

idp

v.3 n.3

72

DEBATES EM ECONOMIA APLICADA

WORKING PAPER

**ANCESTRALIDADE, SOBRENOMES E DISCRIMINAÇÃO:
EVIDÊNCIAS A PARTIR DO JOGO ALEMANHA 7X1
BRASIL**

CAROLINA F. CUSTODIO
LEONARDO M. MONASTERIO

ANCESTRALIDADE, SOBRENOMES E DISCRIMINAÇÃO: EVIDÊNCIAS A PARTIR DO JOGO ALEMANHA 7X1 BRASIL

CAROLINA F. CUSTODIO¹
LEONARDO M. MONASTERIO²

¹ Mestre em Economia pelo Instituto Brasileiro de Ensino, Desenvolvimento e Pesquisa (IDP). E-mail: carolinafcustodio@gmail.com. ORCID: 0000-0001-9616-8825.

² Pesquisador no Instituto de Pesquisa Econômica Aplicada (IPEA) e professor no Instituto Brasileiro de Ensino, Desenvolvimento e Pesquisa (IDP). ORCID: 0000-0001-6692-3236. Os autores agradecem a Daniel Lopes (CNI), Henrique Reichert (UFSC), Mathias Schneid Tessmann (IDP), Patricia Bonini (UFSC) e Roberta Moreira Wichmann (IDP), por sugestões e comentários ao longo do desenvolvimento deste trabalho. Os erros são nossos.

IDP

O IDP é um centro de excelência no ensino, na pesquisa e na extensão nas áreas da Administração Pública, Direito e Economia. O Instituto tem como um de seus objetivos centrais a profusão e difusão do conhecimento de assuntos estratégicos nas áreas em que atua, constituindo-se um *think tank* independente que visa contribuir para as transformações sociais, políticas e econômicas do Brasil.

DIREÇÃO E COORDENAÇÃO

Diretor Geral

Francisco Schertel

Coordenador do Mestrado em Economia

José Luiz Rossi

CONSELHO EDITORIAL

Coordenação

Thiago Caldeira

Renan Holtermann

Milton Mendonça

Supervisão e Revisão

Luiz Augusto Magalhães

Mathias Tessmann

Apoio Técnico

Igor Silva

Projeto Gráfico e Diagramação

Juliana Vasconcelos

www.idp.edu.br

Revista Técnica voltada à divulgação de resultados preliminares de estudos e pesquisas aplicados em desenvolvimento por professores, pesquisadores e estudantes de pós-graduação com o objetivo de estimular a produção e a

DEBATES EM ECONOMIA APLICADA

discussão de conhecimentos técnicos relevantes na área de Economia.

Convidamos a comunidade acadêmica e profissional a enviar comentários e críticas aos autores, visando o aprimoramento dos trabalhos para futura publicação. Por seu propósito se concentrar na recepção de comentários e críticas, a Revista Debates em Economia Aplicada não possui ISSN e não fere o ineditismo dos trabalhos divulgados.

As publicações da Revista estão disponíveis para acesso e download gratuito no formato PDF. Acesse: www.idp.edu.br

As opiniões emitidas nesta publicação são de exclusiva e inteira responsabilidade dos autores, não exprimindo, necessariamente, o ponto de vista do IDP.

Qualquer citação aos trabalhos da Série só é permitida mediante autorização expressa do(s) autor(es).

debates em economia aplicada

SUMÁRIO

1. Introdução	7
2. Estratégia empírica	12
3. Dados	15
4. Resultados	19
5. Conclusão	28
6. Referências	29

RESUMO: Este artigo analisa os mecanismos da discriminação com base na ancestralidade dos trabalhadores brasileiros. O estudo utiliza informações de sobrenomes para identificar grupos de ancestralidades e examina se houve discriminação em favor de indivíduos com ancestralidade germânica após a derrota da seleção brasileira de futebol na Copa do Mundo de 2014. Limitando a nossa análise nos trabalhadores que apenas entraram no mercado de trabalho formal em 2014, antes e depois do jogo em questão, estimamos regressões mincerianas com controles baseados na ancestralidade indicada pelos sobrenomes e modelos probit para a obtenção do primeiro emprego. Os resultados mostram que a vitória da Alemanha teve um impacto positivo nos salários dos trabalhadores com sobrenomes germânicos, mas também resultou em uma redução na probabilidade desses trabalhadores serem contratados. Foram observados impactos menores para trabalhadores com sobrenomes italianos, ibéricos e japoneses, com efeitos positivos para os primeiros e negativo para os japoneses. O estudo contribui para uma compreensão mais ampla dos efeitos econômicos de eventos socioculturais e se relaciona com pesquisas anteriores que exploram os impactos da ancestralidade nos salários e na probabilidade de contratação no mercado de trabalho.

PALAVRAS-CHAVE: Ancestralidade; Salário; Mercado de trabalho.

ABSTRACT: This article analyzes the mechanisms of discrimination based on the ancestry of Brazilian workers. The study uses surname information to identify ancestry groups and examines whether there was discrimination in favor of individuals with Germanic ancestry following the defeat of the Brazilian soccer team in the 2014 World Cup. By limiting our analysis to workers who only entered the formal labor market in 2014, before and after the game in question, we estimate Mincerian regressions with controls based on the ancestry indicated by surnames and probit models for obtaining the first job. The results show that Germany's victory had a positive impact on the wages of workers with Germanic surnames, but it also resulted in a reduction in the likelihood of these workers being hired. Smaller impacts were observed for workers with Italian, Iberian, and Japanese surnames, with positive effects for the former and negative for the Japanese. The study contributes to a broader understanding of the economic effects of sociocultural events and relates to previous research exploring the impacts of ancestry on wages and the likelihood of hiring in the labor market.

KEYWORDS: Ancestry; Wage; Labor market.

RESUMEN: Este artículo analiza los mecanismos de la discriminación basada en la ancestralidad de los trabajadores brasileños. El estudio utiliza información de apellidos para identificar grupos de ancestralidades y examina si hubo discriminación a favor de individuos con ascendencia germánica después de la derrota de la selección brasileña de fútbol en la Copa del Mundo de 2014. Limitando nuestra análisis a los trabajadores que recién ingresaron al mercado laboral formal en 2014, antes y después del juego en cuestión, estimamos

regresiones mincerianas con controles basados en la ascendencia indicada por los apellidos y también modelos probit para obtener el primer empleo. Los resultados muestran que la victoria de Alemania tuvo un impacto positivo en los salarios de los trabajadores con apellidos germánicos, pero también resultó en una reducción en la probabilidad de que esos trabajadores fueran contratados. Se observaron impactos menores para trabajadores con apellidos italianos, ibéricos y japoneses, con efectos positivos para los primeros y negativo para los japoneses. El estudio contribuye a una comprensión más amplia de los efectos económicos de eventos socioculturales y se relaciona con investigaciones anteriores que exploran los impactos de la ascendencia en los salarios y en la probabilidad de contratación en el mercado laboral.

PALABRAS CLAVE: Ascendencia; Salario; Mercado laboral.

CLASSIFICAÇÃO JEL: J15; J31

1. INTRODUÇÃO

Qual a raiz das desigualdades salariais entre indivíduos? A teoria do capital humano, a partir das contribuições seminais de Mincer (1958), Schultz (1961), Becker e Chiswick (1966) e Arrow et al. (1973), costuma ser o ponto de partida das análises econômicas desse fenômeno. Essa abordagem sustenta que os investimentos realizados em educação e qualificação ampliam a capacidade produtiva dos indivíduos, fazendo com que resultem em distintos salários no mercado de trabalho. Assim, em seu modelo mais simples, as diferenças salariais são resultantes das decisões individuais de acumulação de capital humano que, em última instância, decorreriam das diferenças das preferências (Becker, 1962).

Mas há abundantes evidências de que as diferenças salariais vão muito além das discrepâncias entre as dotações de capital humano. De modo a buscar explicar essas desigualdades, surgem abordagens como a teoria da discriminação estatística com trabalhos de Phelps (1972) e Oaxaca e Dickinson (2006), Oaxaca (1973) e Blinder (1973). Assim, a análise da determinação e das discrepâncias salariais dos trabalhadores pode ser conduzida com base em fundamentos teóricos que consideram as variações nos atributos produtivos, a segmentação do mercado de trabalho e a discriminação econômica e estatística.

Na teoria econômica da discriminação, os trabalhadores são tratados de forma desigual com base em características irrelevantes, apenas por pertencerem a determinados grupos (Becker, 1962; Becker e Chiswick, 1966). Dessa forma, a discriminação surge quando um conjunto de indivíduos apresenta habilidades, treinamento, educação, produtividade e experiência semelhantes, mas são remunerados e tratados de forma discrepante em virtude de algum fator não econômico, tal como a cor de pele, gênero e ancestralidade.

A questão da desigualdade salarial no Brasil é um tema amplamente debatido na sociedade e na academia. No entanto, apesar de existir uma vasta literatura sobre a desigualdade salarial, a discussão tem sido predominantemente focada na dimensão da cor, uma vez que registros oficiais no Brasil tradicionalmente utilizam categorias de cor e raça,

como o Instituto Brasileiro de Geografia e Estatística (IBGE), não versando sobre informações de ancestralidade da população.

Em uma abordagem alternativa ao padrão cor/raça, Monasterio (2017) utiliza a informação do sobrenome dos indivíduos para definir esta ancestralidade, o que permite investigações mais precisas sobre diferentes grupos. Dado que existem discrepâncias salariais e de oportunidades de trabalho devido a questões de gênero e cor, é possível que também exista desigualdade a partir da ancestralidade dos indivíduos, afetando suas chances de acesso a oportunidades de trabalho e, conseqüentemente, influenciando seus resultados no mercado de trabalho. Em outras palavras, uma discriminação de ancestralidade no mercado de trabalho pode ser descrita como uma visão viesada por parte dos agentes tomadores de decisão, que tenderiam a optar por indivíduos pertencentes a determinado grupo na expectativa que eles apresentem maior produtividade do trabalho, mesmo não existindo fatores tangíveis para respaldar esta expectativa ou por uma simples preferência pela discriminação.

A discussão do fator ancestralidade ainda é bastante incipiente, como aparece em Monasterio (2017), que apresenta um método para classificar a ancestralidade de sobrenomes brasileiros com base em fontes históricas. O autor apresenta um algoritmo de classificação de sobrenomes que serve como complemento para a categorização da população brasileira contemporânea em grupos ancestrais, diferente do padrão frequentemente empregado de cor/raça. Vale notar que apesar do algoritmo ter sido ampliado e aprofundado por Lopes (2017) e Lopes et al. (2021), mas - tal como a própria categorização oficial do IBGE - é imperfeita e, obviamente, não incorpora todos os aspectos da questão.

A identificação de desigualdade salarial por ancestralidade enseja uma outra questão: qual a sua origem? A explicação mais simples seria que isso decorreria de fatores não observáveis, omitidos nas regressões. Ou seja, nesse caso, não se trataria propriamente discriminação, mas apenas o resultado de variáveis de ancestralidade capturarem o efeito de

dimensões produtivas, como as diferenças na qualidade da educação formal recebida ou outras dimensões comportamentais. Outro mecanismo que explicaria a discriminação, seria a existência de estereótipos racistas por parte dos empregadores que premiariam/puniriam indivíduos.

Em termos empíricos, a dificuldade é separar a atuação desses mecanismos com uma estratégia de identificação econométrica. Neste trabalho, buscou-se identificar um choque exógeno que aumentasse a visão preconcebida sobre uma ancestralidade específica do trabalhador no mercado de trabalho brasileiro. Nesse sentido, investiga-se se a fatídica eliminação da seleção brasileira de futebol na Copa do Mundo de 2014 disparou um efeito de discriminação em favor dos indivíduos com ancestralidade germânica no mercado de trabalho. Dada a inesperada superioridade da seleção alemã frente ao time brasileiro, pode ter havido um choque positivo - sem fundamento produtivo no prestígio dos brasileiros que entravam no mercado de trabalho.

Na tentativa de identificar esse mecanismo, serão estimados os bônus salariais de indivíduos no primeiro emprego a partir de sua ancestralidade - dada pelo último sobrenome - no período de contratação (antes ou depois da Copa do Mundo de 2014). Para tanto, utiliza-se dos dados identificados da Relação Anual de Informações Sociais (RAIS) daquele ano. Optou-se por avaliar apenas os indivíduos de primeiro emprego para eliminar fatores subjetivos de histórico laboral do indivíduo, normalmente atribuído pela experiência e tempo de emprego.

Além da contribuição de avaliar o papel da ancestralidade nas decisões dos agentes, o trabalho objetiva contribuir para uma compreensão mais ampla dos efeitos econômicos causados por eventos não-econômicos. Neste sentido, este estudo está relacionado com os papers de Rubinstein e Brenner (2014), Arai e Skogman Thoursie (2009) e Lopes et al. (2021), que avaliaram os impactos da ancestralidade sobre o mercado de trabalho, bem como Carvalho et al. (2017), Kaplanski e Levy (2010) e Depetris-Chauvin et al. (2020), que

analisaram os efeitos de eventos esportivos no mercado financeiro e no sentimento dos indivíduos.

O estudo israelense de Rubinstein e Brenner (2014) analisou a discriminação nos salários a partir de indivíduos que tiveram casamentos interétnicos entre sobrenomes sefarditas e ashkenazim³. A partir de regressões mincerianas e de Diferenças em Diferenças, descobriu-se que os trabalhadores ashkenazim ganham significativamente menos do que seus colegas sefarditas⁴.

O estudo de Arai e Skogman Thoursie (2009) avaliou o efeito da mudança de sobrenomes que soam suecos sobre os rendimentos de imigrantes asiáticos, africanos e eslavos. Usando um modelo de efeitos fixos com base nos dados individuais em painel anual para o período de 1990 a 2001, verificou-se um aumento substancial nas remunerações após a mudança de sobrenome.

No Brasil, Lopes et al. (2021) estimaram o impacto da cultura familiar nos resultados educacionais de brasileiros. Com uso do algoritmo de classificação de sobrenomes do aluno, de seus pais e avós, foram identificados grupos de ascendência ibérica, japonesa, italiana, alemã, do Leste Europeu e sírio-libanesa e verificado que alunos com ascendência não-ibérica obtêm taxas de aprovação e pontuações estatisticamente e substancialmente mais altas, mesmo após a inclusão de diversos controles estatísticos

No que tange aos eventos esportivos, Carvalho et al. (2017) avaliaram o efeito da derrota do Brasil por 7 a 1 para a Alemanha na Copa do Mundo de 2014 no mercado de ações brasileiro. Kaplanski e Levy (2010) estimaram que o retorno médio no mercado americano

³ Historicamente, judeus da França e Alemanha ficaram conhecidos como “ashkenazim” (palavra hebraica para “alemão”) e os judeus da Espanha ficaram conhecidos como “sefaradim” (palavra hebraica para “espanhol”).

⁴ Os resultados também apontam que a discriminação com base no sobrenome ocorre apenas quando esses nomes fornecem informações adicionais sobre etnia. Quando a etnia pode ser percebida a partir do tom de pele, os sobrenomes não fornecem poder explicativo adicional com relação aos salários.

durante o período de efeito da Copa do Mundo é de -2,58%, em comparação com +1,21% para os retornos médios de todos os dias no mesmo período⁵.

Ainda em um contexto futebolístico, Depetris-Chauvin et al. (2020) revelaram impactos de eventos esportivos no sentimento de orgulho nacional e na identificação étnica. Combinando dados de pesquisa de opinião individual com dados de desempenho de equipes nacionais de futebol em competições internacionais, incluindo modelos de regressão linear, Probit e Diferenças em Diferenças, notou-se que os indivíduos são menos propensos a relatar forte senso de identidade étnica nos dias imediatamente após uma vitória da seleção nacional de futebol, em comparação ao momento imediatamente antes. Os efeitos foram consideráveis, com uma diminuição de 37% na probabilidade média de autoidentificação étnica e parecem ser impulsionados por um aumento genuíno do orgulho nacional, em vez de euforia genérica pós-vitória. A vitória também reduziu a violência civil em até 9% nos meses seguintes.

Esses resultados revelam que a identidade étnica é um fator relevante nas decisões econômicas que pode sofrer algum tipo de discriminação. Além disso, os estudos mostram que existem implicações importantes sobre o comportamento dos indivíduos decorrentes de eventos de natureza não econômica, como jogos e competições.

O trabalho está dividido em cinco seções, sendo a introdução a primeira destas, apresentando a literatura que fundamenta e discute o papel da cor/raça, cultura e ancestralidade nos salários. Na seção seguinte é apresentada a estratégia empírica com a metodologia utilizada. Na terceira seção é explicado o tratamento de dados empregado e posteriormente a análise dos resultados, para a avaliação dos salários e probabilidade de contratação. Adiante, na quinta seção, serão apresentadas as considerações finais, seguidamente pelas referências utilizadas.

⁵ A estratégia se traduz em um movimento de *short selling*, que trata da venda a descoberto pouco antes do início do evento, o que só intensificará a queda dos preços das ações devido ao evento e também pode induzir a queda a começar mais cedo, mesmo antes do início dos jogos. Essa prática funciona como uma aposta na queda dos mercados e é utilizada para se proteger contra fases temporárias de declínio, como por exemplo o evento analisado.

2. ESTRATÉGIA EMPÍRICA

Para este trabalho, avalia-se o efeito do impacto da derrota por 7 a 1 sofrida pelo Brasil na semifinal da Copa do Mundo de 2014 contra a Alemanha como um evento de corte com possível efeitos sobre os salários de pessoas com ancestralidade germânica no mercado de trabalho brasileiro.

Cabe reforçar que o número de gols sofridos pela equipe brasileira não é algo comum em eventos desta magnitude, principalmente se tratando das etapas finais da competição e de duas equipes consideradas de alto nível. No começo do campeonato, o Brasil era o favorito nos mercados de apostas: na casa de apostas William Hill, a maior da Inglaterra, uma vitória do Brasil pagaria quatro libras para cada uma apostada. Já as vitórias de Argentina ou da Alemanha, vistas como mais improváveis, pagariam cinco e sete libras respectivamente (Caulyt, 2014).

O jogo contra a Alemanha que aconteceu no dia 8 de julho de 2014 no estádio do Mineirão, em Belo Horizonte, foi uma das derrotas mais contundentes da história da Copa do Mundo, e uma das maiores vexações esportivas do Brasil. O elástico placar de 7 a 1 da Alemanha sobre o Brasil surpreendeu também casas de apostas, onde somente 8 pessoas no mundo apostaram no improvável placar. Um apostador americano apostou \$20 dólares (cerca de R\$ 44 reais na cotação da época) e com o resultado, aumentou o seu investimento em 231.000% - recebendo cerca de R\$ 102 mil reais⁶.

Os alemães, todavia, jogaram um futebol de alto nível e venceram o jogo de forma decisiva. No que sucedeu o jogo, o elenco alemão se mostrou nobre no jogo e na vitória, parabenizando o Brasil pelo jogo e afirmando que a seleção brasileira deve ser respeitada

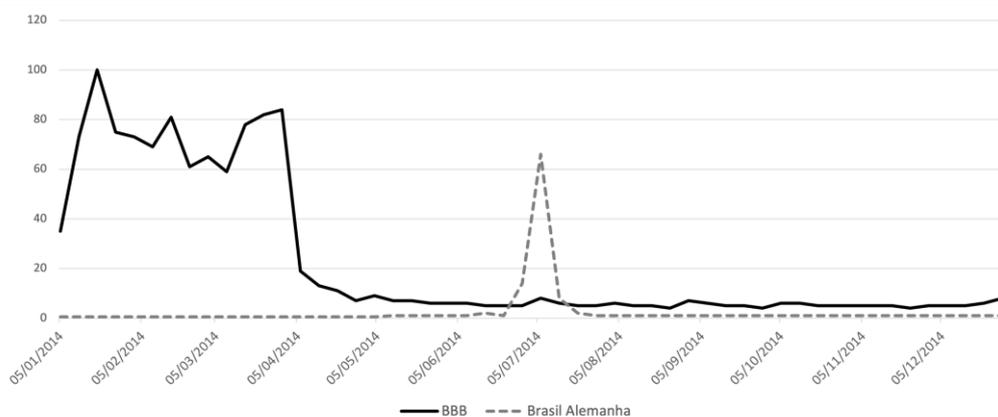
⁶ Notícia veiculada via UOL (2014) no dia seguinte da partida.

como uma grande equipe (UOL, 2015). A edição de Estado de Minas do dia seguinte do jogo sintetizou o sentimento nacional em relação ao time alemão:

“A pior derrota da centenária história da Seleção Brasileira foi algo como um confronto entre um time de garotos perdidos e desequilibrados tentando reagir ao vigoroso oponente atacando-o com uma pluma em vez de espada. Foi com a força de uma tempestade que a Alemanha destruiu o Brasil, no Mineirão (...) De algozes, os comandados de Joachim Löw foram ovacionados em diversos momentos. Os dois últimos gols foram aplaudidos de pé pelos brasileiros, que gritavam “olé” a cada troca de passe do adversário.” (Dias, 2018)

Entre os eventos mais buscados pelos brasileiros no Google, o interesse pelo jogo do Brasil e Alemanha ocorrido na semana do dia 6 de julho, assemelha-se ao reality Big Brother Brasil, termo mais buscado naquele ano, alcançando seu máximo de buscas em 19 de janeiro do mesmo ano, conforme dados da própria empresa.

Figura 1: Buscas no Google pelos eventos BBB e jogo Brasil e Alemanha



Fonte: Elaboração própria com base no Google Trends (2014).

Estes fatos reforçam a relevância de avaliar a temática, que será abordada a partir de uma equação típica minceriana, desenvolvida por Mincer (1974), que integrou a teoria do investimento em capital humano dentro de um contexto empírico, compatível com a teoria econômica. De maneira geral, as equações mincerianas de salários apresentam a seguinte

estrutura log-linear, sendo $\ln W$ o logaritmo natural do salário do indivíduo i ; Θ_j é o coeficiente de algum fator de mercado j que afeta o salário do indivíduo i ; X_{ij} é o valor do fator de mercado j para o indivíduo i e μ é o erro aleatório:

$$\ln W_i = \sum_{j=1}^k \theta_j X_{ij} + \mu$$

Desta maneira, a equação é utilizada para analisar a resposta no salário frente às mudanças no capital humano, podendo ser reescrita como:

$$\ln w = \beta_0 + \beta_1 \cdot educ + \beta_2 \cdot exp + \beta_3 \cdot exp^2 + \gamma' \cdot x + \varepsilon$$

onde:

- $\ln w$ é o salário recebido pelo indivíduo;
- $\beta_1 \cdot educ$ é a escolaridade formal;
- $\beta_2 \cdot exp$ e $\beta_3 \cdot exp^2$ representam a experiência do trabalhador;
- $\gamma' \cdot x$ é um vetor com outras características observáveis do indivíduo;
- ε é um erro estocástico.

Uma condição adotada para a escolha metodológica foi de utilizar apenas registros de primeiro emprego. Ao se utilizar de indivíduos sem nenhuma experiência formal de vínculo empregatício, elimina-se um possível viés da análise relacionada à experiência dos indivíduos. Isso é importante porque a experiência pode afetar a empregabilidade de várias maneiras, como aumentando as chances de conseguir um emprego, aumentando o salário e facilitando a mudança de emprego.⁷

Assim, o modelo funcional adotado por este trabalho considera os fatores apresentados por Mincer (1974) e se adicionam outras características observáveis do indivíduo, como os elementos de ancestralidade, de identificação do evento a ser abordado e variáveis de controle do mês de admissão, da região e da Divisão CNAE, para evitar vieses

⁷ A forma de selecionar os trabalhadores no primeiro emprego será detalhada mais adiante.

setoriais, regionais e temporais. A descrição destas variáveis e as suas estatísticas descritivas são apresentadas na próxima seção.

3. DADOS

Para avaliar uma possível discriminação da ancestralidade no mercado de trabalho, optou-se por utilizar a base da RAIS, de responsabilidade do Ministério do Trabalho e Emprego (MTE) e que contém o registro de todos os vínculos formais de trabalho no Brasil. Para o ano de 2014, a RAIS conta com mais de 76 milhões de observações de vínculos de indivíduos com emprego formal em algum momento do ano, independente do vínculo estar ativo em 31 de dezembro.

Uma vez que a categoria de ancestralidade não é uma condição usual nas fontes de informação, foi necessário utilizar a base de dados identificada, que contém o nome e sobrenome dos indivíduos, para então realizar uma categorização destes trabalhadores a partir do modelo utilizado em Monasterio (2017) e aprofundado por Lopes et al. (2021), apresentado adiante.

Realiza-se o recorte na base, mantendo somente indivíduos com tipo de admissão análogo à condição de primeiro emprego. Adiciona-se que, além de utilizar a variável de primeiro emprego, descrita no dicionário de variáveis da RAIS, também foram realizadas verificações nas bases da RAIS dos quatro anos antecedentes, a fim de certificar que os indivíduos selecionados de fato não apresentaram nenhum vínculo empregatício formal nos anos anteriores. Como resultado, permanecem 1.629.063 observações.

Também são eliminados registros com remuneração inferior a um salário-mínimo (426.787), com sinalização de ensino superior completo, mas com idade inferior a 21 anos (8.035 indivíduos), registros de trabalho em contratos de menos de 20 horas semanais (25.180) e pessoas abaixo de 14 anos (5). Tais procedimentos são necessários para que ocorra a devida limpeza da base de dados, evitando o uso de registros inconsistentes ou com possíveis erros de preenchimento na análise das informações. Após esta etapa de seleção e

limpeza dos registros, restam 1.169.056 registros de trabalho. As variáveis obtidas na base da RAIS estão na Tabela 1.

Tabela 1: Descrição das variáveis

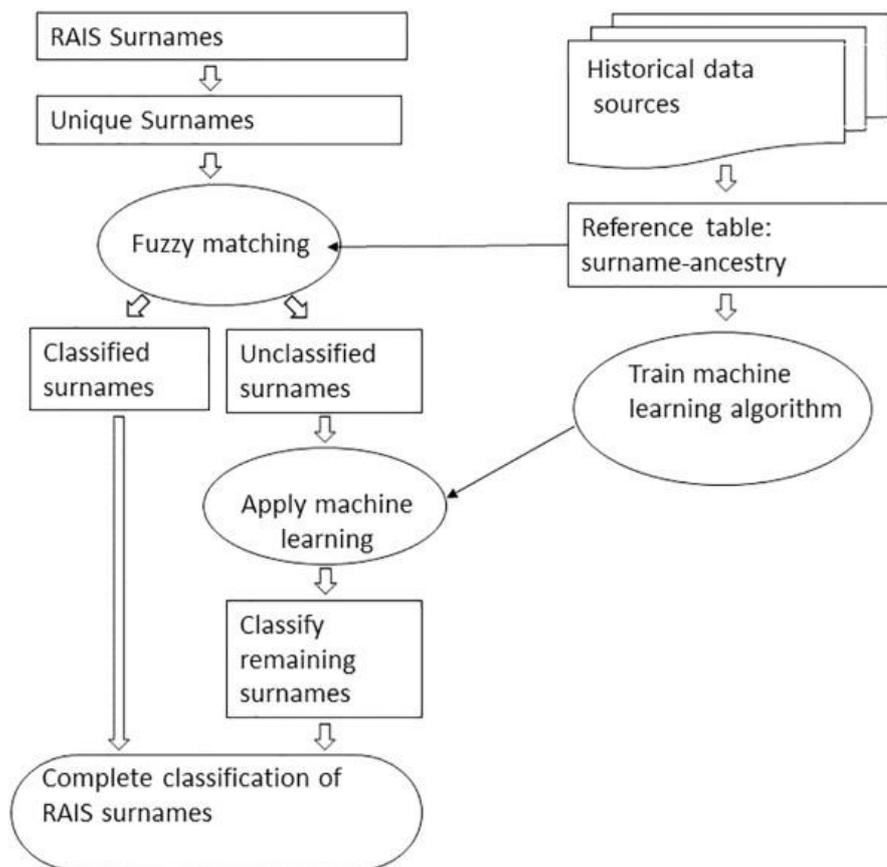
Variável	Tipo	Descrição
Salário	Contínua (salario)	Remuneração mensal em reais. Na regressão utiliza-se do logaritmo do salário mensal
Idade	Discreta (idade)	Idade do trabalhador no momento da contratação
Ensino superior	Categórica – fator	Indica o grau de escolaridade
Região	Categórica – fator	Indica a origem por Grande Região daquela contratação: Norte, Nordeste, Centro-Oeste, Sudeste ou Sul.
Região metropolitana	regiao_metro 0: não metropolitana 1: metropolitana	Indica se a contratação é de região metropolitana ou não
Gênero	genero 0: feminino 1: masculino	Indica a identidade de gênero dos indivíduos binariamente
Mês de admissão	Categórica – fator	Mês de contratação do indivíduo
Ancestralidade do sobrenome	Categórica – fator	Indica a ancestralidade do último sobrenome do indivíduo
Divisão CNAE	Categórica – fator	Nível hierárquico setorial atribuído pela Classificação Nacional de Atividades Econômicas (CNAE)
Pós-jogo da Copa	jogo 0: antes do jogo 1: após o jogo	Indica se a contratação ocorreu antes ou após o jogo

Fonte: Elaboração própria com base na RAIS (2014) e Monasterio (2017).

O algoritmo de classificação de ancestralidade empregado ao sobrenome de cada indivíduo da base foi apresentado em Monasterio (2017). Através de dados históricos e contemporâneos, o autor elaborou um documento de referência que relaciona sobrenomes com suas respectivas ancestralidades. Esse arquivo desempenha um papel fundamental ao

servir como ponto de partida para duas importantes etapas do processo: a correspondência aproximada por meio de *fuzzy matching* e a aplicação de algoritmos de *machine learning*.

Figura 2: Diagrama de fluxo para classificação de sobrenomes



Fonte: Surnames and ancestry in Brazil - Monasterio (2017).

Como mostra a Figura 2, no segundo passo é realizada a correspondência entre a tabela de referência - documento base gerado pelo autor - e os sobrenomes. No entanto, muitos nomes não puderam ser classificados nessa fase inicial, já que não apresentam semelhanças suficientes com os nomes da tabela de referência. Portanto, apenas os sobrenomes sem correspondência serão submetidos à aplicação dos algoritmos de aprendizado de máquina, sendo que esses últimos foram treinados com base na mencionada tabela de referência. Uma vez que todos os sobrenomes únicos tenham sido associados a

uma ancestralidade, o banco de dados RAIS é utilizado mais uma vez para classificar os nomes de todos os indivíduos.

Lopes (2017) e Lopes et al. (2021) ampliaram o banco de dados que associa as ancestralidades aos sobrenomes e alimenta o treinamento do algoritmo de classificação e permitiu uma classificação mais acurada dos sobrenomes. Neste *paper*, a classificação de sobrenomes foi feita com base nessa versão mais recente e sua distribuição consta na Tabela 2.

Tabela 2: Trabalhadores por categoria de ancestralidade

Sigla	Ancestralidade	Indivíduos	%	Salário Médio(R\$)	Média de Idade
IBR	Ibérica	538.040	46,0	1.070	20,0
MXD	Parda	477.937	40,8	973	20,6
BLK	Negra	61.726	5,27	978	20,3
ITA	Italiana	47.426	4,05	1.215	19,9
SLB	Sírio-Libanesa	18.761	1,62	1.094	20,2
GER	Germânica	13.945	1,19	1.140	19,5
JPN	Japonesa	3.770	0,32	1.576	21,5
EAS	Leste Europeu	3.467	0,29	1.097	19,5
NAT	Indígena	2.270	0,19	1.037	20,9
GBR	Grã-Bretanha	597	0,05	1.153	20,1
IND	Indiana	417	0,03	1.076	21,9
FRA	Francesa	353	0,03	1.149	20,2
OTR	Outros	347	0,02	1.509	20,4
Total		1.169.056	100	1.034	20,3

Fonte - Elaboração própria com base na RAIS (2014).

Indivíduos não identificados e aqueles pertencentes às ancestralidades destacadas na Tabela 2 foram removidos do modelo devido à baixa representatividade estatística, restando 1.161.605 observações. Vale notar que a variável de ancestralidade foi construída da seguinte forma: os indivíduos com cor/raça preta, parda e indígenas tiveram essa classificação preservada. Já os que constam na base da RAIS como amarelos ou brancos foram classificados de acordo com o algoritmo de classificação.

Para delimitação temporal do possível efeito do jogo Brasil e Alemanha sobre a discriminação dos trabalhadores de ancestralidade germânica, foram excluídas as observações que ocorreram duas semanas após o jogo (32.016). De acordo com a Pesquisa dos Profissionais Brasileiros, da Catho (2014), o trabalhador brasileiro leva em média 1,4 semanas entre a primeira entrevista para um vaga de emprego e o momento em que a contratação acontece. Assim, tem-se como resultado uma base com 1.129.589 observações, apresentando as seguintes estatísticas:

Tabela 3: Estatísticas descritivas

Variável	Mínimo	Média	Mediana	Máximo	Desvio-padrão
regiao_metro	0	0,5039	1	1	0,4999
genero	0	0,4777	0	1	0,4995
idade	14	20,28	19	29	3,07
salario	724	1.035,5	907,1	108.577	639,06

Fonte - Elaboração própria com base na RAIS (2014).

4. RESULTADOS

Do total de 1,1 milhão de registros de primeiro emprego ao longo de 2014, metade estava vinculada a regiões metropolitanas brasileiras e 47,7% são mulheres. Como se trata de primeiros empregos, a média de idade era de 20,3 anos de idade, sendo que dois terços dos indivíduos tinham até 22 anos e um a cada 10 tinham mais que 25 anos de idade. Apenas 7,8% dos trabalhadores sem experiência prévia cursaram ou pelo menos ingressaram em algum curso de nível superior.

A remuneração média dos registros foi de R\$ 1.034,50, e os menores salários da base de dados foram de R\$ 724, que se iguala ao salário-mínimo da época. Por outro lado, até 95% dos indivíduos da base de dados recebe até R\$ 1.602 mensalmente.

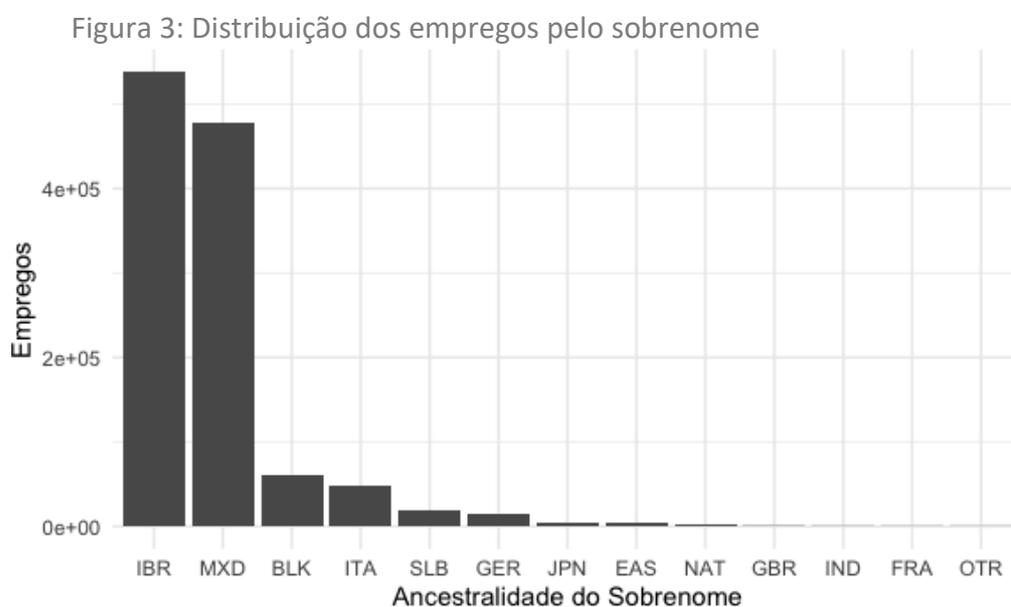
Do total de trabalhadores encontrados nas condições acima, 538 mil (46%) possuem sobrenomes ibéricos (IBR), outros 41% foram classificados com ancestralidade parda (MXD),

sendo estas as duas principais categorias de ancestralidade encontradas no mercado de trabalho brasileiro. Após elas, encontram-se 61,7 mil trabalhadores de ancestralidade negra (BLK), 47,4 mil italianos (ITA), 13,9 mil germânicos (GER) e 3,4 mil japoneses (JPN) - Tabela 2.

Cabe mencionar que do total de trabalhadores de ancestralidade germânica, 70% se encontram na Região Sul do país, outros 20% estão localizados na Região Sudeste. As demais regiões do país contam com 1,3 mil germânicos em primeiro emprego, o que representa aproximadamente 10% do total.

Entre os setores, o comércio varejista foi o principal segmento de primeiro emprego aos germânicos em 2014, com 4,2 mil empregos, seguindo da fabricação de alimentos (791) e serviços de alimentação (653). De modo geral, 65% dos registros de vínculos formais estão na modalidade de serviços e os demais 35% pertencem à agropecuária ou indústria.

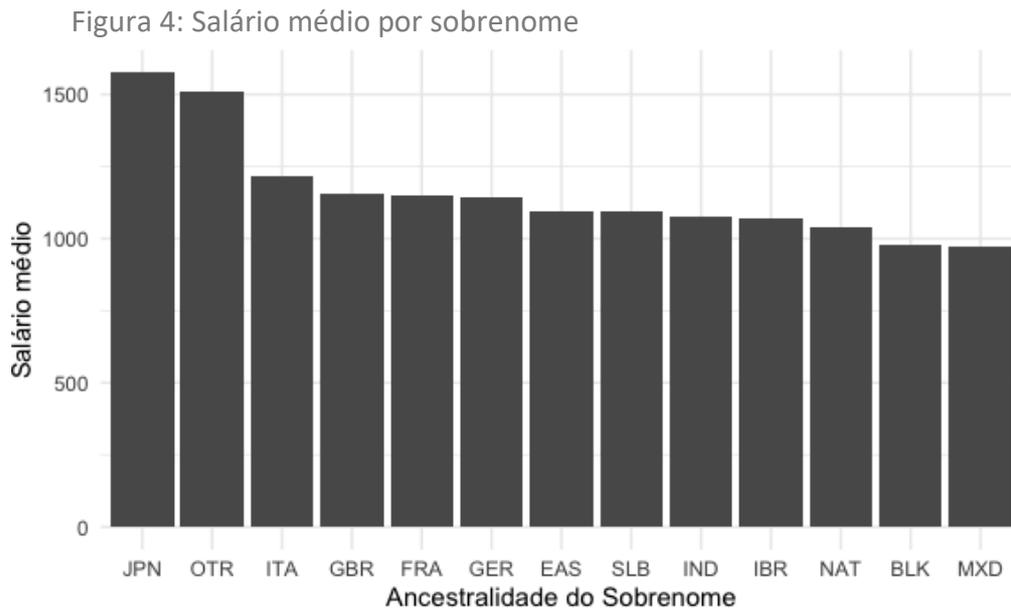
A figura abaixo retrata a composição de outras ancestralidades de menor quantitativo:



Fonte: Elaboração própria com base na RAIS (2014).

Quanto aos salários, indivíduos de ancestralidades parda e negra apresentaram remunerações abaixo da média, de R\$ 973 e R\$ 978, respectivamente. Do outro lado, os

japoneses lideram o ranking de remunerações, com R\$ 1.576, em média, na sequência aparecem italianos, britânicos (GBR), franceses (FRA) e germânicos - Figura 4.



Fonte: Elaboração própria com base na RAIS (2014).

Para analisar o componente de ancestralidade na determinação dos salários, utiliza-se do modelo básico de Mincer para avaliar o incremento desta variável sobre as remunerações. Identifica-se pela Tabela 4 que os níveis de significância estatística justificam a avaliação do fator de discriminação étnica.

Tabela 4: Regressões de Mincer para a ancestralidade*

	Mincer Básico	Mincer com ancestralidade
(Intercepto)	5.932*** (0.012)	5.892*** (0.012)
Idade	0.076*** (0.001)	0.078*** (0.001)
Idade ²	-0.001*** (0.000)	-0.002*** (0.000)
Região Metropolitana	0.043*** (0.000)	0.045*** (0.000)

Gênero	-0.072*** (0.000)	-0.072*** (0.000)
Anc. Germânica		0.047*** (0.002)
Anc. Ibérica		0.028*** (0.001)
Anc. Italiana		0.056*** (0.002)
Anc. Japonesa		0.125*** (0.004)
Anc. Parda		0.007*** (0.001)
Anc. Sírio-Libanesa		0.028*** (0.002)
<hr/>		
Controle do Grau de escolaridade	Sim	Sim
Num.Obs.	1 129 589	1 129 589
R2	0.291	0.293
R2 Adj.	0.291	0.293
AIC	72 101.2	68 762.0
BIC	72 340.0	69 072.3
Log.Lik.	36 030.605	34 354.986
F	25 770.985	19 525.402
RMSE	0.25	0.25

Fonte: Elaboração própria. + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

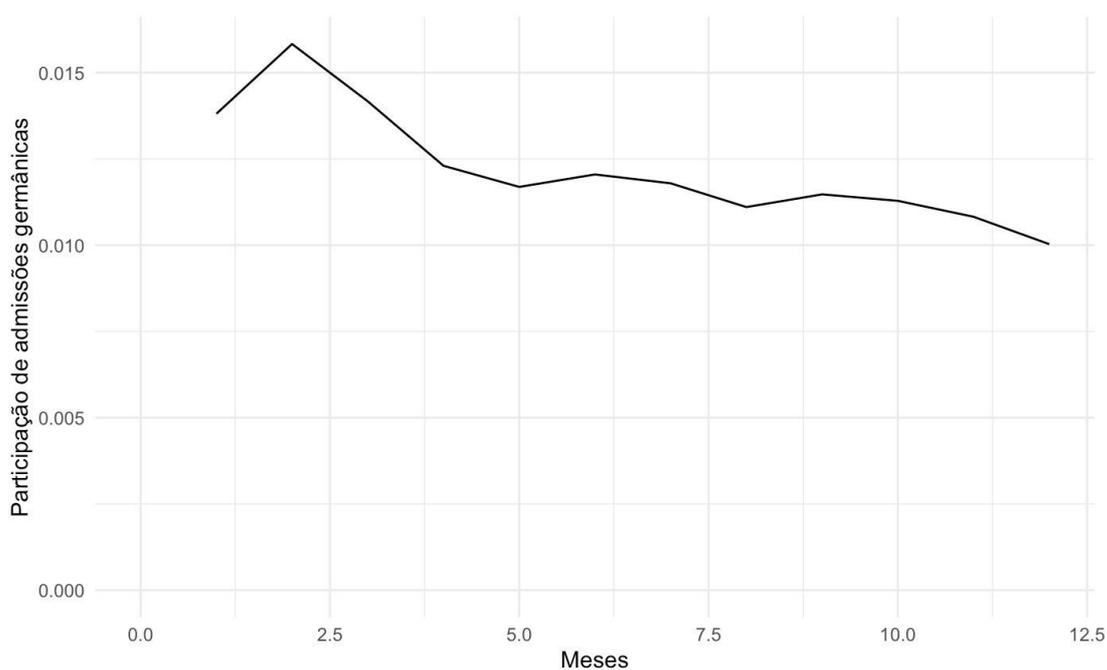
Entre os coeficientes encontrados, destaca-se um prêmio salarial positivo associado a todas as ancestralidades em relação à origem negra. Entre as ancestralidades com maiores coeficientes estão a japonesa, italiana e germânica.

Além disso, destaca-se que as regressões mincerianas apresentam comportamento similar ao observado pela literatura, que sugere que os salários aumentem de acordo com a escolaridade e a experiência. Já o fator de pertencer a uma região metropolitana também se apresentou como significativo e positivo, enquanto a identificação do gênero feminino acabou se apresentando com efeito negativo à formação de salários.

Relacionado ao objetivo específico da análise de um possível impacto da eliminação brasileira na Copa do Mundo de futebol, avalia-se também a evolução da participação germânica nas admissões do mercado de trabalho brasileiro.

A Figura 5 apresenta um claro sinal de queda das participações germânicas logo a partir do início do ano. Isto significa que, relativamente aos demais trabalhadores brasileiros, a entrada de germânicos no mercado de trabalho é muito superior no início do ano e se reduz ao longo dos meses, o que pode estar atrelado a possíveis sazonalidades de setor ou região.

Figura 5: Evolução das admissões germânicas ao longo de 2014



Fonte - Elaboração própria com base na RAIS (2014).

Tais fatores temporais e de segmentação de setores são levados em consideração na avaliação dos efeitos da Copa do Mundo sobre o salário dos indivíduos germânicos e não-germânicos e são acrescentadas ao modelo de regressão adotado como variáveis de controle.

Avançando para a análise do efeito do jogo na definição dos salários de diferentes ancestralidades, são apresentados os resultados de três modelos na Tabela 5⁸.

Tabela 5: Regressões do impacto do jogo sobre os salários*

	Modelo 1	Modelo 2	Modelo 3
(Intercepto)	5.851*** (0.012)	5.851*** (0.012)	6.064*** (0.012)
Idade	0.081*** (0.001)	0.081*** (0.001)	0.064*** (0.001)
Idade ²	-0.002*** (0.000)	-0.002*** (0.000)	-0.001*** (0.000)
Gênero	-0.072*** (0.000)	-0.072*** (0.000)	-0.047*** (0.000)
Região Metropolitana	0.045*** (0.000)	0.045*** (0.000)	0.051*** (0.000)
Anc. Germânica	0.047*** (0.002)	0.043*** (0.003)	0.042*** (0.003)
Anc. Ibérica	0.028*** (0.001)	0.024*** (0.001)	0.024*** (0.001)
Anc. Italiana	0.056*** (0.002)	0.053*** (0.002)	0.052*** (0.002)
Anc. Japonesa	0.126*** (0.004)	0.136*** (0.006)	0.132*** (0.005)
Anc. Parda	0.007*** (0.001)	0.008*** (0.001)	0.010*** (0.001)
Anc. Sírio-Libanesa	0.028*** (0.002)	0.026*** (0.003)	0.028*** (0.003)
Jogo		0.042*** (0.004)	0.034*** (0.004)
Jogo × Anc. GER		0.012* (0.005)	0.008+ (0.005)
Jogo × Anc. IBR		0.009*** (0.002)	0.007*** (0.002)
Jogo × Anc. ITA		0.007* (0.003)	0.005+ (0.003)
Jogo × Anc. JPN		-0.024** (0.009)	-0.032*** (0.008)

⁸ A fim de facilitar a visualização, foram omitidos os coeficientes de controle de mês e Divisão CNAE.

Jogo × Anc. MXD		-0.003 (0.002)	-0.003 (0.002)
Jogo × Anc. SLB		0.004 (0.004)	0.002 (0.004)
Controle do Grau de escolaridade	Sim	Sim	Sim
Controle Mês de Admissão	Sim	Sim	Sim
Controle Região	Sim	Sim	Sim
Controle CNAE	Não	Não	Sim
Num.Obs.	1 129 589	1 129 589	1 129 589
R2	0.294	0.294	0.347
R2 Adj.	0.294	0.294	0.347
AIC	67 404.2	67 031.5	19 825.7
BIC	67 845.9	67 556.7	18 273.9
Log.Lik.	33 665.114	33 471.737	10 042.855
F	13 444.525	11 216.748	4681.679
RMSE	0.25	0.25	0.24

Fonte: Elaboração própria. + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

O Modelo 1 se assemelha a regressão minceriana com ancestralidade obtida na tabela anterior, mas é acrescentado o controle temporal, do mês de admissão, e regional, das grandes regiões brasileiras, ocultas nos resultados para facilidade a visualização. O acréscimo deste controle não altera na significância das variáveis, mas é capaz de marginalmente melhorar o coeficiente de determinação e os critérios de informação de Akaike (AIC) e Bayesiano (BIC).

Já o Modelo 2 apresenta os resultados com a inserção das variáveis de choque, representadas pela variável *dummy* que marca as admissões após o jogo, bem como a sua interação com as ancestralidades avaliadas.

Destaca-se que as métricas de avaliação do modelo pouco se alteram com este incremento, mas o impacto do jogo se apresenta com resultado positivo e significativo estatisticamente tanto para o modelo geral, como sua interação com os germânicos, ibéricos e italianos. Para os japoneses, o impacto do jogo parece ter apresentado efeito contrário, reduzindo a sua disparidade em relação aos salários.

Por fim, o Modelo 3 repete a fórmula com o acréscimo de mais um controle, agora setorial. Optou-se pela classificação dos empregos em divisões CNAE, reduzindo possíveis vieses relacionados a movimentos setoriais. A inserção da Divisão CNAE favorece o coeficiente de determinação e mantém os níveis de significância observados no Modelo 2.

Como um todo, os resultados não são tão claros quanto se esperava em relação à hipótese central. De fato, os resultados indicam que o impacto do jogo do Brasil e Alemanha de fato está associado a salários mais altos para os indivíduos com ancestralidade germânica. Porém, há impactos com valores absolutos menores também para os indivíduos com sobrenomes italianos, ibéricos e, com sinal negativo, para os de ascendência japonesa. Uma vez que foram utilizadas uma ampla gama de variáveis de controle, incluindo regionais, setoriais e temporais, que poderiam estar atuando como confundidoras, não há uma explicação no momento para tais achados.

Complementarmente, buscamos estimar modelos Probit para a probabilidade de o indivíduo ser admitido após o jogo do Brasil e Alemanha de acordo com a sua ancestralidade. Nesta configuração, são apresentados dois modelos, o primeiro sem o incremento da variável de controle de CNAE e o segundo com a inserção desta variável categórica (Tabela 6)⁹.

Tabela 6: Modelos Probit do impacto do jogo sobre a probabilidade de admissão*

	Modelo 1	Modelo 2
(Intercepto)	4.762*** (0.061)	4.787*** (0.062)
Idade	-0.437*** (0.005)	-0.443*** (0.005)
Idade ²	0.009*** (0.000)	0.009*** (0.000)
Escolaridade: Até 5° ano incompleto	0.022 (0.030)	0.026 (0.030)
Escolaridade: Até 5° ano completo	-0.038 (0.030)	-0.036 (0.030)

⁹ A fim de facilitar a visualização, foram omitidos os coeficientes de controle de mês e Divisão CNAE.

Escolaridade: 6° a 9° ano incompleto	0.019 (0.029)	0.023 (0.029)
Escolaridade: Fundamental completo	-0.010 (0.029)	-0.010 (0.029)
Escolaridade: Médio incompleto	-0.024 (0.028)	-0.032 (0.029)
Escolaridade: Médio completo	0.000 (0.028)	-0.016 (0.028)
Escolaridade: Superior incompleto	0.070* (0.029)	0.084** (0.029)
Escolaridade: Superior completo	0.031 (0.029)	0.074* (0.029)
Escolaridade: Mestrado	0.277*** (0.060)	0.391*** (0.060)
Escolaridade: Doutorado	0.526*** (0.173)	0.567*** (0.174)
Gênero	0.079*** (0.002)	0.069*** (0.003)
Anc. Germânica	-0.112*** (0.012)	-0.099*** (0.012)
Anc. Ibérica	-0.055*** (0.005)	-0.051*** (0.006)
Anc. Italiana	-0.072*** (0.008)	-0.068*** (0.008)
Anc. Japonesa	-0.024 (0.022)	-0.029 (0.022)
Anc. Parda	0.019*** (0.006)	0.015** (0.006)
Anc. Sírio-Libanesa	-0.013 (0.011)	-0.014 (0.011)
Controle Região	Sim	Sim
Controle CNAE	Não	Sim
Num.Obs.	1 129 589	1 129 589
AIC	1 528 282.4	1 520 551.9
BIC	1 528 580.8	1 521 877.0
Log.Lik.	764 116.193	760 164.952
F	743.701	231.578
RMSE	0.49	0.49

Fonte: Elaboração própria com base na RAIS (2014) e Monasterio (2017). + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Estimou-se um impacto negativo na probabilidade dos trabalhadores com ancestralidade germânica serem contratados após a derrota para a Alemanha. Um fenômeno do lado da oferta de trabalho pode ajudar a entender esse resultado aparentemente: pode ser que tais indivíduos tenham tido um aumento de sua autoestima e elevaram o seu salário de reserva. Esse mecanismo explicaria - ao mesmo tempo - os impactos positivos estimados para os salários e negativos na contratação.

Mais uma vez, porém, os resultados não foram tão límpidos quando inicialmente desejado. Não há justificativa evidente para que indivíduos de ancestralidade italiana e ibérica também tenham sido menos contratados e pardos mais contratados.

5. CONCLUSÃO

Este trabalho avaliou o impacto emocional da goleada sofrida pelo Brasil contra Alemanha na Copa do Mundo de futebol masculino em 2014 sobre o mercado de trabalho brasileiro, focando a avaliação em um possível efeito nos salários e admissões de trabalhadores de ancestralidade germânica.

Ao estimar o efeito de um choque exógeno de caráter não-econômico no mercado de trabalho, este trabalho contribui com as literaturas relacionadas ao estudo da ancestralidade na dinâmica do mercado de trabalho e na mensuração do efeito de um jogo esportivo sobre as condições econômicas. Fez-se um tratamento na base de dados identificada da RAIS para identificar a ancestralidade dos indivíduos a partir do seu sobrenome, o que permitiu a utilização de um modelo minceriano que avalia o ganho salarial dos trabalhadores a partir das suas qualificações e o efeito da ancestralidade e do jogo do Brasil e Alemanha no mercado de trabalho. Também se avaliou a partir de uma regressão a probabilidade de admissão de indivíduos de ancestralidade germânica antes e depois do jogo.

Os resultados indicaram alguma evidência de que a vitória da Alemanha teve um impacto positivo nos salários de primeiro emprego dos trabalhadores com sobrenomes

germânicos. Ao mesmo tempo, as estimativas sugeriram que a probabilidade destes indivíduos serem contratados caiu após o jogo. Esse resultado pode derivar de um aumento do salário de reserva desses indivíduos, mas não temos outras evidências para corroborar essa conjectura.

Há que se admitir também que há resultados econométricos de difícil interpretação. Identificou-se impactos de menor valor do jogo também para os trabalhadores com sobrenomes italianos, ibéricos e, com o sinal oposto, japoneses. Deve-se ressaltar que as estimações incluíram um conjunto exaustivo de controles, incluindo não só os individuais, mas também regionais, setoriais e temporais. Obviamente, há ainda muito a fazer na literatura sobre desigualdade por ancestralidade no Brasil. Buscou-se aqui examinar apenas um dos possíveis mecanismos de manutenção da desigualdade e discriminação no mercado de trabalho. Novos estudos, utilizando outras estratégias de identificação, talvez possam trazer à luz seus exatos funcionamentos e gravidades.

6. REFERÊNCIAS

Arai, M. and Skogman Thoursie, P. (2009). Renouncing personal names: An empirical examination of surname change and earnings. *Journal of Labor Economics*, 27(1):127–147.

Arrow, K. et al. (1973). The theory of discrimination. *Discrimination in labor markets*, 3(10):3–33. Becker, G. S. (1962). Investment in human capital: A theoretical analysis. *Journal of Political*

Economy, 70(5, Part 2):9–49.

Becker, G. S. and Chiswick, B. R. (1966). Education and the Distribution of Earnings. *The American Economic Review*, 56(1/2):358–369.

Blinder, A. S. (1973). Wage discrimination: reduced form and structural estimates. *Journal of Human Resources*, pages 436–455.

Carvalho, C., Ribeiro, R., and Zilberman, E. (2017). Sentiment, Electoral Uncertainty and Stock Returns. Available at SSRN 2930564.

Catho (2014). Pesquisa dos Profissionais Brasileiros - Um panorama sobre contratação, demissão e carreira. Catho, São Paulo, Brasil.

Caulyt, F. (2014). Seleção brasileira é favorita nas casas de apostas. Deutsche Welle.

Depetris-Chauvin, E., Durante, R., and Campante, F. (2020). Building nations through shared experiences: Evidence from African football. *American Economic Review*, 110(5):1572–1602.

Dias, R. (2018). 8 de julho de 2014, o dia que o Brasil chorou de vergonha no Mineirão 8 de julho de 2014, o dia em que o país chorou de vergonha no Mineirão. Section: 90 anos.

Kaplanski, G. and Levy, H. (2010). Exploitable predictable irrationality: The FIFA World Cup effect on the US stock market. *Journal of Financial and Quantitative Analysis*, 45(2):535–553.

Lopes, D., Silva Filho, G. A. d., and Monasterio, L. (2021). Family Culture and School Achievement in Brazil. Em pré-publicação.

Lopes, D. A. F. (2017). Culture, institutions, and school achievement in Brazil.

Mincer, J. (1958). Investment in human capital and personal income distribution. *Journal of Political Economy*, 66(4):281–302.

Mincer, J. (1974). Schooling, Experience, and Earnings. *Human Behavior & Social Institutions* No. 2.

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

Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International Economic Review*, pages 693–709.

Oaxaca, R. L. and Dickinson, D. L. (2006). Statistical Discrimination in Labor Markets: An Experimental Analysis.

Phelps, E. S. (1972). The statistical theory of racism and sexism. *The American Economic Review*, 62(4):659–661.

Rubinstein, Y. and Brenner, D. (2014). Pride and prejudice: Using ethnic-sounding names and inter-ethnic marriages to identify labour market discrimination. *Review of Economic Studies*, 81(1):389–425.

Schultz, T. W. (1961). Investment in human capital. *The American Economic Review*, 51(1):1–17. UOL (2014). 7 a 1 e gol do Khedira. Apostador “chuta o balde” e ganha R\$100 mil.

UOL (2015). Lahm: No intervalo do 7 a 1, Alemanha combinou de não tripudiar do Brasil.



The background features a dark blue color scheme with various data visualization elements. On the left, there is a bar chart with four bars of increasing height, with values 178, 180, 175, and 190. A line graph with two lines is overlaid on the bars, showing an overall upward trend. To the right, there are two circular progress indicators: the top one shows 68% and the bottom one shows 75%. A dotted line with a triangle points towards the 75% indicator. At the bottom right, there is a network diagram consisting of a sphere of interconnected nodes. The 'idp' logo is centered in the middle of the page.

idp

SGAS Quadra 607 - Módulo 49
Via L2 Sul, Brasília-DF
CEP: 70200-670

  /sejaidp
 (61) 3535-6565
 idp.edu.br