order to avoid information overload and bring coherence to the complexities of daily decision making. According to Schram *et al.* (2009), the likelihood that such cognitive classifications will be ethnically charged is a function of the prevailing cultural stereotypes of ethnic groups, the extent to which policy actors (caseworkers in the present study) hold relevant group stereotypes, and the presence or absence of stereotype-consistent cues.

The Racial Classification Model (RCM) may be a fruitful tool in further cross-national studies of possible ethnic discrimination by caseworkers. The different Norwegian findings may also be investigated within the framework of this model. We have argued that three factors may influence the decisions caseworkers make: (1) differences in cultural stereotypes of ethnic groups, (2) the

https://doi.org/10.1017/S0047279417000113 Published online by Cambridge University Press

extent to which caseworkers (for organisational or other reasons) hold relevant group stereotypes, and (3) the presence or absence of stereotype-consistent cues (as caseworkers, with varying degree of autonomy, decide what to do).

It would be an exciting prospect for future research to investigate the following more exactly and in a larger number of countries: how ethnic discrimination is influenced by differences in cultural stereotypes, the extent to which caseworkers hold relevant group stereotypes, and the presence or absence of stereotype-consistent cues. For this future research, experimental designs will be of great value.

Acknowledgements

This study was made possible by grants from the Research Council of Norway's Welfare, Working Life and Migration (VAM) programme. We gratefully acknowledge the following individuals for their insightful feedback on an earlier version of this article: Jonathan Bradshaw, University of York; Mike Adler, University of Edinburgh; Flemming Larsen, Aalborg University; and Silje B. Fekjær and Anders Molander, Oslo and Akershus University College. We also thank the anonymous referees at the *Journal of Social Policy* for their helpful comments.

Notes

1 We use the terms *front-line worker* and *caseworker* as synonyms for workers who deal directly with welfare claimants. In Norway, eight in ten possess a higher education, the most common type being social work education.

2 The distinction that economists make between taste-based discrimination and statistical discrimination corresponds roughly to a difference between affective-based preferences and cognitive-based preferences (Zajonc and Markus, 1982).

3 See Fix and Turner (1998) for a discussion of audit studies; Bertrand and Mullainathan (2004) for an influential correspondence study; and Heckman and Siegelman (1993) and Neumark (2012) for assessments of how useful these methods are for detecting discrimination in markets.

4 Appendix Table A1 verifies that the user characteristics (ethnicity and gender) in the vignette were randomly allocated to caseworkers. Column 1 shows that caseworker characteristics and factors in their work environment are good predictors of whether the caseworker decides to sanction the participant in the vignette. Column 2 examines whether the ethnicity in the vignette can be predicted by these same characteristics. That is not the case. Nor can they predict the gender in the vignette. Hence, the user characteristics in the vignette do not vary systematically with any of the caseworker/environment variables that predict sanctions. This verifies successful randomisation of ethnicity and gender across caseworkers.

5 In the vignette, the sanction decision (*sanction*) is a binary variable. Caseworkers could decide either to remove the participant from the programme or not to do so. Let *non-native* be an indicator variable that is equal to 1 if the participant in the vignette that caseworker *i* evaluates has a North African name, and X is a vector that contains various caseworker and workplace characteristics. We estimate the following equation:

$$sanction_i = \gamma + Z_i\boldsymbol{\beta} + \alpha\, non\text{-}native_i + \epsilon_i$$

In this case α does indeed capture the causal effect of ethnicity on sanctions. We can use a regression model to test whether gender discrimination exists. We then swap *non-native* with

https://doi.org/10.1017/S0047279417000113 Published online by Cambridge University Press

female, an indicator for female participant. We can also interact *female* and *non-native* to test whether gender-specific ethnic discrimination exists.

To examine whether those who agreed that welfare benefits in Norway encourage migrants to come are more inclined to sanction the non-native participant in the vignette, we constructed a variable *migcome*, which is equal to 1 if caseworkers agree or totally agree with the statement that welfare benefits attract immigrants. We interact this variable with *non-native* and estimate the following equation:

$$sanction_i = \gamma + Z_i\boldsymbol{\beta} + \alpha non - native_i + \delta migcome + \lambda non\text{-}native * migcome + \epsilon_i$$

The interaction term captures a difference in difference (DID) estimate: do those who believe that high welfare transfers attract immigrants to Norway sanction the non-native participant in the vignette harder than those who do not believe this to be the case?

6 The study by Schram *et al.* (2009) considered single female claimants with children (lone mothers) in Florida. By contrast, males have access to the same welfare programmes, on equal terms, as women (with or without children) in Norway.

7 A general worry is that researchers may induce an *experimenter effect* on the survey respondents (caseworkers). The worry here is that those who receive a vignette with an ethnic minority participant (Jamir or Amina) become alerted to the discrimination issue and make decisions accordingly. To test that hypothesis, we examined whether those who received a vignette with a minority participant responded differently to value-laden questions in the survey ('the generous social security benefits in Norway encourage migrants to come to the country' or 'a generous welfare state is costly for the economy'). They did not.

8 An important topic is whether the vignette methodology yields a 'realistic' situation for caseworkers. Audit studies more realistically capture a real-life decision-making situation. However, it is impossible to instruct actors in audit studies to behave in exactly the same way towards a caseworker. Vignette experiments are better at ensuring that everything else is held constant. If errors caused by 'low realism' in vignette experiments are randomly distributed between the experimental groups, differences between these groups can be given a substantial interpretation.

9 More than 80 per cent of caseworkers administering the QP reported having tertiary education at the first degree level or above compared with 60 per cent of the US caseworkers in the study by Schram *et al.* (2009: 404).

References

Alesina, A. and Glaeser, E. L. (2004), *Fighting Poverty in the US and Europe: A World of Difference*, Oxford: Oxford University Press.

Andersson, L., Jakobsson, N. and Kotsadam, A. (2012), 'A field experiment of discrimination in the Norwegian housing market: gender, class and ethnicity', *Land Economics*, 88: 2, 233–40.

Arrow, K. J. (1973), 'The theory of discrimination', in O. Ashenfelter and A. Rees (eds.), *Discrimination in Labor Markets*, Princeton, NJ: Princeton University Press, pp. 3–33.

Bay, A. H. and Stang, E. (2009), 'Politisk svikt eller individuelt ansvar? Pressens portrettering av fattigdom', in A. H. Bay, A. West Pedersen and J. Saglie (eds.), *Når velferd blir politikk*, Oslo: Abstrakt forlag, pp. 159–201

Becker, G. S. (1957), *The Economics of Discrimination*, Chicago: University of Chicago Press.

Bertrand, M. and Mullainathan, S. (2004), 'Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination', *American Economic Review*, 94: 4, 991–1013.

Brodkin, E. Z. (1997), 'Inside the welfare contract: discretion and accountability in state welfare administration', *Social Service Review*, 71: 1, 1–33.

https://doi.org/10.1017/S0047279417000113 Published online by Cambridge University Press

Brodkin, E. Z. (2013), 'Street-level organizations and the welfare state', in E. Z. Brodkin and G. Marston (eds.), *Work and the Welfare State: Street-Level Organizations and Workfare Politics*, Washington, DC: Georgetown University Press,

Bursell, M. (2014), 'The multiple burdens of foreign-named men – evidence from a field experiment on gendered ethnic hiring discrimination in Sweden', *European Sociological Review*, 30: 3, 399–409.

Campbell, T. (2015), 'Stereotyped at seven? Biases in teacher judgement of pupils' ability and attainment', *Journal of Social Policy*, 44: 3, 517–47.

Engebrigtsen, A. (2005), 'Somaliere i Norge. Noen trek fra tilpassning i eksil', in K. Nes, T. Skoug and M. Strømstad (eds.), *Somaliere i Norge: perspektiver på integrering, språk og religion*, Vallset: Opplandske bokforlag, pp. 168–179.

Fangen, K. (2008), *Identitet og praksis: etnisitet, klasse og kjønn blant somaliere i Norge*, Oslo: Gyldendal Akademisk.

Fix, M. and Turner, M. A. (eds.) (1998), *A National Report Card on Discrimination in America: The Role of Testing*, Proceedings of the Urban Institute Conference (Washington, DC, March 1998) ERIC No. ED438348.

Friberg, J. H. and Elgvin, O. (2014), *Når aktivering blir ydmykelse: en studie av møtet mellom somaliske innvandrere og NAV*, Fafo report 2014:43, Oslo: Fafo.

Gilens, M. (1999), *Why Americans Hate Welfare: Race, Media, and the Politics of Antipoverty Policy*, Chicago: University of Chicago Press.

Gjersøe, H. M. (2016), 'Regulating inflow or outflow: a comparison of the work capability assessments in the UK and Norway', *Journal of Social Policy*, 45: 1, 141–58.

Goodin, R. E. (1998), 'Social welfare as a collective social responsibility', in D. Schmidtz and R. E. Goodin (eds.), *Social Welfare and Individual Responsibility*, Cambridge: Cambridge University Press, pp. 97–194.

Gubrium, E., Harsløf, I. and Lødemel, I. (2014), 'Norwegian activation reform on a wave of wider welfare state change: a critical assessment', in I. Lødemel and A. Moreira (eds.), *Activation or Workfare? Governance and Neo-Liberal Convergence*, Oxford: Oxford University Press, pp. 19–46.

Hagelund, A. and Terum, L. I. (2015), 'Fra saksbehandler til veileder. Om iverksetting av velferdspolitikk i arbeidslinjens tid', in A. H. Bay, A. Hagelund and A. Hatland (eds.), *For mange på trygd? Velferdspolitiske spenninger*, Oslo: Cappelen Damm Akademisk, pp. 131–152.

Halvorsen, K., Stjernø, S. and Øverbye, E. (2013), *Innføring i helse- og sosialpolitikk*, Oslo: Universitetsforlaget.

Heckman, J. J. and Siegelman, P. (1993), 'The Urban Institute Audit Studies: their methods and findings', in M. Fix and R. J. Struyk (eds.), *Clear and Convincing Evidence: Measurement of Discrimination in America*, http://works.bepress.com/peter_siegelman/33/ (accessed).

Jewell, C. J. (2007), *Agents of the Welfare State. How Caseworkers Respond to Need in the United States, Germany, and Sweden*. New York: Palgrave Macmillan

Jakobsson, N., Kotsadam, A., Syse, A. and Øien, H. (2015), 'Gender bias in public long-term care? A survey experiment among care managers', *Journal of Economic Behavior & Organization*, available online 14 September 2015.

Lima, I. A. and Naper, S. O. (2013), 'Kommer deltakerne i kvalifiseringsprogrammet i jobb?', *Arbeid og Velferd*, 2: 43–59.

Lipsky, M. (1980/2010) *Street-Level Bureaucracy. Dilemmas of the Individual in Public Services*. New York: Russell Sage Foundation

Midtbøen, A. H. and Rogstad, J. (2012), 'Discrimination. Methodological controversies and sociological perspectives on future research', *Nordic Journal of Migration Research*, 2: 3, 202–13.

Molander, A. and Torsvik, G. (2015), 'Getting people into work: what (if anything) can justify mandatory activation of welfare recipients?', *Journal of Applied Philosophy*, 32: 4, 373–92.

Neumark, D. (2012), 'Detecting discrimination in audit and correspondence studies', *Journal of Human Resources*, 47: 4, 1128–57.

https://doi.org/10.1017/S0047279417000113 Published online by Cambridge University Press

Norwegian Ministry of Labour and Inclusion (2006), *Arbeid, velferd og inkludering*, St. meld nr 9 (2006–2009) [Work, welfare, and inclusion, White Paper No. 9], Oslo: Norwegian Regjeringen.

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

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

Schram, S. F., Soss, J. and Fording, R. C. (eds.) (2003), *Race and the Politics of Welfare Reform*, Ann Arbor: University of Michigan Press.

Schram, S. F., Soss, J., Fording, R. C. and Houser, L. (2009), 'Deciding to discipline: race, choice, and punishment at the frontlines of welfare reform', *American Sociological Review*, 74: 3, 398–422.

Scruggs, L. (2008), 'Social rights, welfare generosity, and inequality', in P. Beramendi and C. J. Anderson (eds.), *Democracy, Inequality, and Representation: A Comparative Perspective*, New York: Russell Sage Foundation, pp. 62–92.

Scruggs, L. and Allan, J. (2006), 'Welfare-state decommodification in 18 OECD countries: a replication and revision', *Journal of European Social Policy*, 16: 1, 55–72.

Terum, L. I. and Hatland, A. (2014), 'Sysselsetting og trygd under arbeidslinja', *Søkelys på Arbeidslivet*, 31: 1–2, 3–22.

Van Der Waal, J., Achterberg, P., Houtman, D. and de Koster, W. (2010), "Some are more equal than others": economic egalitarianism and welfare chauvinism in the Netherlands', *Journal of European Social Policy*, 20: 4, 350–63.

Watkins-Hayes, C. (2013), 'Race, respect, and red tape: inside the black box of racially representative bureaucracies', in E. Brodkin and G. Marston (eds.), *Work and the Welfare State: Street-Level Organizations and Workfare Politics*, Washington, DC: Georgetown University Press, pp. 167–184.

Watts, B., Fitzpatrick, S., Bramley, G. and Watkins, D. (2014), 'Welfare sanctions and conditionality in the UK', in *Research Round-up*, York: Joseph Rowntree Foundation.

Zajonc, R. B. and Markus, H. (1982), 'Affective and cognitive factors in preferences', *Journal of Consumer Research*, 9: 2, 123–31.

https://doi.org/10.1017/S0047279417000113 Published online by Cambridge University Press

Appendix

TABLE A1. Testing for random assignment of the identity of the vignette client

	Dependent variable					
	Sanction		*Minority*		*Female*	
	Coef	*SE*	Coef	*SE*	Coef	*SE*
Age caseworker	0.00145	0.0021	0.0033	0.0021	0.00145	0.0021
Female caseworker	− 0.0679	0.04169	0.0445	0.0602	0.0295	0.0768
Social worker (education)	− 0.0334	0.04756	0.0024	0.0467	− 0.0470	0.0467
Social worker (work experience)	− 0.0780*	0.04623	0.0669	0.0462	0.0447	0.0511
Number of users	0.01396	0.0111	− 0.01745	0.0139	0.0010	0.0121
Fraction on welfare (muni)	0.04407***	0.01661	− 0.02864	0.0227	0.0378	0.0280
Number of caseworkers (office)	− 0.0015***	0.00054	− 0.00062	0.0005	− 0.0007	0.0006
Fraction of African immigrants (muni)	0.02550	0.02469	− 0.0249	0.0377	0.0202	0.0371
Constant	0.6911***	0.165	0.4954	0.0.1896	0.3087	0.1664
F-test for joint significance	7.62		1.47		1.15	
N	454		454		454	
R^2	0.08		0.02		0.02	

Note: Standard errors are clustered at office. muni = municipality.
*$p < 0.1$. **$p < 0.05$. ***$p < 0.01$.

https://doi.org/10.1017/S0047279417000113 Published online by Cambridge University Press