

COMPARANDO OS MICRODADOS DA PNAD CONTÍNUA 2018 E DA PESQUISA DE ORÇAMENTOS FAMILIARES DE 2018

COMPARING MICRODATA FROM THE 2018 CONTINUOUS PNAD AND DE 2018 FAMILY BUDGET SURVEY

Pedro Nessi Snizek Júnior¹
pedro.nessi@univag.edu.br

Carlos Magno da Silva²

Herbert Tadashi Mitsuyuki³

Alessandra Paz⁴

Eduardo Nicolini Dorneles-Bolsista de IC⁵

Resumo

O presente estudo visa comparar os resultados obtidos da pesquisa PNAD contínua 2018 e da POF 2018 através dos microdados. É possível considerar, em especial as variáveis relacionadas aos moradores, que para a grande maioria das variáveis estudadas, que as Pesquisas PNAD Contínua 2018 e POF 2018 produzem amostras com valores bastante aproximados e consistentes. Destacando-se as seguintes situações: métodos de ponderação amostral diferentes no estudo, períodos não coincidentes de coleta e tamanhos amostrais diferenciados. O estudo indica que as amostras de pesquisas demográficas e consequentes extrapolações para as populações devem ser avaliadas com cuidado, sempre considerando se as variáveis escolhidas fazem parte das que são utilizadas para cálculo e dimensionamento amostral.

Palavras-chave: Pof; Pnad Contínua; Microdados.

Abstract

The present study aims to compare the results obtained from the 2018 continuous PNAD survey and the 2018 POF through microdata. It is possible to consider, especially the variables related to residents, that for the vast majority of the variables studied, the Continuous PNAD 2018 and POF 2018 Surveys produce samples with very approximate and consistent values. The following situations stand out: different sample weighting methods in the study, non-coincident collection periods and different sample sizes. The study indicates that demographic research samples and consequent extrapolations to populations must be evaluated carefully, always considering whether the variables chosen are part of those used for sample calculation and sizing.

Keywords: pof, pnad, microdata

Pedro Nessi Snizek Jr.
Doutor em Produção Animal e Pós Doutor em Administração
Tecnologista do IBGE e Professor da Univag

Carlos Magno da Silva
Mestre em Economia e Professor da Univag

Herbert Tadashi Mitsuyuki
Mestre em Engenharia Ambiental e Professor da Univag

Alessandra Pereira Paz de Santana
Mestre em Estudos de Cultura Contemporânea

Eduardo Nicolini Dorneles
Discente do curso de Comércio Exterior.

1. INTRODUÇÃO

Pesquisas domiciliares com representatividade nacional são mundialmente utilizadas como a principal fonte de informações para estudos sobre demografia. O Brasil não constitui uma exceção, e a maior parte dos estudos existentes na literatura nacional apoiam-se na Pesquisa Nacional por Amostra de Domicílios Contínua (PNAD Contínua), coletada anualmente pelo Instituto Brasileiro de Geografia e Estatística (IBGE, 2003).

Pesquisas amostrais domiciliares com cobertura nacional representam a melhor fonte de informação para medir os dados demográficos de um país. No caso do Brasil, a PNAD Contínua constitui a principal fonte de informação sobre demografia, sendo internacionalmente reconhecida como uma fonte de dados de excelente qualidade (Barros et al., 2007; Barros et al., 2008).

Existe uma profusão de bancos de dados de alta qualidade publicamente disponíveis. Essa abundância de dados é extremamente positiva, pois permite que os pesquisadores avaliem a robustez de suas descobertas. Contudo, ela também desperta dúvidas, já que nem sempre os números obtidos são compatíveis entre si (Souza, 2015).

Nenhum dos artigos encontrados na literatura, que utilizaram microdados da PNAD Contínua (Pesquisa Nacional por Amostra de Domicílio) e da POF (Pesquisa de Orçamentos Familiares), realizaram comparações entre estas duas pesquisas (Hecksher, 2020), (Hoffmann, 2019), (Trovão, 2002), (Assunção, 2024), (Gonçalves, 2023), (Braga, 2017), (Deus et al., 2023).

Pela revisão bibliográfica realizada, pode-se considerar que o presente artigo é o primeiro que faz uma comparação quantitativa de microdados de pesquisas demográficas PNAD contínua e POF do IBGE.

A relevância do presente estudo está em uma análise inovadora sobre os microdados divulgados pelo IBGE no seguinte aspecto:

1. Relacionar e comparar microdados oriundos de duas pesquisas demográficas realizadas pelo IBGE. Especificamente as PNAD Contínua e a POF, ambas com referência ao ano de 2018..

Este trabalho procura entender as discrepâncias existentes nas pesquisas domiciliares PNAD Contínua 2018 e POF 2018, regularmente conduzidas pelo IBGE que têm cobertura nacional e coletam informações tanto dos moradores como dos domicílios de residência.

Este artigo está estruturado em 5 sessões incluindo: introdução, seguido por caracterização e metodologias adotadas pela PNAD Contínua 2018 e POF 2018, método de avaliação/pesquisa, resultados e conclusões.

2. AMOSTRA MESTRE PARA AS PESQUISAS DOMICILIARES DO IBGE

O IBGE faz uso de uma amostra mestra (SIPD) para todas as suas pesquisas amostrais domiciliares. A SIPD corresponde a um conjunto de unidades de área selecionadas de um cadastro, segundo um método probabilístico de seleção a partir da qual seja possível selecionar subamostras para atender às diversas pesquisas. (Freitas et. al. 2007; IBGE, 2007b; IBGE, 2008a; Freitas et al., 2007).

Esta amostra mestra corresponde a um conjunto de unidades de área selecionadas de um cadastro, segundo um método probabilístico de seleção, a partir da qual seja possível selecionar subamostras para atender as pesquisas amostrais domiciliares do IBGE (IBGE, 2007b).

Dentre as principais pesquisas do SIPD, já definidas, estão: a pesquisa nacional por amostra de domicílios contínua (PNADC) e a pesquisa de orçamentos familiares (POF) (IBGE, 2007b).

A população alvo da Amostra Mestra inclui toda a população a ser investigada em todas as pesquisas. É constituída pelos moradores residentes em todos os domicílios na área que constitui a abrangência geográfica definida a seguir. A abrangência geográfica da Amostra Mestra considera o âmbito das diversas pesquisas que farão uso dessa amostra comum. Assim, não se pode deixar de incluir qualquer parte do território que seja contemplado por alguma das pesquisas. Portanto, a abrangência geográfica da Amostra Mestra é constituída pelos setores censitários de todo Brasil da base operacional geográfica do último censo demográfico realizado pelo IBGE (IBGE, 2007b).

O plano amostral escolhido foi um desenho estratificado, com seleção das unidades primárias com probabilidade proporcional a uma medida de tamanho. A opção por este plano se justifica pelo fato das metodologias de todas as pesquisas domiciliares do IBGE incluírem planos amostrais que empregam algum tipo de estratificação das UPAs e seleção destas unidades com probabilidade proporcional ao tamanho, conforme descrito em Bianchini e Albieri (2003).

O cadastro para seleção da Amostra Mestra é um cadastro denominado cadastro mestre, que corresponde a um conjunto de unidades de áreas com seus limites bem definidos, que cobrem todo o país. Para cada uma dessas unidades são associadas informações sobre divisões administrativas, contagens populacionais e outras características sociodemográficas.

Como unidades básicas do cadastro foram escolhidos os setores censitários, pois são as menores partições do território para as quais são conhecidas as informações necessárias para sua construção.

O sistema de referência (cadastro) utilizado foi um arquivo contendo informações provenientes do último Censo Demográfico realizado durante o período de realização das pesquisas que fazem parte do sistema SIPD (IBGE, 2007b).

2.1 CARACTERIZAÇÃO E METODOLOGIAS ADOTADAS PELA PNAD CONTÍNUA 2018

O plano amostral adotado para a PNAD Contínua é um plano conglomerado em dois estágios de seleção com estratificação das UPAs (unidades primárias de amostragem). No primeiro estágio são selecionadas UPAs com probabilidade proporcional ao número de domicílios dentro de cada estrato definido. Como a pesquisa é integrante do SIPD, utilizará a infraestrutura amostral comum a todas as pesquisas, deste modo a estratificação das UPAs é adotada para a amostra mestre (IBGE, 2014b).

Como componente deste sistema, a PNAD Contínua integrou a Pesquisa Nacional por Amostra de Domicílios (PNAD) e a Pesquisa Mensal de Emprego (PME), permitindo a produção de indicadores a curto prazo sobre trabalho e rendimento com abrangência nacional e detalhamento por unidade da federação (IBGE, 2008).

A PNAD Contínua é a pesquisa do sistema de pesquisas domiciliares do IBGE que demanda maior amostra, com isso a definição do tamanho da amostra de UPAs da pesquisa corresponde a definição do tamanho da amostra mestre para um trimestre.

No segundo estágio é selecionado um número fixo de domicílios particulares permanentes ocupados dentro de uma UPA da amostra, por amostragem aleatória simples do Cadastro Nacional de Endereços para Fins Estatísticos (CNEFE) atualizado (IBGE, 2014b).

O tamanho da amostra de domicílios dentro de cada UPA foi fixado em 14 amostras de UPAs e de domicílios e dividida pelos 3 meses de um trimestre, seguindo o esquema de rotação 1-2(v), onde v é o número de visitas a serem realizadas em um mesmo domicílio. Neste esquema o domicílio é entrevistado 1 mês e sai da amostra por 2 meses seguidos, sendo este esquema repetido “v” vezes (IBGE, 2014b).

O tamanho da amostra foi determinado considerando a precisão desejada para a estimativa de um parâmetro de interesse. O parâmetro escolhido foi o: “total de pessoas desocupadas no trimestre”, que é um dos principais indicadores da PNAD Contínua (IBGE, 2007b).

A PNAD contínua tem como objetivo produzir indicadores para acompanhar as flutuações de curto prazo e a evolução a médio e longo prazo, da força de trabalho e outras informações necessárias para o estudo e desenvolvimento socioeconômico do País.

A PNAD Contínua tem uma abrangência de coleta de informação nacional em 3500 municípios do país. A abrangência de divulgação que considera: Brasil, Grandes Regiões, Unidades da Federação, Capitais e suas regiões metropolitanas. A periodicidade é mensal, trimestral e anual:

por meio de trimestres móveis, para conjunto restrito de indicadores de trabalho para Brasil (IBGE, 2015).

A cada trimestre, a PNAD Contínua investiga em torno de 211.000 domicílios em aproximadamente 16.000 setores censitários. Cada domicílio selecionado para a pesquisa é visitado cinco vezes, durante cinco trimestres consecutivos. Assim, um domicílio é visitado pela segunda vez três meses após a primeira visita, pela terceira vez três meses após a segunda visita, e assim por diante. Isso quer dizer que a pesquisa segue um esquema de rotação intitulado 1-2(5) onde, de um trimestre para o próximo, há uma sobreposição de 80% dos domicílios e de um trimestre para o mesmo trimestre do ano seguinte, de 20% (IBGE, 2014).

Na PNAD contínua, as informações sobre o tema trabalho são captadas em dois questionários: um reduzido, restrito as informações sobre trabalho que gera rendimentos para o domicílio e outro mais abrangente aplicado na primeira entrevista, concomitante com as características básicas dos moradores e do domicílio (IBGE, 2014). O presente estudo fez uso deste questionário mais abrangente realizado na primeira entrevista.

Os demais temas permanentes da pesquisa como: migração, fecundidade e trabalho infantil poderão ser investigados em um trimestre específico ou, então, aplicados em uma parte da amostra a cada trimestre e acumulados para produzir resultados anuais. Temas suplementares de periodicidade variável, também poderão ser investigados em um trimestre ou acumulados ao longo do ano. (IBGE, 2014).

A abrangência geográfica da PNAD Contínua é todo o território nacional, dividido nos setores censitários da base operacional geográfica do último censo realizado pelo IBGE, excluídas áreas com características especiais, classificadas pelo IBGE como setores de: aldeias indígenas, quartéis, bases militares, alojamentos, acampamentos, embarcações, barcos, navios, penitenciárias, colônias penais, presídios, cadeias, asilos, orfanatos, conventos, hospitais e agrovilas de projetos de assentamentos rurais, e também os setores censitários localizados em Terras Indígenas. E a população alvo é constituída por todas as pessoas moradoras em domicílios particulares permanentes da área de abrangência da pesquisa (IBGE, 2014)

Deve-se ter sempre em mente quando da análise e interpretação dos dados que a unidade de investigação da PNAD Contínua é o domicílio, e a unidade de análise as pessoas residentes nestes domicílios.

O período de coleta das informações da PNAD Contínua 2018 foi entre janeiro e dezembro de 2018.

2.2 CARACTERIZAÇÃO E METODOLOGIAS ADOTADAS PELA PESQUISA DE ORÇAMENTOS FAMILIARES 2018

A POF mensura as estruturas de consumo, dos gastos, dos rendimentos e parte da variação patrimonial das famílias. Possibilita traçar, portanto, um perfil das condições de vida da população brasileira a partir da análise de seus orçamentos domésticos. A alimentação é outro tema de destaque e muito explorado que conta com abordagens não monetárias (IBGE, 2021a).

Para a pesquisa POF 2017 especificamente em termos práticos, num primeiro estágio, foram selecionados de forma aleatória um quantitativo de setores pertencentes à Amostra Mestra, respeitando as estratificações construídas para cada Unidade da Federação. Em um segundo estágio de seleção, também de forma aleatória e para cada setor selecionado no primeiro estágio, foram selecionados em torno de 13 domicílios, no caso das áreas urbanas, e 17 domicílios nas áreas rurais. A soma das amostras das Unidades da Federação corresponde a aproximadamente 75.000 domicílios (IBGE, 2017a).

O período de coleta da POF 2017 foi de 12 meses, entre junho de 2017 e julho de 2018. Desta forma, além das aquisições realizadas continuamente pelos respondentes, durante todo o ano, mede-se também a sazonalidade. (IBGE, 2017a; IBGE, 2017b).

A pesquisa POF 2017 apresentou 7 questionários chamados de POF1 (Características do domicílio e moradores); POF2/POF3 (Caderneta de Aquisição Coletiva - Aquisição de produtos e serviços), POF4 (Questionário de Aquisição Individual); POF5 (Questionário de Trabalho e Rendimento Individual), POF6 e o questionário POF7 (Questionário de Consumo Alimentar Pessoal (IBGE, 2008, 2017a, 2008b)

O preenchimento da Caderneta era feito pelo(a) morador(a) que está mais acostumado a realizar estes tipos de aquisições. (IBGE, 2017a).

O consumo diário de alimentos de cada diário (g) para a pesquisa POF foram obtidos por meio do questionário POF7 ou Bloco de Consumo Alimentar Pessoal.

A POF 7 foi preenchida para uma subamostra de domicílios, selecionados aleatoriamente da amostra da POF, onde é aplicado para todos os indivíduos moradores com 10 anos ou mais de idade, pertencentes à(s) Unidades(s) de Consumo existente(s) no domicílio.

O Bloco de Consumo Alimentar Pessoal (POF 7) é preenchido pelo Agente de Pesquisa mediante entrevista pessoal para preencher dois dias de recordatório com a intenção de pedir ao entrevistado que relate todos os alimentos e bebidas consumidos no dia anterior à entrevista da hora em que acordou até a hora em que foi dormir.

A pesquisa POF 2018 investigou 57.920 domicílios em 5.504 setores censitários (IBGE, 2017a).

A seguir é apresentado o método adotado para realização do objetivo da presente pesquisa.

3. MÉTODO

Foi realizada pesquisa exploratória descritiva com base em dados secundários oriundos dos microdados anuais das pesquisas PNAD Contínua e POF; ambos com referência ao ano de 2018.

Para tratamento, leitura, estruturação e realização das análises estatísticas dos microdados da PNAD Contínua 2018 foi utilizado o pacote PNADcIBGE do software R.

Para tratamento, leitura, estruturação e realização das análises estatísticas dos microdados da POF 2018 foram utilizados os pacotes dplyr e Tidyverse. Considerando a ponderação amostral através da variável PESO_FINAL.

O IBGE já disponibiliza os seguintes pacotes para análises de microdados das pesquisas domiciliares: PNADcIBGE, PNSBIBGE, PNSIBGE e SIPDIBGE no software R. Infelizmente não está disponível o pacote POFIBGE.

As análises estatísticas foram divididas em duas etapas, na primeira foram feitas análises comparativas entre as pesquisas PNAD Contínua 2018 e POF 2018 através de estimativas de quantitativo dos moradores: cor ou raça e gênero e curso mais elevado que frequentou. Na sequência foi calculado a idade média dos moradores para as pesquisas PNAD Contínua 2018 e POF 2018.

Na segunda etapa foram feitas análises comparativas através de estimativas do quantitativo das características dos domicílios de residência: Material predominante nas paredes, material predominante na cobertura, material predominante no piso, principal fonte de abastecimento de água e principal destino do lixo.

Foram realizadas comparações percentuais para mensurar as diferenças entre as pesquisas PNAD Contínua 2018 e POF 2018.

As comparações foram todas realizadas a nível geográfico: Brasil.

As análises estatísticas foram realizadas via software R.

3.1 FONTES DE DADOS E VARIÁVEIS

O estudo fez uso de dados secundários oriundos dos microdados anuais das Pesquisas PNAD Contínua e POF. Ambas com ano de referência de 2018.

Para a Pesquisa PNAD Contínua 2018 os dados foram acessados de forma online via pacote PNADcIBGE instalado no software R. Já para POF 2018 foram usados os pacotes dplyr e Tidyverse, e ponderação amostral através da variável PESO_FINAL.

As variáveis relacionadas as características dos moradores da Pesquisa PNAD Contínua 2018 consideradas foram: Sexo (V2007), Cor ou Raça (V2010) e Qual o curso mais elevado que frequentou anteriormente (V3009A).

As variáveis relacionadas as características dos moradores da Pesquisa POF 2018 consideradas foram: Sexo (V0404), Cor ou Raça (V0405) e qual o curso mais elevado que frequentou anteriormente (V0425A).

As variáveis relacionadas as características dos domicílios considerados na Pesquisa PNAD Contínua foram: Material que predomina na construção das paredes externas (S01002), Material que predomina no telhado (S01003), Material predominante no piso (S01004), qual a principal forma de abastecimento de água utilizada no domicílio (S01007) e qual o principal destino dado ao lixo (S01013).

As variáveis relacionadas as características dos domicílios de moradia consideradas na POF 2018 foram: Material que predomina na construção das paredes externas (V0202), Material que predomina na cobertura (telhado) (V0203), Material predominante no piso (V0204), qual a principal forma de abastecimento de água utilizada no domicílio (V0207) e qual o principal destino dado ao lixo (V02013).

Os códigos das variáveis, indicados entre parênteses, podem ser encontrados no dicionário de microdados fornecido pelo IBGE. Tanto para a pesquisa PNAD Contínua 2018 como para a POF 2018.

As características gerais dos moradores e dos domicílios da Pesquisa PNAD Contínua 2018 foram obtidas nos microdados anuais referentes a primeira entrevista da PNAD contínua do ano de 2018 (IBGE, 2024).

As características gerais dos moradores e dos domicílios da Pesquisa POF 2018 foram obtidas nos microdados da POF do ano de 2018. Os microdados são divididos em cadernos, sendo usados: Morador e Domicílio (IBGE, 2024).

Considerando as características das pesquisas: PNAD Contínua 2018 e POF 2018 de não utilizar setores localizados em terras indígenas, deve-se ter cuidado com relação as análises feitas para a variável cor ou raça (V2010). Todos os resultados obtidos para as pessoas que se consideraram de cor ou raça indígena são para não residentes em terras indígenas.

3.2 PROCESSAMENTO DE DADOS

Para tratamento, leitura, estruturação e realização das análises estatísticas da Pesquisa PNAD Contínua 2018 foi utilizado o pacote PNADcIBGE instalado no software R.

Devido a estrutura amostral da pesquisa PNAD Contínua, foi necessário a utilização do pacote Survey do software R para ponderação amostral dos dados indicada pelo IBGE.

Foram utilizadas as seguintes funções do pacote Survey para análise de dados: Estimativa de totais(svytotal), Estimativa de média (svymean).

Para tratamento leitura, estruturação e realização das análises estatísticas da Pesquisa POF 2018 foram usados os pacotes dplyr e Tidyverse, e ponderação amostral através da variável PESO_FINAL.

Foram utilizadas as seguintes funções dos pacotes dplyr e Tidyverse para análise de dados: Estimativa de totais (count), Estimativa de média (weight.mean).

As variáveis relacionadas aos domicílios na pesquisa PNAD Contínua 2018 tem como referência os moradores e na POF 2018 o domicílio. Foi necessário então dividir os dados das variáveis da pesquisa PNAD Contínua por 3 para ajusta e comparação. O número 3 refere-se ao número médio de moradores por domicílio indicado pela pesquisa PNAD Contínua de 2018.

3.3 MODELO ECONOMETRICO

As análises estatísticas foram divididas em duas etapas, na primeira foram feitas análises comparativas entre as pesquisas PNAD Contínua 2018 e POF 2018, através de estimativas de quantitativo dos moradores: cor ou raça, gênero e curso mais elevado que frequentou. Na sequência foi calculado a idade média dos moradores para as pesquisas PNAD Contínua 2018 e POF 2018.

Na segunda etapa foram feitas análises comparativas entre as pesquisas PNAD Contínua 2018 e POF 2018, através de estimativas comparativas do quantitativo em relação as características dos domicílios de residência: Material predominante nas paredes, material predominante na cobertura, material predominante no piso, principal fonte de abastecimento de água e principal destino do lixo.

As comparações foram todas realizadas a nível geográfico: Brasil

As análises estatísticas foram realizadas via software R.

4. RESULTADOS

4.1 VARIÁVEIS RELACIONADAS AOS MORADORES

Os dados do quadro 1 mostram uma diferença percentual entre os dados da Pesquisa PNAD Contínua 2018 e POF 2018 de apenas 1,44% para homens e -0,65% para mulheres. Destacamos que para todas

as análises, diferença percentual positiva significa que a amostra da PNAD Contínua 2018 apresenta valor numérico maior que a da pesquisa POF 2018. Diferença percentual negativa o contrário.

Quadro 1 – Estimativa de Homens e Mulheres da população Brasileira no ano de 2018, através dos dados das pesquisas Pnad Contínua 2018 e Pesquisa de Orçamentos Familiares 2018.

Gênero	Pnad Contínua 2018	Pesquisa de Orçamentos Familiares 2018	Diferença (%)
Homens	101.657.376	100.210.730	1,44
Mulheres	106.195.917	106.893.060	-0,65

Deve-se destacar que o único artigo encontrado que compara resultados obtidos entre as pesquisas amostrais demográficas PNAD Contínua e POF do IBGE foi realizado por Souza (2015). Entretanto o assunto enfocado foi distribuição de renda e não fica claro na parte do estudo referente a metodologia se as comparações foram feitas com microdados.

Quadro 2 – Estimativa da cor ou raça da população Brasileira no ano de 2018, através dos dados das pesquisas Pnad Contínua 2018 e Pesquisa de Orçamentos Familiares 2018.

Cor ou Raça	Pnad Contínua 2018	Pesquisa de Orçamentos Familiares 2018	Diferença (%)
Branca	88.912.167	91.127.030	-2,43
Preta	19.335.180	21.167.709	-8,66
Parda	97.394.953	92.003.791	5,86
Amarela	1.363.635	1.412.777	-3,47
Indígena	814.280	777.539	4,72

Primeiramente deve ser destacado que o IBGE pesquisa cor ou raça da população brasileira com base na autodeclaração. Ou seja, quando questionada, a pessoa pode se declarar como: preta, parda, branca, amarela ou indígena. Segundo IBGE (2022) a cor ou raça dos moradores deve ser determinada pelo próprio morador, sem qualquer influência do entrevistador.

Como pode ser observado quadro 2, as diferenças encontradas em termos percentuais entre as duas pesquisas para a variável: cor ou raça ficaram entre -2,43% e -8,66%. O percentual mais elevado, refere-se a cor ou raça preta que alcançou uma diferença numérica de 1.832.529 em favor da pesquisa POF 2018.

Importante salientar que segundo IBGE (2022) a variável curso mais elevado que frequentou foi perguntada para as pessoas que já frequentaram a escola.

Como pode ser observado no quadro 3 as diferenças encontradas em termos percentuais entre as duas pesquisas para a variável: curso mais elevado que frequentou ficaram entre 1,63% e 14,07%. O percentual mais elevado, refere-se ao nível de mestrado que alcançou a diferença numérica de

105.437 em favor da pesquisa PNAD Contínua 2018. O nível doutorado também apresentou uma diferença elevada em termos percentuais de -6,87%. Com diferença numérica de 24.123

Quadro 3 – Estimativa do curso mais elevado que frequentou da população Brasileira no ano de 2018, através dos dados das pesquisas Pnad Contínua 2018 e Pesquisa de Orçamentos Familiares 2018.

Curso Mais elevado que frequentou	Pnad Contínua 2018	Pesquisa de Orçamentos Familiares 2018	Diferença (%)
Ensino Fundamental	36.623.103	36.033.756	1,63
Ensino Médio	49.389.598	46.644.085	5,89
Superior	19.248.309	18.638.364	5,88
Especialização	4.691.364	4.558.564	2,91
Mestrado	854.755	749.318	14,07
Doutorado	327.004	351.127	-6,87

Deve ser salientado que para esta variável, as maiores diferenças percentuais foram observadas nos itens com menor valor numérico.

Analisando os dados do quadro 4 pode-se observar diferença de -3,88% entre as idades médias calculadas para as pesquisas PNAD Contínua 2018 e POF 2018. Com diferença numérica de 1,37 anos. Importante também considerar que a amostra da PNAD Contínua 2018 investigou cerca de 210.000 domicílios por trimestre e a da POF 2018 investigou 57.920 domicílios ao longo de 12 meses. Essa diferença de tamanho de amostras, em favor da PNAD Contínua 2018, não deve ser desconsiderada. Mesmo assim a diferença não superou 3,88%.

Quadro 4 – Estimativa da idade média em anos da população Brasileira no ano de 2018, através dos dados das pesquisas Pnad Contínua 2018 e Pesquisa de Orçamentos Familiares 2018.

	Pnad Contínua 2018	Pesquisa de Orçamentos Familiares 2018	Diferença (%)
Idade Média (anos)	33,89	35,26	-3,88

4.2 VARIÁVEIS RELACIONADAS AOS DOMICÍLIOS

Segundo o IBGE (2023) considera-se domicílio como o local estruturalmente separado e independente que se destina a servir de habitação para uma ou mais pessoas, ou que esteja sendo utilizado como tal.

Quadro 5 – Estimativa do quantitativo do material que predomina nas paredes dos domicílios brasileiros no ano de 2018, através dos dados das pesquisas Pnad Contínua 2018 e Pesquisa de Orçamentos Familiares 2018.

Material Predominante nas Paredes	PNAD Contínua 2018	Pesquisa de Orçamentos Familiares 2018	Diferença (%)
Alvenaria com Revestimento	60.315.371	60.281.129	0,057
Alvenaria sem Revestimento	4.894.999	4.750.992	3,03

Taipa sem revestimento	385.821	284.552	35,59
Madeira Adequada para construção	3.277.691	3.253.010	0,76

Analisando o quadro 5 pode-se observar que as diferenças encontradas em termos percentuais entre as duas pesquisas para a variável: material que predomina nas paredes ficaram entre 0,057% e 35,59%. O percentual mais elevado, refere-se a taipa sem revestimento que alcançou a diferença numérica de 101.269 em favor da pesquisa PNAD Contínua 2018. Sendo esta a maior diferença percentual encontrada no estudo.

Deve ser salientado que para esta variável, as maiores diferenças percentuais também foram observadas nos itens com menor valor numérico.

Analisando o quadro 6 pode-se observar que as diferenças encontradas em termos percentuais entre as duas pesquisas para a variável: material que predomina no telhado ficaram entre -0,76% e 9,93%. O percentual mais elevado, refere-se a: Zinco, alumínio ou chapa metálica que alcançou a diferença numérica de 86.682 em favor da pesquisa PNAD Contínua 2018.

Deve ser salientado que para esta variável, a maior diferença percentual também foi observada no item com menor valor numérico.

O item somente laje de concreto também apresentou uma diferença que deve ser destacada de -7,77%. Com valor numérico de 778.046. Esta situação foi diferente da maioria já encontrada, onde as maiores diferenças percentuais ocorreram em itens com menores valores numéricos.

Analisando o quadro 7, pode-se observar que as diferenças encontradas em termos percentuais entre as duas pesquisas para a variável: material que predomina no piso ficaram entre 1,94% e 24,51%. O percentual mais elevado, refere-se a: Terra que alcançou a diferença numérica de 89.300 em favor da pesquisa PNAD Contínua 2018.

Quadro 6 – Estimativa do quantitativo do material que predomina no telhado dos domicílios brasileiros no ano de 2018, através dos dados das pesquisas PNAD Contínua 2018 e Pesquisa de Orçamentos Familiares 2018.

Material Predominante no telhado	PNAD Contínua 2018	Pesquisa de Orçamentos Familiares 2018	Diferença (%)
Telha sem laje de concreto	35.968.684	36.242.967	-0,76
Telha com laje de concreto	22.211.572	20.988.207	5,82
Somente laje de concreto	9.239.195	10.017.241	-7,77
Madeira Adequada para construção	373.212	363.271	2,73

Zinco, alumínio ou chapa metálica	960.018	873.336	9,93
-----------------------------------	---------	---------	------

Deve ser salientado que para esta variável, a maior diferença percentual também foi observada no item com menor valor numérico.

O item madeira apropriada para construção apresentou também uma diferença que deve ser destacada de -21,08%. Com valor numérico de 1.165.345. Esta situação diferente da maioria já encontrada, ocorrendo em item com elevados valores numéricos.

Quadro 7 – Estimativa do quantitativo do material que predomina no piso dos domicílios brasileiros no ano de 2018, através dos dados das pesquisas PNAD Contínua 2018 e Pesquisa de Orçamentos Familiares 2018.

Material Predominante no Piso	PNAD Contínua 2018	Pesquisa de Orçamentos Familiares 2018	Diferença (%)
Cerâmica, Lajota ou Pedra	53.294.536	51.937.632	2,61
Madeira Apropriada para Construção	4.361.526	5.526.871	-21,08
Cimento	10.884.969	10.677.637	1,94
Terra	450.375	361.705	24,51

Analisando o quadro 8, pode-se observar que as diferenças encontradas em termos percentuais entre as duas pesquisas para a variável: principal fonte de abastecimento de água ficaram entre 0,6% e 18,61%. O percentual mais elevado, refere-se a: Fonte ou nascente alcançou a diferença numérica de 229.832 em favor da pesquisa PNAD Contínua 2018.

Para esta variável a maior diferença percentual não ocorreu em item com baixo valor numérico, como nas outras variáveis analisadas.

Analisando o quadro 9, pode-se observar que as diferenças encontradas em termos percentuais entre as duas pesquisas para a variável: principal destino do lixo ficaram entre -1,21% e 23,1%. O percentual mais elevado, refere-se ao item coleta em cacimbado serviço de limpeza que alcançou a diferença numérica de 1.068.868 em favor da pesquisa PNAD Contínua 2018.

Para esta variável a maior diferença percentual não ocorreu em item com baixo valor numérico, como na maioria das outras variáveis analisadas.

Tabela 8 – Estimativa do quantitativo da principal fonte de abastecimento de água dos domicílios brasileiros no ano de 2018, através dos dados das pesquisas PNAD Contínua 2018 e Pesquisa de Orçamentos Familiares 2018.

Principal Fonte de Abastecimento de Água	PNAD Contínua 2018	Pesquisa de Orçamentos Familiares 2018	Diferença (%)
Rede Geral de Distribuição	58.799.231	58.447.643	0,6
Poço profundo ou Artesiano	5.085.268	5.010.189	1,5

Poço Raso, freático ou cacimba	2.374.684	2.426.186	-2,12
Fonte ou nascente	1.464.478	1.234.646	18,61
Água da Chuva Armazenada	497.334	463.593	7,28

Tabela 9 – Estimativa do quantitativo do principal destino do lixo dos domicílios brasileiros no ano de 2018, através dos dados das pesquisas PNAD Contínua 2018 e Pesquisa de Orçamentos Familiares 2018.

Principal destino do lixo	PNAD Contínua 2018	Pesquisa de Orçamentos Familiares 2018	Diferença (%)
Coletado Diretamente por serviço de limpeza	56.871.134	57.567.409	-1,21
Coletado em cacimba do serviço de Limpeza	5.697.310	4.628.442	23,1
Queimado	5.734.914	5.586.374	2,66
Enterrado	258.224	281.616	-8,3
Jogado em Terreno Baldio ou Logradouro	509.742	577.791	-11,78

5. CONCLUSÕES

As maiores diferenças percentuais quando comparadas as duas pesquisas, ocorreram nas variáveis relacionadas aos domicílios.

As maiores diferenças percentuais ocorreram principalmente em itens das variáveis com menores valores numéricos. Excetuando-se os seguintes itens: Cor ou raça preta, somente laje de concreto, madeira apropriada para construção, fonte ou nascente e coleta em caçamba do serviço de limpeza.

A maior diferença percentual ocorreu no item: taipa sem revestimento (35,59%).

As maiores diferenças numéricas ocorreram nos seguintes itens: cor ou raça (1.832.529 pessoas), somente laje de concreto (778.046 domicílios), madeira apropriada para construção (1.165.345 domicílios) e fonte ou nascente (229.832 domicílios).

É possível considerar, em especial nas variáveis relacionadas aos moradores, que para a grande maioria das variáveis estudadas, as pesquisas PNAD contínua 2018 e POF 2018 produzem amostras com valores bastante aproximados e consistentes. Mesmo considerando as seguintes situações: métodos de ponderação amostral diferentes no estudo, períodos não coincidentes de coleta e tamanhos amostrais diferentes.

O estudo indica que se deve ser conservador quando da comparação de amostras de diferentes pesquisas demográficas, e com as consequentes extrapolações para as populações, sempre considerando se as variáveis escolhidas fazem parte ou não das utilizadas para cálculo e dimensionamento amostral. Mesmo que as pesquisas façam parte da mesma Amostra Mestra.

5.1 IMPLICAÇÕES PRÁTICAS DA PESQUISA

Em uma breve análise na base de dados SIDRA do IBGE é possível observar que o IBGE não divulga estimativas dos moradores e características dos domicílios através dos dados da POF 2018, já que a amostra, como já citado, foi dimensionada especificamente para apresentar maior precisão para outras variáveis.

O principal objetivo da pesquisa PNAD Contínua é estudar as variáveis relacionadas a mercado e trabalho e renda. Já a POF 2018 tem como principal finalidade identificar o padrão de gastos dos brasileiros e estabelecer a cestas de produtos que será considerada para o Índice Nacional de Preços ao Consumidor (INPC).

Tanto para a pesquisa PNAD contínua 2018 como para a POF 2018, as amostras são dimensionadas para captar com precisão as variáveis de maior interesse: na PNAD Contínua o número de pessoas ocupadas e taxa de desocupação (%). Já para a pesquisa POF 2018 a amostra é dimensionada para a cesta coletiva de produtos adquiridos. Este estudo comparou variáveis das duas pesquisas que não fazem parte das selecionadas para o dimensionamento das referidas amostras.

Devido a estes fatores, os dados comparados neste estudo devem ser analisados com muita atenção e entendimento do real objetivo das amostras construídas para as pesquisas PNAD Contínua 2018 e POF 2018.

5.2 LIMITAÇÕES DO ESTUDO E SUGESTÕES PARA PESQUISA FUTURA

O presente estudo analisa e compara amostras oriundas de duas pesquisas amostrais, com períodos de referências diferentes. A pesquisa PNAD Contínua 2018 considera dados no período de janeiro e dezembro de 2018. Já a POF considera os dados do período de junho de 2017 a julho de 2018.

As análises dos microdados via software R foram muito facilitadas pelo pacote PNADciIBGE. Como ainda não existe o pacote para a Pesquisa POF 2018, foi necessário o uso dos pacotes dplyr e Tidyverse considerando a variável PESO_FINAL.

O ideal para uma melhor comparação seria a realização da comparação com duas das pesquisas através do uso do pacote do IBGE para a pesquisa POF, que infelizmente ainda não existe.

Com isto existe a dificuldade de expandir as amostras da Pesquisa POF 2018 para a população, já que a função survey que foi utilizada na expansão da Pesquisa PNAD Contínua 2018 facilita muito este processo.

A pesquisa PNAD Contínua 2018 tem uma amostra de cerca de 210.000 domicílios por trimestre, onde as variáveis consideradas no estudo são captadas na primeira das cinco entrevistas realizadas. Já a POF 2018 teve uma amostra perto de 57.920 domicílios e as variáveis consideradas no estudo são captadas através do caderno POF1 ao longo de 12 meses. As diferenças no tamanho das amostras não devem ser desconsideradas como um dos possíveis efeitos sobre os resultados obtidos.

Na PNAD Contínua, as variáveis relacionadas aos domicílios, tem como referência os moradores e na POF 2018 o domicílio. Devido a isto, foi necessário dividir as variáveis da pesquisa PNAD Contínua por 3 para ajusta a comparação. O número 3 refere-se ao número médio de moradores por domicílio indicado pela pesquisa PNAD Contínua de 2018. Este ajuste também não deve ser desconsiderado como um dos possíveis efeitos sobre os resultados obtidos.

Não foi encontrado nenhum estudo na literatura com formato semelhante, para as variáveis analisadas no estudo, que pudesse ser utilizado como comparação e avaliação para o presente estudo.

REFERÊNCIAS

- ASSUNÇÃO, G. **Análise de Microdados da Pnad Contínua**. 2024 RPUBS. Disponível: link:<https://rpubs.com/gabriel-assuncao-ibge/pnadc> Acesso: 09/09/2024.
- BARROS, R.P.; CURY, S. ULYSSEA, G. **A desigualdade de renda no Brasil encontra-se subestimada? Uma análise comparativa com base na Pnad, na POF e nas contas nacionais**. Texto para discussão n° 1263. IPEA, Rio de Janeiro, 2007, 39p.
- BARROS, R.P.; CURY, S. ULYSSEA, G. **A desigualdade de renda no Brasil encontra-se subestimada? Uma análise comparativa usando a Pnad, POF e contas nacionais**. Texto para discussão n° 4. Rio de Janeiro, Departamento de Economia, UFF. 2008, 41P.
- BIANCHINI, Z.M; ALBERI, S. E. **Principais aspectos da amostragem da pesquisas domiciliares do IBGE – revisão 2002**. Rio de Janeiro: IBGE, Departamento de Metodologia. 2003. (Texto para discussão no 8).
- BRAGA, D. **Análise de microdados da Pnad Contínua. Com pacotes PNADcIBGE e survey**. RPUBS..Disponível link:<https://rpubs.com/BragaDouglas/335574>. Acesso em 09/09/2024
- DEUS, G.B.B.F; VIEIRA, H.C.; MELO, T.F.N. **Nota técnica em economia n. 13**. FACE/UFG, 2023. 13P.
- FREITAS, M. P. S.; LILA, M.F. AZEVEDO, R.V.; ANTONACI, G.A. **Amostra mestre para o sistema integrado de pesquisas domiciliares**. Rio de Janeiro: IBGE, Coordenação de Métodos e Qualidade. 2007. 67p.
- GONÇALVES, C. **Indicadores no R: Indicadores de emprego e renda na PnadC**. RPUBS. Disponível link:https://rpubs.com/caiocgonc/pnadc_trimestral. Acesso em 09/09/2024.
- HECKSHER, M. **Valor impreciso por mês exato: microdados mensais baseados na pnad contínua**. Nota Técnica n° 62. Diretoria de Estudos e Política Sociais. IPEA, 2020, 22p.
- HOFFMANN,R. **Distribuição da renda no Brasil em 2017. Uma apresentação didática das principais características da distribuição de renda no Brasil de acordo com os dados da Pnad Contínua de 2017**. Economia & Região, Londrina. v.7, n°2, p 5-28, jul/dez, 2019.
- INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA - IBGE. **Pesquisa de orçamentos familiares – 2002-2003**. Notas Técnicas, Rio de Janeiro 2003,
- INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA - IBGE, **Amostra mestre para o sistema integrado de pesquisas domiciliares. Textos para discussão n° 23**. Rio de Janeiro, 2007a. 67p.
- INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA - IBGE, **Sistema Integrado de Pesquisas Domiciliares - SIPD. Textos para discussão n° 24**. Rio de Janeiro, 2007b. 80p.
- INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA - IBGE, **Sistema integrado de Pesquisas Domiciliares Amostra mestra 2010 e Amostra da Pnad Contínua**. Textos para discussão n° 50. Rio de Janeiro, 2014b, 43p.
- INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA - IBGE, **Sistema Integrado de Pesquisas Domiciliares – SIPD**. XVI Encontro Nacional de Estudos Populacionais, Caxambu-MG. 2008, 20P.
- INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA - IBGE. Pesquisa de Orçamentos Familiares 2008-2009. **Manual do Agente de Pesquisa**. do Planejamento, Desenvolvimento e Gestão. Diretoria de Pesquisas. Coordenação de Orçamento e Gestão . Rio de Janeiro, 2008. 322p.
- INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA - IBGE, **Pesquisa Nacional por Amostra de Domicílios Contínua. Nota Metodológicas**. V.1. Rio de Janeiro, 2014. 47p.
- IBGE, **Sistema integrado de Pesquisas Domiciliares Amostra mestra 2010 e Amostra da Pnad Contínua**. Textos para discussão n° 50. Rio de Janeiro, 2014b, 43p.

- INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA - IBGE. **Principais diferenças metodológicas entre as pesquisas PME, Pnad e Pnad Contínua. Nota técnica.** Rio de Janeiro. 2015. 10p.
- INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA - IBGE. Pesquisa de Orçamentos Familiares 2017-18. **Manual do Agente de Pesquisa.** Ministério do Planejamento, Desenvolvimento e Gestão. Diretoria de Pesquisas. Coordenação de Orçamento e Gestão. Rio de Janeiro, 2017a. 317p.
- INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA - IBGE. Pesquisa de Orçamentos Familiares: 2017:2018: **perfil das despesas no Brasil: Indicadores de qualidade de vida.** Ministério da Economia Diretoria de Pesquisa. Coordenação de Trabalho e Rendimento. Rio de Janeiro, 2021a, 63p.
- INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA - IBGE. **Pesquisa de Orçamentos Familiares 2017-2018. Lançamento da coleta das informações.** Rio de Janeiro, 2017b. 40p.
- INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA - IBGE. **Pesquisa Nacional Por Amostra de Domicílios Contínua. Manual Básico da Entrevista.** Coordenação de Pesquisas Por Amostra de Domicílios. Rio de Janeiro, 2022. 262p.
- INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA - IBGE. **Pesquisa Nacional por Amostra de Domicílios Contínua Notas técnicas.** Versão 1.12. Rio de Janeiro, 2023. 126p.
- R Core Team (2024). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- SOUZA, P.H.G.F. A distribuição de renda nas populações domiciliares brasileiras: harmonização e comparação entre Censos, Pnads e Pofs. Revista Brasileira de Estudos Populacionais. Rio de Janeiro, v.32, n.1, p. 165-188, 2015.
- TROVÃO, C.J.B.M. **Por dentro da Pnad Contínua: uma introdução ao tratamento de dados usando R. Dados eletrônicos** (1 arquivo:78kb). Natal- RN:EDUFRN, 2022. 339p.