

UTILIZAÇÃO DO MODELO LOGIT PARA ANALISAR OS DETERMINANTES DO DESEMPREGO E INFORMALIDADE NO MUNICÍPIO DE SÃO PAULO

Leandro Batista Duarte¹

Resumo: Diante da importância de compreender melhor o mercado de trabalho, o presente estudo avalia o comportamento dos determinantes do desemprego e informalidade através de um modelo econométrico que estima as chances de um indivíduo encontrar-se em tal *status* de emprego, modelo *Logit*. Os determinantes que apresentaram o maior impacto no sentido de reduzir as chances de pertencerem ao setor informal ou estarem desempregados foram: escolaridade, idade, estado civil, posição no domicílio e raça. Visando realizar uma comparação entre os gêneros, estima também as equações separadamente. Dentre as diferenças encontradas, a principal foi que, embora o impacto da educação sobre a probabilidade estudada tenha sido o mais importante dentre os fatores analisados para ambos os gêneros, ele se mostrou mais marcante entre as mulheres.

Palavras-chave: Mercado de trabalho; Informalidade; Desemprego; São Paulo.

Classificação JEL: J40; C01

USE OF THE LOGIT MODEL TO ANALYZE DETERMINANTS OF UNEMPLOYMENT AND INFORMALITY IN THE MUNICIPALITY OF SÃO PAULO

Abstract: In view of the importance of better understanding the labor market, the present study evaluates the behavior of the determinants of unemployment and informality through an econometric model that estimates the chances of an individual being in such employment status, the Logit model. The determinants that had the greatest impact in order to reduce the chances of belonging to the informal sector or being unemployed were: education, age, marital status, position at home and race. In order to make a comparison between the genders, it also estimates the equations separately. Among the differences found, the main one was that, although the impact of education on the probability studied was the most important among the factors analyzed for both genders, it was more striking among women.

Keywords: Labor market; Informality; Unemployment; Sao Paulo.

JEL Classification: J40; C01

1 Professor da Universidade Estadual de Feira de Santana (UEFS). Doutorando em Economia pela Universidade Federal de Pernambuco (UFPE). Mestre em Economia pela Universidade Federal de Viçosa (UFV) e Graduado em Economia pela Universidade Estadual de Santa Cruz (UESC).

1.Introdução

A grande instabilidade da economia brasileira nos anos 1980 e 1990 agravaram os problemas estruturais do mercado de trabalho. Destacam-se a precarização das formas de contratação, o crescimento do desemprego e a reversão da tendência ao aumento da formalização do trabalho assalariado, verificados no período anterior (BRAGA, 2016). Um dos maiores problemas é o desemprego que é um fenômeno que atinge diariamente milhares de pessoas no mundo todo, afetando diretamente a capacidade de obtenção de renda das pessoas. Esse fato torna o problema do desemprego um assunto amplamente debatido do meio acadêmico, principalmente pelo fato de que o desemprego atinge os diferentes grupos sociais em diferentes magnitudes.

Desse ponto de vista, a investigação da informalização da ocupação e do desemprego fornecem preciosas informações acerca da reconfiguração das formas de inserção no mercado de trabalho e do processo de transformações estruturais em andamento na economia e na própria sociedade brasileira. E, sem dúvida, tal afirmação também é válida para o município de São Paulo que constitui o universo de análise deste estudo. O desemprego na capital paulista é de 15%, índice que representa 1.469.545 pessoas desempregadas. Quanto à informalidade, em 2009, 1,7% da população ocupada de São Paulo trabalhava no comércio de rua. Esta participação, embora relativamente pequena, representa cerca de 100 mil pessoas, cuja presença nas ruas, especialmente quando são considerados os ambulantes, tem efeitos urbanos e socioeconômicos bastante importantes (PAMPLONA, 2013).

Com o objetivo de realizar uma discussão mais pormenorizada sobre o fenômeno de interesse, o presente trabalho analisou os determinantes do desemprego e informalidade para o município de São Paulo, avaliando o comportamento atual dos determinantes no mercado de trabalho, quantificando entre os demais fatores, o impacto da educação nas chances do indivíduo estar desempregado ou no setor informal no mercado de trabalho. Além disso, procura-se analisar o comportamento da informalidade e do desemprego considerando cada gênero, visto que principalmente o desemprego não atinge homens e mulheres de forma homogênea.

O estudo utilizou como técnica econométrica o modelo de probabilidade *Logit*. A técnica da regressão logística surgiu por volta de 1960 como uma resposta ao desafio de realizar previsões estatísticas em situações nas quais a variável dependente é qualitativa. O modelo é amplamente utilizado até os dias atuais em diversas situações. Para tanto, foram utilizados os dados do Censo 2010 pelo Instituto Brasileiro de Geografia e Estatística (IBGE) em nível de área de ponderação. Especificamente, procurou-se identificar a existência ou não de papel decisivo na educação para os mercados de trabalho considerados, assim como comparar os resultados encontrados entre os sexos para o município de São Paulo.

Alguns estudos fizeram análises para o desemprego utilizando os modelos de probabilidade. Oliveira, Pazello e Scorzafave (2009) analisaram o desemprego e a inatividade nas metrópoles brasileiras com foco na diferenciação entre homens e mulheres no período de 1995-2004. Costa e Cunha (2010) usaram o modelo *Probit* com a variável dependente sendo a condição do indivíduo estar ou não desempregado com o objetivo de apontar quais foram os principais determinantes do desemprego no Brasil no período 1981-2005 com ênfase nos efeitos da educação. Araújo, Cunha e Lima (2011) realizaram um estudo

sobre os determinantes do desemprego e da inatividade dos jovens no Brasil metropolitano, tendo como base de dados a PNAD do ano de 2007. Para a informalidade Siqueira (2019) utilizando o modelo *Logit* avaliou os determinantes da informalidade no mercado de trabalho brasileiro. Pereira *et al.*, (2018) analisaram os fatores que determinam a escolha de um indivíduo por um emprego secundário e como tal decisão está relacionada com a informalidade.

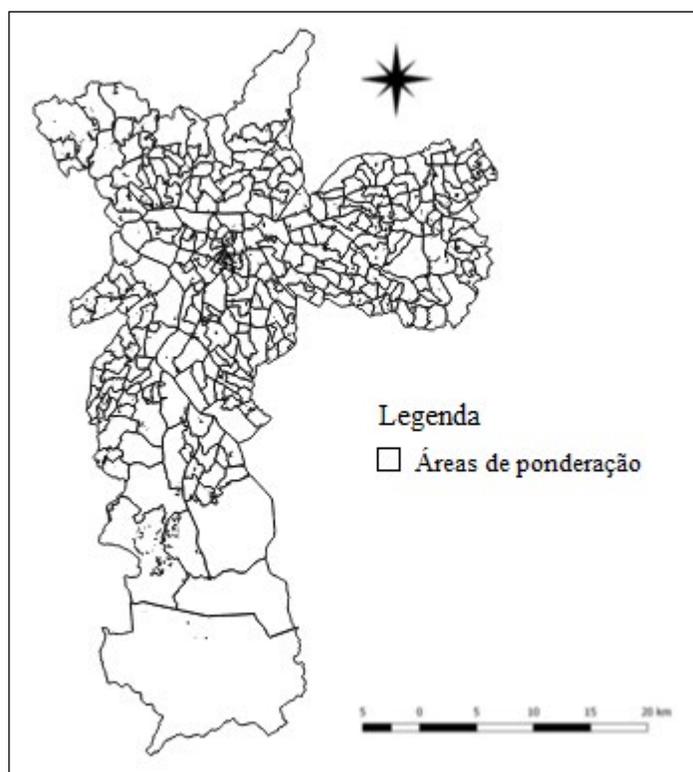
Para cumprir os objetivos propostos, este trabalho está dividido em mais três seções além desta introdução. A segunda seção apresenta a metodologia considerada nessa pesquisa. Na terceira seção, são apresentados os resultados e discussão, que, de modo geral, tratam da probabilidade do indivíduo estar desempregado ou no setor informal no mercado de trabalho. Finalmente, a quarta e última seção exhibe as conclusões do estudo.

2. Metodologia

2.1 Área de Estudo e Dados

A área de estudo é o município de São Paulo, a mais populosa e maior capital brasileira. São Paulo é o principal centro financeiro, corporativo e mercantil da América do Sul. A unidade de análise é o indivíduo dentro da Área de Ponderação (AP), que é definida como uma área composta por um conjunto de zonas mutuamente exclusiva, projetado para dar a robustez estatística necessária à estratégia de amostragem do Censo de População (ver Figura 1). O Censo de 2010 inclui 310 APs para o município de São Paulo. Define-se área de ponderação como sendo uma unidade geográfica, formada por um agrupamento de setores censitários, para a aplicação dos procedimentos de calibração das estimativas com as informações conhecidas para a população como um todo (IBGE, 2010).

Figura 1 – Áreas de Ponderação do município de São Paulo



Fonte: Elaboração própria com base nos dados do Censo 2010.

Os microdados para as características individuais são obtidos no Censo de População de 2010 (Instituto Brasileiro de Geografia e Estatística - IBGE, 2010). O IBGE realiza, a cada dez anos, um Censo Demográfico, com desagregação regional ao nível municipal (ou ao nível da área censitária dos municípios maiores). O Censo Demográfico coleta informações sobre as principais características dos indivíduos e famílias, fornecendo detalhes sobre as condições de vida da população em cada município e servindo como um instrumento de política muito importante em um país. Um questionário mais curto aplica-se a toda a população ao nível do setor censitário, enquanto as características individuais específicas são investigadas em um conjunto mais longo de questões que são dadas a uma amostra e são representativas ao nível das áreas ponderadas (conglomerados de setores censitários com pelo menos 400 domicílios). Para o presente trabalho foi utilizado áreas de ponderação como definição de vizinhança, uma vez que, microdados ao nível individual também estão disponíveis para esta amostra.

2.2 Modelo Econométrico

O objetivo dos modelos de resposta qualitativa é encontrar a probabilidade de acontecimento de um evento, que no caso deste trabalho, é a probabilidade do indivíduo encontrar-se na posição de desempregado ou no setor informal.

Para que a variável dependente seja apresentada de forma qualitativa, faz-se necessário o uso das variáveis binárias ou variáveis *dummies*, como são comumente denominadas na literatura. Esse tipo de variável expressa a presença ou ausência de uma qualidade ou atributo e a forma de se “quantificar” esses atributos, no geral, é formular variáveis artificiais que assumem valores de 0 ou 1, em que 1 indica a presença do atributo e 0 a sua ausência (GUJARATI; PORTER, 2011). Com o objetivo de evitar o problema da perfeita multicolinearidade nas variáveis binárias, ou armadilha da variável *dummy*, como denominam Gujarati e Porter (2011), deve-se excluir uma categoria que passa a ser a referência. Os modelos que consideram o regressando como uma variável qualitativa também são chamados pela literatura de modelos de probabilidade. Neste trabalho é utilizado o modelo chamado *Logit*.

O modelo econométrico a ser apresentado pretende analisar os principais determinantes para o desemprego e informalidade no mercado de trabalho de São Paulo. Para tanto, investigou-se de que forma e com qual intensidade as características pessoais e domiciliares atuavam sobre a probabilidade de se encontrarem desempregado ou no setor informal. As estimativas são realizadas para indivíduos que residem em uma área específica, a fim de capturar o efeito de cada variável dentro de uma estrutura.

Os modelos lineares de regressão são inadequados para estimar as probabilidades, pois, segundo Dias Filho e Corrar (2012), a variável dependente pode assumir valores menores que 0 e maiores que 1. No caso de uma variável dependente na forma binária, o modelo mais adequado é o *Logit*, pois ainda conforme os autores, ele é capaz de converter as observações em razões de chances (*odds ratio*), submetendo cada observação em uma transformação logarítmica. Segundo Dias Filho e Corrar (2012) aplica-se o antilogaritmo sobre os valores estimados dos coeficientes (β). Por exemplo, se o valor estimado de β_1 for 0,5, o seu antilogaritmo será aproximadamente 1,65 ($e^{0,5} \approx 1,65$). Desse modo, para cada unidade de variação em X_{1i} , as chances de que o evento em questão ocorra aumentam em aproximadamente 65% quando todas as demais variáveis são mantidas constantes. O modelo matemático pode ser explicitado da seguinte forma:

$$OddsRatio = \frac{P(sucesso)}{1 - P(sucesso)} \quad (1)$$

Partindo da equação (1), o valor do *Logit* (L_i) é encontrado ao se obter o logaritmo natural da razão de chances, ou *odds ratio*, como será denominada na análise deste trabalho.

$$L_i = \ln \left(\frac{P(sucesso)}{1 - P(sucesso)} \right) = \beta_0 + \beta_1 X_{1i} + \dots + \beta_k X_{ki} \quad (2)$$

Assim, o lado esquerdo da equação (2) apresenta o logaritmo natural da razão de chances e o lado direito as variáveis independentes, que podem ser categóricas (*dummies*) ou métricas, apresentadas em valores correntes. A razão de chances estimada ainda pode ser obtida elevando a constante matemática “e” ao lado direito da equação (2):

$$L_i = \frac{P(sucesso)}{1 - P(sucesso)} = e^{\beta_0 + \beta_1 X_{1i} + \dots + \beta_k X_{ki}} \quad (3)$$

Com base no conceito de razão de chances apresentado na equação (3), tem-se que a probabilidade de ocorrência de um determinado evento é dada pela seguinte fórmula:

$$P(\text{sucesso}) = \frac{e^{(\beta_0 + \beta_1 X_{1i} + \dots + \beta_k X_{ki})}}{1 + e^{(\beta_0 + \beta_1 X_{1i} + \dots + \beta_k X_{ki})}} \quad (4)$$

Dividindo o numerador e o denominador por e com sua potência, é possível simplificar a equação (4) da seguinte forma:

$$P(\text{sucesso}) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_{1i} + \dots + \beta_k X_{ki})}} \quad (5)$$

Segundo Dias Filho e Corrar (2012), na fase de estimação do modelo logístico de regressão não é permitida a utilização do método dos mínimos quadrados ordinários (MQO), cujo objetivo é minimizar a soma dos quadrados das diferenças entre os valores previstos e os observados para a variável dependente. Neste caso, deve-se aplicar o método da máxima verossimilhança, que trata de uma forma de estimar parâmetros de distribuição de probabilidades que maximizem a função verossimilhança.

O modelo utilizado para a operacionalização deste trabalho é o modelo de regressão logística, ou simplesmente *Logit*, como é comumente chamado na literatura. Conforme apresentado acima, trata-se de um modelo de probabilidade de resposta qualitativa, ou seja, a variável dependente Y é qualitativa e se apresenta na forma binária. A variável dependente do modelo apresentado neste estudo é a condição de desemprego ou estar na informalidade, que recebe valor um ($Y=1$) caso o indivíduo esteja na condição de desemprego ou no setor informal e valor zero ($Y=0$) caso o indivíduo não esteja nestas condições. Na prática, o modelo deste trabalho segue a seguinte explicitação matemática:

$$L_i = \frac{P(\text{desemprego})}{1 - P(\text{desemprego})} = \beta_0 + \beta_1 X_i + \beta_2 H_j + \mu \quad (6)$$

$$L_i = \frac{P(\text{informal})}{1 - P(\text{informal})} = \beta_0 + \beta_1 X_i + \beta_2 H_j + \mu \quad (7)$$

Em que X_i denota o vetor de características pessoais; H_j representa o vetor de características domésticas e β são vetores de parâmetros que serão estimados; e μ , o termo de erro aleatório. A razão de chances (*odds-ratio*) é representada pelo lado esquerdo da equação, já que o evento que representa o sucesso ($Y=1$) é a condição de desemprego ou informalidade. A probabilidade do evento de sucesso é representada por $P(\text{desemprego})$ ou $P(\text{informal})$ e a probabilidade de insucesso do evento é dada por $1 - P(\text{desemprego})$ ou $1 - P(\text{informal})$, já que a máxima probabilidade de ocorrência de um evento é igual a um, e neste caso a probabilidade apenas assume os valores 0 (no caso de indivíduo empregado ou no setor formal) ou 1 (para o indivíduo desempregado ou na informalidade). Posteriormente, serão estimados os efeitos marginais das variáveis, que serão considerados nas análises.

Os efeitos marginais representam, por exemplo, o quanto uma mudança na variável educação pode afetar a probabilidade de um indivíduo estar desempregado ou na informalidade. Os efeitos marginais (a variação da probabilidade de o evento ocorrer

quando uma variável independente é modificada) são calculados por meio dos coeficientes estimados β 's.

As variáveis dependentes e explicativas utilizadas na equação podem ser observadas no Quadro 1, a seguir. A variável dependente “desemprego” é possível verificar a condição de desemprego do indivíduo, sendo a posição de empregado a categoria de referência. Do mesmo modo para “informal”, condição de estar na informalidade.

No caso da variável “chefe”, o IBGE disponibiliza oito classificações, a saber: pessoa de referência; cônjuge; filho; outro parente; agregado; pensionista; empregado doméstico e parente do empregado doméstico. Para fins de simplificação, este trabalho considerou apenas as quatro primeiras categorias, de forma que a pessoa de referência recebe valor 1 e é chamada de “chefe de família”, o que caracteriza os indivíduos “não-chefes” como sendo a categoria de referência composta por indivíduos classificados como cônjuge, filho e outro parente. De acordo com a definição do IBGE, a pessoa de referência é aquela responsável pela família ou que assim foi considerada pelos demais membros da família, e espera-se que o desemprego e a informalidade entre esses indivíduos seja relativamente menor, uma vez que são eles os responsáveis pela maior parte da renda familiar.

Quadro 1 – Descrição das variáveis dos modelos econométricos.

Amostra	
VARIÁVEIS	DESCRIÇÃO
Informal (<i>dummy</i>)	Assume valor 1 se a pessoa estar na informalidade
Desemprego (<i>dummy</i>)	Assume valor 1 se a pessoa estar desempregada
Masculino	Variável <i>dummy</i> se for do sexo masculino
Idade	Em anos
Idade ²	<i>Proxy</i> para experiência
Educação (<i>dummies</i>)	
i) Baixo	i) Sem ensino fundamental completo,
ii) Intermediário	ii) Com ensino fundamental completo, mas sem médio completo,
iii) Alto	iii) Com ensino médio completo, mas sem superior,
iv) Bem elevado	iv) Com superior completo ou mais
Estado Civil	Variável <i>dummy</i> se o individuo for casado
Posição no domicílio (<i>dummy</i>)	Chefe Cônjuge Filho
Raça (<i>dummies</i>)	Branco Pardo Negro Amarelo
Localidade	Variável <i>dummy</i> se o domicilio pertence a área urbana
Acesso à infraestrutura (<i>dummies</i>)	Coleta de lixo Saneamento básico

Fonte: Elaboração própria.

Os microdados para as características individuais, incluindo aqueles relacionados a ser um trabalhador informal ou desempregado, são obtidos no Censo de População de 2010 (Instituto Brasileiro de Geografia e Estatística, 2010). Para determinar o *status* de cada trabalhador, utilizamos variáveis sobre *status* ocupacional e tipo de emprego (no trabalho principal).

No que diz respeito aos conceitos de informalidade, as diversas definições podem levar a resultados substancialmente diferentes, dificultando uma comparação entre estudos. Uma classificação tradicional na literatura brasileira enquadra como trabalhadores informais, àqueles que não possuem carteira de trabalho assinada, tal como o trabalho de Fernandes (1996), Carneiro e Henley (2001), Pianto e Pianto (2002) e Menezes-Filho, Mendes e Almeida (2004). Estudos recentes, como os de Dalberto (2014), Cirino e Dalberto (2014) estabelecem como conceito de trabalho informal aquele que não contribui para o Instituto Nacional do Seguro Social. Outros trabalhos como de Machado, Oliveira e Antigo (2008) e Duarte *et al.*, (2018) combina a definição de regulação do trabalho, ou seja, carteira assinada como característica de formalidade e sem carteira assinada como informalidade, e a definição de subordinação. Dessa forma, o trabalhador informal é o que se autodeclara conta própria, exceto as ocupações de profissionais liberais, e os empregados sem carteira assinada.

Apesar de existirem diversas definições para a informalidade, no trabalho foi considerada a definição mais simples possível, o trabalhador apenas sem carteira assinada para representar a informalidade e o trabalhador que está na População Economicamente Ativa (PEA), mas se encontra desocupado para representar o indivíduo desempregado.

Mais especificamente, as variáveis pessoais absorvem as diferentes preferências com relação ao espaço de acordo com certas características. Por exemplo, quanto à *dummy* para gênero, regularmente, observa-se nos trabalhos empíricos que as mulheres tendem a se localizar mais próximos aos locais de trabalho, o que é comumente atribuído à noção de maior responsabilidade para com as tarefas com lar atribuído às mulheres o que podem aumentar a informalidade (LEE & MACDONALD, 2003; CRAINE, 2007). De forma similar, a condição de casado pode atribuir um caráter mais permanente à localização dos domicílios e, assim, menor vinculação imediata ao local de trabalho. As variáveis quanto à idade e de escolaridade traduzem possíveis influências da disposição à locomoção e preferências quanto à vizinhança por parte dos indivíduos.

Discutindo as hipóteses acerca do relacionamento entre as variáveis explicativas consideradas e a probabilidade do indivíduo estar na informalidade ou desempregado (Y_i), espera-se, *a priori*, um relacionamento negativo quanto à escolaridade, já que quanto mais anos de estudo, melhores as oportunidades de emprego e as chances de acesso a rendas elevadas.

Com relação à idade e idade², utilizada como *proxy* da experiência, acredita-se em uma relação negativa também. Tal relacionamento indica que as oportunidades de trabalho e os salários aumentam à medida que a pessoa adquire mais experiência até certo ponto. Em relação à condição no domicílio, é esperado que os chefes apresentem menor probabilidade de serem informais ou desempregados do que os outros grupos, por serem aqueles geralmente os responsáveis por grande parte do sustento de seus membros familiares.

Passando para os coeficientes das variáveis de infraestrutura, acredita-se que eles sejam negativos, visto que ter melhor acesso ao saneamento e coleta de lixo, podem indicar menores as chances de informalidade e desemprego. O mesmo raciocínio vale para as variáveis que compõem a vizinhança.

Existem algumas limitações com os dados utilizados neste estudo. Em primeiro lugar, os dados mais atuais para o Censo é para o ano de 2010, visto que não foram disponibilizados oficialmente os dados do Censo 2020. Em segundo lugar, o *status* de informalidade é derivado dos dados do recenseamento e, portanto, depende de pessoas que relatem com sinceridade seu *status*. Como algumas pessoas podem querer disfarçar que estão trabalhando informalmente, os dados do estudo podem estar superestimando o número de indivíduos formalmente empregados. No entanto, dado o tamanho da amostra relativamente grande, os dados permitem explorar a relação entre informalidade e seus determinantes.

3. Resultados do estudo

As médias e os desvios padrão das variáveis incluídas no modelo para o ano de 2010, levando em conta os dados do Censo, são apresentados na Tabela 1, a seguir. No caso das variáveis *dummies*, a média é a proporção de casos em que a variável assume o valor 1.

Em relação aos anos de estudo, as estatísticas confirmaram que os indivíduos apresentam, em média, nível educacional mais baixo, com maiores taxas para indivíduos sem ensino fundamental completo. Quanto à posição no domicílio, verificou-se, que os chefes e filhos são a maioria, representando, em geral, 31% e 34% do percentual da amostra, respectivamente. Os dados mostrados ainda apresentaram a existência de maior proporção de pessoas que se declaram brancas e pardas no município de São Paulo.

A estimação das equações foram realizadas, separadamente, para homens e mulheres, sendo os resultados gerais apresentados na Tabela 2 e aqueles referentes a gêneros apresentados nas Tabelas 3 e 4. São apresentados os coeficientes das variáveis elencadas como possíveis determinantes do desemprego, o valor do teste Z e o valor da razão de chances de ocorrência do evento. As Tabelas também mostram na parte inferior algumas medidas de avaliação comumente usadas na literatura para modelos logísticos.

Tabela 1 - Estatística Descritiva das variáveis da equação no mercado de trabalho, 2010.

Variáveis	Média	Desvio Padrão	Mínimo	Máximo
Desemprego	0,040	0,196	0	1
Informal	0,074	0,263	0	1
Masculino	0,474	0,499	0	1
Feminino	0,525	0,499	0	1
Idade	33,35	20,47	0	138
Idade ²	1531,8	1615,6	0	19044
Escolaridade				
Baixo	0,467	0,498	0	1
Intermediário	0,161	0,368	0	1
Alto	0,233	0,422	0	1

Variáveis	Média	Desvio Padrão	Mínimo	Máximo
Bem Elevado	0,137	0,344	0	1
Estado Civil				
Casado	0,315	0,464	0	1
Posição no domicílio				
Chefe	0,313	0,464	0	1
Conjuge	0,190	0,392	0	1
Filho	0,343	0,474	0	1
Outro	0,151	0,358		
Raça				
Branca	0,603	0,489	0	1
Negro	0,064	0,245	0	1
Pardo	0,310	0,462	0	1
Amarelo	0,022	0,147	0	1
Localidade				
Urbano	0,989	0,101	0	1
Infraestrutura				
Coleta	0,940	0,236	0	1
Saneamento	0,901	0,297	0	1

Fonte: Elaboração própria com base nos dados da pesquisa.

No que tange à significância, observou-se que a grande maioria das variáveis foi significativa a 1%, ressaltando a importância de tais regressores para a explicação da probabilidade do indivíduo ser informal ou estar desempregado no mercado de trabalho. Observando-se as estimativas, nota-se que, embora o coeficiente de determinação da regressão (R^2) não seja muito elevado, a estimativa para os coeficientes apresentam valores elevados e com sinal esperado como em diversos trabalhos que utilizam tal metodologia. Dentre esses trabalhos pode-se citar o estudo de Souza e Cunha (2012), que utilizou o modelo logit com base nos microdados da PNAD e apresentou valores de Pseudo R^2 de 0,0492; 0,0540 e 0,0374 para a vitimização por furto e/ou roubo, tentativa e agressão, respectivamente, no ano de 1988. O trabalho de Scorzafave (2001) apresentou valores de Pseudo R^2 de 0,1213; 0,1172; 0,1020 e 0,1133 para os anos de 1982, 1987, 1992 e 1997, respectivamente. Costa e Cunha (2010) utilizaram o modelo *probit* e encontraram valores de Pseudo R^2 de 0,0776 e 0,0726 para homens e mulheres, respectivamente no período de 1981 a 2005.

Tabela 2 - Resultado das equações para o município de São Paulo, 2010.

Variáveis	DESEMPREGO			INFORMALIDADE		
	E.M.	D.P	Odds Ratio	E.M.	D.P	Odds Ratio
Masculino	-0,0253***	0,0008	0,6378***	-0,0559***	0,0013	0,6283***
Idade	-0,0057***	0,0001	0,9017***	-0,0067***	0,0002	0,9445***
Idade ²	0,0001***	0,0000	1,0009***	0,0001***	0,0000	1,0006***
Baixo	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Intermediário	0,0004	0,0011	1,0072	-0,0331***	0,0015	0,7412***
Alto	-0,0158***	0,0009	0,7417***	-0,0801***	0,0014	0,4796***
BemElevado	-0,0445***	0,0009	0,3652***	-0,1201***	0,0013	0,2769***
Casado	-0,0093***	0,0010	0,8420***	-0,0322***	0,0015	0,7584***
Chefe	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Cônjuge	0,0074***	0,0013	1,1383***	0,0046**	0,0018	1,0400**
Filho	0,0313***	0,0015	1,6479***	-0,0034*	0,0020	0,9712*
Outro	0,0081***	0,0015	1,1504***	-0,0025	0,004	1,1504
Branco	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Negro	0,0124***	0,0017	1,2309***	0,0173***	0,0026	1,1505***
Pardo	0,0092***	0,0009	1,1752***	0,0119***	0,0015	1,1046***
Amarelo	-0,0048	0,0031	0,9126	-0,0013	0,0048	0,9887
Urbano	-0,0082**	0,0041	0,8690**	-0,0187***	0,0067	0,8605***
Coleta	-0,0052***	0,0017	0,9118***	-0,0061**	0,0028	0,9502**
Saneamento	-0,0076***	0,0014	0,8769***	-0,0222***	0,0024	0,8369***
Constante	-	-	1,5066***	-	-	1,8314***
Observações	289.213			Observações		267.344
Pseudo R ²	0.0817			Pseudo R ²		0.0492
Prob > chi2	0,0000			Prob > chi2		0,0000
LR chi2	12662.66			LR chi2		11245.90
%Predições Corretas	92,44%			%Predições Corretas		84,73%

Fonte: Resultados da pesquisa

Nota: Desvio-padrão entre parênteses. ***p < 0.01, **p < 0.05, *p < 0.1. E.M.: efeito marginal.

Por fim, a estatística LR chi² é estatisticamente significativa e assume um valor que facilmente excede 10. No que diz respeito aos testes de classificação do modelo, a Tabela 2 mostra que em 2010 o modelo para desemprego conseguiu classificar corretamente 92,44% das observações analisadas, enquanto que para a informalidade esse valor foi de 84,73%. A seguir foi realizada uma análise dos resultados estimados com o modelo logit por meio da interpretação dos sinais e magnitudes dos coeficientes, bem como das razões de chances (odds ratio) vinculadas a cada variável. É necessário ressaltar que a análise de uma variável supõe todas as outras constantes (*ceteris paribus*).

Com relação aos resultados, todas as variáveis incorporadas foram importantes para a explicação da situação dos indivíduos no mercado de trabalho. No que se refere ao mercado de trabalho para o município de São Paulo, no geral, a informalidade e desemprego foram influenciadas pela escolaridade, idade, estado civil, raça e infraestrutura (Tabela 2). Os efeitos marginais comportaram-se de acordo com o esperado. Sobre os efeitos marginais, é importante destacar que a sua interpretação em modelos de variável dependente binária,

é a seguinte: i) para variáveis contínuas, representa o impacto médio em termos de pontos percentuais na probabilidade de um indivíduo estar na informalidade ou desempregado para aumento unitário da variável explicativa contínua considerada; e ii) para as variáveis qualitativas, determina a mudança média na probabilidade de um indivíduo estar no setor informal ou desempregado, em pontos percentuais, devido à presença da característica indicada pela *dummy* considerada.

Iniciando a discussão dos determinantes do *status* do indivíduo no mercado de trabalho, verificou-se a ocorrência de sinal negativo para a *dummy* de sexo, refletindo o efeito menor para trabalhadores do sexo masculino em estarem na condição de trabalhador informal ou desempregado (Tabela 2). Observando as razões de chances – *odds ratio* – percebe-se que o sexo masculino obtém chance de estar desempregado ou na informalidade de 36% e 37,2% menor comparado ao sexo feminino. Para fazer a interpretação em termos percentuais, bastou subtrair 1 do coeficiente e multiplicar por 100. Infelizmente, a divisão tradicional das atividades familiares e a consequente maior experiência da mulher em atividades domésticas direcionam essa força de trabalho secundária para o emprego de serviços domésticos (BALSADI, 2000). Outra explicação é que, na maioria dos casos, as mulheres não são chefes de famílias, portanto, buscam empregos parciais, sazonais ou temporários menos remunerados do que empregos em tempo integral. Além disso, o duplo papel de trabalhadora, ou seja, ser dona de casa/mãe restringe os deslocamentos para buscarem empregos no setor formal.

Em relação à idade, os resultados mostraram que quanto mais experiência para o trabalhador, menores as chances de pertencerem na informalidade ou desempregados no mercado de trabalho em São Paulo. Mesmo resultado obtido quando considera o indivíduo casado. Isso pode ser explicado pelo fato de que os melhores salários pagos são os trabalhos formais, dessa forma, os indivíduos estão mais dispostos a buscarem melhores empregos para obter condições para sustento da família. Então ser casado, diminui as chances de pertencer ao setor informal e estarem desempregados em 0,0093 e 0,0322 pontos percentuais com relação ao ponto médio da amostra, respectivamente. Pelas razões de chances, pode-se dizer que o indivíduo casado tem chance de estar no setor informal ou desempregado de 24% e 15,8% a mais do que um indivíduo na condição de não casado.

Com relação à escolaridade, espera-se que seja um fator determinante na obtenção de um emprego, já que é por meio dos estudos que um indivíduo consegue ampliar o seu conhecimento. Conforme esperado, a maioria dos coeficientes apresentaram sinais negativos, já que o grupo-base adotado foram os indivíduos com menor grau de instrução. Dessa forma, em relação a esses últimos, as pessoas pertencentes às demais categorias apresentaram maiores chances de pertencerem ao setor informal ou estarem no desemprego, isto é, esta probabilidade diminui conforme o aumento da escolaridade. Percebe-se que a magnitude dos efeitos marginais também aumenta à medida que aumenta o grau de instrução.

Passando para a variável de posição no domicílio, verificou-se, no mercado diferença estatística entre os trabalhadores chefes e para os outros no domicílio. Analisando o desemprego e a informalidade, percebe-se que o chefe de domicílio tem menor probabilidade de estar desempregado ou no setor informal em comparação quando cônjuge. Na Tabela 2, o fato de o indivíduo chefe ser a pessoa de referência faz que a referida probabilidade seja de

0,0046 e 0,0074 pontos percentuais menores do que se ele fosse cônjuge, respectivamente. Mesma interpretação para filho e outros analisando para o desemprego. Apenas para a informalidade estas mesmas variáveis apresentaram sinais invertidos, porém significativo apenas para quando o indivíduo é filho no domicílio.

Este resultado está em conformidade com o apresentado no trabalho de Oliveira, Scorzafave e Pazello (2009), o qual justifica a menor probabilidade de desemprego do chefe ao fato de que o custo de oportunidade de desemprego para esses indivíduos é maior, uma vez que eles são os maiores provedores dos recursos domiciliares. O resultado também vai de encontro com o trabalho de Silva e Kassouf (2002), que justifica a probabilidade negativa do chefe de família pela busca mais agressiva e persistente de emprego, uma vez que necessita da renda para manter sua família.

Quanto ao impacto da raça sobre a probabilidade dos indivíduos estarem na informalidade ou desempregados, verificou-se que o fato de serem negros ou pardos faz com que a probabilidade seja superior em comparação com o grupo de brancos. Tal resultado para a informalidade pode estar refletindo a maior necessidade de trabalhar daquelas raças em virtude de, geralmente, elas se situarem em classes sociais mais baixas do que as brancas (Cirino & De Lima, 2011), fazendo com que trabalhar no setor informal seja uma forma de não permanecer desempregada. Além do mais, sabe-se que existe a discriminação no mercado de trabalho, onde trabalhadores negros sofrem com isso constantemente.

Por fim, quanto as variáveis que representam a infraestrutura, os domicílios com melhor acesso a coleta de lixo e com rede de saneamento básico, diminui a chance de estarem na informalidade ou desempregados. A localização do domicílio também foi uma variável que mostrou significância, mostrando que indivíduos que residem em áreas urbanas têm maiores chances de estarem desempregados ou na informalidade comparados aos domicílios localizados em áreas rurais.

A relação entre os determinantes da informalidade e desemprego mostram padrões mais distintos quando permite a heterogeneidade entre os gêneros. Nas Tabelas 3 e 4 têm-se efeitos heterogêneos para homens e mulheres.

Tabela 3 - Resultado das equações para informalidade por gênero para o município de São Paulo, 2010.

Variáveis	Homens		Mulheres	
	E.M.	D.P	E.M.	D.P
Idade	-0,0085***	0,0003	-0,0046***	0,0004
Idade ²	0,0001***	0,0000	0,0001***	0,0000
Baixo	Ref.	Ref.	Ref.	Ref.
Intermediário	-0,0242***	0,0020	-0,0450***	0,0024
Alto	-0,0566***	0,0018	-0,1075***	0,0022
BemElevado	-0,0815***	0,0018	-0,1638***	0,0021
Casado	-0,0298***	0,0020	-0,0243***	0,0025
Chefe	Ref.	Ref.	Ref.	Ref.
Conjuge	0,0147***	0,0026	-0,0154***	0,0027
Filho	0,0170***	0,0027	-0,0308***	0,0030

Variáveis	Homens		Mulheres	
	E.M.	D.P	E.M.	D.P
Outro	0,0174***	0,0030	-0,032***	0,0030
Branco	Ref.	Ref.	Ref.	Ref.
Negro	0,0101***	0,0033	0,0259***	0,0043
Pardo	0,0010	0,0019	0,0245***	0,0024
Amarelo	0,0013	0,0061	-0,0064	0,0075
Urbano	-0,0161**	0,0084	-0,0247**	0,0112
Coleta	-0,0017	0,0035	-0,0113**	0,0045
Saneamento	-0,0150***	0,0030	-0,0321***	0,0039
Constante	-	-	-	-
Observações	143.153		124.191	
Pseudo R ²	0,0386		0,0571	
LR chi2	4235,70		6696,96	

Fonte: Resultados da Pesquisa.

Nota: Desvio-padrão entre parênteses. ***p < 0.01, **p < 0.05, *p < 0.1. E.M.: efeito marginal.

Observando todos os efeitos marginais, percebe-se que os sinais parecem corroborar com a literatura para a escolaridade, por exemplo, trabalhadores mais escolarizados tem menor probabilidade de trabalhar no setor informal ou de estarem desempregados. Comparados aos resultados gerais, não houve muita diferença entre os efeitos e sinais esperados. Os resultados para idade, escolaridade, casado, localidade e infraestrutura permanecem no mesmo padrão. Contudo, pode-se observar que algumas variáveis mudam seu efeito quando comparadas entre os sexos.

Tabela 4 - Resultado das equações para desemprego por gênero para o município de São Paulo, 2010.

Variáveis	Homens		Mulheres	
	E.M.	D.P	E.M.	D.P
Idade	-0,0043***	0,0001	-0,0071***	0,0002
Idade ²	0,0001***	0,0000	0,0001***	0,0000
Baixo	Ref.	Ref.	Ref.	Ref.
Intermediário	-0,0023**	0,0012	0,0037**	0,0019
Alto	-0,0149***	0,0011	-0,0172***	0,0016
BemElevado	-0,0323***	0,0011	-0,0582***	0,0016
Casado	-0,029***	0,0013	-0,0026	0,0017
Chefe	Ref.	Ref.	Ref.	Ref.
Conjuge	-0,0005	0,0018	-0,0001	0,0019
Filho	0,0492***	0,0022	0,0103***	0,0021
Outro	0,0264***	0,0022	-0,0116***	0,0020
Branco	Ref.	Ref.	Ref.	Ref.
Negro	0,0099***	0,0020	0,0159***	0,0029
Pardo	0,0040***	0,0011	0,0153***	0,0015
Amarelo	-0,0017	0,0037	-0,0094*	0,0050
Urbano	0,0018	0,0043	-0,0216***	0,0073
Coleta	-0,0055***	0,0021	-0,0048*	0,0028

Variáveis	Homens		Mulheres	
	E.M.	D.P	E.M.	D.P
Saneamento	-0,0048***	0,0017	-0,0110***	0,0024
Constante	-	-	-	-
Observações	152.665		Observações	136.548
Pseudo R ²	0,0910		Pseudo R ²	0,0729
LR chi2	6481,61		LR chi2	6044,44

Fonte: Resultados da Pesquisa.

Nota: Desvio-padrão entre parênteses. ***p < 0.01, **p < 0.05, *p < 0.1. E.M.: efeito marginal.

O tamanho da informalidade e desemprego econômico é uma resposta à incapacidade da economia formal para acomodar todos os trabalhadores, então a estrutura econômica formal determinará quais indivíduos se dedicam a atividades econômicas informais, ao desemprego e onde residem (não onde eles decidem viver). O mercado de trabalho informal para os homens segue um padrão um pouco distinto comparado às mulheres quanto a posição no domicílio, devido que, em geral, cabe ao gênero masculino maior imposição na busca por emprego, sendo que na maioria das vezes não é possível o trabalho formal, tendo que recorrer à informalidade.

Em termos de tamanho do efeito marginal da escolaridade, verificou-se que, de maneira geral, tal variável provocou grande impacto sobre a probabilidade de o indivíduo no mercado de trabalho, principalmente entre as mulheres. Em termos de diferenças entre os gêneros, verificou-se que os efeitos marginais das mulheres se apresentaram superiores aos dos homens, indicando que as mulheres são mais sensíveis. Tal resultado reforça o importante papel da escolaridade para a redução da informalidade feminina no mercado de trabalho (Tabela 4).

4. Conclusão

O presente trabalho analisou os determinantes para o desemprego e informalidade nos mercados de trabalho do município de São Paulo a partir dos dados do Censo 2010 por área de ponderação. Por meio do modelo *Logit* e do cálculo das razões de chances foi possível verificar a direção e a magnitude do impacto de cada variável na determinação da condição de desemprego e informalidade. Os resultados apresentados forneceram uma visão da probabilidade de pertencerem nesses *status* de emprego.

As variáveis escolaridades, idade, estado civil, posição no domicílio, raça, localidade e infraestrutura, apresentaram-se como determinantes significativos para a probabilidade dos indivíduos pertencerem ao setor informal ou estarem desempregados. Quanto aos resultados obtidos por meio do método econométrico e a aplicação do modelo *logit* permitem sugerir que o nível educacional do indivíduo, seja ele homem ou mulher, constitui em um dos principais determinantes da probabilidade de desemprego ou informalidade, principalmente o mais alto nível de escolaridade. Na comparação entre os resultados gerais com os resultados separados por sexos, não se verificou diferença marcante entre os determinantes analisados, apesar de pequenas diferenças nas magnitudes dos coeficientes.

Por outro lado, na comparação entre os sexos, foram notadas algumas diferenças importantes entre as equações: i) o fato de ser a pessoa chefe do domicílio comparado com cônjuge, filho ou outro mostrou impacto negativo para as mulheres na informalidade, uma vez que, geralmente, são as mulheres que assumem, com maior intensidade, o papel secundário no provimento da renda domiciliar, da mesma forma que, quando comparadas com os homens, assumem com mais frequência às atividades domésticas, como o fato também de mulheres casadas geralmente terem filhos, o que diminui as chances de participação nos empregos formais, em virtude de estarem mais associadas às mulheres a educação e os cuidados relativos às crianças. As mulheres acabam buscando empregos mais próximos de casa e acabam recorrendo à informalidade; ii) os efeitos marginais da escolaridade para as mulheres apresentaram-se superiores aos dos homens, indicando que as mulheres são mais sensíveis a informalidade e desemprego comparadas aos homens quando analisadas a escolaridade.

Sobre a variável escolaridade, ressalta-se que ela apresentou grande impacto sobre a probabilidade no mercado de trabalho. Nesse sentido, como sugerido por Cirino & De Lima (2011) as políticas públicas feitas pelo governo federal deve não apenas continuar com os investimentos em educação como aumentá-los, uma vez que tais ações contribuem fortemente para a possibilidade da redução da informalidade e desemprego no mercado de trabalho. Tendo em vista os resultados encontrados, esse estudo é de extrema importância para o meio econômico e social, bem como contribui para novas pesquisas, visto a necessidade da implantação de novos projetos de incentivo a educação, e a continua ampliação das políticas sociais para a educação de nível superior.

Referências

- BALSADI, O. Características do emprego rural no Estado de São Paulo nos anos 90. Campinas: Unicamp, 2000. 169 f. Dissertação (Mestrado em Desenvolvimento Econômico, Espaço e Meio Ambiente) – Programa em Economia Agrícola e Agrária. Campinas, 2000.
- BRAGA, T. S. O setor informal e as formas de participação na produção: os casos das Regiões Metropolitanas de Salvador e Recife. **Anais**, 2016, 1-23.
- CARNEIRO, F.G.; HENLEY, A. Modelling formal vs. informal employment and earnings: micro-econometric evidence for Brazil. Encontro Nacional de Economia. **Anais**. ANPEC, 2001.
- CIRINO, J. F; DE LIMA, J. E. Determinantes da participação feminina no mercado de trabalho: uma comparação entre os sexos e entre os mercados das regiões metropolitanas de Belo Horizonte e Salvador. **Revista Econômica do Nordeste**, 2011, 42.1: 165-182.
- COSTA, J. S; CUNHA, M. S. Determinantes do desemprego no Brasil no período de 1981 a 2005: Uma análise enfatizando a qualificação do indivíduo em um contexto de

maior abertura comercial. **Análise Econômica**, Porto Alegre, ano 28, n. 53, p. 197-220, mar. 2010.

DALBERTO, C.R. Formalidade vs. Informalidade no mercado de trabalho brasileiro: uma investigação dos diferenciais de rendimento. Viçosa, MG: UFV, 2014. 87 f. Dissertação (Mestrado em Economia) – Programa de Pós-Graduação em Economia da Universidade Federal de Viçosa. Viçosa, 2014

DIAS FILHO, J.M.; CORRAR, L.J. Regressão logística. In: CORRAR, L.J; PAULO, E.; DIAS FILHO, J.M. (Orgs.). **Análise multivariada para os cursos de Administração, Ciências Contábeis e Economia**. São Paulo: Atlas, 2012, p. 280-323.

DUARTE, L. B; CIRINO, J. F; SETTE, A. B. P. Diferencial de salários no mercado formal/informal para as regiões metropolitanas do Nordeste. **Revista Econômica do Nordeste**, 2018, 49.1: 9-24.

FERNANDES, R. Mercado de trabalho não-regulamentado: participação relativa e diferenciais de salários. **Pesquisa e Planejamento Econômico**, v.26, n. 3, p. 417-442, 1996.

GUJARATI, D. N; PORTER, D. C. **Econometria Básica**. 5 Ed. Porto Alegre: AMGH Editora, 2011. p. 552-562.

IBGE. Censo Demográfico 2010. Notas Técnicas. Rio de Janeiro, Instituto Brasileiro de Geografia e Estatística, 2010. IBGE. Censo Demográfico 2010, Microdados da Amostra de Domicílios. Rio de Janeiro, Instituto Brasileiro de Geografia e Estatística, 2010.

IBGE. Censo Demográfico 2010, Microdados da Amostra de Pessoas. Rio de Janeiro, Instituto Brasileiro de Geografia e Estatística, 2010.

LEE, B. S; MACDONALD, J. F. (2003). Determinants of commuting time and distance for Seoul Residents: The impact of family status on the commuting of women. *Urban Studies*, Vol. 40, No. 7, 1283-1302,

MACHADO, A.F.; OLIVEIRA, A.M.H.C.; ANTIGO, M. Evolução do diferencial de rendimentos entre setor formal e informal no Brasil: o papel das características não observadas. **Revista de Economia Contemporânea**, v. 12, n. 2, p. 355-388, 2008.

MENEZES-FILHO, N.A.; MENDES, M.; ALMEIDA, E.S. O diferencial de salários formal-informal no Brasil: segmentação ou viés de seleção? **Revista Brasileira de Economia**, v. 58, n. 2, p. 235-248, 2004.

OLIVEIRA, P. R; SCORZAFAVE, L. G; PAZELLO, E. T. Desemprego e inatividade nas metrópoles brasileiras: as diferenças entre homens e mulheres. **Nova Economia**, Belo Horizonte, v. 19, n. 2, p. 291-324, mai/ago. 2009.

PAMPLONA, J. B. Mercado de trabalho, informalidade e comércio ambulante em São Paulo. **Revista Brasileira de Estudos de População**, 2013, 30.1: 225-249.

PEREIRA, R. M; GALVÃO, M. C; MAXIR, H. S. Determinantes do emprego secundário e informalidade: evidências adicionais para o mercado de trabalho brasileiro. **Revista Brasileira de Estudos de População**, 2018, 35.3.

PIANTO, M. E. T; PIANTO, D. M.: “Informal Employment in Brazil – A Choice at the Top and Segmentation at the Bottom: A Quantile Regression Approach”, Texto para discussão nº 236. Brasília, agosto de 2002.

SCORZAFAVE, L. G. D. S. A evolução e os determinantes da participação feminina no mercado de trabalho brasileiro. 2001. 66 f. Dissertação (Mestrado em Economia) – Faculdade de Economia, Administração e Contabilidade, Universidade de São Paulo, São Paulo, 2001. Disponível em: < <http://www.teses.usp.br/teses/disponiveis/12/12138/tde-17022003-101948/pt-br.php>>. Acesso em: 15 abr. 2020.

SILVA, N. D. V; KASSOUF, A. L. A exclusão social dos jovens no mercado de trabalho brasileiro. **Revista Brasileira de Estudos de População**, Rio de Janeiro, v. 19, n. 2, p. 99-115, 2002.

SIQUEIRA, P. B. B. **Determinantes da informalidade no Brasil**: uma abordagem microeconômica. Duque de Caxias – RJ, 2019.

SOUZA, J. P. C; CUNHA, M. S. Uma análise do perfil da vitimização no Brasil. In: XV Encontro de Economia da Região Sul, 2012, Porto Alegre. **Anais eletrônicos...** Porto Alegre: PUC-RGSul, 2012. v. 1. p. 1-20. Disponível em: <<http://www.pucrs.br/eventos/encontroeconomia/?p=aprovados>>. Acesso em: 03 abr. 2020.