

**UNIVERSIDADE ESTADUAL DE PONTA GROSSA
SETOR DE CIÊNCIAS SOCIAIS APLICADAS
DEPARTAMENTO DE ECONOMIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA**

PAULA GABRIELLA AZEVÊDO LEITE

**AVALIAÇÃO DE IMPACTO DO PROGRAMA JOVEM APRENDIZ ENTRE OS
ANOS DE 2014 E 2018**

**PONTA GROSSA
2023**

PAULA GABRIELLA AZEVÊDO LEITE

**AVALIAÇÃO DE IMPACTO DO PROGRAMA JOVEM APRENDIZ ENTRE OS
ANOS DE 2014 E 2018**

Dissertação apresentada para obtenção do título de mestre
na Universidade Estadual de Ponta Grossa, Área de
Economia.

Orientadora: Prof^a Dra Cleise Maria de Almeida Tupich
Hilgemberg.

Corientadora: Prof^a Dra Luma de Oliveira

**PONTA GROSSA
2023**

L533 Leite, Paula Gabriella Azevedo
Avaliação de Impacto do Programa Jovem Aprendiz entre os anos de 2014 e 2018 / Paula Gabriella Azevedo Leite. Ponta Grossa, 2023.
61 f.

Dissertação (Mestrado em Economia - Área de Concentração: Economia),
Universidade Estadual de Ponta Grossa.

Orientadora: Profa. Dra. Cleise Maria de Almeida Tupich Hilgemberg.
Coorientadora: Profa. Dra. Luma de Oliveira.

1. Jovem aprendiz. 2. Mercado - trabalho. 3. Políticas públicas. 4. Propensity
Score Matching. I. Hilgemberg, Cleise Maria de Almeida Tupich. II. Oliveira, Luma
de. III. Universidade Estadual de Ponta Grossa. Economia. IV.T.

CDD: 330.9

PAULA GABRIELLA LETTE

AVALIAÇÃO DE IMPACTO DO PROGRAMA JOVEM APRENDIZ
ENTRE 2014 E 2018

Dissertação aprovada como requisito parcial para obtenção do grau de
Mestre no Programa de Pós-Graduação em Economia, da Universidade
Estadual de Ponta Grossa, pela seguinte banca examinadora:

Ponta Grossa, 28 de novembro de 2023



Cleise Maria de Almeida Tupich Hilgemberg – Universidade Estadual de
Ponta Grossa



Marina Silva da Cunha - Universidade Estadual de Maringá



Luma de Oliveira - Universidade Estadual de Ponta Grossa

AGRADECIMENTOS

Apesar de diversas angustias foi possível chegar ao fim deste ciclo, mas que não seria possível se este caminho fosse solitário.

Agradeço a minha família pelo apoio e incentivo.

À minha estimada orientadora Prof^a Dra Cleise Maria de Almeida Tupich Hilgemberg e à Prof^a Dra Luma de Oliveira que me acolheram, apoiaram e direcionaram no processo de elaboração da dissertação. Minha gratidão pela paciência, sinceridade, acolhimento, seriedade e compromisso destas profissionais da educação. Estendo meus agradecimentos aos docentes do PPGECON-UEPG e aos funcionários do Programa de Pós Graduação em Economia da UEPG, em especial à Vilma que me auxiliou em questões burocráticas relacionadas aos processos administrativos da universidade.

Às agências de fomento à pesquisa CAPES e CNPq que, por determinado período, contribuiu para a minha permanência no programa, assim como o povo brasileiro que, através do pagamento de tributos que pagamos direta ou indiretamente por meio de, financiou parte da minha caminhada nesta pesquisa por meio da bolsa de estudos recebidos durante 12 meses iniciais.

À minha terapeuta, Vida Sousa Prado que não me deixou esquecer o motivo de continuar, ressignificar e buscar sentido para cada etapa do processo da escrita.

Às amigas que encontrei, cultivei, reaproximei e as que deixei ir. Muito obrigada a todas palavras de apoio, motivação e compreensão, em especial Ana Beatriz, Mateus Girola, Camila Bona, Marcionilia Paula, Layanne Marinho.

Por fim, agradeço a Deus que me guia, protege e ampara colocando as pessoas acima citadas para auxiliar nesta jornada.

[...]

Bom mesmo é achar que a gente só passa pelo que
tem que passar

Sol Quadrado – Pitty

RESUMO

Este trabalho analisa uma política pública na área do mercado de trabalho jovem no âmbito brasileiro, no período entre 2014 a 2018 chamada Programa Jovem Aprendiz. Assim, o objetivo desta pesquisa foi avaliar a efetividade de uma política pública por meio de programas e projetos de geração de emprego e renda na sociedade brasileira no referido período. Mais especificamente, verificar o impacto do programa Jovem Aprendiz sobre a renda dos jovens; estimar, de acordo com as características observáveis, a probabilidade do jovem a participar ou não do programa; comparar os jovens que participaram e não participaram do programa jovem aprendiz de acordo com características observáveis, além de contribuir com a literatura acerca da avaliação de políticas públicas. Para avaliar o impacto deste programa foram analisados dados da Relação Anual de Informações Sociais (RAIS) utilizando o método de Pareamento por Escore de Propensão. Os resultados encontrados mostram o efeito médio do tratamento sobre os tratados foi feita utilizando o *Propensity Score Matching (PSM)* a partir de microdados da RAIS dos anos em questão, separados entre jovens que participaram do programa jovem aprendiz e que não participaram. A estimação mostrou que os jovens que participaram do programa jovem aprendiz apresentam na média, ganhos de R\$ 3,14325 a menos por mês trabalhado do que um jovem do grupo de controle, que não participou do programa jovem aprendiz.

Palavras-Chave: *Propensity Score Matching*. Jovem Aprendiz. Mercado de Trabalho. Políticas Públicas.

ABSTRACT

This work analyzes a public policy in the area of the youth labor market in Brazil, in the period between 2014 and 2018 called the Young Apprentice Program. Thus, the objective of this research was to evaluate the effectiveness of a public policy through programs and projects to generate employment and income in Brazilian society in the period between 2014 and 2018. More specifically, check the impact of the Jovem Aprendiz program on the income of young people in the period between 2014 and 2018; estimate, according to observable characteristics, the probability of the young person participating or not in the program; compare young people who participated and did not participate in the young apprentice program according to observable characteristics, in addition to contributing to the literature on the evaluation of public policies. To assess the impact of this program, data from the Annual Social Information List (RAIS) was analyzed using the Propensity Score Matching method. The results found show the average effect of the treatment on those treated. It was done using Propensity Score Matching (PSM) based on microdata from RAIS from 2014 to 2018, where they were separated into young people who participated in the young apprentice program and those who did not participate. . . The research showed that young people who participated in the young apprentice program had, on average, earnings of R\$3.14325 less per month than a young person in the control group, who did not participate in the young apprentice program.

Keywords: Propensity Score Matching. Job market. Public policy.

LISTA DE FIGURAS

Figura 1 - Ciclo da Política.....	16
Figura 2 - Instrumento Legal Programa Jovem Aprendiz de 1942 a 2000.	28
Figura 3 - Instrumento Legal Programa Jovem Aprendiz de 2003 a 2018.	28
Figura 4 - Teoria da Mudança Programa Jovem Aprendiz	30

LISTA DE GRÁFICOS

Gráfico 1 - Tendência da taxa de desocupação entre jovens entre 2014 e 2018	22
Gráfico 2 - Crescente da taxa de desocupação entre jovens por idade	23

LISTA DE TABELAS

Tabela 1 - Quantidade de jovens trabalhadores por ano.....	43
Tabela 2 - Variáveis e sinais esperados para estimação do score de propensão.....	45
Tabela 3 - Média das variáveis de características observáveis por ano.....	47
Tabela 4 - Coeficientes da probabilidade de participar do Jovem Aprendiz estimados pelo modelo Logit.....	50
Tabela 5 - Quantidade de observação após verificação de suporte comum	51
Tabela 6 - Testes de diferenças de Média coletadas por grupos	53
Tabela 7 - Efeito do Programa Jovem Aprendiz sobre o Rendimento	54

LISTA DE ABREVIATURAS E SIGLAS

ATE	<i>Average Treatment Effect</i>
ATT	<i>Average Treatment Effect on The Treated</i>
CBO	Classificação Brasileira de Ocupações
CF	Constituição Federal
CIA	<i>Conditional Independence Assumption</i>
CLT	Consolidação das Leis Trabalhistas
CTPS	Carteira de Trabalho e Previdência Social
ECA	Estatuto da Criança e do Adolescente
IBGE	Instituto Brasileiro de Geografia e Estatística
JPAL	<i>Abdul Latif Jameel Poverty Action Lab</i>
MQO	Mínimos Quadrado Ordinários
MRL	Modelo de Regressão Linear
MTE	Ministério do Trabalho e do Emprego
ONG's	Organizações não Governamentais
OSCs	Organizações da Sociedade Civil
PDTE	Relação Anual de Informações Sociais
PME	Pesquisa Mensal de Emprego
PNAD	Contínua Pesquisa Nacional por Amostra de Domicílios Contínua
PNPE	Programa Nacional de Estímulo ao Primeiro Emprego
PSM	<i>Propensity Score Matching</i>
RAIS	Relação Anual de Informações Sociais
SENAC	Serviço Nacional de Aprendizagem Comercial
SENAI	Serviço Nacional de Aprendizagem Industrial
SESC	Serviço Social do Comércio

SUMÁRIO

	INTRODUÇÃO	12
1	CONCEITOS E FASES DAS POLÍTICAS PÚBLICAS	15
2	POLÍTICAS PÚBLICAS DE INSERÇÃO DO JOVEM NO MERCADO DE TRABALHO NO BRASIL	22
3	METODOLOGIA	32
3.1	MODELO EMPÍRICO	33
3.1.1	Pareamento por <i>Score</i> de Propensão	33
3.1.2	Modelo de Regressão <i>Logit</i> para estimar Escore de Propensão.....	37
3.1.3	Algoritmo de Pareamento	39
3.1.4	Teste de qualidade do Pareamento.....	41
3.1.5	Efeito Médio do Tratamento	42
3.2	BASE DE DADOS	42
4	ANÁLISE E DISCUSSÃO DE RESULTADOS.....	46
4.1	PROBABILIDADE DE PARTICIPAÇÃO DO PROGRAMA JOVEM APRENDIZ - ESCORE DE PROPENSÃO.....	48
4.2	EFEITOS DO PROGRAMA JOVEM APRENDIZ SOBRE O RENDIMENTO.....	54
5	CONSIDERAÇÕES FINAIS	56
	REFERÊNCIAS	59

INTRODUÇÃO

Além de crescimento econômico, uma sociedade que deseja ser desenvolvida economicamente precisa preocupar-se com a qualidade de vida e bem-estar social da população. De acordo com Amartya Sen (2000), o desenvolvimento econômico perpassa pelo crescimento econômico como elemento capaz de propiciar melhorias nas condições de vida do indivíduo e propiciar liberdade. No entanto, o mundo globalizado gera consequências capazes de privar liberdades e reduzir a qualidade de vida através de desigualdade social e econômica, altas taxas de desemprego, pobreza causada pela queda de rendimentos e, consequentemente, fome.

Após a 2ª Guerra Mundial o processo de globalização permitiu a implantação de tecnologias em diversos países, propiciou a modernização da cadeia produtiva e com isso modificou-se também a composição da mão de obra disponível para o mercado de trabalho, consequentemente a composição da taxa de desemprego também se modificou. Apesar de ser um fenômeno grave em países subdesenvolvidos, países desenvolvidos também são atingidos com altas taxas de desemprego, como a Espanha, Grécia, e Itália que apresentaram, em 2020, taxas de desemprego entre jovens maiores que ao desemprego geral (TAVARES, CÂNDIDO E CARMO; 2021).

No Brasil, de acordo com o Instituto Brasileiro de Geografia e Estatística (IBGE, 2021), em 2014 atingiu-se uma taxa de desemprego de 6,8%, destes, 21,1% correspondem ao desemprego entre jovens de 14 a 17 anos e, entre jovens de 18 e 24 anos sendo 13,8%. Nos anos seguintes a taxa de desemprego foi crescente atingindo em 2018 a taxa de 12,3 %, sendo 40,4% para jovens entre 14 e 17 anos e 24,6% para jovens entre 18 e 24 anos. Verifica-se que estes números estão acima da média nacional da taxa de desocupação do mesmo ano. Portanto, nota-se a necessidade de medidas e políticas públicas que propiciem direitos, capacitação e treinamento de jovens para o mercado de trabalho a fim de alocar o grupo etário chamado de jovens.

Nesse sentido, políticas públicas são medidas adotadas com a finalidade de combater algum problema público (SECCHI, 2013). Disto depreende-se que as políticas públicas passam a ser pautadas no projeto de governo e então se concretizam por meio de programas sociais possibilitados pela ação do Estado (HÖFLING, 2001), que devem sempre ser monitoradas e avaliadas após a implantação com o intuito de verificar a efetividade.

Uma avaliação de políticas públicas consiste em medir a eficiência da aplicação dos recursos públicos e o efeito de tal ação em uma sociedade onde foi estabelecida. Além disso, esta avaliação possibilita monitorar a atuação estatal visando melhorias e resultados da gestão. Nesse sentido, quando deseja-se avaliar busca-se também identificar o impacto dessa política na sociedade.

Desse modo, a avaliação é parte essencial e fundamental para a implantação de uma política pública. De acordo com Menicucci (2008) a avaliação de política pública é um instrumento gerenciador, que permite o processo de tomada de decisão do gestor.

Uma política pública no Brasil implementada em 2000 para inserir jovens no mercado de trabalho é o “Programa Jovem Aprendiz”, criado com o intuito de fornecer preparação e capacitação para os jovens. Além disso, possui também como finalidade a regulamentação da atividade laboral e promover a alocação de jovens no mercado de trabalho no Brasil. Em 2000, foi criada pelo Governo Federal a Lei nº 10.097, a Lei da Aprendizagem ou do jovem aprendiz, com o objetivo de promover a estudantes adolescentes e jovens, entre 14 e 24 anos, a oportunidade de conseguir uma vaga no mercado de trabalho, assim como obter experiência (BRASIL, 2000).

No contexto das políticas públicas e mercado de trabalho jovem esta pesquisa procurou responder à questão de qual impacto do programa jovem aprendiz na renda dos jovens que participaram do programa. E para tanto, buscou-se respostas através do tratamento e da análise de dados disponibilizados na forma bruta e com utilização de métodos econométricos que melhor se adequaram aos dados e apresentaram melhores respostas ao questionamento.

Nesse sentido, o objetivo desta pesquisa é avaliar a efetividade de uma política pública por meio de programas e projetos de geração de emprego e renda para os jovens na sociedade brasileira no período entre 2014 a 2018. Mais especificamente pretende-se verificar o impacto do programa Jovem Aprendiz sobre a renda dos jovens no período entre 2014 a 2018; estimar, de acordo com as características observáveis, a probabilidade do jovem a participar ou não do programa; comparar os jovens que participaram e não participaram do programa jovem aprendiz de acordo com características observáveis, além de contribuir com a literatura acerca da avaliação de políticas públicas.

A partir destas considerações o presente trabalho foi estruturado em cinco seções, além da introdução. Inicialmente é feito uma conceitualização sobre políticas públicas. Na segunda seção apresenta-se o referencial teórico com uma breve discussão sobre elaboração, implementação, monitoramento e avaliação de políticas públicas e também um histórico sobre a legislação do Programa Jovem Aprendiz no Brasil. Na seção três encontra-se a metodologia de avaliação de política pública a partir do método de pareamento por escore de propensão e a base de dados utilizada. Os resultados e as discussões encontram-se na quarta seção e, por fim, na quinta e última apresenta-se as considerações finais.

1 CONCEITOS E FASES DAS POLÍTICAS PÚBLICAS

Esta seção aborda a conceituação de políticas públicas, aspectos relevantes na elaboração de uma política eficiente e seus arranjos. Destaca-se que o processo transcende a mera identificação de problemas. Envolve a definição de um ciclo e um design para visualização coesa. Conclui-se que a avaliação de políticas públicas busca mensurar eficácia e eficiência, fornecendo informações sobre a preferibilidade entre diferentes políticas com base em critérios específicos.

As políticas públicas se tornaram objeto de estudo quando, por meio de movimentos sociais, nos anos 50, começaram a organizar as inovações sociais nos Estados Unidos e na Europa, e esses movimentos sociais passaram a cobrar do Estado respostas para problemas sociais da época. A partir de 1970, passou a se cobrar do Estado quais suas funções e a eficácia de ações de intervenção social (TREVISAN; BELLEN, 2008).

Para chegar ao que se conhece hoje por políticas públicas foi necessária a união de definições de quatro grandes cientistas da área da ciência política: H. Laswell, H. Simon, C. Lindblom e D. Easton (SOUZA, 2006). Considerando o conceito destes estudiosos pode-se chegar ao entendimento de que políticas públicas são ações elaboradas e coordenadas por governos afins de interceder em prol de solucionar problemas da sociedade em geral.

Para elaborar Políticas públicas não se define apenas o problema que se deseja resolver. No processo de criação de uma política há a definição de um ciclo de política e um desenho da política para melhor visualização do processo.

Considerando que o objetivo de uma política pública é criar soluções para algum tipo de problema, chega-se ao questionamento de como pode-se medir o sucesso da política. O sucesso do programa não se pode resumir à implantação ou entrega do programa, é necessário cumprir fases e seguir métodos. Dentro do processo de formulação das políticas existem modelos que divergem de acordo com os teóricos, mas que convergem para um ponto em comum: é preciso criar soluções para os problemas.

O ciclo de políticas públicas é considerado um processo que tem a necessidade de trabalhar em fases com diferentes agentes, poderes e ambientes (LASSWEL, 1958). Considera-se que o ciclo das Políticas públicas possui 5 fases de atividades necessárias para cumprir o ciclo.

Analiticamente, as cinco fases são: percepção e definição de problemas, definição de agenda, elaboração de programas e decisão que neste trabalho será chamado de formulação, implantação, por fim, a fase de controle de impacto da política que será chamada neste

trabalho de avaliação (PEDONE,1986). Em cada fase é possível uma correspondência em sequência lógica do processo político administrativo passível de investigações no que diz respeito à esfera de poder, política e social (FREY, 2000). É possível visualizar as fases do ciclo da política na figura 1, enfatizando a fluidez e dinamicidade, em que é possível visitar fases anteriores sem prejuízo às posteriores e vice-versa:

Figura 1– Ciclo da Política



Fonte: A autora a partir de Frey, 2000.

Na figura 1 é possível observar fases que permitem visitá-las independente da etapa do processo. Para este trabalho adotou-se uma análise sobre o ciclo da política composto pelas 5 fases dinâmicas conforme proposto por Frey (2000).

Assim, na primeira fase, perceber e definir um problema social geralmente é o momento em que se identifica uma necessidade de intervenção para melhoria de determinada situação. Desta forma, essa melhoria pode ser através de uma intervenção política, social ou administrativa. Definir o problema pode estar relacionado com a percepção social. Entretanto, é somente na fase de definir agenda que será maturado o problema e alternativas para a solução (PEDONE,1986; FREY, 2000).

Na fase 2, definição de agenda, (agenda *setting*) listam-se os problemas, possíveis alternativas de enfrentamento, agentes governamentais e não governamentais que possam participar. É nessa fase que se levanta questionamentos sociais e formas de enfrentamento do problema, escolhe-se os responsáveis pela definição de temas de construção da agenda, esses agentes podem ser do tipo formais e informais. (ROSA; LIMA; AGUIAR, 2021). Os atores formais são os poderes Legislativo, Executivo e Judiciário e os partidos políticos; já os atores informais são aqueles que atuam na política pública de fato, são os técnicos profissionais do

legislativo, pesquisadores, agências internacionais ou Organizações não Governamentais (ONG's), inclui-se também neste grupo empresários, sindicatos, movimentos sociais e a mídia. Os atores informais atuam em defesa de interesses de acordo com a reivindicação social de cada grupo de interesse.

Quanto à Formulação (fase 3), é nesta etapa em que se constrói o enredo conceitual de compreensão dos problemas sociais, com planejamento, elaboração de programas, diagnósticos, tomada de decisão de qual estratégia a ser adotada. É quando os agentes diretos e indiretos são alocados de acordo com cada agenda. É nesta etapa em que pode ser feita avaliação *ex-ante* que significa uma avaliação antes da implementação, listando resultados esperados, custos, consequências etc. (PEDONE, 1986; FREY, 2000).

Ainda na mesma fase, Secchi (2013) separa em dois momentos: um que envolve a formulação de alternativas, necessitando da especificação dos conflitos em questão, e outro em que ocorre a tomada de decisão pelos agentes envolvidos. Essa decisão é orientada por três modelos: racionalidade, incremental e fluxos múltiplos.

A fase de implementação consiste em aplicar o que foi definido na formulação, verifica-se a disponibilidade de recursos necessários para concretizar a política. Nesse estágio, é possível analisar materiais, programas ou projetos. Para Simões (2018) é nesta etapa que as intenções são convertidas em ações que devem gerar resultados pretendidos pela política. É neste momento que possíveis novos problemas surgem e haverá necessidade de retornar a etapas para fazer alterações (Majone e Wildavsky, 1984 *apud* Pal, 2014).

A implementação (fase 4) pode ser analisada sob duas abordagens: a *top-down* e *bottom-up*. Na *top-down* há uma ênfase às fases de tomar decisão e implementação, pois o primeiro seria formulado por agentes públicos enquanto as burocracias e setores privados são negligenciados. Para a abordagem *bottom-up* o foco passa a ser sobre os atores envolvidos no processo de implementação, e assim verifica-se a necessidade e razão para que haja congruência entre formuladores e implementadores de políticas públicas (SABATIER, 2007).

A quinta fase do ciclo de Políticas públicas é a avaliação. Nela é possível verificar os impactos efetivos do programa e efeitos colaterais indesejados (FREY, 2000). Apesar de ser possível verificar os resultados, por ser a etapa em que constata-se se os objetivos propostos nas etapas anteriores foram alcançados, é a etapa mais negligenciada pelos governistas.

O processo de avaliar uma política pública ou programa social está relacionado com os resultados, entretanto o quanto que se pode considerar eficaz ou eficiente, fracasso ou sucesso

de determinada política pública? Avaliar significa medir o quanto determinada ação foi efetiva ou não, entretanto, efetividade é um termo subjetivo para cada indivíduo.

Considerando que avaliação é medir a contribuição das políticas à sociedade afim de solucionar uma demanda, é necessário quantificar estes resultados (BARROS E LIMA, 2017). Nesse sentido, avaliar no âmbito das políticas públicas significa seguir métodos e metodologias a fim de mensurar, atribuir valor à intervenção feita ou a ser feita.

O principal objetivo de uma avaliação é informar, através de critérios, o que uma política pública se torna preferível a outra. É, portanto, por meio da avaliação que se pode chegar a comparação de efetividade, eficácia e eficiência.

Assim é importante responder a seguinte questão: Por que avaliar?

No âmbito da avaliação de política públicas existem tipologias a serem observadas. Seguindo a lógica do ciclo da política, de acordo com objetivo do programa ou política é possível especificar qual avaliação mais indicada. Rossi (1991) apresenta 5 tipos de avaliações: *Needs Assessment*, *Design Assessment*, *Program Process Assessment*, *Impact Assessment* e *Efficiency Assessment*.

A avaliação *Needs Assessment* refere-se à necessidade da política, se realmente há necessidade do programa social, se os beneficiários serão atendidos de acordo com a necessidade. Quanto ao *Design Assessment* refere-se ao desenho lógico do programa e a teoria utilizada, se o que está sendo proposto tem base teórica e conceitual e atende à demanda. Avaliar do tipo *Program Process Assessment*, refere-se ao processo de implementação do programa ou política, como vai funcionar, quem serão os prestadores de serviço, a quem destina-se o programa. Quando decide-se avaliar observando os impactos e seus resultados enquadra-se na *Impact Assessment*. Por fim, quando já houve todas as avaliações a *Efficiency Assessment* complementa a avaliação de eficiência ou custo-benefício, e torna possível verificar custos da operação e se é possível replicar em outras escalas (ROSSI, 1991).

Considerando que avaliar uma política pública é ter respostas para todas as perguntas dentro dos domínios de avaliação citados anteriormente, o *Abdul Latif Jameel Poverty Action Lab* (JPAL)¹ estruturou um roteiro capaz de nortear todos questionamentos necessários em cada domínio da avaliação colocado por Rossi (1991), que é a utilização da metodologia de

¹ O *Abdul Latif Jameel Poverty Action Lab* (J-Pal) é um centro de pesquisa que tem como objeto de pesquisa a pobreza global e formas de enfrentamento através de avaliação de políticas públicas. Tornou-se referência na avaliação de programas sociais em 2019 quando seus fundadores receberam o Prêmio Nobel de Economia pelo método experimental utilizado com intuito de redução da pobreza global, pela inovação da metodologia utilizada, a metodologia de ensaios controlados randomizados. Os trabalhos premiados passaram a mensagem de que para resolver os problemas sociais é necessário aplicar teorias sobre a pobreza, mas também é importante por meio de observações e experimentos conhecer e verificar se funcionam e, a partir daí, desenhar programas sociais ou Políticas públicas eficazes para resolução de um problema.

ensaios controlados randomizados que consideram o efeito causal com o desenho da Teoria da Mudança.

Através da teoria da mudança é possível construir uma avaliação teórica criando um roteiro indicando os problemas e necessidades da sociedade; avaliar os caminhos que estão sendo utilizados para chegar aos resultados e trazer também o resultado final esperado com determinada intervenção de uma política ou projeto, ou seja, é possível verificar todo o percurso que levará à resolução do problema.

A teoria da mudança, portanto, é um instrumento que facilita a construção de uma cadeia causal teórica e é uma ferramenta bastante utilizada pelas Organizações da Sociedade Civil (OSCs) para planejar e medir o impacto social, funcionando como um roteiro.

Pensando em qual melhor estratégia escolhida quando se tem várias possibilidades a fim de chegar ao mesmo objetivo da mudança, os ganhadores do Prêmio Nobel de Economia em 2019 – J-Pal chamaram de Teoria da Mudança o roteiro que diz para como seguir e para onde ir e como, através de processos, chega-se aos resultados.

A teoria da mudança proposta pelo JPAL tem um formato de cadeia causal, aonde traz o desenho do programa e a implantação com problemas, necessidades, qual o insumo necessário para realização/implantação do programa e o produto que se espera. A partir daí é possível desenhar hipóteses e riscos e chega-se aos impactos do programa, seja ele resultados intermediários e resultado final.

Quando decide-se realizar de uma política pública busca-se prever a contribuição de uma intervenção, seja de determinado programa ou política pública, qual a consequência dessa intervenção para a sociedade. Portanto, ao realizar uma avaliação de impacto busca-se identificar os efeitos causais dessa ação para a sociedade baseada em evidências. Diante disso, através de uma avaliação de impacto pode-se identificar os efeitos de determinada ação gerou sobre os beneficiários, baseado em evidências empíricas que são coletadas durante ou após a ação e que buscam responder questões de causa e efeito da intervenção (GERTLER *et al*, 2018). Uma avaliação de impacto pode ocorrer em diferentes fases do programa ou política pública, de acordo com a necessidade do resultado.

Embora uma avaliação de impacto adote um método econométrico que aponta a correlação entre as variáveis, muitas das vezes a avaliação está equivocada, pois não apresenta a verdadeira relação causal. É importante destacar o real sentido do efeito causal, pois a causalidade pode ser confundida nas análises de regressões pois, através das análises

descritivas ou correlações condicionais é comum serem interpretadas como causalidade (RIBAS, 2008).

Nesse sentido, Dowd e Town (2002) destacaram em seus estudos cinco conceitos para causalidade: causa, determinação, antecedente, razão e ocasião. Com o intuito de simplificar esta pesquisa, optou-se por priorizar os conceitos de causa e determinação. "Causa" refere-se a um resultado decorrente da ocorrência de um evento; já "determinação" implica em uma variável que influencia o resultado de outra variável.

Para Gertler *et al* (2018), a avaliação de impacto pode ser do tipo prospectiva ou retrospectiva. A avaliação de impacto da categoria prospectiva refere-se àquelas que ocorrem concomitante com implementação do programa, já as avaliações de impacto do tipo retrospectivas avaliam o impacto após a implementação da política ou programa social.

Especificamente, a avaliação pode acontecer em três momentos, de acordo com Peixoto *et al* (2017), sendo elas: antes do início do programa e chamadas de avaliações *ex ante*; já a avaliação que acontece concomitante à execução do programa chama-se de avaliação *ex post* de percurso; por fim, a chamada *ex post* de encerramento que ocorre quando ocorre a conclusão da política públicas ou programa social.

Para este trabalho utilizou-se a avaliação de impacto *ex post* de percurso, que apesar de ter o tempo como variável limitadora, a análise do programa pode influenciar e a continuidade e posteriormente a expansão do programa ou até mesmo melhoramento considerando que o programa social ainda não foi finalizado.

Em suma, o objetivo da avaliação de impacto é verificar o efeito causal do programa em quem recebeu o tratamento e qual poderia ser o efeito em quem não recebeu o tratamento, o que será chamado de contrafactual, e qual foi o progresso que o programa ou política trouxeram para a sociedade. O contrafactual é um grupo de indivíduos que não receberam o tratamento e deseja-se saber qual seria o efeito caso estes indivíduos recebessem tratamento, ou seja, os não tratados. Esse grupo também pode ser chamado de grupo controle, sendo aqueles indivíduos que possuem as mesmas características daqueles que receberam tratamento, mas que, por alguma característica não observada, não participaram do programa.

Entretanto, definir o contrafactual é impossível tendo em vista a limitação de verificar o efeito de um tratamento e do não tratamento no mesmo indivíduo, no mesmo período do programa, então como não é possível definir um contrafactual é necessário construir um estatisticamente estimando um contrafactual. Para estimação do contrafactual é necessário

observar as características observáveis e não observáveis e assim separar os grupos de acordo com as características semelhantes.

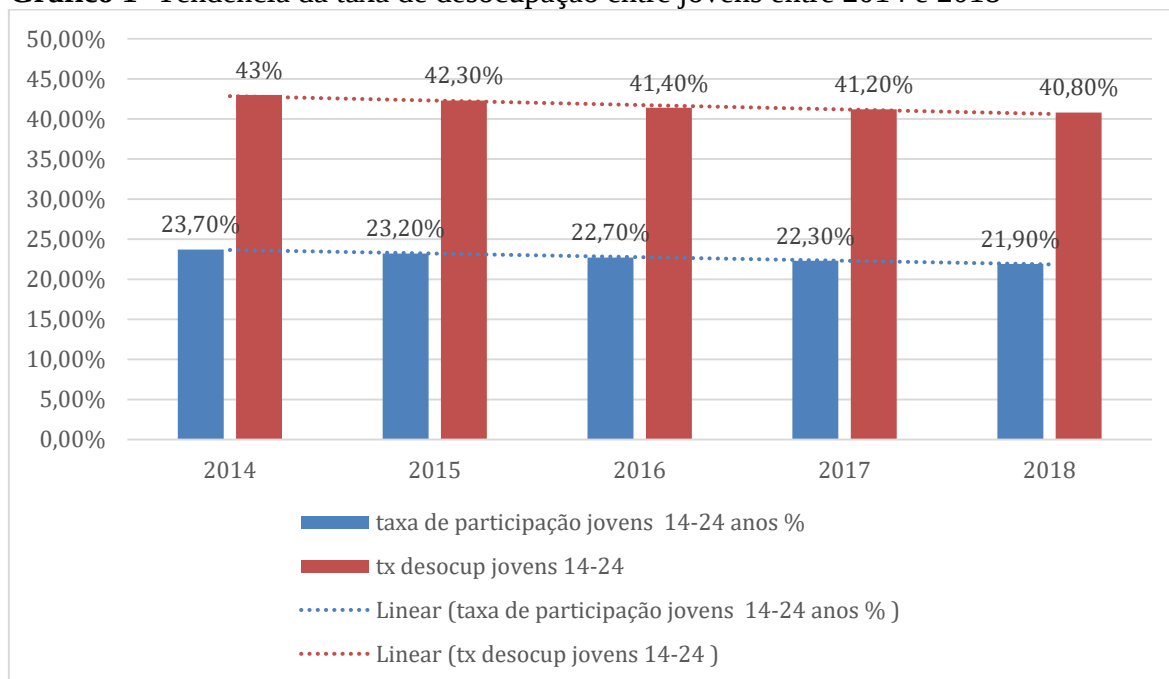
Considerando essa avaliação de impacto, o objeto de estudo são os resultados do programa ou política implementada em determinado meio. Nesse sentido, para que seja possível realizar a avaliação é necessário que se tenha um projeto com ações ou atividades provedoras de resultados que, geralmente, causam mudanças. E, é através da avaliação que é possível medir essas mudanças. A metodologia utilizada nas políticas públicas consiste em seguir uma cadeia causal com um marco lógico ou teoria da mudança, que explicam, partindo de determinado problema até a mudança gerada através de programas ou políticas públicas em questão.

Portanto, na busca de entender a cadeia causal faz-se necessário entender o desenho do Programa Jovem Aprendiz como política pública, desde sua implementação, caminhos que levam aos objetivos, requisitos necessários para participação assim como quais as ferramentas e métricas utilizadas para avaliação do governo e se esta se mostra efetiva de acordo com os resultados esperados. Na seção seguinte apresenta-se o Programa Jovem Aprendiz e o instrumento legal utilizado para implantação do mesmo.

2 POLÍTICAS PÚBLICAS DE INSERÇÃO DO JOVEM NO MERCADO DE TRABALHO NO BRASIL

O Brasil apresentou, de acordo com IBGE, em 2014 uma taxa de desocupação de 4,3%, o menor nível em toda a série histórica da Pesquisa Mensal de Emprego (PME). Entretanto, esta taxa de desocupação é composta majoritariamente de jovens. Especificamente, a taxa de participação de jovens de 14 a 17 anos e jovens de 18 a 24 anos chegou a 23% do total da população geral, e 43 % da soma da população desocupada. Ou seja, apesar de jovens da faixa etária de 14 a 24 anos serem apenas 23% da população em geral, eles representam 43% da população desocupada. No Gráfico 1 é possível verificar a taxa de participação de jovens de 14 a 24 anos na população e a taxa de desocupação da mesma faixa etária.

Gráfico 1- Tendência da taxa de desocupação entre jovens entre 2014 e 2018



Fonte: A autora a partir de IBGE, 2021.

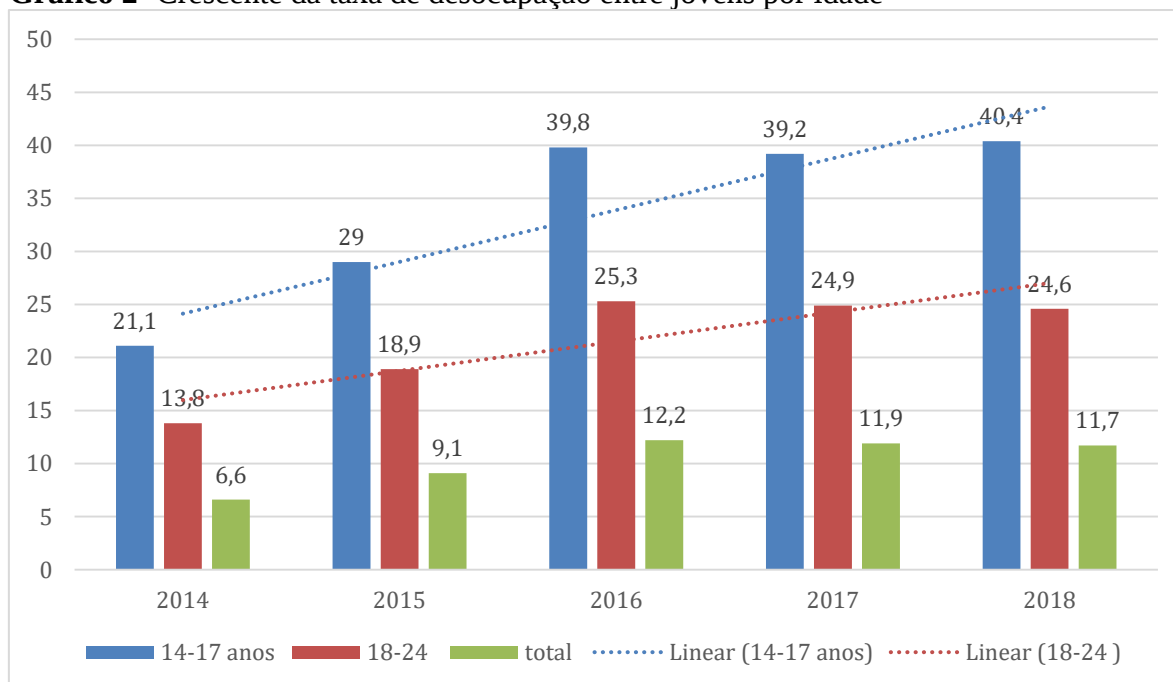
Ao observar o gráfico 1 é possível identificar uma linha em tendência de queda na taxa de desocupação entre jovens de 14 a 24 no período de 2014 a 2018, entretanto, não se pode inferir que essa redução se deu devido à inserção do jovem no mercado de trabalho, pois a taxa de participação desta faixa etária, durante o mesmo período, também diminuiu.

Após 5 anos, em 2018, a taxa de desocupação no Brasil foi de 12,3% de acordo com a Pesquisa Nacional por Amostra de Domicílios Contínua (PNAD Contínua), o melhor ano desde que a pesquisa (PNAD) começou a ser feita. Aproximadamente 4,8 milhões de jovens desocupados no Brasil, na Força de trabalho, representam uma distribuição percentual da taxa

de desocupação entre jovens de 14 a 24 anos de 40%, e uma taxa de participação na população de 21,9% de jovens de 14 a 24 anos. O que se pode verificar a partir do gráfico 1 é o tamanho da parcela que o jovem representa, quase $\frac{1}{4}$ da população em geral, da mesma forma que o desemprego para esta faixa etária também se apresenta com mais intensidade.

No Gráfico 2 pode-se verificar a taxa de desocupação entre os jovens de 14 e 24 anos e percebe-se que é uma parcela considerável da população que se encontra em desemprego.

Gráfico 2- Crescente da taxa de desocupação entre jovens por idade



Fonte: A autora a partir de IBGE, 2021.

No Gráfico 2 destaca-se a tendência crescente da taxa de desocupação entre jovens. A taxa de desocupação entre jovens de 14 a 18 anos se justifica ser maior pois, são jovens que a única forma de vínculo empregatício é por programas de aprendizagem. Entretanto, quando se analisa a faixa etária de jovens de 18 a 24 anos percebe-se um comportamento crescente e superior à taxa de desocupação total.

O desequilíbrio entre oferta de mão de obra e demanda por trabalho gera uma falha de mercado que é o desemprego, e ao considerar o tamanho da população de jovens, o desemprego entre eles tem uma intensidade maior. Com intuito de minimizar os danos causados pelos desequilíbrios, o Governo utiliza as políticas públicas como um mecanismo de resposta para suprir as demandas existentes na sociedade. Portanto, de acordo com Bresser Pereira (2009), o Estado Moderno regula os mercados desde o Estado Absoluto e possui a função de regulador da atividade econômica.

Entende-se o Estado como um conjunto de instituições com regras e com objetivo de proteger e atender necessidades da sociedade, enquanto que o Governo é o conjunto de atores que direcionam as políticas públicas e regulam a sociedade politicamente (SILVA *et al.*, 2015). E, apesar de serem atores do Estado, possuem interesses político partidários e ideológicos.

No Brasil, as políticas públicas passaram a ser destaque para o Governo a partir de 1940, quando iniciou processo de industrialização e se potencializou em 1980 com a redemocratização e descentralização do Estado. De acordo com Araújo (2000) as políticas públicas tinham mais o papel de beneficiar o crescimento econômico do que favorecer melhores condições sociais.

Quando se fala de inserção do jovem no mercado de trabalho brasileiro houveram diversos desafios e mudanças ao longo das décadas. Quando se iniciou o processo de industrialização em 1940 e 1950 as indústrias atraíam os jovens da área rural com expectativa de trabalho e renda. À medida que as indústrias cresciam demandavam trabalhadores qualificados e para suprir esta demanda começou-se a ofertar cursos de ensino técnico e profissionalizantes visando a inserção destes jovens nas fábricas e indústrias em crescimento. O crescimento da participação do jovem no mercado de trabalho fomentou também a participação deles em Sindicatos e Movimentos Trabalhistas lutando por melhores condições de trabalho e melhores salários (SANTOS; GIMENEZ, 2015).

A partir da década de 1940 houve uma forte migração de jovens da área rural para área urbana devido a industrialização, criou-se uma massa de jovens desempregados por falta de qualificação necessária para trabalhar na indústria, em resumo havia uma demanda por mão de obra qualificada enquanto existia a oferta de mão de obra sem qualificação. Iniciou-se então, o protagonismo do Sistema “S” na educação profissionalizante no Brasil. O sistema “S” é um sistema formado por instituições prestadoras de Serviços que se tornou um propulsor do ensino profissionalizante ofertando cursos profissionalizantes que atendessem as necessidades destas instituições. De início, o sistema “S” foi formado a partir da fundação do Serviço Nacional de Aprendizagem Industrial (SENAI) em 1942 e do Serviço Social da Indústria (SESI) em 1946, Serviço Nacional de Aprendizagem Comercial (SENAC) em 1946 e o Serviço Social do Comércio (SESC) em 1946 (KUENZER, 2001).

A legislação do Brasil, acompanhando as mudanças do mercado de trabalho com a Consolidação das Leis Trabalhistas (CLT), estabelecimento do salário mínimo e organizações sindicais, implementou leis trabalhistas que estabeleceram limitações específicas para o

trabalho de jovens, incluindo idade mínima para o emprego e regulamentações para proteger os direitos dos trabalhadores mais jovens, como limites de jornada de trabalho.

A lei do Aprendiz passou também por diversas atualizações e modificações ao longo do tempo. O marco legal foi em 1942 pelo Decreto-Lei nº 4.073, de 30 de janeiro de 1942, que foi chamado de Lei dos Menores e estabelecia a idade dos jovens menores e proteção aos direitos dos trabalhadores, em seguida, em 15 de dezembro 1942, foi promulgada a “Lei do Aprendiz” com o Decreto-Lei nº 5091, que dispôs sobre o conceito de Aprendiz e define a idade que o jovem é considerado jovem aprendiz (BRASIL, 1942). Em 1943, houve a grande conquista da aprovação da Consolidação das Leis do Trabalho, com o Decreto-Lei nº 5.452, de 1º de maio de 1943 que trouxe no capítulo IV a Proteção do trabalhador menor e, na Seção IV os Deveres dos Responsáveis Legais de Menores e dos Empregadores da Aprendizagem ((BRASIL, 1943); Posteriormente em 12 de agosto de 1946, o Decreto-Lei nº 9.576 modificou o Decreto-Lei nº 4.481, que dispôs sobre a aprendizagem dos industriários e estabelecendo deveres dos empregadores e dos aprendizes (BRASIL, 1946). Em 1946 também foi criado o Serviço Nacional de Aprendizagem Comercial (SENAC), pelo Decreto-Lei nº 8.621, de 10 de janeiro de 1946, bem como a aprendizagem dos comerciários foi regulamentada pelo Decreto-Lei nº 8.621, do mesmo dia 10 de janeiro de 1946. No mesmo ano, a Constituição de 1946 definiu que “as empresas industriais e comerciais são obrigadas a ministrar, em cooperação, aprendizagem aos seus trabalhadores menores, pela forma que a lei estabelecer, respeitados os direitos dos professores” (BRASIL, 1946).

Diante do pós-guerra, com a reestruturação produtiva a inserção do jovem no mercado de trabalho passou a ser objeto das políticas públicas do mercado de trabalho brasileiro, pois a década de 1980 no Brasil foi caracterizada por desafios econômicos, políticos e sociais que afetaram a inserção dos jovens no mercado de trabalho. A alta taxa de desemprego devido ao processo de desindustrialização e transição para o setor de serviços e a informalidade eram questões prementes, e os programas de treinamento e educação profissional eram uma tentativa de abordar esses desafios. A década de 1980 também foi marcada por movimentos sociais e trabalhistas que pressionaram por melhorias nas condições de trabalho e direitos dos trabalhadores, incluindo jovens trabalhadores. Isso influenciou a legislação trabalhista e a discussão sobre direitos trabalhistas para os jovens.

O papel do jovem no processo de industrialização no Brasil foi significativo considerando que eles tinham uma elevada participação no mercado de trabalho. A constituição de 1988 trouxe mudanças nas práticas ligadas à aprendizagem de adolescentes e

influenciou aspectos legais de aprendizagem na década de 1990. A Constituição Federal de 1988 trouxe no artigo 7º, parágrafo XXXIII, e artigo 227, inciso 3º a determinação de idade mínima para 14 anos, garantiu direitos previdenciários e trabalhistas e obrigou o acesso do adolescente à escola.

O texto da CF confere ao adolescente a proteção ao trabalho, tornando o jovem um cidadão portador de direitos. Em ato contínuo, através da Lei 8.069/1990 de 13 de julho de 1990, estabeleceu-se um conjunto de direitos e deveres da criança e do adolescente, o Estatuto da Criança e do Adolescente (ECA), trouxe um capítulo versando sobre o Direito à Profissionalização e à Proteção no Trabalho.

Em 2000, foi sancionada a Lei do Aprendiz, a qual tem seu marco legal através da lei de aprendizagem nº 10.097, de 19 de dezembro de 2000, que alterou dispositivos da CLT, revogou a Lei Menores e estabeleceu as bases do programa Jovem Aprendiz como é conhecido atualmente. Esta legislação trouxe em seu conteúdo respaldo legal para contratação de jovens a partir de 14 anos de idade, complementou também a Lei nº 8.069, de 13 de julho de 1990, que versa sobre o ECA e prevê o direito à aprendizagem à criança e ao adolescente (BRASIL, 2000).

Em 2003, criou-se o Programa Nacional de Estímulo ao Primeiro Emprego para os Jovens (PNPE) pela Lei nº 10.748, de 22 de outubro de 2003. O programa visava promover a criação de postos de trabalho para jovens ou prepará-los para o mercado de trabalho e ocupações alternativas, geradoras de renda; e a qualificação do jovem para o mercado de trabalho e inclusão social ((BRASIL, 2003).

A Lei da Aprendizagem foi complementada com o a lei nº 11.180, de 23 de setembro de 2005 que no Art. 18 modificou os Arts. 428 e 433 da CLT, estabelecendo a idade de jovem aprendiz ser entre 14 e 24 anos e contrato de trabalho especial com prazo determinado (BRASIL, 2005).

O Ministério do Trabalho e Emprego (MTE), extinto em 2021, já havia publicado o Decreto nº 9.579, de 22 de novembro de 2018 que consiste no Manual da Aprendizagem, em que se consolidou os parâmetros necessários para cumprimento da legislação e regulação da contratação de aprendizes (BRASIL, 2009).

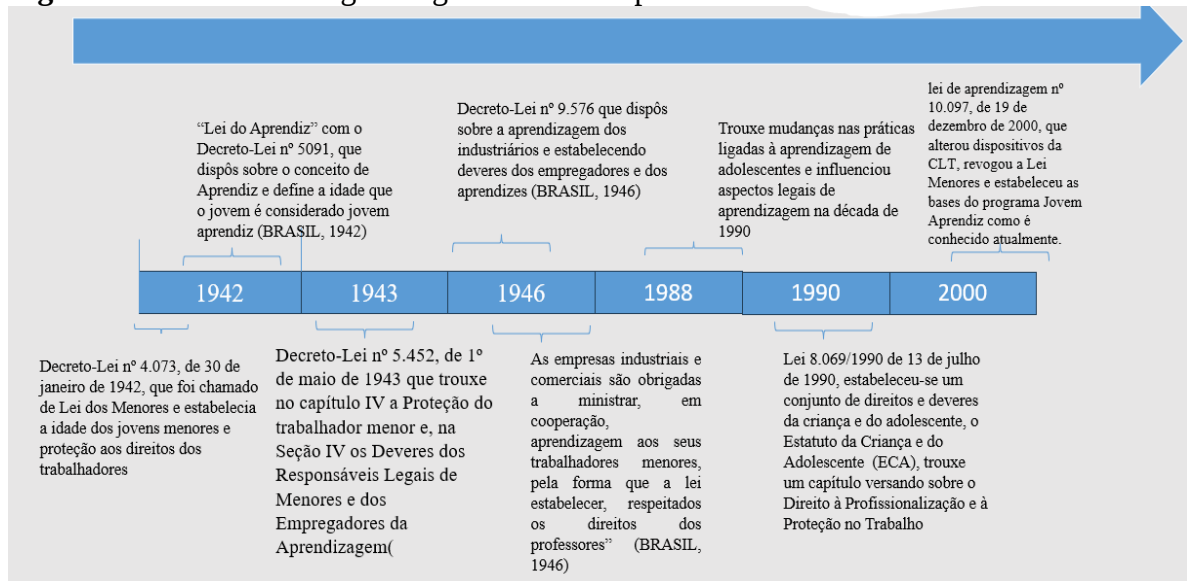
Diante desse aparato legal o Programa Jovem Aprendiz nasce com o intuito de regulamentar a condição de aprendizes através de um contrato de trabalho previamente estabelecido por lei, que foi possível estabelecer idade, condições de trabalho, carga horária de trabalho, remuneração, prazo de validade contratual e exige frequência escolar.

O Programa Jovem Aprendiz é destinado a adolescentes e jovens entre 14 e 24 anos, devidamente matriculados em algum curso de formação metódica e frequentando a escola regularmente ou ter concluído o ensino médio, de acordo com o disposto no Art. 428 da CLT com Redação dada pela Lei nº 11.180, de 2005, que dispõe acerca do Contrato de Aprendizagem como contrato de trabalho especial.

O principal objetivo do Programa Jovem Aprendiz é a capacitação de jovens estudantes inscritos no programa visando a inserção do jovem no mercado de trabalho, proporcionando experiência e conscientização sobre o mercado de trabalho. Entretanto, há uma contrapartida de um determinado salário para o jovem que está inserido no programa.

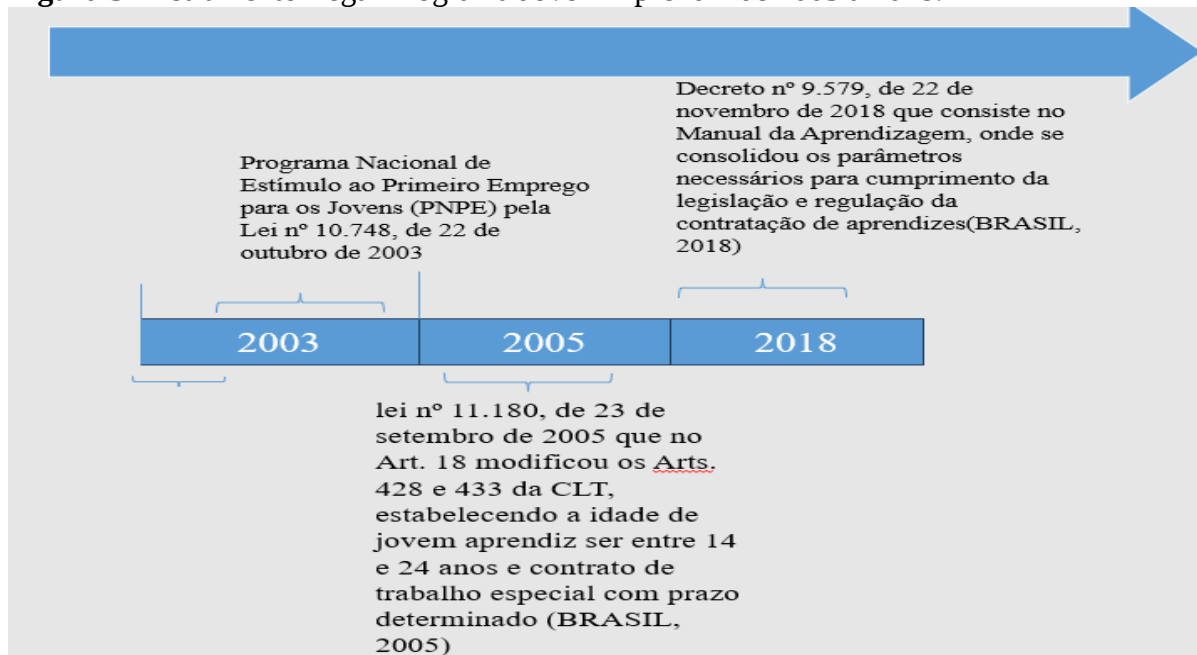
A seleção do jovem aprendiz é feita pelo empregador, entretanto jovens entre 14 e 18 anos são prioridades, considerando os princípios previstos nos artigos 227 da Constituição Federal de 1988 (CF), no Artigo 4º do ECA, e no art. 53 do Decreto nº 9.579/2018. São prioridades desde que as atividades laborais não ocorram em ambiente insalubres ou com periculosidade; não comprometam o desenvolvimento físico, psicológico e moral dos adolescentes ou que não exijam licença ou autorização para jovens com idade inferior a 18 anos.

Resumidamente, na figura 2 tem-se os principais destaques da fundamentação legal do Programa Jovem Aprendiz.

Figura 2- Instrumento Legal Programa Jovem Aprendiz de 1942 a 2000.

Fonte: A autora.

Até o ano de 2000 pode-se perceber que houve um aprimoramento e maior preocupação em oferecer direitos e deveres aos jovens que estariam na situação de jovem aprendiz. Na figura 3, é possível verificar a preocupação com a legislação não foi diferente do que nos anos anteriores, com destaque para criação do PNPE e do Manual da Aprendizagem.

Figura 3- Instrumento Legal Programa Jovem Aprendiz de 2003 a 2018.

Fonte: A autora.

O propósito do Programa Jovem Aprendiz é oferecer uma qualificação técnica profissional com prática e teoria aliadas a fim de possibilitar a construção de uma profissão vinculada a alguma ocupação codificada na Classificação Brasileira de Ocupações (CBO).

O funcionamento do programa é baseado na implantação do programa no estabelecimento, na seleção do aprendiz e por fim na contratação. O Art. 429 da Lei nº 10.097, de 19 de dezembro de 2000, nele se estabelece que os estabelecimentos devem empregar jovens aprendizes e matriculá-los, pelos menos 5% e no máximo 15%, em cursos dos Serviços Nacionais de Aprendizagem, em *ipsis litteris*:

“Os estabelecimentos de qualquer natureza são obrigados a empregar e matricular nos cursos dos Serviços Nacionais de Aprendizagem número de aprendizes equivalente a cinco por cento, no mínimo, e quinze por cento, no máximo, dos trabalhadores existentes em cada estabelecimento, cujas funções demandem formação profissional” (BRASIL, 2020).

Caso os Serviços Nacionais de Aprendizagem não ofereçam os cursos de acordo com a demanda do estabelecimento, pode-se ser matriculado em outras unidades de conhecimento qualificadas em formação técnico-profissional. Por fim, a contratação ocorre pela empresa de forma direta ou indireta, de acordo com os termos estabelecidos nos art. 429 e 431 da CLT. A contratação indireta acontece quando há intermédio de uma entidade profissional que também será a empregadora do jovem. Já a forma de contratação direta é aquela em que a empresa é responsável pela assinatura da Carteira de Trabalho e Previdência Social (CTPS).

A Assinatura da CTPS é na forma de contrato de trabalho com validade de dois anos, onde se estabelece a carga horária de trabalho, e o trabalho do aprendiz não excederá de seis horas diárias, sem prorrogação ou compensação da jornada de trabalho. Quanto à remuneração, a lei garante ao aprendiz o direito ao salário mínimo-hora, inclui-se também em horas trabalhadas aquelas destinadas às atividades teóricas.

Seguindo, ao considerar o funcionamento, implantação e objetivo do programa, é necessário observar a metodologia mais apropriada considerando as características observáveis disponíveis para avaliação da política pública, dado que a proposta estabelecida é resumida como o apresentado na Figura 4 denominada Teoria da Mudança para o Programa Jovem Aprendiz.

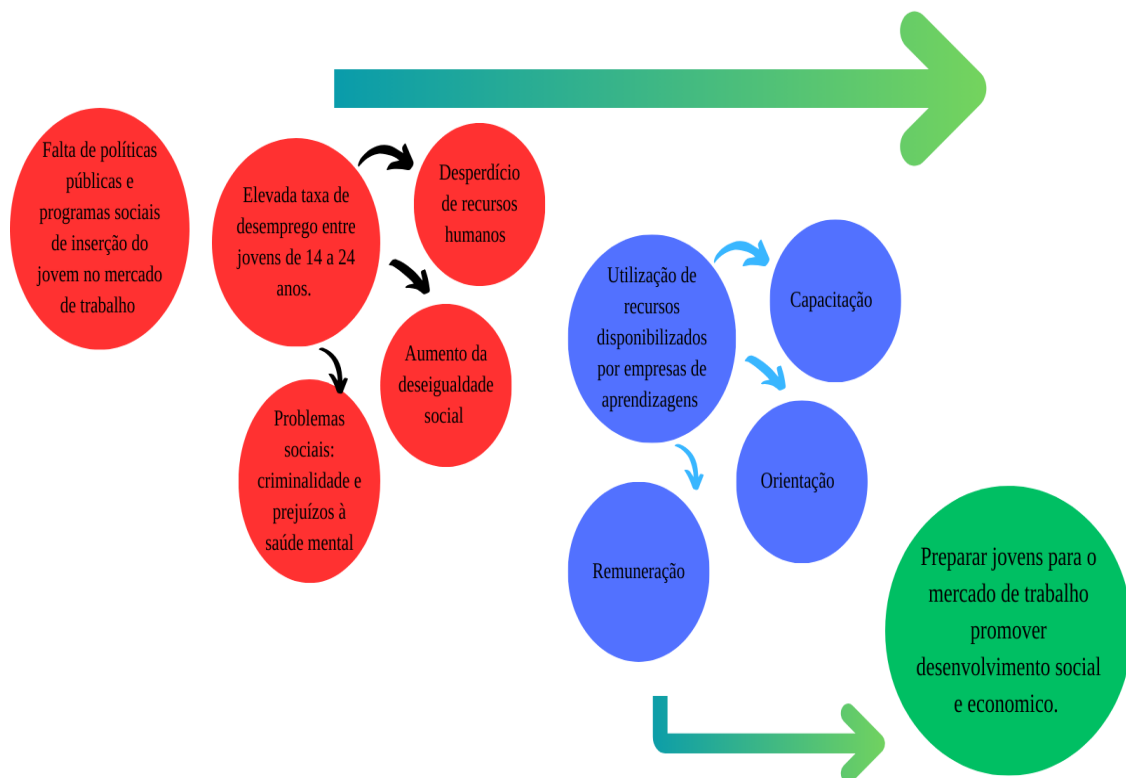
A Teoria da Mudança é uma espécie de roteiro que apresenta uma linha assíncrona de problemas e necessidades sociais, o resultado esperado de determinada política pública ou programa social, e o percurso que levará à resolução do problema em questão.

No caso do Programa Jovem Aprendiz a Teoria da Mudança pode ser observada na Figura 4. O problema que a sociedade enfrenta é a falta de programas e políticas focados na

inserção do jovem no mercado de trabalho assim como nos índices ruins de desocupação entre jovens de 14 a 24 anos, que podem ser desdobrados em três fatores:

i) desperdício de recursos humanos; ii) aumento da desigualdade social, já que jovens baixa renda enfrentam dificuldades ao ingressar no mercado de trabalho; iii) desemprego entre jovens pode levar à problemas sociais como criminalidade e comportamentos de risco à saúde mental.

Figura 4 - Teoria da Mudança Programa Jovem Aprendiz



Fonte: A autora.

O resultado final esperado com o programa é o desenvolvimento profissional dos jovens com a expectativa de que haja retorno educacional, profissional e econômico. Para alcançar tal objetivo, o programa se propõe a utilizar principalmente o recurso disponibilizado por empresas de aprendizagem, com equipes preparadas com orientações e capacitação.

Portanto, com intuito de garantir uma avaliação que absorva todas características inerentes do programa e que apresente resultados contundentes com os objetivos da pesquisa procurou-se utilizar uma metodologia de avaliação de política pública que melhor se adequasse à realidade do programa. Chegando assim ao modelo de Pareamento por escores de propensão que faz um comparativo entre participantes e não participantes, considerando a probabilidade de cada um participar, ou não, de acordo com as características selecionadas.

A escolha da abordagem econométrica de Pareamento por Escore de Propensão é crucial para esta pesquisa, pois permite controlar e ajustar as diferenças observáveis entre os grupos tratados e de controle, oferecendo uma estimativa mais precisa do impacto da política pública. Ao empregar o PSM, a pesquisa visa reduzir os vieses associados a características não controladas e obter uma avaliação mais robusta dos efeitos da política sobre os indivíduos. Na próxima seção, será detalhada a metodologia adotada, com uma justificativa abrangente para a escolha do modelo, levando em conta a especificidade da pesquisa e os desafios associados à análise em um contexto não experimental.

3 METODOLOGIA

Calcular o impacto de uma política pública sobre um indivíduo da sociedade é um trabalho subjetivo que, além de ser afetado por fatores internos, externos e desconhecidos que podem afetar no resultado, há também o problema de verificar o período que ocorre a observação. Portanto, se trata de uma situação não experimental. Considerando esta uma pesquisa quantitativa, em que se utiliza tratamento de dados mediante técnicas estatísticas (MICHEL, 2005), utilizou-se, especificamente, a abordagem econométrica via Pareamento por Escore de Propensão, do inglês *Propensity Score Matching* (PSM).

Ao avaliar uma política pública busca-se mensurar o efeito que determinada intervenção causa, causou ou causará em uma sociedade, ou seja, uma relação de causa e efeito (GERTLER *et al*, 2018). Diante disso, a regressão econométrica clássica por Mínimos Quadrado Ordinários (MQO) se torna inviável, por seguir uma distribuição aleatória. Caso fosse utilizada a regressão por MQO o resultado do valor esperado seria enviesado, em que o coeficiente estimado não seria o verdadeiro valor esperado, pois ao seguir a hipótese de distribuição aleatória implicaria dizer que o efeito de diferentes causas seria o mesmo. Portanto, a causa não pode ser correlacionada com a escolha de ser a causa ou não.

Neste sentido, como objetivo é medir o efeito em uma relação de causa, é necessário estimar o efeito médio de uma política sobre uma população utilizando métodos e metodologias que estimam o efeito médio do tratamento sobre a população minimizando o viés de estimação.

Assim, ao estimar uma relação causal em uma avaliação de políticas públicas, pretende-se obter resultados que informem qual o efeito destas políticas públicas sobre a população em geral e sobre os que receberam o tratamento. Considerando aqui a variável **Y** como a variável de interesse (neste caso, o impacto da política pública), e **T** uma variável binária, de participação ou não da política pública, tem-se, algebricamente:

$$E [T = 1] - E[Y|T = 0] \quad (1)$$

Em que, **T**=1 significa que a observação teve tratamento, caso **T** =0 a observação não obteve tratamento. A relação causal acontece entre **Y** e **T**, já **T** depende das variáveis **X**. Considerando *i* a *i*-ésima observação, tem-se:

$$Y_i = T_i(1) - T_i(0) \quad (2)$$

O resultado observado são dois estimadores: o efeito médio do tratamento para a população, que compara os que receberam tratamento com os que não receberam tratamento, em inglês: *Average Treatment Effect (ATE)*, como pode ser visto na equação 3; e o efeito

médio do tratamento entre os que receberam tratamento, compara entre aqueles que receberam tratamento, em inglês: *Average Treatment Effect on The Treated (ATT)*, como pode se observar na equação 4. Algebricamente descritos, respectivamente:

$$ATE = E[\tau] = E[Y(1) - Y(0)] \quad (3)$$

$$ATT = E[\tau|T = 1] = E[T = 1] - E[T = 1] \quad (4)$$

Quando houver diferença entre estes dois parâmetros, **ATE** e **ATT**, essa diferença é o viés, gerado por diferenças que existem pré-tratamento entre o grupo de tratados e de controle, pode ser descrito formalmente na equação 5.

$$VIÉS = E[T = 1] - E[T = 1] \quad (5)$$

O viés pode acontecer por quando há ausência de suporte comum ou sobreposição, quando há equívocos na seleção das características das variáveis observáveis e por omissão de variáveis selecionadas que não são observáveis (HECKMAN et al., 1997).

Portanto, ao considerar a natureza dos dados disponíveis para avaliação do programa e a consistência de resultados de acordo com o objetivo da avaliação, dentre os métodos de avaliação disponíveis a escolha do Pareamento por *Score* de Propensão deve-se por gerar resultados que permitem inferências de causa e efeito em pesquisas não experimentais em que os experimentos podem ser aleatórios ou não. Para tanto, tratar-se-á do modelo de Pareamento por *Score* de Propensão na subseção seguinte.

3.1 MODELO EMPÍRICO

Apesar da escolha metodológica ser por Avaliação de Políticas Públicas, é necessário escolher também o modelo de estimação que será feita a estimação da probabilidade. Ao observar a composição dos dados, é possível verificar participantes e não participantes do programa jovem aprendiz, o que favorece realizar comparação entre o grupo de tratados e o grupo de controle. Entretanto, a comparação é realizada entre pares que possuem o mesmo escore de propensão obtido por meio da estimação que será apresentada nas subseções a seguir.

3.1.1 Pareamento por *Score* de Propensão

A utilização do método de Pareamento por *Score* de Propensão proposto por Rusebaun e Rubin (1983) consiste em parear indivíduos tratados e não tratados com características

observáveis semelhantes em um experimento, em que a participação tenha sido de forma aleatória com intuito de reduzir o viés nos efeitos do tratamento. Pois, ao comparar os dois grupos, considerando a não aleatorização, pode-se gerar um viés em que não seria possível explicar o real efeito do programa.

Ao utilizar o PSM é possível identificar os reais efeitos do programa naqueles indivíduos que receberam o tratamento e naqueles que não receberam o tratamento, que possuem características observáveis com base nas variáveis de controle selecionadas. Essas características observáveis se tornam o vetor representado pela variável X , e elas são as responsáveis por reduzir o viés.

A teoria do *escore de propensão* é suficiente para diminuir possíveis vieses mesmo se a amostra analisada contiver poucos elementos, pois ao considerar as características observadas no vetor X , fica mais explícito o efeito de X sobre o tratamento T e consequentemente, sobre Y , tem-se formalmente descrito na equação 6 a seguir:

$$T_i = Y_i(1) - Y_i(0) \quad (6)$$

O Pareamento por *Score* de Propensão é uma correspondência discriminante, que se ajusta de acordo com a classificação obtida através do escore de propensão. Assim, após o pareamento é possível representar os resultados das combinações dos algoritmos em um gráfico bidimensional, o que deixa a análise mais fácil de ser interpretada.

O Pareamento por *Score* de Propensão consiste em separar em dois grupos o conjunto de observações, sendo um grupo de observações que receberam tratamento, o grupo de tratados, e o outro grupo com as observações que não receberam tratamento, chamado de contrafactual ou grupo controle. Ao comparar os dois grupos com as mesmas características observáveis, que não interferem na participação do programa, possibilita medir o efeito médio do tratamento mesmo que a seleção não tenha sido aleatória (KHANDKER; KOOLWAL; SAMAD, 2010).

Para iniciar o processo de estimação do impacto pelo método de pareamento por escore de propensão é necessário, inicialmente, obter-se os escores de propensão, que é a probabilidade de a observação estar no grupo que possui tratamento.

Continuamente, faz-se a escolha das variáveis observáveis que serão utilizadas para a estimação. A escolha de variáveis pode ocorrer de diferentes formas, porém devem obedecer a hipótese de independência incondicional, que significa que somente variáveis que não afetam a participação da observação ao tratamento; também é necessário que estejam fixadas no tempo, ou seja, medidas antes da participação, além de basear-se em alguma teoria

econômica, referencial bibliográfico ou aspecto institucional escolhido. Ou seja, é necessário que tenha um critério de escolha das variáveis com características observáveis e que esteja bem definido (MAFFIOLI, 2011).

Quanto aos métodos que as variáveis com características observáveis são escolhidas, um método disponível e utilizado foi o “*hit or miss*” em que as variáveis são escolhidas para maximizar a predição correta da amostra (HECKMAN et al., 1998), e também pode ser combinado com a significância estatística de cada variável do modelo, pois pode aumentar a taxa de predição.

Então, após selecionar as variáveis pode-se estimar o escore de propensão, que é a probabilidade de uma observação participar do programa de acordo com suas características observáveis representado pelo vetor X , podendo ser um número real entre 0 e 1 (GERTLER et al; 2018). Ao considerar a dicotomia de participar ou não do tratamento, normalmente, utiliza-se um modelo de estimação *probit* ou *logit*. Como a amostra é aleatória, há a possibilidade de observações do grupo de tratados não terem pares com escore de propensão semelhantes no grupo de observações que não receberam tratamento durante o pareamento, estas observações sem pares são excluídas da amostra, pois podem gerar uma falta de suporte comum.

Portanto, para que o PSM seja validado há a necessidade de obedecer a duas hipóteses: a de suporte comum e de independência incondicional (CIA, do inglês *Conditional Independence Assumption*). A hipótese de suporte comum se trata do fato de o grupo de tratados e não tratados terem as mesmas características observáveis, pois isto possibilita criar grupos de comparação com consistência para o pareamento.

Existem diferentes estratégias para confirmar a condição de área de sobreposição ou região de suporte comum. A região de suporte comum refere-se a região aonde os parâmetros ATT ou ATE são definidos, ou seja, a região que existe observações com mesmo escore de propensão tanto do grupo de tratados quanto do grupo de controle, e assim, quando se sobrepõe essa região de cada grupo chama-se a região de suporte comum. (CALLIENDO e KOPEINIG, 2005).

Confirmar essa suposição significa que, se existem duas observações com o mesmo escore de propensão cada uma terá a mesma probabilidade de participar do tratamento com base nas variáveis observáveis e assim, se torna possível comparar indivíduos que participam do grupo de tratados e possuem as mesmas características observáveis e mesmo escore de propensão, entretanto não participam do grupo de tratados e sim de controle. Portanto, o

suporte comum ou região de sobreposição pode ser testada pela análise visual, pelo critério de valores mínimo e máximo, por teste inferencial ou por diferença média dos escores de propensão.

A outra hipótese fundamental para que seja possível realizar o pareamento sem vieses é a hipótese de independência incondicional, que significa que as variáveis observáveis, do vetor $\mathbf{X}$, não devem sofrer influência do programa, elas devem ser independentes do tratamento (RIBEIRO; CACCIAMALI, 2012). Caso os dados não sigam estas hipóteses é possível mesclar o pareamento com outros modelos capazes de reduzir os vieses, contudo não será o caso desta pesquisa.

Em casos de avaliações de políticas públicas, a hipótese de independência incondicional ajuda a estimação por PSM, por estimar o efeito da política pública assumindo que, a participação no tratamento é independente dos resultados da política ou programa em questão. Quando há a quebra desta hipótese, ou seja, quando o tratamento é influenciado por variáveis não observáveis gera viés nos impactos estimados.

Neste contexto, insere-se o Programa Jovem Aprendiz que não foi distribuído aleatoriamente, visto que obedeceu a critérios de seleção para a participação do indivíduo no programa e caso a aleatoriedade fosse assegurada, o grupo de controle teria a mesma distribuição de características do grupo de tratamento.

Desta forma, a abordagem do pareamento pelo escore de propensão surge como uma possível solução para o problema do viés de seleção, visto que cada participante seria comparado com um não participante, com base em um escore de propensão, refletindo a probabilidade de participação condicional em suas diferentes características $\mathbf{X}$ observadas, o que representa as duas hipóteses básicas do modelo: independência condicional e suporte comum.

Assumindo que as hipóteses são válidas, e que a população é elegível para participar do tratamento a estimação do efeito médio do tratamento (do inglês **ATT**) pode ser realizada através da diferença média em $\mathbf{Y}$ no suporte comum, com a ponderação das observações que serão comparadas através do escore de propensão, portanto, formalmente tem-se a equação 7:

$$ATT_{PSM} = E[\tau|T = 1]\{E[T = 1, P(X)] - E[T = 0, P(X)]\} \quad (7)$$

Estimar o escore de propensão não é suficiente para estimar o efeito médio do tratamento, pois existe a probabilidade de, pelo menos, duas com o mesmo escore de propensão ser zero, impossibilitando o pareamento. Para tanto, para continuar a estimação escolhe-se os algoritmos ou métodos de correspondência para realizar o pareamento. Esses

algoritmos são considerações sobre a forma de parear as observações do grupo de controle com o grupo de tratados (BECKER e ICHINO, 2002).

Na seguinte subseção apresenta-se uma breve explicação de como o modelo *logit* estima os escores de propensão, o que pode ser chamado de razões de chances de o indivíduo participar ou não do grupo de tratamento.

3.1.2 Modelo de Regressão *Logit* para estimar Escore de Propensão

Para estimar escore de propensão é necessário determinar um modelo de regressão que será estimado o escore de propensão para cada observação e poder prosseguir com o pareamento. A utilização do modelo de regressão *Logit* resulta da possibilidade de estimar a ocorrência ou não de um evento, no caso a participação ou não de um evento.

Em que a variável dependente possui características qualitativas e dicotômicas ou *dummies*, pois assume valores 0 ou 1. Por se tratar de um intervalo limitado não é possível realizar a estimação utilizando um Modelo de Regressão Linear (MRL) tendo em vista o limite de 0 ou 1.

Portanto, quando a variável dependente é binária os modelos de regressão mais indicados são o Modelo de regressão *Logit* ou *Probit*. Em que a estimação destes modelos é simples por terem distribuição *Bernoulli* e, geralmente, por Máxima Verossimilhança (CAMERON E TRIVEDI; 2005).

Diante do exposto, o modelo de regressão de resposta qualitativa a ser utilizado neste trabalho é o modelo de regressão *Logit*, considerando a estimação pelo modelo de regressão logística, além de ser mais simples, a função de distribuição logística possui cauda mais pesada, o que significa que a probabilidade de alcançar os valores de 0 e 1 possui um ritmo mais lento que no modelo *Probit* (GUJARATI E PORTER; 2011).

O modelo de regressão *Logit* busca determinar a probabilidade de ocorrência de um evento, em que a variável dependente Y é binária, só pode assumir valores de 0 ou 1, e o conjunto P é o conjunto de variáveis independentes ($X_1, X_2, X_3 \dots X_p$). Para estimação o modelo pode ser escrito na forma da equação 8:

$$P_i = \frac{1}{1 + e^{-(\beta_1 + \beta_2 X_2 + \dots + \beta_p X_p)}} \quad (8)$$

P_i representa a probabilidade de o indivíduo participar do programa. Substituindo $(\beta_1 + \beta_2 X_2 + \dots + \beta_p X_p)$ por Z_i e simplificando a equação (8) obtém-se a função de densidade logística representada na equação a seguir:

$$P_i = \frac{1}{1 + e^{-Z_i}} = \frac{e^Z}{1 + e^Z} \quad (9)$$

A equação (8) é não linear, então não é possível estimar parâmetros por mínimos quadrados ordinários que seguem o modelo de regressão linear. Para tal é calculado a razão de chances (em inglês *odds ratio*) e realiza-se a linearização. Sendo a razão de chances a razão entre a probabilidade participar ou não do programa; P_i a probabilidade de o indivíduo participar do programa, então $1-P_i$ é a probabilidade de o indivíduo não participar e pode ser escrito da seguinte forma:

$$1 - P_i = \frac{1}{1+e^{Z_i}} \quad (10)$$

Portanto a razão de chances (*odds ratio*) ao ser simplificada pode ser representada por:

$$\frac{P_i}{1-P_i} = \frac{1+e^{Z_i}}{1+e^{-Z_i}} = e^{Z_i} \quad (11)$$

Aplicando o logaritmo natural tem-se a linearização na equação 12:

$$L_i = \ln \ln \left(\frac{P_i}{1-P_i} \right) = Z_i \quad (12)$$

Lembrando que $Z_i = (\beta_1 + \beta_2 X_2 + \dots + \beta_p X_p)$.

Então a partir da razão de chances (*odds ratio*) linearizada é possível realizar a estimação dos parâmetros. Entretanto, apesar de linearizado, não é possível realizar a estimação dos parâmetros por MQO, o método a ser utilizado será por Máxima Verossimilhança.

Para estimar, a função *Logit* L_i passa a ter o formato apresentado na equação 13:

$$L_i = \ln \ln \left(\frac{P_i}{1-P_i} \right) = \beta_1 + \beta_2 X_2 + \dots + \beta_p X_p + u_i \quad (13)$$

Onde u_i representa o termo de erro estocástico do modelo.

A razão de chances (*odds ratio*) matematicamente pode ser representada por:

$$L_i = \ln \ln \left(\frac{P_i}{1-P_i} \right) = \ln \left(\frac{1}{0} \right) = \beta_1 + \beta_2 X_2 + \dots + \beta_p X_p + u_i \quad \text{se o evento acontecer;}$$

$$L_i = \ln \ln \left(\frac{P_i}{1-P_i} \right) = \ln \left(\frac{0}{1} \right) = \beta_1 + \beta_2 X_2 + \dots + \beta_p X_p + u_i \quad \text{se o evento não acontecer.}$$

A probabilidade de ocorrência do evento é P_i e de não ocorrência é $1-P_i$ (FÁVERO E BELFIORE, 2017) Portanto, a probabilidade de ocorrência do evento pode ser escrita como:

$$p(Y_i) = P_i^{Y_i} * (1 - P_i)^{1-Y_i} \quad (14)$$

Sendo $P_i = \frac{e^z}{1+e^z}$, e $1 - P_i = \frac{1}{1+e^z}$ pode-se substituir os valores respectivamente na equação 14 e definir a função de máxima verossimilhança como:

$$L = \prod_{i=1}^n \left[\left(\frac{e^z}{1+e^z} \right)^{Y_i} * \left(\frac{1}{1+e^z} \right)^{1-Y_i} \right] \quad (15)$$

Para resolver a função de verossimilhança é mais conveniente aplicar o logaritmo transformando em somatório e aplicando logaritmo natural no restante da equação. Tem-se, então a equação log verossimilhança:

$$LL = \sum_{i=1}^n \{ [(Y_i) \ln \left(\frac{e^z}{1+e^z} \right)] + [(1 - Y_i) \ln \left(\frac{1}{1+e^z} \right)] \} = \text{máx} \quad (16)$$

A estimação dos parâmetros β , neste trabalho, será realizada através do uso do *software* R de programação linear capaz de encontrar as estimativas para os parâmetros e estimar os escore de propensão das observações tratadas em relação às não tratadas.

Na próxima seção aborda-se os algoritmos de pareamento, uma ferramenta crucial na superação do desafio relacionado à probabilidade de pelo menos duas observações apresentarem o mesmo escore de propensão, resultando em uma probabilidade zero. Estes algoritmos, destacados na literatura, são escolhidos com base na finalidade do pareamento e levam em consideração as características observáveis das amostras.

3.1.3 Algoritmo de Pareamento

Os algoritmos de correspondência ou de pareamento passaram a ser utilizados na literatura como método de superar o problema da existência da probabilidade de pelo menos duas observações com o mesmo escore de propensão, ser zero. Eles são escolhidos de acordo com a finalidade do pareamento e consideram as características observáveis das observações na amostra. A escolha do método de pareamento mais indicado é determinada por características a serem observadas e cada um tem uma peculiaridade específica.

Os algoritmos de pareamento são considerados quanto a reposição das observações; quanto a proximidade das observações, se e como realizar a ponderação; e a quantidade de unidades de comparação correspondente a cada unidade de tratamento (MAFFIOLI et al, 2011).

Quando se considera o Pareamento é necessário decidir se haverá a reposição ou não da observação do grupo de controle que já foi pareada, daí será com ou sem substituição. O pareamento sem reposição acontece quando a combinação é feita com uma observação do grupo de tratados e com outra do grupo de controle, e a observação que não recebeu o tratamento não será comparada com alguma outra observação que pertence ao grupo de tratado, mesmo que tenha características iguais e seja elegível à comparação (AUSTIN, 2012). Ao realizar o pareamento de um para um de cada grupo da amostra é indicado a utilização da amostra do grupo de comparação sem substituição, ou seja, uma vez pareada, a

observação não tem reposição na amostra de origem. Entretanto, discute-se que, em amostras pequenas, a não reposição pode limitar a estimação do PSM, formando grupos de comparação reduzido e as probabilidades podem não estarem tão próximas às probabilidades do grupo de tratados, então se tornou comum realizar o pareamento com reposição em que as observações do grupo de controle com probabilidade mais próxima podem ter correspondência com mais de uma observação do grupo de tratados (DEHEJIA E WAHBA, 2002). A reposição das observações favorece também a redução da variância pois ao não considerar a reposição reduz-se a quantidade de pareamentos, e isso aumenta a variância, ou seja, quanto maior a quantidade de pareamentos menor será a variância.

Quando o pareamento é baseado na proximidade do escore das observações chama-se de “vizinhos mais próximos”. Neste algoritmo de pareamento é necessário que aconteça com reposição, ainda é possível dois tipos de pareamento: pareamento par a par, em que uma observação do grupo de tratados é pareada com uma observação do grupo controle chamado de par ideal e o pareamento completo, onde uma observação do grupo de tratados é pareada com várias observações do grupo de controle ou uma observação do grupo de controle é pareada com várias do grupo de tratados.

Outra opção de método de pareamento é por estratificação, que consiste em separar as observações de cada grupo em intervalos com observações que possuem, na média, mesmo escore de propensão. A literatura mostra que a estratificação em até 5 grupos é suficiente.

O algoritmo de pareamento de kernel é bastante utilizado também no modelo de pareamento por escore de propensão. Esse método compara o escore de cada observação tratada com uma média ponderada das observações do grupo de não tratados. É um método que gera uma menor variância, entretanto pode gerar pareamentos ruins com mais vieses pois ao comparar com uma média ponderada perde-se a especificidade e individualidade de cada observação comparada (MAFFIOLI et al, 2011).

Apesar de diferentes métodos de correspondência, não há regra que especifique qual mais indicado para determinado conjunto de variáveis. Em geral, os métodos de correspondência por escore de propensão devem produzir os mesmos resultados. Entretanto, é importante que o tamanho da amostra seja considerado na escolha do melhor método, sendo importante observar qual algoritmo gera menor viés, assim como, qual algoritmo se encaixa de acordo com as características observáveis das variáveis selecionadas (CALLIENDO e KOPEINIG, 2005).

Após a escolha do algoritmo de pareamento que melhor se adequa às características observáveis das variáveis, a etapa seguinte é testar a qualidade do método de correspondência escolhido.

3.1.4 Teste de qualidade do Pareamento

A avaliação do método de pareamento escolhido é de suma importância, pois é através dele que é possível verificar se houve ou não efeito do tratamento. Para tanto, a avaliação pode ser feita de três maneiras estatística. No primeiro é utilizado o viés de seleção com o teste de significância, o segundo teste utiliza o viés padronizado, já o último utiliza a redução do viés percentual.

Considerando agora viés de seleção como a diferença média entre as condições de tratamento temos a equação 17:

$$VIÉS = E[T = 1] - E[T = 0] \quad (17)$$

Dessa forma, o viés é identificado e, em seguida, é conduzido um teste-t ou teste-z sob a hipótese nula $H_0: VIÉS = 0$, a fim de realizar o teste de significância do viés de seleção. Ao realizar o teste-t, espera-se que não haja diferença significativa entre o viés antes e após o pareamento. No entanto, é importante observar que, apesar de ser um método estatístico, este teste torna-se bastante sensível ao tamanho da amostra, pois em amostras pequenas o viés tende a ser maior (AUSTIN, 2011).

A alternativa então, será encontrar o viés padronizado definidos na equação 18 (ROSENBAUM; RUBIN, 1985):

$$VP = \frac{VIÉS}{\sqrt{\frac{V_1(X_k) + V_0(X_k)}{2}}} \times 100\% \quad (18)$$

Em que $V_0(X_k)$ é a variância do grupo de controle e $V_1(X_k)$ é a variância do grupo de tratados. Então, na literatura quando o viés padronizado for menor que 5% o método de pareamento utilizado pode ser considerado eficaz (CALLIENDO e KOPEINIG, 2008).

Outro teste que pode ser realizado para avaliar o método de pareamento é através da redução do viés padronizado. O referido teste consiste em:

$$RVP = \frac{VIÉS_0 - VIÉS_1}{VIÉS_0} \times 100\% \quad (19)$$

Em que $VIÉS_0$ representa o viés antes do pareamento e $VIÉS_1$ representa o viés após o pareamento. Após esta redução, caso o resultado seja superior a 80% informa que o pareamento foi eficaz (COCHRAN E RUBIN, 1973).

Após testar os pareamentos e encontrar resultados satisfatórios, finalmente pode-se então calcular o efeito do tratamento.

3.1.5 Efeito Médio do Tratamento

Após verificar as condições de independência condicional, estimar o escore de propensão e de suporte comum, realizar o pareamento e os testes dos pareamentos pode-se então calcular, de fato, o efeito médio do tratamento.

Considerando que, verificar o efeito médio do tratamento para a população consiste em comparar o grupo de recebeu tratamento e o que não recebeu tratamento é importante atentar-se para que a comparação aconteça no mesmo ambiente e na mesma posição medida pelo escore de propensão.

Como já foi discutido, o efeito médio do tratamento utilizando o método de pareamento por escore de propensão é a diferença média em Y no suporte comum, com a ponderação das observações que serão comparadas através do escore de propensão, portanto, tem-se equação de Efeito Médio do Tratamento:

$$ATT_{PSM} = E[\tau|T = 1]\{E[T = 1, P(X)] - E[T = 0, P(X)]\} \quad (20)$$

Em ato contínuo, após o pareamento é possível ter o efeito médio do tratamento e então é possível realizar o teste de significância estatística do efeito do tratamento e calcular o erro padrão.

Portanto, ao estimar via pareamento por escore de propensão encontra-se o ATT que é o efeito médio do tratamento para a população através de uma amostra. Para tanto, faz-se necessário ter um conhecimento prévio da base de dados, e por consequência conhecer e selecionar as variáveis com características observáveis que possam afetar a variável de interesse e a seleção para o tratamento. Na seguinte seção será descrita a composição da base de dados e as variáveis selecionadas para a estimação.

3.2 BASE DE DADOS

Nesta seção apresenta-se os dados e as variáveis utilizadas para construção do modelo e como foi feito o processo de seleção utilizado nas estimações.

A base de dados utilizada nesta pesquisa foi formada a partir de microdados da Relação Anual de Informações Sociais (RAIS) vínculos, que consiste em um registro administrativo disponibilizado anualmente pelo Ministério do Trabalho através do Programa de Disseminação das Estatísticas de Trabalho (PDTE), e funciona como instrumento de controle e obtenção de estatísticas que se transformam em informações para entidades governamentais e para a sociedade.

Nessa relação anual de informações sociais encontram-se organizadas informações sobre o mundo do trabalho, principalmente, vínculos empregatícios, número de trabalhadores, tipo de relação contratual, por estabelecimento, por tamanho do estabelecimento, faixa etária, admissões assim como desligamentos, remuneração entre outras informações acerca do mundo do trabalho no território brasileiro.

Considerando o período da pesquisa, foram utilizadas as informações da RAIS vínculos de jovens brasileiros entre os anos de 2014 e 2018. Dentro dessa população, realizou-se um filtro pela idade dos jovens, considerando que o programa jovem aprendiz atua sobre jovens entre 14 e 24 anos. Entretanto, se o corte fosse realizado no intervalo de 14 a 24 anos o viés da estimação seria maior, pois a única relação empregatícia, aqui chamada de tipo de admissão, de jovens entre 14 e 17 anos é do tipo jovem aprendiz. Para tanto, com intuito de evitar maior viés de seleção, optou-se por filtrar jovens de 18 a 24 anos de idade gerando uma amostra de indivíduos com informações socioeconômicas.

Os dados em *cross-section* foram coletados na RAIS, portanto a amostra foi composta por observações que se referem a indivíduos que possuíram algum tipo de vínculo trabalhista no ano e delimitados pela idade. Na tabela 1 verifica-se o total anual de observações por ano.

Tabela 1- Quantidade de jovens trabalhadores por ano

Ano	Quantidade de jovens	Quantidade de jovens		
2014	8.117.679	Controle (0)	7.961.178	98%
		Tratados (1)	156.501	1,9%
2015	7.282.092	Controle (0)	7.104.683	97,5%
		Tratados (1)	177.409	2,49%
2016	6.596.171	Controle (0)	6.402.484	96,9%
		Tratados (1)	193.687	3,02%
2017	6.457.432	Controle (0)	6.230.853	96,5%
		Tratados (1)	226.579	3,5%
2018	6.396.299	Controle (0)	6.127.015	95,8%
		Tratados (1)	269.284	4,2%

Fonte: A autora utilizando dados da RAIS (2021).

A tabela 1 apresenta a quantidade anual de jovens trabalhadores, detalhada por ano de 2014 a 2018, e divide esses trabalhadores em dois grupos: controle (não tratados) e tratados (jovens participantes do programa). Em 2014, havia um total de 8.117.679 jovens trabalhadores, dos quais 7.961.178 (98%) pertenciam ao grupo de controle e 156.501 (1,9%) eram tratados. Em 2015, o total de jovens trabalhadores foi de 7.282.092, com 7.104.683 (97,5%) no grupo de controle e 177.409 (2,49%) no grupo de tratados. Para 2016, o total de jovens trabalhadores caiu para 6.596.171, sendo 6.402.484 (96,9%) do grupo de controle e

193.687 (3,02%) do grupo de tratados. Em 2017, o total foi de 6.457.432 jovens, com 6.230.853 (96,5%) no grupo de controle e 226.579 (3,5%) no grupo de tratados. Finalmente, em 2018, o total de jovens trabalhadores foi 6.396.299, com 6.127.015 (95,8%) no grupo de controle e 269.284 (4,2%) no grupo de tratados. A tabela ilustra uma tendência de aumento gradual na proporção de jovens trabalhadores que são tratados pelo programa ao longo dos anos, enquanto a proporção de jovens no grupo de controle diminui. Em 2014, os tratados representavam apenas 1,9% do total de jovens trabalhadores, mas essa proporção aumentou para 4,2% em 2018. Essa mudança pode refletir um crescimento na cobertura ou na adesão ao programa ao longo dos anos.

Para a construção do grupo de tratamento e grupo de controle, foi necessário levar em consideração o tipo de vínculo. O indivíduo que possui o vínculo de jovem aprendiz foi agrupado no grupo que recebeu o tratamento, e o que não possuía esse vínculo ficou no grupo de controle. Assim, pode-se verificar o grupo de tratados e de controle para cada ano. Como dito anteriormente foram excluídos indivíduos que não possuíam a idade de corte para fazer parte da amostragem, portanto a amostra foi dividida em:

- **Grupo controle:** jovens entre 18 e 24 anos que não possuíam o tipo de vínculo de jovem aprendiz;
- **Grupo tratados:** jovens entre 18 e 24 anos que possuíam o tipo de vínculo de jovem aprendiz.

A variável de interesse utilizada para verificar o efeito do programa jovem aprendiz foi o rendimento. Para esta análise da renda do jovem foi necessário a construção de uma *proxy* rendimento que consiste em o salário recebido em dezembro dividido pela quantidade de horas mensais considerando um mês de 4 semanas.

A seleção das variáveis independentes realizou-se através do método “*hit or miss*” que consiste na seleção de variáveis que aumentam a probabilidade da predição correta da amostra, nesse caso a seleção de variáveis com características observáveis, combinando com a significância estatística de cada (HECKMAN et al., 1998). Na tabela 2 é possível verificar as variáveis com características observáveis selecionadas, assim como os sinais esperados para estimar a probabilidade de participar do programa jovem aprendiz.

Tabela 2- Variáveis e sinais esperados para estimação do score de propensão

Variáveis	Descrição	Sinal esperado
Jovem Aprendiz	Variável binária igual a 1 se o indivíduo possuir o tipo de vínculo jovem aprendiz, e zero se for outro tipo de vínculo.	variável dependente
Rendimento	Variável que utiliza a remuneração do trabalhador em dezembro, dividido pela quantidade de horas de trabalho mensal.	positivo
Idade	Variável que representa a idade do jovem entre 18 e 24 anos.	negativo
Mulher	Variável binária igual a 1 se o indivíduo é mulher e 0 se for homem	negativo
Negro	Variável binária igual a 1 se o indivíduo é negro e 0 se for outra raça;	negativo
	Variável que mede o grau de estudo do indivíduo e pode variar de 1 até 7 de acordo com a quantidade de anos estudados.	positivo
Escola	1- Analfabeto 2 - Possui até o 5º ano incompleta 3 - Possui até o 5º ano completo. 4 - Possui do 6º ao 9º ano incompleto. 5 - Possui o ensino fundamental completo. 6 - Possui o ensino médio incompleto. 7 - Possui o ensino médio completo	positivo

Fonte: A autora com base nos dados da RAIS (2021).

Apesar de existirem mais covariáveis disponíveis para seleção na base de dados bruta disponibilizada pela RAIS e o volume de covariáveis aumentarem a possibilidade de um grupo de tratados mais semelhante ao grupo de controle, a escolha destas e se deu também pelo fato de quanto mais covariáveis no modelo maior a dificuldade de encontrar uma unidade no grupo de controle que seja mais semelhante no grupo de tratamento.

Na próxima seção se discute os resultados obtidos através do pareamento via escore de propensão.

4 ANÁLISE E DISCUSSÃO DE RESULTADOS.

Nesta seção são apresentados os resultados da pesquisa. De início, faz-se uma análise univariada através da estatística descritiva das variáveis utilizadas no modelo de estimação de probabilidade de participação no programa. Em seguida, é analisada, separadamente, a estatística descritiva do grupo de tratados e do grupo de controles e feita uma simples comparação considerando a média. Posteriormente, são discutidos os resultados obtidos na estimação da Probabilidade de participação do Programa Jovem Aprendiz, que consiste em encontrar o escore de propensão de cada observação. Por fim, são analisados os Efeitos do Programa Jovem Aprendiz sobre o Rendimento obtido por meio do pareamento.

Dado o objetivo principal desta pesquisa, que é verificar o impacto do programa jovem aprendiz sobre o rendimento, foi necessário compreender como a amostra estava distribuída de acordo com as características observáveis dos indivíduos. Para isso, foram utilizados dados cuja análise das médias encontra-se resumida na tabela 3.

Considerando ainda a Tabela 1, na seção anterior, é possível verificar que com o passar dos anos, a quantidade de jovens trabalhadores no mercado de trabalho entre 18 e 24 anos diminuiu. O que em 2014 era 8.117.679 jovens trabalhadores, em 2018 chegou a 6.396.299, representando uma queda de 21,20% na quantidade de jovens trabalhadores durante o período observado.

Apesar da queda na quantidade total de jovens, quando se olha para a quantidade de jovens que participaram do Programa Jovem Aprendiz houve um aumento na proporção, O que em 2014 era menos de 2% em 2018 chegou a quase 4,5%, crescendo a cada ano.

Na tabela 3, é viável examinar o percentual médio anual de cada característica observável, culminando em uma média percentual global das características ao longo do período de análise.

Tabela 3– Média das variáveis de características observáveis por ano

Variável	Média 2014	Média 2015	Média 2016	Média 2017	Média 2018	Média Geral
Jovem Aprendiz	0,02418	0,02436	0,03858	0,03509	0,0421	0,032862
Rendimento	4,424	8,549	5,4578	9,494	10,171	7,61916
Idade	21	21	21	21	21	21
Mulher	0,4179	0,4309	0,4225	0,4249	0,4269	0,42462
Negro	0,5214	0,5375	0,5514	0,5603	0,5708	0,54828
Escolaridade	6,493	6,651	6,6218	6,738	6,798	6,66036

Fonte: A autora com base nos dados da RAIS.

No entanto, ao analisar a Tabela 3, percebe-se é possível observar que as características desses jovens permaneceram consistentes ao longo dos anos.

A Tabela 3 apresenta a média das variáveis de características observáveis por ano. No caso da variável "Jovem Aprendiz", a média foi de 0,02418 em 2014, 0,02436 em 2015, 0,03858 em 2016, 0,03509 em 2017, e 0,0421 em 2018, resultando em uma média geral de 0,032862. O "Rendimento" apresentou uma média de 4,424 em 2014, 8,549 em 2015, 5,4578 em 2016, 9,494 em 2017, e 10,171 em 2018, com uma média geral de 7,61916. A "Idade" manteve-se constante em 21 anos durante todos os anos observados. A variável "Mulher" teve médias de 0,4179 em 2014, 0,4309 em 2015, 0,4225 em 2016, 0,4249 em 2017, e 0,4269 em 2018, resultando em uma média geral de 0,42462. A variável "Negro" apresentou médias de 0,5214 em 2014, 0,5375 em 2015, 0,5514 em 2016, 0,5603 em 2017, e 0,5708 em 2018, com uma média geral de 0,54828. Por fim, a variável "Escolaridade" teve médias de 6,493 em 2014, 6,651 em 2015, 6,6218 em 2016, 6,738 em 2017, e 6,798 em 2018, resultando em uma média geral de 6,66036.

Em média, cerca de 3% dos participantes estão envolvidos no programa Jovem Aprendiz. Essa proporção de 3% para os tratados em comparação com os 97% dos controlados revela-se geralmente favorável, promovendo pareamentos eficazes. No que diz respeito ao gênero, a predominância é masculina, com uma média de 42% representando o sexo feminino. Quanto à etnia, aproximadamente metade da amostra se identifica como não branca, totalizando em média 55%, enquanto a outra metade é composta por indivíduos brancos.

Ao analisar a escolaridade, composta por graus de instrução do 1 ao 7, em que 1 é analfabeto, 2 possui até a 5º ano incompleta, 3 possui até a 5º ano completo, 4 possui do 6º ao 9º ano incompleto; 5 possui o ensino fundamental completo, 6 possui o ensino médio incompleto e, por fim, 7 que representa o ensino médio completo. A análise univariada da

variável escolaridade traz que a média de grau de instrução que os jovens da amostra possuem é o ensino médio incompleto, o que é coerente considerando a idade dos jovens da amostra. Jovens com 18 anos estão finalizando o ensino médio e estão no processo de início na graduação.

A análise da Tabela 3 revela algumas tendências importantes ao longo dos anos de 2014 a 2018. A participação de jovens aprendizes aumentou de 0,02418 para 0,0421, enquanto o rendimento médio variou significativamente, com uma média geral de 7,61916, indicando flutuações econômicas ou mudanças no mercado de trabalho. A idade média permaneceu constante em 21 anos, refletindo a homogeneidade do grupo. A participação feminina aumentou ligeiramente de 0,4179 para 0,4269, e a proporção de pessoas negras cresceu de 0,5214 para 0,5708, sugerindo maior inclusão. Por fim, o nível de escolaridade também mostrou um aumento gradual, de 6,493 para 6,798 anos, indicando uma melhoria educacional ao longo do período. Essas tendências podem refletir mudanças sociais, econômicas e políticas ao longo dos anos analisados.

Para estimar o PSM foi utilizado uma base de dados além de separada por anos e foi também separada entre grupo de controle e de tratamento, que serão chamados de jovens aprendiz aqueles que possuem o tipo de vínculo como jovem aprendiz e jovens não aprendiz aqueles jovens que estão em outro tipo de vínculo diferente de jovem aprendiz.

Nesse sentido, para atender as premissas de implementação do pareamento por escore de propensão proposto é necessário seguir as etapas: primeiro estima-se o modelo de escolha discreta, e aqui foi pelo modelo *logit*, pois é através dessa estimação que se obtém o escore de propensão. Em seguida, como descrito na seção anterior, escolhe-se o algoritmo mais indicado e realiza-se o pareamento, e por fim verifica-se a qualidade.

Na próxima seção, portanto, pode-se verificar a probabilidade de participação do Programa Jovem Aprendiz, que consiste na obtenção do escore de propensão através do modelo de estimação de escolha discreta: *logit* e na obtenção da magnitude do efeito da variável de característica observável na variável de interesse.

4.1 PROBABILIDADE DE PARTICIPAÇÃO DO PROGRAMA JOVEM APRENDIZ - ESCORE DE PROPENSÃO

A avaliação de impacto de uma política pública consiste em comparar grupos que receberam o tratamento ofertado pela política pública e aquele grupo que não recebeu o tratamento, mas que possuía todas as características necessárias, porém, por alguma questão

desconhecida, não participou, nesta seção são apresentadas as estimativas utilizando o modelo de regressão *logit* com intuito de obter os escores de propensão necessário para realizar o pareamento.

Para analisar o impacto do programa jovem aprendiz no rendimento destes jovens no Brasil nos anos de 2014 a 2018 por pareamento por escore de propensão é necessário, primeiramente, calcular os escores de propensão, considerando o escore de propensão como a probabilidade de um jovem ter participado ou não do programa jovem aprendiz, ou seja, uma probabilidade em que a variável dependente é dicotômica, utiliza-se o modelo de regressão *logit* para encontrar as estatísticas desta variável.

Portanto, para estimar os escores de propensão, realizar o pareamento e, assim, poder avaliar o impacto do Programa Jovem Aprendiz foi necessário estimar utilizando o modelo de regressão *logit* e os resultados estão na tabela 4.

Tabela 4 – Coeficientes da probabilidade de participar do Jovem Aprendiz estimados pelo modelo Logit

Variáveis explicativas	2014	2015	2016	2017	2018
	Coeficiente	Coeficiente	Coeficiente	Coeficiente	Coeficiente
Intercepto	11,49348*** (0,040847)	10,5544*** (0,1586)	12,64536*** (0,0372)	12,56841*** (0,034218)	12,85887*** (0,031434)
Idade (em anos)	-0,81470*** (0,002046)	-0,79003*** (0,0019)	-0,85737*** (0,00183)	-0,843336*** (0,001670)	-0,83007*** (0,001519)
Mulher	0,345976*** (0,005337)	0,32952*** (0,00508)	0,320458*** (0,00488)	0,316564 *** (0,004558)	0,31862*** (0,004207)
Negro	0,499433*** (0,005388)	0,51222*** (0,00518)	0,394854*** (0,00492)	0,324398*** (0,004588)	0,30169*** (0,004248)
Escolaridade e	0,077227*** (0,002489)	1,50865* (0,7275)	0,127431*** (0,002519)	0,136049*** (0,002459)	0,08915*** (0,002306)

Fonte: A autora com base no tratamento de dados da RAIS, através do software R

Nota: Valores entre parênteses é o desvio padrão;

*** Significância de 1%; ** Significância de 5%; * Significância de 10%.

Após realizar a estimação da probabilidade de participar do programa segundo suas características, através do *software* R, os resultados da estimação obtidos foram apresentados na Tabela 4, em que a variável dependente é uma variável binária, em que se igual a 1 significa que o indivíduo participou do programa jovem aprendiz, e se igual a 0 significa que o indivíduo não participou do programa jovem aprendiz. Observa-se na mesma tabela que todas as variáveis são estatisticamente significantes a menos de 1%, garantido um impacto negativo ou positivo na variável de interesse de acordo com o sinal do coeficiente.

Considerando que os coeficientes obtidos não representam a magnitude da variável no modelo de regressão. Dessa forma, na Tabela 4 apresenta-se os resultados da estimação, em que é possível verificar que a idade reduz a probabilidade de o jovem participar do programa, com significância estatística de 1%. Entretanto, ser mulher aumenta a probabilidade de participar do programa jovem aprendiz, e ser negro aumenta quanto em a probabilidade de participar do programa jovem aprendiz. Já a escolaridade, maior os anos de estudo a probabilidade de participar do jovem aprendiz.

Essa estimação de participação no programa Jovem Aprendiz estima os escores de propensão que indica a probabilidade de o jovem participar ou não do programa jovem aprendiz e, na seção seguinte tem-se os resultados após o pareamento entre tratado e não tratados para então, finalmente, avaliar o efeito do programa jovem aprendiz.

Além disso, é possível realizar uma nova análise de médias das variáveis com os escores de propensão, sendo possível verificar a hipótese de suporte comum. Na tabela 5 tem-se o tamanho do novo conjunto de dados após a verificação da hipótese de suporte comum.

Tabela 5- Quantidade de observação após verificação de suporte comum

Ano	2014	2015	2016	2017	2018
Número de observações original	8.117.679	7.282.092	6.596.171	6.457.432	6.396.299
Trat.	156.501	177.409	193.687	226.579	269.284
N° obs. pareadas Cont.	156.501	177.409	193.687	226.579	269.284
Total	313.002	354.818	387.374	453.158	538.568

Fonte: A autora com base no tratamento de dados da RAIS (2021).

Quando feita a análise nota-se redução na amostra. Pois, na verificação da área de suporte comum algumas observações que receberam tratamento encontravam-se com escore de propensão muito acima ou muito abaixo do escore de propensão de algumas observações que não receberam tratamento então, cria-se um novo grupo de controle que estão dentro da área em comum e os *outliers* são excluídos. E assim, forma-se uma nova amostra após o pareamento e assim se torna possível estimar o efeito do programa jovem aprendiz apresentado no item seguinte.

Com o balanceamento feito, se torna possível realizar o pareamento das observações e assim encontrar a média e desvio padrão das variáveis observadas separadamente entre tratados e grupo de controle, e podem ser verificados na Tabela 6.

Primeiramente, destaca-se que a quantidade média de jovens que participaram do programa jovem aprendiz no período analisado não chega a representar nem 4% da amostra.

Quanto ao gênero dos indivíduos, a proporção entre homens e mulheres no grupo que recebeu o tratamento é próxima ao observada na amostra sem tratamento, sendo composta por aproximadamente 51% por mulheres, enquanto que no grupo controle apenas aproximadamente 42 % são mulheres.

Ao observar a escolaridade do grupo que recebeu tratamento identifica-se que a média de escolaridade é o ensino médio incompleto para os dois grupos, assim como para a amostra completa.

Nesta análise descritiva das variáveis vale salientar que a composição desta amostra é coerente com método escolhido para análise de impacto, pois as características observáveis

dos dois grupos possuem estatísticas descritivas parecidas o que favorece um melhor pareamento entre os grupos, consequentemente, melhores resultados.

Tabela 6- Testes de diferenças de Médias coletadas por grupos

Variável		2014			2015			2016			2017			2018		
		Médias			Médias			Médias			Médias			Médias		
		Tratado	Controle	Dif.	Tratado	Controle	Dif.	Tratado	Controle	Dif.	Tratado	Controle	Dif.	Tratado	Controle	Dif.
Idade (em anos)	Não pareados	19,109	21,402	-2,29	19,13	21,495	-2,37	19,218	21,581	-2,36	19,356	21,446	-2,091	19,457	21,369	-1,91
	Pareados	19,109	19,112	0,00	19,13	19,135	-0,01	19,218	19,222	-0,004	19,356	19,3286	0,027	19,457	19,448	0,01
Mulher	Não pareados	0,5115	0,4244	0,09	0,51375	0,42881	0,08	0,50783	0,4252	0,0826	0,5456	0,429	0,116	0,5124	0,432	0,08
	Pareados	0,5115	0,5082	0,00	0,51375	0,51298	0,00	0,50783	0,50762	0,0002	0,5456	0,548	-0,002	0,5124	0,5109	0,00
Negro	Não pareados	0,6229	0,5234	0,10	0,62921	0,5352	0,09	0,61675	0,5467	0,070	0,632	0,5216	0,1104	0,6287	0,5243	0,10
	Pareados	0,6229	0,6181	0,00	0,62921	0,6317	0,00	0,61675	0,61921	0,002	0,632	0,631	0,001	0,6287	0,6289	0,00
Escolaridade	Não pareados	6,5401	6,6039	-0,06	6,709	6,984	0,275	6,6690	6,8754	0,2064	6,891	6,098	0,793	6,3479	6,1253	0,22
	Pareados	6,5401	6,5251	0,01	6,709	6,699	0,01	6,6690	6,670	0,001	6,891	6,874	-0,001	6,3479	6,3358	0,01

Fonte: A autora com base no tratamento de dados da RAIS feito no software R.

Observa-se que as diferenças médias das variáveis entre o grupo de tratamento e o grupo de controle, após o tratamento, são bem menores. Na tabela 6 é possível verificar as diferenças nas variáveis observadas entre o grupo de tratamento e o grupo de controle. Pode-se observar que nas variáveis observadas, após o pareamento não existe diferença estatisticamente significativa.

Realizado o procedimento de pareamento usando o escore de propensão e atestada a hipótese de suporte comum, pode-se considerar os grupos ajustados e então, estimar a diferença entre eles e verificar a efetividade do programa. Entretanto, para realizar o pareamento é necessário escolher o algoritmo que será usado para o pareamento. Na seção seguinte apresenta-se os resultados obtidos após o pareamento.

4.2 EFEITOS DO PROGRAMA JOVEM APRENDIZ SOBRE O RENDIMENTO

Independente do algoritmo de pareamento utilizado eles devem gerar resultados semelhantes, não há regra que especifique qual mais indicado para determinado conjunto de variáveis, entretanto é importante considerar o tamanho da amostra para a escolha.

Para a pesquisa o algoritmo de pareamento utilizado foi o algoritmo de “vizinho mais próximo”, tanto que na tabela 05, é possível verificar que a quantidade de observações de tratados é igual a quantidade de observações do grupo de controle, garantindo que jovem aprendiz foi comparado com um controle mais próximo com mais similaridade em termos de características observadas.

Portanto, após parear as observações que receberam tratamento com as observações de controle obtém-se as estimativas descritas na tabela 7 que identifica o efeito médio do tratamento para os tratados (ATT).

Tabela 7 - Efeito do Programa Jovem Aprendiz sobre o Rendimento

	2014	2015	2016	2017	2018
Estimativa Rendimento	-3,58	-0.97	-1.05	-1,85	-2,70
Erro padrão	0,12135	0,027307	0,020467	0,0134	0,03145
Teste T	-29,546	-35.386	--51.634	-31,445	-34,235
p,valor	< 2,22e-16	< 2.22e-16	< 2.22e-16	< 2.22e-16	< 2.22e-16

Fonte: A autora através de resultados obtidos através do tratamento de dados da RAIS obtidos pelo software R.

Após combinar indivíduos tratados e de controle, na tabela 7 tem-se a estimativa de rendimento que representa a diferença média entre os participantes do programa jovem aprendiz e os não participantes. Os valores negativos indicaram que os participantes do programa ganharam menos que os controles. O erro padrão indica a precisão da estimativa do rendimento, enquanto o teste T e p valor são os testes de significância estatística das estimativas, em que todos os p-valores são menores que $2,22e-16$, indicando que os resultados são estatisticamente significativos.

Em todos os anos analisados (2014 a 2018), as estimativas são negativas, variando de R\$ -0,97 a R\$-3,58. Isso sugere que a participação no programa "Jovem Aprendiz" está associada a uma redução nos ganhos dos participantes comparados aos indivíduos que não participaram do programa. Em outras palavras, os jovens que não participaram do programa jovem aprendiz, que possuem outro vínculo empregatício que não jovem aprendiz, ganham mais do que aqueles que participaram do programa.

Economicamente, a magnitude do efeito de tratamento é bastante significativa, de acordo com os valores de Teste T são todos bastante elevados em termos absolutos, e os p-valores são extremamente baixos, indicando que as estimativas são estatisticamente significativas. Isso reforça a confiabilidade dos resultados. Isso significa que, em média, os jovens que participaram do programa "Jovem Aprendiz" tiveram uma redução nos ganhos de aproximadamente R\$1 a R\$3,50 em comparação com os jovens que não participaram do programa.

Os resultados indicam que, durante o período analisado, o programa "Jovem Aprendiz" não teve um impacto positivo no rendimento financeiro dos jovens. Na verdade, os participantes do programa parecem ter rendimentos menores do que aqueles que não participaram, o que pode sugerir que o programa pode não estar alcançando seus objetivos de melhorar o rendimento financeiro dos jovens ou que os participantes enfrentam condições de trabalho que não compensam o treinamento recebido.

É relevante considerar essa análise considera apenas a comparação entre participantes e não participantes com base em características observadas. Outras variáveis não observadas podem influenciar os rendimentos, que devido ao tamanho da amostra não foram controladas neste estudo. Salienta-se ainda que isso aponta para a necessidade de uma revisão do programa para entender melhor os fatores que contribuem para esses resultados e para melhorar o impacto econômico do programa sobre os jovens.

5 CONSIDERAÇÕES FINAIS

A temática das políticas públicas abrange diversas áreas do conhecimento, sendo a avaliação dessas políticas por meio de métodos econométricos um campo de estudo relativamente pouco explorado, dado o vasto número de políticas existentes. Este estudo se concentra na avaliação do impacto do Programa Jovem Aprendiz, implementado no Brasil desde a década de 1940 com o objetivo de profissionalizar e formalizar jovens para o mercado de trabalho.

Neste trabalho avaliou-se o impacto do Programa Jovem Aprendiz, implementado no Brasil desde a década de 1940 na forma de Lei, com intuito de profissionalizar e formalizar jovens para o mercado de trabalho. Antes de apresentar os resultados, foi delineado um roteiro para a implementação e avaliação de políticas públicas. A análise se baseou nas características observáveis de dados fornecidos pela RAIS, combinada ao modelo de Pareamento por Escore de Propensão. Isso envolveu a distinção de dois grupos: jovens que participaram do programa e jovens que não participaram (grupo de controle), permitindo uma comparação significativa entre os dois.

As estatísticas revelam que, entre 2014 e 2018, o Programa Jovem Aprendiz não teve um efeito médio positivo no rendimento dos participantes. No entanto, afirmar que o programa não teve impacto é uma análise simplista, considerando a amplitude do programa. Os resultados não surpreendem quando se consideram os objetivos de proporcionar uma fonte de renda aos participantes, especialmente quando comparados a jovens com outros tipos de vínculos empregatícios.

A análise pode divergir das expectativas ao comparar jovens com vínculo formal de aprendizagem com aqueles sem vínculo empregatício formal. Ainda assim, não foi possível estimar os efeitos do programa na empregabilidade a longo prazo e nos salários dos participantes, devido à limitação dos dados da RAIS em seguir os trabalhadores ao longo dos anos.

Embora os resultados financeiros imediatos do programa "Jovem Aprendiz" possam mostrar uma redução nos rendimentos, os efeitos econômicos mais amplos e de longo prazo podem ser positivos. Ao considerar efeitos econômicos do programa "Jovem Aprendiz" os resultados podem ir além dos rendimentos imediatos. A participação no programa pode ter impactos significativos sobre a formação profissional, empregabilidade futura, produtividade, inovação, e desenvolvimento econômico local e nacional. Para uma avaliação completa, é

essencial considerar tanto os impactos diretos quanto os benefícios de longo prazo e realizar análises de custo-benefício que levem em conta todos esses fatores.

O efeito positivo na qualidade do emprego, por exemplo pode melhorar as habilidades e a qualificação dos jovens, aumentando seu potencial de ganhos futuros. Mesmo que o efeito imediato nos rendimentos seja negativo, o desenvolvimento de habilidades pode levar a melhores oportunidades e rendimentos ao longo da carreira. Além de ajudar os jovens a desenvolver habilidades interpessoais, como comunicação, trabalho em equipe e resolução de problemas, que são importantes no ambiente de trabalho e na vida cotidiana.

Outro ponto a ser levantado é a participação no programa pode aumentar a taxa de emprego entre os jovens, mesmo que os rendimentos sejam menores, pois o programa pode ajudar na construção de uma trajetória profissional, favorecer a experiência prática e a exposição a um ambiente profissional e assim aumentar a motivação dos jovens para continuar seus estudos e buscar qualificações adicionais, que podem abrir portas para futuras oportunidades de emprego ou até mesmo para estágios e formações adicionais.

O efeito sobre as empresas também pode ser destacado. Embora as empresas possam enfrentar custos associados ao treinamento e à integração dos jovens aprendizes, elas também podem se beneficiar ao ter funcionários mais engajados e leais, bem como ao preencher lacunas de habilidades com trabalhadores jovens e treinados com novas perspectivas e ideias para as empresas, o que pode fomentar a inovação e melhorar o ambiente de trabalho.

Nesse sentido, se o programa é eficaz em proporcionar oportunidades para jovens de baixa renda, pode ajudar a reduzir a desigualdade econômica ao oferecer uma plataforma para que esses jovens tenham acesso a melhores empregos e rendimentos no futuro. E, a longo prazo estimula o crescimento econômico com o aumento da qualificação da força de trabalho e assim, força de trabalho mais qualificada é geralmente mais produtiva.

É importante ressaltar que esses resultados se aplicam apenas à população pesquisada e não representam jovens fora do mercado de trabalho, Para uma compreensão mais profunda dos efeitos econômicos, seria útil realizar estudos adicionais que acompanhem os participantes ao longo do tempo, considerando não apenas o impacto imediato, mas também o desenvolvimento profissional e as oportunidades futuras que o programa pode ter proporcionado.

Os resultados deste estudo se restringem ao universo pesquisado, destacando que a RAIS é uma relação anual de informações sociais que abrange apenas dados daqueles que estão de alguma forma registrados. Portanto, os resultados não refletem a situação da

população de jovens na faixa etária estudada que está fora do mercado de trabalho. Seria benéfico, em estudos futuros, trabalhar com uma amostra que melhor representasse essa parcela da população, mantendo as mesmas características observáveis selecionadas.

Outra limitação a ser reconhecida nesta pesquisa reside no modelo econométrico de Pareamento por Escore de Propensão. Embora seja um modelo que visa minimizar o viés de seleção ao reduzir as variáveis para aquelas com características observáveis, ele deixa de considerar a endogeneidade das variáveis que não possuem características observáveis. Neste estudo, procurou-se mitigar esse viés ao empregar o modelo *logit* para estimar os escores de propensão, buscando obter o maior número possível de observações de controle para serem pareadas com o par ideal do grupo de tratados.

REFERÊNCIAS

- ARAÚJO, T. B. **Ensaio sobre o Desenvolvimento Brasileiro: heranças e urgências**. Rio de Janeiro: Revan; FASE. 2000.
- BECKER, S. O; ICHINO. A. Estimation of Average Treatment Effects Based on Propensity Scores: **The Stata Journal**. v. 2(4); p.1-19. 2002. Disponível em: < <https://acesse.one/TqQTi>>. Acesso em: 15 Abr. 2023.
- BRASIL. Constituição (1988). **Constituição da República Federativa do Brasil de 1988**. Brasília, DF: Presidência da República, [2016]. Disponível em: http://www.planalto.gov.br/ccivil_03/Constituicao/Constituicao.htm. Acesso em: 14 Ago. 2023.
- BRASIL. **Decreto nº 5.598/2005**. Regulamenta a Contratação de Aprendizizes. Brasília: Diário oficial da União, 2005. Disponível em < http://www.planalto.gov.br/ccivil_03/_ato2004-2006/2005/decreto/d5598.htm> Acesso em: 14 Ago. 2023.
- BRASIL. **Decreto Lei nº 8.621**, de 10 de janeiro de 1946. Dispõe sobre a criação do Serviço Nacional de Aprendizagem Comercial. Brasília/DF, 1946. Disponível em http://www.planalto.gov.br/ccivil_03/decreto-lei/1937-1946/Del8621.htm> Acesso em: 14 Ago. 2023.
- BRASIL. **Lei nº 10.748**, de 22 de outubro de 2003. Cria o Programa Nacional de Estímulo ao Primeiro Emprego para os Jovens – PNPE. Brasília/DF, 2003. Disponível em <http://www.planalto.gov.br/ccivil_03/leis/2003/L10.748impressao.htm> Acesso em: 15 Ago. 2023.
- BRASIL. **Lei nº 10.097**, de 19 de dezembro de 2000. Lei da Aprendizagem. Brasília/DF, 2000. Disponível em < https://www.planalto.gov.br/ccivil_03/leis/l10097.htm> Acesso em: 25 Mai. 2021.
- BRASIL. **Manual da aprendizagem: o que é preciso saber para contratar o jovem aprendiz** – 3. ed. – Brasília: MTE, SIT, SPPE, ASCOM, 2009. Disponível em < http://acesso.mte.gov.br/data/files/8A7C816A454D74C101459564521D7BED/manual_aprendizagem_miolo.pdf> Acesso em: 15 Ago. 2023.
- BRESSER PEREIRA, L. C. **Construindo o Estado republicano: Democracia e reforma da gestão pública**. Rio de Janeiro: FGV Editora, 2009.
- CALLIENDO, M; KOPEINIG, S. Some practical guidance for the implementation of propensity score matching. **Journal of Economic Surveys**. v. 22. p.31–72. 2005. Disponível em: < <https://11nk.dev/052TZ> >. Acesso em: 15 Abr. 2023.
- CAPELLA, A. C. N. Perspectivas Teóricas sobre o Processo de Formulação de Políticas Públicas s. In: HOCHMAN, G.; ARRETCHE, M. et al. (Orgs.). **Políticas Públicas s no Brasil**. Rio de Janeiro: FIOCRUZ. p. 87-121. 2007. Disponível em: < <https://11nk.dev/3L24C>> Acesso em: 20 Mai. 2023.
- COCHRAN, W. G.; RUBIN, D. B. Controlling Bias in Observational Studies: A Review. **Sankhya**. Series A. v. 35. n. 4. p. 417-446. 1973. Disponível em: < <https://11nk.dev/geSrt> >. Acesso em: 20 Mai. 2023.

DEHEJIA, R. H.; WAHBA, S. Propensity score matching methods for non- experimental causal studies. *Review of Economics and Statistics*. v. 84. p. 151–161. 2002. Disponível em: < <https://11nk.dev/bZuvN>>. Acesso em: 20 Mai. 2023.

FREY, K. Políticas públicas: um debate conceitual e reflexões referente à prática da análise de políticas públicas no Brasil. **Revista Planejamento e Políticas Públicas**. Brasília. n. 21. p. 211-259. jun. 2000. Disponível em: < <https://acesse.one/g57Gw>>. Acesso em: 20 Mai. 2023.

FIGUEIREDO, M. F.; FIGUEIREDO, A. M. C. **Avaliação Política e Avaliação de Políticas**: um quadro de referência teórica. São Paulo: IDESP. 1986.

GERTLER, P.; MARTINEZ, S.; PREMAND, P.; RAWLINGS, L. B.; VERMEERSCH, C. M. J. **Impact Evaluation in Practice**. Washington. DC: The World Bank. 2016.

HECKMAN, J. J; ICHIMURA, H; TODD, P. Matching As An Econometric Evaluation Estimator. **The Review of Economic Studies**. v.65. p. 261–294. Abril. 1998. Disponível em: < <https://doi.org/10.1111/1467-937X.00044>>. Acesso em: 23 Fev. 2023.

HEINRICH, C.; MAFFIOLI, A.; VÁZQUEZ, G. A Primer for Applying Propensity-Score Matching. *Impact Evaluation Guidelines. Technical Notes*. **Inter-American Development Bank**. v. 161. 2010. Disponível em: < <https://11nk.dev/mgCVu> >. Acesso em: 23 Fev. 2023.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATISTICA - IBGE. **Painel de Indicadores** (2021). Disponível em:< https://www.ibge.gov.br/estatisticas/sociais/trabalho/9173-pesquisa-nacional-por-amostra-de-domicilios-continua-trimestral.html?=&t=series-historicas&utm_source=landing&utm_medium=explica&utm_campaign=desemprego >. Acesso em: 16 Mar. 2022.

INTERNATIONAL LABOUR OFFICE. **Global Employment Trends for Youth 2020**: Technology and the future of jobs International Labour Office – Geneva: ILO. 2020. Disponível em: < <https://acesse.one/bfjJg>>. Acesso em: 16 Mar. 2022

INTERNATIONAL LABOUR OFFICE. **Global employment trends for youth 2013**: Technology and the future of jobs. International Labour Organisation (ILO). Geneva: OIT. 2013. Disponível em:< <https://acesse.one/pSle5>>. Acesso em: 16 Mar. 2022.

J-PAL. **About Us**. Disponível em: <<https://www.povertyactionlab.org/about-j-pal>> Acesso em: 13 Fev. 2022.

KHANDKER S.. B. KOOLWAL G.. SAMAD H. **Handbook on impact evaluation**: Quantitative methods and practices. Washington. DC: World Bank. 2010.

KUENZER, Acácia. Ensino médio e profissionalizante: as políticas do estado neoliberal. **Coleção questões da nossa época**, v. 63. São Paulo: Cortez, 2001

LASSWEL, H.D. **Politics**: Who Gets What. When. How. Cleveland. Meridian Books. 1958.

MENEZES FILHO, N.A.; PINTO, C. C. X.(org.). **Avaliação Econômica de Projetos Sociais**. 3. ed. São Paulo: Fundação Itaú Social. 2017.

MICHEL, M. H. **Metodologia e Pesquisa Científica**: um guia prático para acompanhamento da disciplina e elaboração de trabalhos monográficos. São Paulo: Atlas. 2005.

NOBEL PRIZE. **The Prize in Economic Sciences 2019**. Disponível em: <<https://www.nobelprize.org/prizes/economic-sciences/2019/summary/>> Acesso em: 10 Out. 2022.

PEDONE. Luiz. **Formulação, implementação e avaliação de políticas públicas**. Brasília: FUNCEP. 1986.

POCHMANN. M. Emprego e desemprego juvenil no Brasil dos anos 90. XI Encontro Nacional de Estudos Populacionais da ABEP. p. 2547-2563.

RIBAS. R. P. **Técnicas Econométricas para Avaliação de Impacto**. Aula: Análise de causalidade e Problema da Avaliação de Impacto. Brasília: International Poverty Centre/UNDP. (2008) Slides. 2008. Disponível em: <<http://www.ipc-undp.org/evaluation/programa.pdf>> Acesso em: 25 Mar. 2023.

RIBEIRO. R. A.; CACCIAMALI. M. C. Impactos do Programa Bolsa Família sobre os indicadores educacionais. **Economia**. Brasília. v. 13. p. 415-446. 2012.

ROSENBAUM P.R; RUBIN D.B. The bias due to incomplete Matching. **Biometrics**. v. 41. p. 103–116. 1985. Disponível em: <<https://l1nk.dev/Oar9T>>. Acesso em: 20 Mai. 2023.

SABATIER. P. A. (ed.). **Theories of the policy process**. Colorado: Westview Press. 2007. Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/4005368/>>. Acesso em: 20 Mai. 2023

SANTOS, A. L. dos; GIMENEZ, D. M. Inserção dos jovens no mercado de trabalho. ESTUDOS AVANÇADOS. 29 (85). Set-Dez 2015. Disponível em <<https://www.scielo.br/j/ea/a/GbtRWPdc95dJYQVnkN5cBNq/abstract/?lang=pt>>. Acesso em: 23 Mai. 2023.

SECCHI. L. **Políticas públicas: conceitos. Categorias de análise. casos práticos**. 2. ed. São Paulo: Cengage Learning. 2013.

SEN. A. **Desenvolvimento como Liberdade**. Tradução de Laura Teixeira Motta. São Paulo: Companhia das Letras. 2000.

SILVA. A. B. da; COSTA. A. V. Políticas e Programas Públicos que Oportunizam a Inserção do Jovem no Mercado de Trabalho. **Id on Line Rev.Mult. Psic.**. 2019. vol.13. N.43. p. 967-981. ISSN: 1981-1179. Disponível em <<https://idonline.emnuvens.com.br/id/article/view/1571>> Acesso em: 20 Mai. 2023.

SIMÕES, A. Avaliação de Políticas Públicas: Tipologias E Técnicas De Análise, 2018.

SOUZA. C. Políticas públicas: uma revisão da literatura. **Sociologias**. v.16. p. 20-45. 2006. Disponível em: <<https://acesse.one/gJgdU>>. Acesso em: 20 Mai. 2023.

TAVARES. I.; CÂNDIDO. A. F.; CARMO. R. M. do. Desemprego e Precariedade Laboral na População Jovem: Tendências Recentes em Portugal e na Europa. **Observatório das Desigualdades**. CIES-Iscte .Lisboa. (2021). Disponível em <<https://l1nk.dev/pZiIr>>. Acesso em: 16 Mai. 2023

TREVISAN. A. P.; Bellen. H. M. V. Avaliação de políticas públicas: uma revisão teórica de um campo em construção. **Revista de Administração Pública - RAP**. v. 42. p. 529–550. 2008. Disponível em <<https://l1nk.dev/9WRiW>>. Acesso em: 16 Mar. 2023.