

# Pre-Analysis Plan for: Barriers in accessing sports: A field experiment at the intersection of sexual orientation and ethnicity

Carlos Gomez-Gonzalez<sup>\*</sup>, Cornel Nessel<sup>†</sup>, Helmut Dietl<sup>†</sup>

May 31, 2024

## Abstract

We are conducting a correspondence study to examine the extent of discrimination based on ethnicity, sexual orientation, and LGBTQ support. We contact amateur sports clubs in Germany. In this analysis plan, we pre-register some key decisions we will follow once the data is collected.

**Keywords:** Correspondence study, discrimination, sports

*JEL Classification:* J15, C92, C93

---

<sup>\*</sup>Corresponding author: carlosgomez.id@gmail.com

<sup>†</sup>University of Zurich, Switzerland

<sup>‡</sup>University of Stavanger, Norway

## 1 Introduction

Discrimination based on race, ethnicity, gender, religion, or sexual orientation is highly persistent in multiple settings. Much research is needed to understand the mechanism behind differential treatment (Bertrand and Duflo, 2017). One of the research areas that require attention is how individuals are affected at the intersection of these traits. In this project, we are interested in the intersection of LGBTQ support and sexual orientation with ethnicity.

We perform a large-scale field experiment in Germany based on a correspondence study. Fictitious applicants contact coaches via email asking to join a trial training with a native- or a foreign-sounding name. The emails are identical, and the response rate to the different groups can provide evidence of discrimination. The experimental design is very similar to previous field experiments we performed in several European countries (Gomez-Gonzalez et al., 2021) and has a twofold aim.

First, we aim to examine if showing support for the LGBTQ movement negatively influences an individual's chance to obtain a response when contacting a sport club. Additionally, we examine if foreign background and gender play a role in this relationship. Our treatment showing support for the LGBTQ movement is a subtle signal in the signature of the email during pride month.

Second, we aim to examine if signaling homosexuality negatively influences an individual's chance to obtain a response when contacting a sport club. Additionally, we examine if foreign background and gender play a role in this relationship. Our treatment signals homosexuality by mentioning that the applicant has just moved to the area with a female/male partner.

In total, we contact 6023 amateur sports clubs.

## 2 Experiment and sample

**Sample.** The dataset includes 6023 amateur sports clubs in Germany from four different sports: 1504 football clubs, 1512 horse riding clubs, 1504 table tennis clubs, and 1503 tennis clubs.

**Experimental Design.** We want to compare the response rate for individuals who contact a sports club. We compare the response rate for individuals with a typical German- and foreign-sounding name. All the clubs in the sample will receive one email only. The content of the email is identical for all clubs except for the name of the applicant and the treatment conditions, which are randomly varied.

1. Name of the applicant: The name is either female or male. We use the three names for each group

for each nationality. Thus, six names for German-sounding names (three for males and three for females) and six names for each of the three largest foreign groups in Germany: Turkish-, Syrian-, and Ukrainian-sounding names. All first and last names are very common in the respective countries. The list of names used is available in Appendix E.

2. Treatment conditions: We include one baseline message and 3 different treatments. The baseline message is a straightforward email in which the fictitious applicants ask for the opportunity to come over to the club. Treatment 1 is a subtle signal in the signature of the email showing support for the LGBTQ movement. Treatment 2 signals heterosexuality by mentioning that the applicant has just moved to the area with a female/male partner, and Treatment 3 does the same to signal homosexuality. The English translation of all messages are available in Appendix A-D.

We use block randomization by state and sport. This ensures that individual names are not overrepresented in sports or geographical areas. We use Stata and the `set seed` command to make the randomization process reproducible.

One unit is a club. For each sport, we split our sample as follows: Approximately 1/4 of clubs are randomly assigned to local name, 1/4 of clubs to foreign group 1, 1/4 of clubs to foreign group 2, and 1/4 of clubs to foreign group 3. Additionally, for each sport, we split our sample evenly among the four conditions: approximately 1/4 of clubs are randomly assigned to the Baseline message, 1/4 of clubs to Treatment 1, 1/4 of clubs to Treatment 2, and 1/4 of clubs to Treatment 3.

**Decision rules for dropping observations.** We will only drop clubs for which we receive an error message after sending the email, e.g., "Address not found".

**Decision rules for dropping variables.** All the data on clubs was collected before the experiment, and no other data except the outcome variable will be collected. We therefore do not intend to drop any variable.

**Missing values.** All the clubs for which we did not have specific information (in particular contact information) were excluded from the initial sample. No other missing values are expected.

### 3 Data and coding of main variables

**Treatment variables.** The main treatment variable is a dummy variable named *Foreign name*. We will also divide the treatment into sub treatments corresponding to the different origins of the fictitious applicants (Germany, Turkey, Syria, and Ukraine).

Additionally, the analyses include three dummy variables for each of the treatments: Treatment 1 (LGBTQ support), Treatment 2 (signal heterosexuality), and Treatment 3 (signal homosexuality.)

**Outcome variable.** The outcome variable is a dummy variable for a positive response.

There are 4 potential outcomes: 1) no response, 2) negative response, 3) response with additional questions (how did you find the club, in what position did the applicant play, at what level, etc) or 4) a positive response (invitation to join club). We will code the outcome variable as 1 for the categories 3 and 4, and 0 for categories 1 and 2. The email shows only interest in joining a trial training session and avoids references to further attributes, as in ([Gomez-Gonzalez et al., 2021](#)).

**Controls.** Controls will include sport-fixed effects and gender. In further analyses, we will also interact the foreign name dummy with the treatments. We might also include controls at the postcode level (population size, share of foreign-born population, and right-win votes).

**Heterogeneity.** No other sources of heterogeneity will be investigated.

### 4 Hypotheses

*H1. Discrimination against individuals with a foreign-sounding name.* Individuals with a foreign-sounding name receive fewer positive responses than individuals with local-sounding names.

*H2. Discrimination against individuals who support the LGBTQ movement.* The LGBTQ support signal negatively influences the probability of receiving a positive response.

*H3. Discrimination against homosexual individuals.* Signaling being in a same sex relationship negatively influences the probability of receiving a positive response.

We have two additional hypotheses for the intersection of ethnicity:

*H2a. Discrimination against individuals who support the LGBTQ movement is more pronounced for applicants with a foreign-sounding name.* The interaction between LGBTQ support signal and foreign

name negatively influences the probability of receiving a positive response.

*H3a. Discrimination against homosexual individuals is more pronounced for applicants with a foreign-sounding name.* The interaction between the homosexuality signal and a foreign name negatively influences the probability of receiving a positive response.

## 5 Empirical Strategy

To test for H1, we will estimate the following probit model, for club  $i$  in sport  $j$ :

$$P(\text{Positive response}_{ij} = 1|X) = \Psi(\alpha_1 + \beta_1 \text{Foreign}_i + \mu_j^1 + \epsilon_{ij}^1) \quad (1)$$

where  $\Psi$  is the cumulative distribution function of the normal distribution,  $\text{Foreign}$  is a dummy variable for whether the applicant has a foreign-sounding name, and  $\mu_j^1$  is a sport-fixed effect.  $\epsilon$  is the error term. We would consider rejecting the null if  $\beta_1 \neq 0$  with  $p < 0.01$ .

To test for H2, we will estimate the following probit model, for club  $i$  in sport  $j$ :

$$P(\text{Positive response}_{ij} = 1|X) = \Psi(\alpha_2 + \beta_2 \text{T1}_i + \mu_j^1 + \epsilon_{ij}^1) \quad (2)$$

where  $\Psi$  is the cumulative distribution function of the normal distribution,  $\text{T1}$  is a dummy variable for whether the applicant shows support for the LGBTQ movement, and  $\mu_j^1$  is a sport-fixed effect. We would consider rejecting the null if  $\beta_2 \neq 0$  with  $p < 0.01$ .

To test for H3, we will estimate the following probit model, for club  $i$  in sport  $j$ :

$$P(\text{Positive response}_{ij} = 1|X) = \Psi(\alpha_3 + \beta_3 \text{T3}_i + \mu_j^1 + \epsilon_{ij}^1) \quad (3)$$

where  $\Psi$  is the cumulative distribution function of the normal distribution,  $\text{T3}$  is a dummy variable for whether the applicant signal homosexuality, and  $\mu_j^1$  is a sport-fixed effect. We would consider rejecting the null if  $\beta_3 \neq 0$  with  $p < 0.01$ .

To test for H2a and H3a, we add the treatments and the interactions with the foreign-sounding name. Specifically, we will estimate the following equation.

$$\begin{aligned}
P(\text{Positive response}_{ij} = 1 | X) = & \Psi(\alpha_4 + \beta_4 \text{Foreign}_i + \beta_5 \text{T1}_i \\
& + \beta_6 \text{Foreign}_i \times \text{T1}_i + \mu_j^4 + \epsilon_{ij}^4)
\end{aligned} \tag{4}$$

where  $T1_i$  is a dummy variable for whether the message includes the LGBTQ-support treatment. The null for hypothesis 4 will be rejected if  $\beta_6 \neq 0$  with  $p < 0.01$ .

Finally, we plan to run regressions 1-4 for each sport individually to examine if the specific results differ between sports.

Note: We include the same specification to test for Foreign and the homosexuality treatment (T3). In a step-wise regression model, we will add the gender of the applicant and include the other controls described above.

## 6 Power Calculation

We assume that the sample will be evenly split among the four groups and treatments. In a previous study, the response rate for German-sounding names was 0.67 and 0.53 for foreign-sounding names ([Gomez-Gonzalez et al., 2021](#)).

Assuming a similar response rate for this study, the parameters were the following: significance level = 0.05, power = 0.95, and expected effect size ( $d$ ) = 0.28. This means we need at least 290 observations for each group. We have four experimental groups: control; treatment 1, treatment 2, and treatment 3. For a basic analysis, we would need  $4 * 290 = 1160$  observations.

## 7 IRB Approval

The project received approval from the Human Subjects Committee of the Faculty of Economics, Business Administration and Information Technology from the University of Zurich on April 23, 2024 (OEC IRB # 2024-045).

We will not ask for informed consent from clubs. However, for all clubs who answer the email, we will reply within a couple of days to tell them that the applicant is no longer interested in becoming a member.

## 8 Archive

The present pre-analysis plan is archived before any data is collected. We archived it at the registry for randomized controlled trials in economics held by the American Economic Association: <https://www.socialscienceregistry.org/> on Feb. 10, 2023. The reference has the title "Barriers in accessing sports: A field experiment at the intersection of sexual orientation and ethnicity".

## References

Bertrand, M. and Duflo, E. (2017). Field experiments on discrimination. In *Handbook of Economic Field Experiments*, volume 1, pages 309–393. Elsevier.

Gomez-Gonzalez, C., Nessler, C., and Dietl, H. M. (2021). Mapping Discrimination in Europe Through a Field Experiment in Amateur Sport. *Humanities and Social Sciences Communications*, 8(1):1–8.

## Appendices

### A English translation of the baseline message

The English translation of the baseline message from applicants is as follows:

Subject: Club membership

Hello,

I recently moved here and would like to join your SPORTS club. I was in a club before I moved and would like to start again. Is it possible to come over?

Thank you

Name

### B English translation of the Treatment 1

Subject: Club membership

Hello,

I recently moved here and would like to join your SPORTS club. I was in a club before I moved and would like to start again. Is it possible to come over?

Thank you

Name

—

PrideMonth2024 

### C English translation of the Treatment 2

Subject: Club membership

Hello,

I recently moved here with my girlfriend (boyfriend) and would like to join your SPORTS club. I was in a club before I moved and would like to start again. Is it possible to come over?

Thank you

Male name (Female name)

## **D English translation of the Treatment 3**

Subject: Club membership

Hello,

I recently moved here with my boyfriend (girlfriend) and would like to join your SPORTS club. I was in a club before I moved and would like to start again. Is it possible to come over?

Thank you

Male name (Female name)

## E List of names by group

<b>Group names</b>	<b>Male name</b>	<b>Female name</b>
Germany	Alexander Müller	Marie Müller
	Leon Fischer	Anna Fischer
	Paul Schneider	Maria Schneider
Turkey	Mehmet Yilmaz	Fatma Yilmaz
	Mustafa Kaya	Aye Kaya
	Ahmet Demir	Emine Demir
Ukraine	Grygoriy Kovalenko	Iryna Kovalenko
	Oleg Bondarenko	Alina Bondarenko
	Ivan Melnyk	Sofiya Melnyk
Syria	Hassan Ismat	Donya Ismat
	Ibrahim Khalil	Amira Khalil
	Abdel Hussain	Fatima Hussain

Table E.1: Names categorized by group and gender