

Área de interesse: Teoria Aplicada

Título do artigo: A SORTE SORRI PARA TODOS? UMA NOVA ABORDAGEM DE DESIGUALDADE DE OPORTUNIDADES.

Autores:

Lauro Nogueira

Professor da Universidade Federal Rural do Semiárido e do Programa de Pós-Graduação em Economia da Universidade do Estado do Rio Grande do Norte.

Rua Portugal, 11, Pau dos Ferros/RN

Email: lauro.nogueira@ufersa.edu.br

Erik Figueiredo

Professor do Departamento de Economia da Universidade Federal da Paraíba

Email: eafigueiredo@gmail.com

A SORTE SORRI PARA TODOS? UMA NOVA ABORDAGEM DE DESIGUALDADE DE OPORTUNIDADES

RESUMO

Este artigo investiga quais as implicações em não considerar – sorte genética – características não observáveis na determinação de um resultado econômico – salário hora – individual. Em outras palavras, buscou-se superar limitações empíricas e estruturais dos atuais modelos de igualdade de oportunidades. Para tanto, adotou-se uma estratégia empírica simples aplicada a uma amostra composta por 680 gêmeos monozigóticos. Os principais resultados mostraram que ao considerar fatores não observáveis como, habilidade cognitiva geral, herdabilidade genética e ambiente compartilhado o efeito médio de uma variável determinante do resultado individual declina em torno 50%.

Palavras Chave: Igualdade de Oportunidades. Sorte Genética. Habilidade Cognitiva Geral. Herdabilidade Genética. Ambiente Compartilhado.

ABSTRACT

This paper investigates the implications of not consider – lucky genetic – unobservable characteristics in determining an economic result - hour salary - individual. In other words, it tried to overcome empirical and structural limitations of current models of equal opportunities. To this end, it adopted a simple empirical strategy applied to a sample of 680 monozygotic twins. The main results showed that when considering unobservable factors such as general cognitive ability, genetic heritability and shared environment the average effect of a variable determinant of individual income declines by around 50%.

Keywords: Equal Opportunities. Lucky Genetics. General Cognitive Ability. Genetic heritability. Shared environment.

JEL classification: I20, I21, D63

1 INTRODUÇÃO

A abordagem clássica tradicional da Teoria de Igualdade de Oportunidades – Roemer, 1998 – estabelece que o sucesso ou fracasso individual é explicado por uma combinação de fatores determinísticos e aleatórios. Em síntese, o resultado individual, é determinado por três conjuntos de variáveis: i) variáveis de circunstâncias, as quais fogem do controle individual, de forma que, toda desigualdade social oriunda destas é considerada nociva, e, portanto, devem ser compensadas; ii) variáveis de esforço ou de responsabilidade individual, nas quais são observadas como origens de desigualdade social justa, e, desse modo, não devem ser anuladas perante esta abordagem; e por fim, iii) por um termo de sorte bruta.

Em relação ao último conjunto, um valor positivo é habitualmente visto como boa sorte. De acordo com Lefranc, Pistolesi e Trannoy (2009), os debates filosóficos sobre a igualdade de oportunidades, o conceito de sorte considera situações em que a responsabilidade do controle individual de escolha ou moral, não tem qualquer relação com a ocorrência dos resultados. É interessante destacar que, duas hipóteses-chaves são assumidas nesse arcabouço teórico: primeira, a independência – exogeneidade – atribuída entre os conjuntos de variáveis de esforço e de circunstâncias. Segundo, o caráter completamente aleatório atribuído ao fator sorte no resultado individual. Essa caracterização abre espaço para, ao menos, duas contestações: i) há independência entre as variáveis de circunstância e de esforço? ii) as componentes da sorte são aleatórias? É importante destacar que caso essas suposições não sejam válidas, então, os resultados empíricos da literatura em questão não expressam a realidade. Isto é, os principais resultados apontados nessa abordagem são no mínimo equivocados.

Destaca-se que em outros trabalhos essas suposições foram questionadas. Por exemplo, em Bourguignon, Ferreira e Menéndez (2007), examina-se a suposição assumida em (i), embora, haja reservas ao tratamento dado a endogeneidade das covariadas, em especial, ao problema referente à omissão de variáveis.

Por outro lado, Lefranc, Pistolesi e Trannoy (2009), postulam diferentes vertentes filosóficas e econômicas referentes à hipótese (ii). Entretanto, em nenhum momento cogitou-se mensurar a participação da sorte no resultado individual. Além disso, os autores caracterizam a sorte, em todas as suas especificidades, como fator de não responsabilidade individual, atribuindo-a apenas a função de captar aspectos que fogem da responsabilidade individual. Posto isso, surge uma questão fundamental a ser investigada. Isto é, há componentes determinísticos que compõem a sorte que é tratada como aleatórias pela moderna teoria de igualdade de oportunidades? Se a resposta for sim, então quais seriam?

Por sua vez, a literatura filosófica e econômica assume a existência de pelo menos quatro tipos de sorte: i) background social, denominada por Rawls de loteria social. Nesse caso os resultados individuais somente diferem em função das diferenças em suas conexões sociais – *rent seeking* – familiares; ii) constitutiva ou genética – loteria natural – a qual designa situações aonde as diferenças nos resultados individuais acontecem em razão da herdabilidade genética; iii) opção, assim denominada por Dworkin. Essa caracterização da sorte ocorre quando o resultado é fruto das escolhas individuais; e por fim, iv) sorte bruta, que contempla as situações que estão completamente fora do controle individual.

Diante disso, podem-se fazer alguns apontamentos. Primeiro, a sorte background social parece ser adequadamente pela teoria tradicional. Por outro lado, a sorte genética ou constitutiva, seja em função da difícil mensuração ou da ausência de fatores que captem seus efeitos é considerada e/ou configurada como sorte bruta. Essa caracterização, ao menos empírica aparenta ser uma grave limitação. Ao mesmo tempo, parece existir também uma clara controvérsia teórica, quanto à nomeada *sorte opção*. Uma vez que, a literatura descreve-a como parte da sorte em que o indivíduo tem controle ou responsabilidade individual de

escolha. Contudo, em trabalhos seminais, por exemplo, Roemer (1998), tais fatores são desconsiderados. Em linhas gerais, nos atuais modelos de Igualdade de oportunidades – IOPs – fatores referentes à genética, opção escolhida e sorte bruta compõe o termo de erro aleatório, e, portanto, origina dúvidas sobre as premissas assumidas dessa especificação.

Diante do exposto, este trabalho tem como principal objetivo investigar o quanto essas supostas limitações – especialmente as características individuais não observáveis – afetam o resultado individual. Para tanto, adota-se uma abordagem simples proposta por Ashenfelter e Rouse (1998), associada a uma amostra composta por 680 gêmeos monozigóticos. Esse procedimento busca captar efeitos de características individuais não observáveis e algumas observáveis, porém, em grande parte não disponíveis, ao invés de simplesmente supor uma aleatoriedade, em especial, justificada pela ausência de variáveis que representem tais fatores. A adoção por uma amostra tão peculiar dar-se em função de dois aspectos fundamentais: i) o controle imediato de diversos fatores consensualmente apontados na literatura como determinantes da desigualdade indesejada ou injusta; ii) o controle substancial de diversas características não observáveis.

O trabalho está dividido em seis seções, além dessa introdução. Na segunda seção faz-se um resumo da literatura, focado essencialmente, em desigualdade de oportunidades, sorte, habilidade cognitiva geral, herdabilidade genética e ambiente compartilhado. Na terceira seção apresenta-se a estratégia empírica. Na quarta descreve-se e analisam-se os dados. A quinta seção é destinada à apresentação e discussão dos principais resultados, e, por fim, a seção final é reservada as considerações finais.

2 REVISÃO DA LITERATURA

De uma maneira geral, faz-se um breve relato sobre a Teoria de Igualdade de Oportunidades. Em seguida, destacam-se algumas variações da componente sorte. Além disso, apresentamos alguns estudos que relatam a importância de fatores como a habilidade cognitiva geral, herdabilidade genética e o ambiente compartilhado.

2.1 Igualdades de Oportunidades

O aparato geralmente utilizado pelos modelos de Igualdade de Oportunidades – IOPs – defende que o resultado individual, w_i , é produto de fatores de responsabilidade e não responsabilidade individual. Em síntese, o resultado obtido pelo agente é composto por uma gama de fatores que representam: i) as circunstâncias, C_i , por exemplo, características familiares, socioeconômicas, etc.; ii) o esforço despendido, E_i , como, horas trabalhadas, anos de estudo, etc. e; iii) por um termo de sorte bruta, u_i . Onde $f(.)$ é uma função desconhecida.

$$w_i = f(C_i, E_i, u_i) \quad (1)$$

Nos últimos anos, há estudos na literatura que criticam a estrutura da equação (1), especialmente, envolvendo a suposta exogeneidade – independência – entre as variáveis de circunstâncias e esforço, como também, o tratamento dado às componentes do termo de erro. Todavia, esses problemas, embora ressaltados, são abordados de forma superficial e insipiente. Por exemplo, em Bourguignon, Ferreira e Menéndez (2007), o estudo propõe uma decomposição da desigualdade de renda em dois aspectos, um que contempla o conjunto de oportunidades, e outro que vislumbra o termo residual. Em síntese, buscam verificar a existência de efeitos diretos e indiretos das circunstâncias sobre o resultado individual, onde os últimos agem sobre o esforço individual despendido. Os dados utilizados referem-se a ganhos salariais de homens e mulheres para o Brasil no ano de 1996. Os resultados indicam um efeito indireto das circunstâncias em torno de 30% sobre a desigualdade observada.

Todavia, a estrutura utilizada para estimar tais efeitos não difere substancialmente da equação (1).

Por outro lado, Lefranc, Pistolesi e Trannoy (2009), relatam a existência e distinção dos componentes do termo de erro, em particular, aos fatores ligados à sorte. Na verdade, o estudo busca ofertar um modelo de igualdade de oportunidades flexível bastante para englobar as mais diversas concepções sobre a teoria de igualdade de oportunidades. Adicionalmente, faz-se uma aplicação empírica referente à renda na França durante o período 1979 a 2000. Os resultados revelam que o nível de desigualdade de oportunidades tende a diminuir e que o grau de risco das distribuições de renda, condicionado à origem social, são semelhantes em todos os grupos de origens sociais observados. Contudo, apesar desses estudos apontarem algumas limitações da equação (1), tais limitações não foram superadas nesses trabalhos. Além disso, em nenhum desses estudos atentou-se para as características individuais não observáveis.

2.2 Sorte

Detalha-se a seguir as quatro principais especificações do fator sorte abordado pela literatura econômica e filosófica. Todavia, a priori, faz-se uma distinção entre sorte bruta *versus* sorte opção. Uma vez que, essa diferenciação, é parte fundamental deste trabalho, pois, em geral, a literatura de igualdade de oportunidade trata a sorte opção como uma complementação da sorte bruta. E, ao que parece, conforme abordado inicialmente e discutido a seguir, esse não é o tratamento mais adequado para o problema.

Nesse contexto, segundo Vallentyne (2002), há em geral, quatro classificações distintas sobre a sorte bruta na literatura, sendo elas: i) razoável não previsibilidade de escolha, ii) razoável inevitabilidade, iii) razoável incapacidade de influenciar, e, por fim, iv) não razoável influência deliberada. A primeira visão, conhecida como, não previsibilidade de escolha, defende que a determinação de um evento para um indivíduo somente é considerada função da sorte bruta, se e somente se, tal ocorrência não seja resultado esperado de suas escolhas.

Por exemplo, admita duas situações, na qual em ambas, há dois agentes similares e identicamente situados. Na primeira, suponha que não existem possibilidades de escolha. Isto é, ambos participam de uma loteria natural – exposição a balas perdidas – na qual o resultado é totalmente aleatório e contabilizado a sorte bruta. Por sua vez, na segunda situação, suponha que ambos possuem duas escolhas – parar ou continuar andando na troca de tiros – embora, suas escolhas não exerçam qualquer efeito sobre as probabilidades de ser atingido ou não por uma bala. Isto é, em ambas as situações, os agentes são expostos à mesma loteria natural. Além disso, admita também, que cada agente seja totalmente consciente do resultado de cada escolha. Suponha também, que cada agente decide parar e somente um deles é atingido por um tiro. Tendo em vista, que a probabilidade de ser atingido por uma bala perdida, dado as suas escolhas, eram razoavelmente previsíveis pelos agentes, o ponto de vista acima qualificaria esta sorte como bruta.

Nas situações acima parece nítido não haver razão para tratar ambos os eventos de forma diferenciada. Pois, não havia nada que os agentes pudessem fazer para alterar a sua exposição às balas perdidas. Por isso, a presença de escolha, mesmo plenamente informada não necessariamente é suficiente para fazer de um resultado consequência de uma opção.

A segunda abordagem é, sem dúvidas, uma forma mais clara de compreender a sorte bruta. Isto é, considerá-la em termos de provável inevitabilidade de um resultado. Neste caso, a influência de um evento no resultado individual é devido à sorte bruta de um indivíduo, se e somente se, o mesmo não possuir razoável possibilidade de evitar sua ocorrência. Em linhas gerais, a sorte bruta, em termos de não evitabilidade, leva em conta se o indivíduo possuía em

algum momento a possibilidade de evitar um determinado resultado. E, portanto, se o mesmo tiver como prever sua ocorrência não faz do resultado uma questão de sorte bruta. Por exemplo, sofrer um ataque de tubarão em área de risco poderia ter sido evitado se o banhista atendesse as placas de sinalização de perigo. Deve-se destacar a distinção entre a abordagem de não previsibilidade de escolha e a de inevitabilidade. Uma vez que, na segunda, a escolha importa e impacta consideravelmente os resultados.

Dito de outro modo, segundo Dworkin (1981), a sorte opção é a forma como um indivíduo leva vantagem ou não, através da aceitação de um risco deliberado e calculado, ou da antecipação de um evento na atenuação de um resultado ruim. Por outro lado, a sorte bruta na sua forma pura não são apostas deliberadas. De forma mais concisa, uma interpretação estrita da abordagem de inevitabilidade, é que, se o indivíduo tem de alguma forma como evitar sua exposição a determinado risco, não se deve considerar a diferença no resultado individual fruto da sorte bruta. Porém, há uma limitação nessa abordagem, pois não incorpora a razoabilidade de evitar o resultado em questão. Entretanto, em uma visão menos estrita, e, talvez mais prudente, compõe a sorte bruta apenas os fatos que influenciam os resultados e que não são razoavelmente evitáveis pelo indivíduo. Todavia, destaca-se que, se existe dois indivíduos identicamente situados, e que ambos fazem escolhas diferentes, então, portanto, o impacto diferencial de suas escolhas é ao que parece, devido à opção escolhida.

Vale ressaltar que, a caracterização da sorte bruta como razoável inevitabilidade tem algumas limitações. Uma das principais seria o fato que é uma simplificação útil e adequada da incapacidade de influenciar. Essa caracterização ignora a possibilidade que os indivíduos possuem de influenciar seus resultados, modificando suas probabilidades. Nesse sentido, segundo a literatura, uma visão mais geral seria a terceira caracterização. Ou seja, sorte bruta como razoável incapacidade de influenciar um resultado. Nesta abordagem, a sorte bruta influencia o resultado individual, se e somente se, o indivíduo não tem como influenciar razoavelmente a possibilidade ou probabilidade de sua ocorrência.

Essa distinção advoga que a evitabilidade é o modo no qual os resultados podem ser influenciados por um indivíduo. Porém, há um problema adicional. Pois, o “se e somente se” nesta especificação é problemática, uma vez que, exige que para determinado resultado seja atribuído à sorte não bruta ou opção, o indivíduo terá que possuir alguma influência mínima sobre o resultado. Contudo, se houver duas escolhas e as mesmas possuírem idênticos *payoffs* e probabilidades, então, deve-se considerar a influência da escolha no resultado como sorte bruta?

Nessa configuração, segundo Vallentyne (2002), uma forma razoável seria substituir o termo “se e somente se” pela expressão “na medida em que” e, em seguida, decompor os resultados decorrentes da sorte bruta e sorte opção de acordo com o nível de participação que o indivíduo exerce sobre o resultado. Embora, essa tarefa não seja trivial. Adicionalmente, há outra questão importante a entender. Pois, até o momento considerou-se um ambiente em que os indivíduos possuem perfeita informação sobre as opções disponíveis. Entretanto, sabe-se que isso não retrata completamente a realidade. Pois, embora, em algumas situações, haja eventos não previsíveis, os mesmos podem ser influenciáveis a um determinado resultado. Todavia, como influenciabilidade é entendida nesse contexto como razoável previsibilidade, assim, portanto, por simplificação, as crenças falsas ou incompletas são ao que parece corretamente tratada como uma questão de sorte bruta. Contudo, o autor ressalva os casos onde os resultados das opções escolhidas por indivíduos idênticos, e que possuem previsibilidade imperfeita, poderia ser atribuída ao diferencial da sorte bruta, mesmo quando eles pudessem influenciar os resultados.

Em outro momento, Vallentyne (2002), destaca que o simples fato de que a possibilidade de um evento seja resultado previsível de escolhas dos indivíduos não é uma condição suficiente para que o resultado seja atribuído à sorte opção. Entretanto, sugere ser

uma condição necessária. Por fim, surge a seguinte caracterização de sorte bruta, isto é, a sorte bruta como não razoável influência deliberada. Em síntese, essa caracterização da sorte defende que a ocorrência ou não de um determinado evento é devido à sorte bruta de um indivíduo, na medida em que, o mesmo não pode ter razoável deliberação que influencie a possibilidade ou probabilidade de sua ocorrência ou não.

Em linhas gerais, a inevitabilidade bruta – núcleo – é a principal distinção entre a sorte bruta e as outras componentes do fator sorte, aqui especificamente, confrontada com a sorte opção. Contudo, também se devem levar em conta, os acontecimentos para os quais os indivíduos não têm a capacidade de influenciar a probabilidade. Como também, aqueles para os quais os mesmos, não têm consciência de sua disposição de impactar suas possibilidades, sejam por causa de problemas de crenças falsas ou incompletas. Adicionalmente, a literatura, em geral, inclui nessa gama, os fatos que não são razoavelmente propícios à influência deliberada, mesmo quando a rigor são tão influenciáveis, embora, não seja claro como isso ocorra. Posto isso, a seguir define-se sinteticamente os quatro tipos de sorte sumariamente citadas.

- I. Sorte background social: caracterizada onde os diferenciais existentes entre dois indivíduos idênticamente talentosos e motivados, é função das conexões e fundo social das suas famílias. Isto é, os diferenciais resultantes são pré-determinados por fatores antecedentes, como origem familiar e social.
- II. Sorte genética: atribuída a situações em que os diferenciais resultantes, por exemplo, entre dois irmãos gêmeos decorrentes de um deles ter uma herdabilidade genética superior, de modo a gerar um talento especial, deve ser configurada como sorte genética ou constitutiva. Vale ressaltar que nesse caso não estar-se considerando gêmeos monozigóticos.
- III. Sorte opção: ocorre, por exemplo, quando a diferença nos resultados existentes é decorrente apenas das escolhas frente às opções disponíveis, excetuando-se, as loterias não evitáveis. Nessa análise considera-se que os indivíduos sejam idênticamente talentosos e motivados, e, que, além disso, possuam origens familiares e sociais similares.
- IV. Sorte bruta, já amplamente enfatizada e sinteticamente conceituada, como o tipo de sorte responsável por um resultado de uma loteria inevitável. Por exemplo, o diferencial no resultado econômico individual atribuído a dois irmãos idênticamente talentosos e situados, onde os efeitos nocivos de uma guerra afeta somente um deles, pois apenas um deles foi designado para batalha.

2.3 Habilidade Cognitiva Geral, Herdabilidade Genética e Ambiente Compartilhado

Seguindo a evolução histórica – Spearman (1904) – a habilidade cognitiva geral (*g*) representa a capacidade que os indivíduos adquirem competências e os permite interagir em seu ambiente. Tais habilidades constituem a estrutura essencial das competências cognitivas individuais. Por exemplo, dominar linguagens, compreender fenômenos, enfrentar situações e problemas, construir argumentações, elaborar propostas, entre outras aptidões individuais.

Nesse contexto, estudos recentes, por exemplo, Heckman (2011), advoga que o diferencial econômico e intelectual individual origina-se a partir da primeira infância – 0 a 6 anos. De acordo com estudo, a desigualdade de oportunidades educacional gera substanciais diferenças na habilidade cognitiva e não cognitiva individual. Além disso, o estudo constata que economias que apresentam ambiente educacional desfavorável tendem a proporcionar baixa produtividade e menores níveis de renda per capita.

Em outro estudo, Herskind *et al.* (1996), relata que o papel da genética varia em torno de 0,5 contra 0,35 do ambiente compartilhado na composição de *g* em indivíduos com idade

entre 4 a 20 anos. O estudo também encontra evidências que as influências ambientais compartilhadas caem virtualmente à zero na idade adulta. Por sua vez, Chipuer, Rovine e Plomin (1990), apontam que a herdabilidade genética é responsável por 50% da composição de g .¹ Entretanto, o estudo afirma que a participação relativa das influências genéticas e ambientais sobre a capacidade cognitiva difere substancialmente ao longo da vida, com o aumento da herdabilidade e queda da influência ambiental compartilhada desde a infância e adolescência até a idade adulta. Nessa mesma direção, Wilson (1983), sugere que o ambiente compartilhado responde por 60% da variância em g na primeira infância.

Nessa mesma temática, Trouton, Spinath e Plomin (2002), investigam as contribuições genéticas e ambientais em gêmeos de 2, 3, 4 e 7 anos de idade, após mudanças e continuidade na linguagem e desenvolvimento cognitivo. Os resultados sugerem que os mesmos genes contribuem largamente para ambos os problemas de linguagem e desenvolvimento cognitivo e afetam o desenvolvimento normal e anormal individual. Conforme podemos perceber não há um consenso empírico sobre esses efeitos.

Entretanto, Ronald, Spinath e Plomin (2002), destacam a divergência entre os defensores e contrários da contabilização do talento individual. Para os adeptos, há fatores biológicos individuais pré-disponíveis como um potencial para o talento que pode ser identificado em algumas crianças e outras não. Para os opostos a esta concepção, a aquisição de habilidades por meio de exposição precoce, prática e outros processos estão mais relacionados a fatores ambientais. Contudo, os principais resultados indicam que mais de 70% da diferença média entre o grupo de pontuação elevada e toda a amostra – crianças de 2 a 4 anos de idade – são determinadas por influências ambientais compartilhadas. E somente um quinto da diferença média é devido a fatores genéticos. Estes resultados corroboram o argumento de que os primeiros "talentos" são movidos por fatores ambientais compartilhados.

Nessa mesma linha de pensamento, Spinath *et al.* (2003), destaca que g é uma das construções mais confiáveis e válidas sobre o comportamento humano. Segundo os autores, a estabilidade de g a longo prazo após a infância é maior do que qualquer outro traço comportamental e, portanto, é um dos principais determinantes de resultados importantes, tais como, níveis educacionais e ocupacionais. Nesse mesmo estudo, os resultados remetem três importantes questões. Primeiro, o fenótipo g é claramente evidente nos primeiros anos de vida. Segundo, a influência genética é em torno de 20% a 30% menor na primeira infância do que a encontrada na infância média. Isto é, algo próximo de 40% inferior ao observado na adolescência, e por volta de 50% menor que o valor estimado após a adolescência. Em terceiro lugar, a influência ambiental compartilhada é maior na infância do que após a adolescência, quando sua importância declina a níveis insignificantes. No entanto, as estimativas captam alta influência do ambiente compartilhado, algo em torno de 60%. Embora, os autores justificam que esse valor pode incluir influências específicas de gêmeos e, talvez, seja produto das medidas utilizadas.

Ratificando os resultados anteriores, Plomin e Spinath (2004), destaca que as duas descobertas genéticas mais importantes são: i) hereditariedade da inteligência aumenta ao longo do tempo de vida; ii) os mesmos genes afetam diversas habilidades cognitivas.

Nessa perspectiva, fica evidente a importância em controlar os fatores genéticos e ambientais compartilhados. Assim sendo, acreditamos haver um considerável ganho nesse estudo comparado aos demais estudos de IOPs. Pois, ao utilizar uma amostra de gêmeos monozigóticos controlamos de modo significativo as limitações a priori estabelecidas. Particularmente, controlar os efeitos da sorte genética – herdabilidade – ou constitutiva,

¹ O fenótipo são as características observáveis ou caracteres de um organismo ou população como, por exemplo: morfologia, desenvolvimento, propriedades bioquímicas ou fisiológicas e comportamento.

habilidade cognitiva geral, ambiente compartilhado, motivação, entre outros fatores não observáveis determinantes do resultado individual é sem dúvidas o maior ganho possibilitado pelo uso da base de dados utilizada.

3 ESTRATÉGIA EMPÍRICA

Este estudo utiliza a princípio o método de inferências contrafatuais *Propensity Score Matching* desenvolvido por Rosenbaum e Rubin (1983), doravante denominado *PSM*. Seguindo Cameron e Trivedi (2005), o *PSM* é um método de correspondência inexata, que ao invés de fazer correspondência com as variáveis explicativas faz correspondência com o escore de propensão. Neste processo, revela a probabilidade condicional de receber tratamento dado X , expressado por $p(X)$, como uma medida de correspondência. Conforme observado, se os dados justificam a correspondência em X , então, a correspondência com base no *PSM* é também justificada.

De acordo com a teoria contrafactual, a situação ideal seria observar o mesmo indivíduo em situações distintas, ou seja, participando e não participando de um determinado evento. Contudo, como não é possível observar o mesmo indivíduo, por exemplo, sendo fumante e não fumante, uma saída para isso é apurar a causalidade e tamanho da influência de determinados fatos através da comparação de dois grupos estatisticamente semelhantes.

Nesse sentido, este método, nada mais é, do que um estimador de correspondência baseado na variância que faz uso do escore de propensão estimado $\hat{P}(X, Z)$ a fim de buscar e impor pesos aos pares simétricos (i, j) com status de tratamento opostos, dados os valores de X e Z .

O objetivo, na verdade, é procurar uma correspondência exata, ou seja, o *PSM* busca encontrar pares para os quais a soma das probabilidades dentro de cada par seja igual a um. Caso isso não ocorra, então, as unidades de comparação são aquelas cujo escore de propensão é suficientemente próximo da unidade tratada. Em outras palavras, o *PSM* procura estabelecer um grupo de controle idêntico ao grupo de tratamento em função de certas características observáveis.

A fundamental hipótese assumida no *PSM* é que o vetor de covariadas X contempla todas as características responsáveis pelo resultado potencial na ausência do tratamento que o indivíduo possui ao decidir ingressar ou não em um determinado tratamento. Essa hipótese é denominada de condição de não confundimento ou seleção nos observáveis.

$$y_i(0) \perp D_i | X_i \quad (2)$$

Desse modo, para obter-se o efeito tratamento médio via *PSM*, faz-se necessário possuir indivíduos no grupo de controle que corresponda identicamente a cada indivíduo no grupo de tratamento, a fim de apurar qual seria seu resultado caso o mesmo não estivesse submetido ao tratamento proposto. Dito de outra forma faz-se imprescindível obter um grupo de controle que possua as mesmas características observáveis contidas no vetor X que as correspondentes do grupo de tratamento. Nessa configuração, deve-se atender a segunda hipótese, denominada hipótese de sobreposição. Em linhas gerais, a mesma estabelece que o grupo de controle contenha indivíduos com características similares aos indivíduos pertencentes ao grupo de tratamento. Tal hipótese é expressa por.

$$P[D_i | X_i < 1] \quad (3)$$

Uma limitação evidente do *PSM* é que ao assumir as hipóteses em (2) e (3), admite-se que o vetor de características observáveis contempla todos os fatores que determinam os indivíduos participar ou não tratamento. Isto é, a decisão de participar do tratamento é independente de fatores não observáveis. Dito de outro modo, se os fatores não observáveis determinarem a decisão de participar do tratamento, como também, influenciar o resultado potencial individual, o *PSM* está sujeito ao viés de seleção.

Contudo, a sensibilidade dos resultados perante o método utilizado não é passível de uma resposta simples. Pois, os resultados podem variar em diferentes amostras, dependendo do grau de sobreposição entre as observações do grupo de tratamento e do grupo de controle. Por um lado, caso os grupos sejam idênticos, em termos de sobreposição substancial dos seus escores de propensão, e se o grupo de controle é grande, então, os matches serão mais acessíveis. De modo que, combinar através do método de substituição é a melhor alternativa. Por outro lado, caso o grupo de controle seja pequeno e díspar em relação às observações do grupo de tratamento, então, os matches podem ser insatisfatórios, e, portanto, deve-se adotar o método sem substituição.

Nessa configuração, denote o grupo de controle do caso tratado i com características X_i como o conjunto $A_j(X) = \{j | X_j \in c(X_i)\}$, onde $c(X_i)$ representa as características da vizinhança de X_i . Além disso, suponha N_c representar o número de casos do grupo de controle e $w(i, j)$ o peso dado ao j th não tratado com o i th caso tratado, onde $\sum_j w(i, j) = 1$, sendo a fórmula geral do *matching* do estimador Efeito Tratamento sobre o Tratado – ATET – expressada por:

$$\Delta^M = \frac{1}{N_T} \sum_{i \in \{D=1\}} y_{1,i} - \sum_j w(i, j) y_{0,j}. \quad (4)$$

Onde $0 < w(i, j) \leq 1$ e $\{D = 1\}$ denota o conjunto dos indivíduos tratados, e j é um elemento do conjunto de *matching* de unidades de comparação. Diferentes *matching* estimadores são gerados por variar a escolha de $w(i, j)$. Por exemplo, o *matching* simples compara células com exatamente a mesma x discreta.

$$\Delta^M = \sum_{k \in \{D=1\}} w_k [\bar{y}_{1,k} - \bar{y}_{0,k}]. \quad (5)$$

Sendo \bar{y}_1 o resultado médio do tratado e \bar{y}_0 representa o resultado médio do não tratado, e w_k é o peso atribuído da k th célula. Ou seja, a fração das observações na célula k . Segundo Cameron e Trivedi (2005), um caso específico é encontrado em Dehejia e Wahba, (2002).

$$\Delta^M = \frac{1}{N_T} \sum_i \left(y_1 - \frac{1}{N_{c,i}} \sum_{j \in \{D=0\}} y_j \right) \quad (6)$$

Em que N_T é o número de casos do grupo de tratamento ($D = 1$) e $N_{c,i}$ corresponde ao número de casos do grupo de controle correspondente a i th observação. Por sua vez, o método de *matching* baseado na vizinhança mais próxima, o conjunto $A_i(X) = \{j | \min_j \|X_i - X_j\|\}$. Aonde $\| \cdot \|$ denota a distância euclidiana entre os vetores. Se $w(i, j) = 1$ na equação (4.1), então, $J \in A_i(X)$, e zero caso contrário, de forma que, esta especificação utiliza apenas um caso para construir o grupo de controle para os casos tratados. Outro tipo de estimador gerado é o *matching kernel*, o qual é expresso pela equação 3.7 a seguir.

$$w(i, j) = \frac{K(X_j - X_i)}{\sum_{j=1}^{N_{c,i}} K(X_j - X_i)}. \quad (7)$$

Onde K representa a função kernel anteriormente citada. Vale ressaltar que, uma das principais vantagens dos métodos apresentados até o momento, e que, os mesmos tem a vantagem de não assumir determinadas suposições da forma funcional para as equações de resultados na estimativa ATET e pode ser estimado por valores específicos de X . Por outro lado, tem a desvantagem de que se X apresenta alta dimensão, então o número de matches pode ser pequeno. Nesses casos, de acordo com a metodologia, o *matching* baseado em uma métrica de valor escalar torna-se mais atrativo.

Por outro lado, o método de estratificação ou *matching* intervalo baseia-se na ideia de dividir o intervalo de variação do *propensity score* em intervalos, de forma que, em cada intervalo da unidade de tratamento e controle há, em média, o mesmo *propensity score*.

Podem-se usar os mesmos blocos identificados pelo algoritmo utilizado para calcular os *propensity score*. Em seguida, calcula-se a diferença entre os resultados médios do grupo tratados e de controle. O ATET, na realidade é a média ponderada destas diferenças, com pesos sendo determinados pela distribuição das unidades tratadas através dos blocos. Uma das desvantagens deste método é que ele descarta observações nos blocos nas quais unidades tratadas ou de controle estão ausentes.

$$ATE_{b}^S = (N_b^T)^{-1} \sum_{i \in I(b)} Y_{1,i} - (N_b^C)^{-1} \sum_{j \in I(b)} Y_{0,j}. \quad (8)$$

Em que $I(b)$ é o conjunto de unidades no bloco b , N_b^T é o número de unidades tratadas no b th bloco. Assim, portanto, o efeito tratamento baseado na estratificação é definido como:

$$ATE^S = \sum_{b=1}^B ATE_b^S \times \left[\frac{\sum_{i \in I(b)} D_i}{\sum D_i} \right]. \quad (9)$$

Onde o termo entre parenteses representa o peso de cada bloco dado à fração correspondente de cada unidade tratada e B é o número total de blocos.

Por sua vez, no *matching* raio o conjunto $A_i(p(X)) = \{P_j | \min_j \|P_i - P_j\| < r\}$, é baseado no *propensity score*. Isto diz que todos os casos pertencentes ao grupo de controle estimados através do *propensity score*, que pertençam ao raio r , é *matching* do i th caso tratado. Adicionalmente, pode-se expressar ATE e ATET em termos em $p(X)$, assumindo a condição de sobreposição $0 < p(X) < 1$. Os dois principais resultados são:

$$ATE = E \left[\frac{(D - p(X))y}{p(x)(1 - p(X))} \right]. \quad (10)$$

$$ATE^S = E \left[\frac{(D - p(X))y}{\Pr[D = 1](1 - p(X))} \right]. \quad (11)$$

Como relatado anteriormente, a fim de superar possível endogeneidade devido à omissão de variáveis que captem características, como, habilidade cognitiva geral, herdabilidade genética, fatores ambientais compartilhados, como também, outros aspectos não observáveis ou não disponíveis adota-se a seguir uma proposta similar à desenvolvida por Ashenfelter e Rouse (1999). Na realidade, faz-me um *matching* – pareamento – simples. Em outras palavras, coparam-se os resultados obtidos entre os próprios irmãos gêmeos. Uma vez que, a estratégia anteriormente empregada, possibilita o uso de *matches* – pareamento de não gêmeos – distintos devido à variável de tratamento não necessariamente ser ponto de diferença entre os gêmeos. Por exemplo, é perfeitamente possível que ambos os gêmeos sejam sindicalizados ou não. Ratificando tais condições, os próprios dados já justificam tal preocupação, pois apenas 23,53% dos indivíduos da amostra são sindicalizados.

Como visto, a principal finalidade dessa segunda abordagem é justamente levar em conta fatores que limitam significativamente as estimações da equação (1). Pois, ao comparar – *matching* simples – irmãos gêmeos monozigóticos entre si controlam-se imediatamente as características não observáveis citadas. Nesse sentido, apresenta-se a seguir uma breve ilustração da estratégia adotada. Isto é, admita que o *matching* recorra ao melhor controle de pareamento entre o grupo de tratados e não tratados, e, portanto, utilizando como unidade de comparação o próprio irmão gêmeo. Para tanto, utilizamos as diferenças existentes entre as observações dos mesmos com o propósito de controlar possíveis efeitos fixos existentes.

Nessa perspectiva, admita que os resultados individuais alcançados – salário hora – expressos por y_{1i} e y_{2i} por cada par de gêmeos sejam respectivamente.

$$y_{1i} = X_i^1 \beta^1 + C_i + u_i. \quad (12)$$

$$y_{2i} = X_i^2 \beta^2 + C_i + u_i. \quad (13)$$

Onde X_i^n denota as variáveis observáveis de cada gêmeo. Como por exemplo, educação individual e parental, idade, sexo, estado civil, sindicalização, etc.; C_i representa os fatores

fixos não observáveis, aqui assumidos como, por exemplo: habilidade cognitiva geral, herdabilidade genética, fatores ambientais compartilhados, como também, outras particularidades individuais não observáveis ou disponíveis. E u_i o termo de erro. Nesse sentido, se tomarmos a diferença entre as variáveis contidas em (12) e (13) obtém-se:

$$y_{1i} - y_{2i} = (X_i^1 \beta^1 - X_i^2 \beta^2) + (C_i - C_i) + u_i. \quad (14)$$

$$dy_i = dX_i \theta + \varepsilon_i. \quad (15)$$

Como assumimos que os fatores contidos em C_i são semelhantes. Isto é, admitimos que não haja diferenças significativas em habilidade cognitiva geral, herdabilidade genética, ambiente compartilhado, entre outros fatores entre os irmãos gêmeos. Logo, $(C_i - C_i) = 0$. Isto é, anulam-se os efeitos fixos não diretamente observáveis. Diante disso, os resultados propostos em (15) apresentam-se como o mais apropriado aos objetivos propostos. Posto isso, faz-se a seguir a descrição e análise parcial dos dados.

4 Descrição e Análise dos Dados

Os resultados reportados neste trabalho foram extraídos de uma amostra de 680 gêmeos monozigóticos – idênticos – coletados no Festival do Dia Anual dos Gêmeos em Twinsburg, Ohio, ocorridos no período de 1991 a 1994. Enfatiza-se que os mesmos estão disponibilizados no *homepage* da Universidade de Princeton. Os dados foram originalmente utilizados no trabalho de Ashenfelter e Krueger (1994).

Quadro 1 – Conjunto de Informações

Variáveis	Descrição
<i>Salário hora</i>	Log do salário hora individual.
<i>Nível Educacional</i>	Anos de estudo individual.
<i>Educação Parental</i>	Anos de estudo da mãe e do pai.
<i>Idade</i>	Idade individual.
<i>Sexo</i>	Dummy para sexo, que assume valor 0 para mulheres e 1 para os homens.
<i>Raça</i>	Dummy para raça, que assume valor 0 para negros e 1 brancos.
<i>Autônomo</i>	Dummy para trabalho, que assume valor 0 para empregado e 1 autônomo.
<i>Sindicalizado</i>	Dummy para sindicalizado, que assume valor 1 para sindicalizados e 0 caso contrário.
<i>Casado</i>	Dummy para união, que assume valor 1 para casados e 0 caso contrário.
<i>Irmãos</i>	Números de irmãos.
<i>Diferenças</i>	Diferença entre as variáveis anteriormente descritas.

Fonte: Elaboração própria.

Todavia, conforme relatado anteriormente, a principal característica para adoção dessa base de dados é a possibilidade de controlar os fatores não observáveis até então ignorados, em especial, em aplicações empíricas. Além disso, o quadro 1 destaca as variáveis utilizadas nesta pesquisa. Adicionalmente foram acrescentadas outras especificidades nas variáveis. Por exemplo, a diferença observada entre os salários horas dos gêmeos e interações entre as covariadas.

Entretanto, cabe ressaltar que foram realizados dois testes de especificação do modelo com a finalidade de elencar o melhor grupo de covariadas.² Os dois testes atestam não haver problemas de variáveis omitidas.³ Além da observância dos resultados dos testes, procurou-se

² Os resultados dos testes de especificação – ovtest e linktest – não rejeitaram a hipótese nula.

³ Maiores detalhes em Chatterjee e Hadi (2013).

optar também por variáveis clássicas utilizadas na literatura.⁴ Guiado por isto, regrediu-se o log do salário individual contra o conjunto de covariadas especificadas no quadro 1.

Posto isso, faz-se agora um breve relato a respeito do conjunto de dados. A tabela 1 sumariza alguns predicados da base.

Tabela 1 – Análise Descritiva

Características Individuais					
Sexo		Raça		Autônomo	Sindicalizados
Masculino 40,29%	Feminino 59,71%	Branco 92,21	Negro 7,79%	13,68%	23,53%
Nível Educacional					
Própria		Pai		Mãe	
Médio 36,48%	Superior 63,52%	Médio 69,56%	Superior 30,44%	Médio 75,59%	Superior 24,41%

Fonte: Elaboração própria, a partir de dados.

Por exemplo, observa-se que aproximadamente 60% dos gêmeos são do sexo feminino; onde somente 8% são declarados negros; 11,02% trabalham por conta própria; 89% são trabalhadores formais, e destes, 21,91% são sindicalizados. Adicionalmente, constata-se que a 24% a 30,5%, respectivamente, mãe e pai, possuem no mínimo nível superior. Adicionalmente, verifica-se que em torno de 64% dos gêmeos possui nível superior. Esse último resultado denota teoricamente uma elevada mobilidade educacional entre as gerações. A seguir apresentam-se os principais resultados encontrados.

5 PRINCIPAIS RESULTADOS

Esta seção designa-se a apresentar e discutir os principais resultados desta pesquisa. Primeiramente, faz-se uma análise das diferenças nos resultados observados – salário hora – dos gêmeos e sua opção de escolha quanto a ser sindicalizado ou não. Uma vez que se torna necessária utilizar uma variável de comparação entre considerar ou não aspectos não observáveis.

Tabela 2 – Diferença nas Médias

Log do Salário	
Sindicalizado	0.2832*** (0.0456)
Intercepto	2.3769*** (0.0286)
<i>Observações</i>	680

Fonte: Elaboração própria. Desvios Padrão (*).

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

Inicialmente, a tabela 2 reporta os resultados do log do salário hora – variável dependente – contra a dummy de filiação sindical – tratamento – que assume valor um caso o trabalhador seja sindicalizado e zero caso contrário. Os resultados sugerem uma diferença salarial média de 28% entre os grupos analisados. Em suma, ser sindicalizado impacta em média 0,28 vezes a mais no resultado econômico individual do que em indivíduos que não são sindicalizados. Contudo, tais resultados não devem ser considerados como verdadeiros, uma

⁴ Por exemplo, Anshelter e Krueger (1994).

vez que, não foi empregada nenhuma forma – variável – de controle. Os resultados foram obtidos usando o método de *bootstrap* com 1.000 replicações.

Por outro lado, a tabela 3 relata a distribuição amostral perante a participação na variável de comparação – Sindicalização – adotada nesta pesquisa. A mesma informa que 76,47% dos indivíduos avaliados não são sindicalizados, ou seja, pertencem ao grupo de controle, contra 23,53% que formam o grupo de tratamento. Destaca-se que a propriedade de balanceamento, isto é, a hipótese 2 é satisfeita em todas as abordagens empregadas.

Tabela 3 – Distribuição do Propensity Score

Grupo	Frequência	%
Controle – Não sindicalizado	520	76,47
Tratamento – Sindicalizado	160	23,53
Observações	680	100

Fonte: Elaboração própria.

Verifica-se também, a partir da tabela 4, que em todos os casos, o escore de propensão foi restrito à região de suporte comum. Outra característica que merece destaque é a representatividade do grupo de tratamento perante o grupo de controle nos blocos estimados, com exceção do último bloco. Posto isto, estima-se a probabilidade de ser sindicalizado, tanto para os participantes, quanto para os que não fazem parte da entidade sindical. As variáveis de controle elegidas para estimar o *PSM* são as mesmas empregadas para delinear os grupos analisados. A tabela 6 reporta os fatores que explicam a sindicalização.

Tabela 4 – Impacto da Sindicalização

Mínimo $\hat{P}(x)$	Não Tratado	Tratado	Total
0.0	302	61	363
0.1	86	20	106
0.2	59	21	80
0.3	36	15	51
0.4	20	16	36
0.5	10	14	24
0.6	06	08	14
0.7	01	03	04
0.8	00	02	02
Total	520	160	680

Fonte: Elaboração própria.

As variáveis as quais os indivíduos não possuem qualquer responsabilidade individual, como idade, raça, gênero, educação parental e número de irmãos foram incluídas em função de representar o background social – sorte social – amplamente difundida na literatura como fonte de desigualdade injusta. Por outro lado, a adoção da idade ao quadrado busca captar uma tendência negativa nos salários à medida que os indivíduos envelhecem. Quanto a outras variáveis, por exemplo, se os indivíduos são trabalhadores autônomos, nível de estabilidade e se os mesmos são casados representam fatores referentes ao esforço individual.

Adicionalmente, destaca-se que a variável dependente – salário hora – empregada na análise, capta características relevantes do esforço individual empregado. Os resultados apresentados informam, por exemplo, que a idade influencia em 0.27 na probabilidade de sindicalizar tanto para o grupo de controle quanto para o grupo de tratamento. Entretanto, um

resultado que chama a atenção é o estado civil – casado – do indivíduo. Pois, realmente parece ser bem intuitivo que indivíduos com essas características busquem uma maior segurança jurídica. Ou seja, o simples fato de ser casado aumenta a probabilidade de possuir filiação sindical. Entretanto, embora o coeficiente estimado seja positivo, o mesmo não é estatisticamente diferente de zero.

Tabela 5 – Propensity Score Matching – Sindicalização

Descrição	B _{PMS}	σ	t	p-value
Idade	0.273**	0.084	3.22	0.001
Idade ²	-0.003**	0.000	-3.37	0.001
Branco	-1.249**	0.398	-3.14	0.002
Feminino	-0.478	0.268	-1.78	0.074
Autônomos	-2.252**	0.649	-3.47	0.001
Casados	0.333	0.328	1.02	0.310
Estabilidade	0.051**	0.017	2.96	0.003
Nº Irmãos	-0.006	0.057	-0.11	0.914
Educ. Pai	-0.123**	0.052	-2.34	0.019
Educ. Mãe	-0.032	0.065	-0.50	0.617
Intercepto	-3.872**	1.860	-2.08	0.037

Fonte: Elaboração própria.

* p<0.10; ** p<0.05; *** p<0.01

Por sua vez, os resultados reportados na tabela 6 revelam que o efeito tratamento sobre o tratado – filiação sindical – referente ao salário hora é em torno 20% maior comparado ao grupo de controle. Em suma, a grandeza econômica do tratamento é bastante expressiva. Sendo mais específico, a decisão de se sindicalizar aumenta em 20% a probabilidade de obter um salário hora 28% maior dos que os indivíduos não sindicalizados. Contudo, o efeito tratamento médio informa um impacto positivo de 0.135 na hora salarial dos tratados.

Outra questão a destacar, é que a influência apurada apresenta-se estatisticamente significativa em todos os modelos adotados. Especificamente, o modelo de distância com $r = 0.01$ e $r = 0.05$ são os que apresentam os coeficientes mais expressivos. Esses resultados parecem evidenciar o principal objetivo desta pesquisa, isto é, permitimos inferir sobre a existência de fatores – *sorte genética* – determinísticos contidos no termo de erro, até então, considerados como aleatórios, ao menos que, empiricamente.

Tabela 6 – Efeito Tratamento Sobre o Tratado – ATT – Sorte Opção

Matching	Tratados	Não Tratados	ATT	ATE	σ	t
<i>PSmatching logit</i>	103	335	0.182***	0.135***	0.069	4.230
	103	335	(0.182)***	-	-	-
<i>Vizinhança</i>	160	256	0.191***	-	0.086	2.234
	160	256	(0.191)***	-	0.080	2.157
<i>Radius r = 0.01</i>	97	286	0.262***	-	0.061	4.299
	97	270	(0.248)***	-	0.068	3.646
<i>Radius r = 0.05</i>	101	335	0.247***	-	0.059	4.218
	101	301	(0.242)***	-	0.058	4.159
<i>Estratificação</i>	101	303	0.198***	-	0.059	3.382
	101	303	(0.198)***	-	0.061	3.255
<i>Kernel</i>	160	520	0.201***	-	-	-
	103	301	(0.199)***	-	0.061	3.244

Fonte: Elaboração própria. * p<0.10, ** p<0.05, *** p<0.01

Valores em parênteses obtidos por bootstrap para 1000 replicações.

Contudo, embora, as estimações apresentem uma convergência em torno dos valores relatados, com exceção, ao método da distância através de um raio. Os coeficientes encontrados neste último modelo estão em torno de 25% mais significativos. Apresenta-se a seguir os resultados da segunda estratégia empírica, a qual busca tratar a omissão das características individuais não observáveis. Isto é, essa estratégia objetiva suplantar importantes limitações empíricas e estruturais presentes na equação (1). Nessa perspectiva, estimamos o modelo expresso pela equação (12) para o mesmo grupo de covariadas. Os resultados mostram que o efeito médio da variável de tratamento – sindicalização – sobre o salário hora se reduziu significativamente, uma vez que, os fatores omitidos superestimam os coeficientes obtidos na tabela 6.

Tabela 7 – Influência da Sorte Opção no Salário Hora

Diferenças	β_{OLS}	β_{1dif}	β_{1difIV}	β_{1dEF}
Tratamento	0.087**	0.091*	0.091*	0.091*
Educação	0.079***	0.081***	0.081*	0.081***
Casado	0.034	0.024	0.024	0.024
Estabilidade	0.023***	0.023***	0.023***	0.023***
Autônomos	-0.007	-	-	-
Intercepto	0.006	-	-	-
Observações	680	680	680	680

Fonte: Elaboração própria. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Os resultados reportados na tabela 7 ratificam as expectativas. Pois, de acordo com os números apresentados, nos quatro modelos estimados a influência média dessa componente da sorte, situa-se próximo a 0.09.⁵ Em suma, a opção de sindicalizar, controlado por fatores observáveis e não observáveis, responde em média por 9% das chances do indivíduo situar-se na distribuição salarial superior. Em linhas gerais, a sorte opção representada aqui pela filiação sindical, faz com que o indivíduo tenha 0.09 maiores oportunidades de estar entre os que têm um salário hora em média 28% superior.

6 ALGUMAS CONSIDERAÇÕES

Nesse trabalho, o problema principal era investigar se qual a importância em considerar – sorte genética – as características não observáveis na determinação de um resultado econômico – salário hora – individual. Isto é, buscar superar algumas limitações empíricas e estruturais na equação fundamental dos atuais modelos de igualdade de oportunidades. Para tanto, regrediu-se a equação (1) através do emprego da equação (12). Em palavras, controlou-se a influência de determinantes não observáveis ou não disponíveis na maioria das bases de dados nos resultados – salário hora – individuais. Em linhas gerais, buscou-se avançar estruturalmente e empiricamente a teoria de igualdade de oportunidades. Pois, fez-se uma aplicação que parece ser mais realista e procura atender tanto a teoria filosófica quanto econômica. Isto é, primeiramente, empregou-se o método *PSM* em cinco abordagens distintas. Adicionalmente, através do emprego de um enfoque alternativo consideraram-se características individuais não observáveis, como, habilidade cognitiva geral, herdabilidade genética, ambiente compartilhado, entre outros fatores. Para tanto, utilizou-se uma amostra composta por 680 gêmeos monozigóticos.

⁵ β_{OLS} – mínimos quadrados ordinários; β_{1dif} – primeiras diferenças; β_{1dEF} – primeiras diferenças variáveis instrumentais e β_{1d} – primeiras diferenças efeitos fixos.

Os resultados indicaram que a variável – filiação sindical – de comparação possui influência substancial no resultado salarial da amostra analisada. Na primeira abordagem, sem considerar os aspectos não observáveis, como alguns fatores observáveis, embora não disponíveis, os valores informam que o grupo de tratamento tem em torno de 20% maiores chances comparado aos indivíduos que compõem o grupo de controle de se localizar na distribuição que recebe em média 28% a mais. Por sua vez, na segunda abordagem, considerando agora, as limitações enumeradas em diversas oportunidades anteriores, essa influência declina para algo em torno de 9%.

Contudo, mesmo assim, a amplitude econômica desse efeito é sem dúvida muito significativa. Sintetizando, os principais objetivos traçados foram alcançados. Isto é, encontraram-se evidências significativas na amostra analisada que parte da sorte, até então desconsideradas nos modelos de IOP's, tem participação significativa no resultado individual. Em outras palavras, o termo de erro dos atuais modelos de IOPs não é formado por componentes totalmente aleatórios. Espera-se que futuras pesquisas sobre o tema em questão, lance luz sobre a participação – quantitativa – das características não observáveis. Por exemplo, o quanto características como habilidade cognitiva geral, herdabilidade genética, ambiente compartilhado respondem pelo resultado individual. Como também, adotar abordagens que busquem decompor o termo de erro, particularmente, sobre os efeitos das componentes do fator sorte.

REFERÊNCIAS

ASHENFELTER, Orley C.; KRUEGER, Alan B. **Estimates of the Economic Returns to Schooling from a New Sample of Twins**. American Economic Review, v. 84, n. 5, p. 1157-73, 1994.

ASHENFELTER, Orley; ROUSE, Cecilia. **Income, schooling, and ability: Evidence from a new sample of identical twins**. The Quarterly Journal of Economics, v. 113, n. 1, p. 253-284, 1998.

BOURGUIGNON, François; FERREIRA, Francisco HG; MENENDEZ, Marta. **Inequality of opportunity in Brazil**. Review of income and Wealth, v. 53, n. 4, p. 585-618, 2007.

CAMERON, A. Colin; TRIVEDI, Pravin K. **Microeconometrics: methods and applications**. Cambridge university press, 2005.

CHATTERJEE, Samprit; HADI, Ali S. **Regression analysis by example**. John Wiley & Sons, 2013.

CHIPUER, Heather M.; ROVINE, Michael J.; PLOMIN, Robert. **LISREL modeling: Genetic and environmental influences on IQ revisited**. Intelligence, v. 14, n. 1, p. 11-29, 1990.

DEHEJIA, Rajeev H.; WAHBA, Sadek. **Propensity score-matching methods for nonexperimental causal studies**. Review of Economics and statistics, v. 84, n. 1, p. 151-161, 2002.

DWORKIN, Ronald. **What is equality? Part 2: Equality of resources.** Philosophy & Public Affairs, p. 283-345, 1981.

HERSKIND, Anne Maria et al. **The heritability of human longevity: a population-based study of 2872 Danish twin pairs born 1870–1900.** Human genetics, v. 97, n. 3, p. 319-323, 1996.

LEFRANC, Arnaud; PISTOLESI, Nicolas; TRANNOY, Alain. **Equality of opportunity and luck: Definitions and testable conditions, with an application to income in France.** Journal of Public Economics, v. 93, n. 11, p. 1189-1207, 2009.

PLOMIN, Robert; SPINATH, Frank M. **Intelligence: genetics, genes, and genomics.** Journal of personality and social psychology, v. 86, n. 1, p. 112, 2004.

ROEMER, John E. **Equality of opportunity.** Cambridge: Harvard University Press, 1998.

RONALD, Angelica; SPINATH, Frank M.; PLOMIN, Robert. **The aetiology of high cognitive ability in early childhood.** High Ability Studies, v. 13, n. 2, p. 103-114, 2002.

ROSENBAUM, Paul R.; RUBIN, Donald B. **The central role of the propensity score in observational studies for causal effects.** Biometrika, v. 70, n. 1, p. 41-55, 1983.

SPEARMAN, Charles. **"General Intelligence," Objectively Determined and Measured.** The American Journal of Psychology, v. 15, n. 2, p. 201-292, 1904.

SPINATH, Frank M. et al. **Phenotypic early in life: On the etiology of general cognitive ability in a large population sample of twin children aged 2–4 years.** Intelligence, v. 31, n. 2, p. 195-210, 2003.

TROUTON, Alexandra; SPINATH, Frank M.; PLOMIN, Robert. **Twins early development study (TEDS): a multivariate, longitudinal genetic investigation of language, cognition and behavior problems in childhood.** Twin Research, v. 5, n. 05, p. 444-448, 2002.

VALLENTYNE, Peter. **Brute Luck, Option Luck, and Equality of Initial Opportunities.** Ethics, v. 112, n. 3, p. 529-557, 2002.