LANANDA

DETERMINANTES DO DESEMPREGO NO BRASIL:

UMA ANÁLISE DE MICRODADOS COM BASE NO MODELO *LOGIT*

2019

Resumo

Este trabalho analisa a probabilidade do empregado, atuando no mercado de trabalho formal, perder o emprego nas cinco regiões brasileiras no ano de 2017. A análise consiste em investigar o papel das características dos indivíduos em influenciar a probabilidade de que ele perca o emprego no referido ano. Tecnicamente, essas características são as variáveis explicativas, que, por sua vez, são escolhidas a partir da literatura pertinente: idade, gênero, escolaridade e cor ou raça. Além dessas características, neste trabalho inclui-se uma variável que descreve a ocupação, de modo a investigar se o fato do trabalhador atuar numa ocupação pertencente às áreas de *Science, Technology, Engineering e Mathematics* (STEM), diminui suas chances de ficar desempregado. Assim, o enfoque principal do trabalho é verificar se indivíduos da força de trabalho STEM possuem menor probabilidade de perderem seus empregos, a partir da literatura revisada. O acrônimo STEM, tem sido objeto de pesquisas que buscam entender o novo comportamento do mercado de trabalho, com importância cada vez maior das ocupações que requerem algum *expertise* técnico ou científico. A metodologia utilizada neste trabalho baseia-se num modelo *logit* de resposta qualitativa, que irá resultar em como cada variável explanatória tem efeito na probabilidade de o indivíduo ficar desempregado. A análise se deu inicialmente de abrangência nacional e, posteriormente, de forma regional, devido às heterogeneidades das regiões brasileiras e para fins de comparação. A base de dados advém da Relação Anual de Informações Sociais (RAIS) 2017, esta, provém do Ministério do Trabalho (MTE). Os resultados para o Brasil mostraram que todas as variáveis do modelo foram significativas para explicar a probabilidade do indivíduo ficar sem emprego, sendo que os indivíduos que atuam em postos de trabalho caracterizado como STEM, possuem 1,5 pontos percentuais a menos de ficarem desempregados, em relação aos demais indivíduos. Ao analisar as regiões brasileiras, detectou-se a presença de heterogeneidade, sendo que as regiões Sul e Sudeste apresentaram resultados similares.

Palavras-chaves: Desemprego. Determinantes Individuais. Modelo *Logit*. STEM.

Abstract

This paper analyzes the probability of employee, working in the formal labor market, losing his job across the five Brazilian regions for the year 2017. The analysis consists of investigating the role of certain individual characteristics, in determining the probability that the individual will lose his job. Tecnically, these characteristics are the explanatory variables which, in turn, are chosen from the relevant literature: age, gender, schooling and color or race. Besides these characteristics, we include an occupation variable to capture the effect of being a STEM (Science, Technology, Engineering and Mathematics) worker on the probability of becoming unemployed. Thus, the main focus of the paper is to verify whether individuals in the STEM workforce are less likely to lose their jobs, according to the revised literature. The acronym STEM has been the subject of research that seeks to understand the new behavior of the labor market with increasing importance of occupations that require some technical or scientific expertise. The methodology used in this paper is based on a qualitative response logit model, which will result in how each explanatory variable has effect on the probability of the individual becoming unemployed. The analysis was initially based on national coverage and later on regionally, due to the heterogeneities of the Brazilian regions and for comparison purposes. The database comes of the Social Information Annual Report (RAIS) 2017, which comes from the Ministry from Labor (MTE). The results for Brazil showed that all model variables were significant in explaining the likelihood of being out of work, and individuals working in jobs characterized as STEM have 1.5 percentage points less than being unemployed, in relation to the other individuals. By analyzing the Brazilian regions, the presence of heterogeneity was detected, and the South and Southeast regions presented similar results.

Keywords: Unemployment. Individual determinants. Logit Model. STEM.

Lista de ilustrações

Figura 1 – Equilíbrio no Mercado de Trabalho 13
Figura 2 – Representação IS-LM e Demanda Efetiva por Trabalho 16
Figura 3 – Organograma da Representação do Mercado de Trabalho 19
Figura 4 – Número de Empregos Formais no Brasil 36
Figura 5 – Número de Empregos Formais por Região no período 2011-2017 37
Figura 6 – Vínculos Empregatício segundo Gênero em 2017 38
Figura 7 – Vínculos Empregatício segundo Idade em 2017 39
Figura 8 – Vínculos Empregatício segundo Faixas de Escolaridade em 2017 40
Figura 9 – Vínculos Empregatício segundo Raça/Cor em 2017 41
Figura 10 – Vínculos Empregatício segundo Ocupações em 2017 42

Lista de tabelas

Tabela 1 – Explanação das variáveis 35
Tabela 2 – Resultado do Modelo *Logit* para o Brasil em 2017 45
Tabela 3 – Efeitos Marginais do Modelo *Logit* para as Regiões do Brasil em 2017 . . . 48
Tabela 4 – Códigos CBO considerados no grupo STEM 55
Tabela 5 – Resultado do Modelo *Logit* para a Região Norte em 2017 56
Tabela 6 – Resultado do Modelo *Logit* para a Região Nordeste em 2017 57
Tabela 7 – Resultado do Modelo *Logit* para a Região Sudeste em 2017 58
Tabela 8 – Resultado do Modelo *Logit* para a Região Sul em 2017 59
Tabela 9 – Resultado do Modelo *Logit* para a Região Centro Oeste em 2017 60
Tabela 10 – Teste de Multicolinearidade para os dados do Brasil 61
Tabela 11 – Teste de Multicolinearidade para os dados da Região Norte 61
Tabela 12 – Teste de Multicolinearidade para os dados da Região Nordeste 62
Tabela 13 – Teste de Multicolinearidade para os dados da Região Sudeste 62
Tabela 14 – Teste de Multicolinearidade para os dados da Região Sul 63
Tabela 15 – Teste de Multicolinearidade para os dados da Região Centro Oeste 63

Sumário

1 INTRODUÇÃO . . . 8

2 REVISÃO BIBLIOGRÁFICA . . . 10
2.1 MERCADO DE TRABALHO . . . 10
2.1.1 Modelo Clássico . . . 10
2.1.2 Modelo Keynesiano . . . 13
2.2 DESEMPREGO: ASPECTOS CONCEITUAIS . . . 16
2.2.1 Pleno Emprego e Taxa Natural do Desemprego . . . 17
2.2.2 Força de Trabalho . . . 18
2.2.3 Taxa de Ocupação . . . 18
2.2.4 Taxa de Desocupação . . . 18
2.2.5 Metodologias de Cálculo . . . 19
2.2.6 Relações do Desemprego com Variáveis Macroeconômicas . . . 20
2.3 PRINCIPAIS TIPOS DE DESEMPREGO . . . 21
2.3.1 Desemprego Friccional . . . 21
2.3.2 Desemprego Estrutural . . . 22
2.3.3 Desemprego Cíclico . . . 22
2.4 ABORDAGEM MICROECONÔMICA . . . 22
2.5 MERCADO DE TRABALHO - CONCEPÇÃO MODERNA . . . 25
2.5.1 Modelo DMP para Desemprego . . . 25
2.5.2 Força de Trabalho nas Áreas de Ciência e Tecnologia: STEM . . . 26

3 METODOLOGIA E DESCRIÇÃO DOS DADOS . . . 28
3.1 CARACTERIZAÇÃO DA PESQUISA . . . 28
3.2 MODELOS DE RESPOSTA QUALITATIVA . . . 28
3.3 MODELO LOGIT . . . 29
3.3.1 Medidas de Avaliação do Modelo . . . 32
3.4 BASE DE DADOS . . . 34
3.5 ANÁLISE DESCRITIVA . . . 35
3.5.1 Estoque do Número de Empregos Formais . . . 35
3.5.2 Estatísticas Descritivas para o Brasil em 2017 . . . 38

4 RESULTADOS ESTATÍSTICOS . . . 43
4.1 ANÁLISE DO MODELO *LOGIT* PARA O BRASIL . . . 43
4.2 ANÁLISE DO MODELO PARA AS REGIÕES BRASILEIRAS . . . 45

5 CONSIDERAÇÕES FINAIS 49

Referências 51

APÊNDICE A – OCUPAÇÕES STEM DE ACORDO COM A CBO . 55

APÊNDICE B – RESULTADO DAS REGRESSÕES 56

APÊNDICE C – TESTES 61

1 INTRODUÇÃO

O mercado de trabalho vem sofrendo grandes mudanças, causadas principalmente pela evolução da globalização e dos avanços tecnológicos. Como consequência, a reestruturação produtiva afeta diretamente o descompasso entre o emprego e o trabalhador. Isso têm influência direta sobre o desemprego, que é uma variável de importância tanto macroeconômica quanto microeconômica.

O avanço da globalização vem revolucionando o mercado de trabalho ao longo dos tempos, principalmente nas duas últimas décadas, em que ocorreu a disseminação de computadores, *smartphones*, *tablets* e diversas outras tecnologias, que aumentam a produtividade em vários setores da atividade econômica. Além disso, as mudanças tecnológicas eliminam alguns tipos de qualificações e criam demandas por novas habilidades e qualificações, o que explica parte do desemprego estrutural das economias modernas. Essa modificação fez com que aumentasse o desemprego de pessoas em determinadas áreas intensificaram a utilização de equipamentos tecnológicos, por isso, esta variável macroeconômica é objeto de estudo de várias pesquisas.

Em termos macroeconômicos, a abordagem teórica sobre o mercado de trabalho ganhou importância a partir da grande depressão americana, iniciada em 1929, que elevou a taxa de desemprego dos Estados Unidos a 25% em 1933. Nesse contexto, surgiu a motivação para se descrever como se davam as relações entre as variáveis agregadas, dentre elas, o desemprego, emprego e produto.

A obra que se sobressaiu abordando o desemprego como variável de destaque foi A Teoria Geral do Emprego, do Juro e da Moeda de Keynes. A partir daí, o pensamento macroeconômico acerca do emprego/desemprego se dividiu em termos de duas teorias principais: clássica e keynesiana. Cada uma destas pressupõe um modelo distinto para analisar o mercado de trabalho. A escola clássica pressupõe que o equilíbrio no mercado de trabalho se dá através das relações de oferta e demanda por trabalho, em um modelo que só existe desemprego friccional e voluntário. Já a escola keynesiana apresenta um modelo em que o mercado de trabalho não está constantemente em equilíbrio de pleno emprego, e propõe o conceito de desemprego involuntário, contrapondo a visão classicista.

Após realizar o estudo da variável de interesse dentro do campo macroeconômico, será realizada uma revisão sobre os fatores que podem influenciar na probabilidade de desemprego, através do enfoque microeconômico. A abordagem microeconômica tem como base examinar o papel das características dos indivíduos e firmas no mercado de trabalho. Assim, conforme detalha, Kon (2012), os resultados encontrados a partir das investigações microeconômicas afetam as variáveis macroeconômicas, havendo desta forma, uma forte relação entre ambas abordagens.

O objetivo deste trabalho se insere na vertente microeconômica da análise do desemprego e consiste em verificar como determinadas características individuais influenciam a probabilidade de um trabalhador brasileiro, que esteja no mercado formal de trabalho, perder o emprego. Os objetivos específicos seguem duas linhas de abordagem. Umas delas consiste em uma análise ocupacional e, a outra, numa análise com foco nas diferenças regionais. A investigação ocupacional é dada pela verificação da diferença de probabilidade entre as ocupações que compõem o acrônimo *Science, Technology, Engineering e Mathematics* (STEM) e Não-STEM. Já em termos regionais, a proposta desse trabalho é examinar em que medida a probabilidade de ficar desempregado é afetada pelo local onde reside o trabalhador, considerando cada uma das cinco regiões brasileiras, devido às suas heterogeneidades.

A justificativa para a execução deste trabalho é dada pela importância tanto macro quanto microeconômica da variável desemprego. A análise acerca dos dados de desemprego pode ser alvo de políticas públicas, principalmente na questão regional, já que cada região comporta características distintas, estas deveriam receber políticas públicas diferenciadas. Além disso, devido à corrida tecnológica que a economia moderna está vivenciando, é de suma importância investigar as diferenças de estabilidade no emprego entre indivíduos que atuam nas áreas de ciência e tecnologia - STEM - e indivíduos fora deste grupo.

Esta pesquisa tem como influência o trabalho de Bachman (2017), que analisou os determinantes da probabilidade de um trabalhador ficar desempregado no estado do Paraná nos anos de 2001 e 2014. O estudo deste trabalho se dará com base nos dados da Relação Anual de Informações Sociais (RAIS) para o ano de 2017 para as regiões brasileiras, sendo que a análise estatística será realizada por meio do modelo de resposta qualitativa *logit*. Além de discorrer sobre a probabilidade do desemprego a partir de determinadas características dos indivíduos, este trabalho investiga ainda se a força de trabalho STEM exibe características semelhantes ao que tem sido relatado na literatura aplicada em economias avançadas. Tem sido observado, nessas economias, que pessoas que estão trabalhando dentro da área STEM possuem uma menor probabilidade de desemprego em relação aos da área Não-STEM. Para a economia brasileira, não existem pesquisas publicadas sobre o grupo de trabalhadores STEM e sua vulnerabilidade ao desemprego, portanto, esta deve ser a contribuição deste trabalho.

Para isso, este estudo está dividido em cinco capítulos, sendo este o primeiro. No segundo capítulo será tratado a respeito da revisão bibliográfica acerca do desemprego. O terceiro e quarto capítulo abordarão, respectivamente, a metodologia e base de dados utilizada, bem como a discussão dos resultados encontrados. No capítulo cinco, são consumadas as conclusões desta pesquisa.

2 REVISÃO BIBLIOGRÁFICA

Neste capítulo, apresenta-se primeiro o modelo teórico sobre determinação de salário e nível de emprego, e em seguida os conceitos relacionados ao mercado de trabalho, num aspecto macroeconômico. Posteriormente, apresenta-se uma abordagem microeconômica, acerca dos fatores que podem influenciar na probabilidade de perder o emprego. Por final, será feita uma análise evolutiva do mercado de trabalho, com foco na mão de obra que atua nas áreas de aplicação de ciência e tecnologia - STEM - e a exposição de uma nova modelagem macroeconômica que premiou três economistas - Diamond, Mortensen e Pissarides - com um Nobel em 2010, estes desenvolveram um modelo para explicar a "falha" de comunicação existente entre a oferta e demanda de emprego em diferentes mercados.

2.1 MERCADO DE TRABALHO

2.1.1 Modelo Clássico

Os modelos macroeconômicos normalmente surgem para contrapor algum outro modelo existente. E assim, surgiu a revolução clássica, para se opor a doutrina mercantilista. Os mercantilistas tinham como pressupostos o metalismo - a riqueza de uma nação era determinada pelo estoque de metal - e a necessidade de intervenção estatal. De maneira oposta, os clássicos defendiam o livre mercado e que, além disso, a riqueza de uma nação deveria ser determinada por variáveis reais, como a produção e o emprego.

A produção é medida através de uma função que relaciona os insumos capital e trabalho. Para os clássicos, no curto prazo o estoque de capital é mantido como constante, tendo variação somente a quantidade de mão de obra. É considerado que todos os agentes estão perfeitamente informados sobre os preços e que o mercado de trabalho sempre está ou tende ao equilíbrio. Assim, no modelo clássico, existe a ideia de que a economia está continuamente em pleno emprego - caso haja flexibilidade salarial -, não havendo desperdício de trabalho, inexistindo a situação de desemprego, conforme afirma Pigou (1968), de modo que o mercado de trabalho está sempre em equilíbrio.

De acordo com a função de produção agregada, dados o estoque de capital e tecnologia disponível, é a quantidade do fator trabalho que determinará o nível ótimo de produto. Assim, os clássicos dão um peso importante para este fator, havendo a necessidade de algumas hipóteses para a construção do modelo. As principais hipóteses são:

- Informação perfeita sobre os preços nos mercados de fatores e de produto;
- Perfeita flexibilidade de preços e salários;

- Homogeneidade dos trabalhadores;

- Mercados são concorrentes, logo as firmas são tomadoras de preço.

A demanda por mão de obra deriva da maximização de lucro das firmas em concorrência total nos mercados de produto e de fatores, em que o trabalho é homogêneo, sendo a função lucro da firma dada pela Equação 2.1, em que o subscrito "i" representa uma firma. O primeiro termo a direita da Equação 2.1 representa a receita de uma firma representativa i e o segundo termo, representa seu custo. Dada a hipótese de concorrência nos mercados de produto e de fatores, as firmas são tomadoras de preços. Assim, a função lucro de uma firma representativa i pode ser escrita:

$$\Pi_i = pY_i - (wN_i + r\bar{K}_i) \tag{2.1}$$

Em que:

– Π_i: lucro da firma i;

– p: preço do produto ofertado pela firma i;

– Y_i: produto de uma firma i;

– w: remuneração do fator trabalho - salário;

– N_i: número de horas de trabalho contratadas pela firma i;

– r: remuneração do fator capital - taxa de juros;

– K_i: fator capital, constante no curto prazo.

Derivando a função lucro em relação ao fator trabalho, e partindo do pressuposto de concorrência perfeita - onde preço se iguala a receita marginal que se iguala ao custo marginal - obtém-se a relação de demanda por mão de obra. A partir da derivação, aufere-se que a condição de maximização[1] do lucro da firma é dada a partir da igualdade de que o produto marginal do trabalho - que é medido em unidades de mercadoria - é equivalente ao salário real do trabalhador. Assim, a demanda por mão de obra da firma, que é traçada contra o salário real, corresponde ao produto marginal do trabalho.

Assim, o equilíbrio no mercado de trabalho está de acordo com a determinação da produção, que, a partir da visão clássica, parte do pressuposto da Lei de Say (1803), em que a oferta cria sua própria procura, assim, a oferta de trabalho será igual a demanda por trabalho e as forças de mercado o conduziriam ao pleno emprego. Isto posto, o modelo classicista não

[1]Essa condição de maximização de lucro da firma i é sujeita a restrição Y=F($\bar{K}$,N).

considera o chamado desemprego involuntário, somente a forma de desemprego transacional, ou seja, os desempregos do tipo friccional e voluntário, estes, são dados de forma voluntária.

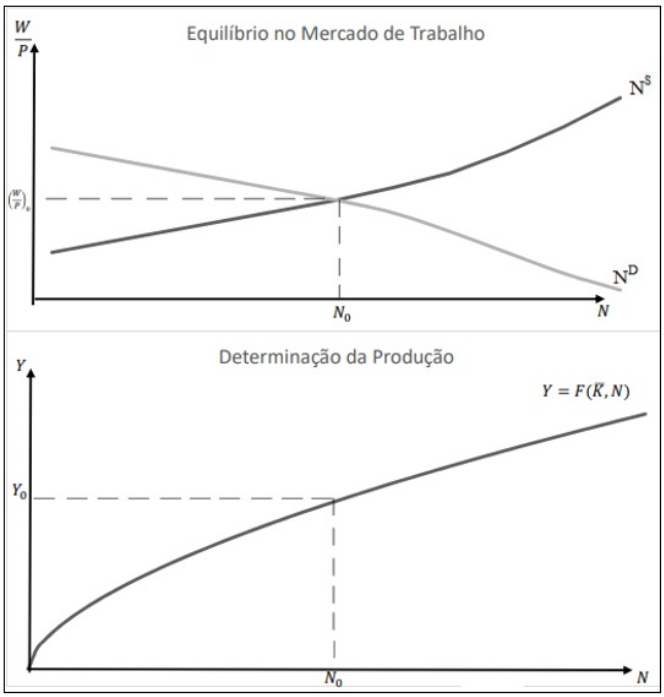
A oferta de mão de obra resulta do *trade-off* entre trabalho e lazer[2]. Supõe-se que os indivíduos podem escolher livremente o quanto de tempo alocar entre lazer e trabalho, escolhendo, através da otimização, a melhor alocação entre esses dois objetivos. De acordo com a teoria econômica, a renda aumentará com mais trabalho, mas, por consequência, diminuirá o tempo a ser gasto em lazer, portanto, a inclinação da curva de oferta por trabalho será positiva.

A perfeita simetria entre as firmas permite que a agregação seja feita em termos de uma firma representativa e, nesse caso, o equilíbrio entre produto e emprego através do mercado de trabalho - oferta e demanda por mão de obra -, pode ser visto na Figura 1. As variáveis explicitadas na Figura 1 são:

- $Y = F(\bar{K}, N)$: função de produção agregada, onde Y é a produção real;
- K: estoque de capital, fixo no curto prazo;
- N: quantidade de mão de obra;
- N^D: quantidade demandada de mão de obra;
- N^S: quantidade ofertada de mão de obra;
- $\frac{W}{P}$: salário real;

[2]No modelo clássico, lazer é explicado como tudo aquilo que não é contabilizado como trabalho.

Figura 1 – Equilíbrio no Mercado de Trabalho



Fonte – Elaboração própria, 2019.

Deste modo, o equilíbrio no mercado de trabalho no modelo clássico se dará entre as curvas de oferta e demanda por mão de obra, significando que quem quer trabalhar ao salário vigente - de equilíbrio - está empregado.

2.1.2 Modelo Keynesiano

A visão keynesiana surge então para contrapor o modelo clássico, pelo fato desta última desconsiderar o chamado desemprego involuntário. Para Keynes (1982), o desemprego não poderia ser atribuído à falta de flexibilidade nos mercados, mas sim, pelo nível de atividade. Assim, neste modelo, o emprego que depende do produto, e não o contrário, e, a partir desta concepção, ocorre o resgate do Princípio da Demanda Efetiva - composta por gastos com investimento e consumo. Como a demanda agregada determina o nível de atividade econômica, ela influenciará e determinará o nível de emprego. Em vista disso, o desemprego involuntário - que pode ser chamado como desemprego cíclico - surge por flutuações da demanda agregada.

Ao contrário dos clássicos, os keynesianos contestam a hipótese de perfeita flexibilidade de salários e preços, proferindo que salários e preços são rígidos e não se ajustam de forma instantânea, implicando que a economia pode ficar distante do seu equilíbrio geral por indeterminados períodos. A partir dessa discussão, sobre o grau de flexibilidade dos preços, e principalmente dos salários reais, é que se iniciou o debate para explicar a existência do desemprego involuntário.

A fundamentação keynesiana para o desemprego foi desenvolvida na década de 1980, a partir dos modelos de rigidez de salários reais - salários de eficiência e teorias de barganha.

Akerlof (1982) expõe uma abordagem moderna sobre o mercado de trabalho, explicando o desemprego involuntário em termos da resposta das empresas ao nível de esforço dos seus trabalhadores . Para que um salário real rígido seja a causa do desemprego, o salário real que as empresas estão pagando aos seus funcionários deve ser mais alto que o salário real de equilíbrio de mercado, no qual demanda e oferta por trabalho se igualam.

As explicações do motivo pelo o qual os salários reais podem ser rígidos são: a existência de um salário mínimo imposto pela legislação trabalhista; o custo de rotatividade, referente ao treinamento de um novo funcionário, assim, se a empresa pagar um salário real mais alto ao seu funcionário já estabelecido, ela poderá mantê-lo por mais tempo na empresa, sem ter que incorrer no treinamento de um novo trabalhador, caso aquele estabelecido saia do emprego; salário de eficiência, representa a ideia de que a produtividade de um trabalhador dependerá positivamente do salário real recebido, assim, trabalhadores que ganham salários reais maiores tendem a ser mais produtivos.

A hipótese do salário de eficiência é uma das principais utilizadas para analisar o mercado de trabalho do ponto de vista keynesiano. A hipótese sanciona que os serviços de um trabalhador é função do salário que ele recebe; e quanto mais bem pago é este trabalhador, maior sua produtividade, chegando ao momento em que um trabalhador pode fazer a função de dois trabalhadores. Além do salário de eficiência, a curva de demanda efetiva de trabalho é de extrema importância para este modelo, indicando a quantidade de trabalho que a firma necessita para atender a demanda do seu produto, mantendo constante a produtividade, o estoque de capital e o esforço. Há, portanto, duas principais diferenças entre este modelo keynesiano e o modelo clássico. A primeira delas é que o pleno emprego da mão de obra é determinado pela intersecção das curvas de demanda efetiva de trabalho com a do salário de eficiência; no modelo clássico, a partir da Figura 1, têm-se que essa intersecção se dá em um nível de salário real acima daquele de equilíbrio, correspondente ao equilíbrio do modelo clássico. A segunda diferença é que, alterações na oferta de trabalho na modelagem de salário de eficiência não afetará o emprego, e consequentemente, não afeta o nível de pleno emprego; no modelo clássico, alterações na oferta de mão de obra afetam o pleno emprego.

O modelo proposto por Keynes (1982) em sua obra A Teoria Geral do Emprego, do Juro e da Moeda, foi sumarizado e formulado matematicamente por Hicks e Hansen (1937) e ficou comumente conhecido como modelo IS-LM (*Investment Saving - Liquidity Money Supply*), que não modelou o mercado de trabalho e fatores de produção. De fato o modelo IS-LM inicialmente, levou ao extremo o princípio da demanda efetiva proposto por Keynes, simplificando o lado da oferta da economia com a forte hipótese de preços fixos.

Mais tarde, em resposta às críticas dos economistas novo-clássicos, a abordagem keynesiana seguiu em direção a fornecer fundamentos microeconômicos para o lado da oferta da economia, a partir dos modelos de rigidez nominal e rigidez real. Neste último caso específico, uma das principais formulações é a que segue as abordagens de salários de eficiência. Um

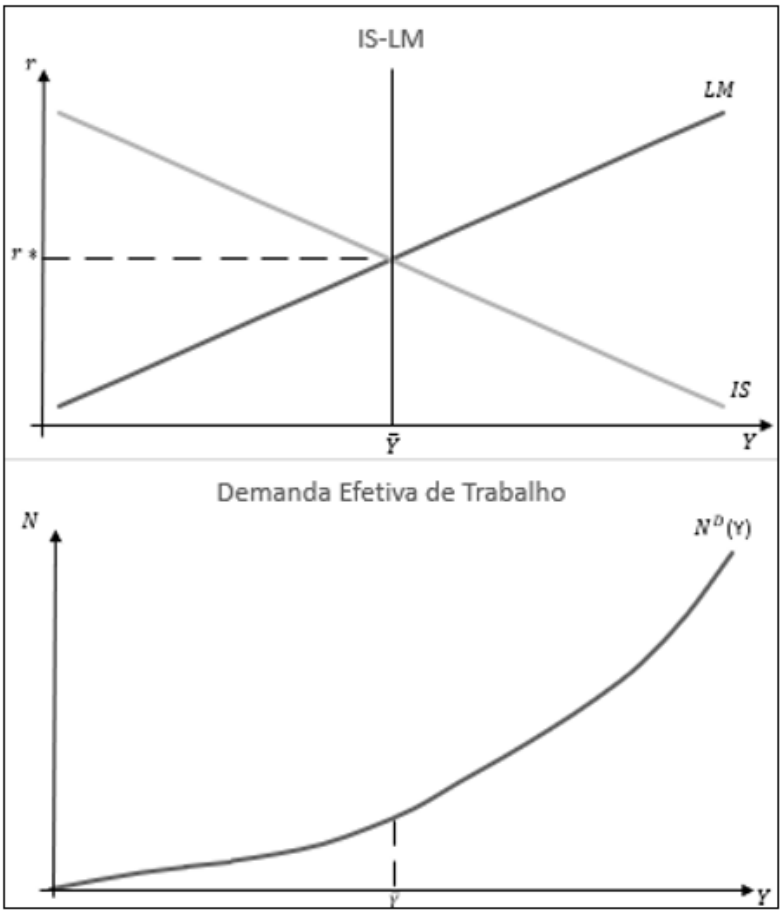
esquema para descrever a economia, sob essa perspectiva keynesiana, é apresentado a seguir, onde a noção de salário de eficiência gera uma curva de demanda efetiva por mão de obra por parte das firmas.

A curva IS representa o mercado de bens e serviços, que advém da igualdade de investimento e poupança. Assim, a demanda por bens é uma função crescente do produto produzido e uma função negativa da taxa de juros[3], logo, a curva IS é negativamente inclinada. A curva LM representa o mercado monetário e de títulos, assim, a oferta de moeda responde positivamente - em mesma direção - a alterações tanto da taxa de juros quanto do produto. A representação do modelo IS-LM em conjunto com a Demanda Efetiva de Trabalho está na Figura 2 - a linha $\bar{Y}$ representa o pleno emprego. Na Figura 2, as variáveis são designadas por:

- Y: produto;
- N: trabalho;
- r: taxa de juros real;
- N^D: quantidade demandada efetiva de trabalho;

[3]Quanto maior a renda de uma economia, maior será a demanda por bens e serviços; quanto maior a taxa de juros de uma economia, menor o nível de investimentos, e, assim, menor a demanda por bens e serviços.

Figura 2 – Representação IS-LM e Demanda Efetiva por Trabalho



Fonte – Elaboração própria, 2019.

A demanda efetiva de trabalho é determinada pela demanda da empresa por seu produto. Existe uma relação positiva entre a quantidade de produto demandada pela empresa e a quantidade de trabalho empregada, relação esta, representada pela curva de demanda efetiva de trabalho. O fator de remuneração do trabalho é o salário, quando o salário de eficiência for mais alto que o salário real de equilíbrio, mas ainda existem trabalhadores desempregados desejando trabalhar ao salário real de equilíbrio, ocorre o caso de desemprego involuntário, assim, as firmas conseguirão alterar o nível de emprego de acordo com a demanda pelo bem produzido por ela, sem alterar o salário.

2.2 DESEMPREGO: ASPECTOS CONCEITUAIS

O desemprego, de acordo, por exemplo, com Dornbusch, Fischer e Startz (2013), é definido como a parcela da força de trabalho em idade ativa que não está empregada, mas que esteve procurando emprego durante as quatro semanas anteriores à pesquisa. Outra explicação de desemprego é dada por Zylberstajn e Balbinotto (1999, p. 131) que definem o desemprego como "[...] uma situação na qual o indivíduo deseja estar empregado a uma dada taxa de salário

prevalecente no mercado de trabalho, mas não encontra emprego". Em contraponto, o emprego é definido como todo indivíduo que na semana de referência da pesquisa executou qualquer tipo de trabalho. Deste modo, a força de trabalho é definida como todos aqueles que estão trabalhando ou procurando trabalho, ou seja, a soma de empregados e desempregados. Portanto a taxa de desemprego, taxa de desocupação ou desemprego aberto expressa a parcela da população desempregada em relação à força de trabalho.

No Brasil, de acordo com o Instituto Brasileiro de Geografia e Estatística (IBGE), uma pessoa economicamente ativa é considerada desempregada quando esteve procurando emprego durante as quatro semanas anteriores ao cálculo da taxa de desemprego. A Relação Anual de Informações Sociais (RAIS), possui a variável de "vínculo empregatício", que contabiliza as relações de emprego do mercado de trabalho formal brasileiro, advindas de trabalho remunerado. O vínculo empregatício, de acordo com a RAIS, é definido como um trabalhador contratado por um CNPJ/CEI vinculado, sendo a situação do indivíduo declarada como ativo (empregado) ou inativo (desempregado) em 31/12 - empregado ou desempregado significa que o indivíduo possui vínculo ou não com o estabelecimento que fez a declaração em 31/12. Dado esse embasamento geral e característico - definição da RAIS - de concepções sobre desemprego, esta seção fornecerá alguns elementos necessários para o melhor entendimento do termo. Esta seção está dividida em seis subseções, estas são relacionadas a definição de conceitos, metodologias utilizadas para o cálculo e, por fim, uma subseção sobre as relações do desemprego.

2.2.1 Pleno Emprego e Taxa Natural do Desemprego

O pleno emprego é definido pela situação no qual todos aqueles que estão dispostos a trabalhar conseguem emprego em um curto período de tempo. O pleno emprego pode ser analisado em termos de durabilidade e da frequência do desemprego. A duração depende da estrutura do mercado de trabalho e está vinculada a fatores cíclicos. Já a frequência, é a média do número de vezes que os trabalhadores ficam desempregados, em determinado período.

Alguns economistas avaliam o desemprego como uma forma de controlar a inflação, a partir do conceito de *Non-Accelerating Inflation Rate of Unemployment* (NAIRU), que traduzindo significa taxa de inflação não aceleradora do desemprego, que associa a taxa de desemprego com inflação estável. Entre estes, Friedman (1968) e Phelps (1970), em que descreveram a NAIRU como sendo a taxa natural de desemprego. O conceito de taxa natural de desemprego é aquele em que a economia tende, no longo prazo, mas que não é imutável, dependendo de fatores estruturais e institucionais. Após se alcançar o pleno emprego, o restante de pessoas desempregadas é calculado como sendo a taxa natural da economia. Segundo Friedman (1968), um nível mais baixo de desemprego - do que a taxa natural - indica que há uma demanda excessiva de trabalho por parte das firmas que pressionará o aumento dos salários reais. Essa explicação é proposta por Friedman ao abordar a relação entre salário e desemprego.

2.2.2 Força de Trabalho

No grupo de estatísticas de desemprego, há algumas variáveis importantes para se chegar na variável de interesse. O conjunto amplo de dados do mercado de trabalho é encabeçado pela População em Idade Ativa (PIA), que significa o número de pessoas que estão aptas para trabalhar. As pessoas que estão disponíveis a trabalhar, mas que não estão trabalhando, fazem parte do grupo População Economicamente Ativa (PEA) ou força de trabalho, que inclui pessoas que ou estão trabalhando ou procurando emprego. Assim, dentro da força de trabalho existem as pessoas empregadas e as desempregadas, estas últimas, são definidas como aquelas que estão sem emprego e, além disso, estão à procura de um trabalho. Desse modo, a taxa de desemprego é representada pela razão entre o número de desempregados e a população economicamente ativa.

As pessoas que estão em idade ativa, mas não estão trabalhando ou procurando emprego, são classificadas como população não economicamente ativa. Segundo Dornbusch, Fischer e Startz (2013), esta inclui aposentados, donas de casa e estudantes em tempo integral. Além destas, também inclui aqueles considerados desalentados, pessoas que gostariam de trabalhar mas que desistiram de procurar emprego. A relação entre o número de pessoas economicamente ativas e o número de pessoas em idade ativas, em determinado período de referência, é chamado como taxa de atividade.

2.2.3 Taxa de Ocupação

De acordo com o IBGE (2019), a população ocupada é conceituada pelas pessoas que em determinado período trabalharam ou tinham trabalho mas estavam de férias. A taxa de ocupação mede - em percentual - o número de pessoas ocupadas em relação à força de trabalho. As pessoas ocupadas são classificadas em:

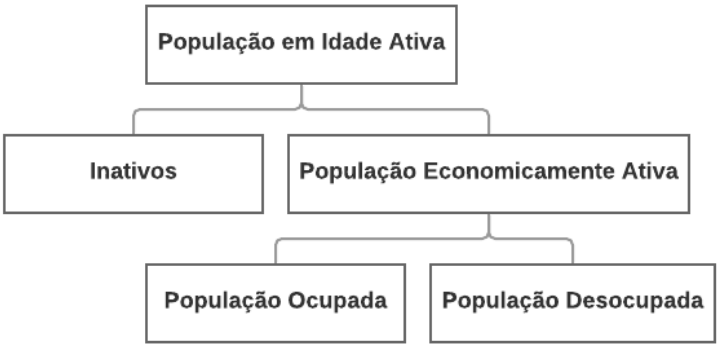
- Empregados: pessoas que trabalham para um empregador ou mais e que recebem, em contrapartida, uma remuneração por esse trabalho fornecido;
- Conta Própria: pessoas que exercem uma profissão mas que não possuem empregados;
- Empregadores: pessoas que exercem uma profissão com o auxílio de um ou mais empregados;
- Não Remunerados: pessoas que exercem uma ocupação econômica sem remuneração, por pelo menos 15 horas semanais.

2.2.4 Taxa de Desocupação

A população desocupada complementa a população ocupada, que na soma, contabilizam a força de trabalho, sendo assim, a taxa de desocupação é o percentual da relação entre o número de desocupados como proporção da população economicamente ativa, também pode ser chamada

como taxa de desemprego aberto. Esse conceito engloba as pessoas que não tinham trabalhado, em determinado período de referência, mas que estavam dispostas a trabalhar nessa semana e tomaram alguma providência para conseguir emprego nos últimos 30 dias de referência. O esquema para o estudo acerca do mercado de trabalho está ilustrado na Figura 3, com as aberturas a partir da compreensão de cada designação.

Figura 3 – Organograma da Representação do Mercado de Trabalho



Fonte – Elaboração própria.

Portanto, a partir da população economicamente ativa é possível medir o número de desempregados, sendo a parcela da força de trabalho classificada como desocupada. De acordo com Larraın e Sachs (2000), as medições da taxa de desemprego irão variar de acordo com os critérios de cada órgão de mensuração, os principais órgãos que avaliam a taxa de desocupação estão listados na próxima seção.

2.2.5 Metodologias de Cálculo

No Brasil, há diversas pesquisas e órgãos que fazem a estimação da quantidade de pessoas desempregadas, os principais são: IBGE, Departamento Intersindical de Estatística e Estudos Econômicos (DIEESE), Ministério do Trabalho (MTE) e Organização Internacional do Trabalho (OIT). O IBGE, por exemplo, possui três pesquisas que apresentam os resultados de forma distinta, mas com o mesmo objetivo. Entre elas tem-se:

- Pesquisa Mensal do Emprego (PME): iniciada em 1980 e encerrada em março de 2016, divulgava resultados mensais referentes ao mercado de trabalho para seis Regiões Metropolitanas (Recife, Salvador, Belo Horizonte, Rio de Janeiro, São Paulo e Porto Alegre);

- Pesquisa Nacional por Amostra de Domicílios (PNAD): iniciada em 1967 com resultados trimestrais e a partir de 1971 com resultados anuais, foi encerrada em 2016, divulgava informações referentes ao emprego e também de características gerais da

população, para o Brasil, Grandes Regiões, Unidades da Federação e nove Regiões Metropolitanas (Belém, Fortaleza, Recife, Salvador, Belo Horizonte, Rio de Janeiro, São Paulo, Curitiba e Porto Alegre);

- PNAD Contínua: iniciada em 2012, com periodicidade mensal, trimestral, anual e variável dependendo das informações, mas sempre tendo como divulgação resultados referentes a força de trabalho, substitui a PNAD, trazendo dados para o Brasil, Grandes Regiões, Unidades da Federação, Regiões Metropolitanas que contêm Municípios das Capitais, Região Integrada de Desenvolvimento (RIDE) e Municípios das Capitais.

O DIEESE em convênio com o Portal de Estatísticas do Estado de São Paulo (SEADE), iniciaram o levantamento de dados mensais a partir de 1984, através da Pesquisa de Emprego e Desemprego (PED), restrita à Região Metropolitana de São Paulo. A partir de 1987, ocorreu a implantação da pesquisa para outras regiões do país, entre elas: Distrito Federal e nas Regiões Metropolitanas de São Paulo, Porto Alegre, Belo Horizonte, Salvador, Recife e Fortaleza.

O MTE inclui informações, em dados estatísticos, do mercado de trabalho, migrações internacionais, entre outros. As duas formas de cálculo referentes a força de trabalho são:

- Relação Anual de Serviços Sociais (RAIS): é um cadastro administrativo contabilizado desde 1975, de âmbito nacional - Brasil, Regiões Naturais, Unidades Federativas, Regiões Metropolitanas e Municípios -, com periodicidade anual e de declaração obrigatória para todos os estabelecimentos do setor público e privado. A base de dados é composta pelo estoque de trabalhadores que se encontram no mercado de trabalho formal;

- Cadastro Geral de Empregados e Desempregados (CAGED): com informações registradas a partir de 1965, o CAGED foi criado como uma fonte de informação para o mercado de trabalho com periodicidade mensal. O registro abrange Brasil, Regiões Naturais, Unidades Federativas, Mesorregiões, Microrregiões e Municípios.

A Organização Internacional do Trabalho (OIT) é um organismo que faz mensuração acerca do trabalho, mas em âmbito internacional, divulgando dados estatísticos sobre inúmeros países. A Organização foi fundada 1919 para promover a justiça social e possui representação no Brasil - em Brasília - desde a década de 1950. A periodicidade dos dados é anual e a referência estatística sobre desemprego inicia em 1991, sendo que o último dado divulgado é referente ao ano de 2018.

2.2.6 Relações do Desemprego com Variáveis Macroeconômicas

Na macroeconomia existem duas relações principais associadas à taxa de desemprego, elas são dadas pela Lei de Okun e pela Curva de Phillips. A Lei de Okun (1962), desenvolvida por

Arthur Melvin Okun, verifica a relação existente entre o crescimento econômico - representado pelo desvio da taxa de crescimento do produto em relação à taxa natural - e a taxa de desemprego - diferença entre a taxa de desemprego natural e taxa de desemprego vigente. A Lei busca estimar quanto deveria crescer o produto de uma economia caso a taxa de desemprego caísse em um ponto percentual, ou quanto o desemprego deve variar dado uma redução de um ponto percentual do produto. Samary (2017) verificou, através do modelo produzido para o período 1996-2016, que dada uma redução em um ponto percentual no hiato do produto, a distância entre a taxa natural e a taxa vigente do desemprego no Brasil tende a variar em -0,28 pontos percentuais.

Outra relação macroeconômica estudada e difundida no mundo é a correlação inversa entre inflação - variável nominal - e o nível de desemprego - variável real -, expressada pela Curva de Phillips (1958), em que se encontrou uma relação negativa entre a taxa de crescimento dos salários e taxa de desemprego. Antes de Phillips, outros autores como Karl Marx (1867) e Irving Fisher em um artigo em 1926, já haviam notado a relação inversa entre essas duas variáveis, conforme afirma Alves Filho (2004). A curva de Phillips é válida somente no curto prazo, quando a moeda não é neutra, pois no longo prazo, a moeda é neutra. Considerar a moeda neutra, segundo Mankiw (2015), significa que alterações no estoque de moeda não impactam variáveis reais, como o desemprego, deste modo, no longo prazo não existe o *trade-off* entre inflação e desemprego. Vieira (2014) fez uma análise desta relação de curto prazo para a economia brasileira, para o período de 2002-2014, porém, não foram encontrados resultados significativos, assim, não se pode afirmar que existe essa relação inversa entre inflação e desemprego para o Brasil.

2.3 PRINCIPAIS TIPOS DE DESEMPREGO

2.3.1 Desemprego Friccional

O mercado de trabalho caracteriza-se por uma busca de emprego tanto por parte dos trabalhadores quanto pelas empresas. Porém, mesmo com vagas disponíveis para trabalhar, uma empresa pode não encontrar um indivíduo com determinadas habilidades requeridas. A explicação dada pelos novos clássicos, é que existe um descompasso entre as capacitações dos indivíduos em relação àquelas que as firmas desejam. Essa diferença que surge quando as empresas procuram por trabalhadores adequados e trabalhadores procuram por empregos apropriados, é chamada de desemprego friccional. O desemprego friccional é mensurado principalmente por aqueles que recentemente saíram de um emprego para procurar outro melhor ou estão entrando no mercado de trabalho, é normalmente classificado como o desemprego de curto prazo.

2.3.2 Desemprego Estrutural

O desemprego estrutural, segundo Larraın e Sachs (2000), ocorre quando a economia está atuando à sua taxa natural de desemprego, é comumente classificado como o desemprego de longo prazo. Muitos indivíduos ficam por um longo período de tempo na situação de desemprego, desta forma, essas pessoas não conseguem encontrar um emprego estável, e, assim, não procuram trabalho com tanta intensidade. De acordo com os novos clássicos, o desemprego do tipo estrutural vai ocorrer por dois principais motivos: pela discrepância entre as habilidades requeridas pelas firmas em relação àquelas apresentadas pelos indivíduos; e pelo crescimento de setores - principalmente tecnológicos - que estão substituindo a mão de obra humana por máquinas, isso cria um determinado nível de desemprego que só vai ser suprimido com a criação de outros empregos inexistentes no momento. O resultado da soma do desemprego friccional e estrutural propiciará a taxa natural de desemprego.

2.3.3 Desemprego Cíclico

Por último, ainda existe a definição de desemprego cíclico, este é associado a empregos perdidos durante uma recessão, geralmente chamado de desvio da taxa natural de desemprego. Essa especificação se refere a forma de desemprego involuntário, em que as pessoas desejam trabalhar, mas devido a mudanças no ciclo econômico, não conseguem encontrar emprego. Essa relação de mudanças no ciclo econômico, entre o aumento do desemprego com a produção da economia decrescendo, pode ser vista pela Lei de Okun, como mencionada na seção 2.2.6.

2.4 ABORDAGEM MICROECONÔMICA

Após entender o desemprego em aspectos teóricos macroeconômicos, acerca dos dois principais modelos de mercado de trabalho, conceitos e tipos de desemprego, esta seção pretende abordar o tema de uma forma microeconômica, a respeito das características individuais de um determinado grupo que podem aumentar as chances de estarem desempregados. Essa abordagem dos fatores microeconômicos que influenciam a probabilidade dos indivíduos se tornarem desempregados, se deu com base em diferentes artigos, estes, com finalidades distintas de impacto - utilizando, por exemplo, para uso de políticas públicas - mas sugerindo variáveis explicativas de modo a corroborar as que serão utilizadas neste trabalho.

Oliveira, Scorzafave e Pazello (2009), através da análise de dados da PNAD de 2004, selecionaram indivíduos de Regiões Metropolitanas com 10 anos ou mais de idade, totalizando 107.376 pessoas. O objetivo do *paper* foi identificar quais fatores atuavam como determinantes do desemprego e da inatividade em 2004, para assim direcionar melhor o uso de políticas públicas. Os autores atualizaram o trabalho de Fernandes e Picchetti (1999), utilizando as mesmas variáveis explicativas, mas diferenciando por gênero, são elas: anos de estudo, idade, renda domiciliar líquida, sexo, raça, condição na família e condição referente à aposentadoria. A

variável dependente, que por sua vez depende das variáveis explicativas, admite três categorias: ocupado, inativo e desempregado. A estimação se deu através do modelo *logit* multinomial e as variáveis que mais tiveram influência, de acordo com a análise, no desemprego foram: educação - até quatro anos de estudo, a probabilidade é similar para homens e mulheres, e a partir disso aumenta para as mulheres -; idade - até os 45 anos a probabilidade de desemprego das mulheres é maior que a dos homens, mas isso se inverte após os 45 anos -; e raça - pessoas negras possuem maior chance de estarem desempregadas em relação a brancos e pardos -.

Já Costa e Cunha (2010), elaboraram um artigo sobre determinantes do desemprego utilizando como principais variáveis educação (como *proxy* de qualificação) e abertura comercial, abrangendo o Brasil no período de 1981 a 2005. Para isso, os autores utilizaram o modelo *probit*, tendo como variável dependente a condição do indivíduo estar ou não desempregado. A pesquisa sugeriu que a qualificação do indivíduo possui efeito significativo num contexto de maior abertura comercial. Além das duas variáveis focalizadas, os fatores individuais como: idade, sexo, condição do domicílio, região geográfica e metropolitana também se mostraram contribuintes para determinar o desemprego.

De acordo com o *paper* de Araújo, Cunha e Lima (2011), em que analisaram os determinantes do desemprego e inatividade para os jovens[4] no Brasil metropolitano, a probabilidade de inatividade é maior para jovens que estudam e que ocupam a posição de cônjuge na família e diminui a probabilidade de desemprego a medida em que se aumenta os anos de experiência do indivíduo. Os resultados foram estimados por um modelo *logit* multinomial, que envolve três categorias que o jovem pode se encontrar: inativo, ativo e empregado ou ativo e desempregado, através de dados da PNAD de 2007.

Uma pesquisa realizada por Sakowski (2012) em cooperação entre o Instituto de Pesquisa Econômica Aplicada (IPEA) e o Ministério do Turismo do Brasil, realizou uma análise das estatísticas de emprego nas indústrias do turismo através da Relação Anual de Informações Sociais (RAIS). A pesquisa, que se dá para o ano de 2010, é apresentada em níveis nacional, regional e local, dando um enfoque em estados nordestinos pela influência turística que estes têm no Brasil. O estudo mostrou que o Brasil deveria explorar melhor as fontes de dados disponíveis, pois estas podem gerar estatísticas de emprego que ajudam na tomada de decisão para a implementação de determinadas políticas públicas, neste caso, focadas no turismo.

Barufi (2014), realizaram um trabalho com dados em painel[5] para seis anos - 2003, 2004, 2005, 2008, 2009 e 2012 - com informações obtidas da RAIS. O objetivo do estudo era verificar como as economias de aglomeração - etapa do processo de industrialização em que acontece a concentração de indústrias produtivas em determinada região - estão presentes no mercado de trabalho formal brasileiro, para isso foi feito uma investigação a nível municipal. Através do estudo de desigualdade salarial, o artigo encontrou, em seus resultados, que economias de

[4]Foi considerado jovem qualquer indivíduo na faixa etária entre 16 e 29 anos de idade.
[5]O termo painel é utilizado para representar várias unidades amostrais acompanhadas ao longo do tempo.

aglomeração parecem aumentar a desigualdade salarial, ou seja, cidades maiores possuem uma maior diferença de remuneração.

Recentemente, Tangtipongkul e Wangmo (2017) publicaram um estudo sobre as características que podem determinar o desemprego para o Butão em 2015. A pesquisa se deu através de um modelo *probit* e, de acordo com os resultados auferidos, os dois fatores que mais influenciaram na probabilidade de se tornar desempregado eram o fato de ser mulher e jovem no mercado de trabalho. As variáveis de controle utilizadas para determinar as características que influenciavam o desemprego foram sexo, idade, educação, estado civil, migração e o patrimônio doméstico.

Muitos estudos, como os de Fernandes e Picchetti (1999), Ramos e Vieira (2001), Oliveira, Scorzafave e Pazello (2009) e Araújo, Cunha e Lima (2011) apontaram a variável raça como sendo determinante do desemprego. Segundo a Pesquisa de Emprego e Desemprego (PED) calculada pelo DIEESE, a taxa de desemprego daqueles considerados negros é maior do que a dos demais grupos raciais, entre os setores de serviços, comércio e indústria, o comércio é onde os negros possuem maior participação em relação aos não negros. Além da taxa de desemprego, a PED também informou que a remuneração para um mesmo serviço é demasiadamente diferente para negros e não negros, em que os negros recebem menos em comparação ao outro grupo. O que se conclui é que "Os negros estão mais presentes em ocupações mais precárias, caracterizadas pela ausência de proteção social, e, por consequência, menores remunerações" (DIEESE, 2015, p. 2), assim, o fator raça é de extrema importância para a determinação da probabilidade de desemprego.

Outra variável de grande destaque nos trabalhos é o fator gênero. Segundo Probst e Ramos (2003), posteriormente às primeira e segunda guerras mundiais, as mulheres aumentaram sua participação no mercado de trabalho. Apesar disso e de todas as conquistas que estas conseguiram, ainda há muita discriminação no mercado de trabalho entre homens e mulheres, sendo, segundo o autor, a taxa de desemprego mais alta para o grupo feminino e existindo principalmente desigualdade salarial. No entanto, para a economia americana, o *Federal Reserve Bank of St. Louis* (FRED) divulgou recentemente dados para o período de 1974-2018 em que analisou a taxa de desemprego nos Estados Unidos separada por gênero e raça, e indicou que ao longo deste período, principalmente após as recessões, ocorreu uma elevação maior da taxa de desemprego dos homens do que das mulheres. Esta relação está em concordância com o *paper* de Wall et al (2009), que analisou o desemprego após a recessão de 2008, confirmando que os homens sofreram um maior impacto no emprego quando comparado com as mulheres.

Barros, Camargo e Mendonça (1997), ao analisar a relação entre as variáveis de desemprego e educação, afirmam que não existe uma relação monotônica entre desemprego e nível educacional. A dependência entre essas duas variáveis se dá em forma de um U invertido - detectando a presença de concavidade. Deste modo, segundo os autores, a taxa de desemprego sobe até sete anos de estudo, e, a partir deste nível de escolaridade, começa a decair. Oliveira,

Scorzafave e Pazello (2009) também encontraram essa relação côncava entre desemprego e educação, tanto para homens quanto para mulheres, mas a partir de quatro anos de estudo a probabilidade de desemprego passou a ser maior para as mulheres.

Monge-Naranjo e Sohail (2015) apontam em seu estudo que a idade também é um fator que aumenta a probabilidade de se tornar desempregado. A pesquisa relata que, independente do gênero considerado, conforme a idade dos indivíduos aumenta, cresce o número de desempregados nos Estados Unidos. O trabalho, feito para duas amostras, antes e depois da recessão de 2008, mostrou que o percentual de desempregados mais do que dobrou tanto para homens quanto para mulheres.

De acordo com os referidos artigos e conforme a percepção acerca dos fatores microeconômicos, as variáveis a serem utilizadas para explicar a probabilidade de desemprego neste trabalho para as regiões brasileiras no ano de 2017 serão: raça, gênero, escolaridade, idade e, por último, uma variável que diferencia um grupo que vem sendo discutido nos países desenvolvidos e que será abordada na seção seguinte, os STEM.

2.5 MERCADO DE TRABALHO - CONCEPÇÃO MODERNA

2.5.1 Modelo DMP para Desemprego

Uma das principais abordagens recentes sobre as causas do desemprego, foi premiada na academia, em 2010. A premiação se deu no Nobel de Economia, concedida aos pesquisadores americanos Peter Diamond e Dale Mortensen e ao britânico Christopher Pissarides. A contribuição destes economistas foi para explicar a "falha" de comunicação existente entre a oferta e a procura de emprego em diferentes mercados. O modelo ficou conhecido como DMP (Diamond-Mortensen-Pissarides), que, além de ajudar a explicar o "ruído" entre oferta e demanda no mercado de trabalho, também fornece sugestões para a elaboração de programas de políticas econômicas, com o objetivo de reduzir o desemprego associado a essas falhas de comunicação.

A formação da modelagem inicia em 1970, com uma publicação de Mortensen (1970), seguida pelo modelo matemático de Diamond (1971). No final dos anos 1970, os professores Diamond, Mortensen e Pissarides se fixaram na análise de políticas públicas acerca de modelos de busca de emprego, chegando a conclusão de que existem externalidades positivas e negativas sobre perder o emprego. Quando uma pessoa perde ou sai do emprego ocorre a geração de uma externalidade positiva para as pessoas que estão desempregadas, dado que existe uma vaga a mais no mercado, mas, em compensação, se esta pessoa estiver atrás de outro emprego, ocorre uma externalidade negativa, pois há mais uma pessoa, além de toda a população que está procurando emprego, em busca de um novo trabalho. Em resumo, aconteceram muitas publicações para se chegar no modelo DMP, mas este modelo utiliza a modelagem matemática feita por Diamond

em 1971 e por trabalhos produzidos principalmente por Mortensen e Pissarides desde 1970 até 2005.[6]

2.5.2 Força de Trabalho nas Áreas de Ciência e Tecnologia: STEM

O termo *Science, Technology, Engineering e Mathematics* (STEM), foi concebido na literatura a partir do final dos anos noventa, pela Fundação Nacional da Ciência dos Estados Unidos. Blackley e Howell (2015), mencionaram que diversas organizações estão dando importância aos STEMs, tais como: a Organização para a Cooperação e Desenvolvimento Econômico (OCDE), Banco Mundial, a Comissão das Nações Unidas para Ciência e Tecnologia e a *International Association for the Evaluation of Educational Achievement* (IEA), tendo como propósito a implementação das disciplinas de ciência, tecnologia, engenharia e matemática nas graduações.

Conforme sugere Davies (2014), não existe uma definição padrão para o conjunto de áreas de formação e atuação profissional STEM. Enquanto áreas como programação de computadores, engenharia mecânica e ciências ambientais são consensualmente pertencentes ao conjunto STEM, ocupações relacionadas à medicina, arquitetura e ciências sociais são mais questionadas (LANGDON et al., 2011). Para o Brasil, Seemann e Bonini (2017) sistematizaram um grupo de ocupações STEM dentro da Classificação Nacional de Atividades Econômicas (CNAE), seguindo os critérios do Departamento de Estatística do Governo Americano - ESA (*Economics and Statistics Administration*), o estudo identificou para o Brasil 164 códigos de ocupações correspondentes ao grupo de interesse.

Em termos de atuação profissional, seguindo os critérios do ESA, Langdon et al. (2011), apresentaram uma lista de cinquenta ocupações específicas do grupo STEM e apontaram que, em 2009, a força de trabalho STEM nos Estados Unidos era formada por 7,4 milhões de trabalhadores, o que representava 5,3% da força de trabalho total americana. Neste mesmo ano, o grupo de empregados STEM, segundo Langdon et al., era composto principalmente por quatro categorias, seguidas respectivamente de seu percentual em relação ao total de empregados nos campos STEM: computação e matemática (47%), engenharia e pesquisa (33%), ciências físicas e naturais (12%) e trabalhos de gerência (8%).

Apesar da parcela de trabalhadores nos campos STEM ser pequena em comparação às áreas complementares, aqueles que ocupam estas áreas costumam apresentar maiores salários e melhores perspectivas de emprego. Langdon et al. (2011), encontraram evidências estatísticas de que trabalhadores que estão ocupados em alguma área STEM, ganham, em média, mais do que trabalhadores fora deste conjunto, independente do seu nível educacional. Esse "ganho maior" dos trabalhadores STEM em relação aos não-STEM é comumente chamado na literatura de vantagem salarial, esta, expressa a diferença salarial - geralmente maior - para os STEM quando comparado ao outro grupo.

[6]Para mais informações sobre o modelo o leitor pode verificar, por exemplo, Zárate (2014)

Como existe essa vantagem salarial para o grupo STEM, existe um alto incentivo para entrar em alguma das áreas que o compõe. Baranyai et al. (2016) apresentaram estatísticas que informam que entre os anos de 2006 e 2011, a Austrália obteve um aumento de 14% do número absoluto de pessoas que estavam empregadas em algum campo STEM, sendo que a área que teve maior crescimento do número de empregados foi a de Tecnologia da Informação - com 25% de aumento. Outro ponto a ser destacado, é que na Austrália, homens que estavam empregados em alguma ocupação STEM estavam mais propensos a trabalhar o dia inteiro do que homens Não-STEM.

A partir do estudo sobre emprego do governo australiano, Baranyai et al. (2016) corrobora que a taxa de desemprego para o grupo STEM é geralmente menor quando comparado ao grupo complementar. De acordo com as estatísticas, entre 2006 e 2011, o número de desempregados STEM aumentou 25%, enquanto para o grupo não-STEM, neste mesmo período, este número aumentou 49%. Destas informações, a taxa de desemprego foi maior para as mulheres do que para os homens, em cada campo STEM. Para os Estados Unidos, Landivar (2013), mostrou estatísticas parecidas com a da Austrália, argumentando que em 2011, a taxa de desemprego dos Estados Unidos para o grupo STEM era menor - cerca de 3,9% - em relação ao grupo de comparação.

Conforme foi apontado pelas pesquisas desenvolvidas para outros países, a taxa de desemprego para os empregados em ocupações STEM é menor do que o grupo complementar. Assim, espera-se que a probabilidade de um empregado em alguma destas áreas ficar desempregado é menor. Portanto, este trabalho possui o intuito de testar se nas regiões brasileiras existe congruência com a literatura de economias modernas.

3 METODOLOGIA E DESCRIÇÃO DOS DADOS

3.1 CARACTERIZAÇÃO DA PESQUISA

De acordo com Gil (2002), as pesquisas podem ser classificadas com base nos seus objetivos de estudo. Elas podem ser classificadas em exploratórias, descritivas e explicativas. Este trabalho é classificado como uma pesquisa descritiva, esta tem como objetivo a descrição das características de determinada amostra ou então pelo estabelecimento da relação entre determinadas variáveis. Além disso, o fato de ser descritiva implica necessariamente no uso de coleta de dados, para posteriormente fazer uma análise descritiva dos resultados gerados pelos dados.

Com base nos procedimentos técnicos utilizados, este trabalho pode ser classificado como uma pesquisa bibliográfica e documental. Bibliográfica por se utilizar de livros e artigos científicos na fundamentação e documental pelo uso de dados que ainda não receberam nenhum tratamento analítico. A partir da fundamentação teórica feita na revisão bibliográfica é possível fazer uma análise satisfatória acerca dos dados.

3.2 MODELOS DE RESPOSTA QUALITATIVA

Os modelos de resposta qualitativa têm se mostrado cada vez mais frequentes em estudos que utilizam dados sociais - tais como pobreza e desemprego - para áreas financeira, de análise de risco na concessão de crédito e até mesmo em medicina. Estes modelos surgiram para situações em que a variável dependente é de natureza binária ou dicotômica, isto é, um aluno pode ser aprovado ou reprovado num exame, um cliente pode confirmar ou cancelar um pedido de compra. Isto significa que dado um conjunto de variáveis explicativas, que podem ser categóricas ou não, o objetivo será explicar qual a probabilidade de determinado evento ocorrer. Para resolver esses tipos de problemas foi desenvolvida a técnica estatística que se dá a partir de modelos de probabilidade.

Os principais modelos de probabilidade utilizados para uma regressão multivariada com uma variável dependente qualitativa são *logit* e *probit*. Costa e Cunha (2010) utilizaram um modelo *probit* para apontar os principais determinantes do desemprego no Brasil no período de 1981-2005, sendo que a variável dependente era a condição do indivíduo ser considerado como desempregado ou não. Os resultados da pesquisa de Costa e Cunha mostraram que o nível de qualificação ou educação do indivíduo é um dos principais fatores que pode determinar o desemprego em um contexto de maior abertura comercial.

Uma exemplificação do modelo *logit* foi a pesquisa de Araújo, Cunha e Lima (2011), estes verificaram que fatores são mais relevantes para explicar o desemprego e a inatividade de jovens entre 16 e 29 anos de idade para as regiões metropolitanas do Brasil. A partir do modelo *logit* multinomial foi encontrado que a probabilidade de desemprego tende a diminuir à medida que aumenta o tempo de experiência dos indivíduos. Além disso, a inatividade é maior nos jovens que ocupam a posição de cônjuge na família e que estudam.

A partir da visão destes dois trabalhos, percebe-se que os trabalhos de modelos de resposta qualitativa vêm sendo cada vez mais utilizados. A diferença desses modelos de probabilidade em relação aos demais modelos de regressão, é que a variável Y é qualitativa. O objetivo de modelos de resposta qualitativa é encontrar a probabilidade de ocorrência de um evento - no caso deste trabalho: a probabilidade do indivíduo se encontrar na posição de desempregado - a partir de um conjunto de variáveis explicativas.

Os modelos que empregam o regressando como uma variável qualitativa são chamados de modelos de probabilidade. O fato de usar a variável dependente como qualitativa implica que o regressando é uma variável binária ou dicotômica. Tais variáveis indicam a presença ou ausência de um atributo, como homem ou mulher, solteiro ou casado, empregado ou desempregado, sendo estas variáveis geralmente nominais. Assim é possível quantificar esses atributos, em que estes assumem valores de 0 ou 1, onde 0 indica a ausência daquele atributo e 1 indica a presença dele (GUJARATI; PORTER, 2011).

A principal diferença entre os modelos *logit* e *probit*, segundo Gujarati e Porter (2011), consiste na distribuição logística que tem caudas mais pesadas. Ou seja, a probabilidade condicional P_i aproxima-se de 0 ou 1 com um ritmo mais lento no *logit* que no *probit*. Para escolher qual dos modelos utilizar, verifica-se o grau de acerto de cada modelo, de acordo com uma matriz de "casos corretamente previstos", produzida pelo *software* STATA. De acordo com o resultado dos dois modelos para cada região, o mais adequado - e, portanto, o que será utilizado - foi o modelo *logit*.

3.3 MODELO LOGIT

Segundo Corrar et al. (2011), a técnica da regressão para variáveis dicotômicas foi desenvolvida por volta de 1960 como uma resposta para explicar a ocorrência de determinados fenômenos quando a variável dependente é binária. A técnica surgiu e se desenvolveu na medicina, primordialmente pelo estudo feito pelo projeto de *Framingham Heart Study*, realizado com a colaboração da Universidade de Boston. O objetivo do projeto foi identificar os fatores que poderiam desencadear doenças cardiovasculares, em que foi reconhecido diversos fatores de risco, como hipertensão arterial, diabetes e entre outros. Além da medicina, muitas aplicações econômicas têm feito uso destes modelos, para explicar a relação entre uma variável dependente dicotômica e um conjunto de variáveis preditoras.

Como o modelo de regressão linear é inadequado para estimar probabilidades, pois a variável dependente não pode assumir valores menores que zero ou maiores que um, efetua-se uma transformação na variável dependente. Apesar da variável Y poder assumir somente os valores zero ou um, torna-se necessário obter valores que possam ser interpretados em termos de probabilidade. Logo, a primeira transformação a ser feita é converter a probabilidade associada a cada observação em razão de chance (*odds ratio*), que representa a probabilidade de sucesso comparada com a de fracasso. Essa associação pode ser expressa da seguinte forma:

$$Razo\ de\ chance = \frac{P(sucesso)}{1 - P(sucesso)} \tag{3.1}$$

Para facilitar a interpretação dos resultados, o próximo passo é obter o logaritmo natural da razão de chance:

$$ln\left(\frac{P(sucesso)}{1 - P(sucesso)}\right) = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2_i} + ... + \beta_k X_{k_i} \tag{3.2}$$

No lado esquerdo da Equação 3.2, tem-se o logaritmo natural da razão de chance. No lado direito, as variáveis independentes - que podem ser categóricas ou métricas - e seus respectivos coeficientes que expressam mudanças no log da razão de chance. Para obter a razão de chance estimada, pode-se aplicar o antilogaritmo aos dois lados da Equação 3.2, como se segue na Equação 3.3.

$$\frac{P(sucesso)}{(1 - P(sucesso))} = e^{\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2_i} + ... + \beta_k X_{k_i}} \tag{3.3}$$

Se a razão de chance estiver estimada, é possível identificar a probabilidade associada à ocorrência de determinado evento. A partir da Equação 3.3 e com base no conceito de chance, obtém-se a equação:

$$P(evento) = \frac{e^{\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2_i} + ... + \beta_k X_{k_i}}}{1 + e^{\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2_i} + ... + \beta_k X_{k_i}}} \tag{3.4}$$

Simplificando a Equação 3.4, dividindo o numerador e o denominador pelo termo $e^{\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2_i} + ... + \beta_k X_{k_i}}$, chega-se na equação do modelo *logit*, em que o evento será a probabilidade de se tornar desempregado.

$$P(desemprego) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2_i} + ... + \beta_k X_{k_i})}} \tag{3.5}$$

De acordo com Corrar et al. (2011), a estimação do modelo não pode ser feita através do método de mínimos quadrados ordinários. Deve se aplicar o método da máxima verossimilhança, que é uma forma de estimar parâmetros de distribuição de probabilidades que maximizem a função verossimilhança.

O modelo de probabilidade terá a variável dependente como binária, sendo esta uma *dummy* - enumerada por 0 ou 1 - que identifica os indivíduos como desempregados (y = 1) ou empregados (y = 0), e como variáveis independentes as características específicas de cada indivíduo. A variável advinda da RAIS que será utilizada como *proxy* de emprego ou desemprego será a de "Vínculo ativo em 31/12", sendo que quando o indivíduo deixa de ter vínculo com a empresa ele passa a se tornar desempregado, assumindo o valor 1. A Equação 3.6 introduz a forma funcional estimada:

$$y_i = \alpha_0 + \eta X_i + \sum_{1}^{4} \gamma_i D_i + \sum_{1}^{4} \lambda_i S_i + \epsilon_i \tag{3.6}$$

Em que:

- y_i: é uma variável *dummy* que assume o valor 1 se o indivíduo não possuía vínculo empregatício ativo em 31 de dezembro e 0 caso contrário;
- X_i: é a matriz de variáveis específicas aos indivíduos, sendo elas: *dummy* de gênero, que assumirá 1 caso o indivíduo seja identificado na base de dados como mulher e 0 caso seja identificado como homem; uma variável *dummy* de cor ou raça[1] (como identificados na RAIS) que assumirá 0 caso o indivíduo for identificado como "branco" ou "amarelo", e 1 caso contrário; uma variável *dummy* de ocupação, que assumirá valor igual a 1 caso o indivíduo pertença ao grupo STEM e igual a 0 caso ele pertença ao grupo Não-STEM;
- α_0: é o parâmetro de intercepto;
- η: é o vetor de parâmetros associados às características individuais;
- D_i: é uma *dummy* que representa as faixas de idades que os indivíduos podem estar. Sendo elas: faixa 1 (16 a 24 anos), faixa 2 (25 a 29 anos), faixa 3 (30 a 49 anos), faixa 4 (50 a 59 anos) e faixa 5 (60 anos ou mais). A variável vai de 1 à 4, a faixa 5 foi omitida para ser a *dummy* de comparação;
- γ_i: representam os parâmetros referentes às faixas de idade;
- S_i: é uma *dummy* que representa as faixas de escolaridade - separadas por grau de escolaridade de acordo com a RAIS - que os indivíduos se encaixam. Sendo elas: faixa 1 (aqueles que possuem grau de escolaridade menor ou igual ao ensino médio incompleto), faixa 2 (ensino médio completo), faixa 3 (ensino superior incompleto), faixa 4 (ensino superior completo) e faixa 5 (pós-graduado, incluindo quem possui

[1] A variável de cor ou raça é composta por "indígenas, brancos, pretos, amarelos, pardos, não identificados e ignorados". Indivíduos que possuíam a variável de cor ou raça com a informação correspondente "ignorada" foram excluídos da amostra.

mestrado e/ou doutorado); A variável vai de 1 à 4, a faixa 5 foi omitida para ser a *dummy* de comparação;

– λ_i: representam os parâmetros referentes às faixas de escolaridade;

– ϵ_i: representam os erros aleatórios do modelo.

O modelo, quando estimado, não pode ser diretamente interpretado como num modelo de regressão linear, interpretam-se apenas os sinais dos coeficientes. De acordo com Pino (2007), para interpretar o modelo *logit*, pode-se utilizar os conceitos de efeitos marginais e de elasticidade. No presente trabalho, é utilizada a interpretação através do efeito marginal. O efeito marginal corresponde à derivada de primeira ordem. O aumento de, por exemplo, uma unidade no valor da variável explicativa, aumenta em tantos pontos percentuais (p.p.) a probabilidade de a variável dependente assumir o valor um.

3.3.1 Medidas de Avaliação do Modelo

Nos modelos de regressão linear existe a necessidade de fazer alguns testes para verificar a capacidade de explicação do modelo e se os resultados gerados são significativos. Assim, no modelo *logit* também se faz necessário o uso de testes estatísticos, pois um modelo não significativo não pode ser usado para boas análises. Como o método utilizado para o modelo *logit* é de máxima verossimilhança, as estratégias de avaliação do modelo são diferentes de um modelo linear, essas medidas de avaliações serão descritas nesta seção.

Likelihood Value

Também conhecido como *Log Likelihood Value*, é um indicador que busca medir o potencial de o modelo estimar a probabilidade associada à ocorrência de determinado evento. Ele é uma medida muito parecida com a estatística F, utilizada em modelos de regressão linear, pois é possível ver a contribuição das variáveis explicativas, calculando primeiro o *Likelihood Value* apenas com a constante no modelo e posteriormente com a inclusão de todas as variáveis independentes. Quanto mais elevada for a diferença entre os dois valores calculados, maior a capacidade dos coeficientes para estimar a probabilidade associada ao acontecimento de determinado evento. De acordo com Corrar et al. (2011), o *Likelihood Value* pode ser representado pela expressão -2LL, que significa o logaritmo natural do *Likelihood Value* multiplicado por -2, este segue uma distribuição Qui-quadrado. Em síntese, pode-se afirmar que o objetivo desta medida é verificar se a regressão é estatisticamente significativa, sendo que quanto mais próximo de zero estiver o *Likelihood Value*, maior o poder de predição do modelo.

Pseudo R^2

O nome R^2 se assemelha ao coeficiente de determinação dos modelos de regressão linear, mas no caso de modelos de probabilidade, o Pseudo-R^2 serve apenas para avaliar se o modelo melhora ou não a qualidade das predições. Segundo Corrar et al. (2011), para esta avaliação existem três possíveis métodos de cálculo, são eles: *McFadden's-R^2* ou R^2logit, *Cox-Snell R^2* e o teste *Nagelkerke R^2*. Não existe hierarquia entre os testes, no entanto, para fins de escolha, este trabalho utilizará o R^2logit. Este, expressa a variação percentual entre o *Likelihood Value* do modelo que considera somente a constante - representado por *lnModConstante* - e o *Likelihood Value* do modelo que incorpora as variáveis explicativas, representado por *lnModCompleto*.

$$R^2logit = 1 - \frac{lnModCompleto}{lnModConstante} \tag{3.7}$$

O numerador evidencia a melhora que se espera ocorrer com a inclusão das variáveis independentes no modelo, já o denominador tende a apresentar um *Likelihood Value* mais elevado, pois reflete o modelo apenas com a constante. O resultado do R^2logit varia entre 0 e 1, este pode ser apropriado para escolher o melhor modelo através da análise da inclusão e/ou exclusão das variáveis explicativas, sendo o melhor modelo aquele que apresentar o maior valor do R^2logit.

Teste de Wald

O *Teste de Wald*, em síntese, é um teste de significância de cada coeficiente do modelo, inclusive a constante. Assim como o teste T nos modelos lineares, o *Teste de Wald* tem por finalidade verificar se cada parâmetro estimado é significativamente diferente de 0. A estatística *Wald* segue uma distribuição Qui-quadrado e pode ser calculada pelo quadrado da divisão do valor estimado do coeficiente de uma determinada variável explicativa (β) pelo seu respectivo erro padrão (ep).

$$Wald = \frac{\beta}{ep}^{2} \tag{3.8}$$

A hipótese nula do teste é de que o coeficiente estimado é igual a zero, logo, o objetivo é rejeitar a hipótese nula, pois se assim feito, o coeficiente é considerado estatisticamente significativo. Se o coeficiente é considerado significativo, poderá ser mantido no modelo, caso contrário deve ser excluído.

Prob > chi2

Esta estatística tem como finalidade testar se todos os coeficientes do modelo são conjuntamente iguais a zero, sendo esta a hipótese nula. A hipótese nula é rejeitada quando a prob > chi2 é menor do que 0,05.

Outro teste importante de se fazer com as variáveis é o teste de multicolinearidade, que se dá usualmente pelo teste FIV. Esse teste serve para verificar se as variáveis explicativas possuem relações lineares exatas ou aproximadamente exatas entre si. Caso isso ocorra, o modelo regredido pode ter problemas de elevados desvios padrão, aumentando a variância dele, ou, em

caso de multicolinearidade perfeita, não pode ser estimado. Geralmente, existe um indicativo de problemas de multicolinearidade se o FIV para cada variável der maior que dez.

3.4 BASE DE DADOS

Este trabalho foi elaborado com base nos microdados da Relação Anual de Informações Sociais (RAIS) de 2017, este, provém do Ministério do Trabalho (MTE). A base abrange o mercado de trabalho formal brasileiro, ou seja, registra somente as pessoas que trabalharam com carteira assinada no período analisado, sendo esta relação divulgada anualmente. Esse levantamento de dados do mercado de trabalho formal é uma solicitação - de cunho obrigatório - do MTE às pessoas jurídicas, tendo estas o dever de declarar anualmente sobre o vínculo empregatício dos funcionários vinculados às mesmas.

O MTE utiliza o Programa de Disseminação das Estatísticas do Trabalho (PDET) para divulgar as informações da RAIS à sociedade civil, esta, abrange todo o território nacional brasileiro. A RAIS utiliza a Classificação Nacional de Atividade Econômica (CNAE), elaborada pelo IBGE, para classificar os dados quanto à ocupação de cada trabalhador e a atividade econômica onde este se insere. Além do CNAE, a RAIS também utiliza a Classificação Brasileira de Ocupações (CBO) 2002, esta última, é um documento que normaliza, nomeia e codifica os títulos e os conteúdos das ocupações do mercado de trabalho brasileiro, sendo ao mesmo tempo uma classificação enumerativa e descritiva. Neste trabalho será utilizado a codificação das ocupações pela CBO 2002.

A variável de interesse da base de dados para representar se o indivíduo está desempregado ou não é a de vínculo empregatício. A variável, designada pela RAIS como "Vínculo Ativo em 31/12", é definida como: "[...] as relações de emprego, estabelecidas sempre que ocorre trabalho remunerado", contabilizando a população economicamente ativa formal do Brasil. Além disso, as observações que apresentaram código CBO correspondente a ocupações da estrutura de "Membros das Forças Armadas, Policiais e Bombeiros Militares" foram retiradas da amostra. O motivo desse corte foi que os membros desta estrutura apresentam uma característica de estabilidade que difere do restante da população economicamente ativa.

Para a elaboração deste trabalho, e classificação quanto a divisão entre os que estão no mercado de trabalho como STEM ou não STEM, será utilizada a CBO 2002. A CBO 2002 segue uma hierarquia e se divide em 10 grandes grupos, além de outras divisões em subgrupos e famílias ocupacionais, que compreendem o total de 2.422 ocupações e de 7.258 títulos sinônimos. Complementarmente, a classificação descritiva esclarece as atividades desempenhadas no trabalho, os requisitos de formação e experiência profissionais.

As variáveis utilizadas neste trabalho, de acordo com a revisão de literatura feita na seção 2.4, estão na Tabela 1 com suas *dummies* discriminadas. A definição do grupo de ocupações que compõe o acrônimo STEM para o Brasil segue Seemann e Bonini (2017), este, é um dos primeiros

estudos com objetivo de sistematizar um grupo de ocupações STEM dentro da classificação CBO 2002, a partir dos critérios do Departamento de Estatística do governo americano - *Economics and Statistics Administration* (ESA).

Tabela 1 – Explanação das variáveis

		Dummy = 0	Dummy = 1
	Vínculo Empregatício	**Ativo ou Empregado**	**Inativo ou Desempregado**
Faixas de Idade	Faixa 1	*	16-24 anos
	Faixa 2	*	25-29 anos
	Faixa 3	*	30-49 anos
	Faixa 4	*	50-59 anos
	Faixa 5	*	60-90+ anos
	Raça	Branco	Não Branco
Faixas de Escolaridade	Menor Médio	*	1-6
	Médio	*	7
	Superior Incompleto	*	8
	Superior	*	9
	Pós Graduado	*	10-11
	Ocupação	Não-STEM	STEM
	Gênero	Homem	Mulher

Fonte – Elaboração própria, 2019.
Nota: * significa que o indivíduo não pertence ao grupo referido.

Os números relacionados às faixas de escolaridade representam o grau de instrução do indivíduo, conforme determinado pela RAIS. O conjunto de ocupações STEM utilizados neste trabalho está detalhado na Tabela 4 do APÊNDICE A. Dentre as 2.422 ocupações CBO, foram identificados 164 códigos de ocupações que correspondem ao grupo STEM, sendo os códigos restantes formados pelo grupo Não-STEM.

3.5 ANÁLISE DESCRITIVA

Esta seção consiste na exposição dos dados descritivos do Brasil, mostrando o comportamento do número de empregos formais - para o Brasil e posteriormente separado por regiões. Além disso, será analisado o comportamento da variável de interesse - vínculo empregatício - a partir das características dos indivíduos. Assumindo que a população economicamente ativa se mantém constante ao longo do tempo, quanto maior o número de empregos formais, menor será o número de desempregados neste mercado.

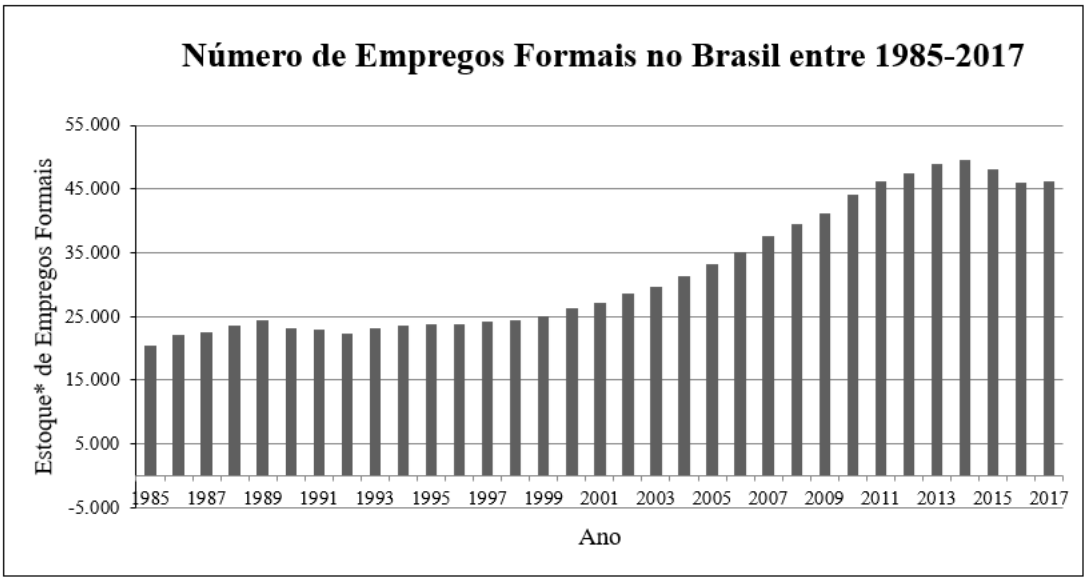
3.5.1 Estoque do Número de Empregos Formais

Os resultados divulgados pela RAIS designam a relação de quem está empregado - vínculo ativo - ou desempregado - vínculo inativo - dentro do mercado de trabalho formal. O

estoque do número de empregos formais no Brasil aumentou consideravelmente, se comparado ao ano de 1985. No período de 1985 a 2017, de acordo com a RAIS, a variação relativa foi quase por completo positiva. No entanto, nos períodos de 1990-1992 e, mais recentemente, em 2015 e 2016, o número de pessoas ativas decresceu, entre 2015 e 2016, com uma variação de aproximadamente 4 pontos percentuais negativos. A Figura 4 representa o número de empregos formais no Brasil no período 1985-2017.

No ano de 2017, o número total de empregos formais registrados na RAIS foi de 46.281.590, indicando um aumento de 0,48% em relação ao estoque de emprego em 2016. Como aumentou o número total de empregos, pode-se inferir, mantendo tudo o mais constante, que de 2016 para 2017 o número de desempregados no mercado de trabalho formal diminuiu. Os setores econômicos que apresentaram maior expansão em empregos formais, que reagiram mais rapidamente ao ciclo econômico, foram os de Administração Pública, Serviços e Agropecuária. No quesito de gênero, tanto homens quanto mulheres foram aclamados com expansão no emprego no ano de 2017, tendo uma maior participação da força de trabalho feminina.

Figura 4 – Número de Empregos Formais no Brasil



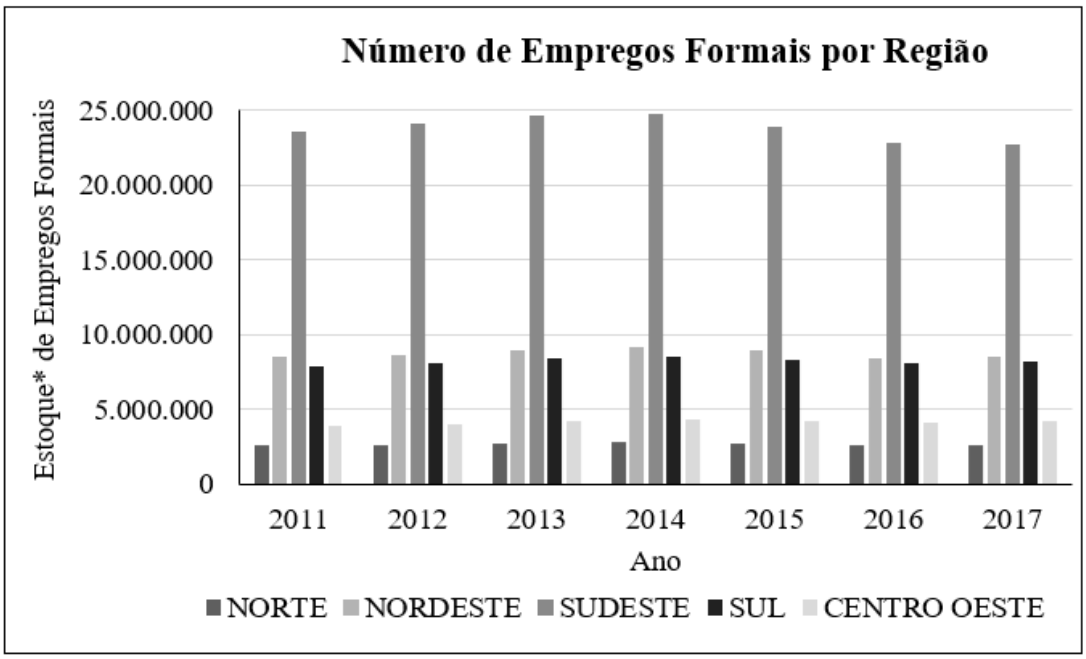
Fonte – Elaboração própria com base nos dados da RAIS, 2019.
*Estoque na base mil.

Para tornar a análise mais relevante, com a possibilidade de comparação, é possível examinar os dados estratificados por região. As regiões brasileiras são muito heterogêneas, segundo Matteo (2013), isso é explicado pela distribuição de cada região ao longo do país, tendo características históricas, geográficas, culturais, políticas e econômicas diferentes. Essa heterogeneidade cria uma desigualdade da distribuição territorial das atividades econômicas, sendo geralmente a região Sudeste mais influente na economia brasileira em relação às demais.

Conforme pode ser verificado na Figura 5, não houve uma mudança drástica sobre as regiões que detém maior participação do mercado de trabalho formal brasileiro. No ano de

2017, houve expansão do emprego em todas as regiões, com exceção do Sudeste. Apesar de não haver expansão do emprego na região Sudeste, esta deteve em 2017 cerca de 49,2% do estoque de emprego formal do Brasil. As Unidades da Federação que tiveram maior crescimento no estoque de empregos em 2017, em comparação a 2016, foram Minas Gerais, Goiás, Bahia e Santa Catarina.

Figura 5 – Número de Empregos Formais por Região no período 2011-2017



Fonte – Elaboração própria com base nos dados da RAIS, 2019.
*Estoque na base mil.

Como apresentado na sessão 2.2.6, em que foi verificado as relações com o desemprego, é possível utilizar a relação inversa do produto com o desemprego. Matteo (2013), ao analisar as regiões brasileiras, examina a participação dos setores de atividade econômica na distribuição do valor adicionado de cada região, considerando os setores: agropecuária, indústria, construção, comércio, serviços e administração pública. O Produto Interno Bruto (PIB) do Brasil é composto principalmente pelos setores de serviços e indústria, segundo dados do IBGE. A partir da relação da Lei de Okun, espera-se que a população de regiões onde esses setores tenham maior participação no seu PIB regional, tenham uma probabilidade menor de ficarem desempregadas.

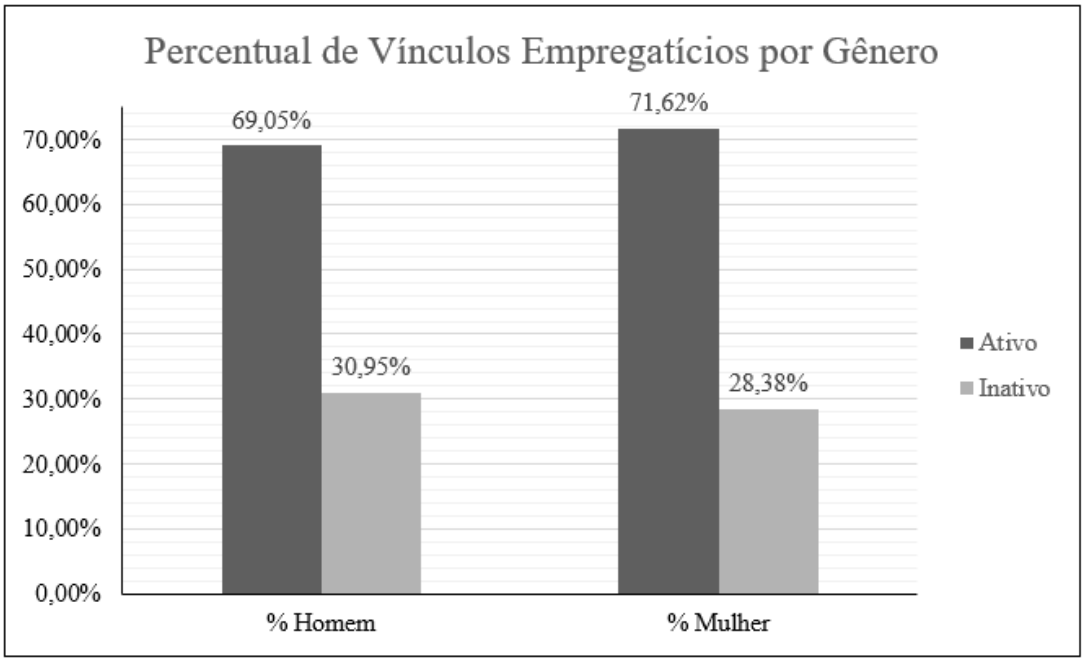
As regiões Norte e Nordeste, apresentam grande participação da administração pública como participação do seu PIB, que, por ser normalmente composta por concursos públicos, possui menor rotatividade se comparada às demais. A economia das regiões Sul e Sudeste são notavelmente parecidas, sendo que os setores que possuem maior destaque em ambas são o da indústria e o de serviços. As diferenças entre elas são verificadas pelo tamanho de cada economia e pelo fato de a região Sul ter forte presença da agropecuária em relação à região Sudeste. Já a região Centro Oeste, tem os setores de serviços e de administração pública como mais influentes

na participação do PIB, sendo que a participação do último setor é devida a presença do Distrito Federal.

3.5.2 Estatísticas Descritivas para o Brasil em 2017

Ao explorar os dados segundo gênero, percebe-se uma proporção muito similar na distribuição de homens e mulheres dentro do mercado de trabalho formal. Conforme mencionado no início desta seção, no ano de 2017 houve crescimento do número de empregos, tanto para homens quanto para mulheres, em relação ao ano de 2016. Do total de vínculos ativos, os homens correspondiam a 56% do estoque de empregos, e as mulheres a 44%. Ao examinar o percentual de homens ativos/inativos sobre o total de homens no mercado de trabalho, e a mesma relação para as mulheres, apesar da diferença ser mínima, o percentual de mulheres desempregadas é menor que o dos homens. Essa última relação pode ser vista na Figura 6.

Figura 6 – Vínculos Empregatício segundo Gênero em 2017

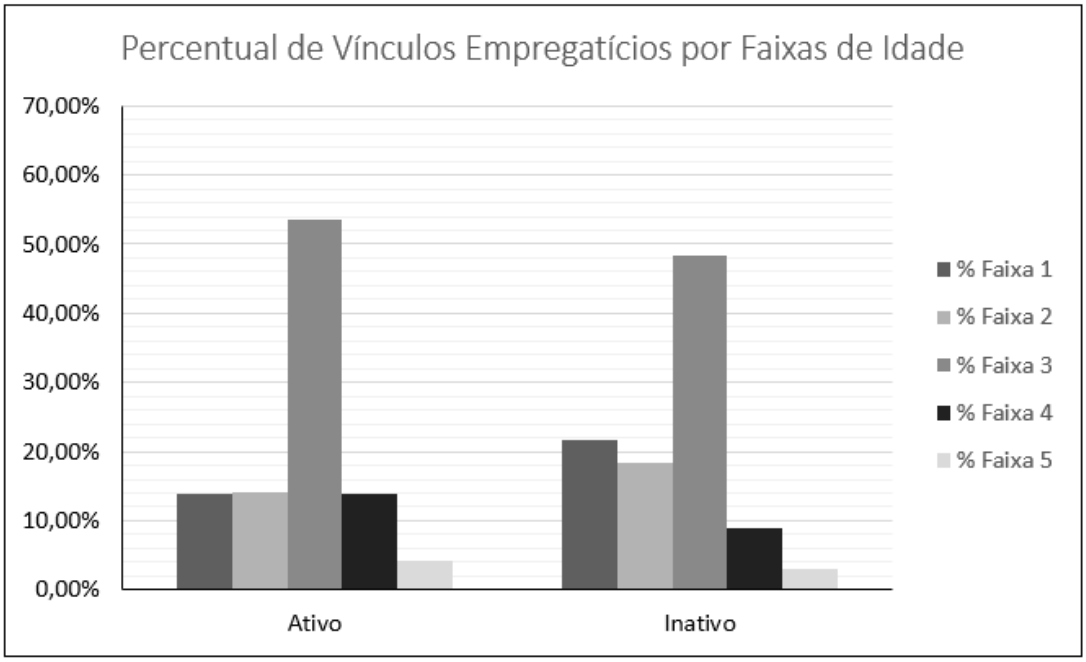


Fonte – Elaboração própria com base nos dados da RAIS, 2019.

Segundo o relatório de resultados da RAIS de 2017, ao observar a variável emprego por faixa etária e compará-la com o ano de 2016, verificou-se que as faixas de idade acima de 30 anos tiveram crescimento no estoque de empregos, enquanto as faixas até 29 anos registraram retração. Esta estatística está de acordo com a literatura, pois espera-se que pessoas mais jovens tenham maiores chances de perderem seus empregos, quando comparadas ao grupo que possui mais de 30 anos de idade. Além disso, ao verificar os dados para o ano de 2017, notou-se que indivíduos que possuem entre 30-49 anos - Faixa 3 - representam maioria tanto em percentual

de vínculo empregatício ativo, como inativo. Já os jovens, Faixa 1, representam a segunda faixa com maior percentual de inativos, como pode ser visto na Figura 7.

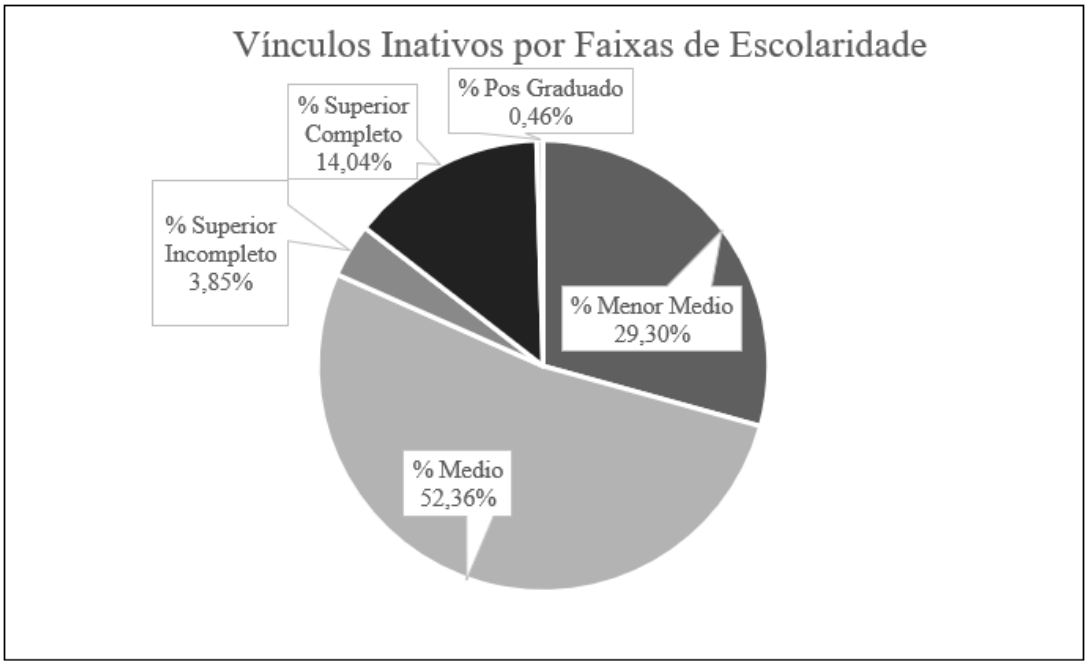
Figura 7 – Vínculos Empregatício segundo Idade em 2017



Fonte – Elaboração própria com base nos dados da RAIS, 2019.

As faixas de escolaridade também demonstram grande importância no mercado de trabalho. No ano de 2017, segundo o relatório de resultados da RAIS, comparado ao ano de 2016 as faixas de escolaridade mais elevada apresentaram crescimento do estoque de número de empregos, enquanto as faixas com escolaridade mais baixa registraram retração na quantidade de vínculos ativos. Ao analisar o percentual de desempregados por faixas de escolaridade na Figura 8, percebe-se que os grupos com menor escolaridade dominam a inatividade, enquanto os grupos formado por superior incompleto, superior completo e pós graduado representam apenas 18,35% do total de inativos no Brasil em 2017.

Figura 8 – Vínculos Empregatício segundo Faixas de Escolaridade em 2017

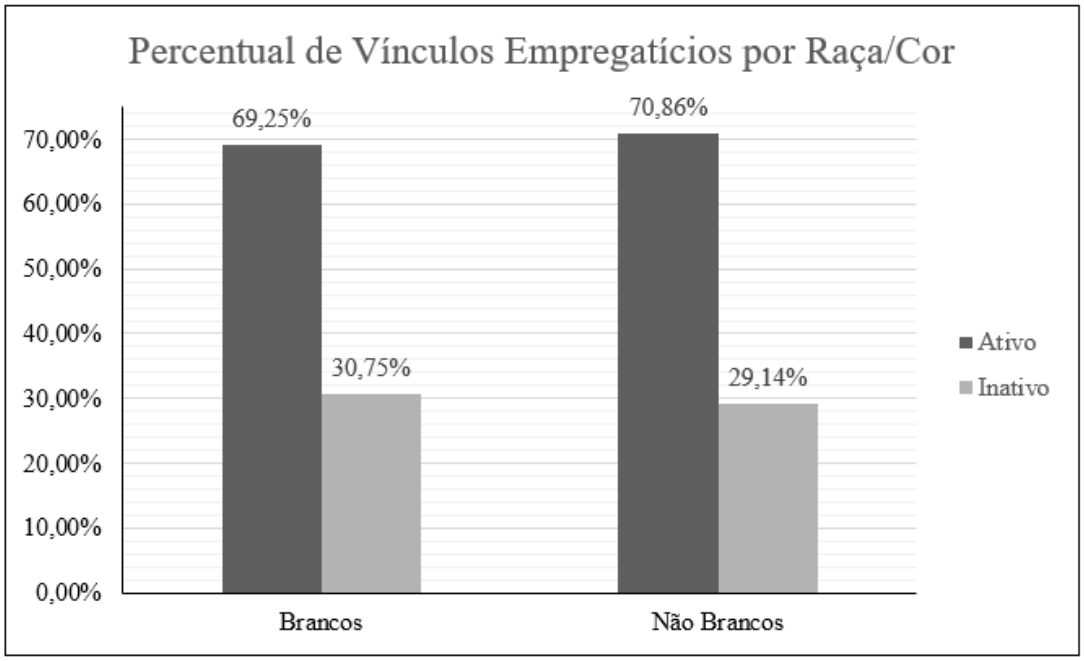


Fonte – Elaboração própria com base nos dados da RAIS, 2019.

Outra característica que pode influenciar na empregabilidade é a raça/cor do indivíduo. Em 2017, em comparação com o ano de 2016, as modalidades de raça/cor preta e parda expandiram o estoque de empregos, ao passo que as modalidades branca, amarela e indígena reduziram o número de vínculos ativos. A variável raça/cor pode apresentar certos problemas, como, por exemplo, quando a variável de pessoas que não declaram sua etnia[2] for alta. No ano de 2017, o número de pessoas que optou por não declarar sua raça/cor correspondeu a 27,5% do total de empregados. A proporção de indivíduos declarados como "não brancos" foi maior que os classificados como "brancos" em 2017. Quando comparado seus vínculos empregatícios, percebe-se uma relação similar, tanto de vínculos ativos quanto inativos, entre "brancos" e "não brancos". A Figura 9 mostra o percentual de vínculos empregatícios por raça/cor, sendo que a raça/cor classificada como "ignorado" foi excluída da amostra.

[2] A palavra etnia será usada como sinônimo de raça/cor.

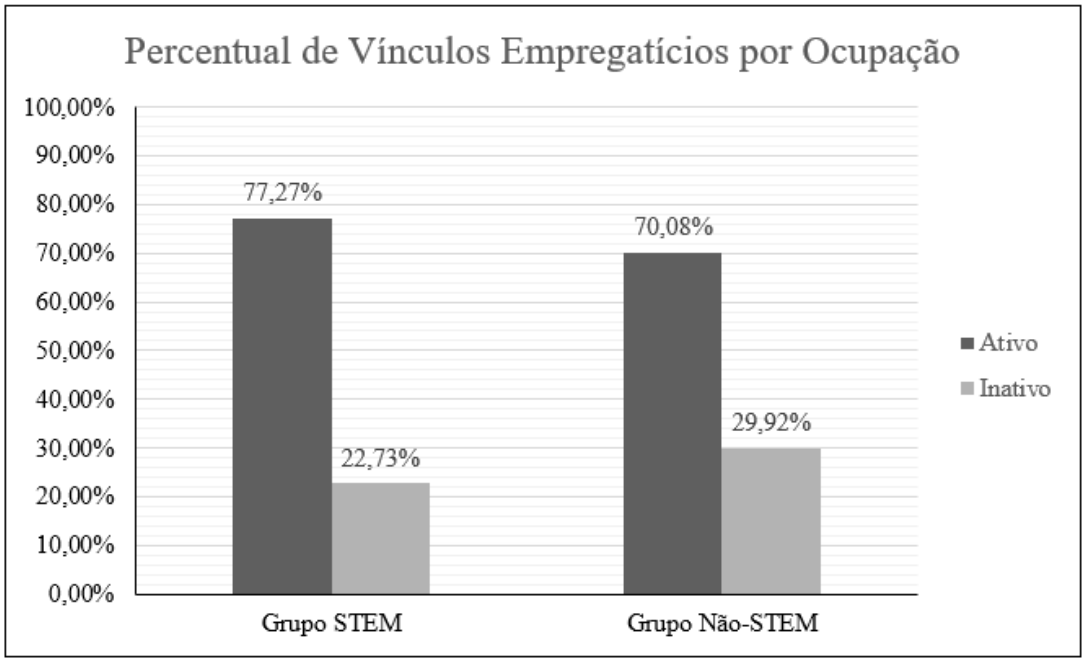
Figura 9 – Vínculos Empregatício segundo Raça/Cor em 2017



Fonte – Elaboração própria com base nos dados da RAIS, 2019.

A última variável de interesse para esta pesquisa são as ocupações dos indivíduos. Esta variável é empregada para atingir um dos objetivos deste trabalho, que é atestar, de acordo com a literatura revisada, que indivíduos empregados em ocupações STEM possuem menores chances de perderem seus empregos, ou seja, de ficarem inativos, se comparado ao grupo Não-STEM. Ao verificar essa relação para o Brasil no ano de 2017, percebe-se um resultado que corrobora com a literatura. Conforme mostra a Figura 10, em 2017, indivíduos classificados como STEM possuíam maior percentual de vínculos ativos - empregados - e menor percentual de desempregados.

Figura 10 – Vínculos Empregatício segundo Ocupações em 2017



Fonte – Elaboração própria com base nos dados da RAIS, 2019.

Nesta seção foram apresentadas as estatísticas descritivas das variáveis explicativas que podem influenciar na probabilidade de ficar sem emprego. Conforme visto, a proporção de homens e mulheres com vínculos inativos é similar, mas ainda assim as mulheres possuem um menor percentual. Ao analisar por faixa de idade, observou-se que a massa da força de trabalho consiste nos indivíduos pertencentes a Faixa 3 (30-49 anos), possuindo alto percentual de ativos quanto de inativos. A variável raça/cor não apresentou diferenças notáveis entre os dois grupos analisados. Corroborando o que a literatura prediz, indivíduos com maior grau de escolaridade representam a menor parte dos inativos, quando comparados aos de menor grau de escolaridade. Por fim, indivíduos pertencentes ao acrônimo STEM no ano de 2017, correspondiam a um menor percentual de inativos, quando comparados aos Não-STEM. Na próxima seção, será examinado os resultados do quanto cada variável pode influenciar na probabilidade de um indivíduo ficar sem vínculo empregatício.

4 RESULTADOS ESTATÍSTICOS

Na economia brasileira foram fechados 20.832 postos de trabalho com carteira assinada em 2017, segundo o MTE. Os números do CAGED mostram que em 2017, as mulheres, em comparação aos homens, foram mais atingidas pelo desemprego. A diferença entre o número de demissões e contratações de mulheres foi negativa, em cerca de 42.526 -, enquanto a dos homens foi positiva, ou seja, o número de contratações para os homens superou o de demissões. Isso pode estar relacionado com a recuperação da economia. No momento em que a crise econômica atingiu o Brasil, as mulheres estavam com o saldo de desemprego "melhor" que o dos homens, mas em 2017, quando a recuperação já havia começado, ocorreu o aumento do número de vagas de setores ocupados tipicamente por homens.

O mercado de trabalho, ao iniciar sua recuperação, começou a melhorar de forma generalizada em todos os setores. O número de pessoas dispostas a trabalhar por um salário menor do que a sua qualificação exigida aumentou, acelerando o crescimento da população economicamente ativa. Isso aumentou a razão da população desocupada sobre a força de trabalho, podendo ter reduzido a probabilidade de permanência no desemprego.

Neste capítulo são apresentados e analisados os resultados do ajuste do modelo de probabilidade descrito na seção 3.3 pela Equação 3.6, bem como os efeitos marginais de cada variável explicativa. O capítulo está dividido em duas seções, sendo a primeira para detalhar os resultados do Brasil, e, a segunda, exibindo os resultados diferenciados para cada região brasileira. Os modelos foram estimados pelo *software* STATA e a apresentação dos *outputs* completos estão no APÊNDICE B.

4.1 ANÁLISE DO MODELO *LOGIT* PARA O BRASIL

Nesta seção são apresentados os resultados do modelo de probabilidade obtidos pelo ajuste dos dados ao modelo econométrico da Equação 3.6. Ao analisar os resultados do modelo, de acordo com a RAIS de 2017, pode-se inferir como cada característica do indivíduo pode influenciar na probabilidade de ele ficar sem emprego. O resultado do modelo, bem como o efeito marginal das variáveis, está na Tabela 2. Ao examinar a influência das variáveis na probabilidade de vínculo empregatício inativo, é considerada a interpretação através do efeito marginal.

Todas as variáveis do modelo foram estatisticamente significativas, tendo elas tanto significância individual, verificado pelo teste de *Wald*, quanto significância conjunta, verificado pela prob > chi2. Além desses testes, foi feito o teste de multicolinearidade FIV[1], em que não foi constatado a presença de multicolinearidade entre as variáveis.

[1] O resultado do teste FIV está no APÊNDICE C

A Tabela 2 resume os resultados do modelo *logit* para os dados do Brasil em 2017. Na coluna 1 encontram-se as variáveis utilizadas no modelo, bem como os testes realizados. Nas colunas 2 e 3 encontram-se, respectivamente, o resultado do modelo - em que os parâmetros não podem ser diretamente interpretados - e o desvio padrão de cada variável. Por último, nas colunas 4 e 5, encontram-se os efeitos marginais de cada variável e seus respectivos desvios padrão. A interpretação do modelo se embasa nos efeitos marginais.

Para a variável explicativa idade, espera-se que indivíduos mais jovens tenham maior probabilidade de ficarem desempregados, quando comparados ao grupo mais velho. Como a idade dos indivíduos foi separada por faixas, e estas são, por construção, binárias, compara-se elas em relação a faixa omitida - Faixa 5 -. As Faixas 1, 2 e 3 possuem maiores probabilidade de ficarem sem emprego, quando comparadas aos indivíduos da Faixa 5. Já a Faixa 4, possui 1,28 pontos percentuais a menos de chance de ficar desempregada, em relação a Faixa 5.

Para a variável explicativa Raça, constatou-se que no Brasil como um todo, o indivíduo declarado como "não branco" possui 1,14 pontos percentuais a menos de chance de perder o seu emprego, quando comparado ao indivíduo declarado como "branco". Este resultado quanto à variável Raça, pode ser devido à omissão de declaração dos indivíduos quanto a sua raça/cor - cerca de ¼ da amostra de vínculos ativos no ano de 2017.

O nível educacional do trabalhador está representado por *dummies* de faixas de escolaridade, logo devem ser interpretadas com relação à *dummy* omitida, neste caso, a de indivíduos que possuem escolaridade inferior ao ensino médio completo. Analisando os efeitos marginais da Tabela 2, infere-se que indivíduos que possuem ensino médio completo têm 2,6 pontos percentuais a menos de chance de ficarem desempregados; enquanto que indivíduos que possuem ensino superior completo têm 9,92 pontos percentuais a menos de possibilidade de ficarem desempregados; todos os resultados comparados em relação ao indivíduo base, ou seja, ao indivíduo que possui escolaridade inferior ao ensino médio completo. Dessa maneira, os resultados corroboram o que foi evidenciado na literatura, quanto maior o nível de escolaridade do indivíduo, menor a probabilidade de ele ficar desempregado.

Quanto à variável de ocupação, representada pela *dummy* STEM, os resultados se referem à comparação do grupo de indivíduos que atuam em ocupações STEM (*dummy* = 1) comparados aos demais indivíduos do mercado de trabalho (*dummy* = 0). Os resultados do modelo - apresentados pelos efeitos marginais na Tabela 2 - reforçam o que a literatura afirma. Como o coeficiente marginal gerado pelo modelo é -0,0150, pode-se concluir que os indivíduos que se encontram em ocupações do grupo STEM possuem 1,5 pontos percentuais a menos de probabilidade de ficarem desempregados, quando comparados aos indivíduos - de mesmas características base - Não-STEM.

A variável gênero está em concordância com o que tem sido relatado na literatura moderna, contestando as literaturas passadas, estas últimas, comprovavam que mulheres tinham maior probabilidade de ficarem desempregadas do que homens. Wall et al. (2009), em seu

paper sobre análise da recessão de 2008, constatou que após recessões as mulheres ficam menos desempregadas do que os homens. Na economia brasileira, foi evidenciado esta tendência mais moderna no ano de 2017, as mulheres tinham 1,3 pontos percentuais a menos de chance de perderem seus empregos em relação aos homens.

Tabela 2 – Resultado do Modelo *Logit* para o Brasil em 2017

Variáveis	**Modelo**		**Efeito Marginal**	
Desemprego	Coeficiente	Erro Padrão	dy/dx	Erro Padrão
Faixa 1	0,7881***	0,0016	0,1612***	0,0003
Faixa 2	0,6597***	0,0017	0,1349***	0,0003
Faixa 3	0,3113***	0,0016	0,0636***	0,0003
Faixa 4	-0,0627***	0,0017	-0,0128***	0,0004
Raça	-0,0560***	0,0006	-0,0114***	0,0001
Médio	-0,1271***	0,0007	-0,0260***	0,0001
Superior Incompleto	-0,2766***	0,0015	-0,0565***	0,0003
Superior	-0,4853***	0,0009	-0,0992***	0,0002
Pós Graduado	-0,7877***	0,0037	-0,1611***	0,0008
STEM	-0,0734***	0,0027	-0,0150***	0,0006
Gênero	-0,0631***	0,0006	-0,0130***	0,0001
Constante	-1,0292***	0,0016		
Log Likelihood Value		-38750040		
Pseudo R^2		0,0189		
Prob >chi2		0,0000		
Observações		64808575		

Fonte – Elaboração própria com base nos dados da RAIS, 2019.
Nota: * denota significância ao nível < 5%, ** denota significância ao nível < 1%, *** denota significância ao nível < 0,1%.

Ao examinar as variáveis para o Brasil no ano de 2017, percebe-se que todas estão em concordância com a literatura revisada, com exceção da variável Raça. Como existe uma heterogeneidade regional no Brasil, as variáveis podem mudar de interpretação para cada região. Por isso, na próxima seção serão analisados os resultados da influência das variáveis em perder o emprego, para cada região do Brasil, de forma a verificar se as regiões são afetadas diferentemente pelas características dos indivíduos.

4.2 ANÁLISE DO MODELO PARA AS REGIÕES BRASILEIRAS

Ao analisar os resultados do modelo, pode-se constatar como cada característica do indivíduo pode influenciar na probabilidade de ele ficar desempregado. A Tabela 3 expõe o efeito

marginal das variáveis, separado por cada região. O resultado completo do modelo, assim como os testes realizados para o modelo de cada região, estão dispostos no APÊNDICE B.

As variáveis do modelo foram estatisticamente significativas, com exceção da variável Raça na região Sul. A partir do teste de *Wald*, detectou-se a significância individual das variáveis do modelo, e pela prob > chi2, verificou-se a significância conjunta das variáveis de forma a explicar a variável dependente. Além desses testes, foi realizado o teste de multicolinearidade FIV[2], em que não foi constatado a presença de multicolinearidade entre as variáveis.

Ao examinar os efeitos marginais do modelo, percebe-se uma alta similaridade entre os resultados das regiões Sul e Sudeste, possivelmente explicada pela semelhança da composição do PIB de cada uma delas, composto principalmente por indústria e serviços. Ao analisar as variáveis correspondentes à faixas de idade, segue-se a mesma lógica anterior, pelo motivo de serem *dummies*, compara-se cada faixa de idade com àquela omitida. As faixas iniciais, representadas por 1, 2 e 3 possuem o mesmo comportamento, verificado pelo sinal, em todas as regiões, a diferença se dá pela magnitude. Esse comportamento, indica que quanto mais jovem for o indivíduo, maior é a probabilidade de ficar desempregado. Este resultado está de acordo com o *paper* de Araújo, Cunha e Lima (2011), estes analisaram e atestaram que, nas regiões metropolitanas brasileiras em 2007, os jovens possuíam maiores chances de ficarem desempregados, quando comparados aos mais velhos.

Constata-se que as regiões Norte e Centro Oeste são as que têm maior influência das variáveis relacionadas à idade, na probabilidade de ficar desempregado, quando comparado às outras regiões. Nessas regiões, o fato de um indivíduo ter entre 16 e 24 anos de idade (Faixa 1), aumenta em 20 pontos percentuais as chances de ficar desempregado, quando comparado àqueles que estão na Faixa 5.

Ao avaliar a variável Raça, percebe-se que na região Sul ela não teve significância e nas regiões Sudeste e Centro Oeste ela foi menos significante que as regiões Norte e Nordeste. No geral, o indivíduo declarado como "não branco" possui uma probabilidade menor de perder o emprego com relação ao indivíduo "branco", havendo exceção apenas na região Centro Oeste. O resultado auferido pela maior parte das regiões, contrapõe os trabalhos de Fernandes e Picchetti (1999), Oliveira, Scorzafave e Pazello (2009) e Araújo, Cunha e Lima (2011), estes encontraram a variável raça de forma significativa e com impacto positivo no desemprego. Para estes autores, um indivíduo classificado como "negro" teria maior chance de ficar desempregado. Porém, a amostra utilizada pelas três pesquisas, abrange somente as regiões metropolitanas do Brasil, podendo trazer resultados viesados para as variáveis.

Os resultados do efeito marginal para as faixas de escolaridade, por serem *dummies*, são interpretados da mesma forma das faixas de idade, em relação à *dummy* omitida, neste caso, em relação ao indivíduo que possui escolaridade inferior ao ensino médio completo. De acordo com os efeitos marginais, pode-se concluir aquilo auferido pela literatura, de que quanto maior o grau

[2]O resultado do teste FIV está no APÊNDICE C

de escolaridade do indivíduo, menor será a probabilidade de ele ficar desempregado. Barros, Camargo e Mendonça (1997) chamam essa relação entre desemprego e níveis educacionais de U-invertido, ou seja, as faixas de educação iniciais estão ligadas a um maior nível de desemprego, e, as faixas educacionais mais avançadas - como a de pós graduação - estão ligadas a um menor nível de desemprego. Os efeitos marginais, a partir da Tabela 3, mostram que as faixas de educação possuem maiores magnitudes nas regiões Norte, Nordeste e Sul. As regiões Sudeste e Centro Oeste possuem também resultados significativos, mas de menor magnitude. Assim, um indivíduo da Região Sul em 2017, com pós graduação, possuía 16,27 pontos percentuais a menos de chance de perder seu emprego, se comparado ao indivíduo com nível de escolaridade inferior ao ensino médio completo; em comparação à este, um indivíduo pós graduado da Região Centro Oeste, possuía em 2017, 11,46 pontos percentuais a menos de chance de perder seu emprego.

Ao analisar a variável gênero, percebe-se que sua interpretação é de igual forma para todas as regiões brasileiras. Pode-se afirmar então que as mulheres, no ano de 2017, possuíam uma probabilidade menor de perderem seus empregos, em relação aos homens. Sendo a variável de magnitude maior na região Nordeste, em que as mulheres possuem 3,15 pontos percentuais a menos de chance de ficarem desempregadas em comparação com os homens. Isso mostra uma reversão do resultado de diversos *papers*, que afirmavam e certificavam, que as mulheres possuíam maiores chances de perderem seus empregos, em relação aos homens. Atualmente, e conforme pode ser visto pelos resultados apontados, as mulheres possuem uma probabilidade menor que a dos homens, de ficarem desempregadas. Esta mudança do mercado de trabalho pode ser explicada pela oferta e demanda de determinados "serviços", em que somente mulheres se dispõe a fazer determinado trabalho, como, por exemplo, os "serviços gerais"; e principalmente quando a economia passa por uma recessão, estes "serviços" ainda são demandados, ocasionando assim, uma menor saída do mercado de trabalho das mulheres.

A variável de ocupação, representada pela *dummy* STEM, apresenta sinais diferentes entre as regiões. Nas regiões Sudeste, Sul e Centro Oeste, os indivíduos STEM possuem menor probabilidade de ficarem desempregados, em relação aos Não-STEM, sendo que na região Sul essa estatística é 3,84 pontos percentuais a menos de chance. Nas regiões Norte e Nordeste, os indivíduos que estão dentro do grupo STEM tem uma probabilidade maior de perderem seus empregos em relação ao grupo Não-STEM, isso pode ser possivelmente explicado pela baixa demanda de trabalhadores STEM nestas regiões. De acordo com Silva e Bonini (2019), as regiões brasileiras que mais empregaram trabalhadores em ocupações STEM em 2015 foram respectivamente o Sudeste, seguido pelo Nordeste, Sul, Centro Oeste e Norte. Estas autoras fizeram uma relação do prêmio salarial STEM - mencionado na seção 2.5.2 - para as regiões, e, curiosamente, as regiões Nordeste e Norte, possuíam respectivamente os segundo e terceiro maiores prêmio STEM no Brasil. Isso pode ser uma explicação para a variável ter sido positiva nestas regiões, representando que indivíduos STEM possuem maior probabilidade de desemprego, o prêmio é maior devido, possivelmente, ao menor número de indivíduos qualificados para as áreas STEM e, por conta disso, a probabilidade de perder o emprego tende a ser maior.

Tabela 3 – Efeitos Marginais do Modelo *Logit* para as Regiões do Brasil em 2017

	Efeito Marginal				
Variáveis	**dy/dx**				
Desemprego	Norte	Nordeste	Sudeste	Sul	Centro Oeste
Faixa 1	0,2192***	0,1886***	0,1605***	0,1631***	0,2072***
Faixa 2	0,1898***	0,1602***	0,1312***	0,1314***	0,1801***
Faixa 3	0,0973***	0,0745***	0,0539***	0,0518***	0,0889***
Faixa 4	0,0114***	-0,0062***	-0,0181***	-0,0197***	0,0029*
Raça	-0,0335***	-0,0144***	-0,0003*	-0,0001	0,0009*
Médio	-0,0334***	-0,0381***	-0,0181***	-0,0214***	-0,0284***
Superior Incompleto	-0,0567***	-0,0601***	-0,0419***	-0,0647***	-0,0658***
Superior	-0,1263***	-0,1036***	-0,0840***	-0,1009***	-0,0807***
Pós Graduado	-0,1586***	-0,1450***	-0,1189***	-0,1627***	-0,1146***
STEM	0,0111***	0,0095***	-0,0177***	-0,0384***	-0,0168***
Gênero	-0,0224***	-0,0315***	-0,0051***	-0,0112***	-0,0204***

Fonte – Elaboração própria com base nos dados da RAIS, 2019.
Nota: * denota significância ao nível < 5%, ** denota significância ao nível < 1%, *** denota significância ao nível < 0,1%.

Nas regiões em que a presença dos STEM é maior, como o Sudeste, Sul e Centro Oeste, esta pesquisa gerou resultados que estão em linha com o que foi revisado na literatura, de que os trabalhadores STEM possuem maior estabilidade no mercado de trabalho, sendo menos demitidos, e, por conseguinte, tendo menor probabilidade de ficarem desempregados. Assim, como afirma Noonan (2017), que fez um estudo acerca de trabalhadores STEM e Não-STEM nos Estados Unidos, com relação à taxa de desemprego no período 1994-2015, além da força de trabalho STEM possuírem ganhos salariais mais altos que os Não-STEM, eles também experimentam, em média, taxas de desemprego mais baixas que estes últimos trabalhadores. Silva e Bonini (2019) também encontraram estes resultados, de ganhos salariais maiores para os STEM, em relação aos Não-STEM, para todas as regiões brasileiras no ano de 2015, sendo os prêmios STEM de forma diferenciada em cada região.

Como foi observado, as regiões brasileiras apresentaram características distintas para algumas variáveis, como as relacionadas à idade, raça e para aqueles indivíduos pertencentes às ocupações STEM. Isso pode ser explicado pela heterogeneidade e peculiaridade de cada região. Quando acontece algo inesperado na economia brasileira, como, por exemplo, uma recessão, as regiões Norte e Nordeste são, geralmente, as mais afetadas, especialmente no âmbito do mercado de trabalho. As regiões Sul, Sudeste e Centro Oeste também sofrem com os choques econômicos, mas conseguem suavizar eles de forma mais rápida. Essa heterogeneidade do mercado de trabalho está diretamente ligada às atividades econômicas desenvolvidas em cada região. Por isso, é de extrema importância conhecer as especificidades de cada região, para que políticas públicas possam ser empregadas de modo mais eficaz no combate ao desemprego, à luz da heterogeneidade do mercado de trabalho.

5 CONSIDERAÇÕES FINAIS

O principal objetivo deste trabalho foi analisar como determinadas características individuais podem influenciar na probabilidade de um trabalhador brasileiro - que esteja no mercado de trabalho formal - perder o emprego, em âmbito nacional e regional. Foi utilizada a base de dados advinda da RAIS, para o ano de 2017. As características consideradas determinantes do desemprego, foram escolhidas com base na literatura pertinente, são elas: idade, gênero, escolaridade, cor ou raça e ocupação. Esta última, foi utilizada para investigar se o fato do trabalhador atuar numa ocupação pertencente às áreas de STEM, diminui suas chances de ficar desempregado. A caracterização da variável *dummy* de ocupação STEM foi feita com base na Classificação Brasileira de Ocupações (CBO). Assim, foi possível analisar se indivíduos dentro da força de trabalho STEM possuíam menor probabilidade de ficar desempregado.

No capítulo 2, foi apresentado o arcabouço teórico que, inicialmente, fez uma revisão acerca da modelagem de mercado de trabalho, seguido por conceitos relacionados ao desemprego, e posteriormente, a fundamentação teórica para a escolha das variáveis utilizadas na explicação do desemprego. Ainda neste capítulo, foi feito uma breve revisão sobre as concepções modernas do mercado de trabalho, a respeito do modelo DMP de desemprego - que conquistou prêmio Nobel em 2010 - e sobre a força de trabalho nas áreas de ciência e tecnologia - STEM.

No capítulo 3, foi descrito o procedimento metodológico, utilizado para atingir os objetivos propostos do trabalho, que consistiu no uso de um modelo de resposta qualitativa *logit*. A variável de interesse, que representa a *proxy* de desemprego, foi a de ”vínculo empregatício”. As variáveis explicativas utilizadas, conforme mencionando no início deste capítulo, foram: idade, escolaridade, cor ou raça, gênero e, uma variável utilizada para representar as ocupações dos indivíduos, STEM. As variáveis de idade e escolaridade foram distribuídas em faixas, para medir como cada faixa impacta na probabilidade de desemprego, comparado com a faixa base.

A análise descritiva das variáveis explicativas, observando, por exemplo, o número de empregos formais regionalmente, percebe-se uma distribuição desigual do número de empregos ao longo do território brasileiro. Constatou-se, ao longo da análise dos resultados encontrados, que as regiões brasileiras são muito heterogêneas. Além disso, ao relacionar o PIB das regiões, verificou-se que as regiões que possuem em seu PIB regional fatores que têm maior influência no PIB nacional, possuem uma relação menos agressiva com o desemprego, quando comparado às outras regiões.

Ainda no capítulo 3, analisando as estatísticas descritivas para o Brasil, no ano de 2017, verificou-se, para as variáveis utilizadas que existe uma proporção similar de homens e mulheres com vínculos inativos, além disso, a faixa de idade entre 30-49 anos é a base do mercado de trabalho formal brasileiro, possuindo tanto número maior de vínculos ativos, quanto de

inativos. Constatou-se ainda que, indivíduos que possuem até ensino médio completo são os que têm maior número de vínculos inativos. Quando observado a variável de raça/cor, não se verificou diferença significativa no percentual de vínculos, já em relação à variável relacionada às ocupações, constatou-se uma "inclinação" do Brasil rumo às economias desenvolvidas, pois o percentual de brasileiros STEM inativos é, consideravelmente menor que o grupo Não-STEM.

No capítulo 4, foram analisados os resultados da análise estatística. O ajuste do modelo aos dados para o Brasil, resultou que todas as variáveis foram significativas para explicar a probabilidade de desemprego. Confirmando o que foi visto na literatura, indivíduos mais jovens possuem maior probabilidade de perder o emprego, assim como indivíduos que possuem menor nível de escolaridade. Além disso, tanto mulheres quanto indivíduos pertencentes ao grupo ocupacional STEM, possuem chances menores de ficar desempregados, quando comparados aos seus grupos base.

A análise dos dados recortados regionalmente, indicou a presença da heterogeneidade entre as regiões brasileiras. Os resultados indicaram uma alta similaridade entre as regiões Sul e Sudeste, que pode ser possivelmente explicada pela composição semelhante do PIB. As variáveis relacionadas à idade, têm maior influência na probabilidade de desemprego nas regiões Norte e Centro Oeste, sendo que indivíduos mais jovens possuem maior probabilidade de perderem seus empregos. A variável raça se mostrou mais significativa nas regiões Norte e Nordeste, sendo que apresentou menor significância nas regiões Sudeste e Centro Oeste, e, nenhuma significância na região Sul. O resultado das variáveis referentes à escolaridade, comprovaram o que a literatura aborda, que quanto maior o nível de escolaridade do indivíduo, menor será a probabilidade de ele ficar desempregado. Ao analisar a variável gênero, percebeu-se a reversão apresentada em diversos trabalhos, pois foi possível demonstrar que mulheres possuem uma menor chance de perderem seus empregos, em relação aos homens.

Além disso, ao analisar a variável que representa as ocupações, STEM, verificou-se um comportamento heterogêneo ao longo das regiões brasileiras. Nas regiões Sul, Sudeste e Centro Oeste, confirmou-se aquilo evidenciado nas economias modernas, de que indivíduos empregados em ocupações STEM, possuem menor probabilidade de ficarem desempregados. Mas ao fazer esta análise para as regiões Norte e Nordeste, ocorreu um resultado positivo, demonstrando que trabalhadores STEM possuem maior probabilidade de perderem seus empregos. Este último resultado pode ser explicado pelo elevado prêmio STEM que estas regiões possuem, e, este prêmio é maior devido, possivelmente, à menor disponibilidade de indivíduos qualificados para trabalhar nas áreas de ciência e tecnologia nas regiões Norte e Nordeste.

Assim, observou-se que a variável desemprego é afetada diferentemente por cada característica do indivíduo, além de, afetar cada região brasileira de forma distinta. Deste modo, o uso de políticas públicas seria mais eficiente se fossem analisadas as heterogeneidades regionais aparentes, de forma a entender a especificidade de cada região e, assim, reduzir o desemprego ao longo do território brasileiro.

Referências

AKERLOF, G. A. Labor contracts as partial gift exchange. **The quarterly journal of economics**, MIT Press, v. 97, n. 4, p. 543–569, 1982.

ALVES FILHO, O. S. **A curva de salário para a Região Metropolitana de Salvador:** uma análise econométrica a partir dos dados da PED de 1997 a 2003. Bahia, Brasil, 2004. Dissertação (Economia), UFBA (Universidade Federal da Bahia).

ARAÚJO, A. A.; CUNHA, D. A.; LIMA, J.E. Determinantes do desemprego e inatividade de jovens no Brasil metropolitano. **Revista de Economia e Agronegócio**, v. 9, n. 3, p. 369–392, 2011.

BACHMAN, R. **Determinantes da probabilidade do desemprego no Paraná:** uma aplicação do modelo logit (2001 e 2014). Paraná, Brasil, 2017. Monografia (Economia), UEM (Universidade Estadual de Maringá).

BARANYAI, K. et al. **Australia's STEM workforce**: science, technology, engineering and mathematics. [S.l.]: Australian Government, Canberra, Australian Capital Territory, 2016.

BARROS, R. P.; CAMARGO, J. M.; MENDONÇA, R. **A estrutura do desemprego no Brasil**. Rio de Janeiro: IPEA, 1997. Texto para discussão n. 478.

BARUFI, A. M. B. **Agglomeration economies in the Brazilian formal labor market**. University of São Paulo TD Nereus, São Paulo, p. 03–2014, 2014.

BLACKLEY, S.; HOWELL, J. A STEM Narrative: 15 Years in the Making. **Australian Journal of Teacher Education**, v. 40, n. 7, p. 8, 2015.

CORRAR, L. et al. **Análise multivariada para os cursos de administração, ciências contábeis e economia**. São Paulo: Atlas, 2011.

COSTA, J. S.; CUNHA, M. S. Determinantes do desemprego no Brasil no período de 1981 a 2005: uma análise enfatizando a qualificação do indivıduo em um contexto de maior abertura comercial. **Análise Econômica**, Porto Alegre, v. 28, n. 53, p. 197–220, 2010.

DAVIES, P. Getting to the root of STEM. **Regional Business & Economics Newspaper**, Minneapolis, p. 1–8, 2014. Disponível em: <https://www.minneapolisfed.org/~/media/files/pubs/fedgaz/1404/stem_fedgazette_april2014.pdf>.

DIAMOND, P. A. A model of price adjustment. **Journal of Economic Theory**, Academic Press, v. 3, n. 2, p. 156–168, 1971.

DIEESE. Departamento Intersindical de Estatística e Estudos Socioeconômicos. **Os negros nos mercados de trabalho metropolitanos**. [S.l.]: Pesquisa de Emprego e Desemprego (PED), 2015.

DORNBUSCH, R.; FISCHER, S.; STARTZ, R. **Macroeconomia, 10ª edição**. [S.l.]: São Paulo: McGraw-Hill, 2013.

FERNANDES, R.; PICCHETTI, P. Uma análise da estrutura do desemprego e da inatividade no Brasil metropolitano. **Pesquisa e Planejamento Econômico**, Instituto de Pesquisa Econômica Aplicada (IPEA), Rio de Janeiro, v. 29, n. 1, p. 87–112, 1999.

FRIEDMAN, M. The Role of Monetary Policy. **The American Economic Review**, v. 58, n. 1, p. 102–110, 1968.

GIL, A. C. **Como elaborar projetos de pesquisa**. São Paulo: Atlas, v. 5, 2002.

GUJARATI, D. N.; PORTER, D. C. **Econometria Básica**. 5. ed. [S.l.]: Porto Alegre: AMGH Editora, 2011.

HICKS, J. R.; HANSEN, A. Mr. Keynes and the"classics"; a suggested interpretation. **Econometrica: journal of the Econometric Society**, JSTOR, p. 147–159, 1937.

IBGE. **Instituto Brasileiro de Geografia e Estatística**. [S.l.: s.n.], 2019. Disponível em: <http://www.ibge.gov.br>. Acesso em: 30 out. 2019.

KEYNES, J. M. **A teoria geral do emprego, do juro e da moeda [1936]**. [S.l.]: São Paulo: Atlas, 1982.

KON, A. Pleno emprego no Brasil: interpretando os conceitos e indicadores. **Revista Economia & Tecnologia**, v. 8, n. 2, 2012.

LANDIVAR, L. C. Disparities in STEM employment by sex, race, and hispanic origin. **Education Review**, v. 29, n. 6, p. 911–922, 2013.

LANGDON, D. et al. Women in STEM: A gender gap to innovation. **Economics and Statistics Administration Issue Brief**, p. 4–11, 2011.

LARRAIN, F.; SACHS, J. **Macroeconomia em uma economia global**. [S.l.]: São Paulo: Pearson Education do Brasil, 2000.

MANKIW, N. G. **Macroeconomia**, 8ª edição. São Paulo: LTC, 2015.

MATTEO, M. **Heterogeneidade regional**. CEPAL-IPEA, 2013. Texto para discussão.

MONGE-NARANJO, A.; SOHAIL, F. Age and gender differences in long-term unemployment: before and after the great recession. **Economic synopses**, n. 26, 2015.

MORTENSEN, D. T. Job search, the duration of unemployment, and the Phillips curve. **The American Economic Review**, JSTOR, v. 60, n. 5, p. 847–862, 1970.

NOONAN, R. STEM Jobs: 2017 Update. ESA Issue Brief# 02-17. **US Department of Commerce**, ERIC, 2017.

OKUN, A. M. Potential GNP: Its Measurement and Significance In Proceedings of the Business and Economics Statistics Section. **American Statistical Association**, 1962.

OLIVEIRA, P. R.; SCORZAFAVE, L. G.; PAZELLO, E. T. Desemprego e inatividade nas metrópoles brasileiras: as diferenças entre homens e mulheres. **Nova economia**, SciELO Brasil, v. 19, n. 2, p. 291–324, 2009.

PHELPS, E. S. **Microeconomic foundations of employment and inflation theory**. Norton, 1970.

PHILLIPS, A. W. The relation between unemployment and the rate of change of money wage rates in the United Kingdom, 1861–1957. **Economica**, Wiley Online Library, v. 25, n. 100, p. 283–299, 1958.

PIGOU, A. C. **Theory of unemployment**. New York: Augustus M. Kelley, 1968. (Reprints of Economic Classics).

PINO, F. A. Modelos de decisão binários: uma revisão. **Rev. de Economia Agrıcola**, v. 54, n. 1, p. 43–57, 2007.

PROBST, E. R.; RAMOS, P. A evolução da mulher no mercado de trabalho. **Instituto Catarinense de Pós-Graduação**, Santa Catarina, p. 1–8, 2003.

RAMOS, L.; VIEIRA, M. L. **Desigualdade de rendimentos no Brasil nas décadas de 80 e 90:** evolução e principais determinantes. Rio de Janeiro: IPEA, 2001. Texto para discussão n. 803.

SAKOWSKI, P. Measuring employment in the tourism industries in Brazil: from national to regional and local level. In: **UNWTO/INROUTE 1ST SEMINAR ON REGIONAL TOURISM, VENICE**. [S.l.: s.n.], 2012.

SAMARY, A. P. F. **Crescimento econômico e desemprego**: uma estimativa da Lei de Okun pós Plano Real. 2017. Tese (Doutorado).

SAY, J. B. **A Treatise on Political Economy**, 3 rd US Edition. Philadelphia: John Grigg, 1803.

SEEMANN, M. W.; BONINI, P. **Trabalho STEM no Brasil de acordo com a CBO**. [S.l.]: Seminário de Iniciação Científica UDESC, 2017. Disponível em: <https://www.udesc.br/arquivos/udesc/id_cpmenu/6217/TRABALHO_STEM_NO_BRASIL_DE_ACORDO_COM_A_CBO_15033981416899_6217.pdf>.

SILVA, F.; BONINI, P. **STEM nas Regiões Brasileiras**. [S.l.]: Seminário de Iniciação Científica UDESC, 2019. Disponível em: <https://www.udesc.br/arquivos/udesc/id_cpmenu/10577/31_15657187077558_10577.pdf>.

TANGTIPONGKUL, K.; WANGMO, D. Determinants of Unemployment: Characteristics and Policy Responses in Bhutan. **Southeast Asian Journal of Economics**, v. 5, n. 2, p. 27–48, 2017.

VIEIRA, P. V. **Análise da aplicação da Curva de Phillips e da Lei de Okun para a economia brasileira no perıodo de 2002 a 2014**. Brasília, Brasil, 2014. Monografia (Economia), UNB (Universidade de Brasília).

WALL, H. et al. The "man-cession"of 2008-2009: It's big, but it's not great. **The Regional Economist**, Federal Reserve Bank of St. Louis, v. 18, n. 4, p. 4–9, 2009.

ZÁRATE, D. Nobel 2010: fricciones y costos de relación entre los agentes de mercado. **Equilibrios y Conjeturas**, p. 33–40, 2014.

ZYLBERSTAJN, H.; BALBINOTTO, G. N. As teorias de desemprego e as polıticas públicas de emprego. **Estudos Econômicos (São Paulo)**, v. 29, n. 1, p. 129–149, 1999.

APÊNDICE A – Ocupações STEM de acordo com a CBO

Tabela 4 – Códigos CBO considerados no grupo STEM

Família CBO	Descrição Família CBO	Códigos
1425	Gerentes de Tecnologia da Informação	142505, 142510, 142515, 142520, 142525, 142530, 142535
1426	Gerentes de Pesquisa e Desenvolvimento e Afins	142605, 142610
2011	Profissionais da Biotecnologia	201105, 201110, 201115
2012	Profissionais da Metrologia	201205, 201210, 201215, 201220, 201225
2021	Engenheiros de Controle e Automação, Engenheiros Mecatrônicos e Afins	202105, 202110, 202115, 202120
2030	Pesquisadores das Ciências Biológicas	203005, 203010, 203015, 203020, 203025
2031	Pesquisadores das Ciências Naturais e Exatas	203105, 203110, 203115, 203120, 203125
2032	Pesquisadores de Engenharia e Tecnologia	203205, 203210, 203215, 203220, 203225, 203230
2034	Pesquisadores das Ciências da Agricultura	203405, 203410, 203415, 203420
2111	Profissionais da Matemática	211105, 211110, 211115, 211120
2112	Profissionais de Estatística	211205, 211210, 211215
2122	Engenheiros em Computação	212205, 212210, 212215
2123	Administradores de Tecnologia da Informação	212305, 212310, 212315, 212320
2124	Analistas de Tecnologia da Informação	212405, 212410, 212415, 212450
2131	Físicos	213105, 213110, 213115, 213120, 213125, 213130, 213135, 213140, 213145, 213150, 213155, 213160, 213165, 213170, 213175
2132	Químicos	213205, 213210, 213215
2133	Profissionais das Ciências Atmosféricas e Espaciais e de Astronomia	213305, 213310, 213315
2134	Geólogos, Oceanográficos, Geofísicos e Afins	213405, 213410, 213415, 213420, 213425, 213430, 213435, 213440
2140	Engenheiros Ambientais e Afins	214005, 214010
2142	Engenheiros Civis e Afins	214205, 214210, 214215, 214220, 214225, 214230, 214235, 214240, 214245, 214250, 214255, 214260, 214265, 214270, 214275, 214280
2143	Engenheiros Eletricistas, Eletrônicos e Afins	214305, 214310, 214315, 214320, 214325, 214330, 214335, 214340, 214345,214350, 214355, 214360, 214365, 214370
2144	Engenheiros Mecânicos e Afins	214405, 214410, 214415, 214420, 214425, 214430, 214435
2145	Engenheiros Químicos e Afins	214505, 214510, 214515, 214520, 214525, 214530, 214535
2146	Engenheiros Metalurgistas, de Materiais e Afins	214605, 214610
2147	Engenheiros de Minas e Afins	214705, 214710, 214715, 214720, 214725, 214730, 214735, 214740, 214745,214750
2148	Engenheiros Agrimensores e Engenheiros Cartógrafos	214805, 214810
2149	Engenheiros de Produção, Qualidade, Segurança e Afins	214905, 214910, 214915, 214920, 214925, 214930, 214935, 214940
2211	Biólogos e Afins	221105
2212	Biomédicos	221205
2221	Engenheiros Agrossilvipecuário	222105, 222110, 222115, 222120
2222	Engenheiros de Alimentos e Afins	222205, 222215

Fonte – Elaboração própria com base nas informações da CBO e com base em Seemann e Bonini (2017).

APÊNDICE B – Resultado das Regressões

Tabela 5 – Resultado do Modelo *Logit* para a Região Norte em 2017

Variáveis	**Modelo**		**Efeito Marginal**	
Desemprego	Coeficiente	Erro Padrão	dy/dx	Erro Padrão
Faixa 1	1,0100***	0,0077	0,2192***	0,0017
Faixa 2	0,8906***	0,0077	0,1898***	0,0017
Faixa 3	0,5177***	0,0073	0,0973***	0,0014
Faixa 4	0,0584***	0,0082	0,0114***	0,0016
Raça	-0,1683***	0,0035	-0,0335***	0,0007
Médio	-0,1723***	0,0029	-0,0334***	0,0006
Superior Incompleto	-0,3125***	0,0075	-0,0567***	0,0012
Superior	-0,7162***	0,0040	-0,1263***	0,0006
Pós Graduado	-1,0617***	0,0165	-0,1586***	0,0017
STEM	0,0569***	0,0158	0,0111***	0,0031
Gênero	-0,1158***	0,0025	-0,0224***	0,0005
Constante	-1,1184***	0,0079		
Log Likelihood Value		-2058115,6		
Pseudo R^2		0,0294		
Prob >chi2		0,0000		
Observações		3586240		

Fonte – Elaboração própria com base nos dados da RAIS, 2019.
Nota: * denota significância ao nível < 5%, ** denota significância ao nível < 1%, *** denota significância ao nível < 0,1%.

Tabela 6 – Resultado do Modelo *Logit* para a Região Nordeste em 2017

Variáveis	**Modelo**		**Efeito Marginal**	
Desemprego	Coeficiente	Erro Padrão	dy/dx	Erro Padrão
Faixa 1	0,8908***	0,0043	0,1886***	0,0009
Faixa 2	0,7730***	0,0043	0,1602***	0,0009
Faixa 3	0,4053***	0,0041	0,0745***	0,0007
Faixa 4	-0,0334***	0,0045	-0,0062***	0,0008
Raça	-0,0762***	0,0018	-0,0144***	0,0003
Médio	-0,2029***	0,0016	-0,0381***	0,0003
Superior Incompleto	-0,3463***	0,0040	-0,0601***	0,0006
Superior	-0,6055***	0,0023	-0,1036***	0,0004
Pós Graduado	-1,0025***	0,0096	-0,1450***	0,0009
STEM	0,0503***	0,0090	0,0095***	0,0017
Gênero	-0,1684***	0,0014	-0,0315***	0,0003
Constante	-1,1580***	0,0044		
Log Likelihood Value		-6387552,9		
Pseudo R^2		0,0235		
Prob >chi2		0,0000		
Observações		11390070		

Fonte – Elaboração própria com base nos dados da RAIS, 2019.
Nota: * denota significância ao nível < 5%, ** denota significância ao nível < 1%, *** denota significância ao nível < 0,1%.

Tabela 7 – Resultado do Modelo *Logit* para a Região Sudeste em 2017

Variáveis	**Modelo**		**Efeito Marginal**	
Desemprego	Coeficiente	Erro Padrão	dy/dx	Erro Padrão
Faixa 1	0,7095***	0,0023	0,1605***	0,0005
Faixa 2	0,5963***	0,0023	0,1312***	0,0005
Faixa 3	0,2639***	0,0021	0,0539***	0,0004
Faixa 4	-0,0890***	0,0024	-0,0181***	0,0005
Raça	-0,0018*	0,0008	-0,0003*	0,0002
Médio	-0,0879***	0,0009	-0,0181***	0,0002
Superior Incompleto	-0,2117***	0,0021	-0,0419***	0,0004
Superior	-0,4291***	0,0013	-0,0840***	0,0003
Pós Graduado	-0,6687***	0,0051	-0,1189***	0,0007
STEM	-0,0873***	0,0034	-0,0177***	0,0007
Gênero	-0,0251***	0,0008	-0,0051***	0,0002
Constante	-1,0398***	0,0021		
Log Likelihood Value		-19366904		
Pseudo R^2		0,0163		
Prob >chi2		0,0000		
Observações		32177865		

Fonte – Elaboração própria com base nos dados da RAIS, 2019.
Nota: * denota significância ao nível < 5%, ** denota significância ao nível < 1%, *** denota significância ao nível < 0,1%.

Tabela 8 – Resultado do Modelo *Logit* para a Região Sul em 2017

Variáveis	**Modelo**		**Efeito Marginal**	
Desemprego	Coeficiente	Erro Padrão	dy/dx	Erro Padrão
Faixa 1	0,7220***	0,0038	0,1631***	0,0008
Faixa 2	0,5882***	0,0039	0,1314***	0,0008
Faixa 3	0,2471***	0,0037	0,0518***	0,0008
Faixa 4	-0,0946***	0,0041	-0,0197***	0,0008
Raça	-0,0005	0,0013	-0,0001	0,0002
Médio	-0,1014***	0,0015	-0,0214***	0,0003
Superior Incompleto	-0,3238***	0,0031	-0,0647***	0,0006
Superior	-0,5078***	0,0021	-0,1009***	0,0004
Pós Graduado	-0,9463***	0,0095	-0,1627***	0,0012
STEM	-0,1880***	0,0071	-0,0384***	0,0014
Gênero	-0,0531***	0,0013	-0,0112***	0,0002
Constante	-0,9437***	0,0037		
Log Likelihood Value		-7231774,4		
Pseudo R^2		0,0192		
Prob >chi2		0,0000		
Observações		11798893		

Fonte – Elaboração própria com base nos dados da RAIS, 2019.
Nota: * denota significância ao nível < 5%, ** denota significância ao nível < 1%, *** denota significância ao nível < 0,1%.

Tabela 9 – Resultado do Modelo *Logit* para a Região Centro Oeste em 2017

Variáveis	**Modelo**		**Efeito Marginal**	
Desemprego	Coeficiente	Erro Padrão	dy/dx	Erro Padrão
Faixa 1	0,8987***	0,0057	0,2072***	0,0013
Faixa 2	0,7887***	0,0057	0,1801***	0,0013
Faixa 3	0,4225***	0,0054	0,0889***	0,0011
Faixa 4	0,0136*	0,0060	0,0029*	0,0013
Raça	0,0041*	0,0020	0,0009*	0,0004
Médio	-0,1323***	0,0021	-0,0284***	0,0005
Superior Incompleto	-0,3227***	0,0049	-0,0658***	0,0009
Superior	-0,3893***	0,0028	-0,0807***	0,0006
Pós Graduado	-0,5976***	0,0107	-0,1146***	0,0017
STEM	-0,0791***	0,0090	-0,0168***	0,0018
Gênero	-0,0950***	0,0019	-0,0204***	0,0004
Constante	-1,0372***	0,0056		
Log Likelihood Value		-3637836		
Pseudo R^2		0,0189		
Prob >chi2		0,0000		
Observações		5855507		

Fonte – Elaboração própria com base nos dados da RAIS, 2019.
Nota: * denota significância ao nível < 5%, ** denota significância ao nível < 1%, *** denota significância ao nível < 0,1%.

APÊNDICE C – Testes

Tabela 10 – Teste de Multicolinearidade para os dados do Brasil

Variável	FIV	1/FIV
Faixa 3	3,13	0,32
Médio	2,84	0,35
Raça	2,23	0,45
Gênero	1,85	0,54
Superio Completo	1,82	0,55
Faixa 2	1,68	0,59
Faixa 1	1,67	0,60
Faixa 4	1,43	0,70
Superior Incompleto	1,16	0,86
STEM	1,07	0,94
Pós-Graduado	1,03	0,97
FIV Médio	**1,81**	

Fonte – Elaboração própria com base no *software* STATA, 2019.

Tabela 11 – Teste de Multicolinearidade para os dados da Região Norte

Variável	FIV	1/FIV
Faixa 3	5,27	0,19
Médio	3,14	0,32
Raça	6,43	0,16
Gênero	1,83	0,55
Superior Completo	1,95	0,51
Faixa 2	2,22	0,45
Faixa 1	2,12	0,47
Faixa 4	1,86	0,54
Superior Incompleto	1,13	0,89
STEM	1,03	0,97
Pós-Graduado	1,04	0,96
FIV Médio	**2,55**	

Fonte – Elaboração própria com base no *software* STATA, 2019.

Tabela 12 – Teste de Multicolinearidade para os dados da Região Nordeste

Variável	**FIV**	**1/FIV**
Faixa 3	4,53	0,22
Médio	3,11	0,32
Raça	5,07	0,20
Gênero	1,86	0,54
Superior Completo	1,83	0,55
Faixa 2	2,02	0,49
Faixa 1	1,91	0,52
Faixa 4	1,77	0,56
Superior Incompleto	1,15	0,87
STEM	1,03	0,97
Pós-Graduado	1,04	0,96
FIV Médio	**2,30**	

Fonte – Elaboração própria com base no *software* STATA, 2019.

Tabela 13 – Teste de Multicolinearidade para os dados da Região Sudeste

Variável	**FIV**	**1/FIV**
Faixa 3	3,07	0,33
Médio	2,89	0,35
Raça	1,92	0,52
Gênero	1,85	0,54
Superior Completo	1,85	0,54
Faixa 2	1,68	0,60
Faixa 1	1,68	0,60
Faixa 4	1,41	0,71
Superior Incompleto	1,16	0,86
STEM	1,09	0,92
Pós-Graduado	1,03	0,97
FIV Médio	**1,78**	

Fonte – Elaboração própria com base no *software* STATA, 2019.

Tabela 14 – Teste de Multicolinearidade para os dados da Região Sul

Variável	**FIV**	**1/FIV**
Faixa 3	2,58	0,39
Médio	2,55	0,39
Raça	1,55	0,65
Gênero	1,88	0,53
Superior Completo	1,72	0,58
Faixa 2	1,55	0,64
Faixa 1	1,58	0,63
Faixa 4	1,31	0,76
Superior Incompleto	1,18	0,85
STEM	1,06	0,94
Pós-Graduado	1,03	0,97
FIV Médio	**1,64**	

Fonte – Elaboração própria com base no *software* STATA, 2019.

Tabela 15 – Teste de Multicolinearidade para os dados da Região Centro Oeste

Variável	**FIV**	**1/FIV**
Faixa 3	3,52	0,28
Médio	2,64	0,38
Raça	3,37	0,3
Gênero	1,83	0,55
Superior Completo	1,88	0,53
Faixa 2	1,81	0,55
Faixa 1	1,84	0,54
Faixa 4	1,50	0,66
Superior Incompleto	1,15	0,87
STEM	1,06	0,95
Pós-Graduado	1,04	0,96
FIV Médio	**1,97**	

Fonte – Elaboração própria com base no *software* STATA, 2019.