Boys' Club: Cultural Inheritance and Social Networks and their Effects on Labor Market Outcomes

Nicole de Sousa Alves[†] Lorena Hakak[‡] Daniel A Feitosa Lopes[§]

August 2, 2022

Abstract

This study aims to identify the causal effect of cultural inheritance on formal labor market variables via the network effect. The identification strategy is based on an epidemiological approach to understanding cultural effects on economic variables that allows us to separate cultural aspects between groups that share the same institutions and economic environments. For this, a rich database is organized based on the identification of individuals of Jewish and non-Jewish Polish descent. The results of the income difference between these two groups are explained by the labor network effect. Furthermore, we find an important difference in the role of gender in the formation of these networks, such that the network effect is mainly produced by the network formed by men.

JEL codes: J01, J15, J62, D85.

Keywords: labor market, gender, networks, family ties, culture.

*The authors thank M. Theodora Barbosa for having made available data from the Memory Center of the Jewish Museum of Sao Paulo and Maurício Reis for having made available the pairings between job skills and occupations. The authors acknowledge the comments from Emerson Marçal and seminar participants at the 2nd Brazilian Meeting in Family and Gendar Economics (GeFam/2021). We also thank the Center for Regional and Urban Economics (NEREUS) at USP (http://www.usp.br/nereus/) for their aid and computational support.

[†]FEA - São Paulo University

[‡]UFABC: Federal University of ABC Email: lorena.hakak@ufabc.edu.br. (Corresponding author)

[§]SESI/CNI.

1 Introduction

New methodological approaches that provide more evidence of cultural influences on economic outcomes and the diversity among countries have been developed based on data from immigrants of different backgrounds and cultural traits. This is because the studies based on data from these individuals offer new empirical possibilities and a completely different perspective of observing the cultural impact on economic variables, as in Fernandez and Fogli (2009), Fernández (2011) e Figlio et al. (2016).

An example of a methodology that provides a new perspective on the cultural impact on economic outcomes based on immigrant data is the epidemiological approach. This was widely explored by Fernández (2011) and seeks to separate cultural aspects from the environmental aspects of an institutional and economic nature, based on the study of individuals who have different cultural backgrounds but share the same economic and institutional apparatus. Through a formal comparative analysis of the characteristics of these individuals, this approach allows us to disentangle the effects of culture or the institutional economic environment on variations in economic outcomes at different levels.

Brazil is a good candidate for this type of analysis, as the country received a massive flow of immigrants up to the 1960s, including the participation of immigrants from different ethnic, racial, and religious backgrounds. Among the flows that make up this immigration was the Jewish immigration that started in the middle of the nineteenth century, which was part of the Brazilian cultural formation. The flows of this migration occurred in different regions of the country and in different periods of time. (Decol (1999))

This ethnic-religious group is an important example of the persistence of religious and ethnic traits and the difficulties of cultural assimilation outside the original environment. For Herberg (1955), this group clearly has one of the three great faiths for which religious assimilation tends to fail, and Glazer and Daniel (1963), from a study of the main ethnic groups in New York, noticed that Jews tended to preserve their economic, political, and cultural patterns even long after immigration. Investment in human capital is an important asset for this group.[1]

The aim of this study is to analyze the causal effect of Jewish culture (via social networks) on formal labor market outcomes. In addition, the authors constructed a rich new database of identified socioeconomic data for individuals from 2009 to 2017 by pairing the surnames of

[1] Lehrer (2004) shows that Jews are more likely to invest in human capital because it is an asset more portable than investments in physical capital.

Jewish and non-Jewish Polish immigrants. Despite being from the same country, the surnames of individuals of Jewish origin are completely different from those of non-Jewish origin. This is a key point of our empirical strategy.[2]

The main contribution of this work is to show that social networks arising from cultural traits can affect labor market variables. We find that the difference in average income between Jewish and non-Jewish Poles, even when controlling for other variables, comes from the social network effect. The Jewish network seems to explain the difference between the two groups.

A second contribution of this work is to add to the literature on gender gaps and social networks. We find that the social network effect is greater for men than women; that is, on average, the income effect is higher for men. Moreover, when we consider network composition, it seems that the general effect we find in labor market outcomes is due to the social network being composed of men. These results are in line with those of the two different strains in the literature. Agarwal et al. (2016) and Cullen and Perez-Truglia (2019) find different effects in labor market outcomes when considering gender and social networks. It seems that the social networks formed from sports activities or smoking together during breaks have a significant effect on men's careers at the expense of women. There is a lower likelihood of women engaging in activities outside of working hours, which diminishes the network effect.

Cultural and social norms can benefit the formation of social networks between men and their entry and permanence in the labor market. In the case of Jews, the organization of the family in more traditional patterns and the greater interaction of its male members in religious activities may explain the different effects between genders that we find in the results. The network formed only by women has no significant effect on explaining the income difference between the two groups; however, the network formed by men explains the effects we find.[3]

The migratory flow in Brazil comprised several groups from different countries, including Italians, Germans, Japanese, and Iberians. However, the Brazilian migratory flow ceased from the 60s of the last century. Thus, in this study, we aimed to explore long-term cultural inheritance. Polish migratory waves had different characteristics. Between 1889 and 1914, peasants predominantly migrated.[4] This pattern changed in the following migration waves before and after the 2^{nd} World War. At that time, the destination was mainly the city of São

[2]There is a difference in the etymology of surnames between Jewish and non-Jewish Poles. Abramowicz and Dacewicz; Beider (1996); KOwALsKI and ŚLEsZyńsKI (2010); Walkowiak (2012)

[3]According to Jayachandran (2021), there is evidence that women can suffer some type of restriction on entering or remaining in the labor market, which decreases the likelihood of forming a social network as men do.

[4]Decol (2016).

Paulo.[5]

According to Bastos and Salles (2014), more than half of the immigrants in the post-war period were people with technical training and, according to Salles (2008), of the occupations performed by immigrants in their first job, more than half of the Poles were qualified and technical workers with mid- and high-level professionals. According to Decol (1999), Polish immigration was intense before and after the 2^{nd} World War, among both Jewish and non-Jewish Poles. Their density was highest in urban areas. Half of the Polish immigration that entered the country after the 2^{nd} World War remained in São Paulo. According to Bastos and do Rosário Rolfsen Salles (2012), this is related to the shortage of labor due to industrial expansion in the 1940s.[6] In this sense, we compare two groups that theoretically only diverge by religion, some with Jewish background and the others not, and we use data for their descendants that remain in São Paulo.

This study is divided as follows. The next section is a literature review concerning Jewish immigration and ancestry. The third section describes the database used. The fourth section presents the study's methodology. The fifth and sixth sections present the descriptive statistics and empirical strategy. Finally, we present the results and the main findings.

2 Literature review

2.1 The Jewish immigration

An important reference for understanding the key characteristics of Jewish immigration in Brazil is Decol (1999). The author initially characterizes the Brazilian Jewish immigration flow within the urban immigration flow, in contrast to the rural flows, which have in general been more explored in the literature. The Jewish migratory flows mainly came from large and small cities of Central Europe and the capitals of countries in North Africa and Asia. In general, the most common destinations for these flows were urban parts of the cities. Decol (1999) highlights that this immigration occurred more intensely from the 1920s onward and flowed into more urban areas because most of this group came from regions with advanced urban development.

[5] According to Decol (1999), in 1968, 30% of the Jews residing in the city of São Paulo declared Poland their country of birth.

[6] According to Decol (1999), Jewish Poles came in greater numbers before the 2^a World War, while non-Jewish Poles came in greater numbers after the war. This change was probably due to the Holocaust, in which 90% of Polish Jews were killed.

According to the author, international urban immigration flows brought in groups with different social and educational backgrounds. Those from urban areas had already been changing their fertility patterns, a cultural component relevant to Brazil's demographic history.

Regarding European Jewish immigrants, census data show that they arrived in Brazil at advanced stages of their fertility transition process. Decol (1999) understands that because most of them came from cities, European Jews had already experienced a period of intense structural changes in Europe. Their fertility pattern transition could be understood as part of their adaptation process to this new situation that, with time, would spread to the rest of the world.

Since we aim to study the influence of the Jewish cultural component on individuals' economic outcomes, it seemed appropriate to understand which determinants could explain the intergenerational persistence of Jewish traits. Chiswick (1999) constructs an economic model of ethnic-religious identity and discusses the benefits and costs of Jewish assimilation. This could be defined as the loss of Jewish cultural traits over time after contact with non-Jewish individuals.

In this model, religion and ethnicity are seen as single quasi-public goods self-produced at home and in the community, and the efficiency of its production can be increased by investing in group-specific human capital. In this arrangement, group-specific human capital plays a key role in determining the identity and productivity of ethnic-religious consumption. The groups in this model are differentiated according to the specificity of their human capital, the level of group-specific investment, and the degree of complementarity between group-specific and general human capital.[7]

The author explains that Judaism is a religion of intensive human capital via its own religious activity. It is an ethnic-religious group that can be properly represented by the model. Chiswick (1999) argues that the greater the intensity of specific Jewish human capital, the higher the rates of return for other forms of human capital and, therefore, there are long-term social rewards. Whether these rewards lead to assimilation depends on the characteristics of the whole society.

In this study, Jewish continuity is defined as the consequence of voluntary participation in Jewish life in order to ensure the intergenerational transmission of Jewish human capital. Individuals with high levels of Jewish human capital receive and transmit their capital more effectively, and those with little Jewish human capital have less to lose by leaving the community.

[7]Group-specific investment is related to the study of the Torah and general investment in formal education.

Finally, the model predicts that Jewish-specific human capital is a dominant factor in Jewish continuity. Furthermore, for a Jewish culture developed in a hospitable environment, assimilation is more likely to be chosen by the socioeconomically vulnerable and those with less access to education. However, as Jewish culture developed in a hostile environment, it is wealthy Jews with high levels of general education and culture who are more susceptible to cultural assimilation. According to the author, through a brief historical review, it is possible to verify examples of both extreme situations and some intermediate possibilities.

2.2 Ancestry

We ran an algorithm that allows us to pair the surnames of immigrant individuals with the surnames of the contemporary ones, in order to find people of the same ancestry. Several papers have used similar tools with similar objectives. Monasterio (2017) presents a method for classifying the ancestry of Brazilian surnames based on historical sources via matching methods using machine learning. The author applies these methods to more than 46 million individuals and was able to classify these individuals into different backgrounds: Iberian, Italian, Japanese, German, Polish, and Eastern European, reflecting the countries of immigrants who arrived in Brazil after 1872. In this study, contemporary surnames were obtained from individuals from the 2013 RAIS database.[8] The first two surnames of each individual were considered and the rest excluded; in addition, rare surnames that appeared a few times in the database were removed; insignificant parts and first names were also excluded. This procedure yielded approximately 530,000 unique surnames.

Finally, the work of Monasterio (2017) demonstrates the potential of using historical immigrant and contemporary databases with text pairing and machine learning techniques to classify individuals by the ancestry of their surnames, which is an important foundation for the work developed in the present study. A final comparison of the results with data from foreigners from the 1920 census and with the geographic distribution of non-Iberian surnames indicates the effectiveness of the procedure. The study demonstrates that surname-related ancestry is associated with significant differences in wages and education.

Lopes (2018b) applies the method developed by Monasterio (2017) to explain the role played by different cultural backgrounds in educational outcomes on standardized tests among Brazilian primary school students. Lopes (2018b) exploits the different cultural traits transmitted from parents to child as a complementary source of the persistence of educational

[8]The Annual List of Social Information (RAIS) is a national administrative list, mandatory for all establishments in the Brazilian formal labor market. More information is provided in Section 3.1.

outcomes. Given the interest in analyzing cultural differences, two main methods are used to isolate the effect of culture from those generated by other social and institutional influences.

First, the epidemiological approach was followed, as treated by authors, such as Fernandez and Fogli (2009) and Figlio et al. (2016), to study descendants belonging to generations later than the second Brazilian generation of immigrants analyzed in the work. It was possible to analyze children who had ancestors who migrated to Brazil in the same location, thus distinguishing between cultural factors and other institutional and economic factors. Another method used followed the line of work in Guiso et al. (2006) and Tabellini (2008), in which the historical episode of mass migration was explored to decouple the potential persistent effects of culture from those of local institutions, controlling for colonial factors and the presence of children from families that varied their geographic location recently during the period of work analysis. This was possible because the data from school censuses from which the children's surnames were obtained also contained data on their individual addresses.

Lopes (2018b) results suggest that there is an ancestry influence on educational performance. Using Iberian Brazilians as a control group, it was observed that students with non-Iberian European ancestry (Germans, Italians, Eastern Europeans) or Japanese achieved statistically significant higher academic results on the mathematics test of the National Literacy Assessment and on the Brazil test conducted for third and fifth graders, even after controlling for a broad set of individual, family, and municipal variables.

3 Database

3.1 Annual List of Social Information (RAIS)

RAIS was created in 1975 as a national administrative list with mandatory registration by all establishments in the Brazilian formal labor market. The purpose of the registration is to meet the needs for control, statistics, and information of government entities in the social area. It constitutes an essential instrument for the fulfillment of legal norms and is of fundamental importance for the monitoring and characterization of the formal labor market.

Declarations are provided annually, generally from January to March, and refer to the previous year. The information is provided according to the stock (number of jobs) and the flow of employed labor (admissions and dismissals) by gender, age group, level of education, average income, and income ranges in minimum salaries. Taking these cuts into consideration, data on wage mass were also obtained.

The variables that we obtained through RAIS and that were used in this work are described in the Appendix, in section A - Personal Characteristics Variables, B - Labor Market Variables, and C - Cultural Relation Variables.

We also structured two databases to achieve the goal of better understanding the influence of the cultural component of individuals of Jewish origin, on the variables of interest (income, education, and others). With this goal in mind, we constructed a database containing Polish non-Jewish and Jewish descent by surname for the years between 2009 and 2017.

3.2 Data of Descendants of Jewish Poles

The database with individuals of Polish ancestry and Jewish culture was structured from the list of immigrants entered by Santos Harbor and registered by the Israeli Beneficent Society "Erza" of São Paulo from Kremnitzer (1998), considering, however, only individuals of cultural traits identified with Judaism and Polish ancestry. These data were later processed to identify individuals in the RAIS database, which contains individual data on those working in the Brazilian formal labor market, such as income, education, and occupation.[9]

3.3 Data of Descendants of Non-Jewish Poles

The database of Polish ancestry and non-Jewish culture was structured in Lopes (2018a). The data were organized and made available by Lopes (2018a). Thus, we processed the data for later matching with the information contained in the RAIS database in a similar way to the previous section.[10]

4 Methodology

4.1 Search Algorithm for Surname Descendants

The database was constructed using Stata software codes. We only consider the last two surnames of each individual listed in the RAIS. For the database of individual non-Jewish

[9]We only considered the surnames of these immigrants with capital letter formatting. We excluded all surnames with fewer than four characters in order to avoid pairings that were not relevant to the search. Finally, we created an index variable for each last name to help with the *Matchit* function. These changes were then saved.

[10]We only considered the surnames of these immigrants with capital letter formatting. We excluded all surnames with fewer than four characters in order to avoid pairings that were not relevant to the search. Finally, we created an index variable for each last name to guide the action of the Matchit function. These changes were then saved.

Poles originating from RAIS, we excluded all pairings obtained with a similarity degree lower than 0.8. We excluded the surnames Adelson, Miller, and Salomão, as they generated pairs with few pairings.

Similarly, to match the data of Polish Jews with RAIS, we ended by excluding all matches obtained with a degree of similarity lower than 0.8, thereby excluding the following surnames: Martins, Ferreira, Agostinho, Valente, Agostinhac, and Frederico.

The next step, after fully processing the two surname databases, was to access the RAIS databases for the years 2009 to 2017. We restricted our analysis to the data from the city of São Paulo, the capital of the State of São Paulo. We considered the last two surnames of each individual, which in Brazil, in general, are inherited from the mother and father, respectively. We disregard all surnames with fewer than three characters.

After having the two surname lists properly processed and the RAIS surname list assembled for each year, we accessed the list of first and last names originally obtained from both Jewish and non-Jewish Poles and performed the final pairing.[11] We ended with a panel of individuals who are, with high probability, descendants of Jewish and non-Jewish Polish immigrants.

After creating the panel data, we performed final manipulations on the base to create the variables to be used in the next sections of our empirical models. Details of these variables are described in the Appendix, in the sections B Labor Market Variables and C Cultural Relation Variables.

Finally, we grouped the firms identified by their number in the Brazilian Internal Revenue Service (CNPJ, Portuguese acronyms) from the panel data to create identifiers capable of clustering the data based on similar characteristics of these companies. This information is used in the empirical models.

5 Descriptive Statistics

The panel has 52,667 observations, of which 37.74% (19,876) are related to individuals descending from non-Jewish Poles' surnames and 62.26% (32,791) to individuals descending from Jewish Poles' surnames.

The distribution by gender of the data panel is well balanced for all datasets: 47.28% male and 52.72% female. In 2009, this proportion was 47.20% for males and 52.80% for females, and in 2017, 46.41% (male) and 53.59% (female).

[11] The pairing is explained in the section 4.1.

Considering the distribution of gender and culture, we found that in the whole database, 9,146 (17.36%) were male non-Jewish Poles and 10,724 (20.36%) were female. Of these, 15,754 (29.91%) were male Jewish Poles and 17,043 (32.35%) were female. In 2009, these proportions were non-Jewish Poles, 951 (17.05%) men and 1,131 (20.28%) women, and Jewish Poles, 1,683 (30.18%) men and 1,811 (32.47%) women. In 2017, non-Jewish Poles included 973 (17.41%) men and 1,141 (20.42%) women, and Jewish Poles 1,620 (28.99%) men and 1,853 (33.16%) women.

Table 1 presents the distribution of individuals in the panel according to their cultural inheritance. In every year of the panel we have more individuals descending from Jewish Poles than from non-Jewish Poles. Table 1 shows that our panel is unbalanced since the number of individuals monitored varies from year to year (given the movement of workers entering and leaving the formal labor market).

[insert Table (1) here]

Table 2 shows statistical averages for age, nominal income, and the percentage of individuals with management positions in the data panel.

[insert Table (2) here]

Table 3 shows the average number of individuals with Polish Jewish surname ancestry and Table 4 contains the statistics of individuals with Polish non-Jewish surname ancestry by firm. In both cases, the average number of individuals per company is around two and the standard deviation is high.

[insert Table (3) here]
[insert Table (4) here]

6 Empirical Strategy

According to Levy (1974), 6.2% of the Brazilian population in 1900 was made up of foreigners. In the early 1930s, the federal government started to adopt measures that restricted immigration to reduce these international migratory flows to the country. In 1946, the new Brazilian constitution reduced these restrictions, and migratory flows began to grow again. After 1964, with the military government, Brazil again faced a significant decline in this migratory flow. In 2010, the number of foreigners in Brazil decreased to only 0.3% of the

population, as shown in Lopes (2018a). Based on this background, we explore the transmission of cultural traits between immigrants and their descendants and whether these aspects persist over time.

Migratory flows ceased in the 60s of the last century. Our data primarily contain the second-generation and subsequent generations of immigrants. The individuals in the sample do not face problems such as learning a new language or any other aspects that can affect the first generation. We explore the cultural persistence that can be assumed to be predetermined over time for two different groups: descendants of Jewish Poles and non-Jewish Poles. Thus, we investigate whether the cultural differences between these groups have an effect on their labor market outcomes in the formal market.

We estimate the empirical model using an epidemiological approach Fernández (2011). This approach allows us to disentangle the effect of culture from the original economic and institutional environment, given that both groups are from the same country and region, in addition to having migrated to Brazil, specifically in the city of São Paulo, in close periods.

The main variable of interest is the logarithm of income. The Mincer equation we estimate is as follows:.

$$LnIncome_{icjt} = \theta_j + \theta_t + \ \tau_c Dculture_{ic} + \beta_c Network_{icjt} + \beta_3 X_{ijt} + \varepsilon_{ijt} \tag{1}$$

where i denotes the individual, c is the individual's culture (1 if Polish Jew and 0 if Polish non-Jew), j is the firm, and t is the year from 2009 to 2017. $LnIncome$ represents the logarithm of the income. θ_j is the firm fixed effect, θ_t is the year fixed effect, $Dculture_{ic}$ is 1 if the individual is of Polish Jewish descent c and 0 otherwise, $Network_{icjt}$ is 1 if there are more than two people of the same culture c in firm j and 0 otherwise. X_{ijt} is a vector of controls, such as education level, race, sex, age, first job, managerial position, public or private institution position, and period of employment in the current job (logarithm); all variables are at the individual level.

We constructed the network variable as follows: We grouped the panel by firm and year, and we added the individuals by cultural background, thus obtaining the number of individuals by cultural background, firm, and year. The network variable captures the extent to which an individual works in an organization that contains other individuals with this cultural component. The dummy network has the value 1 when the individual works in a company with two or more individuals with the same cultural background, and 0 otherwise.

It is important to emphasize that cultural traits can be passed on from parents to chil-

dren, but they can coexist through other links within the community itself. Therefore, it is important to understand and control for the role played by networks in the labor market results. The results are presented in Tables 5 and 6.

In addition, we estimate the following regression:

$$Y_{icjt} = dummy_t + \tau_c Dculture_{ic} + \beta_c Network_{icjt} + \beta_3 X_{ijt} + \varphi_{ijt} \quad (2)$$

The results are presented in Tables 8, 7, and 11, which were estimated using a probit model with three different dependent variables: probability of an individual having a managerial position, probability of first job compensation of an individual being higher than the average of all individuals in their first job, and the probability of income being above the 90th percentile. We used all the controls available and year dummies. The manager position variable enters as a control in the regressions with the dependent variables wage at first job and wage at the 90th percentile.

To the best of our knowledge, this type of investigation has not been applied to Brazilian data, given the complexity of concatenating all the datasets, as described in Section 3. Therefore, rigorous econometric analysis of this unique and rich database can provide important insights into how cultural traits can help explain labor market outcomes.

7 Results

We present our main results in Table 5. The culture variable has a positive effect of 8.3% (column (1)) and is statistically significant. This means that, on average, people with Polish and Jewish ancestry earn 8.3% higher than those of non-Jewish Polish ancestry. When we include the skills variables, this effect drops to 7.5%. In column (3), we add the Jewish network variable, which explains why descendants of Jewish Poles earn more, on average, than the control group. When we include this variable, culture loses statistical significance. The Jewish network has a positive and statistically significant effect of 5.4% on average income. All three columns include firm and year fixed effects, as well as the other controls.

[insert Table (5) here]

We run the regression (1) considering the network variable formed by men and women separately in order to understand whether the effect of the network variable comes from a specific gender. In Table 6, we observe in columns (1) and (2) the positive effect of the

Jewish network composed of men and women on average income with and without controls. In column (3), we consider the effect of the Jewish network, considering that this network is composed only of women. The effect was slightly negative and not statistically significant. In this case, the culture variable remained positive and statistically significant.

Furthermore, when we look at column (4), the Jewish network variable considering only men has a positive and statistically significant effect of 6.7%, and the effect is greater than that of the Jewish network formed by men and women (5.4%), however in the former the Culture variable is positive and statistically significant. The results of columns (3) and (4) present some evidence that the network effect comes from the male side. This result is in line with that found in Agarwal et al. (2016), where the authors argue that women can be penalized for not being part of a labor network like men. In that study, the network was formed through sports by playing golf. Cullen and Perez-Truglia (2019) find an effect on promotions where it seems the channel comes from male employees and managers who smoke together.

[insert Table (6) here]

One possible situation in which networks and backgrounds can play an important role is in the first job. Thus, we estimate a regression with the individuals declared to be in their first job against the variables of culture, Jewish network, firm and year fixed effects, and the other controls. In Table 8, we find a positive and statistically significant effect in which the Jewish network variable explains the probability of the Jewish group being above the top 50% in relation to the average income in the entry salary. This group is 9% more likely to be in the top 50 than the control group.

[insert Table (8) here]

In Table 7, we find that the probability of being above average in relation to the average income in the 90th percentile due to Jewish network is positive and 4.5% greater in the Jewish group than the control group. This result may be related to the increase in the probability of having a higher-than-average income in the first job. According to Brunner and Kuhn (2014), worse conditions for entering the labor market, such as a lower first wage, may explain a trajectory of lower earnings throughout life.

[insert Table (7) here]

7.1 Robustness analysis

Heterogeneity Tests: In Table 9 we run the same regressions as in Table 6 but considering the database composed only by women. The Jewish network variable composed only of women has a positive effect, but it is not statistically significant. The Jewish network variable composed of men and women has a positive effect (5%) on women's income and has a statistically significant effect at 10%. In both cases, the Jewish culture variable loses significance. This corroborates the findings of the previous section, in which the Jewish network explains the higher average income among the descendants of this group, but the results are heterogeneous between genders. In Table 10, we run the regressions by considering a database composed only of men. The variables of Jewish network (only composed of men) and Jewish network (composed of men and women) are positive by the same amount of 6.4% (significant at 5%) in columns (3) and (4). It seems that the positive effect from the network we find in our work comes from men's side, and the effect on income is higher for men than for women, 6.4% and 5%, respectively.

Furthermore, comparing columns (3) of Tables 9 and 10, the result is only positive and significant for the man-to-man network. These results outline that the network effect works more among man-to-man than woman-to-woman channel. These results are in line with those of Agarwal et al. (2016), where they find a heterogeneous effect between genders, showing that the network effect can reduce women's chances of serving on the board of companies. Cullen and Perez-Truglia (2019) find that "the male-to-male advantage explains a third of the gender gap in promotions at this firm."

[insert Table (9) here]
[insert Table (10) here]

Leadership Position: We find an increase in the probability of 3.5% of the individual having a management position, given that they have a culture equal to 1 (Jewish culture). This effect is positive and statistically significant at the 1% level, as presented in Table 11 column (1). The Jewish network variable increases by 1% the individual's probability of having a head position (column (2)). When we include both variables, what prevails is the culture variable. The network variable has a negative value. In this case, it seems that the culture variable has a more important effect than the network.

[insert Table (11) here]

Placebo Test: We randomize the culture variable, and the effect on income is not statistically significant, as shown in Table 12. We did the same with the Jewish network variable, which also showed a statistically non-significant effect.

[insert Table (12) here]

8 Final Remarks

This study aimed to analyze whether persistent cultural effects can have an impact on labor market outcomes. We found a positive difference in average income between the two groups of Polish descendants, Jews and non-Jews, explained by the culture variable. However, this difference disappears when we include network variables. It seems that the average difference between these groups is explained by the Jewish network effect, which is positive and statistically significant. Another labor outcome analyzed was entry salary. We find that the probability of being above the top 50% in relation to the average income in the entry salary is 9% higher in the Jewish than the control group.

Furthermore, when we analyzed the effects of social networks by gender on average income, it seems that the positive effect of the network comes from the male side. The effect of the female network variable in the sample composed only of women was not statistically significant. Considering the male network variable, there is a positive effect of 6.7% on average income in the sample with only men. In addition, when we consider the social network effect on each sample by gender, the results are greater for the men's sample (6.4%) than for the women's sample (5%) (Tables 10 and 9 respectively). These results corroborate Agarwal et al. (2016) and Cullen and Perez-Truglia (2019), who find robust results that the social network is greater among men and benefits them compared to women.

This work contributes to the literature by measuring the persistent cultural effects on labor market outcomes, in addition to pointing out that these effects mostly come from the effects of the social network formed within the group. Moreover, the effects are heterogeneous between men and women.

References

Abramowicz, Z. and L. Dacewicz, 'The development of Jewish anthroponymic system in north-eastern Poland–the historical perspective'.

Agarwal, S., W. Qian, D. M. Reeb, and T. F. Sing: 2016, 'Playing the boys game: Golf buddies and board diversity'. *American Economic Review* **106**(5), 272–76.

Bastos, S. and M. d. R. R. Salles: 2014, 'A imigração polonesa para São Paulo no pós-Segunda Guerra Mundial no quadro das entradas dos" deslocados de guerra": 1947 a 1951'. *Revista Brasileira de Estudos de População* **31**(1), 151–167.

Bastos, S. R. and M. do Rosário Rolfsen Salles: 2012, 'Polish postwar migration to Brazil, 1945-55'. *Polish American Studies* **69**(1), 55–73.

Beider, A.: 1996, *A dictionary of Jewish surnames from the Kingdom of Poland*. Avotaynu Teaneck, NJ.

Brunner, B. and A. Kuhn: 2014, 'The impact of labor market entry conditions on initial job assignment and wages'. *Journal of Population Economics* **27**(3), 705–738.

Chiswick, C. U.: 1999, 'The economics of Jewish continuity'. *Contemporary Jewry* **20**(1), 30–56.

Cullen, Z. B. and R. Perez-Truglia: 2019, 'The old boys' club: Schmoozing and the gender gap'. Technical report, National Bureau of Economic Research.

Decol, R. D.: 1999, 'Imigrações urbanas para o Brasil: o caso dos judeus'.

Decol, R. D.: 2016, 'Uma história oculta: a imigração dos países da Europa do Centro-Leste para o Brasil'. *Anais* pp. 1–12.

Fernández, R.: 2011, 'Does culture matter?'. In: *Handbook of social economics*, Vol. 1. Elsevier, pp. 481–510.

Fernandez, R. and A. Fogli: 2009, 'Culture: An empirical investigation of beliefs, work, and fertility'. *American economic journal: Macroeconomics* **1**(1), 146–77.

Figlio, D., P. Giuliano, U. Özek, and P. Sapienza: 2016, 'Long-term orientation and educational performance'. Technical report, National Bureau of Economic Research.

Glazer, N. and P. Daniel: 1963, 'Beyond the melting pot: the Negroes, Puerto Ricans, Jews, Italians and Irish of New York City'.

Guiso, L., P. Sapienza, and L. Zingales: 2006, 'Does culture affect economic outcomes?'. *Journal of Economic perspectives* **20**(2), 23–48.

Herberg, W.: 1955, 'Catholic, Protestant, Jew. An Essay in American Religious Sociology'.

Jayachandran, S.: 2021, 'Social norms as a barrier to women's employment in developing countries'. *IMF Economic Review* **69**(3), 576–595.

KOwALsKI, M. and P. ŚLEsZyńsKI: 2010, 'The migration of Poles to Germany in the context of the most common Polish surnames'. *Przegląd Zachodni* **2**, 119–135.

Kremnitzer, A.: 1998, *Lista de Imigrantes Entrados pelo Porto de Santos entre 1928 e 1932 e registrados pela "Ezra" de São Paulo*. Arquivo Histórico Judaico Brasileiro.

Lehrer, E. L.: 2004, 'Religion as a determinant of economic and demographic behavior in the United States'. *Population and development review* **30**(4), 707–726.

Levy, M. S. F.: 1974, 'O papel da migração internacional na evolução da população brasileira (1872 a 1972)'. *Revista de Saúde Pública* **8**, 49 – 90.

Lopes, Daniel e Monasterio, L.: 2018a, 'Brasil sem imigrantes: estimativas de longo prazo baseadas em microdados'. Mpra paper, University Library of Munich, Germany.

Lopes, Daniel, S. F. G. M. L.: 2018b, 'Culture, Institutions and School Achievement in Brazil'.

Monasterio, L.: 2017, 'Surnames and ancestry in Brazil'. *PloS one* **12**(5), e0176890.

Reis, M.: 2018, 'Measuring the mismatch between field of study and occupation using a task-based approach'. *Journal for labour market research* **52**(1), 1–15.

Salles, M. d. R. R.: 2008, 'Território e experiência imigratória: os refugiados em São Paulo no pós-Segunda Guerra Mundial'. *Cadernos Metrópole* (20), 179–196.

SEBRAE: 2006, 'Nota metodológica para definição dos números básicos de MPE'. *Brasília, DF: SEBRAE*.

Tabellini, G.: 2008, 'Institutions and culture: presidential address'. *Journal of the European Economic Association* **6**(2-3), 255–94.

Walkowiak, J. B.: 2012, 'Feminine Surnames in Polish: Two Policies and the Practice'. *Onoma: Journal of the International Council of Onomastic Sciences* (47), 223–246.

Appendix

Tables and Figures

Table 1: Cultural Distribution by Year

	Individuals with Surname Ancestry	
Year	Polish Jewish	Polish non-Jewish
2009	2,008	3,488
2010	2,136	3,590
2011	2,188	3,718
2012	2,312	3,785
2013	2,322	3,787
2014	2,334	3,762
2015	2,232	3,631
2016	2,152	3,555
2017	2,112	3,475

Table 2: Descriptive Statistics: Age, Average Nominal Income, and Percentage in Management Positions

Variable	2009-2017	2009	2017
Age (years)	37.48	36.41	38.54
Average Nominal Income (Reais)	5,458	3,849	7,127
Management Position (%)	12.88	10.40	15.03

Table 3: Descriptive Statistics: Individuals with Surnames Indicating Jewish Polish Ancestry by Company

	Average	Standard Deviation	Minimum	Maximum
Individuals by company	2.10	9.06	0	466

Table 4: Descriptive Statistics: Individuals with Surnames Indicating Polish Non-Jewish Ancestry by Company

	Average	Standard Deviation	Minimum	Maximum
Individuals by Company	1.18	11.91	0	1,262

Table 5: Results: Effect on Average Income

	(1) Average Income	(2) Average Income	(3) Average Income	(4) Average Income
Culture (=1, Jewish)	0.083***	0.075***	0.034	
	(0.016)	(0.016)	(0.022)	
Skill 2		0.200***	0.201***	
		(0.040)	(0.040)	
Skill 1		0.004	0.004	
		(0.095)	(0.095)	
Skill 3		-0.189***	-0.189***	
		(0.035)	(0.035)	
Jewish Network			0.054**	0.087***
			(0.022)	(0.016)
Observations	32882	32882	32882	32882
F-stat	25.37	20.97	2.383	26.94
Average Income	8.050	8.050	8.050	8.050

Standard errors in parentheses

In column (1), we consider the culture of individuals according to the Culture dummy variable (defined in Appendix A.1). In column (2), we add the skills dummy variables, where Skills 1 corresponds to the group of cognitive non-routine skills, Skills 2 corresponds to the group of routine skills, and Skills 3 corresponds to the group of non-routine manual skills (groups defined in Appendix B.5); in column (3), we add the variable Jewish Network defined in Appendix C.1, and column (4) we only test the effect of the Jewish Network on Average Income. Average Income is the logarithm of the Average Nominal Income of individuals, as defined in Appendix B.1. The equations include year and company fixed effects. The control variables omitted from the equation are: Job Change (Appendix B.8), Number of contractual hours (RAIS), Ln of period of employment (RAIS), Education (Appendix A.3), Age (RAIS), Gender (Appendix A.4), First Job (Appendix B.2), Private (Appendix B.3), Leadership position (Appendix B.4), Year variables, and Constant.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Jewish Network Heterogeneity Test: Effect on Average Income

	(1)	(2)	(3)	(4)
	Average income	Average income	Average income	Average income
Culture (=1, Jewish)	0.034		0.079***	0.058***
	(0.022)		(0.021)	(0.016)
Skill 2	0.201***		0.199***	0.200***
	(0.040)		(0.040)	(0.040)
Skill 1	0.004		0.004	0.002
	(0.095)		(0.095)	(0.094)
Skill 3	-0.189***		-0.189***	-0.188***
	(0.035)		(0.035)	(0.035)
Jewish Network	0.054**	0.087***		
	(0.022)	(0.016)		
Jewish Female Network			-0.010	
			(0.026)	
Jewish Male Network				0.067**
				(0.029)
Observations	32882	32882	32882	32882
F-stat	2.383	26.94	13.86	12.54
Average income	8.050	8.050	8.050	8.050

Standard errors in parentheses

In column (1), we consider the individuals' culture according to the Culture dummy variable (defined in Appendix A.1) and add the Jewish Network defined in Appendix C.1 and the skill dummies, where Skills 1 corresponds to the group of non-routine cognitive skills, Skills 2 corresponds to the group of routine skills, and Skills 3 corresponds to the group of non-routine manual skills (groups defined in Appendix B.5); in column (2), we tested only the effect of Jewish Network on Average Income; in column (3), we tested the effect of the Skills dummy variable, Culture, and Jewish Female Network, which considers Jewish women working in companies with other Jewish women; and in column (4) we test the effect of Culture, skill dummies, and Jewish Male Network, which considers Jewish men working in companies with other Jewish men. Average Income is the Ln of the Average Nominal Income of individuals, as defined in Appendix B.1. The equations include the fixed effects of years and company. The control variables omitted from the equation are: Job Change (Appendix B.8), Number of contractual hours (RAIS), Ln of Employment Time (RAIS), Education (Appendix A.3), Age (RAIS), Gender (Appendix A.4), First Job (Appendix B.2), Private (Appendix B.3), Leadership position (Appendix B.4), Year variables, and Constant. .

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 7: Jewish Network Heterogeneity Test: Effect on Average Income

	(1)	(2)	(3)	(4)
	Average income	Average income	Average income	Average income
Culture (=1, Jewish)	0.034		0.079***	0.058***
	(0.022)		(0.021)	(0.016)
Skill 2	0.201***		0.199***	0.200***
	(0.040)		(0.040)	(0.040)
Skill 1	0.004		0.004	0.002
	(0.095)		(0.095)	(0.094)
Skill 3	-0.189***		-0.189***	-0.188***
	(0.035)		(0.035)	(0.035)
Jewish Network	0.054**	0.087***		
	(0.022)	(0.016)		
Jewish Female Network			-0.010	
			(0.026)	
Jewish Male Network				0.067**
				(0.029)
Observations	32882	32882	32882	32882
F-stat	2.383	26.94	13.86	12.54
Average income	8.050	8.050	8.050	8.050

Standard errors in parentheses

In column (1), we consider the individuals' culture according to the Culture dummy variable (defined in Appendix A.1) and add the Jewish Network defined in Appendix C.1 and the skill dummies, where Skills 1 corresponds to the group of non-routine cognitive skills, Skills 2 corresponds to the group of routine skills, and Skills 3 corresponds to the group of non-routine manual skills (groups defined in Appendix B.5); in column (2), we tested only the effect of Jewish Network on Average Income; in column (3), we tested the effect of the Skills dummy variable, Culture, and Jewish Female Network, which considers Jewish women working in companies with other Jewish women; and in column (4) we test the effect of Culture, skill dummies, and Jewish Male Network, which considers Jewish men working in companies with other Jewish men. Average Income is the Ln of the Average Nominal Income of individuals, as defined in Appendix B.1. The equations include the fixed effects of years and company. The control variables omitted from the equation are: Job Change (Appendix B.8), Number of contractual hours (RAIS), Ln of Employment Time (RAIS), Education (Appendix A.3), Age (RAIS), Gender (Appendix A.4), First Job (Appendix B.2), Private (Appendix B.3), Leadership position (Appendix B.4), Year variables, and Constant. .

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 11: Effect of Culture on the Probability of Being in a Leadership Position

	(1)	(2)	(3)
	Leadership Probability	Leadership Probability	Leadership Probability
Culture (=1, Jewish)	0.035***		0.041***
	(0.003)		(0.004)
Jewish Network		0.007*	-0.012***
		(0.003)	(0.004)
Observations	33351	33351	33351

Standard errors in parentheses

We consider the Leadership Probability, the probability of the individual being in a leadership position, as defined in Appendix B.4, and the Culture dummy variable defined in Appendix A.1. The probabilistic relationship was estimated using a probit model with year dummies. The table presents the marginal values of the variables. In column (1), we assess the culture effect on the probability of being in a leadership position; in column (2), the Jewish Network effect on the probability of being in a leadership position; and in column (3) the culture effect and the Jewish Network on the probability of being a manager. Control variables omitted from the equation are: Change of Employment (Appendix B.8), Number of contractual hours (RAIS), Ln of Employment Time (RAIS), Education (Appendix A.3), Age (RAIS), Gender (Appendix A.4), First Job (Appendix B.2), Private (Appendix B.3), Year Variables, and Constant.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 8: Effect of Culture and Jewish Network on the Probability of Being above Average on Average Income in Entry Wage

	(1) Top 50% First Job	(2) Top 50% First Job	(3) Top 50% First Job
Culture (=1, Jewish)	0.034		-0.003
	(0.025)		(0.029)
Jewish Network		0.091***	0.093***
		(0.027)	(0.031)
Observations	875	875	875

Standard errors in parentheses

In this table, we consider only the observations of individuals registered in their first job. We obtained the probabilistic result of being above average on Average Income in the entry salary from the variable 10% richer in the first job as defined in Appendix B.11, the Culture dummy variable as defined in Appendix A.1, and Jewish Network as defined in Appendix C.1. The Table presents the marginal values of the variables. In column (1), we evaluate the Culture effect on the probability of the individual being above average on Average Income in the first job; in column (2), we evaluate the Jewish Network effect on the probability of the individual being above average on Average Income in the first job; and in column (3) we evaluated the Jewish Network effect and Culture effect on the probability of the individual being above average on Average Income in the first job. The probabilistic relationship was estimated using a probit model with year dummies. The control variables omitted from the equation are: Number of contractual hours (RAIS), Education (Appendix A.3), Age (RAIS), Gender (Appendix A.4), Private (Appendix B.3), Leadership position (Appendix B.4), Year Variables, and Constant.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 9: Base Heterogeneity Test: Effect on Average Income (Sample with Women Only)

	(1) Average Income	(2) Average Income	(3) Average Income	(4) Average Income	(5) Average Income
Culture (=1, Jewish)	0.064***	0.058**	0.037	0.016	
	(0.022)	(0.022)	(0.031)	(0.033)	
Skill 2		0.212***	0.212***	0.213***	
		(0.075)	(0.075)	(0.074)	
Skill 1		-0.070	-0.070	-0.070	
		(0.082)	(0.082)	(0.082)	
Skill 3		-0.187***	-0.187***	-0.187***	
		(0.049)	(0.049)	(0.049)	
Female Jewish Network			0.027		
			(0.031)		
Jewish Network				0.050*	0.068***
				(0.030)	(0.021)
Observations	17561	17561	17561	17561	17561
F-stat	8.324	6.461	1.436	0.231	10.65
Average Income	7.982	7.982	7.982	7.982	7.982

Standard errors in parentheses

In this table, we consider only observations corresponding to female gender. In column (1), we consider the culture of individuals according to the dummy variable Culture (defined in Appendix A.1). In column (2), we add the skills dummy variables, where Skills 1 corresponds to the group of cognitive non-routine skills, Skills 2 corresponds to the group of routine skills, and Skills 3 corresponds to the group of non-routine manual skills (groups defined in Appendix B.5); in column (3), we add the Female Jewish Network variable (similar to the Jewish Network variable defined in Appendix C.1), which considers only Jewish women who work in companies with other Jewish women; in column (4), we evaluate the effect on the Average Income of the Jewish Network combined with the variables of culture and skills; and in column (5) we evaluated the isolated effect of the Jewish Network. Average Income corresponds to the Ln of the Average Nominal Income of individuals, as defined in Appendix B.1. The equations include year and company fixed effects. The control variables omitted from the equation are: Change of Employment (Appendix B.8), Number of contractual hours (RAIS), Ln of Employment Time (RAIS), Education (Appendix A.3), Age (RAIS), First Job (Appendix B.2), Private (Appendix B.3), Leadership position (Appendix B.4), Year Variables, and Constant.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 10: Base Heterogeneity Test: Effect on Average Income (Sample with Males Only)

	(1) Average Income	(2) Average Income	(3) Average Income	(4) Average Income	(5) Average Income
Culture (=1, Jewish)	0.115***	0.101***	0.059*	0.053	
	(0.028)	(0.026)	(0.031)	(0.035)	
Skill 2		0.196***	0.199***	0.198***	
		(0.052)	(0.052)	(0.052)	
Skill 1		0.093	0.092	0.092	
		(0.149)	(0.149)	(0.149)	
Skill 3		-0.212***	-0.213***	-0.213***	
		(0.044)	(0.044)	(0.044)	
Male Jewish Network			0.064**		
			(0.030)		
Jewish Network				0.064**	0.113***
				(0.031)	(0.026)
Observations	15324	15324	15324	15324	15324
F-stat	16.94	14.29	3.554	2.289	18.48
Average Income	8.130	8.130	8.130	8.130	8.130

Standard errors in parentheses

In this table, we consider only observations corresponding to male gender. In column (1), we consider the culture of individuals according to the dummy variable Culture (defined in Appendix A.1). In column (2), we add the skills dummy variables, where Skills 1 corresponds to the group of cognitive non-routine skills, Skills 2 corresponds to the group of routine skills, and Skills 3 corresponds to the group of non-routine manual skills (groups defined in Appendix B.5); in column (3), we add the Male Jewish Network variable (similar to the Jewish Network variable defined in Appendix C.1), which considers Jewish men who work in companies with other Jewish men; in column (4) we evaluate the effect on the Average Income of the Jewish Network combined with the variables of culture and skills; and in column (5) we evaluate the isolated effect of the Jewish Network. Average Income is the logarithm of the Average Nominal Income of individuals, as defined in Appendix B.1. The equations include year and company fixed effects. The control variables omitted from the equation are: Change of Employment (Appendix B.8), Number of Contractual Hours (RAIS), Ln of Employment Time (RAIS), Education (Appendix A.3), Age (RAIS), First Job (Appendix B.2), Private (Appendix B.3), Leadership position (Appendix B.4), Year Variables, and Constant. .

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 12: Placebo Test: Effect on Average Income

	(1) Average Income	(2) Average Income	(3) Average Income
Randomized Culture	0.009	0.008	0.001
	(0.006)	(0.006)	(0.008)
Skill 2		0.204***	0.204***
		(0.040)	(0.040)
Skill 1		0.004	0.004
		(0.097)	(0.097)
Skill 3		-0.196***	-0.196***
		(0.035)	(0.035)
Randomized Jewish Network			0.013
			(0.012)
Observations	32882	32882	32882
F-stat	2.203	1.857	0.0362
Average Income	8.050	8.050	8.050

Standard errors in parentheses

In column (1), we enter as Randomized Culture the variable defined in Appendix A.2. In column (2), we add the skills dummy variables, Skills 1 corresponding to the group of cognitive non-routine skills, Skills 2 corresponding to the group of routine skills, and Skills 3 corresponding to the group of non-routine manual skills (groups defined in Appendix B.5), and in column (3), we add the Randomized Jewish Network variable defined in Appendix C.2. Average Income is the logarithm of the Average Nominal Income of individuals, as defined in Appendix B.1. The equations include year and company fixed effects. Control variables omitted from the table are: Change of Employment (Appendix B.8), Number of contractual hours (RAIS), Log of Employment Time (RAIS), Education (Appendix A.3), Age (RAIS), Gender (Appendix A.4), First Job (Appendix B.2), Private (Appendix B.3), Leadership position (Appendix B.4), Year variables, and Constant.

.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Variables

In this section, we describe in more detail the variables used throughout this work.

A Variables - Individual Characteristics

A.1 Culture

In the search algorithm for individuals with similar surnames to Jewish and non-Jewish Poles, the pairing takes place sequentially so that we initially search for a cultural group and then for the other; thus, at the end of the pairings between databases (initials: Polish Jews and non-Jews) with RAIS database, we created the variable Culture to indicate the cultural aspect of the database treated. It is, therefore, a dummy variable assigned 0 for the group of individuals descended from Polish non-Jewish surnames and 1 for the group of individuals descended from Polish Jewish surnames.

A.2 Random Culture

Random dummy variable created for the placebo test. Randomization based on the assignment of a uniform random distribution ranging between 0 and 1 for all observations; observations with assigned value greater than 0.5 receive the value of 1 for the dummy variable Randomized Culture.

A.3 Education

The education levels of each of the surname descendants obtained through the matching algorithm originate from the variable level of education or education of the RAIS database. The variable educ1 covers illiterate individuals or those who have up to the 5th year incomplete or complete in elementary school; educ2 refers to individuals who have incomplete 6th to 9th grade of elementary school or have completed elementary school; educ3 refers to people with complete or incomplete secondary education; educ4 covers individuals with complete or incomplete higher education; and Graduate refers to individuals with a complete master's or doctoral degree.

A.4 Gender

A dummy variable obtained through the Sex variable of the RAIS database, equal to 1 for female individuals and 0 for male individuals.

B Variables - Labor Market

B.1 Average Nominal Income

The Average Nominal Income variable was obtained from the income information contained in the RAIS database. Considering the existence of individuals with more than one active relationship in the formal market per year registered in the RAIS database, we only considered the formal activity with the highest average annual nominal income obtained by removing all CPFs repeated per year in the database, keeping the corresponding observation with the highest paying activity. Finally, we apply the Ln transformation to the average annual nominal income obtained per individual and per year.

B.2 First Job

Variable obtained from the RAIS database information. This is a dummy equal to 1 for all individuals with the Type of admission variable corresponding to 1 - Admission of an employee in the first job or appointment of a permanent or commission civil servant in the first job. All other possibilities receive 0.

B.3 Private Sector Worker

Variable obtained from RAIS information. This is a dummy variable that marks 0 for all cases in which the information contained in the variable Type of employment in RAIS database indicates a public nature in the employment relationship. These cases are: Employee governed by the Single Legal Regime (federal, state, and municipal) and military, linked to the Own Social Security Regime (30), Servant governed by the Single Legal regime (federal, state, and municipal) and military, linked to the General Social Security Regime (31), and Non-effective civil servant (dismissible ad nutum or admitted by special legislation, not governed by the CLT) (35). For all other types of relationship, the private variable receives the value 1, indicating the private sector nature of the established employment relationship.

B.4 Leadership Position

Variable obtained from RAIS database information. We analyzed the information contained in the CBO Occupation 2002 variable of the RAIS database that listed all the codes associated with leadership positions (Management and Board); equal to 1 for all individuals who had these codes listed; otherwise 0.

B.5 Skill Groups: Cognitive non-routines, manual routines and non-routine manual skills

Variables were obtained based on the work of Reis (2018). We built a database that relates the occupations of formal market workers (based on CBO 2010) with three groups of activities outlined by Reis (2018): i) non-routine cognitive skills, ii) routine skills, and iii) non-routine manual skills. The type of task framed in the first group, cognitive non-routine tasks, can be characterized, for example, by the use of creativity and the ability to analyze and solve problems, as well as negotiation, coordination, and personnel administration. Routine activities, on the other hand, can be characterized by repetitive production processes or by monitoring activities that involve both bureaucratic tasks in the office and machine operations in a factory. As non-routine manual tasks, we have tasks that require the use of physical capacity with some kind of perception or adaptation to different situations. Cleaning and repair services are examples of non-routine manual activities. Based on the schematization proposed by Reis (2018), we created a database that relates the CBO code of occupations with the three skill groups, indicating for each occupation which skill groups are related. Thus, we created three dummy variables for each skill group (cognitive non-routine, non-routine manual, and non-routine), assigning 1 if a given skill group is associated with the activities of the assessed occupation and 0 otherwise. Finally, we added the dummies created to our main base of descendants by the surname of Polish immigrants and Polish Jews, assigning for each descendant the skill groups that would be associated with the occupation exercised in each year.

B.6 Number of Employees per CNPJ

Variable created from information from the RAIS 2009 to 2017. First we aggregated all CNPJs present in the RAIS database from 2009 to 2017, adding up the observations for each CNPJ, thus obtaining the number of workers per CNPJ. Later, we paired this relationship with our main base of descendants by the surname of non-Jewish and Jewish Polish immigrants, thus obtaining for each descendant the number of workers employed in the company in which they are employed (identified by CNPJ) per year.

B.7 Company Size

Variable created from the Number of workers per CNPJ variable (Appendix B.6) considering the classification established by SEBRAE (2006), as shown in Table **??**. Thus, we created a

categorical company size variable that assigns 1 to companies associated to micro-enterprises via the number of employed people, 2 to small companies, 3 to medium-sized companies, and 4 to large companies.

B.8 Job Change

Variable created from the formed panel according to Subsection 4.1. We grouped the panel obtained by CPF and by year and counted the number of times the individuals' CNPJ varied over the available years (2009 to 2017). From this, we create a dummy variable that equals 1 if the individual has at least one change in CNPJ over the years and 0 otherwise.

B.9 Entry Wage

Variable created from the formed panel according to Subsection 4.1. First, we grouped the base by CPF and by year and created a reference variable to mark the year in which each individual changed employment status (CNPJ change in relation to the previous year). This reference variable takes 1 for observations corresponding to the years in which individuals changed jobs and 0 otherwise. If the year marked for each individual by the reference variable is greater than 2009 (after the end of 2009, we cannot infer that the individual started the CNPJ in the given year, considering the lack of previous history), the entry salary variable corresponds to the value of the reference variable multiplied by the value of the variable Ln of the individual's Average Nominal Annual Income (defined in Appendix B.1).

B.10 Top 10 Income

Variable created from the formed panel according to Subsection 4.1. We have arranged the base by year and created a dummy variable that equals 1 if the Ln of the income of each observation (variable defined according to Appendix B.1) is greater than the 90th percentile.

B.11 Top 10 Income for the First Job

Variable created from the formed panel according to Subsection 4.1. We kept in the database only individuals with valid employment relationships in their first working experience (the variable defined in Appendix B.2). We have arranged the base by year and created a dummy variable that equals 1 if the Ln of the income of each observation (variable defined according to Appendix B.1) is greater than the 90th percentile.

C Variables - Cultural Relationship

C.1 Jewish Polish Network

Variable created from the formed panel according to Subsection 4.1. We grouped the panel obtained by CNPJ and by year and summed the number of individuals with Jewish Polish surname descent, that is, those with the dummy Culture (Appendix A.1) equal to 1; thus, we obtained the number of individuals with this surname descent by CNPJ and by year. The Jewish Polish Network variable marks 1 for all cases in which the individual is Jewish Polish by surname descent and in his company (identified via CNPJ) there is more than one position occupied by a Jewish Pole.

C.2 Randomized Jewish Polish Network

Random dummy variable created from the Random Culture dummy variable (Appendix A.2) for the fake test. The Randomized Jewish Polish Network dummy variable equals 1 for cases in which there is more than one Jewish Poles in the individual's company and in which the individual's Randomized Culture dummy indicates the value 1.

C.3 Polish Network

Variable created from the formed panel according to Subsection 4.1. We grouped the panel obtained by CNPJ and by year and then added the number of individuals with Polish surname descent, that is, who have the dummy Culture (Appendix A.1) equal to 0, obtaining the number of individuals with this surname descent by CNPJ and by year. The Polish Network variable equals 1 for all cases in which the individual is Polish by surname descent and in his company (identified via CNPJ) there is more than one employed Pole.

C.4 Polish Jewish Network Proportion

Variable created from the formed panel according to Subsection 4.1. For cases in which there is a Jewish Polish network (variable defined according to Appendix C.1), we attribute the Jewish Polish Network Proportion variable, which corresponds to the number of Jewish Poles employed in each individual's company per year divided by the total number of workers in the company related to the individual per year (variable defined in Appendix B.6).

C.5 Polish Network Proportion

Variable created from the formed panel according to Subsection 4.1. For cases in which there is a Polish network (variable defined according to Appendix C.3), we attribute to the Polish Network Proportion variable the value that corresponds to the number of Poles employed in each individual's company per year divided by the total number of workers in the company related to the individual per year (variable defined in Appendix B.6).

C.6 Polish Jewish Worker-Management Cultural Relationship

Variable created from the formed panel according to Subsection 4.1. We grouped the panel obtained by CNPJ by year and created a dummy variable that equals 1 for observations whose culture (variable defined in Appendix A.1) corresponds to 1 (Jewish Polish) and the individual is in a leadership position (variable defined in Appendix B.4 corresponds to 1). Then, by CNPJ and per year, we add up the number of Polish Jewish leaders on the panel. The variable Polish Jewish Worker-Management Cultural Relationship equals 1 if the individual works in a company with at least one Jewish Polish leader and the individual is not a leader itself and is classified as a Polish Jew.

C.7 Polish Worker-Management Cultural Relationship

Variable created from the formed panel according to Subsection 4.1. We grouped the panel obtained by CNPJ by year and created a dummy variable that equals 1 for observations whose culture (variable defined according to Appendix A.1) corresponds to 0 (Polish) and the individual is in a leadership position (variable defined according to Appendix B.4 corresponds to 1). Then, by CNPJ and per year, we add up the number of Polish leaders in the panel. The Polish Worker-Management Cultural Relationship variable equals 1 if the individual works in a company with at least one Polish manager and the individual is not a leader and is classified as Polish.

C.8 Categorical Polish Jewish Network

Variable created from the formed panel according to Subsection 4.1. We grouped the panel obtained by CNPJ and by year and summed the number of individuals with Jewish Polish surname descent, that is, those with the dummy Culture (Appendix A.1) equal to 1; thus, we obtained the number of individuals with this surname descent by CNPJ and by year. The Categorical Jewish Polish Network variable equals 1 for cases in which the individual is

Jewish Pole by surname descent and in his company (identified via CNPJ) there is more than one and equal to or less than 20 Jewish Poles, equals 2 where the individual is a Jewish Pole and in his company there are more than 20 and equal to or less than 50 Jewish Poles, and equals 3 where the individual is a Jewish Pole and there are more than 50 Jewish Poles in his company. We also created a lagged version of this variable, in which we consider the earlier value of the Categorical Jewish Polish network variable.

C.9 Categorical Polish Network

Variable created from the formed panel according to Subsection 4.1. We grouped the panel obtained by CNPJ and by year and added the number of individuals with Polish surname descent, that is, those with the dummy Culture (Appendix A.1) equal to 0; thus, we obtained the number of individuals with this surname descent by CNPJ and by year. The Categorical Polish Network variable equals 1 for cases in which the individual is Polish by surname descent and in his company (identified via CNPJ) there is more than one and equal to or less than 20 Poles, equals 2 where the individual is Polish and there are more than 20 and equal to or less than 50 Poles, and equals 3 where the individual is Pole and there are more than 50 Poles in his company. We also created a lagged version of this variable, in which we consider the lagged value of the Categorical Polish Network variable.