Roma and Bureaucrats: A Field Experiment in the Czech Republic*

Štěpán Mikula and Josef Montag†

This paper tests for discriminatory treatment of Roma minority by public officials in the Czech Republic. Our focus is on public servants at local job centers whose job is to advise unemployed individuals and process applications for unemployment benefits. Our experimental design facilitates testing for the presence of each of the two key (but intertwined) sources of discrimination: ethnic animus and socioeconomic status. We find substantial evidence for the presence of both. Because Roma tend to have lower socioeconomic status, the two sources of discrimination compound for them.

JEL codes: J15, D73, H55.

Keywords: Discrimination, Roma, socioeconomic status, public services, social security, field experiment.

1 Introduction

Roma people are one of the largest etnic minorities in the European Union. According to the European Commission (2010), a significant part of ten to 12 million European Roma live in extreme marginalization. Roma generally lag behind in employment, wages, poverty risk, and education (European Union Agency for Fundamental Rights 2012; O'Higgins and Brüggemann 2014). This likely results in stigma and discrimination. However, rigorous studies on discrimination of European Roma are lacking. One exception is a recent field experiment by Bartoš, Bauer,

* For helpful comments and criticisms we thank Libor Dušek and seminar participants at the Institute of Law and Economics at Hamburg University and participants at the 8th International Meeting in Law and Economics at Paris Nanterre University. Study pre-registration: Mikula, Stepan and Josef Montag. 2019. "Testing for ethnic discrimination within the Czech Social Security system." AEA RCT Registry. November 08. https://doi.org/10.1257/rct.4873-1.0. Replication package is available at https://dataverse.harvard.edu/privateurl.xhtml?token=3e558c9c-1484-4a21-b50e-6e046a7a6c0b.

† Mikula: Department of Economics, Masaryk University. Address: Lipová 41a, 60200 Brno, Czech Republic. Phone: +420 724 512 946. Email: stepan.mikula@econ.muni.cz. Web: www.sites.google.com/site/stepanmikula. Montag: Corresponding author. Department of Economics, Faculty of Law, Charles University, nám. Curieových 7, 11640 Prague, Czech Republic. Phone: +420 221 005 543. Email: josef.montag@gmail.com. Web: josefmontag.github.io.

Chytilová, and Matějka (2016) showing that Czech Roma are heavily discriminated against in the housing market and labor market.

In this paper, we report the results of a field experiment designed to study ethnic and socioeconomic discrimination of Roma in the public sector domain in the Czech Republic. With an estimated population of 200 to 300 thousand (two to three percent of the total), Roma are the largest ethnic minority in the Czech Republic (Czech Statistical Office 2014; Liégeois 2005; Liégeois, Gheorghe, Shuinear, and Phillips 1995). Discrimination of Czech Roma in the public sector domain aparently begins with their access to primary education due to ethnic segregation of schools and historically disproportionate numbers of Roma pupils placed in special schools designated for mentally challenged children, resulting in a *de facto* institutional segregation from an early age (Cviklova 2015; European Court for Human Rights 2007, Public Defender of Rights 2012). According to a research of the office of the Public Defender of Rights (2012) conducted in 67 special schools in 2011 and 2012, about one third of their pupils were Roma -- approximately ten times their share in the population.

In addition, only about one third of Czech Roma children aged four to six attend kindergarten, compared to over 70 percent non-Roma, and only about 30 percent of Roma aged 20 to 24 complete upper-secondary education, compared to over 80 percent of non-Roma population (European Union Agency for Fundamental Rights 2012). Roma thus typically exhibit lower educational attainment, which limits their labor market opportunities. This implies lower human capital accumulation and low socioeconomic status, which then likely perpetuates stigma and fuels discrimination.

Learning about the extent, patterns, and sources of discrimination of Roma is thus of high importance. We contribute to this goal by testing for discrimination of Roma in the realm of public services, namely their treatment when requesting unemployment benefits. We do so by sending three email queries to a sample of 457 public servants in job centers in the Czech Republic and randomly varying signals of putative ethnicity and socioeconomic status, and observing their responses. This allows us to tap into the potential mechanisms driving discrimination, which is our second main contribution.

We focus on two key mechanisms that may drive discrimination: ethnicity and socioeconomic status. We believe that this distinction is useful and important because ethnicity is fixed at birth whereas socioeconomic status is tightly linked to human

capital accumulation throughout an individual's life.[1] As a consequence, each of these two sources of discrimination likely requires different policy responses. For instance, policies addressing ethnic discrimination (xenophobia) should primarily focus on those who discriminate, whereas policies addressing socioeconomic discrimination should primarily focus on those who are being discriminated against.

However, because these two sources of dicrimination are intertwined (Roma tend to have lower socioeconomic status), telling them apart is a challenge. We therefore present a simple framework that clarifies how the signals of ethnicity and socioeconomic status mix and yields conditions under which the two sources of discrimination can be identified separately in our experiment. The implied testable hypotheses map neatly to our experimental design.

Better understanding what is behind a possible differential treatment is also important because the patterns of discrimination in the public domain are *a priori* ambiguous and various mechanisms may compound or cancel out.

Specifically, the standard (neoclassical) models of discrimination predict higher prevalence of discriminatory attitudes in the absence of market forces (Alchian and Kessel 1962; Becker 1957). To the extent negative attitudes towards Roma may be shared by public servants, one can therefore expect Roma facing similar, or even intensified, degree of discrimination in the public domain. This may be aggravated by socioeconomic discrimination, if public officials dislike dealing with low socioeconomic status individuals.

On the other hand, individuals who opt to become public servants may be intrinsically motivated to help the disadvantaged (Banuri and Keefer 2016; Dur and Zoutenbier 2014; Tonin and Vlassopoulos 2015). Such motivation in individuals who self-select into these professions could in principle ballance out the preferences for ethnic discrimination (if any), or possibly lead to positive discrimination.

Relatedly, in various countries there seems to be a common belief that minorities are often favoured and that various social security services, subsidy programmes, and

[1] In the public sector context, we view both sources of discrimination as essentially preferential, since there is no legitimate link between ethnicity or socioeconomic status and eligibility to access social security services. This is because the entitlement to receive unemployment benefits is orthogonal to ethnicity as well as socioeconomic status of an individual. This is different from a market setting, where socioeconomic background affects individual's (expected) productivity and statistical discrimination is an efficient response to imperfect information about individual's productivity.

NGO programs are more readily available to them than to the majority population, giving rise to *de facto* positive discrimination.[2] Such beliefs are often reflected in statements by politicians, on the political left and right, populist as well as moderate. However, if such beliefs are incorrect, they may play down the true discrimination and lead to misguided policy responses. Thus, apart from contributing to the general understanding of discrimination in the provision of public and legal services, our research provides a factual input into the public debate.

To summarize our results, we find strong evidence of both sources of discrimination, ethnic animus against Roma and negative socioeconomic discrimination. Queries sent by putatively Roma senders were seven percentage points less likely ($p < 0.01$) to be responded to, compared to queries sent by putatively Czech majority senders (response rates 0.60 and 0.53, respectively). Queries sent by putatively low-literacy senders were 25 percentage points less likely ($p < 0.01$) to be responded to, compared to queries sent by putatively high-literacy senders (response rates 0.64 and 0.48), respectively). These effects of discrimination are substantively important and appear significantly larger than effects found in field experiments on discrimination of minorities by public officials in other countries discussed in the next section.

Using a simple conceptual framework, we show that having found the socioeconomic discrimination to be negative, the differential treatment of Roma in our experiment cannot be explained by their socioeconomic status. In other words, ethnic discrimination is due to animus and not due to statistical discrimination. The implication is that because real-life Roma tend to have lower socioeconomic status than Czechs, the two sources of discrimination compound for them.

2 Related literature

Much of the existing research on discrimination, including correspondence studies, has focused on labor markets (see Baert 2018; Bertrand and Duflo 2017; and Neumark

[2] In the Czech Republic, for instance, 58 percent of individuals surveyed in April 2019 (n = 1052) state that Roma face better opportunities than non-Roma population when dealing with public administration, 29 percent stated that Roma face about equal opportunities, and 11 percent stated that they face worse opportunities (Public Opinion Research Centre 2019). In the same survey, 49 percent stated that Roma face better opportunities than non-Roma population to defend their interests in civil conflicts and disputes. A similar point is raised by Distelhorst and Hou (2014), who study discrimination by public officials in China, suggesting such beliefs are more general. By contrast, less than 15 percent believed that Roma have better opportunities in the job market, in education, or when obtaining qualifications.

2018 for recent comprehensive reviews). Discrimination in the public sector domain is significantly less researched than discrimination in private markets. In Political Science, there are several experimental studies of discrimination focusing on election officials and state politicians in the United States (see, e.g., Broockman 2013; Buttler and Broockman 2011; Hughes et al. 2019; White et al. 2015; for a meta-analysis see Costa 2017). These papers do find discrimination against minorities, particularly Latinos.

The closest study to ours is by Giulietti, Tonin, and Vlassopoulos (2019) who employed the correspondence approach to study discrimination against African Americans in local public services (education, libraries, sheriff offices, public administration, taxation, and job centers) across the US. They find that African Americans face a "penalty" of about four percentage points in the response rate (relative to 72 percent baseline response rate for whites) and suggest that the differential is likely to be driven by animus, rather than by "statistical discrimination" (due to African Americans' lower socioeconomic status). In the unemployment services domain, however, they find that African Americans are as likely as whites to receive a response to a query about unemployment benefits.

Several other studies report results relevant for this paper. Carnes and Holbein (2019) test for differential treatment of rich and poor constituents by state legislators, public school principals, and mayors -- finding null effects. In a similar vein, Einstein and Glick (2017) test for discriminatory behavior by street-level bureaucrats dealing with affordable housing programs in the US. They do not find any evidence of discrimination towards African Americans and only limited evidence of discrimination towards Hispanics (particularly in the tone of the responses). Outside the US, Distelhorst and Hou (2014) find evidence of discrimination by local officials against Muslims in China. Adman and Jansson (2017) and Ahmed and Hammarstedt (2020) tested for discrimination against muslims by municipalities in Sweden. Both found discrimination in the quality of response but not in response rates.

3 Institutional setting and research design

3.1 Job centers and unemployment specialists

Our correspondence experiment tests for potential discriminatory treatment of Roma minority by public servants employed at job centers in the Czech Republic. Job centers are local branches of the Labor Office of the Czech Republic (Úřad práce České Republiky) and are primarily tasked with providing information about job openings, registration of the unemployed, and provision of unemployment benefits. Job centers are a natural setting for our investigation and have the practical advantages that they are numerous (overall there are 198 local job centers), spread throughout the country, and each has a standardized website with contacts to its employees and their area of specialization.[3]

We focus on the public servants whose job is to assist unemployed individuals and process applications for unemployment benefits (henceforth "unemployment specialists"). Specifically, we test for differential treatment of Roma minority versus Czech majority at the stage of initial contact of a potential application for unemployment benefits.

The advantage of this setting is that the eligibility criteria for the unemployment benefits are objective and exactly specified by the law, and are of course orthogonal to ethnicity and socioeconomic status. Specifically, an individual is *entitled* to receive unemployment benefits if she becomes unemployed and has been contributing to the pension insurance (i.e. was employed or self-employed) for at least 12 months during the preceding two years. As a consequence, the unemployment specialists have zero discretion when administering them.[4]

[3] For instance, contacts to public servants employed at the job center in Ústí nad Labem, a town of 90,000, can be accessed at https://www.uradprace.cz/web/cz/kontakty-na-zamestnance-81 (accessed July 8, 2021).

[4] Formally, unemployment benefits in the Czech Republic are a form of government social insurance. Every employed or self-employed individual is obliged to contribute to the pension insurance, which is the determinant of eligibility for unemployment benefits. The specific amount of the unemployment benefits is then determined based on the unemployed person's previous wage with a cap at 58 percent of the (gross) average wage in the economy (in 2019, the maximum unemployment benefits were about 18,000 Czech crowns or 710 euro). Specifically, during the first two months of unemployment, the benefits are set at 65 percent of her previous wage, then 50 percent during the third and fourth month, and 45 percent for the remaining period. The maximum length of eligibility depends on the age of the unemployed person: five months for those aged up to 50, eight months for those between 50 and 55, and 11 months for those above 55 years of age.

However, the unemployment specialists do have discretion with respect to the degree of helpfulness and advice they offer to unemployed individuals, whether in direct contact or via electronic communication. We believe that in this limited-discretion setting, any discriminatory treatment cannot be explained as being policy-related and is likely to reflect broader discriminatory attitudes in the public sector domain.

3.2 Sample construction and interventions

Up to three unemployment specialists from each job center were included in our sample. At job centers with three or fewer unemployment specialists, we have included all. If there were more than three unemployment specialists at a job center, we have randomly sampled three of them. We have capped the number of unemployment specialists at three per job center in order to mitigate the burden that this study created for the job centers' employees and to limit the risk of raising suspicion when different officials receive messages with similar content. This resulted in a sample of 457 unemployment specialists (an average of 2.3 per job center). We have calculated that this sample size should facilitate identification of five-percent effect of discrimination with the power of 0.80 using the McNemar's test (see Appendix for power calculations).

In our interventions, fictitious applicants sent brief emails stating that they just lost their job, that they would like to get unemployment benefits, and asking what they should do. Public servants in the Czech Republic have a general duty to provide information. This is a simple query to which the recipient may just respond that the sender needs to show up at their local job center and register as unemployed and that he (all our senders were male) has the right to obtain help with their job search and is eligible for unemployment benefits, provided they worked at least 12 months during the last two years. More details may be provided, such as links to the local branch, excerpts from the related laws, or links to websites with information for unemployed.

Prior to the actual experiment, we have tested our queries by sending emails to several unemployment specialists and have received genuine responses from individual public servants. We have also discussed our queries with two senior public servants working under the Ministry of Labour and Social Affairs and concluded that while they are standard, there is no prescribed (mechanical) response to them and that public servants should respond to each such query individually. However, such queries are not

considered to be an official administrative communication, are not centrally registered, and responses are not subject deadlines. Hence, the unemployment specialists have a discretion over the response, without an obvious consequence to them in case of nonresponse.

Each email query carried two distinct signals, resulting in two-by-two variation: Ethnicity was signaled by the sender's name and socioeconomic status was signaled by literacy, i.e. by the formal and linguistic quality of the query. We believe literacy is a natural signal of basic educational attainment, and thus of socioeconomic status, that is relevant for low socioeconomic status minorities.[5]

In order to obtain names signaling ethnicity, we use a sample of names extracted from a convenience survey of poor families in Brno (the second largest city in the Czech Republic).[6] From these, we have selected ten putative Roma minority names and ten typical Czech names. We have tested the ethnicity signal associated with these names at the end of a lab experiment (unrelated to this project) in which we asked the participants (mostly students of Masaryk University in Brno) to assign one of four ethnicities (Czech, Slovak, Roma, or Hungarian) to each name. For the two names most strongly associated with Roma ethnicity (Mario Lakatoš and Jakub Gaži), 70 percent of our participants believed they belong to Roma (2-5 percent thought they are Czech and between 10 to 15 percent stated they are Slovak or Hungarian). For the two names most strongly associated with the Czech majority (Jakub Svoboda and Pavel Pospíšil), over 95 percent stated they belong to the Czech majority. We therefore use these four names to signal the putative ethnicity of our fictitious senders.

In order to obtain patterns of errors signaling low literacy, we have asked several clients of the Salvation Army in Brno to draft several email queries about unemployment benefits. Grammatically correct equivalents were drafted by ourselves.

[5] We have opted in favor of this type of intervention to signal socioeconomic status over somewhat more direct signals of socioeconomic status, such as academic titles (a standard practice in the Czech Republic to signal tertiary education in written communication) or occupations (as in Giulietti, Tonin, and Vlassopoulos, 2019). Because there are very few Czech Roma with tertiary education, it is not clear what signal would this send and whether such intervention would have much external validity. Similar issues arise with occupations. For instance, most if not all of the occupations used in Giulietti, Tonin, and Vlassopoulos (2019) require tertiary or at least secondary education. Occupations are also numerous, so that nontrivial discretion would be needed in choosing them, and they may carry distinct signals, perhaps interacting with ethnicity. All in all, we believe that the linguistic and formal quality of the message is a genuine signal of socioeconomic status that is most appropriate in the context of our study.

[6] Because of legal constraints, neither the Census nor any other administrative dataset containing ethnicity and names is available in the Czech Republic.

We note that all queries were polite, starting with a neutral greeting, followed by the query itself, and closed with thanking the recipient and the name of the sender. Table A1 in the Appendix shows the queries (in Czech) spell-checked with Microsoft Word.

3.3 Implementation details

Each unemployment specialist included in our study received three distinct email queries. In the first email, each official is randomly assigned one of four possible combinations of ethnicity and literacy of the first sender. The second email differed in the putative ethnicity of the sender, keeping the literacy signal constant. In the third email, we changed the literacy signal, and randomized the putative ethnicity.

In order to prevent a situation in which two unemployment specialists from the same job center receive identical queries, we have set up eight fictitious personas, four Roma and four Czech, by swapping the first names and surnames between the two Roma names and between the two Czech names. Additionally, we have used 12 distinct draft emails, with identical content but different wordings, six carrying a low-literacy signal and six with a high-literacy signal (see Table A1 in the Appendix). This allowed us to assign emails to the unemployment specialists so that out of the maximum of nine emails arriving at a job center: each email had unique wording and at most two emails came from the same sender. Of course, each of the three emails any individual official received had a distinct wording and distinct persona.

We sent our emails in batches twice a week (on Tuesday and Thursday mornings) over the course of six weeks of our implementation phase (November 11, 2019 until December 20, 2019), 12 batches in total.[7] In each batch, at most one unemployment specialist from each job center would receive our email query and the minimum span between any two emails sent to an unemployment specialist was ten days.

4 Hypotheses and conceptual framework

Our null hypotheses are of course the absence of ethnic and socioeconomic discrimination. Non-discrimination and equality before the law are core tenets human rights rooted in international law as well as in most countries' constitutions (see, e.g.,

[7] The seasonally adjusted unemployment rate in the Czech Republic in November and December 2019 was 2.2 and 2.0 percent, respectively.

Cassola et al. 2016).[8] Social security and unemployment benefits are public policies that should not be discriminatory (safe for the possibility that their aim is to help the discriminated and disadvantaged). It is thus legitimate and reasonable to expect that public servants will not discriminate against minorities or individuals with low socioeconomic status. In addition, the previous literature, reviewed in Section 2, found only limited evidence of discrimination in settings similar to ours.

Despite these baseline expectations, the absence of discrimination is not gurarranited *per se* and needs to be verified. We note that, even though absence of any discrimination can be thought of as the first best, not all patterns of discrimination are equally bad. Consider the possibility that public servants discriminate against Roma because of their ethnicity. Since Roma tend to have lower socioeconomic status (as discussed in the Introduction), then socioeconomic discrimination in favor of low socioeconomic status individuals is preferable as it may compensate some of the ethnic discrimination. However, discrimination against individuals with low socioeconomic status would mean that Roma get hit twice. This latter pattern of discrimination is therefore most problematic.

4.1 Naive tests of ethnic and socioeconomic discrimination

As we have indicated in the Introduction, the specific mechanisms behind discrimination, if any, are likely to be intertwined. This creates a challenge in studying the channels that may drive discrimination. Specifically, we are concerned that socioeconomic status is not directly observable in genuine email communication. In our experiment, it is signaled through literacy. However, because real-life Roma tend to have lower socioeconomic status than Czechs, socioeconomic status is correlated with ethnicity. Hence, the two signals may get mixed and separate identification of ethnic animus and socioeconomic discrimination may not be guaranteed in simple group-wise comparisons. Put differently, differential treatment of Roma may be potentially explained by preference for socioeconomic discrimination and Roma ethnicity serving as a signal of low socioeconomic (statistical discrimination).

[8] The first sentence of the Universal Declaration of Human Rights declares: "Whereas recognition of the inherent dignity and of the equal and inalienable rights of all members of the human family is the foundation of freedom, justice and peace in the world[.]" This is followed by Article 7: "All are equal before the law and are entitled without any discrimination to equal protection of the law. All are entitled to equal protection against any discrimination in violation of this Declaration [...]."

To fix ideas, let $e = C, R$ be the sender's putative ethnicity (Czech and Roma), and $l = L, H$ the signal about the sender's literacy (low and high). Our main outcome of interest is an indicator whether an applicant for unemployment benefits i received a response or not $Y = 0, 1$.

Consider two following (naive) conditions for absence of discrimination:

Hypothesis 1 (naive test of ethnic discrimination): $E(Y \mid e = R) = E(Y \mid e = C)$, i.e. Roma and Czech senders receive responses with the same probability.

Hypothesis 2 (naive test of socioeconomic discrimination): $E(Y \mid l = L) = E(Y \mid l = H)$, i.e. high and low-literacy types receive response with the same probability.

The problem with these two hypotheses arises because literacy is an imperfect signal of socioeconomic status and thus the signal of an individual's ethnicity is likely to be contaminated with some signal of his socioeconomic status. In addition, the signal about socioeconomic status contained in the literacy signal is likely to be different for each ethnicity. To be more precise, from the unemployment specialist's point of view, the lower average socioeconomic status of Roma makes a low literacy signal together with putative Roma identity consistent with the stereotype, whereas a high-literacy Roma would be considered an "overachiever." In the case of Czech majority, a high literacy would be the norm, whereas a low-literacy Czech would suggest an "underachiever." Testing Hypothesis 1 within each literacy level then does not immediately help, because socioeconomic status is not kept constant across the two ethnicity signals. Put differently, Hypothesis 1 woud constitute a clean test of ethnic discrimination only in the absence of socioeconomic discrimination, i.e. under Hypothesis 2 being true.

Nonetheless, we note that despite being naive from the point of view of identification of ethnic and socioeconomic discrimination, Hypotheses 1 and 2 are relevant from the policy viewpoint as their rejection suggests the presence some type of discrimination.

4.2 A framework for thinking about ethnicity, literacy, and socioeconomic status

In order to formulate valid hypotheses that test for ethnic and socioeconomic discrimination, the points just discussed need to be developed more precisely. Suppose

the socioeconomic status of an individual $s(a, l)$ is an increasing function of two variables, innate aptitude *a* and acquired human capital, which we proxy with literacy *l*. Suppose there are three levels of innate aptitude $a = 1, 2, 3$ distributed independently from ethnicity and let B_a measure the potential benefits from acquiring literacy, which are increasing in *a*. Unlike aptitude, literacy is principally a choice variable (determined by the parents and fundamentally by the costs and benefits). The costs of acquiring literacy are $c + e\,c_R$, where *c* is the baseline cost of literacy, $e = 0, 1\ (= C, R)$ is the indicator of ethnicity, and c_R is the additional cost of acquiring literacy for Roma and can be thought of as representing worsened access to quality education due to segregation and discrimination. It can also represent the costs of job market discrimination which reduces human capital returns for Roma and thus impedes investment in human capital.

Literacy is acquired if the benefits exceed the costs, i.e. if $B_a > c + e\,c_R$, and is therefore a function of aptitude and ethnicity, $l(a, e)$. As a consequence, socioeconomic status becomes $s(a, l(a, e))$.

Suppose $B_1 < c < B_2 < B_3$ and $B_1 < B_2 < c + c_R < B_3$, so that Czechs with $a = 2, 3$ invest in literacy, while only Roma with $a = 3$ invest. Then the following relations obtain

$$\begin{aligned} &s(3, l(3, R)) = s(3, l(3, C)) > s(2, l(2, C)) \\ &> s(2, l(2, R)) > s(1, l(1, R)) = s(1, l(1, C)). \end{aligned} \tag{1}$$

In our experiment, only literacy and ethnicity are observed by the unemployment specialists. As a result, middle and high aptitude Czechs cannot be distinguished (both will exhibit high literacy). Similarly, middle and low aptitude Roma cannot be distinguished (both will exhibit low literacy). At the same time, from the unemployment specialist's point of view, the lower average socioeconomic status of real life Roma, compared to the Czech majority, make a low-literacy signal together with Roma identity consistent with the norm, whereas a low-literacy Czech would suggest an "underachiever". Similarly, high-literacy Roma would be considered an "overachiever", whereas a high literacy Czech would fit the norm.

To summarize, the real life population of Roma exhibit *lower* average literacy than Czechs and thus lower socioeconomic status on average (even if their innate aptitude is the same). At the same time, Roma in our experiment have the *same* literacy as Czechs and so must have a higher aptitude, and thus higher socioeconomic status on average. The latter statement follows from (1), i.e.

$$s(3,R) + \frac{s(2,R)+s(1,R)}{2} > \frac{s(3,C)+s(2,C)}{2} + s(1,C),$$

writing $s(a, l(a,e))$ as $s(a,e)$ for short.

4.3 Tests of ethnic and socioeconomic discrimination

The key implication of the discussion in the previous section is that if we observe socioeconomic discrimination that is negative, ethnic discrimination (negative) of Roma must be driven by animus and cannot be explained by statistical discrimination of Roma because of their socioeconomic status. This is because, under negative socioeconomic discrimination, statistical discrimination should *favor* Roma (due to their higher socioeconomic status in our experiment).

Because of correlation between ethnicity and unobserved determinants of socioeconomic status in our experiment, socioeconomic discrimination can be primarily identified within ethnicities, particularly within the Czech ethnicity.[9] This implies the following hypothesis:

Hypothesis 3 (negative socioeconomic discrimination):

$$E(Y \mid e = C, l = L) < E(Y \mid e = C, l = H).$$

Under Hypothesis 3, we then have two tests of ethnic discrimination:

Hypothesis 4 (ethnic discrimination):

$$E(Y \mid e = R) < E(Y \mid e = C),$$

since from (1) directly follows that Roma in our experiment have higher *average* aptitude (are more often overachievers) than Czechs and thus higher socioeconomic

[9] We note that this test requires that the ethnicity signal is independent of literacy. This condition is likely to be satisfied for Czechs as Czech names are overwhelmingly associated with Czech ethnicity (as discussed in Section 3.2). However, Roma surnames and Roma ethnicity are not linked as tightly. Typical Roma names often originate from Slovak or Hungarian. (In the case of our two names, Lakatoš comes from Hungarian, whereas Gaži actually is from the Roma language itself.) It is therefore possible that emails from high-literacy putative Roma senders are less likely to be perceived as Roma than emails from low-literacy putative Roma senders. Depending on the sign of socioeconomic discrimination, potential presence of ethnic discrimination would then lead to underestimation of positive socioeconomic discrimination or overestimation of negative socioeconomic discrimination for Roma.

status on average. Hence, observing ethnic discrimination under Hypothesis 3 rules out the possibility that it is explained by socioeconomic discrimination.

Hypothesis 5 (strong ethnic discrimination):

$$E(Y \mid e = R, l = H) < E(Y \mid e = C, l = L),$$

since from (1) directly follows that all low-literacy Czechs (who are all underachievers) have lower socioeconomic status than all high-literacy Roma (all are overachievers).

Hypotheses 3 through 5 thus identify the worst case scenario: ethnic discrimination against Roma and socioeconomic discrimination against individuals with low socioeconomic status, noting that under Hypothesis 3, the magnitude of ethnic discrimination (if any) will be underestimated in the presence of statistical discrimination.

5 Data and descriptive statistics

As planned, we have sent out 1371 email requests to 457 unemployment specialists. We have received 905 responses altogether. The original recipient responded in 614 cases, 189 queries were responded to by another unemployment specialist after being forwarded. Some (102) of the initially received responses were automated (e.g. when the recipient was on vacation), and several emails received notifications about the message being forwarded with no followup. After removing these non-responses, we end up with 773 genuine responses to our queries (overall rate of response 56.4 percent).[10] About half of the responses arrived within two hours from the time the respective query was sent out, the last arrived after 21 days.

Our key outcome variable is an indicator taking on the value of one if the query was responded to, else it is coded as zero. Table 1 reports the descriptive statistics of the resulting dataset, broken down by the four treatment arms of our experiment.[11] The response rate varies widely across our four treatment arms, from 71 percent in case of queries from senders carrying signals of Czech ethnicity and high literacy to 46 percent in case of queries from putative low-literacy Roma senders. The last column of Table 1 reports the resulting p-value from F-test testing for systematic differences across the

[10] More than one genuine response arrived in several cases. We code these situations as $Y = 1$, i.e. the query was responded to.

[11] We note that we relied on pure randomization and did not impose uniformity of sample sizes across the four treatment arms. The exactly replicable randomization code is in the replication package.

four treatment arms. For the response outcome, the null (i.e. no systematic differences) is easily rejected.

[Table 1 about here.]

Across the four treatment arms, we do not observe systematic differences in the time to response, length of response, or its distinctiveness (i.e. whether the unemployment specialist copy-pasted the same response once or twice). High literacy senders are twice more likely to be greeted by name, but there is no difference between ethnicities in name greeting.

The last three variables capture the proportion of responses in which our query was marked as spam, the proportion of queries forwarded and responded by someone else, and the proportion of queries receiving an automated response. It is reassuring that there are no systematic differences in these variables across our treatment arms (all p-values are above 0.05). This suggests that the receiving email servers did not recognize any systematic differences between our queries and any differential treatment thus must be due to human behavior.

6 Statistical approach and results

Figure 1 plots the response rates across our four treatment arms, together with 83 percent confidence intervals allowing for visual evaluation of differences at five percent level of significance (Goldstein and Healy 1995). The figure shows two key patterns: (i) differential treatment of Roma, particularly within the high-literacy category and (ii) negative socioeconomic discrimination within both ethnicities. Although there appears to be a differential treatment of Roma within the low-literacy category, the difference is comparatively small and not statistically significant.

[Figure 1 about here.]

6.1 Statistical approach

For each subject (unemployment specialist) we have two observations with varying ethnicity (and constant literacy) and two observations with varying literacy (and constant ethnicity). McNemar's test (i.e. paired binomial test) is therefore the relevant

non-parametric test for our main hypotheses.[12] Because we test five core hypotheses, the reader may wish to set the threshold for rejecting the null at $\alpha_5 = 0.01$, which in the absence of any discrimination corresponds to Type I error probability of $\alpha = 0.05$.

Table 2 accompanies these nonparametric tests with estimates of random effects regressions testing the key relationships postulated in our main hypotheses (standard errors are clustered at the unemployment specialist level). We note that the identifying assumption in these regressions is that individual error terms are not correlated with the right-hand side variables. This requirement is satisfied by construction as the putative ethnicity and literacy signals are assigned randomly. We also checked this formally and the Hausman test never rejects consistency of random effects.

The McNemar tests exploit paired observations only (i.e. for each unemployment specialists we have the first two queries varying ethnicity and the last two queries varying literacy), which implies reductions in sample sizes available, especially for testing Hypotheses 3 through 5. The regressions use all available data and may thus provide more power. We treat both sets of results as complementary and interpret the findings from paired data as conservative.

6.2 Main results

Testing for Hypothesis 1, using the pairs of the first two emails sent to each recipient ($n = 2 \times 457$), in which we varied ethnicity, keeping the literacy signal constant, we obtain 44 pairs of queries in which only the Roma sender was responded to and 79 pairs in which only the Czech sender was responded to, yielding rejection of the null (McNemar's Test, $p = 0.0016$). This corresponds to the estimated coefficient on Roma indicator in specification (1) of Table 2, suggesting a seven percentage points reduction in response rate, compared to the 60 percent response rate to putatively Czech senders. Thus we obtain:

Result 1 We find evidence for differential treatment of queries from putatively Roma and putatively Czech senders.

[12] The McNemar test is a variant of the nonparametric sign test for matched pairs of binary observations (see Conover 1998, ch 3.4 and 3.5). Fagerland, Lydersen, and Laake (2013) recommend the mid-*p* approach to calculate the *p*-value of McNemar test as giving the best tradeoff between preservation of the significance level and power (in about 9,600 simulation scenarios, the mid-*p* McNemar test never violated the nominal level of significance while being almost as powerful as the asymptotic test).

[Table 2 about here.]

Testing for Hypothesis 2, using the pairs of the second and the third email sent to each recipient ($n = 2 \times 457$), in which we varied literacy, we obtain 33 pairs where only the low-literacy sender was responded to and 103 where only the high-literacy sender was responded to, yielding rejection of the null (McNemar's Test, $p = 0.93 \times 10^{-9}$). This corresponds to the estimated coefficient on low-literacy indicator in specification (2) of Table 2, suggesting 14 percentage points reduction in response rate, compared to the 64 percent response rate to putatively high literacy senders. Thus we have:

Result 2 We find evidence for differential treatment of queries from putatively high-literacy and putatively low-literacy senders.

Testing for Hypothesis 3, the differential treatment using the pairs of queries with the Czech ethnicity and varying literacy signals ($n = 2 \times 241$), we obtain 13 pairs where only the low-literacy sender received a response and 68 pairs where only the high-literacy senders was responded to, yielding rejection of the null hypothesis of non-negative socioeconomic discrimination (one-sided McNemar's Test, $p = 0.11 \times 10^{-9}$). This results corresponds to the estimated coefficients on low-literacy indicator in specification (3) of Table 2, suggesting 21.4 percentage points reduction in response rate to low-literacy Czech senders. These findings thus yield:

Result 3 We find evidence for negative socioeconomic discrimination within the sample of putatively Czech senders.

Because the Result 3 rules out positive socioeconomic discrimination, we can go ahead and test Hypotheses 4 and 5. Testing for Hypothesis 4, the negative discrimination of Roma, uses the same data as those used for testing Hypothesis 1 and yields rejection of the null hypothesis of non-negative ethnic discrimination (one-sided McNemar's Test, $p = 0.0008$). This again corresponds to the coefficient on Roma indicator in specification (1) of Table 2, suggesting 7 percentage points lower response rate to high-literacy Roma senders, compared with the 60 percent baseline response rate to Czech senders.

Result 4 We do find evidence for differential treatment of putatively Roma senders due to ethnic animus.

Finally, testing for Hypothesis 5, a differential treatment using the pairs of queries with the high-literacy Roma senders and low-literacy Czech senders ($n = 2 \times 242$), we obtain 55 pairs where only the Roma sender received a response and 25 pairs where only the Czech sender was responded to. We note that the result is in the opposite direction to that postulated by Hypothesis 5, hence the null hypothesis of strong non-negative discrimination cannot be rejected. This corresponds to the coefficient on Roma indicator in specification (7) of Table 2, suggesting a 9.1 percentage points higher response rate to high-literacy Roma senders, compared to 49.8 percent baseline response rate to low-literacy Czech senders. Thus we conclude:

Result 5 We do not find evidence for the hypothesis of strong ethnic discrimination, whereas low-literacy Czech senders would be preferred to high-literacy Roma senders. On the contrary, negative socioeconomic discrimination aparently dominates ethnic animus, resulting in preferential treatment of high-literacy Roma.

6.3 Complementary results

We note, that evaluating Hypothesis 3 within the pairs of queries with Roma ethnicity ($n = 2 \times 216$), we obtain 30 pairs where only the low-literacy sender received a response and 46 pairs where only the high-literacy sender was responded to, yielding rejection of the null hypothesis of non-negative socioeconomic discrimination (one-sided McNemar's Test, $p = 0.034$). This results corresponds to the estimated coefficient on low-literacy indicators in specification (4) of Table 2, suggesting 11.2 percentage points reduction in response rate to low-literacy Czech senders and low-literacy Roma senders, respectively. Hence, the important Result 3 of negative socioeconomic discrimination is corroborated in the Roma subsample as well.

When evaluating Hypothesis 4 within each literacy level, we find that within the pairs of queries with the high literacy and varying ethnicity signals ($n = 2 \times 238$), we obtain 22 pairs where only the Roma sender received a response and 45 pairs where only the Czech sender was responded to, yielding rejection of the null hypothesis of non-positive ethnic discrimination (one-sided McNemar's Test, $p = 0.0025$). This corresponds to the coefficient on Roma indicator in specification (5) of Table 2, suggesting 11.5 percentage points lower response rate to high-literacy Roma senders, compared with the 69.7 percent baseline response rate to high-literacy Czech senders.

Within the pairs of queries with low-literacy signal ($n = 2 \times 219$), we obtain 22 pairs where only the Roma sender received a response and 34 pairs where only the Czech sender was responded to. However, we cannot reject the null hypothesis of non-negative socioeconomic discrimination (one-sided McNemar's Test, $p = 0.056$). This corresponds to the coefficient on Roma indicator in specification (6) of Table 2, suggesting 4.4 percentage points lower response rate to low-literacy Roma senders, which is not statistically different from the 49.8 percent baseline response rate to low-literacy Czech senders.

6.4 Estimates using only between-subject variation

As a robustness check, in Table 3 we report results of OLS regressions analogous to those in Table 2, but limited to the subsample of first queries sent to each unemployment specialist. These regressions thus rely on the between-subject variation only, although the reduction in sample size to one third limits precision and power. We interpret the estimates in Table 3 so that the key patterns in our main results are corroborated and none of the five results is overturned.

[Table 3 about here.]

One marginal exception is that the negative estimate of differential treatment of low literacy Roma (statistically not significant) in specification (6) of Table 2 is replaced with a positive estimate (statistically not significant) in specification (6) of Table 3. However, the difference between the two estimates is not itself statistically significant (two-sample *z*-test, $p = 0.35$).

7 Concluding remarks

Our study, designed to test for the presence of ethnic animus and socioeconomic discrimination of Czech Roma minority in the public sector domain yield substantial evidence of both types of discrimination. We note that, on ballance, socioeconomic discrimination seems to be more significant driver of differential treatment. However, we suggest caution with this interpretation. We suggest that the low literacy signal in our experiment is highly salient to the recipients and so they respond more intensively

to it -- in a similar way a reviewer would act upon receiving a badly written paper. By contrast, Roma ethnicity signals are more subtle and somewhat noisy.

Nonetheless, taken at face value, these results suggest that public policy programmes aimed at improving Roma socioeconomic status can go a long way in reducing discrimination of Roma. In particular, more effort is needed to eliminate institutional discrimination in the access to education and compensate for Roma disadvantages in schooling (language defficiency, family background), all of which plausibly has detrimental effects on their socioeconomic status.

With regard to ethnic discrimination, we note that the findings may be partially attenuated by imperfect signalling of ethnicity as Roma often have "standard" Czech names and the names of the personas may not have been perceived as Roma by all recipients (recall about 70 percent of surveyed students identified these names as belonging to Roma). This suggests, we may be underestimating the magnitude of ethnic discrimination. We also note, that while our estimate of ethnic discrimination within the low-literacy senders are smaller and not statistically significant, the point estimate of -4.4 percent is significant substantively and comparable to the estimates in the previous literature from other countries.

Taken together, our findings demonstrate that Roma face non-trivial discrimination when dealing with the public sector. Because of their ethnicity and typically lower socioeconomic status, they end up discriminated against twice. In particular, our results are clearly inconsistent with the idea that Roma (and possibly other minorities) substantially benefit from preferential treatment by public sector officials.

Finally, we note that the standard criticism of audit/correspondence studies in labor market discrimination is that people are frequently employed via social connections and that these studies do test discrimination in average firms and not at the relevant margin (Heckman 1998). This criticism does not quite apply to our setting as unemployment benefits can only be obtained via standard bureaucratic application procedure and in this sense all public servants are "marginal" and discrimination by an average public servant is the relevant quantity. Thus, our study identifies discrimination at the relevant margins.

References

Adman, Per, and Hanna Jansson. 2017. A field experiment on ethnic discrimination among local Swedish public officials. *Local Government Studies* 43: 44-63.

Ahmed, Ali, and Mats Hammarstedt. 2020. Ethnic Discrimination in Contacts with Public Authorities: A Correspondence Test Among Swedish Municipalities. *Applied Economics Letters* 27: 1391-1394.

Baert, Stijn. 2018. Hiring discrimination: an overview of (almost) all correspondence experiments since 2005. In *Audit studies: Behind the scenes with theory, method, and nuance*. Cham, Germany: Springer 63-77.

Bancroft, Angus. 1999. 'Gypsies to the Camps!': Exclusion and Marginalisation of Roma in the Czech Republic. *Sociological Research Online* 4: 1-14.

Banuri, Sheheryar, and Philip Keefer. 2016. Pro-social motivation, effort and the call to public service. *European Economic Review* 83: 139-164.

Bartoš, Vojtěch, Bauer, Michal, Chytilová, Julie, and Filip Matějka. 2016. Attention discrimination: Theory and field experiments with monitoring information acquisition. *American Economic Review* 106: 1437-1475.

Bertrand, Marianne, and Esther Duflo. 2017. Field experiments on discrimination. In *Handbook of Economic Field Experiments*, vol. 1:, pp. 309-393. North-Holland.

Broockman, David E. 2013. Black politicians are more intrinsically motivated to advance blacks' interests: A field experiment manipulating political incentives. *American Journal of Political Science* 57: 521-536.

Butler, Daniel M., and David E. Broockman. 2011. Do politicians racially discriminate against constituents? A field experiment on state legislators. *American Journal of Political Science* 55: 463-477.

Cameron, A. Colin, and Douglas L. Miller. 2015. A practitioner's guide to cluster-robust inference. *Journal of Human Resources* 50: 317-372.

Carnes, Nicholas, and John Holbein. 2019. Do public officials exhibit social class biases when they handle casework? Evidence from multiple correspondence experiments. *PLOS One* 14: e0214244.

Conover, William Jay. 1998. *Practical nonparametric statistics*. New York: John Wiley & Sons.

Costa, Mia. How responsive are political elites? A meta-analysis of experiments on public officials. *Journal of Experimental Political Science* 4: 241-254.

Cviklova, Lucie. 2015. Direct and indirect racial discrimination of Roma people in Bulgaria, the Czech Republic and the Russian Federation. *Ethnic and Racial Studies* 38: 2140-2155.

Czech Statistical Office. 2014. Národnostní struktura obyvatel. [Ethnic structure of the population.] Online at https://www.czso.cz/documents/10180/20551765/170223-14.pdf (last accessed December 11, 2020).

Dur, Robert, and Robin Zoutenbier. 2015. Intrinsic motivations of public sector employees: Evidence for Germany. *German Economic Review* 16: 343-366.

European Court of Human Rights. 2007. Case of D. H. and Others v. Czech Republic. Online at http://hudoc.echr.coe.int/eng?i=001-83256 (last accessed December 11, 2020).

Einstein, Katherine Levine, and David M. Glick. 2017. Does race affect access to government services? An experiment exploring street‐level bureaucrats and access to public housing. *American Journal of Political Science* 61: 100-116.

European Union Agency for Fundamental Rights. 2012. The situation of Roma in 11 EU Member States: Survey results at a glance. Online at https://fra.europa.eu/sites/default/files/fra_uploads/2099-FRA-2012-Roma-at-a-glance_EN.pdf (last accessed December 11, 2020).

European Commission. 2020. The social and economic integration of the Roma in Europe. Online at https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52010DC0133 (last accessed December 11, 2020).

Fagerland, Morten W., Stian Lydersen, and Petter Laake. 2013. The McNemar test for binary matched-pairs data: mid-p and asymptotic are better than exact conditional. *BMC Medical Research Methodology* 13: 91.

Giulietti, Corrado, Mirco Tonin, and Michael Vlassopoulos. 2019. Racial discrimination in local public Services: A field experiment in the United States. *Journal of the European Economic Association* 17: 165-204.

Goldstein, Harvey, and Michael J. R. Healy. 1995. The graphical presentation of a collection of means. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 158: 175-177.

Greene, William H. 1981. On the asymptotic bias of the ordinary least squares estimator of the Tobit model. *Econometrica* 49: 505-513.

Heckman, James J. 1998. Detecting discrimination. *Journal of Economic Perspectives* 12: 101-116.

Hughes, D. Alex, Micah Gell-Redman, Charles Crabtree, Natarajan Krishnaswami, Diana Rodenberger, and Guillermo Monge. 2019. Persistent Bias Among Local Election Officials. *Journal of Experimental Political Science*: 1-9.

Liégeois, Jean-Pierre, and Nicolae Gheorghe. 1995. *Roma/gypsies: a European minority*. London: Minority Rights Group International.

Liégeois, Jean-Pierre. 2005. Les Roms au cœur de l'Europe. *Le Courrier des pays de l'Est* 6: 19-29.

Neumark, David. 2012. Detecting discrimination in audit and correspondence studies. *Journal of Human Resources* 47: 1128-1157.

Neumark, David. 2018. Experimental research on labor market discrimination. *Journal of Economic Literature* 56: 799-866.

O'Higgins, Niall, and Christian Brüggemann. 2014. The consequences of cumulative discrimination: How special schooling influences employment and wages of Roma in the Czech Republic. *European Educational Research Journal* 13: 282-294.

Public Defender of Rights. 2012. Výzkum veřejného ochránce práv k otázce etnického složení žáků bývalých zvláštních škol. [Research of the Public Defender of Rights with respect to the question of the ethnic composition of pupils of the former special schools.] Online at www.ochrance.cz/fileadmin/user_upload/DISKRIMINACE/Vyzkum/Vyzkum_skoly- zprava.pdf (last accessed December 11, 2020).

Public Opinion Research Centre. 2019. Romové a soužití s nimi očima české veřejnosti - duben 2019. [Roma and co-habitation with them through the lenses of Czech Society - April 2019.] Prague: Academy of Sciences of the Czech Republic. Online at https://cvvm.soc.cas.cz/media/com_form2content/documents/c2/a4924/f9/ov190517.pdf (last accessed December 11, 2020).

Suissa, Samy, and Jonathan J. Shuster. 1991. The 2 × 2 matched-pairs trial: exact unconditional design and analysis. *Biometrics* 47: 361-372.

Tonin, Mirco, and Michael Vlassopoulos. 2015. Are public sector workers different? Cross-European evidence from elderly workers and retirees. *IZA Journal of Labor Economics* 4: 11.

White, Ariel R., Noah L. Nathan, and Julie K. Faller. 2015. What do I need to vote? Bureaucratic discretion and discrimination by local election officials. *American Political Science Review* 109: 129-142.

Administrative information

Funding: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Institutional Review Board (ethics approval): No IRB approval was necessary for this project.

Declaration of interest: None.

Figures and Tables

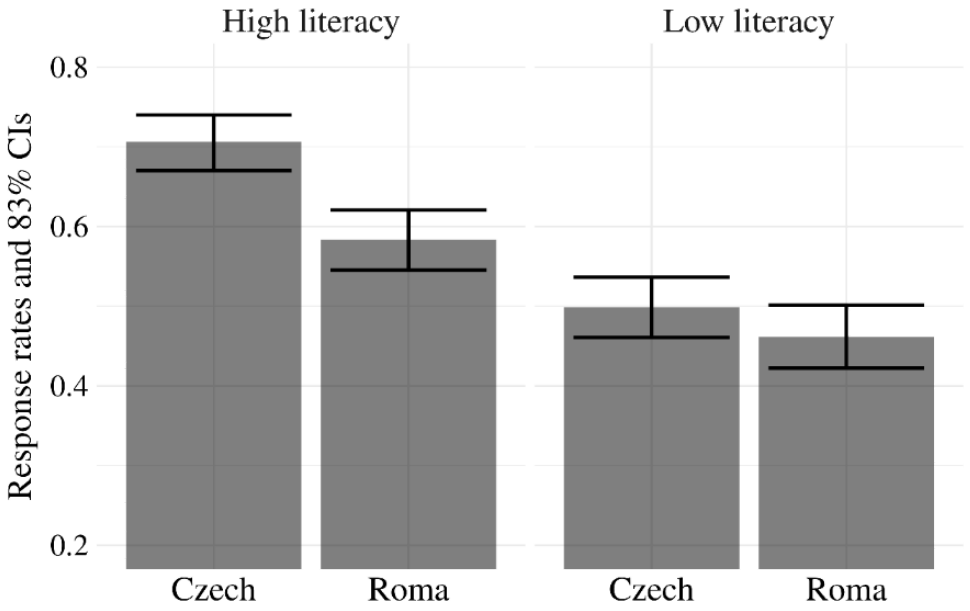

Figure 1: Email response rates by ethnicity signal and literacy signals. Confidence intervals are computed for the binomial proportion. Non-overlapping 83% CIs indicate statistically significant difference at $\alpha = 0.05$ (Goldstein and Healy 1995).

Table 1: Descriptive statistics by ethnicity and literacy signals (means and std. deviations)

		High literacy		Low literacy		F-test (p-value)
	All	Czech	Roma	Czech	Roma	
Response rate	0.56	0.71	0.58	0.50	0.46	$p < 0.01$
	(0.50)	(0.46)	(0.49)	(0.50)	(0.50)	
Responses:						
Time to response (hours)	13.79	14.42	11.18	18.18	11.22	0.23
	(38.63)	(35.43)	(39.31)	(49.48)	(26.32)	
Distinct response	1.04	1.05	1.04	1.04	1.03	0.55
	(0.20)	(0.24)	(0.19)	(0.20)	(0.18)	
Greeted by name	0.10	0.14	0.13	0.07	0.06	0.01
	(0.31)	(0.35)	(0.34)	(0.26)	(0.23)	
Length of resp. (words)	69.77	70.20	75.58	71.98	58.94	0.22
	(79.84)	(79.94)	(83.23)	(83.65)	(69.68)	
Query marked as spam	0.30	0.28	0.33	0.34	0.26	0.29
	(0.46)	(0.45)	(0.47)	(0.48)	(0.44)	
Query forwarded	0.20	0.17	0.22	0.22	0.21	0.45
	(0.40)	(0.38)	(0.41)	(0.42)	(0.41)	
Automatic response first	0.08	0.07	0.08	0.09	0.11	0.52
	(0.28)	(0.25)	(0.27)	(0.28)	(0.31)	
Observations	1371	347	348	351	325	

Note: F-tests test for systematic differences across the four treatment arms.

Table 2: Query responses, ethnicity, and literacy (random effects regressions)

Subsets:			Ethnicity =		Literacy =		Low-lit. Czechs, High-lit. Roma
			Czech	Roma	High	Low	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Roma (=1)	−0.070***				−0.115***	−0.044	0.091**
	(0.021)				(0.030)	(0.030)	(0.031)
Low literacy (=1)		−0.161***	−0.214***	−0.112**			
		(0.023)	(0.031)	(0.034)			
Intercept	0.598***	0.643***	0.707***	0.581***	0.697***	0.498***	0.498***
	(0.020)	(0.020)	(0.024)	(0.026)	(0.024)	(0.026)	(0.026)
Observations	1,371	1,371	698	673	695	676	699
Adjusted R^2	0.006	0.038	0.067	0.023	0.018	0.004	0.021

Note: Specifications (1) and (2) contain the full data, specifications (3) and (4) are run on samples restricted to putative Czech and Roma ethnicities, respectively, specifications (5) and (6) are run on samples restricted to putative high- and low-literacy senders, respectively, and specifications (7) is estimated on sample restricted to putative low-literacy Czechs and high-literacy Roma. Standard errors clustered at the recipient level are in parentheses: $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Table 3: Query responses, ethnicity, and literacy, first emails subsample (between-subject, OLS)

Subsets:			Ethnicity =		Literacy =		Low-lit. Czechs, High-lit. Roma
			Czech	Roma	High	Low	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Roma (=1)	−0.100*				−0.232***	0.059	0.094
	(0.046)				(0.059)	(0.068)	(0.068)
Low literacy (=1)		−0.183***	−0.326***	−0.035			
		(0.045)	(0.061)	(0.066)			
Intercept	0.648***	0.685***	0.797***	0.565***	0.797***	0.471***	0.471***
	(0.032)	(0.030)	(0.037)	(0.047)	(0.037)	(0.049)	(0.049)
Observations	457	457	227	230	238	219	219
Adjusted R^2	0.008	0.032	0.111	-0.003	0.058	-0.001	0.004

Note: Specifications (1) and (2) contain the full data, specifications (3) and (4) are run on samples restricted to putative Czech and Roma ethnicities, respectively, specifications (5) and (6) are run on samples restricted to putative high- and low-literacy senders, respectively, and specifications (7) is estimated on sample restricted to putative low-literacy Czechs and high-literacy Roma. Robust standard errors in are parentheses: $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Appendix

Power calculations

For each subject (unemployment specialist) we have two observations with varying ethnicity (and constant literacy) and two observations with varying literacy (and constant ethnicity). McNemar's test (paired binomial test) is therefore the relevant non-parametric test for our data.

Denote p_{11}, p_{10}, p_{10}, p_{00} the sampled probabilities that a subject responds to both Czech and Roma, only Czech, only Roma, and neither of the two ethnicities, respectively. We have $p_{11} + p_{10} + p_{10} + p_{00} = 1$.

Let $p_C = p_{11} + p_{10}$ and $p_R = p_{11} + p_{01}$ be the overall response probabilities of receiving a response for the putative Czech and Roma senders, respectively. Finally, let $\delta = p_C - p_R$ be the response differential between the two ethnicities (the discrimination effect), which after substituting yields $\delta = p_{10} - p_{01}$.

Let n be the number of subjects (paired observations), then McNemar's test statistic is

$$S = \frac{(p_{10}n - p_{01}n)^2}{(p_{10}n + p_{01}n)} = \frac{\delta^2 n}{p_{10} + p_{01}},$$

which under $H_0: \delta = 0$ asymptotically follows a chi-squared distribution with one degree of freedom.

Fagerland, Lydersen, and Laake (2013) investigate Type I errors frequencies and power of alternative methods to compute the *p*-values. Under a wide range of parameter scenarios, the Exact unconditional McNemar test and McNemar mid-*p* test, Type I errors frequency never exceeds five percent and are almost as powerful as the asymptotic McNemar test. We therefore base our power calculations on the Exact unconditional McNemar test (Suissa and Shuster 1991).

In our notation, the power of the test depends on three parameters, n, δ, p_{01}. In our case $n = 457$ and we consider $\delta = 0.05$ a substantively significant discrimination

coefficient (Giulietti, Tonin, and Vlassopoulos, 2019, found four percentage points differential between whites and blacks).

In order to gauge p_{01}, the baseline response rate in Giulietti et al. (2019) was 70 percent, setting our expectation for $p_C = 0.7$ and implying a constraint $p_{01} = 0.3 - p_{00}$. One now has to make a judgment about the actual size of p_{01}. Responses to only Roma senders may result because of two main reasons, positive discrimination of Roma by some subjects, and the fact that some subjects may respond to emails randomly. We believe that positive discrimination of Roma should not be too frequent, but random responses can. If we set $p_{01} = 0.05$ (randomness in the response occurs with the same frequency as discrimination), δ implies $p_{10} = 0.1$.

The power for one-sided Exact unconditional McNemar test with the rejection criterion $\alpha = 0.05$ under the stated parameters is 0.85. If we set $p_{01} = 0.06$, the corresponding power will be 0.80.

Spell checked email queries

Table A1: Spellchecked email queries, by the signal of socioeconomic status

Grammatically correct queries (High socioeconomic status signal)	Grammatically incorrect queries (Low socioeconomic status signal)
Dobrý den, byl jsem propuštěn ze svého zaměstnání. Můžete mi poradit, jak mám dále postupovat, abych získal státní podporu? Děkuji za odpověď,	Dobry den, Sef mi vyhodil a nemam prci. Podadite mne co potrebuju udelat aby ste mi daly potporu? Moc diky
Dobrý den, přišel jsem o práci a potřeboval bych vědět, co mám dělat, abych od vás dostal podporu pro nezaměstnané. Děkuji za odpověď.	Dobry den Sem odedneżka bes práce a že můžu dostat nějakou podporu co mám pro to udělat??? Diki
Dobrý den, chtěl bych se zeptat, co musím udělat, abych dostal dávky pro nezaměstnané, když jsem v práci dostal výpověď. Mockrát děkuji,	Dobry den, Minuli tyden sem skoncil v praci a potreboval bych potpru. Diky za radu.
Dobrý den, jsem bez práce. Chci si zažádat o dávky pro nezaměstnané. Poradíte mi, prosím, co musím udělat? Děkuji.	Dobry den Uz nemam praci a chtel byjsem podtpou. Jak to muzu zaridit? Dekuji,
Dobrý den, Můžete mi, prosím, poradit, když si potřebuji požádat o podporu pro nezaměstnané? Co pro to mám udělat? Děkuji za Váš čas,	Dobrý den ! Předvčíre mi vyhodili z práce. Potřeboval bych potporu. Co mam delaz? Porate!!! Dekuji.
Krásný den, píši Vám, protože jsem dostal výpověď, a tak bych si chtěl požádat o podporu v nezaměstnanosti. Můžete mi, prosím, říct, co a jak? Velmi děkuji,	Dobrý den propustiili mně dnes s mého zamjestnání a kamarádi mi poradily abych se vás zeptal na podporu. Poraďte jak na to. Díky moc