

Pre-Analysis Plan for: A large-scale field experiment to disentangle sources of statistical discrimination in a social setting*

Carlos Gomez-Gonzalez^{1,*}, Gwen-Jirō Clochard^{2,3,*}, Helmut Dietl¹, and Juan Cruz Duhalde^{3,4}

¹Department of Business Administration, University of Zurich

²Institute of Social and Economic Research, Osaka University

³Joint Initiative for Latin American Experimental Economics

⁴Universidad del CEMA

*Corresponding authors: carlos.gomez@business.uzh.ch; gjclochard@iser.osaka-u.ac.jp

February 22, 2024

Abstract

We are conducting a correspondence study to investigate the extent of discrimination of immigrants in the context amateur football clubs in 15 Latin American countries. In this analysis plan, we pre-register some key decisions we will follow once the data is collected.

Keywords: correspondence study, discrimination, football

JEL Classification: J15, C92, C93

*The project was approved by the Human Subjects Committee of the Faculty of Economics, Business Administration and Information Technology from the University of Zurich on February 12, 2024 (OEC IRB # 2024-013).

1 Introduction

The gold standard to measure the magnitude of discrimination is correspondence studies (Verhaeghe, 2022), in which researchers send fictitious applications, changing only the characteristics of interest (e.g., name, gender, level of education, origin). To cite a few examples, correspondence studies have shown that prejudice against minorities exists in contexts such as employment (Bertrand and Mullainathan, 2004), housing (Auspurg et al., 2019), transportation (Ge et al., 2020), and the sharing economy (Liebe and Beyer, 2021).

In the discrimination literature, a long-standing debate is whether observed differences between groups stem from statistical discrimination – in a presence of uncertainty, using group characteristics could be rational (Arrow, 1972) – or taste-based discrimination, where differences arise from an innate animus against (or fondness for) a particular group (Becker, 1957). Correspondence studies have the advantage of a flexible design that allows varying *only* certain attributes that can influence the outcome, and therefore reduce the prevalence of statistical discrimination. However, even in a context where the research design could control all signals sent by the applicant, some uncertainty still remains about the quality of the applicants, their background, or their motivations, leaving room for statistical discrimination. Furthermore, there can still exist statistical discrimination against certain groups if part of the production function is due to interactions between members of the organization (e.g., colleagues).

In this study, we argue that there can actually exist two forms of statistical discrimination: individual and collective. Using a correspondence study design, we aim to identify their relative importance. *Individual statistical discrimination* would occur when there exist doubts about someone's individual abilities, such as whether a particular worker would be able to perform a specific technical task. *Collective statistical discrimination* would occur when the uncertainty does not lie with the individual's abilities but rather with their capacity to work as a team with other workers. Information on sources and underlying causes of discrimination is essential to direct policy efforts.

We would test for the relevance of these two sources of statistical discrimination in the context of discrimination against immigrants applying to play in amateur football clubs in 15 Latin American countries (see the list in Appendix A). This allows us to standardize the design across several countries with no variations, include a large number of existing clubs (approximately 15,000 men's clubs), and identify small effects of the magnitude of 2 percentage points (See Section 6).

We would send applications via email, phone, and social media, asking to join the club, the core of which would be identical for all players. However, we would randomly vary the name and the country

of origin of the applicant (for many Latin American countries, the most common names are identical, so the application will be explicit about the origin, see Section 2). Also varied are sentences aimed at distinguishing between the two types of statistical discrimination. To test for individual statistical discrimination, we add a sentence, saying that the player has played in the academy of a second division club in Italy, thus signaling the player's (considerable) ability. To test for collective statistical discrimination, we add a statement that the player speaks the local language (Spanish or Portuguese).

Furthermore, one drawback of correspondence studies is that they can only shed light on the initial discrimination - that is, to get to the first round of interviews ([Bertrand and Duflo, 2017](#)). In this paper, for the subset of clubs who reply asking the applicant for further information, we will use a standardized response (the same for each respondent), to see if clubs further discriminate in the second round.

2 Experiment and sample

Sample The sample will consist of approximately 15,000 men's amateur football clubs in 15 Latin American countries for which we were able to gather contact information.¹

Experimental Design There are two dimensions of treatment in our experiment. The first dimension is relative to the *content of the message*. We will randomize the content of the application message that clubs receive. The reference treatment (T_0) is a core *neutral* message (Appendix B). For the other treatment groups, we include an additional message to the core of the email.

For the *Individual* treatment (T_1), the following sentence will be added: "When I was young, I played in the academy of Cremonese in Italy." For the amateur clubs in our sample, the treatment would probably signal a higher level than their average player, so we expect it to have an effect on individual statistical discrimination (Hypothesis 4).

For the *Collective* treatment (T_2), the added sentence is "I did part of my high-school curriculum in Spanish [Portuguese in Brazil] so I speak Spanish [Portuguese] well." The *Collective* treatment is designed to signal the player's ability to communicate with others in the team, i.e., to have an effect on collective statistical discrimination (Hypothesis 5).

Lastly, we include a treatment that combines the individual and collective treatments (T_1+T_2). The order of the treatments is randomized in the text of the email to avoid order effects.

¹We were only able to find information for approximately 700 female clubs for the entire region, so for power considerations, we decided not to include female clubs in the analysis.

The second dimension of treatment is relative to the origin of the applicants. In each country, census data are used to determine the countries used. We use the four largest countries in each of the following three regions: Latin America, Europe, and Asia.² The list of countries of origin for each club country in our data is presented in Appendix D. We used census data to get the two most common first and last names for children born in 2000.

Treatments are summarized in Figure 1. Each club will be randomly placed in one of the 16 (4x4) treatment cells. All the analyses are **between-clubs**. Between-clubs analysis means that the clubs do not receive similar-sounding applications from two different players, and we are restrained from identifying individual measures of discrimination of non-consenting participants.

Assignment to treatment will be stratified by country, i.e., there will be approximately the same share of clubs in each treatment cell for all countries. Wherever applicable, we will also stratify treatments by region/province.

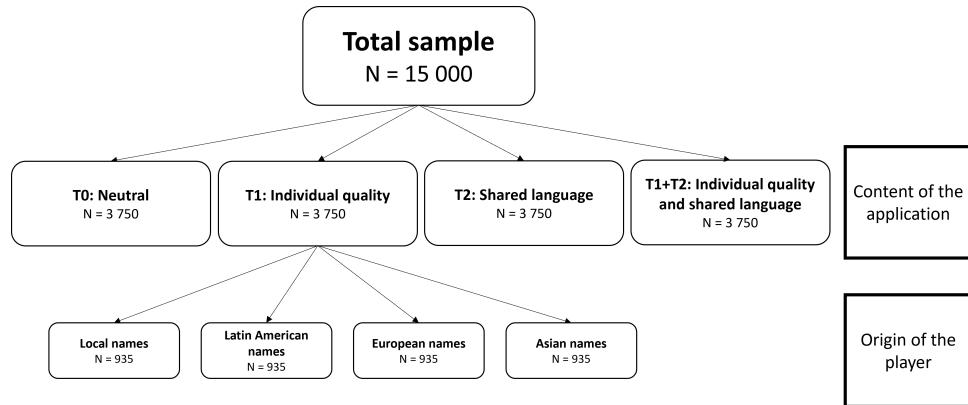


Figure 1: Treatment arms for the main experiment. *Note:* The applicant's origin (Local, Latin American, European, and Asian names) are applied to each content treatment (*Neutral*, *Individual quality* and *Shared language*).

Data collection and analysis will not be performed blind to the conditions of the experiment, because of feasibility constraints. The main reason for this is that we will track the responses to the applications sent by email or message, so we could always see the name of the applicant and the content of the application.

Decision rules for dropping observations. We will drop clubs for which the email comes back with an error message.

²Many Latin American countries have a too small a population of individual African countries, or sometimes do not report the disaggregated statistics by country. We therefore decided not to include African countries in the analysis.

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. There will be two dimensions of treatment: Content of the application and Origin of the applicant.

Treatments as well as expected sample sizes in each group are presented in Figure 1.

Outcome variable. The outcome variable will be a dummy variable for whether the email receives a positive response.

To be specific, there are 4 potential outcomes: 1) no response, 2) negative response, 3) response with additional questions (in what position did the applicant play, at what level, etc) or give the number of the coach for a call, or 4) a positive response (invitation to training). We will code the variable as 1 for categories 3 and 4, and 0 for categories 1 and 2.

Furthermore, for clubs who asked for additional information, we will provide a reply with the content displayed as Appendix C. We will classify the outcome variable "Second round" similarly to the response in the first round.

Controls. Controls will include country-fixed effects. We might also include controls at the city level. From our hypotheses, we will also test for the presence of discrimination as a function of whether the applicant comes from a country that shares the language of the club (Spanish or Portuguese), and with the FIFA ranking of the country.

Heterogeneity. No other sources of heterogeneity will be investigated.

4 Hypotheses

Hypothesis 1 (presence of discrimination against immigrants in the amateur football context in Latin America): Foreign applicants receive fewer positive responses than local applicants.

Hypothesis 2 (presence of individual statistical discrimination): There is a negative correlation between positive response rate and the FIFA ranking of the country, i.e., applicants coming from countries with a lower FIFA ranking receive fewer positive responses.

Hypothesis 3 (presence of collective statistical discrimination): There is a negative correlation between positive response rate and applicants with a different first language, i.e., applicants from countries that do not share a common language with the country of the club receive fewer positive responses.

Hypothesis 4 (reduction in individual statistical discrimination): The treatment signaling individual quality has a positive effect for applicants from countries with a lower FIFA ranking.

Hypothesis 5 (reduction in collective statistical discrimination): The collective treatment, signaling that the player is fluent in the language of the club country increases positive response rate.

We have an additional hypothesis for the subsample who will reply asking for more information. This hypothesis is exploratory.

Hypothesis 6 (second round discrimination): There exists discrimination against foreign applicants in the second round of applications.

5 Empirical Strategy

To test for Hypothesis 1, i.e., whether the positive response rate for foreign applicants is lower than for natives, we will estimate the following probit model, for club i in country 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 Ψ is the cumulative distribution function of the normal distribution, Foreign is a dummy variable for whether the player is a foreigner, and μ_j^1 is a country-fixed effect. We would consider rejecting the null if $\beta_1 \neq 0$ with $p < 0.01$.

To test for Hypothesis 2, we will estimate the following equation:

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

where $Ranking_i$ is the value of the FIFA ranking of the country of the applicant. The null for hypothesis 2 will be rejected if $\beta_2 \neq 0$ with $p < 0.01$.

To test for hypothesis 3, we will similarly estimate for the following equation:

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

where $Language_i$ is a dummy variable for whether the language spoken in the applicant's country of origin is different from the language spoken in the club's country. The null for hypothesis 3 will be rejected if $\beta_3 \neq 0$ with $p < 0.01$.

For hypothesis 4, we would add the treatments and interactions to test for the effect of treatment T1 (individual quality treatment). Specifically, we will estimate the following equation.

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

where $T1_i$ is a dummy variable for whether the application belongs to treatment T1. The null for hypothesis 4 will be rejected if $\beta_6 \neq 0$ with $p < 0.01$.

For hypothesis 5, we would similarly interact with the treatment T2 (collective treatment):

$$\begin{aligned} P(\text{Positive response}_{ij} = 1|X) = & \Psi(\alpha_5 + \beta_7 \text{Language}_i + \beta_8 \text{T2}_i \\ & + \beta_9 \text{Language}_i \times \text{T2}_i + \mu_j^5 + \epsilon_{ij}^5) \end{aligned} \quad (5)$$

where $T2_i$ is a dummy variable for whether the club belongs to the collective treatment. The null for hypothesis 5 will be rejected if $\beta_9 \neq 0$ with $p < 0.01$.

To test for Hypothesis 6, i.e., whether the positive response rate for foreign applicants is lower than for natives in the second round, we will estimate the following probit model, for club i in country j :

$$P(\text{Positive response}_{ij}^2 = 1|X) = \Psi(\alpha_6 + \beta_{10} \text{Foreign}_i + \mu_j^6 + \epsilon_{ij}^6) \quad (6)$$

where $\text{Positive response}_{ij}^2$ is the positive variable dummy in round 2. The null will be rejected if $\beta_{10} \neq 0$ with $p < 0.05$.

6 Power Calculation

The power analysis was conducted with the Optimal Design software ([Raudenbush et al., 2011](#)).

For the 5 first hypotheses, the parameters were the following: significance level $\alpha = 0.01$ (corresponding to 5 hypotheses, following Bonferroni's correction), power $\delta = 0.80$ and sample size $N = 15,000$. According to the calculations, the Minimum Detectable Effect is 0.05 standard deviation.

With data from the pilot experiment in Argentina, where we had an average positive response rate of 20%, the Minimum Detectable Effect is approximately 2 percentage points.

For Hypothesis 6, we use the sample size from the pilot data in Argentina to compute the sample size. In the pilot data, approximately 4% of the responses were asking for follow-up information. This means that, assuming that the proportion would be similar in the entire sample, we would have a sample of approximately 600 clubs. Using $p = 0.05$ and $\delta = 0.80$, we obtain a minimum detectable size of 0.23 standard deviation. We do not have a clear benchmark of the number of answers in the second round, but assuming 75%, this gives a MDE of approximately 10 percentage points.

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 February 12, 2024 (OEC IRB # 2024-013).

We will not ask for informed consent from clubs. However, for all clubs who answer the email, we will reply within 2 days to tell them that the applicant is no longer interested.

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. 25, 2024. The reference is XXX.

References

- Arrow, K. J. (1972). Some mathematical models of race discrimination in the labor market. In *Racial discrimination in economic life*, pp. 187–204. Lexington Books.
- Auspurg, K., A. Schneck, and T. Hinz (2019). Closed doors everywhere? A meta-analysis of field experiments on ethnic discrimination in rental housing markets. *Journal of Ethnic and Migration Studies* 45(1), 95–114.
- Becker, G. S. (1957). *The economics of discrimination*. University of Chicago press.
- Bertrand, M. and E. Duflo (2017). Field experiments on discrimination. In *Handbook of economic field experiments*, Volume 1, pp. 309–393. Elsevier.
- Bertrand, M. and S. Mullainathan (2004). Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *American Economic Review* 94(4), 991–1013.
- Ge, Y., C. R. Knittel, D. MacKenzie, and S. Zoepf (2020). Racial discrimination in transportation network companies. *Journal of Public Economics* 190, 104205.
- Liebe, U. and H. Beyer (2021). Examining discrimination in everyday life: A stated choice experiment on racism in the sharing economy. *Journal of Ethnic and Migration Studies* 47(9), 2065–2088.
- Raudenbush, S. W., J. Spybrook, R. Congdon, X.-f. Liu, A. Martinez, H. Bloom, and C. Hill (2011). Optimal design software for multi-level and longitudinal research (version 3.01)[software].
- Verhaeghe, P.-P. (2022). Correspondence studies. In *Handbook of labor, human resources and population economics*, pp. 1–19. Springer.

Appendices

A Countries to be included in the experiment



Figure A.1: Countries to be included in the experiment

Table A.1: List of countries and sample size for the main experiment

Country	Expected number of clubs in sample
Argentina ¹	1 300
Bolivia	594
Brazil	4 718
Chile	3 399
Colombia	2 138
Costa Rica	154
Ecuador	818
El Salvador	372
Honduras	139
Mexico	449
Panama	260
Paraguay	287
Peru	709
Uruguay	376
Venezuela	460
Total	16 173

Note: The sample size used in our estimations is $N = 15\,000$. ¹ Argentina is only counted as a pilot without the content treatments.

B Text of the email

The English translation of the text of the application from applicants is as follows:

Hello,

I have recently moved to the city. I come from Y (capital city of the country) and would like to play football in your club to meet people and enjoy playing football. I have already played at a similar level.

Could I come for a trial training session with the senior team?

Thank you very much,

Name

C Reply for additional information

Thank you for your reply. Yes, of course. I am 26 years old and in the last few years, I have played in several teams at the regional level. I typically play as a midfielder; where I enjoy the most is as an attacking midfielder. Now I am looking for a club and I decided to contact you to see if I could join a training session.

D List of countries of origin

Table D.1: List of countries of origin per country

Country of clubs	Latin America	Europe	Asia
Argentina	Bolivia	Italy	China
	Peru	Spain	South Korea
	Chile	France	Japan
	Paraguay	Germany	Taiwan
Bolivia	Argentina	Spain	Japan
	Brazil	Germany	South Korea
	Peru	Italy	
	Mexico	France	
Brazil	Venezuela	Portugal	Japan
	Paraguay	Italy	China
	Bolivia	Spain	Lebanon
	Argentina	Germany	South Korea
Chile	France	Bolivia	China
	Germany	Colombia	
	Italy	Peru	
	Spain	Venezuela	
Colombia	Venezuela	Spain	China
	Ecuador	Italy	Japan
	Argentina	France	
	Peru	Germany	
Costa Rica	Nicaragua	Spain	Hong Kong
	Venezuela	Italy	
	Colombia	Germany	
	El Salvador	France	
Ecuador	Colombia	Spain	China
	Peru	Germany	South Korea
	Chile	Italy	Japan
	Venezuela	France	Israel
El Salvador	Honduras	Spain	China
	Guatemala	Italy	South Korea
	Nicaragua	Germany	Japan
	Mexico	France	

Note: The table presents the countries of origins of the largest number of immigrants in each country, with a maximum of four by continent.

Table D.2: List of countries of origin per country, continued

Country of clubs	Latin America	Europe	Asia
Honduras	Guatemala	Spain	China
	El Salvador	Germany	Palestine
	Nicaragua	Italy	Japan
	Mexico	France	Israel
Mexico	Guatemala	Spain	Japan
	Venezuela	France	China
	Colombia	Germany	India
	Honduras	Italy	Israel
Panama	Venezuela	Spain	China
	Colombia	United Kingdom	India
	Nicaragua	Italy	Philippines
	Dominican Republic	Germany	Israel
Paraguay	Brazil	Germany	South Korea
	Argentina	Spain	Japan
	Uruguay	Italy	China
	Chile	Poland	
Peru	Venezuela	Spain	China
	Bolivia	Germany	Japan
	Argentina	Italy	South Korea
	Colombia	France	India
Uruguay	Argentina	Spain	China
	Venezuela	Italy	
	Brazil	Germany	
	Chile	Poland	
Venezuela	Colombia	Spain	Syria
	Peru	Italy	China
	Ecuador	Germany	Lebanon
	Chile	France	Saudi Arabia