

# PESOS LONGITUDINAIS PARA A PESQUISA NACIONAL POR AMOSTRA DE DOMICÍLIOS CONTÍNUA (PNAD CONTÍNUA)

Antônio Etevaldo Teixeira Júnior<sup>1</sup>  
Eduardo Santiago Rosseti<sup>2</sup>  
Paula Alves de Almeida<sup>3</sup>  
Denise Britz do Nascimento Silva<sup>4</sup>

## 1 INTRODUÇÃO

Estudos longitudinais visam à observação de características de indivíduos ou empresas ao longo do tempo e sua realização depende da existência de informações para as mesmas unidades em repetidas ocasiões. No âmbito das estatísticas públicas brasileiras, análises longitudinais podem ser realizadas a partir de registros administrativos ou de pesquisas amostrais cuja coleta de dados repetida no tempo possibilite a identificação e o acompanhamento das unidades. Este estudo tem como foco as pesquisas amostrais, cuja representatividade é assegurada pela incorporação de pesos amostrais nos procedimentos de expansão dos resultados.

Em estudos longitudinais, apenas as unidades amostrais com resposta em todas as ocasiões no escopo da análise podem ser utilizadas nos procedimentos de estimação. Tal situação faz com que os pesos amostrais usualmente publicados (pesos transversais) deixem de cumprir seu objetivo de produzir estimativas válidas para a população-alvo. Além disso, a ausência de resposta em uma ou mais ocasiões (não repostas longitudinais) pode produzir vieses nas estimativas, interferindo também na sua precisão. Um desafio importante em estudos longitudinais é manter a representatividade da amostra no decorrer do estudo. Sendo assim, institutos produtores de estatísticas oficiais de Estados Unidos, Canadá, Alemanha, Austrália, entre outros, implementam procedimentos para ajustar os pesos transversais de suas pesquisas domiciliares longitudinais,<sup>5</sup> de forma a torná-los válidos para estudos longitudinais, produzindo então pesos longitudinais (Teixeira Júnior, 2015).

---

1. Doutorando em população, território e estatísticas públicas na Escola Nacional de Ciências Estatísticas (Ence).

2. Analista da Fundação Getúlio Vargas (FGV).

3. Doutora em população, território e estatísticas públicas pela Ence.

4. Professora e pesquisadora da Ence.

5. Pesquisas que realizam o acompanhamento de todas as pessoas que residem no domicílio quando este foi selecionado para a amostra.

A elaboração de pesos longitudinais é também importante para pesquisas domiciliares que possuem a estrutura longitudinal no nível do domicílio por meio da sobreposição parcial de sua amostra, como é o caso de duas pesquisas realizadas pelo Instituto Brasileiro de Geografia e Estatística (IBGE): a Pesquisa Mensal de Emprego (PME) e a Pesquisa Nacional por Amostra de Domicílios Contínua (PNAD Contínua). Uma metodologia para a produção de pesos longitudinais para a PME foi proposta por Teixeira Júnior (2015) e Teixeira Júnior, Silva e Veiga (2016a; 2016b). No caso da PNAD Contínua, pesos longitudinais foram desenvolvidos por Teixeira Júnior e Rosseti (2017). Um exemplo de utilização desses pesos pode ser encontrado em Teixeira Júnior, Alves e Rosseti (2018), para análise da relação da situação ocupacional dos jovens com a quantidade de horas despendidas em afazeres domésticos.

## 2 A ESTRUTURA LONGITUDINAL DA PNAD CONTÍNUA

A PNAD Contínua, um dos mais importantes levantamentos de informações conjunturais realizados no país e principal fonte de estatísticas públicas sobre o mercado de trabalho, é uma pesquisa cujo planejamento amostral possui um esquema de rotação 1-2(5), com sobreposição parcial da amostra ao longo dos trimestres. O domicílio é entrevistado um mês, retirado da pesquisa por dois meses consecutivos, e reinserido no trimestre seguinte, sendo esta sequência repetida cinco vezes, quando então o domicílio é retirado definitivamente da amostra (IBGE, 2014). Sendo assim, os microdados trimestrais da pesquisa permitem a realização de estudos longitudinais em dois trimestres consecutivos, com 80% de sua amostra; e em trimestres idênticos de anos consecutivos, com 20% de sua amostra.

Além de disponibilizar microdados trimestrais da pesquisa, o IBGE divulga também microdados anuais referentes à primeira e à última participação do domicílio na pesquisa. Bases de dados como essas podem ser obtidas por meio da seleção, em cada trimestre do ano, dos domicílios que estão na entrevista de interesse, totalizando cerca de 450 mil registros, constituindo importante fonte de informação para análises longitudinais sobre as condições socioeconômicas no país.

Este estudo apresenta a seguir a metodologia para construção de pesos longitudinais para uma base de dados longitudinal composta pela junção da base de dados dos domicílios que tiveram sua primeira participação na PNAD Contínua em 2017 com aqueles que tiveram sua quinta e última participação na pesquisa em 2018.

## 3 PAREAMENTO

Como já mencionado, a PNAD Contínua foi planejada para acompanhar domicílios da amostra durante cinco trimestres consecutivos. No caso de haver mudança de moradores, a entrevista é realizada com os novos moradores do domicílio. Logo, a PNAD Contínua não foi planejada como uma pesquisa longitudinal de indivíduos. Adicionalmente, o IBGE não disponibiliza, até o momento, junto com os microdados da pesquisa, informações que possibilitem a identificação dos indivíduos em diferentes ocasiões. Entretanto, é possível parear os moradores dos domicílios da amostra em ocasiões distintas da pesquisa a partir de algumas informações sociodemográficas, de forma semelhante à adotada por outros autores com dados da PME

(Lopes, 2002; Ribas, 2008; Rojas, 2014). O procedimento de pareamento de indivíduos na PNAD Contínua considerou primeiro o pareamento dos domicílios e, em cada um, utilizou quatro critérios exigindo igualdade nas variáveis/características dos moradores da seguinte forma:

- número de ordem, posição no domicílio, sexo e data de nascimento;
- número de ordem, sexo e data de nascimento;
- posição no domicílio, sexo e data de nascimento; e
- sexo e data de nascimento.

A taxa de pareamento obtida, considerando a amostra de todo o país, foi de 67,3%; resultado muito satisfatório, tendo em vista a ausência de uma chave de identificação e, principalmente, a distância no tempo entre as duas ocasiões de interesse (quatro trimestres). Para efeito de comparação, as taxas de pareamento anuais da PME variavam em torno de 50% (Teixeira Júnior, 2015). A tabela 1 apresenta estatísticas da distribuição das taxas de pareamento segundo unidade da federação (UF) e pós-estratos.<sup>6</sup>

TABELA 1  
**Taxas de pareamento segundo UF e pós-estratos – Brasil (2017-2018)**  
(Em %)

| Localidade   | Mínimo | 1º quartil | Mediana | 3º quartil | Máximo | Média |
|--------------|--------|------------|---------|------------|--------|-------|
| UF           | 38,7   | 61,1       | 68,4    | 71,7       | 81,3   | 66,1  |
| Pós-estratos | 12,0   | 60,9       | 67,9    | 72,6       | 86,5   | 64,3  |

Fonte: PNAD Contínua Anual (primeira entrevista em 2017 e segunda entrevista em 2018) – IBGE.

Os resultados podem ser considerados satisfatórios. Entretanto, existem algumas localidades com taxas de pareamento consideravelmente baixas. Dentre elas destaca-se o pós-estrato da cidade de Maceió, onde a taxa de pareamento foi de apenas 12%. Esse resultado decorre da ausência da informação relativa à data de nascimento em grande parte dos registros, informação crucial para o procedimento de pareamento.

#### 4 TRATAMENTO DO PESO TRANSVERSAL

Um peso longitudinal é construído a partir de um peso transversal inicial, comumente denominado “peso-base”. O IBGE divulga dois pesos amostrais na PNAD Contínua, o peso proveniente do desenho amostral<sup>7</sup> (V1031) com correção para a não resposta dentro do setor censitário, e um peso amostral ajustado por pós-estratificação (V1032) para que as estimativas de totais populacionais coincidam com as estimativas populacionais divulgadas pelo IBGE.

Entretanto, em nenhum desses pesos há tratamento para a não resposta diferenciada segundo a estrutura por idade e sexo da população. É comum que em pesquisas domiciliares a estrutura por idade e o sexo dos entrevistados seja diferente da encontrada na população, devido ao viés de disponibilidade (Vasconcellos, Silva e Szwarcwald, 2005). Sendo assim, neste trabalho, para a obtenção dos pesos-base, utilizou-se a técnica de calibração (Deville, Särndal e Sautory, 1993) que, por meio de multiplicadores conhecidos como

6. Domínios para os quais o IBGE ajusta os pesos amostrais da PNAD Contínua.

7. Para mais informações, ver: <<https://bit.ly/2kX20NU>>.

fatores de calibração, iguais as estimativas de totais populacionais da pesquisa a totais conhecidos de variáveis de interesse (no caso a distribuição conjunta de idade e sexo). Para o sucesso do procedimento, é importante que esses totais sejam amplamente aceitos como corretos e provenientes de fontes conhecidas e confiáveis (Brick e Montaquila, 2009). Dessa forma é possível ajustar a estrutura por idade e sexo dos respondentes da pesquisa à correspondente estrutura definida pelas projeções populacionais do IBGE, fonte esta que atende à restrição de qualidade sugerida por Brick e Montaquila (2009).

O método de calibração escolhido para esse ajuste denomina-se *generalized raking* (Deville, Särndal e Sautory, 1993), no qual são utilizadas as frequências marginais das variáveis auxiliares. Nesse estudo, a primeira frequência marginal foi composta pelos totais populacionais de cada UF, segundo sexo e grupos de idade; e a segunda pela população dos pós-estratos. O procedimento de calibração foi realizado considerando o planejamento amostral da PNAD Contínua, para tal foi utilizado o pacote *survey* do *software R* (Lumley, 2013).

Para avaliar o procedimento de calibração adotou-se a metodologia proposta por Silva (2004), que sugere critérios relativos aos fatores de calibração e uma aferição da eficiência média de um conjunto de indicadores de interesse da pesquisa. Todos os resultados foram considerados adequados, com destaque para os fatores de calibração que, devido às altas taxas de resposta da PNAD Contínua (em torno de 90%), foram muito baixos. A média dos fatores de calibração foi 1,17, enquanto que os valores mínimo e máximo foram, respectivamente, 0,61 e 1,93, dentro dos limites estipulados por Deville, Särndal e Sautory (1993) e Silva (2004). O coeficiente de variação dos fatores de calibração foi de 12,4%, abaixo de 15%, valor definido pelo IBGE para indicar boa precisão de estimativas amostrais.

A eficiência média foi obtida calculando-se a média dos coeficientes de variação de doze indicadores selecionados (tabela 2). Resultados semelhantes foram obtidos ao calcular as estimativas em nível Brasil e para as UFs, com diferentes pesos amostrais, sendo que o peso ajustado por idade e sexo apresentou melhor desempenho quanto à eficiência média. O ajuste pela distribuição de idade e sexo teve efeito diverso nos indicadores. Os indicadores demográficos apresentaram redução em seus níveis e os indicadores relacionados ao mercado de trabalho apresentaram elevação, em ambos os casos, devido à distribuição por idade e sexo dos respondentes da pesquisa ser mais envelhecida que a distribuição das projeções populacionais. Por sua vez, os indicadores socioeconômicos apresentaram estabilidade.

TABELA 2

**Estimativas e coeficientes de variação dos indicadores selecionados, por peso amostral utilizado – Brasil (2017-2018)**

| Indicador  | Peso IBGE sem calibração |        | Peso IBGE com calibração |        | Peso com calibração para a estrutura idade x sexo |        |
|--|--------------------------|--------|--------------------------|--------|---|--------|
|  | Estimativa               | cv (%) | Estimativa               | cv (%) | Estimativa  | cv (%) |
| <b>Eficiência média (nível Brasil)</b>                 | <b>0,94</b>              |        | <b>0,91</b>              |        | <b>0,88</b>                                       |        |
| Proporção de pessoas na PIA*                           | 81,3                     | 0,12   | 81,2                     | 0,12   | 79,9  | 0,03   |
| Proporção de pessoas vivendo em domicílios unipessoais | 5,1                      | 1,07   | 5,1                      | 1,05   | 4,8   | 1,00   |
| Proporção de pessoas brancas                           | 44,1                     | 0,55   | 43,7                     | 0,50   | 43,6  | 0,50   |
| Proporção de outros parentes e não parentes            | 10,5                     | 0,84   | 10,5                     | 0,82   | 10,7  | 0,82   |

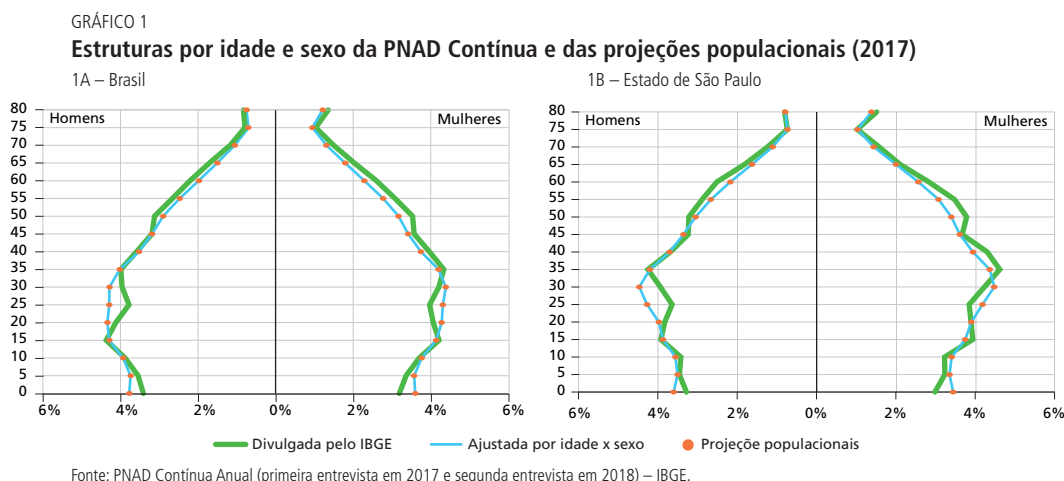
(Continua)

(Continuação)

| Indicador   | Peso IBGE sem calibração |        | Peso IBGE com calibração |        | Peso com calibração para a estrutura idade x sexo |        |
|---|--------------------------|--------|--------------------------|--------|---|--------|
|   | Estimativa               | cv (%) | Estimativa               | cv (%) | Estimativa  | cv (%) |
| Rendimento médio domiciliar <i>per capita</i> (R\$)     | 1.282                    | 1,33   | 1.269                    | 1,27   | 1.254   | 1,26   |
| Taxa de atividade                                       | 62,3                     | 0,23   | 62,2                     | 0,22   | 63,3  | 0,18   |
| Proporção de analfabetos                                | 5,8                      | 1,18   | 5,9                      | 1,14   | 5,5   | 1,13   |
| Proporção de pessoas com ensino superior                | 16,1                     | 1,58   | 15,9                     | 1,54   | 16,1  | 1,55   |
| Taxa de desocupação                                     | 12,5                     | 1,07   | 12,5                     | 1,03   | 12,7  | 0,99   |
| Rendimento médio no trabalho principal (R\$)            | 2.125                    | 1,43   | 2.110                    | 1,36   | 2.098   | 1,33   |
| Proporção de ocupados que contribuem para a previdência | 63,9                     | 0,34   | 63,6                     | 0,33   | 63,8  | 0,32   |
| Proporção de subocupados                                | 7,0                      | 1,52   | 7,0                      | 1,48   | 7,1   | 1,49   |

Fonte: PNAD Contínua Anual (primeira entrevista em 2017 e segunda entrevista em 2018) – IBGE.  
Obs.: PIA – população em idade ativa.

Para finalizar a descrição do tratamento da não resposta transversal são apresentadas as pirâmides etárias do Brasil e do estado de São Paulo. Verifica-se que a estrutura de idade e sexo obtida na PNAD Contínua com o peso divulgado pelo IBGE difere, em ambos os casos, da estrutura divulgada nas projeções populacionais (IBGE, 2018). Ao utilizar o peso ajustado, a estrutura de idade e sexo obtida com os dados da PNAD Contínua coincide com a estrutura das projeções populacionais corrigindo, principalmente, a representação na população do grupo de pessoas com idade entre 20 e 34 anos.



## 5 PROPENSÕES DE RESPOSTA

Os modelos de propensão de resposta foram incorporados ao tratamento do problema da não resposta em pesquisas amostrais por David *et al.* (1983) e constituem uma extensão direta da teoria do escore de propensão de Rosenbaum e Rubin (1983). A utilização desses modelos com respostas binárias para estimar as propensões de resposta pode ser mais efetiva que outros métodos (GRAU, 2006). Dentre os modelos para variáveis com respostas binárias, a regressão logística é a técnica mais utilizada para prever as probabilidades ou propensões de resposta (Kalton, Lepkowski e Lin, 1987; Gouskova *et al.*, 2008).

Consideremos que a variável dependente na regressão logística seja  $R_j$  que assume os valores:  $R_j = 1$ , se a unidade amostral  $j$  respondeu e  $R_j = 0$ , caso contrário. Define-se então um modelo para a propensão de resposta  $\theta_{j,c|p}$  para cada unidade  $j$  na ocasião de pesquisa  $c$  da forma:

$$\ln\left(\frac{\hat{\theta}_{j,c|p}}{1 - \hat{\theta}_{j,c|p}}\right) = \mathbf{x}_{pj}^T \hat{\boldsymbol{\beta}}_p,$$

onde  $\theta_{j,c|p}$  é a probabilidade (propensão) de resposta da unidade amostral  $j$  na ocasião  $c$  dado que a unidade respondeu na ocasião passada  $p$ ,  $\hat{\theta}_{j,c|p}$  é a estimativa da propensão obtida a partir dos valores previstos do modelo,  $\mathbf{x}_{pj}^T$  é um vetor de variáveis explicativas para cada  $j$  de uma ocasião passada  $p$ ,  $\hat{\boldsymbol{\beta}}_p$  é um vetor de coeficientes estimado e  $c$  é a ocasião para a qual a propensão está sendo estimada. As propensões de resposta estimadas para cada unidade respondente são então dadas por:

$$\hat{\theta}_{j,c|p} = \{1 + \exp(-\mathbf{x}_{pj}^T \hat{\boldsymbol{\beta}}_p)\}^{-1}$$

Lepkowski (1987) cita três tipos de estratégias para o cálculo dos fatores de ajuste. Como neste trabalho o objetivo é produzir pesos longitudinais para apenas pares de ocasiões, os fatores de ajuste são obtidos pelo inverso das propensões de resposta estimadas nos modelos de regressão logística, tal que:  $f_j = 1/\hat{\theta}_{j,c|p}$ . Obtidos os fatores de ajuste  $f_j$ , os pesos longitudinais  $w_j^{long}$  são obtidos diretamente pela expressão:  $w_j^{long} = w_j^{base} * f_j$ , onde  $w_j^{base}$  é o peso-base da unidade  $j$ .

Nos procedimentos de modelagem para estimação das propensões de resposta foram utilizadas doze variáveis, todas consideradas significativas conforme resultados apresentados na tabela 3. Assim como no procedimento de calibração, os modelos foram ajustados considerando o planejamento amostral da PNAD Contínua.

TABELA 3

**Estimativas dos coeficientes, estatísticas de teste e razões de chances, segundo variáveis explicativas – Brasil (2017-2018)**

| Variáveis explicativas        | Categoria de referência | Coefficiente | Erro padrão | Razão de chances | p-valor |
|-------------------------------|-------------------------|--------------|-------------|------------------|---------|
| Intercepto                    |                         | 0,405        | 0,036       | 1,50             | < 0,001 |
| Região metropolitana          | Não metropolitana       | -0,346       | 0,021       | 0,71             | < 0,001 |
| Ensino médio completo ou mais | Até médio incompleto    | 0,135        | 0,014       | 1,14             | < 0,001 |
| Filho(a)/enteado(a)           | Condição no domicílio:  | -0,157       | 0,015       | 0,85             | < 0,001 |
| Outro parente/não parente     | responsável ou cônjuge  | -0,673       | 0,022       | 0,51             | < 0,001 |
| De 30 a 59 anos               | Até 29 anos             | 0,371        | 0,015       | 1,45             | < 0,001 |
| 60 anos ou mais               |                         | 0,636        | 0,022       | 1,89             | < 0,001 |
| Norte/Nordeste/Centro-oeste   | Sul/Sudeste             | -0,267       | 0,019       | 0,77             | < 0,001 |
| Desocupado                    |                         | 0,053        | 0,021       | 1,05             | 0,0119  |
| Inativo                       | Ocupado                 | 0,150        | 0,013       | 1,16             | < 0,001 |
| Fora da PIA                   |                         | 0,435        | 0,017       | 1,54             | < 0,001 |
| Não frequenta escola          | Frequenta escola        | -0,258       | 0,014       | 0,77             | < 0,001 |
| Mulheres                      | Homens                  | 0,030        | 0,008       | 1,03             | 0,0038  |
| Rural                         | Urbano                  | 0,260        | 0,023       | 1,30             | < 0,001 |
| Nuclear                       |                         | 0,189        | 0,050       | 1,21             | < 0,001 |
| Estendida                     | Espécie de domicílio:   | 0,055        | 0,053       | 1,06             | 0,3078  |
| Composta                      | Unipessoal              | -0,472       | 0,075       | 0,62             | < 0,001 |

(Continua)

(Continuação)

| Variáveis explicativas                | Categoria de referência                        | Coefficiente | Erro padrão | Razão de chances | p-valor |
|---------------------------------------|--|--------------|-------------|------------------|---------|
| 2 moradores c/14 anos ou mais         | 1 morador com 14 anos ou mais                  | 0,155        | 0,046       | 1,17             | < 0,001 |
| Mais de 2 moradores c/14 anos ou mais |  | 0,301        | 0,047       | 1,35             | < 0,001 |
| Mais de 0,5 até 1 s.m.                | Rendimento domiciliar per capita: Até 0,5 s.m. | 0,101        | 0,021       | 1,11             | < 0,001 |
| Mais de 1 até 2 s.m.                  |  | 0,244        | 0,023       | 1,28             | < 0,001 |
| Mais de 2 s.m.                        |  | 0,380        | 0,030       | 1,46             | < 0,001 |

Fonte: PNAD Contínua Anual (primeira entrevista em 2017 e segunda entrevista em 2018) – IBGE.

## 6 AJUSTE FINAL DOS PESOS

Para finalizar o procedimento de produção de pesos longitudinais para a PNAD Contínua, uma etapa adicional de calibração foi necessária, desta vez realizada por meio de pós-estratificação (Deville, Särndal e Sautory, 1993). O objetivo dessa calibração final foi manter os totais populacionais dos pós-estratos, domínios de divulgação da PNAD Contínua, iguais aos divulgados pela pesquisa.

A tabela 4 apresenta os resultados finais dos principais indicadores do processo de ajuste dos pesos. Verifica-se que a propensão de resposta mínima ficou abaixo de 20%, entretanto o primeiro quartil da distribuição ficou acima de 62%, indicando que poucos casos tiveram propensão de resposta muito baixa. Os fatores de ajuste finais dos pesos longitudinais encontram-se dentro dos limites desejáveis e o peso longitudinal máximo, apesar de alto, não possui valor muito acima do peso máximo transversal da pesquisa.

TABELA 4

### Estadísticas das distribuições das propensões de resposta, dos fatores de ajuste e dos pesos longitudinais

| Variáveis                    | Mínimo | 1º quartil | Mediana | Média | 3º quartil | Máximo |
|------------------------------|--------|------------|---------|-------|------------|--------|
| Propensão de resposta        | 0,194  | 0,628      | 0,698   | 0,683 | 0,753      | 0,911  |
| Fator de ajuste longitudinal | 0,695  | 0,896      | 0,947   | 0,937 | 1,015      | 2,843  |
| Peso longitudinal            | 4      | 185        | 312     | 450   | 549        | 10.438 |

Fonte: PNAD Contínua Anual (primeira entrevista em 2017 e segunda entrevista em 2018)/IBGE.

Nota: † Considera as propensões de resposta e o pós-estratificação.

O principal objetivo na produção de pesos longitudinais é a eliminação do viés de não resposta longitudinal. Para avaliar se os pesos longitudinais permitem atingir esse objetivo, foram calculadas estimativas para o conjunto de indicadores selecionados, utilizando as bases de dados longitudinais (contendo os respondentes de ambas as ocasiões no escopo de análise) e as bases de dados transversais (contendo todos os respondentes da primeira ocasião no escopo de análise). Na base de dados longitudinal, foram aplicados os pesos longitudinais e os pesos transversais. A ideia da utilização dos pesos transversais na base longitudinal é ilustrar o comportamento das estimativas que seriam calculadas caso os pesos longitudinais não fossem produzidos. A suposição é de que quanto mais próximas as estimativas calculadas com o peso longitudinal na base longitudinal estiverem das estimativas calculadas com o peso-base na base transversal, menor será o viés de não resposta longitudinal.

A tabela 5 contém as estimativas calculadas na base de dados transversal com pesos-base – ajustados pela estrutura de idade e sexo das projeções populacionais, isto é, com tratamento para o viés de disponibilidade –, assim como as estimativas calculadas na base de dados longitudinal com os pesos publicados pelo IBGE e com os pesos longitudinais.

TABELA 5

**Estimativas e coeficientes de variação dos indicadores selecionados, por base de dados e peso amostral utilizado, segundo indicadores selecionados – Brasil (2017-2018)**

| Indicador   | Base transversal               |        | Base longitudinal |        |                   |        |
|---|--------------------------------|--------|-------------------|--------|-------------------|--------|
|   | Peso ajustado por idade e sexo |        | Peso IBGE         |        | Peso longitudinal |        |
|   | Estimativa                     | cv (%) | Estimativa        | cv (%) | Estimativa        | cv (%) |
| Proporção de pessoas na PIA                             | 79,9                           | 0,03   | 81,6              | 0,14   | 79,9              | 0,15   |
| Proporção de pessoas vivendo em domicílios unipessoais  | 4,8                            | 1,00   | 5,0               | 1,22   | 4,8               | 1,24   |
| Proporção de pessoas brancas                            | 43,6                           | 0,50   | 45,7              | 0,61   | 45,1              | 0,62   |
| Proporção de outros parentes e não parentes             | 10,7                           | 0,82   | 8,2               | 1,10   | 10,5              | 1,12   |
| Rendimento médio domiciliar <i>per capita</i> (R\$)     | 1.254                          | 1,26   | 1.343             | 1,43   | 1.269             | 1,39   |
| Taxa de atividade                                       | 63,3                           | 0,18   | 61,3              | 0,27   | 63,3              | 0,26   |
| Proporção de analfabetos                                | 5,5                            | 1,13   | 5,6               | 1,37   | 5,1               | 1,40   |
| Proporção de pessoas com ensino superior                | 16,1                           | 1,55   | 16,7              | 1,66   | 16,3              | 1,66   |
| Taxa de desocupação                                     | 12,7                           | 0,99   | 11,8              | 1,29   | 12,6              | 1,29   |
| Rendimento médio no trabalho principal (R\$)            | 2.098                          | 1,33   | 2.213             | 1,59   | 2.106             | 1,51   |
| Proporção de ocupados que contribuem para a previdência | 63,8                           | 0,32   | 64,5              | 0,38   | 63,9              | 0,39   |
| Proporção de subocupados                                | 7,1                            | 1,49   | 7,2               | 1,65   | 7,5               | 1,67   |

Fonte: PNAD Contínua Anual (primeira entrevista em 2017 e segunda entrevista em 2018)/IBGE.

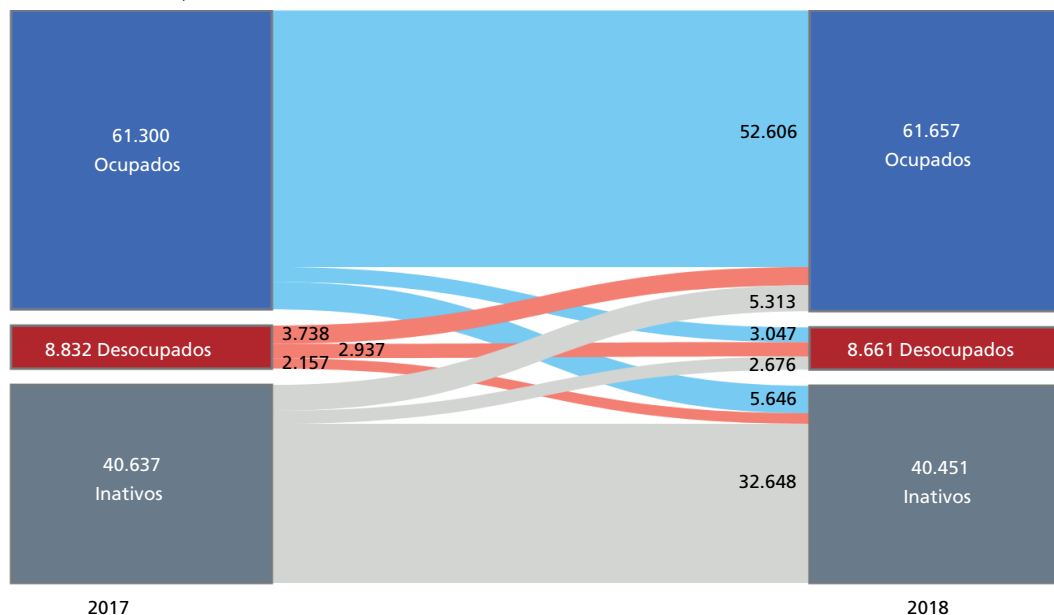
Os resultados apresentam padrões esperados, primeiro devido ao fato dos coeficientes de variação das estimativas na base longitudinal serem mais elevados do que os obtidos na base transversal, em razão da amostra disponível ser menor. Adicionalmente, e com maior relevância, pelo fato da distância encontrada entre as estimativas calculadas na base transversal e na base longitudinal, obtidas com os pesos longitudinais, ser inferior às diferenças encontradas ao comparar as estimativas calculadas na base transversal e as obtidas com os pesos divulgados pelo IBGE na base longitudinal.

## 7 FLUXOS NO MERCADO DE TRABALHO

Todos os meses, milhares de pessoas mudam de condição de ocupação ou de atividade, ou seja, tornam-se ativos ou inativos, ocupados ou desocupados. Estudos de fluxo da força de trabalho possibilitam, por exemplo, perceber quando uma queda no desemprego de um mês a outro se deve a um incremento no número de desocupados que se tornam ocupados, ou a um aumento no número de desocupados que passam à situação de inativos (U.S. Census Bureau, 2015). A figura 1 apresenta o fluxo das condições de atividade no mercado de trabalho brasileiro entre 2017 e 2018.



FIGURA 1  
**Fluxo das condições de atividade na população em idade ativa – Brasil (2017-2018)**  
 (Em milhares de pessoas)



Fonte: PNAD Contínua Anual (primeira entrevista em 2017 e segunda entrevista em 2018) – IBGE.  
 Obs.: Figura cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

O estudo dos fluxos possibilita compreender diferenças no comportamento da dinâmica do mercado de trabalho em relação a mulheres e homens, ou entre pessoas de diferentes pertencas raciais, de diferentes grupos de estudo ou etário, diferenças regionais, por exemplo. Segundo Ribeiro (2010), o estudo do fluxo de trabalhadores entre estados no mercado de trabalho permite a avaliação da capacidade de realocação de recursos de uma economia. As taxas de rotatividade e realocação do mercado de trabalho brasileiro são muito altas, o que pode gerar “grande insegurança para os trabalhadores, pela redução do tempo de permanência em um emprego ou perda de bem-estar por custos de ajustamento e de oportunidade, além de perdas salariais dos trabalhadores na troca do emprego” (Ribeiro, 2010, p. 2). Esses estudos auxiliam na compreensão da dinâmica do mercado de trabalho e, especialmente, do desemprego, contribuindo para o planejamento de políticas públicas eficientes de redução da rotatividade.

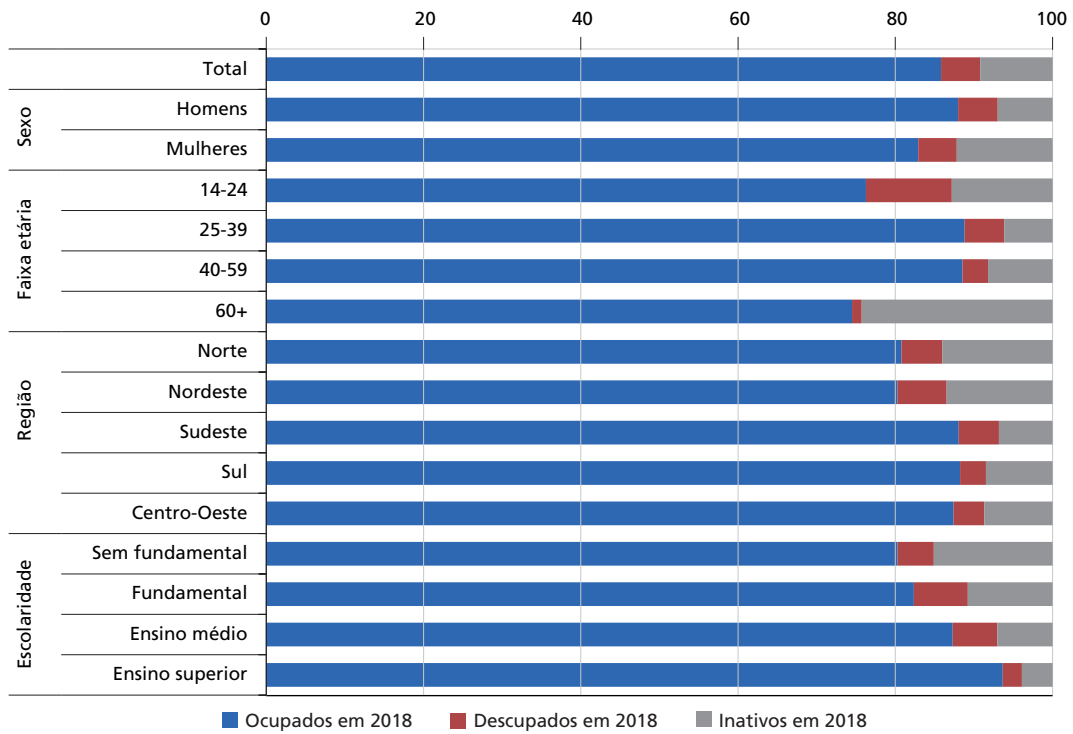
No Brasil, 20,4% da PIA mudou a sua condição de atividade entre 2017 e 2018. No gráfico 2 são apresentadas as movimentações de ocupados e desocupados na PIA para diferentes grupos populacionais. Evidencia-se maior movimentação entre os desocupados, que se distribuem em proporções próximas nas três condições. Considerando todos os grupos analisados, pelo menos 80% permaneceram como ocupados no ano seguinte, com exceção de dois grupos: pessoas com 60 anos ou mais e pessoas entre 14 e 24 anos de idade. No primeiro grupo, a maioria dos que mudaram de condição migraram para a inatividade, enquanto no segundo grupo migraram de maneira praticamente igual para a desocupação e para a inatividade. Por outro lado, as pessoas com ensino superior completo apresentaram maior estabilidade, com 95% das pessoas se mantendo como ocupadas.

GRÁFICO 2

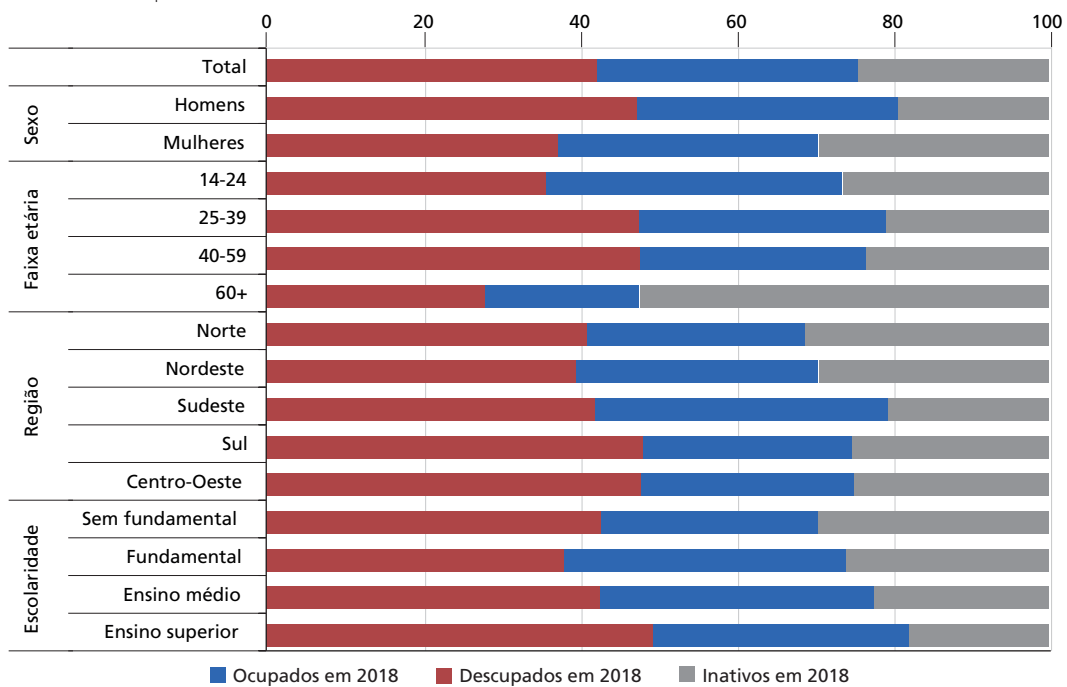
**Mudança na condição ocupacional de ocupados e desocupados na PIA – Brasil (2017-2018)**

(Em %)

2A – Ocupados



2B – Desocupados



Fonte: PNAD Contínua Anual (primeira entrevista em 2017 e segunda entrevista em 2018)/IBGE.

Considerando as pessoas que estavam desocupadas em 2017, os grupos compostos por homens, pessoas com idade entre 25 e 59 anos, residentes nas regiões Sul e Centro-Oeste e de pessoas com nível superior são os que se destacam no que se refere à persistência na condição de desocupados. Por outro lado, pessoas com 60 anos ou mais, mulheres, residentes nas regiões Norte e Nordeste, sem nível fundamental completo, movimentaram-se para fora da população econômica ativa, sinalizando um possível desalento na procura por trabalho. Uma análise mais detalhada do fenômeno requer, entretanto, a consideração de vários fatores conjuntamente. Para isso, se a estratégia de análise incorporar modelos estatísticos, recomenda-se o uso de pesos longitudinais como propostos neste artigo.

## REFERÊNCIAS

- BRICK, J. M.; MONTAQUILA, J. M. Nonresponse and weighting. *In*: PFEFFERMANN, D. C.; RAO, C. R. (Orgs.). **Sample surveys: theory, methods and inference**. Amsterdam: Elsevier B.V., 2009. p. 163-185. (Handbook of Statistics, v. 29a).
- DAVID, M. *et al.* **Nonrandom nonresponse models based on the propensity to respond**. Alexandria: American Statistical Association, 1983. p. 168-173. (Survey Research Methods Section).
- DEVILLE, J. C.; SÄRNDAL, C. E.; SAUTORY, O. Generalized raking procedures in survey sampling. **Journal of the American Statistical Association**, Alexandria, v. 88, n. 423, p. 1013-1020, 1993.
- GRAU, E. *et al.* **Nonresponse adjustment using logistic regression: to weight or not to weight?** Alexandria: American Statistical Association, 2006. p. 3073-3080. (Survey Research Methods Section).
- IBGE – INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **Pesquisa Nacional por Amostra de Domicílios Contínua: notas metodológicas**. Rio de Janeiro: IBGE, 2014. v. 1.
- \_\_\_\_\_. **Projeções de população**. 2. ed. Rio de Janeiro: IBGE, 2018. (Série Relatórios Metodológicos, n. 40).
- KALTON, G.; LEPKOWSKI, J. M.; LIN, T. K. Compensating for wave nonresponse in the 1979 ISDP research panel. *In*: KALTON, G. *et al.* **The treatment of person-wave nonresponse in longitudinal surveys**. Washington: U.S. Bureau of the Census, 1987. p. 4-21.
- LEPKOWSKI, J. M. The Treatment of Wave Nonresponse in Panel Surveys. *In*: KALTON, G. *et al.* **The treatment of person-wave nonresponse in longitudinal surveys**. Washington: U.S. Bureau of the Census, 1987.
- LOPES, M. D. **Avaliação de desgaste de painéis em estudos longitudinais: uma aplicação na pesquisa mensal de emprego (PME/IBGE)**. 2002. Dissertação (Mestrado em Estudos Populacionais e Pesquisas Sociais) – Escola Nacional de Ciências Estatísticas, Rio de Janeiro, 2002.
- LUMLEY, T. **Survey: analysis of complex survey samples**. CRAN, 2013. Disponível em: <<https://bit.ly/2m2fhFd>>. Acesso em: 1º jul. 2015.
- RIBAS, R. P.; SOARES, S. S. D. **O atrito nas pesquisas longitudinais: o caso da pesquisa mensal de emprego (PME) do IBGE**. Rio de Janeiro: Ipea, 2008. (Texto para Discussão, n. 1347).
- RIBEIRO, E. P. Fluxo de empregos, fluxo de trabalhadores e fluxo de postos de trabalho no Brasil. **Revista de Economia Política**, v. 30, n. 3, p. 401-419, jul.-set./ 2010.

ROJAS, H. A. G. **Modelos para estimar cambios brutos en encuestas rotativas con ausencia de respuesta en diseños de muestreo complejos**. 2014. Tese (Doutorado) – Universidad Nacional de Colombia, Bogotá, 2014.

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

SILVA, P. L. N. **Calibration estimation: when and why, how much and how**. Rio de Janeiro: IBGE, 2004. (Texto para Discussão da Diretoria de Pesquisas, n. 14).

TEIXEIRA JÚNIOR, A. E. **Produção de pesos longitudinais para estimação e análise de dados da pesquisa mensal de emprego do IBGE**. 2015. Dissertação (Mestrado em Estudos Populacionais e Pesquisas Sociais) – Escola Nacional de Ciências Estatísticas, Rio de Janeiro, 2015.

TEIXEIRA JÚNIOR, A. E.; SILVA D. B. N.; VEIGA A. C. Produção de pesos longitudinais para estimação e análise de dados da pesquisa mensal de emprego do IBGE. *In*: SIMPÓSIO NACIONAL DE PROBABILIDADE E ESTATÍSTICA, 22., 2016, Porto Alegre, Rio Grande do Sul. **Anais...** Porto Alegre: SINAPE, 2016a.

\_\_\_\_\_. A importância do uso de pesos longitudinais na produção de indicadores para análise da evolução do mercado de trabalho. *In*: CONGRESO DE LA ASOCIACIÓN LATINOAMERICANA DE POBLACIÓN, 7.; ENCONTRO NACIONAL DE ESTUDOS POPULACIONAIS, 20., 2016, Foz do Iguaçu, Paraná. **Anais...** Foz do Iguaçu: ABEP; ALAP, 2016b.

TEIXEIRA JÚNIOR, A. E.; ROSSETI E. S. Metodologia para a construção de pesos longitudinais para a Pesquisa Nacional por Amostragem de Domicílios Contínua (PNAD Contínua). *In*: ESCOLA DE AMOSTRAGEM E METODOLOGIA DE PESQUISA, 5., Cuiabá, Mato Grosso T. **Anais...** Cuiabá: ESAMP, 2017.

TEIXEIRA JÚNIOR, A. E.; ALVES P. A.; ROSSETI E. S. Análise longitudinal do tipo de atividade de jovens e a permanência ou não na categoria ‘nem-nem’ entre 2015 e 2016: diferenciais de gênero e responsabilidades domésticas. *In*: ENCONTRO NACIONAL DE ESTUDOS POPULACIONAIS, 21., Poços de Caldas, Minas Gerais. **Anais...** Poços de Caldas: ABEP, 2018.

U.S. CENSUS BUREAU. **Research series on labor force status flows from the Current Population Survey**. Washington: Division of Labor Force Statistics, 2015.

VASCONCELLOS, M. T. L.; SILVA, P. L. N.; SZWARCOWALD, C. L. Aspectos de amostragem da pesquisa mundial de saúde no Brasil. **Cadernos de Saúde Pública**, Rio de Janeiro, n. 21, p. 589-599, 2005.