Pre-Analysis Plan for:

Arriving in Argentina: A Field Experiment With Football Clubs

Carlos Gomez-Gonzalez[*], Gwen-Jirō Clochard[‡], Helmut Dietl[*]

February 10, 2023

Abstract

We are conducting an audit study to investigate the extent of discrimination of foreigners in Argentinian amateur football clubs. In this analysis plan, we pre-register some key decisions we will follow once the data is collected.

Keywords: Audit study, discrimination, football

JEL Classification: J15, C92, C93

[*]University of Zurich

[†]Corresponding author: gjclochard@uchicago.edu

[‡]University of Chicago and Universidad del CEMA's Joint Initiative for Latin American Experimental Economics - CREST

1 Introduction

Differential treatment of immigrants has been observed in many countries and in many contexts, be it employment (Bertrand and Mullainathan, 2004), housing (Auspurg et al., 2019) or transportation (Liebe and Beyer, 2021).[1] This has in particular been identified in a number of studies involving football - or soccer - (Gomez-Gonzalez et al., 2021).

This study aims at investigating the extent to which there exists discrimination against immigrants in Argentinian football. Furthermore, we plan to establish heterogeneity by countries involved.

To do so, we intend to run an audit study with Argentinian amateur football clubs. We plan to send them fictitious application emails from prospective players, randomly varying the origin of the applicant (Argentina vs. other country).

2 Experiment and sample

Sample The sample will consist of approximately 1028 male and 413 female amateur football clubs in Argentina for which we were able to gather email addresses. The clubs are present in all 23 provinces of the country, as well as the federal capital.

Experimental Design All the clubs in the sample will receive one email. The content of the email will in large part be similar for all clubs (see Appendix A), with the exception of certain details, which will be randomly varied.

In particular, the main distinction is whether the fictitious applicant is coming from Argentina or a foreign country.

Foreign applicants will be from the countries with the highest counts of immigration, according to the last available census data in Argentina. The countries from which the immigrants come from will be: Bolivia, Chile, Paraguay and Peru for South America, France, Germany, Italy and Spain for Europe and China, Japan, South Korea and Taiwan for Asia.

The sample will be evenly split between the four regions of treatment: Argentina, Asia, Europe, South America.

For every country, we used census data (or the best available information) to get the two most common first (male and female) and last names for children born around the year 2000. For Argentina, we

[1] For a review of experiments on discrimination, see Bertrand and Duflo (2017).

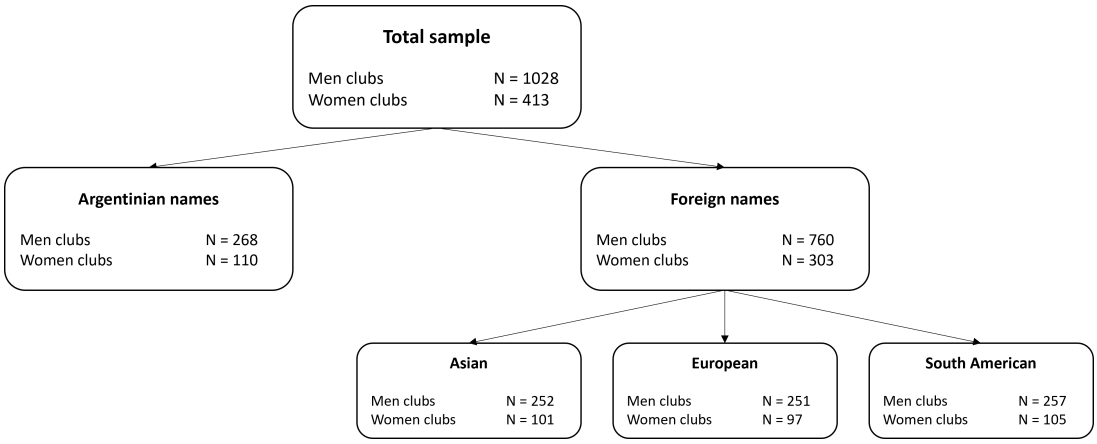

Figure 1: Treatments arms and sample size

will use the five most common names. The list of names used is available in Appendix B.

The treatments will be stratified by province and by team gender.

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

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 will be a dummy variable named *Foreign name*.

We will also divide the treatment into sub treatments corresponding to the different origins of the fictitious applicants (*South America*, *Europe* and *Asia*).

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

Outcome variable The outcome variable is going to 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 4) a positive response (invitation to training). We will code the variable as 1 for the categories 3 and 4, and 0 for categories 1 and 2.

Controls Controls will include the province, level of division in which the team plays, a dummy variable for male or female clubs. We will also add controls at the city level (e.g., population or political affiliation of the city mayor).

Heterogeneity Other than the main treatment (foreign versus Argentinian name), we will investigate whether the callback rate is different depending on the region of origin of the applicant. We will also investigate which potential characteristics of countries could explain differences (e.g. geographical information, socioeconomic status, football success, etc.).

4 Empirical Strategy

The empirical strategy is standard for a between-subject design, as in Equation 1. Y_{ig} corresponds to the outcome variable for club i in province p. The main coefficient of interest is β_1.

$$Y_{ip} = \beta_1 \text{Foreign name}_{ip} + \beta X_{ip} + women_{ip} + \nu_p + \epsilon_{ig} \quad (1)$$

To test for the heterogeneity of treatment by origin of the applicant, we will use a dummy variable for each continent of foreign applicants (South America, Europe and Asia). To do so, we will estimate the following regression.

$$Y_{ig} = \gamma_0 + \gamma_1 \text{SouthAmerica}_{ig} + \gamma_2 \text{Europe}_{ig} + \gamma_3 \text{Asia}_{ig} + \gamma X_{ig} + women_{ig} + \delta_g + \epsilon_{ig} \quad (2)$$

To test for the heterogeneity based on other dimensions, we will interact the treatment variable with dimensions of interest (see above).

5 Hypotheses

There are mainly four null hypotheses that we will test, for β_1, γ_1, γ_2 and γ_3.

6 Power Calculation

We test for four main hypotheses, so we correct for 4 independent tests using Bonferroni's correction.

We assume that the sample will be evenly split among the four groups (Argentinian names, South American, European and Asian). We also assumed that 10% of the variance can be explained by control variables, and use the conventional $\delta = 0.80$ value for power.

According to the calculations made with the Optimal Design software (Raudenbush et al., 2011), the Minimum Detectable Effect is 0.32.

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 the 12 September 2022 (OEC IRB # 2022-073).

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 XXX. The reference is XXX.

References

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.

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.

Gomez-Gonzalez, C., C. Nesseler, and H. M. Dietl (2021). Mapping discrimination in europe through a field experiment in amateur sport. *Humanities and Social Sciences Communications 8*(1), 1–8.

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].

Appendices

A Text of the email

The text of the email from applicants was the following:

Hello,

I have just arrived in X (city of the club) . 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

For Argentinian applicants, the email was the same, except that the Y (capital city of the country) is replaced by either “Buenos Aires” (capital city of Argentina) for clubs in any region but the federal capital, and “Córdoba” (second largest city) for clubs in the capital.

B List of names per country

Table B.1: List of names used

Zone	Country	Male names			Female names		
		First name	Last name	Id	First name	Last name	Id
Argentina	Argentina	Juan Ignacio	González	1	Valentina	González	30
		Agustín	Rodríguez	2	Camila	Rodríguez	31
		Tomás	Gómez	3	Martina	Gómez	32
		Santiago	Fernández	4	Agustina	Fernández	33
		Juan Pablo	López	5	María Belén	López	34
South America	Paraguay	Ramón	González	22	María	González	51
		Juan	Benítez	23	Elizabeth	Benitez	52
	Chile	Agustín	González	24	Sofía	González	53
		Mateo	Muñoz	25	Agustina	Muñoz	54
	Bolivia	Juan Carlos	Mamani	26	Martha	Mamani	55
		José Luis	Quispe	27	Roxana	Quispe	56
	Peru	José	Quispe	28	María	Quispe	57
		Luis	Flores	29	Rosa	Flores	58
Europe	Spain	Alejandro	García	14	María	García	43
		Pablo	Rodríguez	15	Lucía	Rodríguez	44
	Italy	Francesco	Rossi	16	Alessia	Rossi	45
		Andrea	Russo	17	Chiara	Russo	46
	Germany	Alexander	Müller	18	Marie	Müller	47
		Maximilian	Schmidt	19	Sophie	Schmidt	48
	France	Thomas	Martin	20	Léa	Martin	49
		Lucas	Bernard	21	Manon	Bernard	50
Asia	Japan	Hiroto	Sato	6	Hina	Sato	35
		Daiki	Takahashi	7	Sakura	Takahashi	36
	China	Tāo	Wáng	8	Tíng	Wáng	37
		Hào	Lǐ	9	Xīnyí	Lǐ	38
	South Korea	Min-jun	Kim	10	Seo-yeon	Kim	39
		Ji-hoon	Lee	11	Min-seo	Lee	40
	Taiwan	Chia-hao	Chen	12	Shu-fen	Chen	41
		Chih-ming	Lin	13	Shu-hui	Lin	42