FUNDAÇÃO GETULIO VARGAS
ESCOLA DE ECONOMIA DE SÃO PAULO

PEDRO LOPES DA SILVA

AFFIRMATIVE ACTION POLICIES AND MARRIAGE DIVERSITY: EVIDENCE FROM BRAZIL

São Paulo

2023

PEDRO LOPES DA SILVA

AFFIRMATIVE ACTION POLICIES AND MARRIAGE DIVERSITY: EVIDENCE FROM BRAZIL

Dissertação apresentada à Escola de Economia de São Paulo como pré-requisito à obtenção de título de mestre em Economia de Empresas.

Orientador: Fernanda Estevan.

São Paulo

2023

Silva, Pedro Lopes da.

Affirmative action policies and marriage diversity : evidence from Brazil / Pedro Lopes da Silva. - 2023.

50 f.

Orientador: Fernanda Gonçalves de La Fuente Estevan.

Dissertação (mestrado CME) – Fundação Getulio Vargas, Escola de Economia de São Paulo.

1. Programas de ação afirmativa na educação. 2. Ensino superior - Brasil. 3. Economia - Aspectos sociológicos. 4. Casamento. I. Fuente Estevan, Fernanda Gonçalves de La. II. Dissertação (mestrado CME) – Escola de Economia de São Paulo. III. Fundação Getulio Vargas. IV. Título.

CDU 33::378(81)

Ficha Catalográfica elaborada por: Isabele Oliveira dos Santos Garcia CRB SP-010191/O

Biblioteca Karl A. Boedecker da Fundação Getulio Vargas - SP

PEDRO LOPES DA SILVA

AFFIRMATIVE ACTION POLICIES AND MARRIAGE DIVERSITY: EVIDENCE FROM BRAZIL

Dissertação apresentada à Escola de Economia de São Paulo como pré-requisito à obtenção de título de mestre em Economia de Empresas.

Data de aprovação: 18/04/2023

Banca examinadora:

Prof. Fernanda Estevan
FGV-EESP (Orientador)

Prof. Bruno Ferman
FGV-EESP

Prof. Ursula Mello
PUC-Rio

Agradecimentos

A conclusão deste trabalho representa o fim de uma etapa importante em minha vida. Nela, vivenciei um processo de aprendizagem e amadurecimento tanto científico quanto pessoal. O sentimento que carrego comigo neste momento é o de gratidão por todas as experiências e pessoas em minha trajetória. Em especial, pelos meus pais, Marcia e Joaquim, que sempre acreditaram no potencial transformador da educação e me apoiaram incondicionalmente na decisão de realizar o mestrado. Pelas minhas irmãs, Marina e Julia, que me ajudaram nos momentos necessários e que são fonte de excelente companhia. Pela minha orientadora, Fernanda, que me guiou no processo de ideação e edificação dessa pesquisa de maneira muito gentil, solícita e paciente. Por sua qualidade técnica, posso afirmar que foi um privilégio poder discutir ideias com ela. Por seu jeito de ser, posso afirmar também que essa atividade foi bastante divertida. Por fim, sou grato aos demais participantes da banca avaliadora deste trabalho, Bruno e Ursula. É uma honra poder contar com os comentários de pesquisadores tão qualificados.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

Resumo

Este artigo investiga se políticas de ação afirmativa implementadas no sistema de ensino superior público brasileiro afetam a diversidade de casamentos. Analisamos duas dimensões da diversidade. A primeira é casamento inter-tipo de ensino médio (privado/público) - um *proxy* para a origem socioeconômica dos indivíduos - e a segunda é interétnico. Para definir a exposição dos indivíduos às políticas de ação afirmativa, exploramos a variação no tipo de faculdade que os indivíduos frequentaram e a heterogeneidade da implementação da política entre as universidades e ao longo do tempo. Nossos resultados sugerem que as políticas de ação afirmativa aumentam os casamentos entre indivíduos que frequentaram diferentes tipos de ensino médio e diminuem os casamentos interétnicos. No geral, mostramos que o contato inter-grupos promovido pela política afeta os relacionamentos de longo prazo em que os indivíduos se envolvem.

Palavras-chave: Ações Afirmativas, Educação Superior, Casamento, Diversidade

Abstract

This paper investigates whether affirmative action policies implemented in the Brazilian public higher education system affect the diversity of marriages. We analyze two dimensions of diversity. The first is high school inter-type (private/public) - a proxy for the socioeconomic background of individuals - and the second is interethnic marriage. To determine individuals' affirmative action exposure, we explore the variation in the type of college individuals attended and the heterogeneity of the policy implementation across universities and over time. Our results suggest that affirmative action policies increase high-school inter-type marriages and decrease interethnic marriages. Overall, we show that the inter-group contact promoted by the policy affects the long-term relationships individuals engage in.

Keywords: Affirmative Action, Higher Education, Marriage, Diversity

JEL Classification: I23, J12, Z13

List of Figures

List of Tables

Contents

1 Introduction

Sociologists commonly use intermarriage to measure tolerance, reciprocity, trust, and social ties among groups. In this context, Gordon (1964) posits that there is an indissoluble connection between structural assimilation and marital assimilation.[1] For the author, high intermarriage rates reflect that a minority group has achieved widespread acceptability at the familial level, removing the main obstacles preventing their full acceptance into the dominant society.

A pre-requisite for intermarriage is the exposure among groups, which depends on their spatial distribution. However, as shown by Reardon and Bischoff (2011), a robust relationship exists between income inequality and income segregation, characterized by the uneven spatial distribution of income groups within certain areas. Moreover, given the strong correlation between ethnicity and income, income segregation is also often empirically linked with ethnic segregation.

Since Brazil presents one of the highest levels of income inequality in the world, exposure among individuals of distinct socioeconomic levels is often limited, which decreases the chances of intermarriages between socioeconomic groups. Furthermore, although miscegenation presents high rates - demonstrating interethnic sociability - the ethnic mixture is not uniform in all sectors of Brazilian society. As Telles (2006) discusses, it is common to observe Brazilians of lower socioeconomic classes of all ethnicities associate among themselves without salient regard for color. Still, there is little interethnic marriage among those from more advantaged classes. He reasons this may be because whites are in various tiers of Brazilian society while most non-whites are in the lower socioeconomic classes. Therefore, interethnic contact among the higher socioeconomic classes would be limited, affecting interethnic marriages in this tier.

The implementation of affirmative action policies in Brazilian public universities provides an interesting context to examine how exposure among groups translates into intermarriage. There is substantial evidence in the literature that affirmative action in Brazil succeeded in increasing college attendance for disadvantaged groups targeted by the policy,

[1] Gordon (1964) defines structural assimilation as the "entrance of the minority group into the social cliques, clubs, and institutions of the core society at the primary group level." (pag. 80)

specifically public school and non-white students.[2] Changing the (hitherto homogeneous) college-admitted students' profiles created a new university social environment. Yet, little is known about the effect of affirmative action on social interactions between those whose policy targets and those whose not. Furthermore, this setting is particularly interesting because, as Kirkebøen et al. (2021) shows, college is a local marriage market. According to the authors, graduates are likely to marry within their own institution not because of the admitted students' predetermined traits but due to attending a specific institution at a given time.

In this article, we investigate whether the implementation of affirmative action policies at Brazilian universities affects this core indicator of inter-group contact, intermarriage. We consider two dimensions of diversity. The first is the high school inter-type (private/public) - a proxy for the socioeconomic background of individuals - and the second is interethnic marriage. Importantly, this proxy for the background is appropriate since, in Brazil, the quality of the high school and students' backgrounds differ systematically between private and public schools.[3]

We explore the introduction of affirmative action across the universe of state and federal public higher education institutions in Brazil. The variation we use to identify the policy effects arises from the type (private/public) of college individuals attended and from the heterogeneity of affirmative action implementation across universities and over time. However, we are in a challenging research setup. The difficulty we face is investigating the impact of affirmative action on a long-term outcome without longitudinal data on individuals.

Crucially, we exploit a cross-sectional household survey that includes information on the types of college and high school individuals attended and their ethnicity, age, and metropolitan statistical area (MSA) of residence.[4] Yet, we do not observe when individuals attended college and in which institution. The lack of these variables implies we must make some assumptions to determine which individuals were affected by the policy and by

2 Mello (2022), Vieira and Arends-Kuenning (2019), and Francis and Tannuri-Pianto (2012) provide evidence in this direction.

3 Cavalcanti et al. (2010) provide evidence in this direction by studying the profile of private and public high school students attending a Brazilian public university entrance exam. The authors show not only that the test score in the exam from those students of public schools is lower but also that they have lower family incomes and parents' educational levels.

4 Other papers use cross-sectional data to investigate the long-term impacts of a policy. Duflo (2001), for example, use cross-sectional data to study the e<0xEF><0xAC><0x80>ects of building primary schools on the wages of those directly benefited by the policy.

how much. To proxy individuals' exposure to treatment, we use their birth cohort, type of college attended, and MSA. More specifically, we assume individuals entered college by the year they were 19 and attended college in the same MSA as their current residence. For those individuals who attended a public college, we build a measure of affirmative action exposure by linking each pair of college admission year and MSA with data on the share of vacancies reserved for affirmative actions by year and MSA. For those who attended private universities, we assume they were unaffected by the policy.[5]

Furthermore, studying marriages in an applied work imposes another challenge due to the necessity of accommodating the stable unit treatment value assumption in the empirical strategy, which is necessary for identifying causal effects in a potential outcomes framework. This assumption requires that the potential outcome of each unit does not depend on the treatment status of other units. The difficulty we face is that marriage is a bilateral decision, so it depends on the treatment status of both individuals of the couple. To overcome this challenge, we use the household as the unit of analysis, and therefore its treatment status depends on the affirmative action exposure of both individuals of the couple.

Using a fixed-effect model, we employ different specifications to assess the effect of affirmative actions on marriage diversity. Regardless of the specification we use, when analyzing high school inter-type marriages, the point estimates of our regressions suggest that affirmative action exposure increased marriage between individuals that attended different types of high school. On the other hand, when analyzing interethnic marriages, they suggest the policy decreased marriage between individuals of different ethnicities. However, the significance level of our estimates is very sensitive to the specification choice for both outcomes. When we use a continuous measure of affirmative action exposure and adopt the most flexible model, we find that an increase of one percentage point in affirmative action exposure increases high school inter-type marriages by 0.19 percentage points and decreases interethnic marriages by 0.27 percentage points. In these models, we can reject the null hypothesis of no effect.

Moreover, we examine assortative matching on the type of high school (couples who both attended the same high school type) and ethnicity to understand the compositional

5 Although private universities usually did not implement affirmative action, there were federal initiatives that may have stimulated private college enrollment for disadvantaged groups. Policies such as the *Programa Universidade para Todos* and the *Fundo de Financiamento ao Estudante do Ensino Superior* may have changed the ethnic and socioeconomic composition of students attending private institutions.

changes in marriage diversity. As before, in this analysis, the coefficients' significance level in which we can reject the null hypothesis of no effect is sensitive to the specification choice. However, based on the point estimates of the regressions, we find that the increase in high school inter-type marriages is mostly associated with the decrease in assortative matching between private high school individuals. In addition, we find that the decrease in interethnic marriages is mostly associated with the increase in assortative matching on ethnicity for white individuals.

Overall, our results suggest a two-sided story of the effects of affirmative action on marriage. On the one hand, it increases marriage among individuals who attended different types of high schools by decreasing the likelihood of assortative matching on high school type by private high school individuals. On the other hand, it decreases marriage among individuals of different ethnicities by increasing the likelihood of assortative matching on ethnicity by white individuals.

Our paper contributes to three strands of the economic literature. First, it adds to the literature on affirmative action in higher education. Researchers have already examined various topics related to this theme. For example, Mello (2022), Vieira and Arends-Kuenning (2019), and Francis and Tannuri-Pianto (2012) study this policy's effects on students' college enrollment. Moreover, Francis-Tan and Tannuri-Pianto (2018) focus on the labor market outcomes of the target group and Estevan et al. (2019) on the major choice. However, as far as we know, this is the first paper to study the effect of affirmative action on marriage directly. We contribute to this literature by showing that affirmative action affects which relationships individuals engage in.

Second, it adds to the literature on the effects of intergroup contact on social ties. Particularly, two papers in this area deeply relate to our research context. First, Merlino et al. (2019) studies whether interethnic contact in childhood impacts adult romantic relationships. The authors provide evidence that more black school peers of the same gender causes white to have more relationships with blacks as adults. Second, Billings et al. (2021) examines whether behaviors change when people of distinct ethnicities are randomly assigned to live together during their first college year. The authors find that white students randomly assigned to black roommates are more likely to support affirmative action and have personal contact with individuals of other ethnicities after their first year. These articles suggest that white individuals' contact with non-white individuals leads to higher social interaction. In contrast, we show that social exposure among groups does not always increase social interactions, as measured by marriages.

Finally, our paper contributes to the literature on marriage markets and assortative matching. The marriage literature concerned with drawing causal inferences is in its infancy. Usually, papers are most concerned with the assortative matching on educational outcomes, such as the major choice in college, as Kirkebøen et al. (2021) and Artmann et al. (2018), or assortative matching on ethnicity, as Merlino et al. (2019). Our research context takes advantage of the fact that, in Brazil, there are labels of income level associated with the type of high school individuals attended, which in turn, represents the socioeconomic level of individuals before they marry. Therefore, we also contribute to the marriage literature by drawing causal inferences on a new topic, inter-socioeconomic background marriages.

Moreover, many empirical papers in the marriage literature have documented the tendency of individuals to sort into homogeneous marriages through empirical work. Eika et al. (2019) finds evidence of this pattern for several developed countries and Pereira and Santos (2017) for Brazil. In this last paper, the authors show that between 1970 and 2010, marriages in the country become increasingly selective since individuals marry partners with more similar socioeconomic characteristics over the years. By doing counterfactual exercises, the authors show that less selective marriages could have decreased income inequality across the period. We also contribute to the marriage literature by providing evidence that public policies that stimulate intergroup contact, such as affirmative action policies, impact whom one marries, a subject that may have important implications for income inequality.

We organize this paper as follows. Section 2 describes the institutional context of affirmative actions in Brazilian public universities. Section 3 describes the data, method for design treatment assignment, sample selection, main variables, and unit of analyses. The empirical strategy, results, and sensitivity analyses are present in Section 4, 5, and 6, respectively. Section 7 concludes.

2 Institutional Context

The Brazilian Higher Education System combines private and public institutions. According to the 2010 Higher Education Census, there are 2099 private and 278 public institutions.[1] In most cases, public universities are extremely selective. On average, they are known for their superior quality than their private counterparts. Moreover, by law, they are free of charge, while private institutions are not. Only 10.5% of the nearly 4.75 million students enrolled in undergraduate courses study at public institutions.

Due to the high competitiveness of the college admission process, access to public higher education in Brazil has historically been achieved mostly by privileged students. Those from wealthier families that could offer better pre-college educational opportunities, such as private high schools and private courses aimed at university entrance exams, used to have a great advantage in obtaining university vacancies. In this setting, most disadvantaged students did not attend the Brazilian public higher education system for long.

In the early 2000s, some public universities started to adopt affirmative action policies to mitigate the lack of representation of disadvantaged groups. The default policy was a quota system that guaranteed a share of vacancies reserved to students based on their ethnicity, family income, or type (public/private) of high school attended.[2] During this decade, no national initiatives obligated universities to implement affirmative action, and each institution could determine whether and how to implement it, resulting in high heterogeneity in policy adoption.

By 2012, the Brazilian government enacted Law 12.711/2012 - the National Law of Quotas - imposing that until 2016, each major in federal universities should reserve 50% of its vacancies for students that attended only public high school. Half of the spots reserved for quotas should be for students with a family gross per capita income of a maximum of 1.5 minimum wages. In addition, from the reserved vacancies, a minimum share is

[1] The public higher education system comprises 99 federal, 108 state, and 71 municipal public institutions. This article focuses only on the a□rmative actions implemented at federal and state institutions, which account for approximately 87% of the spots in the public higher education system.

[2] Few higher education institutions implemented alternative policies, such as bonuses on the students' scores at the universities' admission exams. This article focuses only on the quota policy because its harder to measure how bonus strategies translated into disadvantaged groups' college enrollment.

destined for indigenous, black, and mixed students, proportionally to the percentage of these ethnic groups in the population of each state (measured by the 2010 Population Census). Moreover, the National Law of Quotas allowed for different policy implementation timing. The federal universities were obligated to increase the share of vacancies reserved for affirmative actions by at least 12.5% in each year between 2013 and 2016. Still, they could implement a faster transition to the 50% postulated by the federal government. Finally, until the end of our analysis period, 2015, the state universities remained without a common guideline.

3 Data

3.1 Data and Method for Design Treatment Assignment

This work uses data from distinct sources. The Continuous National Household Sample Survey (henceforth, PNADC) from 2016 to 2019 provides household data.[1] We use the Higher Education Census, the replication package of Mello (2022), and both federal and state university's entrance exam regulation norms to construct a database containing the share of vacancies reserved for affirmative action (AA) in each metropolitan statistical area (MSA) from 2004 to 2015. Finally, the 2010 Brazilian Population Census provides data on migration.

In the PNADC database, individuals may declare themselves as the household head or as their spouse or partner. Therefore, we consider that the household contains a married couple when we observe both a household head and a spouse or partner. PNADC also contains several individuals' educational and demographic variables. We observe their highest course attended previously, whether they concluded it, their ethnicity, age, and region of residence by state and MSA.

We use PNADC data from each second quarter of 2016 to 2019 because only in those editions the survey collected supplementary educational data about the type (private/public) of high school attended.[2] In addition, if she attended an undergraduate or graduate program, there is also information about the university's type (private/public).

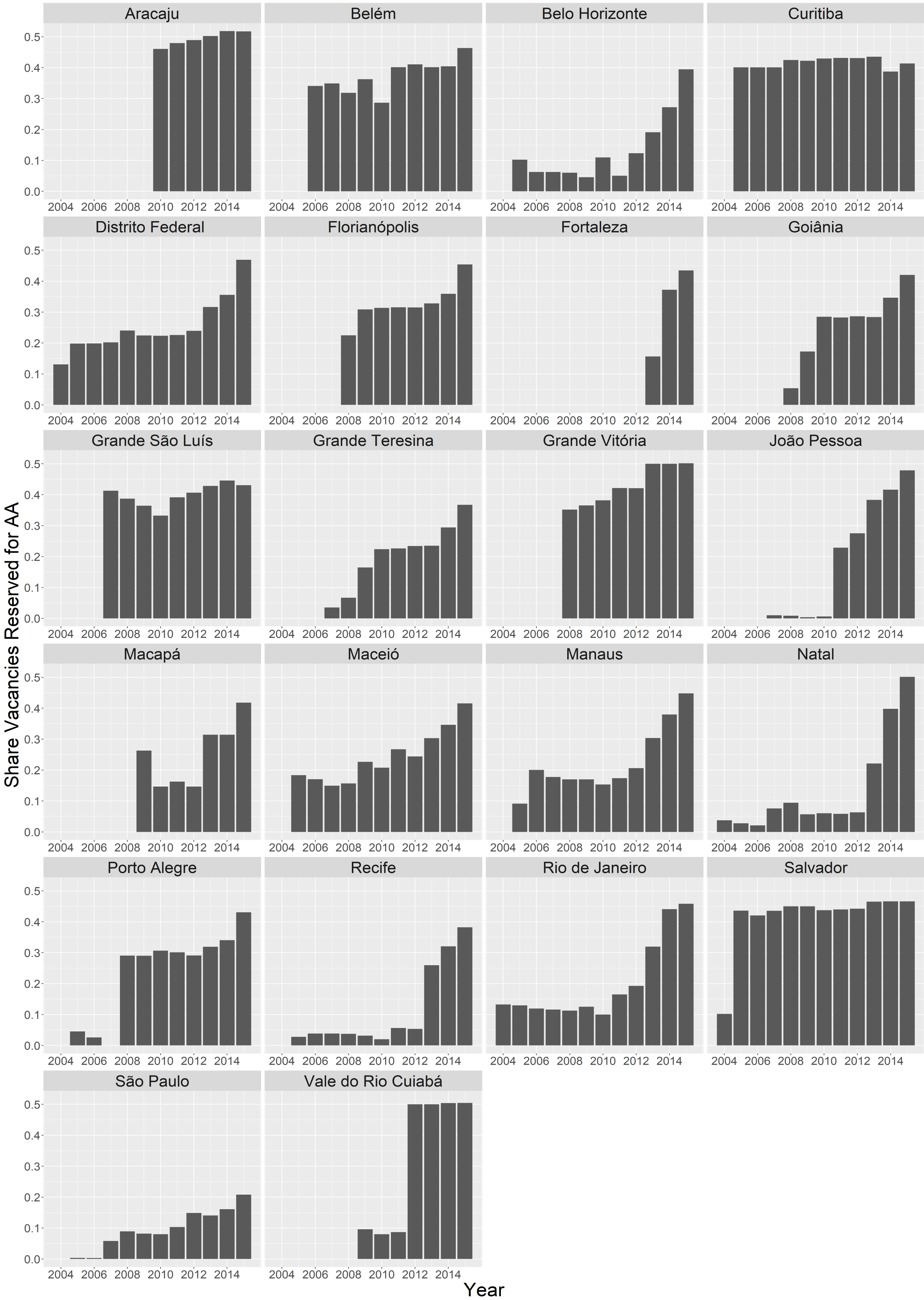
To obtain data on the share of vacancies reserved for AA at higher education institutions, we use the replication package of Mello (2022) and the university's entrance exam regulation norms. The former provides the share of vacancies reserved for AA for each state and federal university from 2010 to 2015. Through the latter, we obtain the same information from 2004 to 2009. Additionally, we use the Higher Education Census to collect the total number of vacancies offered for each of these universities by the metropolitan

1 *Pesquisa Nacional por Amostra de Domicilíos Contínua* in Portuguese.

2 The PNADC collects data every quarter of the year, and each household answers the survey through five consecutive quarters. The sample planning has a rotation scheme, partially overlapping the sample over the quarters. Thus, a fraction of households answer the survey in consecutive years. We only use the data from the first time the household answers the survey in the second quarter.

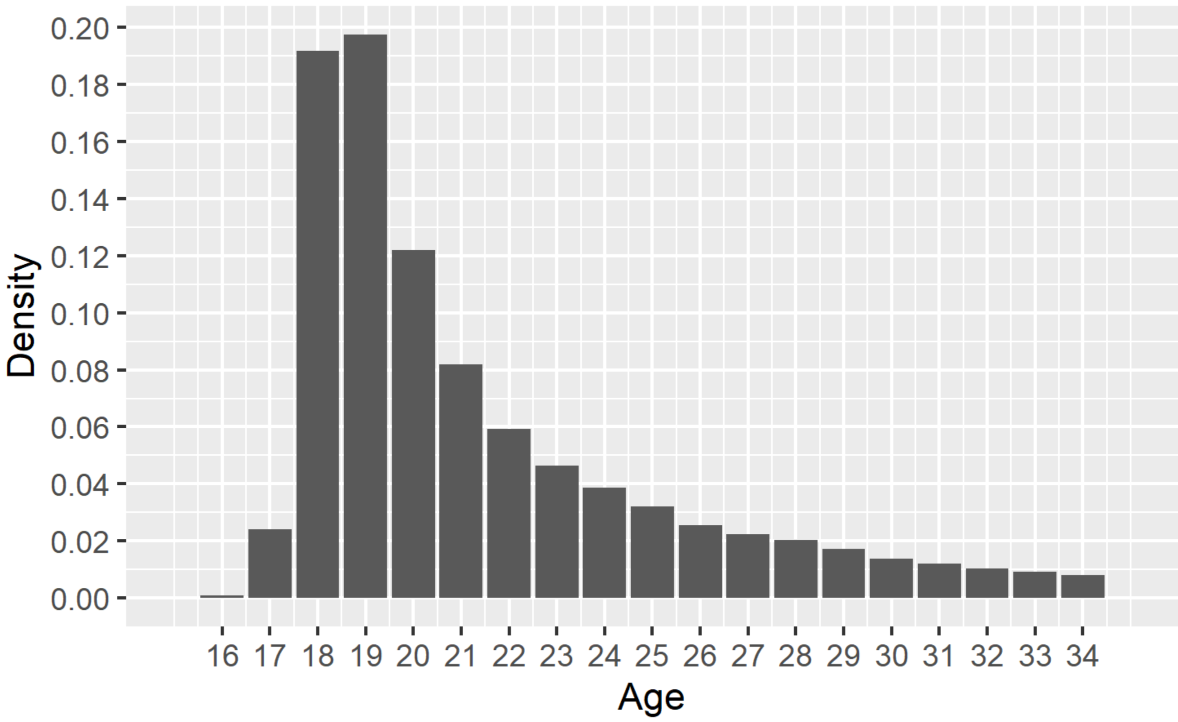
statistical area over the entire period. Then, for each MSA and year from 2004 to 2015, we sum universities' shares of vacancies reserved for AA weighted by their total number of vacancies offered. By doing so, we built a database containing the share of vacancies reserved for AA per MSA and year, represented graphically in Figure 1.

Figure 1 – Share of Vacancies Reserved For AA Per MSA And Year



We use the AA database and PNADC to identify which individuals were affected by the policy and its intensity. We consider that only the individuals who attended a public university could have AA exposure. However, not all of them had because of the heterogeneity of the policy implementation across universities and over time. Since there is no available information on when these individuals attended a college and in which institution, we build a measure of AA exposure based on their birth cohort and MSA. We assume all individuals entered college by 19, the mode age of public university admission, 19 as presented in Figure 2. Furthermore, we assume they attended a university in their current MSA of residence. With this, it is possible to proxy their AA exposure by the proportion of vacancies reserved for AA in their MSA of residence when they were 19.

Figure 2 – Age of College Admission In Public Universities



Source: Higher Education Census Dataset (2010)

To check the plausibility of our assumption about the MSA of residence being the same as the MSA of college attendance, we use migration data from the 2010 Population Census. Table 1 presents descriptive statistics on married individuals who live with their partner in a metropolitan area, are between 23 and 34, and whose highest course attended previously by them and their partner was an undergraduate course.[3] Since these individuals were 19, about 85% of them still live in the same state, and nearly 70% still live in the same municipality. Furthermore, we find that as much as 73% of them never migrated or had their municipality of residence in the same MSA, and almost 79% never migrated or had their municipality of residence five years ago in the same MSA. We conclude that our assumption holds for the majority of these individuals.

Table 1 – Migration Statistics

	%
Lives in the same state since 19	85.63
Lives in the same municipality since 19	70.59
Never migrated or last residence municipality is in the same MSA	73.34
Never migrated or municipality of residence five years ago was in the same MSA	79.24

Source: Population Census Dataset (2010)

3.2 Unit of Analysis, Main Variables, and Sample

Our unit of analysis is the household, and the sample of interest consists of households with married individuals. The choice of the unit of analysis is due to the stable unit treatment value assumption (SUTVA) used to identify causal effects in a potential outcomes framework. This assumption implies that there are no externalities between units, i.e., the outcome of each unit does not depend on whether other units received treatment. Since marriage is a bilateral decision, it depends on the treatment status of both individuals of the couple. Therefore, it prevents us from examining causal effects in marriage diversity at another unit of analysis than the relationship (represented by the household in our setup).

The main outcome of interest is marriage diversity. In this article, this variable assumes two forms. The first is a dummy variable that assumes value one when the married individuals attended different types (public/private) of high school and zero otherwise.[4] By examining marriage diversity at the high school level, we aim to capture diversity in the socio-economic background of the couple. The second is a dummy variable that assumes

[3] These are some criteria for being eligible for our sample, as discussed in the next subsection.
[4] Individuals who attended both types of high school are a small fraction of the sample. For simplification, we consider them as having attended only a public institution.

value one when the married individuals have a different ethnicity, specifically white and non-white, and zero otherwise.[5]

Our treatment variable is AA exposure, assuming continuous and binary forms. To define the former, let $h \in \{1, ..., H\}$ be the index of households, and $p \in \{1, 2\}$ be the index of married individuals within a given household. We assume each p receives a dose d_p such that d_p equals the share of vacancies reserved for AA at the MSA of residence of person p in the year of her college admission if she attended a public college and zero otherwise. Additionally, we assume that each h receives a dose d_h where $d_h := max\{d_1, d_2\}$. Since the maximum number of quotas in an MSA is 50 percent, we have that $0 \leq d_h \leq 0.5$. The binary treatment variable is a dummy that receives value one if $d_h > 0$ and zero otherwise.

We keep in our sample only couples whose highest courses attended previously are undergraduate programs.[6] Moreover, we restrict the sample to households whose both individuals are strictly older than 22 and whose youngest member of the couple was born after 1984. Appendix's Figure A1 illustrates the age criteria for the sample. The restriction on the individuals' minimal age of 23 is to avoid putting into the sample the households where the treatment exposure happened after the outcomes variables definition. Moreover, we restrict the sample to households whose both individuals are strictly older than 22 and whose youngest member of the couple was born after 1984. The restriction on the individuals' minimal age of 23 is to avoid putting into the sample the households where the treatment exposure happened after the outcomes variables definition.[7] In turn, the constraint related to the maximum age of the youngest member of the couple is to guarantee that each household is susceptible to affirmative action exposure. Indeed, when the first universities started to implement AA in 2004, the youngest member of the couple had a maximum age of 19, as they were born in or after 1985.[8] Furthermore, by imposing this restriction towards the maximum age of the youngest member of the couple, we also

5 We consider as white those individuals who declare themselves white or Asian. No a□rmative action specifically targeted these ethnicities.

6 A□rmative actions happened only in undergraduate programs. Since we do not observe the type of undergraduate institution people who attended graduate school studied at, we do not know whether they had AA exposure. Thus, we decide to keep them out of the sample.

7 If we do not impose this restriction, we would, for example, include in our sample individuals who married at 20 but were college admitted at 22. Note that we adopt a higher minimum age to be eligible for our sample than the age our treatment definition would require, i.e., 20. We opt to be cautious to avoid misleading results in our analyses.

8 For the sake of clarity, note that if the youngest member of the couple born in 1984 or earlier, she would be 20 or older in 2004. Hence, accordingly to our individuals' treatment exposure definition, she would not have been treated.

avoid putting in our sample households whose birth cohorts of their members are far different from those susceptible to AA exposure.

Finally, in the sample, we set the region of residence as the MSA. This constraint implies that the AA exposure measure relies on the share of vacancies reserved in the MSA and that we restrict our sample to households in their state's metropolitan region. By doing so, we expect our measure of treatment exposure to be more accurate than it would be if we did the analysis keeping households from the entire state. To be more specific, given that are more universities at the state level than at the MSA and that there was substantial heterogeneity in AA across universities, we tend to have more precision in the level of household's exposure to the treatment by restricting the geographical area in which we observe the sample.

4 Empirical Strategy

4.1 Empirical Strategy

As discussed in Section 3, the household treatment status depends on whether at least one individual of the couple had AA exposure. In turn, an individual's exposure depends on the type of college she attended, MSA of residence, and birth cohort. The type of college is the primary source of variation in individual exposure to affirmative action since only public universities implemented AA policies. Other sources of variation come from the MSA and birth cohort, as the universities implemented this policy differently across time and place.

In our empirical strategy, we consider that each household is subject to the same sources of variation as the individuals of the couple. Importantly, since each household has two individuals, it belongs to two birth cohorts. For simplification, we consider that each household belongs to only the birth cohort of the youngest member of the couple. This decision relies on the fact that younger cohorts at public universities are more likely to have AA exposure.

We use the following fixed-effect model to estimate the impact of affirmative action on marriage diversity:

$$Y_{hrts} = \lambda_r + \lambda_t + \lambda_s + \lambda_{rt} + \lambda_{st} + \beta D_{hrt} + \varepsilon_{hrts} \tag{4.1}$$

where Y_{hrts} is a dummy that takes value one when the household h at MSA r who answered the survey in the year s and with the youngest member of the couple belonging to the birth cohort t has a diverse marriage. λ_r and λ_t account separately for specific unobserved confounders in the MSA and youngest birth cohort, while λ_{rt} accounts for the unobserved birth cohorts' confounders that may vary across MSA. We include λ_s since households answered the PNADC survey in different years and use the interaction term λ_{st} to account for observing each birth cohort at different ages. The treatment variable is D_{hrt}, continuous or binary, depending on the specification. The error term, ε_{hrts}, captures the unobserved determinants of marriage diversity. We cluster standard errors at the MSA level.

We do not use household weights in the regressions. The PNADC weights are designed using the entire sample, while in this research, we use only a subsample. Moreover, we aggregated different survey years to increase the sample size, so whether the sample weights remain valid after this procedure is not obvious. Also, according to Solon et al. (2013), unweighted regressions provide unbiased and more precise estimates if no endogenous sampling exists. The authors also suggest reporting both weighted and weighted regressions in cases where there are many nuances in deciding whether and how to weigh. We report results using PNADC survey weights in Appendix A.

4.2 Identification

We need the conditional independence assumption to hold to identify the average treatment effect on the treated. In our setup, this assumption states that conditional on the fixed effects variables of our model, AA exposure is independent of potential outcomes. In other words, we must assume that no household characteristics correlate with the treatment status and outcome variable after conditioning on the control variables (i.e., there is no omitted variable bias). Moreover, we also must assume the absence of sample self-selection.

Our sample has only households with married individuals. However, if AA exposure changes the probability of marriage, households may self-select into or out of the sample due to the treatment status. If that happened, observed treated households would be systematically different from the control group, leading to bias in the estimates. Thus, to identify the impact of AA on marriage diversity, we must ensure that there is no sample self-selection.

To check whether this is a concern, we test whether AA exposure impacts marital status. To do so, we construct a new sample containing households with married couples (our main sample) and those whose household head has no spouse or partner. Then, we run equation (3.1) using as the outcome variable a dummy variable that receives value one when the household contains married individuals and zero otherwise.

Table 2 presents AA exposure's impacts on the probability of marriage. In addition to the model presented in equation (3.1), we also use alternative specifications to evaluate the sensitivity of the results to the choice of control variables. In models (1) and (4), we exclude the interaction between MSA and the birth cohort. Models (2) and (5) substitute the interaction term with a linear MSA trend. Models (3) and (6) correspond to our main specifications using the interacted fixed effect between the youngest birth cohort and the

MSA.

For all specifications, we cannot reject the null hypothesis of no effect of AA exposure on the probability of marriage. This result suggests that the treated households in our sample are not likely to be systematically different from those in the control group regarding their decision to marry. With this finding, our identification strategy becomes more credible, even if it is not guaranteed valid (it still may be confounders, correlated to the outcome and the treatment variable after conditioning on the covariates).

Table 2 – Affirmative Action and Probability of Marriage

Dependent Variable:	Marriage					
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
AA Exposure Binary	-0.0239	-0.0274	-0.0367			
	(0.0219)	(0.0230)	(0.0250)			
AA Exposure Continuous				-0.0449	-0.0658	-0.0918
				(0.0869)	(0.0912)	(0.1034)
p-value	0.2887	0.2470	0.1566	0.6107	0.4783	0.3849
Control Mean	0.5263	0.5263	0.5263	0.5263	0.5263	0.5263
MSA Linear Trend						
		Yes			Yes	
Fixed-e ects						
MSA	Yes	Yes	Yes	Yes	Yes	Yes
Youngest Birth Cohort	Yes	Yes	Yes	Yes	Yes	Yes
PNADC Year	Yes	Yes	Yes	Yes	Yes	Yes
PNADC Year-Youngest Birth Cohort	Yes	Yes	Yes	Yes	Yes	Yes
MSA-Youngest Birth Cohort			Yes			Yes
Fit statistics						
Observations	3,768	3,768	3,768	3,768	3,768	3,768
MSA Clusters	22	22	22	22	22	22

Clustered (MSA) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

5 Results

5.1 Marriage Diversity

We show the effects of AA exposure on the high-school type and ethnic marriage diversity in Tables 3 and 4, respectively. Models (1) to (3) use the binary treatment variable, and (4) to (6) use the continuous treatment variable. In models (1) and (4), we do not include the interaction term between MSA and the birth cohort but only the other terms in equation (3.1), such as the MSA and birth cohort fixed effects separately. In this case, we assume the effect of being in a given MSA on marriage diversity is equal to everyone in that MSA, independent of their birth cohort. Analogously, we assume the effect of being in a given birth cohort is the same for everyone in that birth cohort, independent of their MSA. These are restrictive assumptions. It may be that the effect of being in a certain MSA varies over birth cohorts. We loosen these restrictive assumptions in models (2) and (6) to account for this possibility by including an MSA linear trend. In these models, we let the effect of being in a given MSA to vary over the birth cohort. However, we impose that these variations are linear over the birth cohort. In our main specifications, presented in models (3) and (6), we include the interaction term between the birth cohort and MSA. Hence, we let the effect of being in a certain MSA vary arbitrarily over birth cohorts.

All point estimates in Table 3 are positive, suggesting that affirmative action increases high school inter-type marriages. However, the significance levels of the coefficients are sensitive to the specification choice. Models (3) and (6), including the interaction term between MSA and birth cohort fixed effects, are the only ones in which we can reject the null hypothesis of no effect at conventional significance levels. When using the binary treatment variable and interaction between MSA and birth cohort fixed effects, the point estimate is about 0.039, which means that AA exposure increases high school inter-type marriages by 3.9 percentage points. This is a large increase since it represents nearly 21% of the mean marriage diversity in the control group. Other specifications that use the binary treatment variable - models (2) and (3) - present point estimates of about 0.0317 and 0.0315, respectively.

Regarding the models with the continuous treatment variable, we should interpret them cautiously. Considering the model with the interaction between birth cohort and MSA, it has a point estimate of about 0.1935, which means an increase by one percentage point of AA exposure increases linearly, the chance of having an inter-high school marriage by nearly 0.1935 percentage points. The other specifications using the continuous treatment variable have point estimates considerably lower than the previous one. They are about 0.1216 and 0.1084 in models (4) and (5), respectively.

Table 3 – Affirmative Action and High School Diversity

Dependent Variable:	HS Diversity					
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
AA Exposure Binary	0.0317	0.0315	0.0390			
	(0.0213)	(0.0227)	(0.0225)			
AA Exposure Continuous				0.1216	0.1084	0.1935
				(0.0820)	(0.0930)	(0.0917)
p-value	0.1507	0.1804	0.0978	0.1530	0.2568	0.0470
Control Mean	0.1858	0.1858	0.1858	0.1858	0.1858	0.1858
MSA Linear Trend						
		Yes			Yes	
Fixed-e ects						
MSA	Yes	Yes	Yes	Yes	Yes	Yes
Youngest Birth Cohort	Yes	Yes	Yes	Yes	Yes	Yes
PNADC Year	Yes	Yes	Yes	Yes	Yes	Yes
PNADC Year-Youngest Birth Cohort	Yes	Yes	Yes	Yes	Yes	Yes
MSA-Youngest Birth Cohort			Yes			Yes
Fit statistics						
Observations	1,956	1,956	1,956	1,956	1,956	1,956
MSA Clusters	22	22	22	22	22	22

Clustered (MSA) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

The point estimates regarding inter-ethnic marriages are negative in all specifications suggesting that AA exposure decreases inter-ethnic marriages. In models (1) and (2), the coefficients are statistically insignificant and small, -0.0165 and -0.0119, respectively. The point estimate of the model (3) is about -0.04, which is a large effect since it represents almost 15% of the mean diversity of the control group. Moreover, its p-value is 0.1064, being almost significant at conventional levels. Models (4) and (5) have point estimates of -0.1354 and -0.1254, respectively. However, we can reject the null of no effect in either. In model (6), we find that an increase in AA exposure by one percentage point linearly decreases the probability of having an ethnically diverse marriage by 0.271 percentage points. This effect is statistically significant at 5% since its p-value is 0.0117.

Table 4 – Affirmative Action and Marriages with Ethnic Diversity

Dependent Variable:	Ethnic Diversity					
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
AA Exposure Binary	-0.0165	-0.0119	-0.0402			
	(0.0233)	(0.0251)	(0.0238)			
AA Exposure Continuous				-0.1354	-0.1254	-0.2748
				(0.0898)	(0.1016)	(0.0995)
p-value	0.4869	0.6396	0.1064	0.1467	0.2310	0.0117
Control Mean	0.2710	0.2710	0.2710	0.2710	0.2710	0.2710
MSA Linear Trend						
		Yes			Yes	
Fixed-e ects						
MSA	Yes	Yes	Yes	Yes	Yes	Yes
Youngest Birth Cohort	Yes	Yes	Yes	Yes	Yes	Yes
PNADC Year	Yes	Yes	Yes	Yes	Yes	Yes
PNADC Year-Youngest Birth Cohort	Yes	Yes	Yes	Yes	Yes	Yes
MSA-Youngest Birth Cohort			Yes			Yes
Fit statistics						
Observations	1,956	1,956	1,956	1,956	1,956	1,956
MSA Clusters	22	22	22	22	22	22

Clustered (MSA) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

5.2 Assortative Matching

In the previous subsection, we provide evidence that affirmative action decreases assortative matching on the high-school type and increases assortative matching on ethnicity. In other words, AA exposure increases marriage diversity at the high school type level and decreases marriage diversity at the ethnic level. A relevant exercise in understanding these findings is examining the compositional changes within each group (high school and ethnic) on assortative matching. Specifically, if AA exposure indeed affects marriage diversity, we should examine which categories (e.g., public and private) within each group explains the aggregate outcome. For example, consider the increase in high school marriage diversity. In that case, we should examine whether it happens together to a decrease in assortative matching between those from public, private, or both types of schools. With this purpose, we run the same regressions as before, except for the outcome variables. Our dependent variable is a dummy variable, equal to one when both individuals within a household are from the same category (e.g., private-private or public-public) and zero otherwise.

In Tables 5 and 6, we examine whether the increase in high school marriage diversity occurred among those who studied at public or private high schools. Even though we cannot reject the null hypothesis of no effect for all specifications in both tables, the point estimates usually suggest that those from private high schools have played a more prominent role. The results in Table 5 are volatile - points estimates using the continuous treatment variable oscillate of sign according to the specification. Nonetheless, columns (4) to (6) of Table 6 shows that assortative matching of individuals who attended private high school decreases in a range from nearly 0.1871 to 0.1394 percentage points when AA exposure increases by 1 percentage point.

Table 5 – Affirmative Action and Assortative Matching of Individuals Who Attended Public High School

Dependent Variable:	Both Public					
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
AA Exposure Binary	-0.0108	-0.0066	-0.0152			
	(0.0452)	(0.0453)	(0.0453)			
AA Exposure Continuous				0.0537	0.0787	-0.0541
				(0.1390)	(0.1422)	(0.1424)
p-value	0.8137	0.8857	0.7410	0.7034	0.5860	0.7076
Control Mean	0.5643	0.5643	0.5643	0.5643	0.5643	0.5643
MSA Linear Trend						
		Yes			Yes	
Fixed-e ects						
MSA	Yes	Yes	Yes	Yes	Yes	Yes
Youngest Birth Cohort	Yes	Yes	Yes	Yes	Yes	Yes
PNADC Year	Yes	Yes	Yes	Yes	Yes	Yes
PNADC Year-Youngest Birth Cohort	Yes	Yes	Yes	Yes	Yes	Yes
MSA-Youngest Birth Cohort			Yes			Yes
Fit statistics						
Observations	1,956	1,956	1,956	1,956	1,956	1,956
MSA Clusters	22	22	22	22	22	22

Clustered (MSA) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 6 – Affirmative Action and Assortative Matching of Individuals Who Attended Private High School

Dependent Variable:	Both Private					
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
AA Exposure Binary	-0.0209	-0.0249	-0.0238			
	(0.0433)	(0.0430)	(0.0409)			
AA Exposure Continuous				-0.1752	-0.1871	-0.1394
				(0.1352)	(0.1355)	(0.1398)
p-value	0.6342	0.5689	0.5667	0.2091	0.1819	0.3302
Control Mean	0.2498	0.2498	0.2498	0.2498	0.2498	0.2498
MSA Linear Trend						
		Yes			Yes	
Fixed-e ects						
MSA	Yes	Yes	Yes	Yes	Yes	Yes
Youngest Birth Cohort	Yes	Yes	Yes	Yes	Yes	Yes
PNADC Year	Yes	Yes	Yes	Yes	Yes	Yes
PNADC Year-Youngest Birth Cohort	Yes	Yes	Yes	Yes	Yes	Yes
MSA-Youngest Birth Cohort			Yes			Yes
Fit statistics						
Observations	1,956	1,956	1,956	1,956	1,956	1,956
MSA Clusters	22	22	22	22	22	22

Clustered (MSA) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Tables 7 and 8 present the assortative matching on ethnicity results. They provide evidence that marriage diversity decreases mainly due to increased assortative matching among whites with affirmative action exposure. While in Table 8, the point estimates change their sign according to the specification we use, the point estimates in Table 7 are all positive, being far more consistent. In this table, we can reject the null hypothesis of no effect in both models (4) and (6). Our main specification, model (6), shows that an increase of 1 percentage point in affirmative action exposure increases marriages among whites by 0.2103 percentage points.

Table 7 – Affirmative Action and Assortative Matching of Whites

Dependent Variable:	Both White					
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
AA Exposure Binary	0.0306	0.0257	0.0328			
	(0.0254)	(0.0258)	(0.0270)			
AA Exposure Continuous				0.1813	0.1635	0.2103
				(0.1002)	(0.1069)	(0.1175)
p-value	0.2420	0.3300	0.2367	0.0848	0.1409	0.0878
Control Mean	0.4183	0.4183	0.4183	0.4183	0.4183	0.4183
MSA Linear Trend						
		Yes			Yes	
Fixed-effects						
MSA	Yes	Yes	Yes	Yes	Yes	Yes
Youngest Birth Cohort	Yes	Yes	Yes	Yes	Yes	Yes
PNADC Year	Yes	Yes	Yes	Yes	Yes	Yes
PNADC Year-Youngest Birth Cohort	Yes	Yes	Yes	Yes	Yes	Yes
MSA-Youngest Birth Cohort			Yes			Yes
Fit statistics						
Observations	1,956	1,956	1,956	1,956	1,956	1,956
MSA Clusters	22	22	22	22	22	22

Clustered (MSA) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 8 – Affirmative Action and Assortative Matching of Non-Whites

Dependent Variable:	Both Non White					
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
AA Exposure Binary	-0.0141	-0.0138	0.0074			
	(0.0152)	(0.0159)	(0.0181)			
AA Exposure Continuous				-0.0459	-0.0382	0.0645
				(0.0660)	(0.0707)	(0.0867)
p-value	0.3634	0.3944	0.6877	0.4944	0.5949	0.4652
Control Mean	0.3108	0.3108	0.3108	0.3108	0.3108	0.3108
MSA Linear Trend						
		Yes			Yes	
Fixed-e ects						
MSA	Yes	Yes	Yes	Yes	Yes	Yes
Youngest Birth Cohort	Yes	Yes	Yes	Yes	Yes	Yes
PNADC Year	Yes	Yes	Yes	Yes	Yes	Yes
PNADC Year-Youngest Birth Cohort	Yes	Yes	Yes	Yes	Yes	Yes
MSA-Youngest Birth Cohort			Yes			Yes
Fit statistics						
Observations	1,956	1,956	1,956	1,956	1,956	1,956
MSA Clusters	22	22	22	22	22	22

Clustered (MSA) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

6 Sensitivity Analysis

Our results rely on the untestable assumption that no omitted variable bias exists in our empirical strategy to identify the causal effect of AA exposure on marriage diversity. In this setting, an important exercise is to examine how fragile our results are against the possibility of unobserved confounding. In this section, we conduct a sensitivity analysis using the tools developed by Cinelli and Hazlett (2019).

In their article, the authors recommend that researchers report the robustness values, RV and RV_α. These values provide a reference point to assess the robustness of the coefficient of interest and its significance level, respectively, to unobserved confounders. If the confounders' association with the treatment and the outcome - given by $R^2_{D \sim Z|X}$ and $R^2_{Y \sim Z|X,D}$, respectively - are both assumed to be less than then RV then such confounder cannot bring the point estimate of interest to zero.[1] Analogously, the same is valid for the RV_α, but in this case, the confounder cannot make the coefficient statistically different from zero at a significance level α. We may interpret these summary statistics as the strength necessary for a confounder to have for its existence to be problematic and fully explain our regression results. Furthermore, the authors also suggest reporting the proportion of variation in the outcome uniquely explained by the treatment, i.e., the R^2 of the treatment with the outcome ($R^2_{Y \sim D|X}$). This statistic reflects how strong the association of the confounder with the treatment must be to eliminate all the observed coefficient effect in an extreme scenario where the confounder explains all the residual variation of the outcome.

Judging whether confounders with such associations are plausible is a practical difficulty. To address this concern, the authors suggest using relative claims, i.e., we reason about an unobserved confounder strength plausibility by analyzing how many times it must have a stronger association than an observed (group of) covariate(s) important to the research context to explain away all the observed effects of the variable of interest.

In our research setup, we face the scarcity of available covariates to provide bounds on the plausible strength of the unobserved confounder. No sole covariate can explain both the variation of treatment assignment and the outcome. Given such restriction, we use a group of covariates that - when used together - may capture the variations of interest.

[1] To be clear, we let Y be the outcome variable, D be the treatment variable, Z be the confounder, and X be the set of covariates.

We use the birth cohort of the youngest member of the couple to explain variation in treatment assignment. This choice relies on the fact that affirmative actions increased from 2004 to 2015, so younger individuals are more susceptible to AA exposure, and hence, the birth cohort may capture the variation in treatment. Still, we need covariates that explain the variation in the outcome. In this context, social class, cultural values, and religion may influence individuals' relationships and can be closely related to ethnicity. Thus, ethnicity might be important in explaining marriages among individuals who attended different high school types. On the other hand, the type of high school individuals attended might be important in explaining marriage with ethnic diversity by a similar reason in which AA exposure might affect ethnically diverse marriages; that is, since public schools tend to be more ethnically diverse than private ones, the type of high school attended might influence the social networks individuals attain and, in turn, their interaction with people of distinct ethnicities.

Considering those arguments, we use in our regressions variables that we previously used to build our outcome of interest: the type of high school and the ethnicity of the youngest (when analyzing ethnicity and high school diversity, respectively). Therefore, the benchmark covariates we use lie on binary variables that assign the birth cohort of the youngest member of the couple and whether she is white or attended a public high school, depending on the outcome of interest. Moreover, we must answer how strong potential confounders are compared to our benchmark covariates to make sense of our sensitivity analyses. Unfortunately, this question has no definitive answer. However, imagining an unobserved confounder just as strong as the benchmark covariates in explaining the variation in the outcome or the treatment variable is no easy task. Thus, we adopt the benchmark covariates' strengths as a reasonably higher bound for the unobserved confounder strengths.

Table 9 presents the results of the effects of AA exposure on marriage diversity when using our main specifications with the continuous treatment definition and controlling for the type of ethnicity and type of high school of the youngest member of the couple. In these regressions, we follow Cinelli and Hazlett (2019) and use i.i.d. standard errors. Compared to their previous analogous regressions, including controls does not change the point estimate of high school diversity and has almost no effect on changing the point estimate of the ethnic diversity variable. In columns (6) of Table 4, we get that an increase of one percentage point in AA exposure decreases ethnically diverse marriages by 0.27

percentage points. By including the controls, column (2) of Table 9 present that for the same amount of increase in AA exposure, the effect is a decrease of 0.28 percentage points for ethnically diverse marriages. Moreover, the coefficients of the ethnicity of the youngest member and her type of high school attended are large. The former is about 0.03, and the latter is 0.06. This finding is evidence that these covariates are important to explain the variation of the outcomes of interest in our research context.

Table 9 – Affirmative Action and Marriage Diversity Including Controls

Dependent Variables:	HS Diversity	Ethnic Diversity
Model:	(1)	(2)
Variables		
AA Exposure Continuous	0.1935	-0.2814
	(0.1168)	(0.1315)
Youngest is White	0.0316	
	(0.0208)	
Youngest Attend a Public HS		0.0625
		(0.0236)
Fixed-e ects		
MSA	Yes	Yes
Youngest Birth Cohort	Yes	Yes
PNADC Year	Yes	Yes
PNADC Year-Youngest Birth Cohort	Yes	Yes
MSA-Youngest Birth Cohort	Yes	Yes
Fit statistics		
Observations	1,956	1,956

IID standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

The sensitivity analyses for high school diversity are in Table 10. Panel A shows that the robustness value for the coefficient is 4.01% and for the significance level is 0.08%. The partial R^2 of the treatment with the outcome is 0.17%. The first summary statistic points that unobserved confounders (orthogonal to the covariates) that explain more than 4% of the residual variance of both the treatment and the outcome are strong enough to bring the point estimate to 0. Analogously, the second summary statistic points that unobserved cofounders that explain more than 0.08% of the residual variance of both the treatment and the outcome are strong enough to bring the estimate to a range where it is no longer 'statistically different' from 0 at the significance level of $\alpha = 10$. The third summary statistic indicates that an extreme confounder (orthogonal to the covariates) that explains 100% of the residual variance of the outcome would need to explain at least 0.17% of the residual variance of the treatment to account for the estimated effect fully.

Panel B of Table 10 let us assess whether these values of the summaries statistics are high or low by presenting the association of the unobserved confounder to the treatment and the outcome variable for different strength levels when compared to the benchmark covariates, in which this case are the ethnicity of the youngest member of the couple and dummy variables indicating which year she was born. Note that if the association of a confounder is three times as stronger as our benchmark covariates, it would still be far from the RV, indicating that it could not eliminate the effect. However, the significance level is considerably more sensitive than the estimate. A confounder as strong as our chosen benchmark covariates could make it no longer significant at 10%. Furthermore, considering the extreme scenario, a confounder as strong as our chosen benchmark covariates could bring the point estimate to zero.

Table 10 – Sensitivity Analyses High School Diversity

A. Sensitivity Statistics		
Robustness Value	4.01%	
Robustness Value, $\alpha = 10\%$	0.08%	
Partial R^2 of Treatment With Outcome	0.17%	
B. Bounds	$R^2_{D\ Z\mid X}$	$R^2_{Y\ Z\mid X,D}$
1x White and Birth Cohort Dummies	0.88%	0.69%
2x White and Birth Cohort Dummies	1.76%	1.38%
3x White and Birth Cohort Dummies	2.64%	2.07%

Table 11 presents the sensitivity analyses for ethnic diversity. Panel A indicates that the RV is 5.15%, the RV_α is 1.27%, and the partial R^2 of the treatment with the outcome is 0.28%. Panel B provides the R^2 with the outcome and the treatment for benchmark covariates, which are the type of high school of the couple's youngest member and dummy variables indicating the year she was born. As before, a confounder three times as strong as our benchmark covariates could not bring the point estimate to zero since its association with the treatment and the outcome would be below the robustness value. Moreover, our variable of interest would remain significant at 10% if there were a confounder just as stronger as the benchmark covariates. However, our point estimate would lose significance at the conventional levels if the confounder was twice as strong. Finally, considering the extreme scenario, a confounder as strong as our chosen benchmark covariates could bring the point estimate to zero.

Table 11 – Sensitivity Analyses Ethnic Diversity

A. Sensitivity Statistics		
Robustness Value	5.15%	
Robustness Value, $\alpha = 10\%$	1.27%	
Partial R^2 of Treatment With Outcome	0.28%	
B. Bounds	$R^2_{D\ Z\mid X}$	$R^2_{Y\ Z\mid X,D}$
1x Public HS and Birth Cohort Dummies	0.91%	1.22%
2x Public HS and Birth Cohort Dummies	1.83%	2.44%
3x Public HS and Birth Cohort Dummies	2.74%	3.66%

In summary, considering that the unobserved confounder cannot be stronger than the benchmark covariates, we got evidence that the estimate of our variable of interest when analyzing the high school inter-type and the interethnic marriage is robust to unobserved confounding when not considering an extreme scenario. However, we should not take these results for granted due to the sensitivity of our significance level in the case of high school inter-type marriages and the sensitivity of both point estimates when considering an extreme scenario in which the unobserved confounder explains all the residual variation of the outcome.

7 Conclusions

In this article, we investigate whether affirmative action policies implemented in Brazilian public universities affected marriage diversity, as measured by the high school inter-type (private/public) and interethnic marriages. We provide evidence that this policy affects which relationships individuals engage in. However, it has different effects depending on the type of diversity analyzed.

On the one hand, increasing by one percentage point in affirmative action policies exposure increases marriages among individuals with distinct types of high school by 0.19 percentage points. To understand the compositional changes in marriage diversity, we examine assortative matching on the type of high school. The results suggest that the increase in inter- type of high school marriages is mostly associated with the decrease in assortative matching on the type of high school among individuals who attended private high schools. On the other hand, increasing by one percentage point affirmative action policies exposure decreases marriages among individuals with distinct ethnicities by 0.27 percentage points. Examining assortative matching on ethnicity suggests that decreased interethnic marriage is most likely associated with increased assortative matching on ethnicity among white individuals.

Importantly, the heterogeneity of results across diversity dimensions shows that increasing exposure among groups does not necessarily lead to higher interaction. Understanding why distinct groups have distinct behaviors when exposed to affirmative action policies is crucial to comprehend the effects of such policies on social ties. In other words, it is essential to disentangle the channels in which affirmative actions affect inter-group contact. This is an important avenue for future research.

Bibliography

Artmann, E., N. Ketel, H. Oosterbeek, and B. van der Klaauw (2018). Field of study and family outcomes. (11658).

Billings, S. B., E. Chyn, and K. Haggag (2021, September). The long-run effects of school racial diversity on political identity. *American Economic Review: Insights 3*(3), 267–84.

Cavalcanti, T., J. Guimaraes, and B. Sampaio (2010, November). Barriers to Skill Acquisition in Brazil: Public and Private School Students Performance in a Public University Entrance Exam. *The Quarterly Review of Economics and Finance 50*(4), 395–407.

Cinelli, C. and C. Hazlett (2019, 12). Making Sense of Sensitivity: Extending Omitted Variable Bias. *Journal of the Royal Statistical Society Series B: Statistical Methodology 82*(1), 39–67.

Duflo, E. (2001, September). Schooling and labor market consequences of school construction in indonesia: Evidence from an unusual policy experiment. *American Economic Review 91*(4), 795–813.

Eika, L., M. Mogstad, and B. Zafar (2019). Educational assortative mating and household income inequality. *Journal of Political Economy 127*(6), 2795–2835.

Estevan, F., T. Gall, and L.-P. Morin (2019, May). Can Affirmative Action Affect Major Choice? (dp-324).

Francis, A. M. and M. Tannuri-Pianto (2012). Using brazil's racial continuum to examine the short-term effects of affirmative action in higher education. *The Journal of Human Resources 47*(3), 754–784.

Francis-Tan, A. and M. Tannuri-Pianto (2018). Black movement: Using discontinuities in admissions to study the effects of college quality and affirmative action. *Journal of Development Economics 135*, 97–116.

Gordon, M. M. (1964). *Assimilation in American life: The role of race, religion, and national origins*. Oxford University Press, USA.

Kirkebøen, L., E. Leuven, and M. Mogstad (2021, April). College as a marriage market. (28688).

Mello, U. (2022, August). Centralized admissions, affirmative action, and access of low-income students to higher education. *American Economic Journal: Economic Policy 14*(3), 166–97.

Merlino, L. P., M. F. Steinhardt, and L. Wren-Lewis (2019). More than just friends? school peers and adult interracial relationships. *Journal of Labor Economics 37*(3), 663–713.

Pereira, L. and C. Santos (2017). Casamentos seletivos e desigualdade de renda no brasil. *Revista Brasileira de Economia - RBE 71*(3).

Reardon, S. F. and K. Bischoff (2011). Income inequality and income segregation. *American Journal of Sociology 116*(4), 1092–1153.

Solon, G., S. J. Haider, and J. Wooldridge (2013, February). What are we weighting for? *National Bureau of Economic Research* (18859).

Telles, E. E. (2006). *Race in another America*. Princeton University Press.

Vieira, R. S. and M. Arends-Kuenning (2019). Affirmative action in brazilian universities: Effects on the enrollment of targeted groups. *Economics of Education Review 73*, 101931.

Appendix

APPENDIX A – Tables and Figures

A.1 Tables

Table A1 – Affirmative Action and Probability of Marriage

Dependent Variable:	Marriage					
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
AA Exposure Binary	-0.0177	-0.0209	-0.0244			
	(0.0229)	(0.0238)	(0.0257)			
AA Exposure Continuous				-0.0426	-0.0688	-0.0739
				(0.0911)	(0.0961)	(0.1147)
p-value	0.4481	0.3890	0.3534	0.6450	0.4820	0.5264
Weighted Control Mean	0.5516	0.5516	0.5516	0.5516	0.5516	0.5516
MSA Linear Trend						
		Yes			Yes	
Fixed-e ects						
MSA	Yes	Yes	Yes	Yes	Yes	Yes
Youngest Birth Cohort	Yes	Yes	Yes	Yes	Yes	Yes
PNADC Year	Yes	Yes	Yes	Yes	Yes	Yes
PNADC Year-Youngest Birth Cohort	Yes	Yes	Yes	Yes	Yes	Yes
MSA-Youngest Birth Cohort			Yes			Yes
Fit statistics						
Observations	3,768	3,768	3,768	3,768	3,768	3,768
MSA Clusters	22	22	22	22	22	22

Clustered (MSA) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A2 – Affirmative Action and High School Diversity

Dependent Variable:	HS Diversity					
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
AA Exposure Binary	0.0219	0.0229	0.0288			
	(0.0192)	(0.0203)	(0.0178)			
AA Exposure Continuous				0.1268	0.1163	0.1925
				(0.0931)	(0.0987)	(0.0951)
p-value	0.2673	0.2729	0.1203	0.1878	0.2518	0.0558
Weighted Control Mean	0.1768	0.1768	0.1768	0.1768	0.1768	0.1768
MSA Linear Trend						
		Yes			Yes	
Fixed-e ects						
MSA	Yes	Yes	Yes	Yes	Yes	Yes
Youngest Birth Cohort	Yes	Yes	Yes	Yes	Yes	Yes
PNADC Year	Yes	Yes	Yes	Yes	Yes	Yes
PNADC Year-Youngest Birth Cohort	Yes	Yes	Yes	Yes	Yes	Yes
MSA-Youngest Birth Cohort			Yes			Yes
Fit statistics						
Observations	1,956	1,956	1,956	1,956	1,956	1,956
MSA Clusters	22	22	22	22	22	22

Clustered (MSA) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A3 – Affirmative Action and Ethnic Diversity

Dependent Variable:	Ethnic Diversity					
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
AA Exposure Binary	0.0027	0.0125	-0.0047			
	(0.0411)	(0.0435)	(0.0430)			
AA Exposure Continuous				-0.1762	-0.1347	-0.2717
				(0.1060)	(0.1240)	(0.1076)
p-value	0.9477	0.7769	0.9136	0.1111	0.2897	0.0197
Weighted Control Mean	0.2699	0.2699	0.2699	0.2699	0.2699	0.2699
MSA Linear Trend						
		Yes			Yes	
Fixed-e ects						
MSA	Yes	Yes	Yes	Yes	Yes	Yes
Youngest Birth Cohort	Yes	Yes	Yes	Yes	Yes	Yes
PNADC Year	Yes	Yes	Yes	Yes	Yes	Yes
PNADC Year-Youngest Birth Cohort	Yes	Yes	Yes	Yes	Yes	Yes
MSA-Youngest Birth Cohort			Yes			Yes
Fit statistics						
Observations	1,956	1,956	1,956	1,956	1,956	1,956
MSA Clusters	22	22	22	22	22	22

Clustered (MSA) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A4 – Affirmative Action and Assortative Matching of Individuals Who Attended Public High School

Dependent Variable:	Both Public					
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
AA Exposure Binary	-0.0054	-0.0008	-0.0110			
	(0.0322)	(0.0318)	(0.0311)			
AA Exposure Continuous				0.0201	0.0458	-0.1082
				(0.1286)	(0.1299)	(0.1283)
p-value	0.8679	0.9806	0.7262	0.8772	0.7279	0.4084
Weighted Control Mean	0.5787	0.5787	0.5787	0.5787	0.5787	0.5787
MSA Linear Trend						
		Yes			Yes	
Fixed-e ects						
MSA	Yes	Yes	Yes	Yes	Yes	Yes
Youngest Birth Cohort	Yes	Yes	Yes	Yes	Yes	Yes
PNADC Year	Yes	Yes	Yes	Yes	Yes	Yes
PNADC Year-Youngest Birth Cohort	Yes	Yes	Yes	Yes	Yes	Yes
MSA-Youngest Birth Cohort			Yes			Yes
Fit statistics						
Observations	1,956	1,956	1,956	1,956	1,956	1,956
MSA Clusters	22	22	22	22	22	22

Clustered (MSA) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A5 – Affirmative Action and Assortative Matching of Individuals Who Attended Private High School

Dependent Variable:	Both Private					
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
AA Exposure Binary	-0.0165	-0.0221	-0.0178			
	(0.0278)	(0.0263)	(0.0272)			
AA Exposure Continuous				-0.1469	-0.1621	-0.0843
				(0.1131)	(0.1066)	(0.1294)
p-value	0.5604	0.4091	0.5205	0.2079	0.1433	0.5218
Weighted Control Mean	0.2444	0.2444	0.2444	0.2444	0.2444	0.2444
MSA Linear Trend						
		Yes			Yes	
Fixed-e ects						
MSA	Yes	Yes	Yes	Yes	Yes	Yes
Youngest Birth Cohort	Yes	Yes	Yes	Yes	Yes	Yes
PNADC Year	Yes	Yes	Yes	Yes	Yes	Yes
PNADC Year-Youngest Birth Cohort	Yes	Yes	Yes	Yes	Yes	Yes
MSA-Youngest Birth Cohort			Yes			Yes
Fit statistics						
Observations	1,956	1,956	1,956	1,956	1,956	1,956
MSA Clusters	22	22	22	22	22	22

Clustered (MSA) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A6 – Affirmative Action and Assortative Matching of Whites

Dependent Variable:	Both White					
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
AA Exposure Binary	0.0258	0.0222	0.0291			
	(0.0338)	(0.0340)	(0.0304)			
AA Exposure Continuous				0.2405	0.2211	0.2842
				(0.1046)	(0.1093)	(0.1049)
p-value	0.4539	0.5213	0.3498	0.0318	0.0561	0.0131
Weighted Control Mean	0.4574	0.4574	0.4574	0.4574	0.4574	0.4574
MSA Linear Trend						
		Yes			Yes	
Fixed-e ects						
MSA	Yes	Yes	Yes	Yes	Yes	Yes
Youngest Birth Cohort	Yes	Yes	Yes	Yes	Yes	Yes
PNADC Year	Yes	Yes	Yes	Yes	Yes	Yes
PNADC Year-Youngest Birth Cohort	Yes	Yes	Yes	Yes	Yes	Yes
MSA-Youngest Birth Cohort			Yes			Yes
Fit statistics						
Observations	1,956	1,956	1,956	1,956	1,956	1,956
MSA Clusters	22	22	22	22	22	22

Clustered (MSA) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A7 – Affirmative Action and Assortative Matching of Non-Whites

Dependent Variable:	Both Non White					
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
AA Exposure Binary	-0.0285	-0.0347	-0.0244			
	(0.0248)	(0.0256)	(0.0283)			
AA Exposure Continuous				-0.0643	-0.0864	-0.0124
				(0.0989)	(0.1118)	(0.1330)
p-value	0.2622	0.1909	0.3986	0.5227	0.4486	0.9264
Weighted Control Mean	0.2727	0.2727	0.2727	0.2727	0.2727	0.2727
MSA Linear Trend						
		Yes			Yes	
Fixed-e ects						
MSA	Yes	Yes	Yes	Yes	Yes	Yes
Youngest Birth Cohort	Yes	Yes	Yes	Yes	Yes	Yes
PNADC Year	Yes	Yes	Yes	Yes	Yes	Yes
PNADC Year-Youngest Birth Cohort	Yes	Yes	Yes	Yes	Yes	Yes
MSA-Youngest Birth Cohort			Yes			Yes
Fit statistics						
Observations	1,956	1,956	1,956	1,956	1,956	1,956
MSA Clusters	22	22	22	22	22	22

Clustered (MSA) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

A.2 Figures

Figure A1 – Age Of The Individual In Each Year According To Her Birth Cohort

	AGE OF THE INDIVIDUAL IN EACH YEAR ACCORIDNG TO HER BIRTH COHORT																			
	1985	1986	1987	...	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
BIRTH COHORT																				
1984	1	2	3	...	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35
1985	-	1	2	...	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34
1986	-	-	-	...	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33
1987	-	-	-	...	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
1988	-	-	-	...	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
1989	-	-	-	...	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
1990	-	-	-	...	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
1991	-	-	-	...	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28
1992	-	-	-	...	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
1993	-	-	-	...	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
1994	-	-	-	...	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1995	-	-	-	...	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1996	-	-	-	...	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1997	-	-	-	...	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1998	-	-	-	...	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21

Eligible individuals for our sample are those within the space outlined by the black line. However, as we only observed data from 2016 to 2019, we only observed individuals in the space highlighted in blue. Importantly, we observe individuals from the same birth cohort at different ages.