**FATORES DETERMINANTES DA PERMANÊNCIA NO EMPREGO FORMAL NO SUL DO BRASIL: Uma Análise Empírica com Modelos Logit e Probit**

Miguel das Neves Lopes[[1]](#footnote-0)

Eduardo Andre Tillmann[[2]](#footnote-1)

Gibran

**Resumo:** Este estudo investiga os fatores determinantes da permanência no emprego formal na região Sul do Brasil, utilizando dados da RAIS e modelos Logit e Probit. Os resultados indicam que variáveis individuais, como idade e escolaridade, têm um impacto positivo e significativo na probabilidade de permanência no emprego. A idade, representada em grupos quinquenais, apresentou uma relação linear crescente, enquanto níveis mais elevados de escolaridade elevaram as chances de permanência. Em contrapartida, o trabalho intermitente e vínculos empregatícios precários reduziram substancialmente a probabilidade de retenção no emprego, evidenciando o papel adverso de condições ocupacionais vulneráveis. Características organizacionais, como o tamanho do estabelecimento, apresentaram impacto positivo relevante, sugerindo que empresas maiores oferecem maior estabilidade laboral. A faixa de remuneração também contribuiu positivamente, porém com efeito marginal relativamente pequeno. Os modelos apresentaram desempenho preditivo robusto, com cerca de 70,8% de acerto geral, reforçando a adequação das variáveis explicativas. Os resultados sugerem a necessidade de políticas públicas que fortaleçam as relações de trabalho, reduzam as desigualdades ocupacionais e incentivem o investimento em educação e qualificação profissional.

**Palavras-chave:** Permanência no emprego, mercado de trabalho, trabalho intermitente, estabilidade laboral, modelos Logit e Probit..

**Abstract:** This study investigates the determinants of job retention in the formal labor market in the South of Brazil using RAIS data and econometric Logit and Probit models. The results show that individual variables, such as age and education, have a positive and significant impact on the probability of job retention. Age, represented in five-year groups, exhibited a linear increasing relationship, while higher education levels enhanced the likelihood of retention. Conversely, intermittent work and precarious employment contracts significantly reduced the probability of job retention, highlighting the adverse effects of vulnerable occupational conditions. Organizational characteristics, such as firm size, demonstrated a substantial positive impact, suggesting that larger firms provide greater job stability. Additionally, wage levels had a positive yet relatively small marginal effect. Both models demonstrated robust predictive performance, with an overall accuracy of 70.8%, confirming the relevance of the explanatory variables. The findings underscore the importance of public policies aimed at strengthening labor relations, reducing occupational inequalities, and fostering investments in education and professional training to improve job stability.

**Keywords:** Job retention, labor market, intermittent work, employment stability, Logit and Probit models.

**Classificação JEL:** J21, J63, J71, C25.

**Área temática:** Microeconometria.

# INTRODUÇÃO

A relação entre trabalho e educação representa uma correlação forte em senso comum. Afinal, um indivíduo não pode desempenhar uma função sem antes adquirir o conhecimento necessário para realizá-la, seja por meio da instrução formal ou do aprendizado prático. Saviani (2007) argumenta que, nas sociedades primitivas, a educação estava intrinsecamente ligada ao processo produtivo, tornando aprender e produzir atividades indissociáveis e essenciais para a sobrevivência coletiva. No entanto, com o advento da propriedade privada e a divisão do trabalho, ocorreu uma cisão histórica entre essas esferas, resultando em desigualdades no acesso à formação. Enquanto uma parte da população era direcionada para a educação voltada ao trabalho manual, outra recebia instrução para funções intelectuais, reforçando a estratificação social (Saviani, 2007, p. 154-157).

Nessa perspectiva, Becker (1994), trazendo a teoria do capital humano, enfatiza que a formação educacional não apenas aumenta os ganhos individuais, mas também eleva a produtividade ao fornecer conhecimento, habilidades e ferramentas analíticas fundamentais (Becker, 1994, p. 19). O autor destaca que, mesmo após a conclusão do ensino formal, a necessidade de treinamento persiste, seja por meio de programas estruturados ou do aprendizado informal no ambiente de trabalho (Becker, 1994, p. 20). Além disso, a desigualdade de rendimentos está fortemente associada às diferenças nos níveis de escolaridade e qualificação profissional (Becker, 1994, p. 12). A relação entre qualificação e crescimento econômico também se mostra evidente, uma vez que poucos países conseguiram alcançar um desenvolvimento sustentado sem investimentos significativos na capacitação de sua força de trabalho (Becker, 1994, p. 12; p. 24). Ademais, os retornos da educação ao longo do ciclo de vida são expressivos: embora o período inicial de treinamento possa implicar menores rendimentos, o investimento em capital humano tende a resultar em aumentos substanciais na remuneração futura (Becker, 1994, p. 37).

No âmbito do mercado de trabalho, Stigler (1962) argumenta que a busca por informações é um processo que envolve custos e retornos tanto para trabalhadores quanto para empregadores. A duração esperada do emprego é um fator determinante na intensidade dessa busca, pois períodos mais longos justificam um maior investimento na coleta de informações sobre salários e condições de trabalho. Além disso, o autor ressalta que a variabilidade salarial reflete tanto diferenças na qualidade dos trabalhadores quanto imperfeições no acesso à informação, influenciando a capacidade dos indivíduos de permanecerem em empregos específicos (Stigler, 1962, p. 97-98; p. 101-102).

Dessa forma, a permanência no emprego não se restringe apenas a questões individuais, mas envolve também a eficiência econômica e a redução de custos sociais relacionados à alta rotatividade. Neste estudo, o foco recai sobre a identificação dos fatores que condicionam a permanência dos indivíduos no mercado de trabalho, considerando as influências de gênero, raça e escolaridade na estabilidade profissional.

# REFERENCIAL TEÓRICO

Burdett (1978) apresenta um modelo teórico em que as taxas de saída de empregos dependem diretamente da distribuição salarial entre as empresas e dos custos associados à busca por novas oportunidades. O autor argumenta que trabalhadores com empregos de longa duração tendem a permanecer em suas posições devido aos custos crescentes de mudança e ao valor percebido do emprego atual em comparação com outras ofertas disponíveis (Burdett, 1978, p. 214-216).

Nesse sentido, a pesquisa de Abraham e Farber (1987) sugere que a relação entre a antiguidade e os rendimentos está fortemente influenciada pela qualidade do trabalhador, do emprego e da correspondência entre ambos. Esse viés ocorre porque trabalhadores com empregos de maior duração tendem a possuir características não observadas que elevam seus salários desde o início do vínculo empregatício (Abraham; Farber, 1987, p. 279). Para tal perspectiva, "*The finding that workers in longer jobs earn more in every year on the job than workers in shorter jobs is verified*"[[3]](#footnote-2) (Abraham; Farber, 1987, p. 290).

No entanto, De Oliveira (2013) argumenta que esse tipo de análise não é suficiente para explicar a realidade do mercado de trabalho, especificamente no Brasil. Para o autor, a verificação dos dados

evidencia a impertinência de análises que supervalorizam o incremento educacional como elemento determinante da posição dos indivíduos no mercado de trabalho. Em síntese, a escolarização não tem sido o determinante da entrada e permanência no mercado de trabalho, e muito menos determina salários e forma de ocupação (De Oliveira, 2013, p. 711).

O autor observa que as condições de permanência no mercado de trabalho variam significativamente em função de gênero, raça e escolaridade, com mulheres e negros enfrentando maiores dificuldades, independentemente de seu nível de qualificação (De Oliveira, p. 687-689).

Seguindo a mesma mesma linha, Silva, De Faria e Teixeira (2021) apontam que, enquanto o aumento da escolaridade eleva as chances de inserção no mercado formal, o impacto é substancialmente menor para trabalhadores negros em comparação aos brancos. Além disso, as mulheres negras enfrentam uma dupla discriminação, de gênero e racial, que as posiciona como o grupo mais vulnerável dentro do mercado de trabalho formal (Silva; De Faria; Teixeira, 2021, p. 61-64).

Já Do Monte et al. (2005) mostram que indivíduos à procura de reemprego possuem maior probabilidade de inserção ocupacional em relação aos que buscam o primeiro emprego, devido, em grande parte, à experiência acumulada no mercado de trabalho, que atua como um diferencial competitivo (Do Monte et al., 2005, p. 159).

N’outra linha, Gonzaga, Maloney e Mizala (2003) afirmam que o sistema FGTS, ao permitir o acesso ao fundo somente em casos de demissão sem justa causa, gera incentivos para demissões fraudulentas e altos índices de rotatividade, prejudicando a estabilidade no mercado de trabalho formal (Gonzaga; Maloney; Mizala, 2003, p. 166). O estudo também aponta que alterações legislativas, como as de 1988 e 2001, que aumentaram os custos de demissão, tiveram impacto significativo na duração do emprego, especialmente para trabalhadores formais (Gonzaga; Maloney; Mizala, 2003, p. 201).

Nessa mesma ideia, contudo, Pontes et al. (2018) demonstram que a escolaridade elevada está associada à redução da rotatividade, enquanto características como idade jovem, residência em área urbana e chefia de família aumentam a probabilidade de alternância entre empregos (Pontes et al., 2018, p. 18-20). Segundo o estudo, trabalhadores sindicalizados apresentaram menor probabilidade de rotatividade, sugerindo que a proteção aos direitos trabalhistas influencia positivamente a permanência no mercado de trabalho (Pontes et al., 2018, p. 20).

Por fim, Sarathy e Barbosa (1981) destacam que a produtividade de um trabalhador aumenta significativamente com o tempo de serviço em uma empresa, pois a experiência acumulada resulta em melhor desempenho e menores custos operacionais para o empregador (Sarathy; Barbosa , 1981, p. 17). O estudo sugere que fatores externos, como oportunidades de emprego e laços sociais, influenciam a decisão de permanência, enquanto fatores internos, como clima organizacional e condições de trabalho, determinam a satisfação no emprego (Sarathy; Barbosa , 1981, p. 19-21). Além disso, a permanência no emprego é influenciada por variáveis como idade, estado civil e nível de instrução. Casados tendem a permanecer mais tempo empregados, possivelmente devido às responsabilidades financeiras familiares (Sarathy; Barbosa , 1981, p. 23).

A partir desse arcabouço teórico, o presente estudo busca integrar as dimensões econômicas e sociais na análise da permanência no emprego, verificando a continuidade das hipóteses levantadas no referencial. É social no sentido de incorporar características individuais imutáveis, como raça e sexo; e econômica por incorporar questões de ocupação, como a área de atuação do indivíduo.

# MÉTODO

## Modelo

A análise dos determinantes da permanência no emprego formal fundamentou-se em modelos econométricos voltados a variáveis dependentes binárias, estimados a partir de microdados da Relação Anual de Informações Sociais (RAIS). A variável de interesse, denominada , assume valor 1 para vínculos empregatícios com duração superior a 24 meses e 0 caso contrário. Com isso, buscou-se investigar de que forma atributos individuais, ocupacionais e organizacionais afetam a probabilidade de um trabalhador manter-se formalmente empregado ao longo desse período.

### Modelo Logit

Para sistematizar a estimação, foram especificados dois tipos de modelos: Modelo Logit e Modelo Probit. A seguir, são apresentadas as equações que embasam cada um desses testes, bem como as expressões das funções de verossimilhança e derivadas relevantes para a interpretação dos efeitos marginais.

No modelo Logit, supõe-se a existência de uma variável latente dada por

|  |  |
| --- | --- |
|  | (1) |

A variável observada (0 ou 1) indica se é positiva. A probabilidade de permanência no emprego, condicionada ao vetor , segue a função logística acumulada:

|  |  |
| --- | --- |
|  | (2) |

Se essa probabilidade estimada for maior ou igual a 50%, conclui-se que o modelo está prevendo uma maior chance de “permanência” do que de “não permanência”, para tal 0.5 é o *threshold*. Portanto, a especificação completa do modelo Logit é:

com

A função de log-verossimilhança do modelo Logit, dada uma amostra de indivíduos, é:

|  |  |
| --- | --- |
|  | (3) |

Na equação, . A estimação de por máxima verossimilhança equivale a encontrar os valores que maximizam .

### **Efeito Marginal no Modelo Logit**

Para uma variável contínua , o efeito marginal da mudança em sobre a probabilidade de permanência é:

|  |  |
| --- | --- |
|  | (4) |

No Stata, esse efeito marginal pode ser calculado em torno da média das variáveis ou em outros pontos específicos.

### **Razões de chance**

No Logit, também se interpretam os coeficientes em termos de *odds ratios*:

indicando a multiplicação das chances de para um incremento de uma unidade em , mantidas fixas as demais variáveis.

### Modelo Probit

No modelo Probit, a diferença essencial é a distribuição assumida para o termo de erro. Em lugar de ter distribuição logística, supõe-se . A variável latente permanece:

|  |  |
| --- | --- |
|  | (5) |

Assim, a probabilidade condicional de assume a forma:

|  |  |
| --- | --- |
|  | (6) |

Aqui, é a função de distribuição acumulada normal padrão. A log-verossimilhança do modelo Probit é:

|  |  |
| --- | --- |
|  | (7) |

A maximização dessa função permite obter os estimadores .

### **Efeito Marginal no Modelo Probit**

A derivada da probabilidade em relação a é:

|  |  |
| --- | --- |
|  | (8) |

Acima, é a densidade da Normal padrão (a derivada de ). Também aqui, o efeito marginal pode ser calculado em diversos pontos das covariadas ou como média marginal na amostra.

### Modelo estimado

A especificação comum aos modelos Logit e Probit agrupa as mesmas variáveis explicativas, resultando em:

em que corresponde à função logística no Logit ou à função de distribuição acumulada normal no Probit, e representa o termo aleatório inerente a cada especificação.

Como a variável dependente é binária, a classificação de cada observação baseia-se em um ponto de corte (0,5) aplicado sobre a probabilidade estimada *[[4]](#footnote-3)*. Se , classifica-se como permanência = 1; caso contrário, permanência = 0. Isso possibilita a construção de uma matriz de confusão e o cálculo de métricas como sensibilidade, especificidade e taxa de acerto geral.

A comparação entre os resultados dos modelos (Logit e Probit) contempla não apenas a significância estatística e o sinal dos coeficientes, mas também a magnitude dos efeitos marginais e as medidas de qualidade preditiva. Em particular, a proximidade entre os coeficientes Logit e Probit (quando ajustados às escalas apropriadas) reforça a robustez dos achados. Além disso, as regressões com termos de interação ou variáveis transformadas (como idade ao quadrado) foram estimadas em análises adicionais para verificar eventuais não linearidades ou efeitos moderadores entre variáveis, embora não constem necessariamente em todas as especificações finais.

Dessa forma, a abordagem metodológica delineada — composta pela construção de uma variável binária para permanência e pela aplicação de Modelos Logit e Probit — permite examinar de maneira abrangente quais fatores influenciam a estabilidade no emprego formal.

## Dados

Foram utilizados registros do Sul do país a partir de microdados da Relação Anual de Informações Sociais (RAIS) de 2022, uma base administrativa que reúne informações detalhadas sobre vínculos empregatícios formais no Brasil, abrangendo dados sociodemográficos e características do vínculo empregatício, com o objetivo de identificar fatores que influenciam a permanência no emprego.

A base de dados foi importada e tratada para garantir consistência e padronização. As variáveis selecionadas incluem idade, escolaridade, raça, sexo, tipo de vínculo empregatício, remuneração média nominal, faixa horária contratada, trabalho intermitente, tamanho do estabelecimento, setor econômico (classificado pela CNAE 2.0) e adesão ao Simples Nacional. Os dados também contêm informações sobre o tempo de vínculo empregatício, que foram fundamentais para a definição da variável dependente.

**Tabela 1 –** Estatísticas descritivas das variáveis

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Variável | Observações | Média | Desvio Padrão | Mínimo | Máximo |
| permanencia | 5,954,105 | 0.3723 | 0.4834 | 0 | 1 |
| idade\_grupo | 5,954,105 | 36.07 | 12.17 | 10 | 100 |
| escolaridade\_após\_2005 | 5,954,105 | 6.81 | 1.56 | 1 | 11 |
| raçacor | 5,954,105 | 5.29 | 3.34 | 1 | 9 |
| sexotrabalhador | 5,954,105 | 1.40 | 0.49 | 1 | 2 |
| tipovínculo | 5,954,105 | 21.73 | 20.56 | 10 | 97 |
| faixahoracontrat | 5,954,105 | 6.02 | 6.14 | 1 | 99 |
| indtrabintermitente | 5,954,104 | 0.0072 | 0.0848 | 0 | 1 |
| tamanhoestabelecimento | 5,954,105 | 5.87 | 2.79 | 1 | 10 |
| cnae20classe | 5,954,105 | 52,102.01 | 25,857.88 | 977 | 99,008 |
| indsimples | 5,954,105 | 0.3062 | 0.4609 | 0 | 1 |
| vlremun\_faixa | 5,954,105 | 6.73 | 19.72 | 0 | 5,356 |

Fonte: Produzido pelos autores.

O tratamento inicial envolveu a recodificação de valores inválidos, como aqueles marcados como "99" em variáveis categóricas, que foram substituídos pela categoria "Não Identificado". Variáveis originalmente em formato string, como remuneração média nominal e tempo de vínculo, foram convertidas para valores numéricos após padronização de separadores decimais. Adicionalmente, foram criadas variáveis derivadas para facilitar a análise, como faixas etárias arredondadas a cada cinco anos e faixas de remuneração categorizadas em intervalos de R$500.

A variável dependente, permanência no emprego, foi construída como binária, assumindo o valor 1 para trabalhadores com tempo de vínculo superior a dois anos e 0 para os demais. Esse recorte operacional reflete o interesse em analisar fatores associados à estabilidade no mercado de trabalho formal. Como variáveis explicativas, foram consideradas tanto características individuais, como idade, escolaridade, raça/cor e sexo, quanto características do vínculo empregatício e do empregador, como tipo de contrato, remuneração, porte do estabelecimento e setor econômico.

**Tabela 2 –** Distribuição da Variável Dependente (Permanência no Emprego)

|  |  |  |  |
| --- | --- | --- | --- |
| Categoria | Frequência | Percentual | Acumulado (%) |
| 0 (Não permaneceu) | 3,737,428 | 62.77 | 62.77 |
| 1 (Permaneceu) | 2,216,677 | 37.23 | 100.00 |
| Total | 5,954,105 | 100.00 | - |

Fonte: Produzido pelos autores.

A análise estatística foi conduzida por meio de modelos econométricos de regressão, com destaque para os modelos Logit e Probit, adequados para variáveis dependentes binárias. O modelo Logit estima a probabilidade de permanência no emprego, sendo os resultados interpretados em termos de *odds ratios*. O modelo Probit, que utiliza a função de distribuição acumulada normal, foi empregado como abordagem complementar, permitindo maior robustez na interpretação dos resultados.

Os efeitos marginais foram calculados para interpretar a magnitude dos impactos das variáveis explicativas sobre a probabilidade de permanência no emprego. Esses cálculos foram realizados tanto em torno da média das variáveis quanto para valores específicos, como faixas etárias e níveis de remuneração. Modelos com termos de interação também foram estimados para explorar possíveis relações não lineares, como a interação entre idade e escolaridade.

Por fim, as probabilidades preditas foram analisadas graficamente, destacando como a permanência no emprego varia em função de diferentes características individuais e ocupacionais. Toda a análise foi conduzida no software Stata, com a implementação de scripts que asseguram a replicabilidade dos resultados e incluem desde o tratamento dos dados até a estimação dos modelos e a apresentação dos resultados.

# RESULTADOS

## Resultados dos Modelos Logit e Probit

O modelo Logit mostrou que a idade, medida em grupos quinquenais, tem um impacto positivo e significativo na probabilidade de permanência no emprego, com um coeficiente de 0,062 (erro padrão: 0,00009). Isso indica que trabalhadores mais velhos apresentam maior probabilidade de permanecerem empregados. A escolaridade, representada pela variável escolaridadeapós2005, também foi um determinante importante, com coeficiente de 0,155 (erro padrão: 0,00067), sugerindo que níveis mais altos de educação estão associados a maior estabilidade no emprego. Por outro lado, a variável raçacor apresentou coeficiente negativo (-0,031; erro padrão: 0,00030), o que aponta para uma menor probabilidade de permanência entre trabalhadores não brancos.

Outras variáveis explicativas também se mostraram estatisticamente significativas. O sexo do trabalhador teve um coeficiente positivo (0,179; erro padrão: 0,00205), indicando que mulheres apresentam maior probabilidade de permanência. Em contrapartida, o tipo de vínculo empregatício apresentou um impacto negativo (-0,046; erro padrão: 0,00008), reforçando que vínculos mais precários estão associados a menores probabilidades de estabilidade no emprego. A intermitência no trabalho teve um efeito marcadamente negativo (-1,202; erro padrão: 0,01318), evidenciando o impacto adverso desse regime na permanência.

O modelo Probit apresentou resultados qualitativamente semelhantes, com coeficientes ligeiramente menores devido às diferenças na escala inerentes ao método. A idade em grupos continuou sendo um fator positivo, com coeficiente de 0,037 (erro padrão: 0,00005), assim como a escolaridade (0,093; erro padrão: 0,00040). A variável raçacor manteve seu impacto negativo (-0,019; erro padrão: 0,00018), enquanto o sexo do trabalhador permaneceu como um fator positivo (0,106; erro padrão: 0,00123). Novamente, a intermitência no trabalho se destacou como um dos fatores mais negativos, com coeficiente de -0,686 (erro padrão: 0,00736).

Os dois modelos apresentaram resultados consistentes, indicando a robustez das análises. As variáveis idade, escolaridade e sexo destacaram-se como fatores positivos associados à permanência no emprego, enquanto características como raça, intermitência no trabalho e tipos precários de vínculo apresentaram efeitos negativos. O ajuste dos modelos também foi moderadamente adequado, com valores de Pseudo R² de 0,1896 para o modelo Logit e 0,1915 para o Probit.

**Tabela 3 –** Comparação dos Resultados dos Modelos Logit e Probit

|  |  |  |
| --- | --- | --- |
| Variável | Logit Coeficiente (Erro Padrão) | Probit Coeficiente (Erro Padrão) |
| idade\_grupo | 0.06199 (0.00009) | 0.03721 (0.00005) |
| escolaridadeapós2005 | 0.15469 (0.00067) | 0.09265 (0.00040) |
| raçacor | -0.03141 (0.00030) | -0.01921 (0.00018) |
| sexotrabalhador | 0.17912 (0.00205) | 0.10649 (0.00123) |
| tipovínculo | -0.04571 (0.00008) | -0.02715 (0.00004) |
| faixahoracontrar | 0.00180 (0.00017) | 0.00192 (0.00009) |
| indtrabintermitent | -1.20159 (0.01318) | -0.68637 (0.00736) |
| tamanhoestabelecimento | 0.11824 (0.00042) | 0.06991 (0.00025) |
| cnae20classe | 3.14e-06 (4.13e-08) | 1.81e-06 (2.47e-08) |
| indsimples | -0.29523 (0.00253) | -0.17739 (0.00150) |
| vlremun\_faixa | 0.00319 (0.00007) | 0.00159 (0.00004) |
| \_cons | -3.93130 (0.00679) | -2.35475 (0.00398) |
| Pseudo R² | 0.1896 | 0.1915 |

Fonte: Produzido pelos autores.

## Marginal Effects dos Modelos Logit e Probit

As margens estimadas para os modelos Logit e Probit representam os efeitos marginais das variáveis explicativas sobre a probabilidade de permanência no emprego. A análise incluiu tanto os efeitos condicionais (calculados na média das variáveis explicativas) quanto os efeitos médios marginais (calculados ao longo de toda a distribuição dos dados). Esses resultados fornecem uma interpretação direta sobre como as variáveis independentes afetam incrementalmente a probabilidade de um trabalhador permanecer no emprego.

No modelo Logit, os efeitos marginais calculados na média mostraram que um aumento em uma unidade na idade (grupos quinquenais) aumentou a probabilidade de permanência em 1,35 pontos percentuais. Similarmente, a escolaridade elevou a probabilidade em 3,36 pontos percentuais por nível adicional. Por outro lado, a raça/cor teve um impacto negativo de 0,68 pontos percentuais, indicando uma menor probabilidade de permanência para trabalhadores de grupos raciais historicamente desfavorecidos. O tipo de vínculo empregatício também contribuiu negativamente, reduzindo a probabilidade em 0,99 pontos percentuais, enquanto o trabalho intermitente mostrou-se o fator mais adverso, diminuindo a probabilidade de permanência em 26,10 pontos percentuais.

Outras variáveis apresentaram impactos positivos, como o tamanho do estabelecimento, que aumentou a probabilidade em 2,57 pontos percentuais, e a faixa de remuneração, com um efeito menor, mas significativo, de 0,07 pontos percentuais por unidade. A faixa horária contratada teve um impacto positivo discreto de 0,04 pontos percentuais, enquanto a adesão ao regime simplificado apresentou efeito negativo, reduzindo a probabilidade de permanência em 6,41 pontos percentuais. Por fim, o setor econômico (CNAE 2.0) teve um impacto marginal positivo, mas estatisticamente significativo, com um aumento de 0,000068 pontos percentuais por unidade.

**Tabela 4 –** Marginal Effects (Modelo Logit)

|  |  |  |
| --- | --- | --- |
| Variável | Conficional [at means] | Média Marginal |
| Idade em grupos | 0.0135 (0.00002) | 0.0112 (0.00001) |
| Escolaridade | 0.0336 (0.00015) | 0.0280 (0.00012) |
| Raça/Cor | -0.0068 (0.00006) | -0.0057 (0.00005) |
| Sexo | 0.0389 (0.00045) | 0.0325 (0.00037) |
| Tipo de Vínculo | -0.0099 (0.00002) | -0.0083 (0.00001) |
| Faixa horária contratada | 0.0004 (0.00004) | 0.0003 (0.00003) |
| Trabalho intermitente | -0.2610 (0.00286) | -0.2177 (0.00238) |
| Tamanho do estabelecimento | 0.0257 (0.00009) | 0.0214 (0.00007) |
| CNAE 2.0 | 6.82e-07 (8.97e-09) | 5.69e-07 (7.48e-09) |
| Regime simples | -0.0641 (0.00055) | -0.0535 (0.00046) |
| Faixa de remuneração | 0.0007 (0.00002) | 0.0006 (0.00001) |

Fonte: Produzido pelos autores.

No modelo Probit, os resultados seguiram a mesma direção, com magnitudes ligeiramente diferentes devido à função de ligação utilizada. A idade em grupos quinquenais aumentou a probabilidade de permanência em 1,34 pontos percentuais, e a escolaridade resultou em um incremento de 3,33 pontos percentuais. A intermitência no trabalho manteve-se como o fator mais adverso, reduzindo a probabilidade de permanência em 24,68 pontos percentuais. O tipo de vínculo empregatício, a raça/cor e a adesão ao regime simplificado também tiveram impactos negativos consistentes, enquanto o tamanho do estabelecimento e a faixa de remuneração contribuíram positivamente, como observado no modelo Logit.

**Tabela 5 –** Marginal Effects (Modelo Probit)

|  |  |  |
| --- | --- | --- |
| Variável | Conficional [at means] | Média Marginal |
| Idade em grupos | 0.0134 (0.00002) | 0.0112 (0.00001) |
| Escolaridade | 0.0333 (0.00014) | 0.0280 (0.00012) |
| Raça/Cor | -0.0069 (0.00006) | -0.0058 (0.00005) |
| Sexo | 0.0383 (0.00044) | 0.0321 (0.00037) |
| Tipo de Vínculo | -0.0098 (0.00001) | -0.0082 (0.00001) |
| Faixa horária contratada | 0.0007 (0.00003) | 0.0006 (0.00003) |
| Trabalho intermitente | -0.2468 (0.00265) | -0.2071 (0.00222) |
| Tamanho do estabelecimento | 0.0251 (0.00009) | 0.0211 (0.00007) |
| CNAE 2.0 | 6.52e-07 (8.88e-09) | 5.47e-07 (7.44e-09) |
| Regime simples | -0.0638 (0.00054) | -0.0535 (0.00045) |
| Faixa de remuneração | 0.0006 (0.00001) | 0.0005 (0.00001) |

Fonte: Produzido pelos autores.

Os efeitos médios marginais apresentaram magnitudes menores em comparação aos efeitos condicionais, mas mantiveram a consistência nos sinais e na significância estatística das variáveis explicativas. Isso reforça que os determinantes da permanência no emprego possuem impactos robustos e que as direções dos efeitos se mantêm inalteradas, independentemente do método de estimação utilizado.

## *Odds Ratios* dos Modelos Logit e Probit

As estimativas do modelo logístico, apresentadas em termos de *odds ratios*, permitem interpretar como cada variável explicativa influencia as chances de permanência no emprego. O impacto é quantificado de forma multiplicativa: valores superiores a 1 indicam aumento das chances, enquanto valores inferiores a 1 sugerem redução.

A idade em grupos apresentou *odds ratio* de 1,064 (Intervalo de Confiança [IC] de 95%: 1,0638 a 1,0641), indicando que cada incremento nos grupos quinquenais de idade aumenta as chances de permanência em aproximadamente 6,4%, mantendo as demais variáveis constantes. A escolaridade após 2005 teve um impacto semelhante, com *odds ratio* de 1,167 (IC: 1,1658 a 1,1688), revelando que níveis mais altos de escolaridade aumentam em 16,7% a probabilidade de permanência, evidenciando o papel positivo do capital humano.

Por outro lado, a variável raça/cor apresentou um *odds ratio* de 0,969 (IC: 0,9685 a 0,9696), sugerindo que trabalhadores de grupos raciais historicamente desfavorecidos têm, em média, 3,1% menos chances de permanecer no emprego. Esse resultado reflete, possivelmente, as desigualdades estruturais no mercado de trabalho do Sul brasileiro. De forma similar, o tipo de vínculo empregatício apresentou um impacto negativo relevante, com *odds ratio* de 0,955 (IC: 0,9552 a 0,9555), indicando uma redução de 4,5% na probabilidade de permanência para trabalhadores com vínculos mais precários.

Entre os fatores mais expressivos, destaca-se o trabalho intermitente, com *odds ratio* de 0,301 (IC: 0,2930 a 0,3086), o que implica que trabalhadores nessa modalidade têm cerca de 70% menos chances de permanecer empregados em comparação aos demais. Este resultado reforça o impacto adverso de vínculos menos estáveis na probabilidade de retenção.

Em contraste, variáveis relacionadas às características do estabelecimento apresentaram efeitos positivos. O tamanho do estabelecimento mostrou um *odds ratio* de 1,126 (IC: 1,1246 a 1,1264), indicando que trabalhadores vinculados a empresas de maior porte têm 12,6% mais chances de permanência. O setor econômico (representado por CNAE 2.0) teve um efeito positivo marginal, com *odds ratio* de 1,000003 (IC: 1,000003 a 1,000003), sugerindo uma influência estatisticamente significativa, mas de pequena magnitude.

A variável faixa de remuneração, embora com impacto menor, apresentou *odds ratio* de 1,003 (IC: 1,0030 a 1,0033), o que sugere que aumentos incrementais na remuneração média aumentam ligeiramente as chances de permanência em 0,3%. Já o regime simplificado (Simples Nacional) teve efeito negativo, com *odds ratio* de 0,744 (IC: 0,7407 a 0,7481), indicando uma redução de 25,6% na probabilidade de permanência em comparação a trabalhadores em empresas fora desse regime.

Por fim, a faixa horária contratada teve um impacto positivo discreto, com *odds ratio* de 1,002 (IC: 1,0015 a 1,0021), demonstrando que aumentos na jornada contratual estão associados a pequenas elevações nas chances de permanência.

**Tabela 6 –** Odds Ratios – Modelo Logístico

|  |  |  |  |
| --- | --- | --- | --- |
| Variável | Odds Ratio | Intervalo de Confiança (95%) | Interpretação |
| Idade em grupos | 1.064 | [1.0638 - 1.0641] | +6,4% nas chances de permanência. |
| Escolaridade | 1.167 | [1.1658 - 1.1688] | +16,7% nas chances. |
| Raça/Cor | 0.969 | [0.9685 - 0.9696] | -3,1% nas chances. |
| Sexo | 1.196 | [1.1914 - 1.2010] | +19,6% para homens. |
| Tipo de Vínculo | 0.955 | [0.9552 - 0.9555] | -4,5% nas chances. |
| Trabalho Intermitente | 0.301 | [0.2930 - 0.3086] | -70% nas chances. |
| Tamanho do Estabelecimento | 1.126 | [1.1246 - 1.1264] | +12,6% nas chances. |
| CNAE 2.0 | 1.000003 | [1.000003 - 1.000003] | Pequeno efeito positivo. |
| Regime Simples | 0.744 | [0.7407 - 0.7481] | -25,6% nas chances. |
| Faixa de Remuneração | 1.003 | [1.0030 - 1.0033] | +0,3% nas chances. |
| Faixa horária | 1.002 | [1.0015 - 1.0021] | Pequeno impacto positivo. |

Fonte: Produzido pelos autores.

Os resultados indicam que fatores individuais, como idade e escolaridade, e características do estabelecimento, como porte e setor econômico, têm influência positiva na permanência no emprego. Em contrapartida, vínculos precários e regimes intermitentes são determinantes negativos significativos. As variáveis raça/cor e regime simplificado também sugerem a existência de desigualdades estruturais que afetam negativamente a estabilidade no mercado formal de trabalho.

## Predição das Probabilidades e Avaliação do Desempenho dos Modelos

As probabilidades previstas e as taxas de classificação dos modelos Logit e Probit foram utilizadas para avaliar a capacidade desses modelos de reproduzir a distribuição empírica da variável dependente e de prever corretamente os eventos observados de permanência no emprego.

Os valores das probabilidades previstas mostram que os modelos Logit e Probit capturam adequadamente a média empírica da variável permanencia, que é de 37,23%. As médias das probabilidades previstas foram 0,3723 para o Logit e 0,3705 para o Probit, com desvios padrão próximos (0,2291 e 0,2288, respectivamente). Esses resultados sugerem que ambos os modelos ajustam-se bem aos dados observados. Em contraste, o modelo de regressão linear (OLS) apresentou médias semelhantes, mas com previsões fora do intervalo [0,1], o que demonstra sua inadequação para variáveis dependentes binárias.

**Tabela 7 –** Probabilidades Previstas pelos Modelos

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variável | Média Prevista | Desvio Padrão | Mínimo | Máximo |
| Permanência | 0.3723 | 0.4834 | 0.0000 | 1.0000 |
| Logit (plogit) | 0.3723 | 0.2291 | 0.0016 | 0.9999 |
| Probit (pprobit) | 0.3705 | 0.2288 | 0.0001 | 1.0000 |

Fonte: Produzido pelos autores.

A avaliação das taxas de classificação revela que os modelos Logit e Probit apresentaram desempenho semelhante na capacidade de prever corretamente a permanência no emprego. Com um ponto de corte de 0,5, ambos os modelos classificaram corretamente cerca de 70,8% das observações. A sensibilidade foi de aproximadamente 52%, indicando uma precisão moderada na identificação dos trabalhadores que permaneceram no emprego. Em contrapartida, a especificidade foi elevada, atingindo 81,7% no Logit e 81,9% no Probit, o que reflete uma alta capacidade de prever corretamente os trabalhadores que não permaneceram no emprego.

**Tabela 8 –** Probabilidades Previstas pelos Modelos

|  |  |  |
| --- | --- | --- |
| Métrica | Logit (%) | Probit (%) |
| Taxa de acerto geral | 70,83 | 70,82 |
| Sensibilidade | 52,49 | 52,21 |
| Especificidade | 81,71 | 81,86 |
| Valor preditivo positivo | 62,99 | 63,06 |
| Valor preditivo negativo | 74,36 | 74,28 |

Fonte: Produzido pelos autores.

.Os resultados indicam que, embora os modelos apresentem dificuldades em identificar corretamente todos os casos de permanência (sensibilidade moderada), eles possuem alta precisão na identificação dos casos de não permanência (especificidade elevada). A taxa de classificação geral próxima de 70,8% demonstra que os modelos capturam, de forma robusta, os padrões associados à probabilidade de permanência no emprego.

Portanto, tanto as probabilidades previstas quanto as taxas de classificação confirmam a adequação dos modelos Logit e Probit para a análise da variável dependente. O desempenho semelhante entre os modelos reforça que ambos são apropriados para esse tipo de análise.

## Análise Gráfica dos Efeitos Marginais: Probabilidade de Permanência por Idade

A análise dos efeitos marginais foi realizada para avaliar a probabilidade predita de permanência no emprego ao longo de diferentes faixas etárias, com incrementos de cinco anos, variando de 20 a 60 anos. O objetivo foi identificar como a idade influencia a probabilidade de um trabalhador permanecer empregado, mantendo constantes as demais variáveis do modelo.

Os valores preditos das probabilidades foram obtidos através do comando margins, que estima as probabilidades médias condicionais de permanência para diferentes valores da variável idade\_grupo. Os resultados indicam um aumento contínuo e linear na probabilidade de permanência conforme o grupo de idade avança.

A tabela abaixo apresenta os valores estimados das probabilidades e os respectivos intervalos de confiança a 95%:

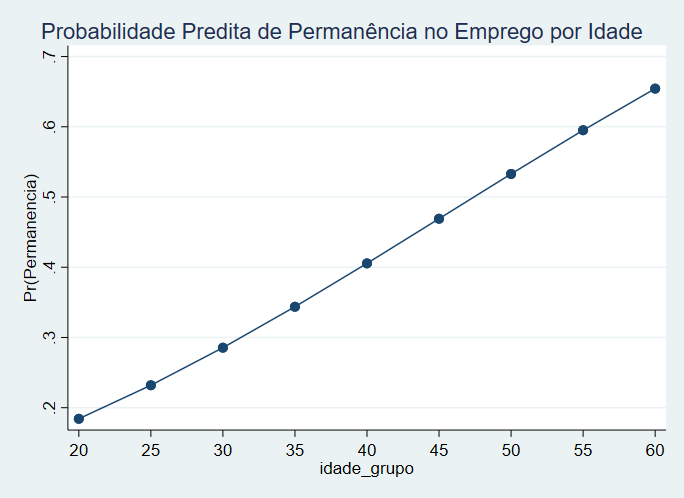
**Tabela 9 –** Probabilidade Predita de Permanência no Emprego por Faixa Etária

|  |  |  |  |
| --- | --- | --- | --- |
| Faixa Etária (idade\_grupo) | Probabilidade Predita | Erro Padrão | Intervalo de Confiança (95%) |
| 20 | 0,1842 | 0,00025 | [0,1837 - 0,1847] |
| 25 | 0,2319 | 0,00023 | [0,2315 - 0,2324] |
| 30 | 0,2855 | 0,00021 | [0,2851 - 0,2859] |
| 35 | 0,3438 | 0,00019 | [0,3434 - 0,3441] |
| 40 | 0,4055 | 0,00020 | [0,4051 - 0,4059] |
| 45 | 0,4691 | 0,00024 | [0,4686 - 0,4696] |
| 50 | 0,5328 | 0,00029 | [0,5323 - 0,5334] |
| 55 | 0,5951 | 0,00035 | [0,5944 - 0,5958] |
| 60 | 0,6543 | 0,00040 | [0,6536 - 0,6551] |

Fonte: Produzido pelos autores.

O gráfico resultante, apresentado a seguir, ilustra visualmente essa tendência:

**Gráfico 1 –** Probabilidade Predita de Permanência no Emprego por Idade



Fonte: Produzido pelos autores.

Os resultados evidenciam uma relação positiva e aproximadamente linear entre a idade e a probabilidade predita de permanência no emprego. Aos 20 anos, a probabilidade estimada é de 18,42%, crescendo gradativamente até atingir 65,43% aos 60 anos. Esse comportamento sugere que trabalhadores mais velhos têm maiores chances de permanecer empregados, possivelmente em função de fatores como experiência acumulada e estabilidade no mercado de trabalho.

Adicionalmente, os erros padrão baixos e os intervalos de confiança estreitos reforçam a precisão das estimativas, garantindo a robustez dos resultados. A tendência observada é consistente com as expectativas teóricas, que associam o avanço da idade à maior estabilidade ocupacional.

## Análise Gráfica dos Efeitos Marginais: Probabilidade de Permanência por Faixa de Remuneração

A relação entre a faixa de remuneração e a probabilidade predita de permanência no emprego foi analisada utilizando os valores preditos médios gerados pelo modelo Logit. As probabilidades foram estimadas para valores da variável vlremun\_faixa que variam de 1 a 10, com incrementos unitários, representando diferentes faixas de remuneração.

Os resultados demonstram um aumento suave e gradual na probabilidade de permanência conforme a faixa de remuneração aumenta. Esse comportamento é consistente com a expectativa de que níveis salariais mais elevados estejam associados a uma maior estabilidade no emprego formal.

**Tabela 10 –** Probabilidade Predita de Permanência no Emprego por Faixa Etária

|  |  |  |  |
| --- | --- | --- | --- |
| Faixa de Remuneração (vlremun\_faixa) | Probabilidade Predita | Erro Padrão | Intervalo de Confiança (95%) |
| 1 | 0,3675 | 0,00019 | [0,3671 - 0,3679] |
| 2 | 0,3680 | 0,00018 | [0,3676 - 0,3683] |
| 3 | 0,3685 | 0,00018 | [0,3681 - 0,3688] |
| 4 | 0,3689 | 0,00018 | [0,3686 - 0,3693] |
| 5 | 0,3694 | 0,00018 | [0,3691 - 0,3698] |
| 6 | 0,3699 | 0,00017 | [0,3696 - 0,3703] |
| 7 | 0,3704 | 0,00017 | [0,3700 - 0,3707] |
| 8 | 0,3709 | 0,00017 | [0,3705 - 0,3712] |
| 9 | 0,3714 | 0,00018 | [0,3710 - 0,3717] |
| 10 | 0,3718 | 0,00018 | [0,3715 - 0,3722] |

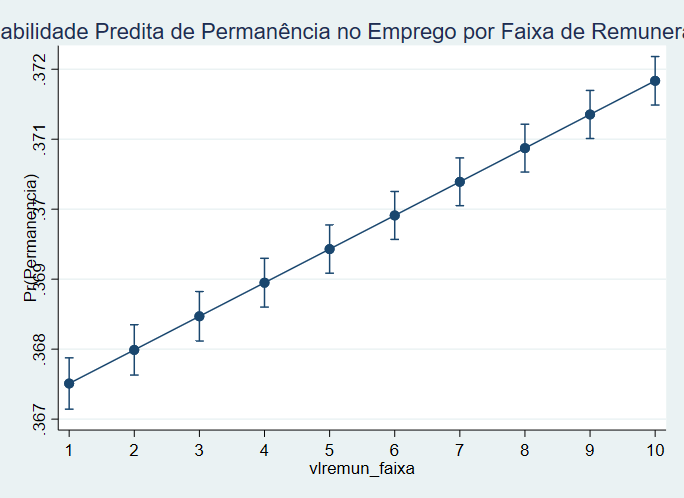
Fonte: Produzido pelos autores.

O gráfico gerado pelo comando marginsplot apresenta a tendência ascendente na probabilidade de permanência conforme a faixa de remuneração aumenta. Aos trabalhadores com menor faixa de remuneração (nível 1), a probabilidade de permanência é estimada em 36,75%. Essa probabilidade cresce gradualmente e atinge 37,18% na faixa de remuneração mais alta analisada (nível 10).

Embora o efeito da remuneração seja positivo e estatisticamente significativo, ele é relativamente pequeno em magnitude, conforme evidenciado pelo crescimento discreto nas probabilidades preditas. Isso sugere que, embora salários mais altos estejam associados a uma maior probabilidade de permanência, outros fatores podem desempenhar papéis mais importantes, como idade, tipo de vínculo empregatício e características do estabelecimento.

Os erros padrão baixos e os intervalos de confiança estreitos indicam que as estimativas são precisas e robustas.

**Gráfico 2 –** Probabilidade Predita de Permanência no Emprego por Faixa de Remuneração



Fonte: Produzido pelos autores.

O gráfico confirma a tendência linear positiva, com um crescimento estável e previsível na probabilidade de permanência conforme a remuneração aumenta. Este resultado é consistente com teorias econômicas que associam salários mais elevados a maior retenção no mercado de trabalho, seja por maior satisfação do trabalhador ou por características dos empregos melhor remunerados.

A probabilidade de permanência aumenta mais rapidamente com a idade do que com a faixa de remuneração, sugerindo que experiência desempenha um papel mais relevante.

# CONCLUSÃO

O presente estudo analisou os fatores que influenciam a probabilidade de permanência no emprego formal no Sul do Brasil, utilizando dados da RAIS e modelos econométricos como Logit e Probit. Os resultados revelaram evidências robustas sobre o papel de variáveis individuais, ocupacionais e organizacionais na determinação da permanência dos trabalhadores no mercado de trabalho formal.

As estimativas dos efeitos marginais demonstraram que a idade e a escolaridade são fatores determinantes positivos e significativos. O aumento na idade, representado por grupos quinquenais, apresentou uma relação linear com a probabilidade de permanência, evidenciando o impacto cumulativo da experiência e da estabilidade ao longo da vida profissional. De maneira similar, níveis mais elevados de escolaridade aumentaram as chances de permanência, destacando a importância do capital humano na retenção de trabalhadores.

Por outro lado, variáveis associadas a vulnerabilidades ocupacionais exerceram efeitos negativos significativos. O trabalho intermitente destacou-se como o fator mais adverso, reduzindo drasticamente a probabilidade de permanência, ao passo que o tipo de vínculo empregatício e o regime simplificado (Simples Nacional) também apresentaram impactos negativos relevantes. Esses resultados sugerem que trabalhadores em condições mais precárias enfrentam maior risco de rotatividade e instabilidade.

Adicionalmente, características dos estabelecimentos, como o tamanho da empresa, mostraram-se relevantes. Estabelecimentos de maior porte aumentam significativamente a probabilidade de permanência, sugerindo que empresas maiores oferecem maior estabilidade e estrutura ocupacional. Embora a faixa de remuneração tenha apresentado um efeito positivo estatisticamente significativo, sua magnitude foi relativamente pequena, indicando que os salários, apesar de relevantes, não são o principal fator determinante da permanência no emprego.

Os resultados dos modelos Logit e Probit foram consistentes, com diferenças marginais em termos de sensibilidade e especificidade. Ambos os modelos demonstraram bom desempenho preditivo, com 70,8% de acerto geral, destacando a robustez dos fatores explicativos na análise da permanência.

Em suma, o estudo evidencia a importância de políticas públicas voltadas à redução das desigualdades ocupacionais e ao fortalecimento das relações de trabalho, com foco em grupos mais vulneráveis, como trabalhadores intermitentes e em vínculos precários. Investimentos em educação e qualificação profissional também se mostram fundamentais para aumentar a retenção no mercado formal de trabalho, ao passo que o estímulo à estabilidade organizacional, especialmente em empresas de menor porte, pode contribuir para a redução da rotatividade laboral.

# REFERÊNCIAS

ABRAHAM, Katharine G.; FARBER, Henry S. Job Duration, Seniority, and Earnings. **The American Economic Review**, [*s. l.*], v. 77, n. 3, p. 278-297, 1987.

BECKER, Gary Stanley. **Human Capital**: A Theoretical and Empirical Analysis with Special Reference to Education. 3. ed. Chicago: The University of Chicago Press, 1994. 390 p. ISBN 0-226-04119-0.

BURDETT, Kenneth. A Theory of Employee Job Search and Quit Rates. **The American Economic Review**, [*s. l.*], v. 68, n. 1, p. 212-220, 1978.

DE OLIVEIRA, Ramon. Educação, pobreza e emprego: uma análise a partir das categorias escolaridade, gênero e cor. **Perspectiva**, *[S. l.]*, v. 31, n. 2, p. 687–719, 2013.

DO MONTE, Paulo Aguiar; DE ARAÚJO, Tarcisio Patrício; DE LIMA, Roberto Alves. Primeiro emprego e reemprego: análise de inserção ocupacional e duração do desemprego no Brasil metropolitano. **Economia e Desenvolvimento**, Recife, v. 3, n. 1, p. 133-171, 2005.

GONZAGA, Gustavo; MALONEY, William F.; MIZALA, Alejandra. Labor Turnover and Labor Legislation in Brazil [with Comments]. **Economía**, [*s. l.*], v. 4, n. 1, p. 165-222, 2003.

PONTES, Raquel Pereira; BARBOSA, Marcio Nora; SILVEIRA, Leonel dos Santos; COSTA, Lorena Vieira. O comportamento de jovens e adultos no mercado de trabalho brasileiro com relação à permanência no emprego. **Revista de Economia Contemporânea**, [*s. l.*], v. 22, n. 3, p. 1-23, 2018.

SARATHY, Ravi; BARBOSA, Jenny Dantas. Fatores explicativos da permanência e saída do emprego. **Revista de Administração de Empresas**, [*s. l.*], v. 21, n. 1, p. 17-29, 1981.

SAVIANI, Dermeval. Trabalho e educação: fundamentos ontológicos e históricos. **Revista Brasileira de Educação**, [*s. l.*], v. 13, n. 34, p. 152-180, 2007.

SILVA, Lucas Adriano; DE FARIA, Antônio Cláudio Lopes; TEIXEIRA, Evandro Camargos. Desigualdade racial no mercado de trabalho formal brasileiro. **Revista Perspectivas Online: Humanas & Sociais Aplicadas**, [*s. l.*], v. 11, n. 30, p. 51-67, 2021.

STIGLER, Geroge J. Information in the Labor Market. **Journal of Political Economy**, [*s. l.*], v. 70, n. 5, p. 94-105, 2007.

# APÊNDICES

**APÊNDICE A —** Script do Stata

\* Script para Analisar Fatores que Influenciam a Permanência no Emprego usando Modelos Logit/Probit

\* Limpar memória de trabalho

clear all

set more off

\* Definir diretório de trabalho

cd "C:/Users/Miguel/Desktop/Trabalho do Tillmãe/"

\* Importar a base de dados da RAIS, mantendo apenas as variáveis necessárias

import delimited "RAIS\_VINC\_PUB\_SUL.txt", delimiter(";") varnames(1) clear

keep idade escolaridadeapós2005 raçacor sexotrabalhador vlremunmédianom tipovínculo faixahoracontrat indtrabintermitente tamanhoestabelecimento cnae20classe indsimples tempoemprego

\* Substituir valores inválidos (99) por um valor apropriado (9 - Não Identificado)

replace raçacor = 9 if raçacor == 99

\* Verificar a distribuição após a correção

tabulate raçacor

\* Limpeza e conversão da variável 'vlremunmédianom'

\* Criar uma nova variável 'vlremunmédianom\_clean' para preservar o valor original

clonevar vlremunmédianom\_clean = vlremunmédianom

\* Substituir vírgulas por pontos, para padronização de separadores decimais

replace vlremunmédianom\_clean = subinstr(vlremunmédianom\_clean, ",", ".", .)

\* Remover zeros à esquerda, que podem causar problemas de conversão

replace vlremunmédianom\_clean = regexr(vlremunmédianom\_clean, "^0+", "")

\* Converter a variável de string para numérica, forçando a conversão

destring vlremunmédianom\_clean, replace force

\* Limpeza e conversão da variável 'tempoemprego'

\* Criar uma cópia 'tempoemprego\_clean' para preservar o original

clonevar tempoemprego\_clean = tempoemprego

\* Substituir vírgulas por pontos

replace tempoemprego\_clean = subinstr(tempoemprego\_clean, ",", ".", .)

\* Converter a variável de string para numérica

destring tempoemprego\_clean, replace force

\* Verificar se as conversões foram bem-sucedidas

summarize vlremunmédianom\_clean tempoemprego\_clean

\* Criar a variável dependente binária 'permanencia' para indicar se o trabalhador permaneceu no emprego por mais de 2 anos

gen permanencia = 0

replace permanencia = 1 if tempoemprego\_clean > 24

\* Criar faixas etárias (arredondadas a cada 5 anos) para análise de efeitos marginais

\* A multiplicação por 5 permite que as variáveis fiquem de 5 em 5 anos

gen idade\_grupo = round(idade/5)\*5

\* Criar faixas de remuneração para facilitar a análise

gen vlremun\_faixa = ceil(vlremunmédianom\_clean / 500)

\* Definir variáveis dependente e independentes como globais

global ylist permanencia

global xlist idade\_grupo escolaridadeapós2005 raçacor sexotrabalhador tipovínculo faixahoracontrat indtrabintermitente tamanhoestabelecimento cnae20classe indsimples vlremun\_faixa

\* Estatísticas descritivas

describe $ylist $xlist

summarize $ylist $xlist

tabulate $ylist

\* Regressão linear como comparação inicial

reg $ylist $xlist

\* Modelos Logit e Probit

logit $ylist $xlist

probit $ylist $xlist

\* Marginal effects (margens ao redor da média e médias marginais)

quietly logit $ylist $xlist

margins, dydx(\*) atmeans

margins, dydx(\*)

quietly probit $ylist $xlist

margins, dydx(\*) atmeans

margins, dydx(\*)

\* Exibir odds ratios usando modelo logístico

logistic $ylist $xlist

\* Predição de probabilidades

quietly logit $ylist $xlist

predict plogit, pr

quietly probit $ylist $xlist

predict pprobit, pr

quietly reg $ylist $xlist

predict pols, xb

summarize $ylist plogit pprobit pols

\* Classificação percentual correta

quietly logit $ylist $xlist

estat classification

quietly probit $ylist $xlist

estat classification

\* Análise gráfica de efeitos marginais

\* Probabilidade por faixa etária

margins, at(idade\_grupo=(20(5)60)) predict(pr)

marginsplot, title("Probabilidade Predita de Permanência no Emprego por Idade")

\* Probabilidade por faixa de remuneração

margins, at(vlremun\_faixa=(1(1)10)) predict(pr)

marginsplot, title("Probabilidade Predita de Permanência no Emprego por Faixa de Remuneração")

\* Modelo com interação para análise adicional

logit $ylist c.idade##c.escolaridadeapós2005 $xlist

\* Estimar modelos com robustez

logit $ylist $xlist, vce(robust)

probit $ylist $xlist, vce(robust)

\* Salvar os resultados em log

log using "analise\_permanencia\_emprego.log", replace

log close

1. Mestrando no Programa de Pós-Graduação em Economia Aplicada da Universidade Federal do Rio Grande. E-mail: lopesmiguel100@gmail.com. [↑](#footnote-ref-0)
2. Professor Adjunto do Instituto de Ciências Econômicas, Administrativas e Contábeis e do Programa de Pós-Graduação em Economia Aplicada da Universidade Federal do Rio Grande. E-mail: etillmann@gmail.com. [↑](#footnote-ref-1)
3. “A constatação de que trabalhadores em empregos de maior duração ganham mais a cada ano no emprego do que trabalhadores em empregos de menor duração é confirmada.” (N.T.) [↑](#footnote-ref-2)
4. podendo ser no caso do Logit ou para o Probit. [↑](#footnote-ref-3)