

DISCUSSION PAPER SERIES

IZA DP No. 16218

**Roma and Bureaucrats:
A Field Experiment on Ethnic and
Socioeconomic Discrimination**

Štěpán Mikula
Josef Montag

JUNE 2023

DISCUSSION PAPER SERIES

IZA DP No. 16218

Roma and Bureaucrats: A Field Experiment on Ethnic and Socioeconomic Discrimination

Štěpán Mikula

Masaryk University, IZA and SYRI National Institute

Josef Montag

Charles University

JUNE 2023

Any opinions expressed in this paper are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but IZA takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The IZA Institute of Labor Economics is an independent economic research institute that conducts research in labor economics and offers evidence-based policy advice on labor market issues. Supported by the Deutsche Post Foundation, IZA runs the world's largest network of economists, whose research aims to provide answers to the global labor market challenges of our time. Our key objective is to build bridges between academic research, policymakers and society.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Roma and Bureaucrats: A Field Experiment on Ethnic and Socioeconomic Discrimination*

This paper tests for discriminatory treatment of the Roma minority by public officials in the Czech Republic at the stage of initial contact preceding a potential application for unemployment benefit. Our correspondence experiment facilitates testing for the presence of each of two intertwined drivers of discrimination: ethnic animus and socioeconomic status prejudice. We find substantial evidence for the presence of discrimination based on both of these sources. Since Roma tend to have lower socioeconomic status, the two sources of discrimination compound for them.

JEL Classification: J15, D73, H55

Keywords: discrimination, Roma, ethnicity, socioeconomic status, public services, social security, field experiment

Corresponding author:

Josef Montag
Department of Economics
Faculty of Law
Charles University
nám. Curieových 7
11640 Prague
Czech Republic
E-mail: josef.montag@gmail.com

* We thank Libor Dušek, Katerina Linos, seminar participants at the Institute of Law and Economics at Hamburg University, and participants at conferences of the American Law and Economics Association, the European Association of Law and Economics, the German Law and Economics Association, the International Meeting in Law and Economics at Paris Nanterre University, and workshop participants from the Ministry of Labor and Social Affairs and the Ministry of Regional Development for their helpful comments on our paper. We are also grateful to Michal Kozel, who assisted us with this research. This study was pre-registered at the AEA RCT Registry: 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>. A replication package with data and code is available at Harvard Dataverse at <https://doi.org/10.7910/DVN/SOJYQM>. Mikula acknowledges the financial support from the grant NPO "Systemic Risk Institute" (LX22NPO5101).

1 Introduction

Roma people are one of the largest ethnic minorities in the European Union. According to the European Commission (2020), a significant proportion of Europe's 10 to 12 million Roma live in extreme marginalization. About 80 percent of Roma in nine EU countries with the largest Roma populations live in poverty (European Union Agency for Fundamental Rights 2016). Roma not only lag behind in education, employment, and wages but also in access to health insurance, nutrition, and even tap water (European Union Agency for Fundamental Rights 2012, 2016; O'Higgins and Brüggemann 2014).

All this likely feeds into stigma and discrimination. Indeed, according to the Eurobarometer, 20 percent of respondents stated that they would feel uncomfortable having a colleague who is Roma (European Commission 2015). In the Czech Republic, the corresponding figure is 52 percent. Based on this survey evidence, Roma are the most stigmatized minority group in the EU.

Contrasting with these facts is the lack of rigorous research to date on discrimination against European Roma. One exception is a recent field experiment by Bartoš, Bauer, Chytilová, and Matějka (2016) showing that Czech Roma are heavily discriminated against in the housing market and labor market. Most recently, Linos, Jakli, and Carlson (2021) report that donations in a text-to-give campaign experiment in Greece decreased by one half when a Roma child was referenced.

In this paper, we report the results of a pre-registered correspondence experiment designed to study discrimination against Roma in the public sector in the Czech Republic on ethnic and socioeconomic grounds. Specifically, we test for differential treatment of Roma minority versus Czech majority at the stage of initial contact preceding a potential application for unemployment benefit. The outcome corresponds

to whether a job center official responded to an informal email inquiry by an unemployed person who is entitled to receive an unemployment benefit. We ran this study in November and December 2019.

With an estimated population of 260 thousand (2.4 percent of the country's total population), Roma people constitute the largest ethnic minority in the Czech Republic (Office of the Government 2019). Public sector discrimination against Czech Roma apparently begins with their access to primary education: ethnic segregation of schools and the historical practice of placing disproportionate numbers of Roma pupils in special schools designated for mentally challenged children have resulted in *de facto* institutional segregation from an early age (Cviklova 2015; European Court for Human Rights 2007, Public Defender of Rights 2012). According to research conducted by the office of the Public Defender of Rights (2012) in 67 special schools, 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 of non-Roma children, and only about 30 percent of Roma aged 20 to 24 have completed upper-secondary education, compared to over 80 percent of the 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 discriminatory treatment of Roma in the realm of public services, specifically when requesting unemployment benefit. We do so by sending email queries to a sample of 457 public servants in job

centers in the Czech Republic containing 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.

Distinguishing between ethnic animus and discrimination on socioeconomic grounds as potential drivers of discrimination 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. As a consequence, each of these two sources of discrimination likely requires a different policy response. For instance, policies addressing discrimination on ethnic grounds (xenophobia) should primarily focus on those who discriminate, whereas policies addressing discrimination on socioeconomic grounds should also focus on those who are discriminated against.¹

In the public sector context, we view both sources of discrimination as taste-based, since there is no legitimate link between ethnicity or socioeconomic status and eligibility to access social security services. This is because an individual's entitlement to unemployment benefit is orthogonal to both their ethnicity and their socioeconomic status (it depends only on their unemployment status and preceding professional activity).

However, because the signals of ethnicity (senders' names) and of socioeconomic status (signaled via the senders' literacy level, i.e. the formal and linguistic quality of the query, which is an innovation this paper brings to the field experimental literature) in our interventions are imperfect, discriminatory preferences along one dimension may also trigger "statistical discrimination" along the other dimension. Specifically,

¹ We note that the term socioeconomic status is not an exactly delineated concept. In this paper we conceive it in a relatively narrow sense, as reflecting individuals' education and cognitive skills (but not necessarily income level or professional success). Arguably, these are the dimensions of socioeconomic status that are relevant from the public servants' viewpoint as they directly affect the expected costs (time and effort) of dealing with an individual.

differential treatment of Roma may not only take place in reaction to direct signals of ethnicity and low economic status themselves, but may also be indirectly driven by ethnicity taken as a proxy for low socioeconomic status, or a combination of all of these.²

Given this possible intertwinement, telling these sources of discrimination apart is a challenge. We therefore present a simple framework that clarifies how 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 onto our experimental design.

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

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

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

² Bartoš et al. (2016) document that Roma-sounding names are associated with substantially lowered expectations of high school or university education (see their table S1 in the Appendix at http://vojtechbartos.net/wp-content/uploads/Papers/20151105_Attention_discrimination_SOM.pdf).

self-select into these professions could in principle balance out any preferences for ethnic discrimination, or possibly even lead to positive discrimination.

Indeed, there seems to be a common belief in various countries that minorities are often favored and that various social security services, subsidy programs, and NGO programs are more readily available to them than to the majority population, giving rise to *de facto* positive discrimination.³ Such beliefs are often reflected in statements by politicians, both left- and right-wing, populist as well as moderate. However, if such beliefs are incorrect, they may play down the true level of 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 much-needed factual input to this public debate.

To summarize our results, we find strong evidence of both sources of discrimination: ethnic animus against Roma and negative discrimination on socioeconomic grounds. Queries sent by putatively Roma senders were seven percentage points less likely to be responded to than queries sent by putatively Czech majority senders (response rates 0.53 and 0.60, respectively, $p < 0.01$). Queries sent by putatively low socioeconomic status senders were 16 percentage points less likely to be responded to than queries sent by putatively higher socioeconomic status senders (response rates 0.48 and 0.64, respectively, $p < 0.01$). These effects are substantively

³ In the Czech Republic, for instance, 58 percent of individuals surveyed in April 2019 ($n = 1052$) stated that the Roma people have better opportunities than non-Roma when dealing with public administration, 29 percent stated that Roma have 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 have better opportunities than non-Roma to defend their interests in civil conflicts and disputes. A similar point was raised by Distelhorst and Hou (2014), who studied discrimination by public officials in China; this suggests such beliefs are more general. By contrast, less than 15 percent of the surveyed Czech respondents believed that Roma have better opportunities in the job market, in education, or when obtaining qualifications.

important and appear significantly larger than the effects found in field experiments on public officials' discrimination against minorities in other countries, which we discuss in the next section.

Using our conceptual framework, we show that despite our finding that there is negative discrimination on socioeconomic grounds, the differential treatment of Roma in our experiment cannot be explained by their perceived socioeconomic status. This is because in order to achieve the same level of literacy as an ethnic Czech person, a Roma individual needs more inputs facilitating human capital accumulation. Hence when we fix the level of literacy for both ethnicities, Roma signal higher socioeconomic status than Czechs in our experiment. Put differently, high-literacy Roma are more likely to be perceived as “overachievers” than high-literacy Czechs, whereas low-literacy Czechs are more likely to be perceived as “underachievers” than low-literacy Roma. As a consequence, we interpret the differential treatment of Roma in our study as being due to ethnic animus rather than 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 (for recent comprehensive reviews see Baert 2018; Bertrand and Duflo 2017; and Neumark 2018). Discrimination in the public sector has been the subject of significantly less research than discrimination in private markets. In political science, there have been several experimental studies focused on discrimination by 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 have found evidence of discrimination against minorities, particularly Latinos. A recent paper by Crawford and Ramli (forthcoming) finds discrimination against both Jews and Muslims by elected local government representatives affiliated with the two major political parties in the UK.

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 (schools, libraries, sheriffs' offices, public administration, taxation, and job centers) across the US. They found that African Americans face a penalty of about four percentage points in response rate (relative to a 72 percent baseline response rate for whites) and suggest that the differential is likely driven by animus rather than by statistical discrimination (driven by African Americans' lower socioeconomic status). In the unemployment services domain, however, they find that African Americans are equally as likely as whites to receive responses to queries about unemployment benefits.

Several other studies have reported results relevant for this paper. Carnes and Holbein (2019) tested for differential treatment of rich and poor constituents by US state legislators, public school principals, and mayors, and found null effects. In a similar vein, Einstein and Glick (2017) tested for discriminatory behavior by street-level bureaucrats dealing with affordable housing programs in the US. They did not find any evidence of discrimination towards African Americans and only limited evidence of discrimination towards Hispanics (primarily in the tone of the responses). Outside the US, Distelhorst and Hou (2014) found evidence of discrimination against Muslims by local officials 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 responses but not in response rates.

3 Institutional setting and research design

3.1 Job centers, unemployment specialists, and sample construction

Our correspondence experiment tests for potential discriminatory treatment of the 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), tasked with the administration and provision of state social services. A key part of their agenda is registering individuals as unemployed, processing applications for unemployment benefit, and providing information about job vacancies. Job centers have the practical advantages that they are numerous, spread throughout the country, and each has a standardized website with contact details for its employees and their job descriptions.⁴ We focus on the public servants whose job is to assist unemployed individuals and process applications for unemployment benefit (henceforth “unemployment specialists”).

In this setting, we test for differential treatment of Roma minority versus Czech majority service users at the stage of initial contact preceding a potential application for unemployment benefit. Public servants in the Czech Republic have a general duty to provide information. However, they do have significant discretion with respect to the degree of helpfulness and advice they offer to unemployed individuals, whether in direct contact or via electronic communication.

⁴ For instance, contact details for 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).

An important feature of our setting is that the formal eligibility criteria for unemployment benefit 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 benefit if he or she becomes unemployed, has paid pension insurance contributions (i.e. been employed or self-employed) for at least 12 months during the preceding two years, and registers as unemployed. Unemployment specialists have zero discretion when registering such individuals as unemployed and processing their applications for unemployment benefit.⁵

We believe that in this limited-discretion environment, any discriminatory treatment cannot be rationalized as being policy-related and is likely to reflect broader attitudes towards minorities in the public sector.

Our sample included up to three unemployment specialists from each job center, identified using the published job descriptions. At job centers with three or fewer unemployment specialists, we included all of them. Where there were more than three unemployment specialists at a single job center, we randomly sampled three of them. We 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 from the same job center received messages with similar content. This resulted in a sample of 457 unemployment

⁵ Formally, unemployment benefit in the Czech Republic is a form of state social insurance. Every employed or self-employed individual is obliged to contribute to pension insurance, which is the determinant of eligibility for unemployment benefit. The specific amount of unemployment benefit 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 benefit per month was about 18,000 Czech crowns or 710 euro). Specifically, during the first two months of unemployment, the benefit amounts to 65 percent of the individual's previous wage, during the third and fourth month 50 percent, and for the remaining period 45 percent. The maximum duration of eligibility for unemployment benefit depends on the age of the unemployed person: five months for those aged up to 50, eight months for those aged between 50 and 55, and 11 months for those above 55 years of age.

specialists from 198 job centers (an average of 2.3 per job center).⁶ We calculated that this sample size should facilitate the identification of a five-percent discrimination effect with the power of 0.80 using McNemar's test (see Appendix for power calculations).

3.2 Interventions

In our interventions, fictitious applicants sent brief emails stating that they had lost their job and would like to receive unemployment benefit, and asking what they should do. This is a simple query to which the recipient may simply respond by informing the sender that he (all our fictitious senders were male) should come in to his local job center to register as unemployed and that he is eligible for unemployment benefit provided he has worked for at least 12 months during the last two years.⁷ More details may be provided, such as the possibility to receive help with their job search, links to the local job center, excerpts from the related laws, or links to websites with information for unemployed people.

Prior to the actual experiment, we tested our queries by sending emails to several unemployment specialists and receiving genuine responses from individual public servants. We also discussed our queries with two senior public servants working for the Ministry of Labor and Social Affairs and concluded that while they are standard types of enquiry, there is no prescribed (mechanical) response to them and public servants would be expected to respond to each such query individually. However, such queries

⁶ Our data covers only local job centers dealing with the unemployment agenda, identified using the employees' job descriptions. We were able to identify unemployment specialists at 198 of the country's 244 local job centers. Some local job centers specialize in other state social services, while some unemployment specialists or job centers may have not been identified by us because the published job descriptions are not standardized.

⁷ While it is not mandatory to register as unemployed, doing so is a necessary condition in order to be eligible for unemployment benefit. It is also necessary in order for the state to cover the mandatory monthly health insurance contributions, which the unemployed individual would otherwise have to cover themselves (about 70 euro per month as of 2019).

are not considered to be an official administrative communication, are not centrally registered, and responses are not subject to any deadlines. Hence, the unemployment specialists have discretion over the timing and content of their responses, with no obvious consequence to them should they fail to respond.

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 that literacy is a natural signal of basic educational attainment, and thus of socioeconomic status, which is relevant for low socioeconomic status minorities. As noted above (see footnote 1), it is also a relevant signal for the public servants because education and cognitive skills arguably affect the expected costs (time and effort) of dealing with an individual.⁸

We note that, in our setting there is a limited scope for variation in the unemployment specialists' perceptions of other potential determinants of socioeconomic status. Notably, our queries imply that the senders have been continuously in work and that they do not have a previous experience as recipients of unemployment benefit. This should also minimize the possibility that they may be some sort of "abusers" of the social security services. Second, they all possess at most secondary education: this is

⁸ A recent study of recruiters by Sterkens et al. (2021) reported that individuals perceive grammatical errors to be a signal of low mental abilities, conscientiousness, and interpersonal skills.

We opted in favor of this type of intervention to signal socioeconomic status over somewhat more broader 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 and is appropriate in the context of our study.

implicitly signaled by the absence of any academic titles in our queries. As a norm, academic titles are used in formal communication in the Czech Republic. Last, the fact that they lost a job and that this creates a problem for them (otherwise they would not be contacting a job center to apply for benefit), suggests they are likely to be “marginal” workers in manual or low-level administrative jobs (i.e. unlikely to be secured company insiders or highly skilled professionals with immediate alternative job offers). We note that our experiment took place against the backdrop of a very low unemployment rate at the time, making these implications firmer.⁹

In order to obtain names signaling ethnicity, we used a sample of names extracted from a convenience survey of poor families in Brno (the second largest city in the Czech Republic).¹⁰ From these, we selected ten putative Roma minority names and ten typical Czech names. We 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 belonged to Roma (2-5 percent thought they were Czech and between 10 and 15 percent stated they were 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 used these four names to signal the putative ethnicity of our fictitious senders.

⁹ The seasonally adjusted unemployment rate in the Czech Republic in November and December 2019 was 2.2 and 2.0 percent, respectively.

¹⁰ Because of legal constraints, neither the Census nor any other administrative dataset containing ethnicity and names is available in the Czech Republic.

In order to obtain patterns of writing errors signaling low literacy, we asked several clients of the Salvation Army in Brno (i.e., people of low socioeconomic status, often homeless) to draft email queries about unemployment benefit. We drafted grammatically correct equivalents ourselves. All the queries were polite; each opened with a neutral greeting followed by the query itself, and closed by thanking the recipient and signing off with the sender's name. Table A1 in the Appendix shows the queries (in Czech) spell-checked using Microsoft Word. The low-literacy queries contain grammatical errors, typos, and untidy formatting.

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

3.3 Implementation details

In order to prevent any situation arising in which two unemployment specialists at the same job center might receive identical queries, we created eight fictitious personas – four Roma and four Czech – by combining each first name with both surnames for each ethnicity. Additionally, we used 12 distinct message texts, with identical content but different wording, six carrying a low-literacy signal and six with a high-literacy signal (see Table A1 in the Appendix). This enabled us to assign emails to the unemployment specialists in such a way that each of the maximum of nine emails arriving at a given job center was uniquely worded and at most two of those nine emails came from the same sender. We ensured that all messages received by a particular official were distinct in both wording and sender persona.

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

4 Hypotheses and conceptual framework

Our null hypotheses are the absence of any discrimination on ethnic and socioeconomic grounds. Social security and unemployment benefit are public policies that should not be discriminatory (except in those instances in which they are targeted 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, reported only limited evidence of discrimination in settings similar to ours.

Despite these baseline expectations, the absence of discrimination is not guaranteed *per se* and needs to be verified. Moreover, although the absence of any discrimination is desirable, 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, discrimination in favor of low socioeconomic status individuals would be preferable as it might compensate for some of the negative discrimination driven by ethnic animus. However, discrimination against individuals with low socioeconomic status would mean that Roma got hit twice. This latter pattern of discrimination is therefore most problematic.

4.1 Naive tests of ethnic and socioeconomic discrimination

As we indicated in the Introduction, the specific mechanisms behind discrimination, if any, are likely to be intertwined. This creates a challenge when studying the forces that may drive discrimination. Specifically, because real-life Roma tend to have lower socioeconomic status than Czechs, socioeconomic status is correlated with ethnicity. At the same time, socioeconomic status is not directly observable in genuine email communication; in our experiment, it is signaled through literacy. As a result, signals of ethnicity and socioeconomic status 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 the presence of negative discrimination on socioeconomic grounds (statistical discrimination), while Roma ethnicity serves as a signal of low socioeconomic status.

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

Consider the two following (naive) conditions for the presence of discrimination:

Hypothesis 1 (naive hypothesis of discrimination based on ethnicity):

$E(Y | e = R) \neq E(Y | e = C)$, i.e. Roma and Czech senders are treated differently and are not equally likely to receive responses.

Hypothesis 2 (naive hypothesis of discrimination based on socioeconomic status):

$E(Y | l = L) \neq E(Y | l = H)$, i.e. high- and low-literacy senders are not equally likely to receive responses.

The problem with these two hypotheses is that socioeconomic status is not constant in comparisons under Hypothesis 1. Testing Hypothesis 1 within each literacy level does not help either, because socioeconomic status is not kept constant across the two ethnicity signals even within literacy levels. In this particular case, the lower average socioeconomic status of Roma people makes a low literacy signal together with a putative Roma identity consistent with the stereotype, from the unemployment specialist's perspective, whereas a high-literacy Roma sender would be considered an "overachiever." In the case of a Czech majority sender, high literacy would be the norm, whereas low literacy would suggest an "underachiever." In other words, Hypothesis 1 captures purely ethnic discrimination only in the absence of socioeconomic discrimination.

However, we note that despite being naive from the point of view of identification of ethnic and socioeconomic discrimination, Hypotheses 1 and 2 remain relevant from the policy viewpoint as the potential rejection of the corresponding null hypotheses suggests the presence of 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 an individual's socioeconomic status $s(a, l)$ 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 increase in a . Unlike aptitude, literacy is principally a choice variable (initially determined by the parents and later by the individual, fundamentally depending on the

costs and benefits). The costs of acquiring a high level of 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. c_R can be thought of as representing worsened access to quality education due to segregation and discrimination, but also family environment less favorable to academic performance (e.g. due to parents' low educational attainment). It can also represent the costs of job market discrimination which reduces human capital returns for Roma and thus impedes incentives to invest 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

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

In our experiment, the unemployment specialists only observe ethnicity and literacy level; socioeconomic status is inferred. As a result, they cannot distinguish between middle and high aptitude Czechs (both will exhibit high literacy). Similarly, they cannot distinguish between middle and low aptitude Roma (both will exhibit low literacy). As a result, since the Roma enquirers in our experiment have the *same* literacy as Czechs they are perceived as having higher aptitude, implying 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. This inference holds also within each literacy level.

4.3 Tests of ethnic and socioeconomic discrimination

The key implication of this is that if we observe non-positive discrimination driven by socioeconomic status, then any discrimination against Roma must be driven by ethnic animus, i.e. it cannot be explained as statistical discrimination against Roma because of their socioeconomic status. This is because, if there is negative discrimination on socioeconomic grounds, statistical discrimination should *favor* Roma in our experiment (due to their relatively higher socioeconomic status in our experiment).

Because of the correlation between ethnicity and unobserved determinants of socioeconomic status in our experiment, we are primarily able to identify discrimination based on socioeconomic status within each ethnicity, particularly within the Czech ethnicity (because of the clear ethnic signal).¹¹ This implies the following hypothesis:

Hypothesis 3 (negative socioeconomic discrimination):

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

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

¹¹ We note that this test requires that the ethnicity signal is independent of literacy. This condition is likely to be satisfied for Czechs, since the Czech names we use are overwhelmingly associated with Czech ethnicity (as discussed in Section 3.2). However, the Roma surnames are not linked as tightly with Roma ethnicity. Typical Roma names often originate from Slovak or Hungarian. (In the case of our two names, Lakatoš comes from Hungarian, whereas Gaži is from the Roma language.) It is therefore possible that high-literacy Roma-named senders are less likely to be perceived as Roma than low-literacy Roma-named senders. Depending on the sign of socioeconomic discrimination, the potential presence of ethnic discrimination would then lead to an underestimation of positive socioeconomic discrimination or an overestimation of negative socioeconomic discrimination towards Roma.

Hypothesis 4 (ethnic discrimination):

$$E(Y | e = R) < E(Y | e = C),$$

since it follows directly from (1) that Roma senders in our experiment have higher *average* aptitude (are more often overachievers) than Czechs and thus higher socioeconomic status on average. Hence, observing ethnic discrimination under Hypothesis 3 rules out the possibility that it is explained by socioeconomic discrimination operating statistically.

Hypothesis 5 (strong ethnic discrimination):

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

since it follows directly from (1) directly that *all* low-literacy Czech senders (who are all underachievers) have lower socioeconomic status than all high-literacy Roma senders (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. We note, that under Hypothesis 3, the magnitude of ethnic discrimination (if any) will be underestimated in the presence of statistical discrimination on socioeconomic grounds operating via ethnicity.

5 Data and descriptive statistics

As planned, we sent out 1371 email requests to 457 unemployment specialists. We received 905 responses altogether. The original recipient responded in 614 cases, while 189 queries were responded to by a different unemployment specialist after being forwarded by the initial recipient. Some of the responses received were automated (e.g. when the recipient was out of the office at the time), and several queries received responses notifying the sender that their message had been forwarded, but with no

followup. After removing these non-responses, we obtained 773 genuine responses to our queries (an overall response rate of 56.4 percent).¹²

About half of the responses arrived within two hours of the time the respective query was sent out, the latest response arrived 21 days after the query was sent. A typical response contained one or two sentences (70 words on average, including greeting and signature), most commonly advising the sender to register as unemployed at his local job center and confirming that he is entitled to benefit provided he worked for 12 months during the previous two years. Further details were often included, such as details of the required forms and documents (e.g. national ID card).

Our key outcome variable is an indicator that takes the value of one if the query was responded to, and zero otherwise. Table 1 reports the descriptive statistics of the resulting dataset, broken down by the four treatment arms of our experiment.¹³ The response rate varies widely across our four treatment arms, from 71 percent for queries from senders carrying signals of Czech ethnicity and high literacy to 46 percent for queries from senders carrying signals of low-literacy and Roma ethnicity. The last column of Table 1 reports the resulting p -value from an F -test for systematic differences across the four treatment arms. For the response outcome, the null hypothesis is easily rejected.

We note that the vast majority of the unemployment specialists (95 percent) contacted in our experiment are females. Also, most of the job centers contacted (84 percent) are located in municipalities with populations of below 50,000. Our four treatment arms are well balanced across the gender and geographic dimensions.

¹² Several enquiries received more than one genuine response. We code these situations as $Y = 1$, i.e. the query was responded to.

¹³ Our experiment relied on pure randomization and we did not impose uniformity of sample size across the four treatment arms. The exactly replicable randomization code is in the replication package.

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

	All	High literacy		Low literacy		<i>F</i> -test (<i>p</i> -value)
		Czech	Roma	Czech	Roma	
Response (=1)	0.56 (0.50)	0.71 (0.46)	0.58 (0.49)	0.50 (0.50)	0.46 (0.50)	$p < 0.01$
Unemployment specialists (recipients):						
Female (=1)	0.95 (0.22)	0.95 (0.21)	0.95 (0.23)	0.95 (0.21)	0.94 (0.23)	0.90
Located in Prague (=1)	0.06 (0.24)	0.05 (0.22)	0.07 (0.25)	0.05 (0.21)	0.07 (0.26)	0.54
Located in a city above 100k pop. (=1)	0.05 (0.22)	0.05 (0.22)	0.05 (0.22)	0.06 (0.23)	0.05 (0.22)	0.97
Located in a city above 50k pop. (=1)	0.05 (0.22)	0.05 (0.22)	0.05 (0.21)	0.06 (0.23)	0.06 (0.23)	0.92
Responses:						
Time to response (hours)	13.79 (38.63)	14.42 (35.43)	11.18 (39.31)	18.18 (49.48)	11.22 (26.32)	0.23
Distinct response (=1)	0.96 (0.20)	0.95 (0.22)	0.96 (0.19)	0.96 (0.20)	0.97 (0.18)	0.66
Greeted by name (=1)	0.10 (0.31)	0.14 (0.35)	0.13 (0.34)	0.07 (0.26)	0.06 (0.23)	0.01
Length of resp. (words)	69.77 (79.84)	70.20 (79.94)	75.58 (83.23)	71.98 (83.65)	58.94 (69.68)	0.22
Query marked as spam (=1)	0.30 (0.46)	0.28 (0.45)	0.33 (0.47)	0.34 (0.48)	0.26 (0.44)	0.29
Query forwarded (=1)	0.20 (0.40)	0.17 (0.38)	0.22 (0.41)	0.22 (0.42)	0.21 (0.41)	0.45
Automatic response first (=1)	0.08 (0.28)	0.07 (0.25)	0.08 (0.27)	0.09 (0.28)	0.11 (0.31)	0.52
Observations	1371	347	348	351	325	

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

We do not observe any systematic differences in time to response, length of response, or response distinctiveness (i.e. whether the response was unique or identical to another) across our four treatment arms. High-literacy senders are twice as likely to be greeted by name than low-literacy senders, but there is no difference between the ethnicities in the named greeting rate.

The last three variables reported in Table 1 capture the proportion of responses in which our query was marked as spam, the proportion of queries forwarded and responded to by someone else, and the proportion of queries that received automated responses.¹⁴ 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

¹⁴ The spam filter at job centers' email server flags potential spam in the email subject, making it observable for us.

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 also appears to be differential treatment of Roma within the low-literacy category, this difference is comparatively small and not statistically significant.

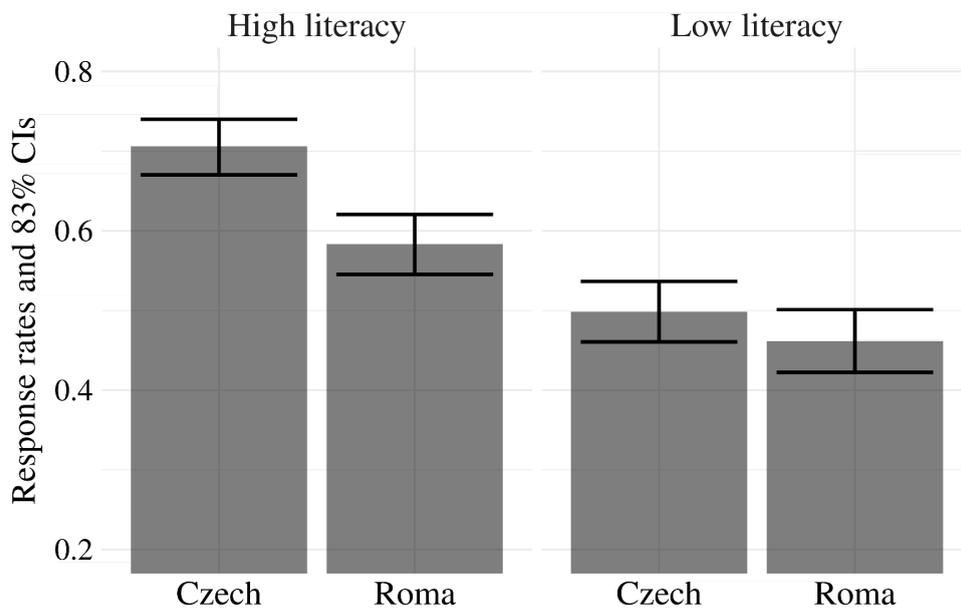


Figure 1: Email response rates by ethnicity signal and literacy signals. Confidence intervals are computed for the binomial proportion. Non-overlapping 83% CIs indicate statistically significant difference at $\alpha = 0.05$ (Goldstein and Healy 1995).

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 (in accordance with the pre-registration).¹⁵ 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 a 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, since the putative ethnicity and literacy signals are assigned randomly.¹⁶

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

¹⁵ 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.

¹⁶ We also checked this formally and the Hausman test never rejected the consistency of our random effects regressions (results available in the replication package).

6.2 Main results

Testing for Hypothesis 1, using pairs of the first two emails sent to each recipient ($n = 2 \times 457$), in which we varied ethnicity while 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 a rejection of the null (McNemar's Test, $p = 0.0016$). This corresponds to the estimated coefficient on the 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 according to the sender's putative ethnicity.

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

Subsets:			Ethnicity =		Literacy =		Low-lit. Czechs, High-lit. Roma
	(1)	(2)	Czech (3)	Roma (4)	High (5)	Low (6)	
Roma (=1)	-0.070*** (0.021)				-0.115*** (0.030)	-0.044 (0.030)	0.091** (0.031)
Low literacy (=1)		-0.161*** (0.023)	-0.214*** (0.031)	-0.112** (0.034)			
Intercept	0.598*** (0.020)	0.643*** (0.020)	0.707*** (0.024)	0.581*** (0.026)	0.697*** (0.024)	0.498*** (0.026)	0.498*** (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: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Testing for Hypothesis 2, using pairs of the second and third emails sent to each recipient ($n = 2 \times 457$), in which we varied literacy while keeping the ethnicity signal constant, we obtain 33 pairs of queries in which only the low-literacy sender was responded to and 103 pairs in which only the high-literacy sender was responded to,

yielding a rejection of the null (McNemar's Test, $p = 0.93 \times 10^{-9}$). This corresponds to the estimated coefficient on the low-literacy indicator in specification (2) of Table 2, suggesting a 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 according to the sender's putative level of literacy.

Testing for Hypothesis 3, using the pairs of queries with Czech ethnicity and varying literacy signals ($n = 2 \times 241$), we obtain 13 pairs in which only the low-literacy sender received a response and 68 pairs in which only the high-literacy sender was responded to, yielding a rejection of the null hypothesis of non-negative socioeconomic discrimination (one-sided McNemar's Test, $p = 0.11 \times 10^{-9}$). This result corresponds to the estimated coefficient on the low-literacy indicator in specification (3) of Table 2, suggesting a 21.4 percentage points reduction in response rate to low-literacy Czech senders compared with the 70.7 percent baseline response rate to high-literacy Czech senders. These findings thus yield:

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

Because Result 3 rules out positive socioeconomic discrimination, we can go ahead and test Hypotheses 4 and 5. To test for Hypothesis 4, negative discrimination against Roma, we use the same data as we used to test Hypothesis 1 and this yields a 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 the Roma indicator in specification (1) of Table 2, and suggests a seven percentage points lower

response rate to high-literacy Roma senders compared with the 60 percent baseline response rate to Czech senders.

Result 4 We find evidence of discrimination of putatively Roma senders due to ethnic animus.

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

Result 5 We do not find evidence for strong ethnic discrimination, such as would result in more favourable treatment of low-literacy Czech senders than high-literacy Roma senders. On the contrary, negative socioeconomic discrimination apparently dominates over ethnic animus, resulting in the preferential treatment of high-literacy Roma over low-literacy Czechs.

6.3 Complementary results

We note that by evaluating Hypothesis 3 using pairs of queries with Roma ethnicity signals ($n = 2 \times 216$), we obtain 30 pairs in which only the low-literacy sender received a response and 46 pairs in which only the high-literacy sender received a response, yielding a rejection of the null hypothesis of non-negative socioeconomic discrimination (one-sided McNemar's Test, $p = 0.034$). This result corresponds to the

estimated coefficient on low-literacy indicators in specification (4) of Table 2, and suggests an 11.2 percentage points reduction in response rate to low-literacy Roma senders compared to high-literacy Roma senders. Hence, our key Result 3, confirming negative socioeconomic discrimination, is corroborated in the Roma subsample.

By evaluating Hypothesis 4 within each literacy level, we find that within the pairs of queries with high literacy and varying ethnicity signals ($n = 2 \times 238$), we obtain 22 pairs in which only the Roma sender received a response and 45 pairs in which only the Czech sender was responded to, yielding a rejection of the null hypothesis of non-positive ethnic discrimination (one-sided McNemar's Test, $p = 0.0025$). This corresponds to the coefficient on the Roma indicator in specification (5) of Table 2, and suggests an 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 signals ($n = 2 \times 219$), we obtain 22 pairs in which only the Roma sender received a response and 34 pairs in which 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 the Roma indicator in specification (6) of Table 2, and suggests a 4.4 percentage points lower response rate to low-literacy Roma senders compared with the 49.8 percent baseline response rate to low-literacy Czech senders, which is not a statistically significant difference.

6.4 Estimates using only between-subject variation

As a robustness check, in Table 3 we report the 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 between-subject variation only

(and the associated reduction in sample size to one third limits the precision and power of these estimates). 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.

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, $p = 0.35$).

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

Subsets:			Ethnicity =		Literacy =		Low-lit. Czechs, High-lit. Roma
	(1)	(2)	Czech	Roma	High	Low	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Roma (=1)	-0.070*** (0.021)				-0.232*** (0.059)	0.059 (0.068)	0.094 (0.068)
Low literacy (=1)		-0.183*** (0.045)	-0.326*** (0.061)	-0.035 (0.066)			
Intercept	0.598*** (0.020)	0.685*** (0.030)	0.797*** (0.037)	0.565*** (0.047)	0.797*** (0.037)	0.471*** (0.049)	0.471*** (0.049)
Observations	1,371	457	227	230	238	219	219
Adjusted R ²	0.006	0.032	0.111	-0.003	0.058	-0.001	0.004

Note: OLS regressions, analogous to those reported in Table 2, using only the first email queries sent to each unemployment specialist. Robust standard errors in are parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

6.5 Exploring heterogeneity in discrimination

We also explored potential sources of heterogeneity in the observed discriminatory behavior, focusing on unemployment specialists' gender, urban location, and exposure to individuals from socially excluded areas. A summary of the results is reported in Table 4.

With respect to gender, 95 percent of unemployment specialists in our sample are females, as reported in Table 1. Dropping males from the sample and re-estimating the regressions reported in Table 2 yields marginally higher coefficient estimates (in the

absolute sense), as reported in Panel A of Table 4. However these differences in coefficient estimates are not statistically significant (using a two-sample z-test). We interpret this result as indicating that our findings generalize to female unemployment specialists (which is the relevant population) and that the presence of male unemployment specialists does not appear to alter the intensity or the pattern of discrimination against Roma.

Table 4: Exploring heterogeneity in discrimination (random effects regressions)

Subsets:			Ethnicity =		Literacy =		Low-lit. Czechs, High-lit. Roma
	(1)	(2)	Czech (3)	Roma (4)	High (5)	Low (6)	
A: Dropping male unemployment specialists							
Roma (=1)	-0.074*** (0.021)				-0.125*** (0.031)	-0.043 (0.031)	0.091** (0.032)
Low literacy (=1)		-0.168*** (0.023)	-0.223*** (0.031)	-0.116** (0.035)			
Observations	1,302	1,302	666	636	660	642	664
B: Dropping unemployment specialists located in Prague							
Roma (=1)	-0.081*** (0.021)				-0.128*** (0.031)	-0.052 (0.031)	0.076* (0.032)
Low literacy (=1)		-0.159*** (0.023)	-0.212*** (0.031)	-0.104** (0.035)			
Observations	1,290	1,290	663	627	654	636	659
C: Dropping unemployment specialists in cities with population above 100k							
Roma (=1)	-0.084*** (0.022)				-0.128*** (0.032)	-0.052 (0.032)	0.072* (0.034)
Low literacy (=1)		-0.160*** (0.024)	-0.215*** (0.032)	-0.100** (0.037)			
Observations	1,218	1,218	626	592	619	599	621
D: Controlling for exposure to socially excluded locations							
Roma (=1)	-0.109*** (0.025)				-0.193*** (0.036)	-0.053 (0.038)	0.048 (0.040)
Low literacy (=1)		-0.156*** (0.028)	-0.239*** (0.040)	-0.079 (0.040)			
Excluded share (%)	-0.020 (0.011)	-0.018 (0.011)	-0.029* (0.014)	-0.010 (0.015)	-0.030* (0.013)	-0.008 (0.016)	-0.009 (0.016)
Roma × Excluded share	0.016 (0.009)				0.024 (0.016)	0.010 (0.018)	0.0002 (0.015)
Low literacy × Excluded share		0.014 (0.010)	0.020 (0.021)	0.014 (0.016)			
Observations	1,116	1,116	574	542	564	552	572

Note: Random effects regressions analogous to those reported in Table 2. For description of the data on socially excluded locations see Section 6.5. Standard errors clustered at the recipient level are in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Similarly, our results do not change appreciably if we exclude unemployment specialists located in Prague (the capital city, with a population of 1.3m and nine job centers, i.e. six percent of the sample) or all cities with populations above 100k (five additional cities, eight job centers: five percent of the sample), as reported in Panels B and C, respectively. We obtain similar results if we drop towns with populations above 50k (eight cities, eight job centers: five percent of the sample, results available in the replication package). We interpret these results as indicating that discriminatory behavior does not vary significantly between job centers located in more densely populated urban areas and those in the rest of the country.

Lastly, we explored whether discrimination may be explained by the job center staff's factual experience with the disadvantaged Roma minority. We use data on socially excluded locations collected for the Ministry of Labour and Social Affairs in 2014 (Čada et al. 2015).¹⁷ For each of the 205 administrative districts in the Czech Republic, this data provides the number of identified excluded locations and estimates their respective populations. While socially excluded locations are not exclusively inhabited by the Roma population, the associated report states that Roma represent the majority in about 75 percent of them (Čada et al. 2015, p. 47). The report also states that excluded locations tend to be perceived as "Roma" locations irrespective of whether Roma constitute the majority there or not (p. 19). Hence, the data should pick up some of the local variation in presence of the most disadvantaged Roma minority.

Using census data on the population sizes of the administrative districts, we compute the shares of each district's population living in socially excluded locations and merge this data with the data from our field experiment. There is not a perfect overlap

¹⁷ See online (in Czech) at <https://www.esfcr.cz/mapa-svl-2015/www/index2679.html> (last accessed December 30, 2021).

between job centers and administrative districts (some districts contain more than one job center), and, as a result, only 146 out of 197 job centers can be fully matched with administrative districts. The resulting dataset thus covers 79 percent of unemployment specialists from our main dataset (notably excluding the two large cities of Prague and Ostrava).¹⁸

We then interact the percentage share of the population living in socially excluded locations (min: 0, median: 0.4, mean: 1.3, max: 15.9) with the indicators for Roma ethnicity and low literacy and re-estimate the regressions reported in Table 2 with these interactions and the main effect of the share of excluded population. Panel D of Table 4 reports the results.

The estimated effects of the exposure to the socially excluded population and the interactions tend not to be statistically significant. Importantly, controlling for exposure to the socially excluded population does not “explain away” our main results, as the coefficients on the Roma and low-literacy indicators remain qualitatively and quantitatively similar (the differences between the corresponding estimates are not statistically significant, using a two-sample *z*-test).

7 Concluding remarks

Our study, designed to test for the presence of discrimination driven by ethnic animus or socioeconomic prejudice against the Czech Roma minority in the public sector, yields substantial evidence of both types of discrimination. We note that, on balance, discrimination on socioeconomic grounds seems to be the more significant of these two drivers. However, we suggest caution over this interpretation. The

¹⁸ Reestimating the regressions reported in Table 2 using this subset of the data produces very similar results (the differences in the coefficient estimates are not statistically significant, using a two-sample *z*-test).

low-literacy signal in our experiment is highly salient to the recipients and so they may simply respond more intensively to it -- much as a reviewer would act upon receiving a badly written paper. By contrast, the Roma ethnicity signals in our experiment are more subtle and somewhat noisy.

Hence, our findings regarding ethnic discrimination, may be partially attenuated by imperfect signalling of ethnicity, as Roma often have “standard” Czech names and the names used in our experiment may not have been perceived as Roma by all recipients (recall that only about 70 percent of surveyed students identified the names we used as belonging to the Roma minority). This suggests that we may be underestimating the magnitude of ethnic discrimination. We also note that, while our estimate of ethnic discrimination against low-literacy senders is smaller and not statistically significant, the point estimate of -4.4 percent is substantively significant 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 public services. They are affected by both ethnic and socioeconomic prejudices, and hence end up being discriminated against twice over. Notably, our results are clearly inconsistent with the idea that members of the Roma minority (and possibly other minorities) benefit substantially from preferential treatment by public sector officials.¹⁹

Our results suggest that public policy programs aimed at improving the socioeconomic status of Roma people could also help to reduce discrimination against them. In particular, more effort is needed to eliminate institutional discrimination in

¹⁹ Relatedly, Linos et al. (forthcoming) found that disadvantaged groups tend to avoid phone- and Internet-based communication channels for bureaucratic inquiries. Increasing use of these channels by the public as well as private sector thus creates new administrative burdens for them. Discrimination likely aggravates this burden further.

access to education and to compensate for the Roma population's disadvantages in schooling (language deficiency, family background), all of which plausibly has a detrimental effect on their socioeconomic status.

Finally, we note that a 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, i.e. not at the relevant margin (Heckman 1998). This criticism does not quite apply to our setting, as unemployment benefit can only be obtained via a single, standardised 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 does identify 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. 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.
- Čada, Karel, Daniela Büchlerová, Zuzana Korecká, and Tomáš Samec. 2015. Analysis of socially excluded localities in the Czech Republic. European social fund in the

- Czech Republic. Online at <https://www.esfcr.cz/documents/21802/791224/Analysis+of+Socially+Excluded+Localities+in+the+CZ.pdf> (last accessed December 30, 2021).
- Crawford, Lee, and Ramli Ukasha. Forthcoming. Discrimination by politicians against religious minorities: Experimental evidence from the UK. *Party Politics*.
- Dur, Robert, and Robin Zoutenbier. 2015. Intrinsic motivations of public sector employees: Evidence for Germany. *German Economic Review* 16: 343-366.
- 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 Commission. 2015. Eurobarometer on Discrimination in the EU in 2015 - Factsheets Czechia. Online at <https://webgate.ec.europa.eu/ebsm/api/public/deliverable/download?doc=true&deliverableId=51296> (last accessed October 28, 2021).
- 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 30, 2021).
- 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 30, 2021).
- 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 30, 2021).

- European Union Agency for Fundamental Rights. 2016. Second European Union Minorities and Discrimination Survey Roma -- Selected findings. Online at https://fra.europa.eu/sites/default/files/fra_uploads/fra-2016-eu-minorities-survey-roma-selected-findings_en.pdf (last accessed October 28, 2021).
- 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.
- Linos, Katerina, Melissa Carlson, Laura Jakli, Nadia Dalma, Isabelle Cohen, Afroditi Veloudaki, and Stavros Nikiforos Spyrellis. Forthcoming. How Do Disadvantaged Groups Seek Information About Public Services? A Randomized Controlled Trial of Communication Technologies. *Public Administration Review*.

- Linos, Katerina, Laura Jakli, and Melissa Carlson. 2021. Fundraising for Stigmatized Groups: A Text Message Donation Experiment. *American Political Science Review* 115: 14-30.
- 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.
- Office of the Government. 2019. Zpráva o stavu romské menšiny v České republice za rok 2019. [Report on the state of the Roma minority in the Czech Republic in 2019.] Online at www.vlada.cz/assets/ppov/zalezitosti-romske-komunity/aktuality/Zprava-o-stavu-romske-mensiny-2019.pdf (last accessed December 30, 2021).
- 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 30, 2021).
- 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 30, 2021).
- Sterkens, Philippe, Ralf Caers, Marijke De Couck, Michael Geamanu, Victor Van Driessche, and Stijn Baert. 2021. Costly Mistakes: Why and When Spelling Errors

- in Resumes Jeopardise Interview Chances. GLO Discussion Paper Series 899, Global Labor Organization (GLO). Online at <https://www.econstor.eu/bitstream/10419/235907/1/GLO-DP-0899.pdf> (last accessed December 30, 2021).
- 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.

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_{01} , 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 $p_{11} + p_{10} + p_{01} + p_{00} = 1$.

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 error frequencies and the 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, the 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 happen for two main reasons: positive discrimination in favor of Roma enquirers by some subjects, and the fact that some subjects may respond to emails randomly. We believe that positive discrimination of Roma is not likely very frequent, but random responses may be. If we set $p_{01} = 0.05$ (randomness in the response occurs with the same frequency as discrimination), $\delta = 0.05$ implies $p_{10} = 0.1$.

The power for the 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 is 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, Seř mi vyhodil a nemam práci. Podadíte mne co potrebuj udelat aby ste mi daly potporu? Moc díky
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 odednežka 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 práci 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 práci a chtel bysem 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! Předvěře 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. Porad'te jak na to. Díky moc