

Metodologia de Retropolação da Pesquisa Nacional por Amostra de Domicílios Contínua de 1992 a 2012

Bruno Ottoni Eloy Vaz¹

Tiago Cabral Barreira²

Área ANPEC 13 – Economia do Trabalho

Sumário: A Pesquisa Mensal de Emprego (PME), divulgada mensalmente pelo Instituto Brasileiro de Geografia e Estatística (IBGE) desde 1980, era reconhecida como a principal fonte de informações acerca do mercado de trabalho brasileiro, porém, foi descontinuada em 2016. Ela foi substituída pela Pesquisa Nacional por Amostra de Domicílios Contínua (PNADC), de abrangência nacional e também divulgada pelo referido instituto. A substituição da PME pela PNADC criou inúmeras dificuldades relacionadas a continuidade das séries. O estudo aqui apresentado visa contornar o problema da extensão temporal relativamente limitada das novas séries da PNADC, iniciada em março de 2012. Utilizando os dados da PME e de outra pesquisa de emprego divulgada pelo IBGE – a Pesquisa Nacional por Amostra de Domicílios (PNAD) – desenvolvemos um método de retropolação da PNADC, que permite produzir séries temporais mais longas. Enquanto os dados originais podem ser utilizados para retroagir somente até março de 2012, as novas séries permitem voltar até setembro de 1992. Com o novo método de retropolação, é possível não somente contornar o problema da incapacidade de gerar séries longas para a PNADC, como também fornecer um importante ferramental para compreendermos o comportamento histórico do emprego e da renda nacionalmente.

Palavras-chave: Retropolação, Desemprego, Força de Trabalho, População em Idade Ativa

Classificação JEL: C82, J11, J21

Abstract: The Monthly Employment Survey (PME), released monthly by the Brazilian Institute of Geography and Statistics (IBGE) since 1980, and recognized as the main source of information about the Brazilian labor market, was discontinued in 2016. It was replaced by the National Survey by Continuous Household Sample (PNADC), nationwide and also released by that institute. The replacement of PME by PNADC created numerous difficulties related to the continuity of the series. The study presented here aims to circumvent the problem of the relatively limited time span of the new PNADC series, started in March 2012. Using data from the PME and another job survey released by IBGE - the National Household Sample Survey (PNAD) - we developed a PNADC retropolation method, which allows us to produce longer time series. While the original data can be used only since March 2012, the new series are estimated back to September 1992. With the new retropolation method, it is possible not only to circumvent the problem of the inability to generate long series for the PNADC, but also also provide an important tool for understanding the historical behavior of employment and income nationally.

Keywords: Retropolation, Unemployment, Workforce, Working Age Population

JEL Classification: C82, J11, J21

¹ Pesquisador do IDados

² Pesquisador do IDados

1. Introdução

A Pesquisa Mensal de Emprego (PME), divulgada mensalmente pelo Instituto Brasileiro de Geografia e Estatística (IBGE), foi descontinuada no início de 2016. A Pesquisa Nacional por Amostra de Domicílios Contínua (PNADC) é, desde então, a única pesquisa mensal relacionada ao mercado de trabalho brasileiro a ser disponibilizada ao público. Esta nova pesquisa de emprego, apesar de possuir maior abrangência territorial (a PME selecionava somente seis regiões metropolitanas), é frequentemente criticada por fornecer séries temporais relativamente curtas. Não existem dados para a PNADC anteriores a março de 2012, como havia para a PME, cuja série histórica tem início em 1980.

Este estudo visa contornar o problema da extensão temporal relativamente limitada das novas séries geradas a partir da PNADC. Utilizando os dados da PME e de outra pesquisa de emprego divulgada pelo IBGE – a Pesquisa Nacional por Amostra de Domicílios (PNAD) – desenvolvemos um método de retroposição da PNADC, que permite produzir séries temporais mais longas³. Enquanto os dados originais podem ser utilizados para retroagir somente até março de 2012, as novas séries construídas aqui permitem voltar até setembro de 1992.

Podemos enumerar sete principais indicadores de mercado de trabalho medidos pela PNAD e PNADC, entre os quais: i) o indicador de ocupação⁴, representado pela População Ocupada (PO); ii) o indicador de desocupação⁵, representado pela População Desocupada (PD); iii) o indicador de força de trabalho⁶, representado pela População Economicamente Ativa (PEA); iv) o indicador demográfico de população em idade apta para trabalhar⁷, representado pela População em Idade Ativa (PIA); v) o indicador de inatividade⁸, representado pela População Não-Economicamente Ativa (PNEA); vi) o indicador de taxa de desemprego (TD); vii) o indicador de rendimento médio dos trabalhadores ocupados.

Um dos principais obstáculos encontrados aqui, para a construção de séries temporais longas, é a existência de inúmeras diferenças metodológicas entre a PNAD e a PNADC. Mais precisamente, estas pesquisas possuem diferenças importantes, por exemplo, quanto à definição de idade mínima de ocupação e com relação à classificação de inatividade. Também divergem quanto ao tamanho e a distribuição geográfica das amostras coletadas.

³ Vale ressaltar que a PNAD foi descontinuada em 2015

⁴ Segundo o IBGE, são classificadas como ocupadas as pessoas que, em um determinado período de referência, trabalharam em trabalho remunerado ou em trabalho sem remuneração direta em ajuda à atividade econômica de membro do domicílio, ou, ainda, as pessoas que tinham trabalho, mas estavam temporariamente afastadas dele.

⁵ Segundo o IBGE, são classificadas como desocupadas as pessoas não ocupadas em um determinado período de referência, que tomaram alguma providência efetiva para conseguirem um trabalho no mesmo período.

⁶ Segundo o IBGE, a força de trabalho abrange o total de pessoas ocupadas e desocupadas em um determinado período de referência.

⁷ A idade mínima de aptidão ao trabalho varia entre as pesquisas do IBGE, oscilando entre os 10 e 15 anos.

⁸ Segundo o IBGE, são classificadas como inativas as pessoas fora da força de trabalho que não estavam ocupadas nem desocupadas em um determinado período de referência.

Tomando-se como exemplo o ano de 2014, tais diferenças geram uma discrepância, entre as pesquisas, de aproximadamente 10 milhões na PEA e de tamanho semelhante na PO. As divergências metodológicas entre as duas pesquisas também geram discrepâncias importantes nos níveis da PIA, da PD e da PNEA.

Resumidamente, pode-se dizer que o processo de construção de séries temporais longas, apresentado neste estudo, visa lidar, primeiro, justamente com as incompatibilidades entre a PNADC e a PNAD, para apenas em seguida, realizar a retropolação. Portanto, o procedimento desenvolvido aqui pode ser dividido em duas etapas. Na primeira etapa, chamada de Processo de Compatibilização, buscamos utilizar os microdados, tanto da PNAD quanto da PNADC, para minimizar os possíveis desníveis existentes entre as séries das duas pesquisas. Mais especificamente, buscamos ajustar os dados da PNAD para reproduzir os números da PNADC. Realizamos esta primeira etapa para o período, compreendido entre 2012 e 2014, em que ambas as pesquisas coexistem. A vantagem de realizar a primeira etapa para um período em que ambas as pesquisas coexistem é poder verificar em que medida os dados da PNAD Ajustada são capazes de reproduzir os números da PNADC.

Na segunda etapa, chamada de Processo de Retropolação, também utilizamos os microdados para ajustar a PNAD, visando obter números compatíveis com a PNADC. Porém, neste segundo passo, produzimos os dados ajustados da PNAD para os anos compreendidos entre 1992 e 2012, período anterior a existência da PNADC. Portanto, não é possível verificar, neste segundo momento, se os dados da PNAD Ajustada são capazes de reproduzir os números da PNADC. Em seguida, utilizando dados da PNAD Ajustada, geramos as séries retropoladas da PNADC, para o período anterior a março de 2012. Finalmente, produzimos a variação mensal das séries retropoladas a partir dos dados da PME.

O presente artigo está composto, além desta introdução, de mais 4 seções. A seção 2 faz uma revisão da literatura relacionada à estimação de séries temporais. A seção 3 explica a metodologia de compatibilização de pesquisas abordada neste estudo. A seção 4 apresenta os resultados das séries retropoladas. E a seção 5, por fim, levanta algumas considerações finais dos resultados encontrados aqui.

2. Revisão da Literatura

A construção e estimação de séries históricas busca atender a três finalidades distintas: i) a retropolação de dados, consistindo no preenchimento de séries históricas anteriores ao período inicial do indicador de interesse; ii) a extrapolação de dados, que preenche séries posteriores ao período final do indicador de interesse; iii) a interpolação de dados, que preenche intervalos de dados faltantes de uma série histórica, para um dado indicador de interesse. Diversas técnicas estatísticas de estimação têm sido desenvolvidas visando atender estas finalidades.

Podemos definir dois tipos de metodologia para a estimação (retropolação, extrapolação ou interpolação) de séries históricas de indicadores, segundo a literatura econômica. O primeiro método estima séries sem depender de indicadores auxiliares e baseando-se unicamente em critérios matemáticos. Harvey e Pereira (1985), por exemplo, realizou a interpolação de séries de fluxos de estoque através de um modelo ARMA univariado.

O segundo método estima séries históricas através de modelos de minimização que tomam como *inputs* indicadores de alta frequência relacionados à série construída, buscando replicar tanto a tendência da referida série histórica quanto às oscilações do indicador de alta frequência. Para este método de estimação, destacam-se os estudos de Denton (1971), que estabeleceu um modelo de minimização quadrática para a interpolação de séries trimestrais.

Outros trabalhos recentes de estimação de séries aplicaram modelos de espaço-estado multivariados, tais como o modelo de componentes não-observados (Harvey e Chung, 2000) (Moauero e Savio, 2005). Este último estudo, em particular, utilizou o modelo de componentes não-observadas para interpolar as séries de Produto Interno Bruto (PIB) de 12 países da Organização para a Cooperação e Desenvolvimento Econômico (OCDE). As séries interpoladas de PIB tomam como referência as variações de 8 indicadores de alta frequência relacionados, tais como taxa de juros, oferta monetária e índice de atividade industrial. Já Mönch e Uhlig, (2005) construiu um método de interpolação híbrido que incorpora 6 modelos distintos de estado-espaço.

As referidas técnicas de estimação vêm sendo aplicadas para retropolação e interpolação de diversas séries de indicadores econômicos brasileiros. O modelo de Mönch e Uhlig foi utilizado para a interpolação de séries mensais de indicadores fiscais de estados e municípios brasileiros (Dias et al., 2010). Também encontramos aplicações destas técnicas de estimação em indicadores do mercado de trabalho brasileiro, entre os quais a PNADC. A técnica de interpolação de minimização quadrática de Denton foi utilizada em Sacchet (2016) para a construção de séries longas e de alta frequência consistentes com a PNADC. O método de Denton foi aplicado para as variáveis de ocupação, desocupação e inatividade entre 1992 e 2014, a partir dos dados da PNAD e da PME.

Outros estudos também têm aplicado metodologias distintas de retropolação para o mercado de trabalho brasileiro. Dias et al. (2010) reconstruíram as séries de emprego e renda da PME anteriormente a 2002, ano em que a pesquisa passou por uma grande revisão de metodologia. Para a retropolação da PNADC, Bacciotti e Marçal (2020) construiu uma taxa de desemprego nacional e de frequência trimestral desde 1976, a partir de modelos em estado-espaço de componentes não-observados. Este método utilizou como inputs a série histórica da Pesquisa de Emprego e Desemprego do Dieese e da PNAD Anual. E por fim, Alves e Fasolo (2015) partiu de um método de vetores auto regressivos de frequência mista (MF-VAR) para simular como teria evoluído a taxa de desemprego da PNADC desde maio de 1999.

O estudo aqui apresentado buscará estimar séries retropoladas de indicadores da PNADC, baseando-se no uso de microdados de pesquisas de mercado de trabalho disponíveis nos anos 1990 e 2000, dentre os quais a PME e a PNAD Anual. A metodologia utilizada neste estudo apresenta vantagens e desvantagens com relação a estudos similares sobre o tema. O presente artigo aborda um método mais descritivo e qualitativo e menos quantitativo, baseado em compatibilizações de definições de pesquisas. Este método contrasta com os demais abordados na literatura, que são baseados em métodos quantitativos de estimação e modelagem econométrica. A metodologia de construção de séries retropoladas a partir da compatibilização de pesquisas, portanto, tende a evitar eventuais vieses e erros de estimação que poderiam ocorrer em metodologias que utilizam modelos econométricos. O processo de compatibilização utilizado neste artigo é explicado em maiores detalhes na próxima sessão.

3. O Processo de Compatibilização

Existem três fontes principais de divergências entre a PNAD e a PNADC. O primeiro responsável pela existência de diferenças entre as duas pesquisas é o método de construção da amostra. Por exemplo, a PNADC possui uma amostra expressivamente maior do que a da PNAD.⁹ Dado que esta primeira fonte divergências entre PNAD e PNADC nasce no processo de coleta, não há nada que possa ser feito para compatibilizar as duas pesquisas nesta dimensão. Portanto, depois de realizado o processo de compatibilização entre PNAD e PNADC, descrito nesta seção do presente artigo, é possível que haja diferenças remanescentes justamente em virtude das distinções iniciais na forma como a amostra é construída nas duas pesquisas.

A segunda divergência entre PNAD e PNADC é decorrente do universo considerado. Por um lado, a PNADC produz uma amostra representativa do Brasil como um todo. Por outro lado, a PNAD não inclui, antes de 2004, os municípios das zonas rurais da Região Norte (compreendida por Acre, Rondônia, Roraima, Amazonas, Pará, Amapá), o que gera distorções de níveis importantes para o período pré- 2004. Esta segunda fonte de divergências é corrigida aqui, no presente artigo. Estimamos a população das zonas rurais da Região Norte entre 1992 e 2004 a partir de dados disponíveis do Censo Demográfico de 1991 e 2000. Foram encontrados os níveis correspondentes de PIA, PEA e PO da população rural norte em tais anos. Os demais anos foram estimados a partir de projeções de crescimento médio demográfico (para encontrar a PIA) e de proporções médias de taxas de participação (PEA/PIA, para encontrar a força de trabalho média) e taxa de ocupação (PO/PEA, para encontrar o nível médio de ocupação).

A terceira divergência entre PNAD e PNADC decorre de definições distintas que são adotadas nas duas pesquisas. Por exemplo, a PNAD inclui na definição de ocupados aquelas pessoas que trabalham uma hora, ou menos, por semana. Alternativamente, a PNADC não considera indivíduos que trabalham período igual,

⁹ Na PNAD são 1100 municípios coletados, contra 3500 pela PNADC.

ou inferior, a uma hora por semana como ocupados. Esta terceira fonte de diferenças pode ser corrigida com auxílio dos microdados. O Processo de Compatibilização, descrito na presente seção, procura lidar justamente com esta terceira fonte de divergências entre PNAD e PNADC.

Ao todo notamos quatro diferenças nas definições utilizadas em ambas pesquisas. A Tabela 1, abaixo, resume as incompatibilidades encontradas. Em seguida, propomos quatro passos capazes de compatibilizar as definições presentes nas duas pesquisas. No primeiro eliminamos os indivíduos de idade inferior a 14 anos da PNAD. No segundo excluimos da PNAD, indivíduos que trabalham habitualmente um período igual ou inferior a 1 hora por semana. No terceiro eliminamos o trabalho não-remunerado para o próprio consumo/uso da PNAD. Por fim, no quarto, excluimos da PNAD os indivíduos afastados do trabalho. O Apêndice 1 deste artigo apresenta uma descrição mais detalhada dos quatro passos necessários para compatibilizar as definições utilizadas em ambas pesquisas.

Tabela 1: Principais Divergências Entre as Definições da PNAD e da PNADC

Critérios	Subcritérios	Definição da PNAD	Definição da PNADC	Massa de indivíduos alocados*
População em Idade Ativa (PIA)		A partir de 10 anos	A partir de 14 anos	Redução de -12,9 milhões na PIA
Trabalho Remunerado	Tempo de trabalho	Sem restrições ao número de horas trabalhadas na semana	Pelo menos 1 hora trabalhada por semana	Nenhuma alteração nos níveis de PO e PEA
Trabalho na Produção para próprio consumo/uso	Próprio consumo	Apenas aqueles na produção de bens primários (agricultura, pesca, pecuária, etc.) para o próprio uso de pelo menos um membro na unidade domiciliar	Não considera como ocupado	Queda de -4,1 milhões na PO, dos quais +3,9 milhões alocados para a PNEA e +210 mil para a PD
	Próprio uso	Apenas aqueles na construção de edificações, estradas privativas, poços e outras benfeitorias para o próprio uso de pelo menos um membro da unidade domiciliar	Não considera como ocupado	
Trabalhador afastado temporariamente		Não há restrições quanto ao tempo de afastamento para considerar a pessoa como ocupada	Há algumas restrições quanto ao tempo de afastamento para considerar a pessoa como ocupada	Queda de -1,6 milhão na PO, do qual +1,5 milhão alocado para a PNEA e + 100 mil para a PD

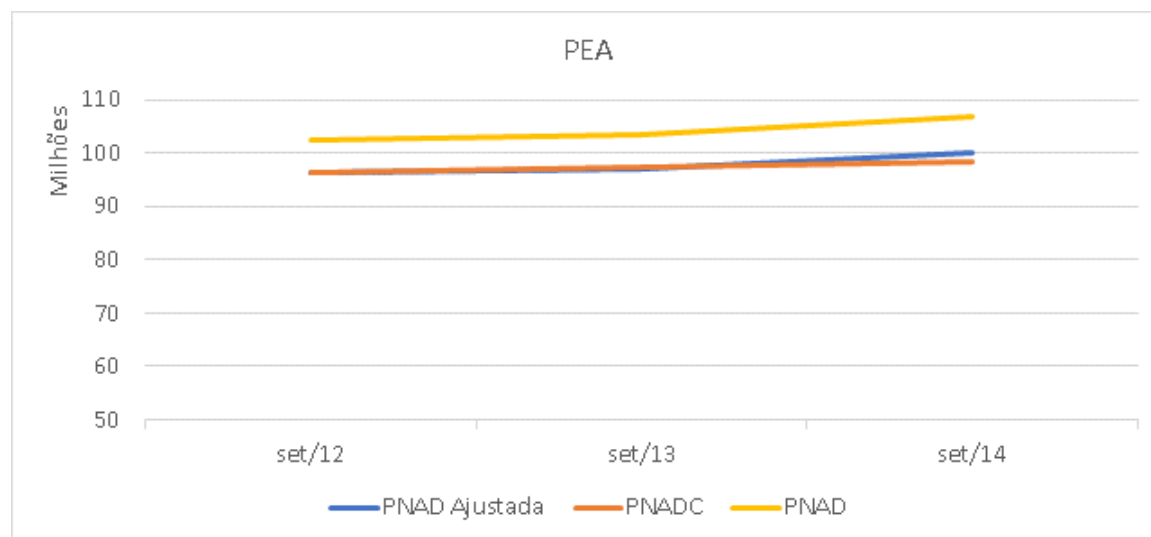
Fonte: Notas Metodológicas (2015), IBGE

*Valores em Média 2012-2014

Por último, uma quinta fonte de divergências entre as duas pesquisas PNAD e PNADC merece ser destacada. Os microdados da PNADC sofreram uma revisão pelo IBGE em meados de 2019, levando a um descolamento entre os níveis da PNAD pré-2012 e os níveis PNADC pós-2012¹⁰. Este descolamento foi corrigido por um novo ajuste nos níveis finais da PNAD compatibilizada. Com isso, os indicadores da PNAD sofreram uma revisão para baixo, com a eliminação de 390 mil para a PO e de 435 mil para a PEA. A seguir, aplicamos os passos para compatibilização, descritos acima, aos dados da PNAD para gerar novas séries de emprego e renda. O objetivo é produzir séries ajustadas da PNAD que sejam mais semelhantes aos dados na PNADC.

Nos gráficos 1, 2 e 3 abaixo, observamos que os resultados da PNAD Ajustada estão bem próximos daqueles apresentados pela PNADC. No gráfico 1, conseguimos reduzir expressivamente a diferença inicial na PEA entre a PNAD e a PNADC. Por exemplo, os dados da PNAD apontam para uma PEA de 106,8 milhões em setembro de 2014. Já a PNAD Ajustada sugere, para o mesmo mês, uma PEA de 100,1 milhões, número que é significativamente mais próximo dos 98,4 milhões apresentados pela PNADC. No gráfico 2, o mesmo procedimento é realizado para PO. A PNAD aponta para uma PO de 99,4 milhões em setembro de 2014, contra 91,8 milhões na PNADC. Com o processo de compatibilização, obteve-se uma diminuição desta diferença em 7,0 milhões, com a PO da PNAD Ajustada reduzindo-se para 92,4 milhões. Por fim, no gráfico 3, a Renda Real também apresenta modificação significativa, tornando-se mais próxima da PNADC. Em setembro de 2014, os R\$2149,00 da PNAD sobem para R\$ 2330,00 na PNAD Ajustada (o nível da PNADC é de R\$2316,00)^{11 12}.

Gráfico 1



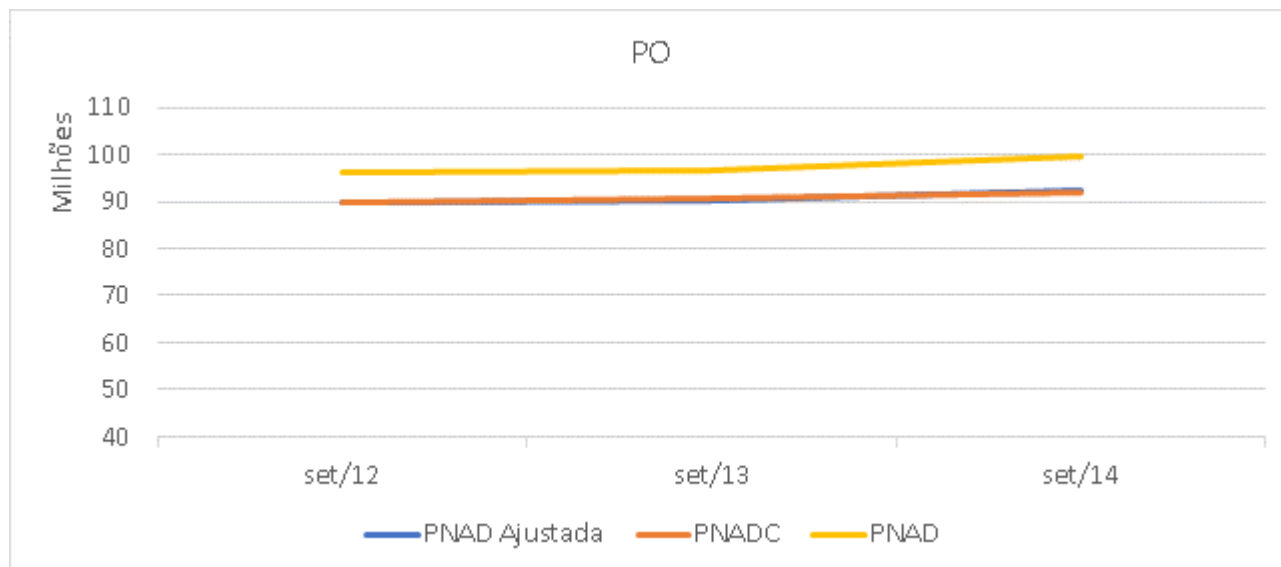
Fonte: PNAD e PNADC (IBGE)

¹⁰ Maiores detalhes, ver Apêndice 1

¹¹ Vale ressaltar que a Renda Real é na verdade a Renda Média Real do Trabalho

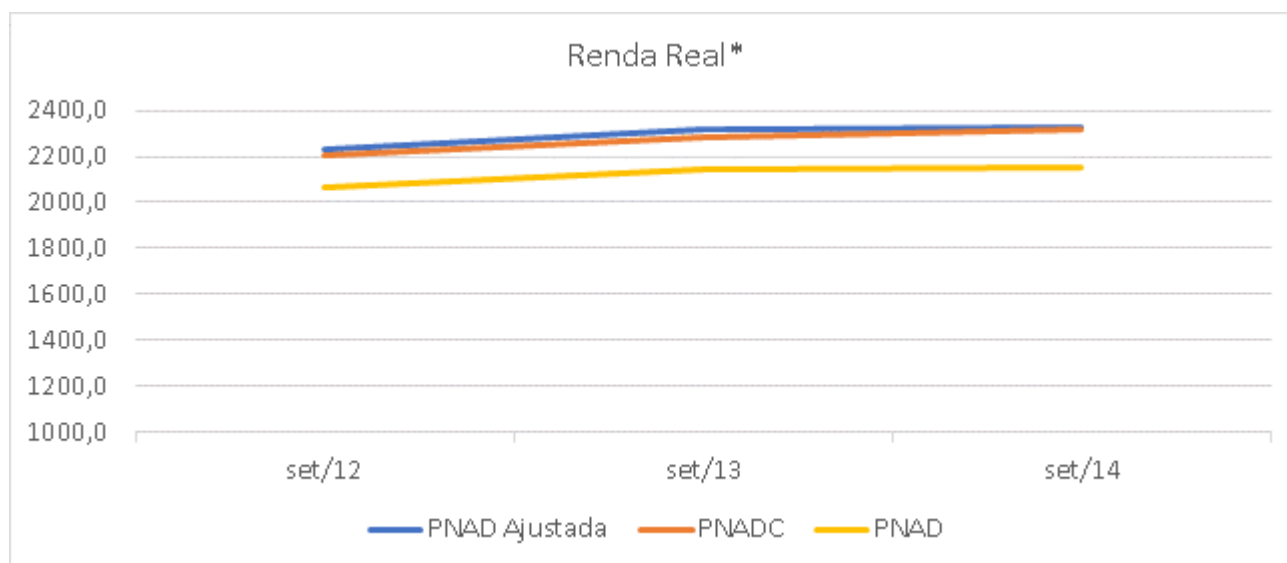
¹² Os rendimentos reais da PNAD ajustada são deflacionados segundo um índice de preços composto tanto pelas variações do IPCA quanto pelas variações do INPC. Na sessão Apêndice, descrevemos com maiores detalhes este método de deflacionamento.

Gráfico 2



Fonte: PNAD e PNADC (IBGE)

Gráfico 3



Fonte: PNAD e PNADC (IBGE)

*A preços correntes de outubro/2019

Na tabela 2, abaixo, apresenta-se a contribuição relativa de cada um dos quatro passos propostos, acima, para a minimização das diferenças entre PNAD e PNADC. Por exemplo, pode-se verificar na tabela que existia uma diferença de 6,77 milhões entre a PEA apresentada na PNAD e aquela fornecida pela PNADC (a PEA apresentada na tabela é na verdade uma média para os meses de setembro de todos os anos entre 2012 e 2014). Dentre os critérios de ajuste utilizados para a eliminação desta diferença, os principais foram a retirada de trabalhadores para o próprio consumo/uso (que contribui para uma queda de -3,8 milhões na variação absoluta em relação à PNAD) e pela exclusão de trabalhadores afastados (queda de -1,5 milhões).

Ao todo, os ajustes corrigem 94% da diferença inicial entre a PEA da PNAD e aquela da PNADC (ou - 6.361.730), porém acabam por produzir uma variação levemente maior do que a necessária. Mais precisamente, os ajustes produzem uma variação total que representa um pequeno erro de 6,0% (ou - 411.249) da diferença inicial entre a PEA de ambas pesquisas

Ainda na tabela 2, apresentamos o mesmo tipo de comparação para os demais indicadores. A PO é menor em 6.688.760 na PNADC com relação à PNAD. Neste caso, o trabalho para o próprio consumo/uso e os trabalhadores afastados constituem os principais componentes explicativos para as discrepâncias de nível entre as pesquisas (que contribuem para queda absoluta de -4,1 milhões e -1,6 milhões, respectivamente). Os ajustes eliminam 98,7% das diferenças iniciais (-6.688.760) e produzem um erro de 1,3% (86.578). Na mesma tabela, verifica-se que é possível eliminar grande parte da diferença absoluta de 148 reais entre a Renda Real da PNADC e da PNAD, a partir da exclusão do trabalho para o próprio consumo/uso e da eliminação de indivíduos de idade inferior a 14 anos (97,2 e 10,1 reais, respectivamente). No total, os ajustes eliminam 116,2% das diferenças (101 reais), com uma componente não explicada de erro de 16,24% (24 reais).

Tabela 2: Decomposição da compatibilização da PNAD na PNADC segundo critérios (Média 2012-14)

Níveis por metodologia			
Critérios	PEA	PO	Renda Real
Total PNAD Contínua (A)	97.456.667	90.713.667	2267,7
Total PNAD (B)	104.229.645	97.402.427	2119,3
Diferença (C = A-B)	-6.772.979	-6.688.760	148
Variação por procedimento			
Critérios	PEA	PO	Renda Real
14 anos (D)	-503.816	-469.610	10,1
S/Trabalhadores próprio consumo/uso (E)	-3.894.493	-4.104.911	97,2
S/Trabalhadores afastados (F)	-1.527.640	-1.636.933	4,8
S/Trabalho inferior a 1 h (G)	0	0	0
Ajuste Reponderação PNADC 2019 (H)	-435.780	-390.729	60,3
Parte Explicada da Variação (I= D+E+F+G+H)	-6.361.730	-6.602.182	172
Parte Não Explicada (J-I)	411.249	86.578	24

Fonte: PNAD e PNADC (IBGE)

4. O Processo de Retropolação

As séries, analisadas na seção acima, estão restritas ao período, compreendido entre 2012 e 2016, em que PNAD e PNADC coexistem. Resta agora criar as séries retropoladas para o período pré-2012. Ao todo, o processo de construção das séries retropoladas para os anos anteriores a 2012 utiliza três pesquisas. A primeira é a PME, de divulgação mensal e disponibilidade de dados a partir de jan/1991. A segunda é a PNAD, de divulgação anual e disponível de 1992 a 2014. Por último, a PNADC, divulgada mensalmente desde março de 2012.

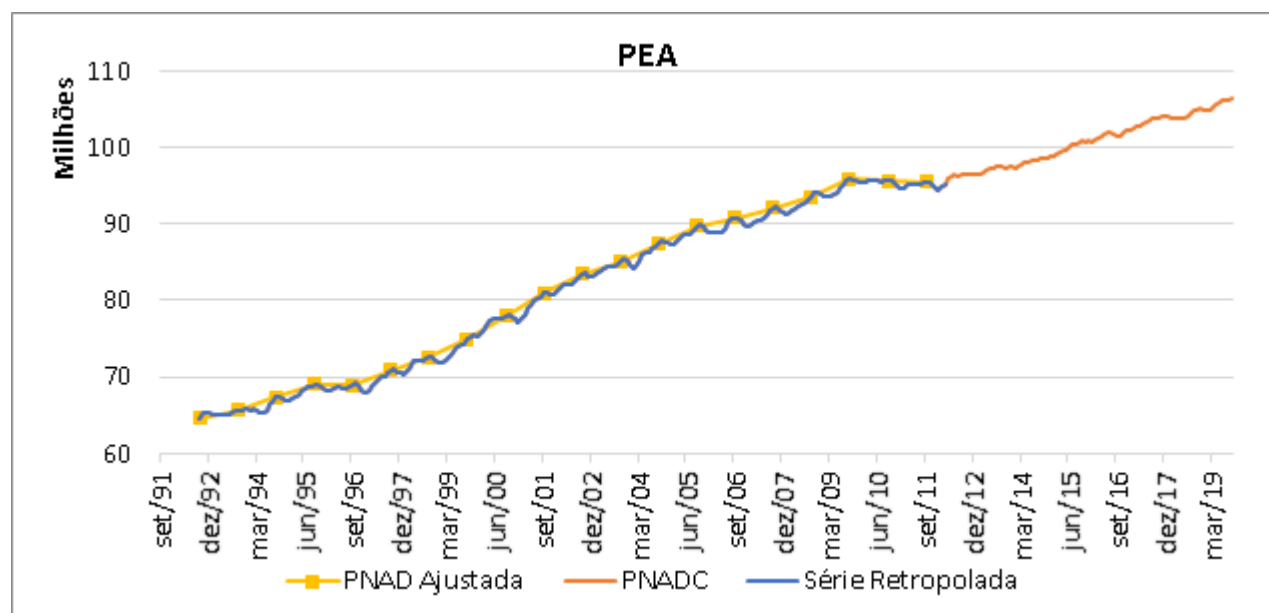
A retropolação divide-se em duas etapas. Na primeira, os níveis retropolados são construídos de modo a coincidir, a cada setembro, com os níveis da PNAD Ajustada, para todos os anos entre 1992 e 2011. Em seguida, a variação mensal das séries retropoladas é produzida a partir dos dados da PME, redefinindo sempre as tendências de modo a garantir que os níveis de setembro, determinados pela PNAD Ajustada, sejam atingidos.¹³¹⁴

Nos gráficos 7, 8, 9 e 10, traçamos as Séries Retropoladas para PEA, PO, Taxa de Desemprego (TD) e Renda Real, respectivamente. Como já dito anteriormente, a partir da metodologia de retropolação, apresentada aqui, conseguimos gerar séries expressivamente mais longas do que as inicialmente fornecidas pela PNADC. Enquanto as séries originais, disponibilizadas pela PNADC, começam em março de 2012, as novas séries, produzidas a partir da interpolação dos dados da PNAD Ajustada, têm início em setembro de 1992.

13 Para replicar de forma mais precisa os movimentos sazonais da PNADC detectamos a necessidade de utilizar as variações mensais da PME, porém defasadas em 1 mês com relação ao mês corrente. Ver o Apêndice 2 para maiores detalhes.

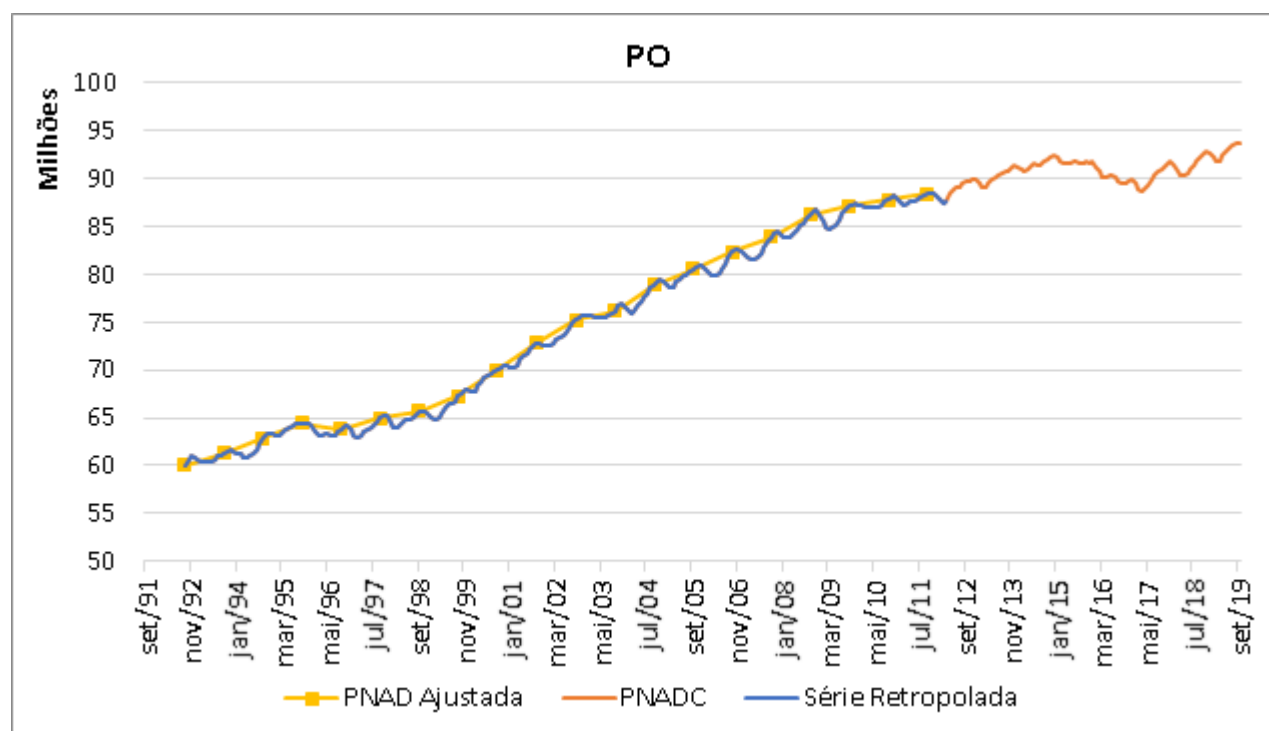
14 Para gerar séries coincidentes com as séries da PNAD Ajustada, nos meses de setembro, realizamos uma correção nas variações mensais da PME. Mais precisamente, padronizamos a volatilidade da PME de forma a seguir a tendência de crescimento anual da PNAD Ajustada. Assim, conseguimos construir séries retropoladas que captam tanto a tendência anual presente na PNAD Ajustada como os ruídos das variações mensais existentes na PME. Ver o Apêndice 3 para maiores detalhes.

Gráfico 7



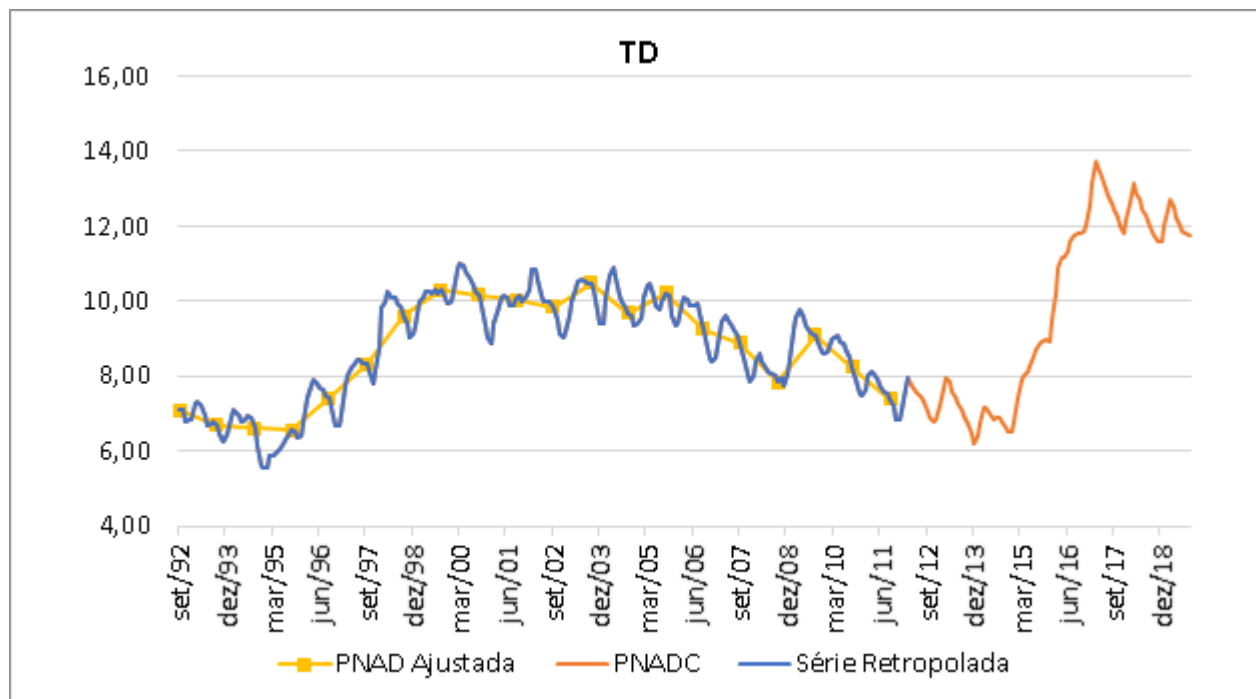
Fonte: PNAD, PME e PNADC (IBGE)

Gráfico 8



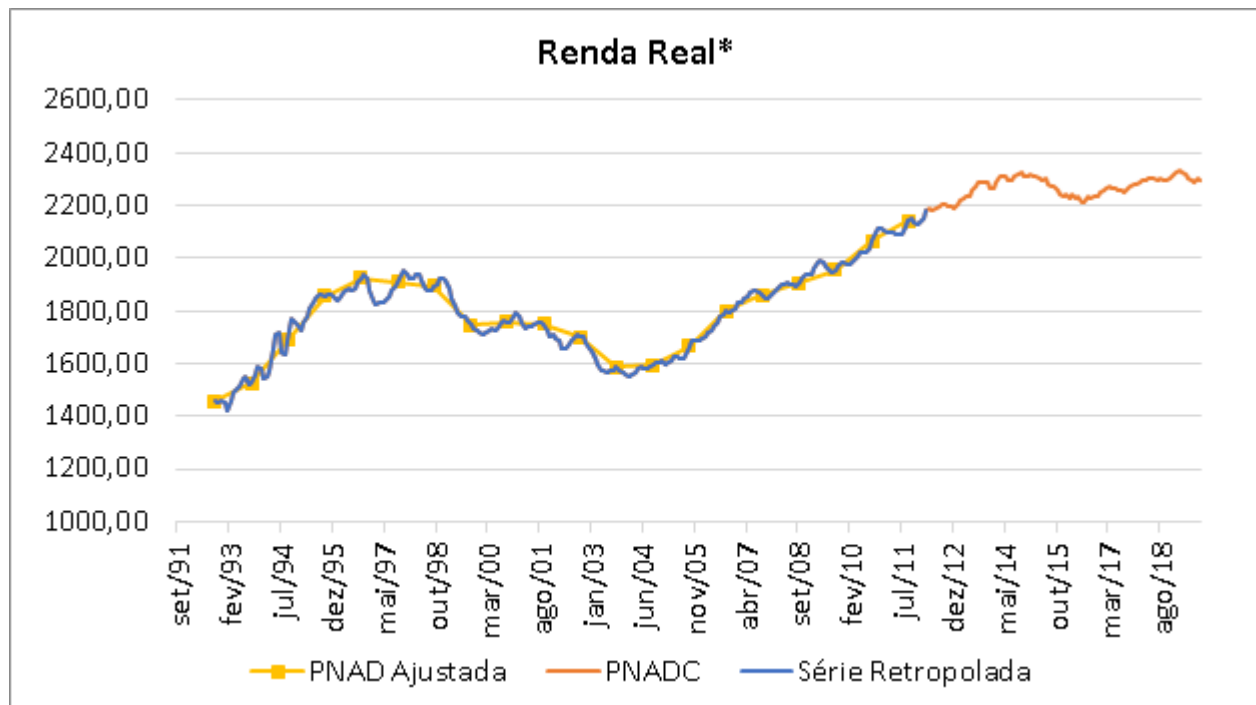
Fonte: PNAD, PME e PNADC (IBGE)

Gráfico 9



Fonte: PNAD, PME e PNADC (IBGE)

Gráfico 10



Fonte: PNAD, PME e PNADC (IBGE)

*Preços correntes de outubro/2019

Observamos que as novas séries apresentam trajetórias bastante razoáveis, segundo os contextos vividos na

economia brasileira. Para a série retropolada da taxa de desemprego, no gráfico 9, observamos picos nos períodos recessivos de 1998-2000, 2002-2004 e no ano de 2009, momentos que a referida taxa alcançou os valores de 10,9%, 10,8% e 9,7%, respectivamente. Os resultados indicam ainda, que a taxa de desemprego da PNADC alcançou, em março de 2017, o seu recorde em toda a série histórica (no referido mês a taxa de desemprego chegou ao patamar de 13,7%).

Também observamos um comportamento mais volátil da renda real nos anos 90 e início dos 2000 (gráfico 10). Nas crises de 1998-2000 e 2002-2003, a renda apresentava quedas bruscas de nível, um fenômeno que não se repetiu em um cenário de crise de maior magnitude como o atual. Isso corrobora a tese de aumento da rigidez dos salários reais na economia brasileira, dado a crescente formalização dos trabalhadores ao longo da década de 2000 e 2010.

5. Observações Finais

O fim da PME, restrita às 6 principais regiões metropolitanas, criou inúmeras dificuldades relacionadas a continuidade das séries de emprego e renda. Com a metodologia de retropolação, descrito neste artigo, acreditamos ter conseguido não somente contornar o problema da incapacidade de gerar séries longas a partir da PNADC, como também fornecer um importante ferramental para compreendermos o comportamento histórico do emprego e da renda nacionalmente.

Finalmente, vale ressaltar que a metodologia de retropolação apresentada aqui pode ser estendida às demais variáveis categóricas da PNADC. Portanto, é possível aplicar os mesmos processos para gerar séries retropoladas para os seguintes recortes da PO: (i) por vínculo, (ii) por região e (iii) por setor. Também é possível estender a série para períodos anteriores a 1992, uma vez que há dados da PNAD e da PME disponíveis desde 1980. No entanto, deixaremos estes e outros exercícios para trabalhos futuros.

6. Bibliografia

- ALBIERI, S. e BIANCHINI, Z.M (2015), “Principais Aspectos de Amostragem das Pesquisas Domiciliares do IBGE - Revisão 2015 / IBGE “. Rio de Janeiro: IBGE. 46p (Texto para Discussão, nº 55).
- ALVES, L. e FASOLO, A. (2015).” Not Just Another Mixed Frequency Paper”. (No. 400).
- BACCIOTTI, R. e MARÇAL, E. (2020). “Taxa de Desemprego no Brasil em quatro décadas: retropolação da PNAD contínua de 1976 a 2016.” São Paulo: FGV EESP. (Textos para Discussão, nº 522)
- DIAS, V., ISSLER, J. e RODRIGUES, C. (2010). “Interpolação de variáveis fiscais brasileiras usando representação de espaço de estados.” 32º Meeting of the Brazilian Econometric Society.
- DENTON, F. (1971). “Adjustment of monthly or quarterly series to annual totals: an approach based on

quadratic minimization”. *Journal of the American Statistical Association*, 66(333), 99-102.

HARVEY, A., e CHUNG, C. (2000). “Estimating the underlying change in unemployment in the UK”. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 163(3), 303-309.

HARVEY, A. e PEREIRA, P. (1985). “The estimation of dynamic models with missing observations”. *Brazilian Review of Econometrics*, 5(2), 81-95.

MOAURO, F., e SAVIO, G. (2005). “Temporal disaggregation using multivariate structural time series models”. *The Econometrics Journal*, 8(2), 214-234.

MÖNCH, E. e UHLIG, H. (2005). "Towards a Monthly Business Cycle Chronology for the Euro Area". *Journal of Business Cycle Measurement and Analysis* 2(1).

CORSEUIL, C.H. e FOGUEL, M. (2002), “Uma sugestão de deflatores para rendas obtidas a partir de algumas pesquisas domiciliares do IBGE”. Rio de Janeiro: IPEA (Texto para Discussão, nº897)

MOREIRA, A. CORSEUIL, C.H. e FOGUEL, M. (2016), “Um Modelo de Transição entre Estados da Força de Trabalho para a Previsão do Desemprego e a Ocupação Formal”. Rio de Janeiro: Ipea. (Carta de Conjuntura nº32 jul-set 2016)

NOTAS Metodológicas (2015), “Diferenças Metodológicas das pesquisas PNAD, PME e PNADC”. Rio de Janeiro: IBGE.

PESQUISA Nacional por Amostra de Domicílios Contínua (2018), “Deflacionamento dos rendimentos do trabalho dos trimestres móveis da PNADC”. Rio de Janeiro: IBGE (Notas Técnicas)

PESQUISA Mensal de Emprego (2014), Rio de Janeiro: IBGE (Série Relatórios Metodológicos, vol.1. 1ª Edição). 40p.

PESQUISA Nacional por Amostra de Domicílios (2014), Rio de Janeiro: IBGE (Síntese dos Indicadores da PNAD 2014). 97p.

PESQUISA Nacional por Amostra de Domicílios Contínua (2007), Rio de Janeiro: IBGE (Notas Técnicas, vol.23. 2ª Edição). 89p.

SACCHET, S. (2016), “Construção de séries longas de alta frequência de indicadores do mercado de trabalho com a PME e a PNADC”. Rio de Janeiro: Ipea. (Carta de Conjuntura nº32 jul-set 2016)

SACCHET, S. (2016), “Metodologia para a construção de séries longas de mercado de trabalho”. Rio de Janeiro: Ipea. (Carta de Conjuntura nº32 jul-set 2016).

Apêndice

- 1) Descrevendo, de forma mais detalhada, os quatro passos necessários para compatibilização das definições presentes na PNADC e na PNAD: Primeiro, é necessário compatibilizar a idade mínima

para trabalhar, ou a PIA mínima, que difere em ambas as metodologias. Diferentemente da PNAD, a PNADC possui uma idade mínima para participação na PIA mais restrita, de 14 anos, contra 10 anos da primeira. Para a realização deste critério de compatibilização, descarta-se nos dados da PNAD todas as observações de indivíduos com idade inferior a 14 anos. Deste modo, conseguimos gerar uma PIA para a PNAD idêntica à PIA da PNADC. Como exemplo, este critério eliminou 12,9 milhões de indivíduos na PIA da PNAD correspondente à média do período 2012-2014.

Segundo, realoca-se o trabalho remunerado inferior a 1 hora semanal. Enquanto para a PNAD não há restrições quanto ao tempo de trabalho para integrar a PO, na PNADC exige-se um mínimo de 1 hora de trabalho semanal para o indivíduo ser nela classificado. Assim, na variável correspondente ao número de horas trabalhadas na semana (v0713), todas as observações que responderam um valor em horas inferior a 1 são eliminadas da PO. Os indivíduos eliminados da PO podem ser realocados tanto para desocupados como para inativos. Caso eles estejam tomando providência para procurar emprego, são realocados para a PD. Caso não estejam, passam a integrar a PNEA, tornando-se inativos. Ao todo, este critério não resultou em nenhuma alteração nos níveis de PO e PEA.

Em terceiro, realoca-se trabalhadores não-remunerados para o próprio consumo e próprio uso. Na PNAD, o trabalho na produção para o próprio consumo/uso é definido como uma atividade de ocupação, estando os indivíduos nela envolvidos integrados na PO. Já na PNADC tais trabalhadores não fazem parte da PO, estando classificados seja como inativos, seja como desocupados. Como critério de compatibilização, elimina-se na PO da PNAD todas as observações correspondentes a indivíduos envolvidos no trabalho familiar agrícola não-remunerado (para o próprio consumo) e trabalho na construção civil (para o próprio uso). Mais especificamente, a variável v4716 do dicionário metodológico da PNAD classifica esses dois regimes de trabalho em categorias separadas, podendo ser facilmente identificadas e eliminadas da PO total. Em seguida, tais indivíduos são realocados tanto para a PD como para a PNEA, assim como na reclassificação acima para os indivíduos que trabalham uma hora ou menos. Na prática, este critério eliminou em média 4,1 milhões de indivíduos na PO da PNAD, no período 2012-2014, dos quais 3,9 milhões foram alocados para a PNEA e 210 mil para a PD.

Em quarto, realoca-se trabalhadores afastados temporariamente do trabalho na semana de referência. Os trabalhadores afastados são considerados ocupados pela PNAD, enquanto na PNADC há restrições quanto à sua classificação na PO, integrando o grupo de inativos. Como critério de compatibilização, a variável v0705 da PNAD identifica trabalhadores afastados, podendo ser descartados da PO todas as observações que respondem sim à pergunta. Neste caso, diferentemente da eliminação dos trabalhadores remunerados com trabalho semanal inferior a 1 hora e dos não-remunerados para próprio consumo/uso, que são direcionados tanto para a PD como para a PNEA, os indivíduos integrantes do

trabalho afastado são exclusivamente realocados para a PNEA. Este critério eliminou 1,6 milhões de indivíduos da PO da PNAD, sendo 1,5 milhões realocados como inativos na PNEA e 100 mil realocados como desocupados na PD.¹⁵

Um último ajuste realizado corresponde à revisão dos dados da PNADC realizados pelo IBGE em meados de 2019. Os microdados da PNADC sofreram uma mudança na ponderação dos pesos, promovendo uma revisão para baixo nos níveis PNADC de PEA, PO e PD para o período pós-2012. Esta revisão dos níveis PNADC levaram a um descolamento entre os níveis da PNAD ajustada pré-2012 e os níveis PNADC pós-2012. Com isso, foi necessário revisar os níveis finais da PNAD compatibilizada, aplicando-as como fator uma constante. Esta constante, aplicada aos níveis PNAD pré-2012, corresponde à razão entre os níveis médios PNADC pré-revisão e PNADC pós-revisão. Com isso, os indicadores da PNAD sofreram um ajuste para baixo, com a eliminação de 390 mil para a PO e de 435 mil para a PEA.

Existem outras pequenas diferenças de metodologia, que foram ignoradas neste método de compatibilização. Entre elas estão a forma como são classificadas algumas categorias de trabalho não-remunerado presentes na PNAD, como aprendiz/estagiário (considerado trabalho remunerado na PNADC), ajuda a instituição religiosa/beneficente (considerado não- ocupado pela PNADC) e trabalho auxiliar familiar (restrito a atividades agrícolas e de construção pela PNAD). Entretanto, tais grupamentos são muito pequenos em relação à população total, tendo assim pouco impacto nas discrepâncias de níveis das duas pesquisas.

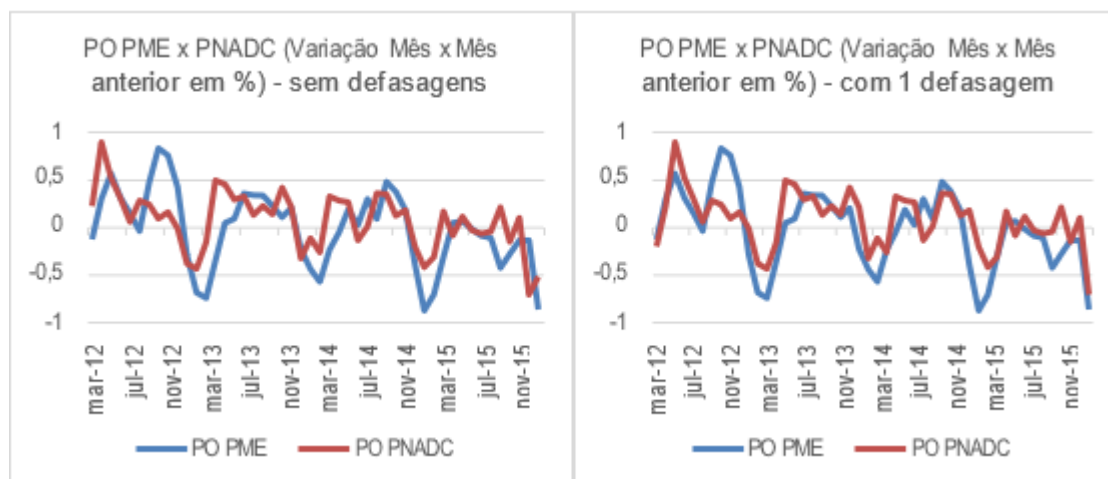
- 2) Para replicar de forma mais precisa os movimentos sazonais da PNADC detectamos a necessidade de utilizar as variações mensais da PME, porém defasadas em 1 mês com relação ao mês corrente. Mais precisamente, realizamos este procedimento para os seguintes indicadores: (i) PO, (ii) PEA, (iii) PIA e (iv) Renda Real. Como exemplo, da necessidade do uso da defasagem em 1 mês nas variações da PME, para replicar os movimentos sazonais da PNADC, vamos analisar o caso da PO. Tivemos que aplicar uma defasagem de 1 mês nas variações da PO advinda da PME, principalmente, em virtude das sazonalidades mensais do período entre dezembro e março (meses em que ocorrem as festas de fim de ano e as férias). Estes meses impactam significativamente na PO, acarretando em grandes oscilações deste indicador. É possível notá-las nos dois gráficos abaixo. Estes gráficos mostram que a PME, ajustada em 1 defasagem, é melhor correlacionada com as variações da PNADC. Dado que o período de festas de fim de ano também afeta os demais indicadores de mercado de trabalho (como PEA, PIA

¹⁵ Esta reclassificação difere um pouco da definição da PNADC. Nesta pesquisa, apenas alguns grupos que compõem a categoria de trabalho afastado são excluídos da PO, enquanto outros, como o grupamento de afastamento por férias remuneradas, são mantidos. Assumimos por aproximação que todos os trabalhadores afastados do trabalho estão fora da PO, incluindo aqueles afastados por motivos de férias ou licença médica. Não existe a possibilidade de separar internamente na PNAD estas categorias das demais.

e Renda Real) tivemos que replicar, nestes casos, o procedimento de adotar a defasagem de 1 mês nas variações da PME.

- 3) Para replicar de forma mais precisa os movimentos sazonais da PNADC detectamos a necessidade de realizar algumas adaptações às variações mensais observadas na renda PME, entre as quais: i) defasagem das variações em 1 mês com relação ao mês corrente da PNADC; ii) eliminação dos picos sazonais dos meses novembro-janeiro nas variações PME do período 1992-2002 . Para o procedimento de defasagem de 1 mês, na etapa i, buscamos ajustar os seguintes indicadores: (i) PO, (ii) PEA, (iii) PIA e (iv) Renda Real. Como exemplo da necessidade do uso da defasagem em 1 mês nas variações da PME, para replicar os movimentos sazonais da PNADC, vamos analisar o caso da PO. Tivemos que aplicar uma defasagem de 1 mês nas variações da PO advinda da PME, principalmente, em virtude das sazonalidades mensais do período entre dezembro e março (meses em que ocorrem as festas de fim de ano e as férias). Estes meses impactam significativamente na PO, acarretando em grandes oscilações deste indicador. É possível notá-las nos dois gráficos abaixo. Estes gráficos mostram que a PME, ajustada em 1 defasagem, é melhor correlacionada com as variações da PNADC. Dado que o período de festas de fim de ano também afeta os demais indicadores de mercado de trabalho (como PEA, PIA e Renda Real) tivemos que replicar, nestes casos, o procedimento de adotar a defasagem de 1 mês nas variações da PME.

Gráfico 1A



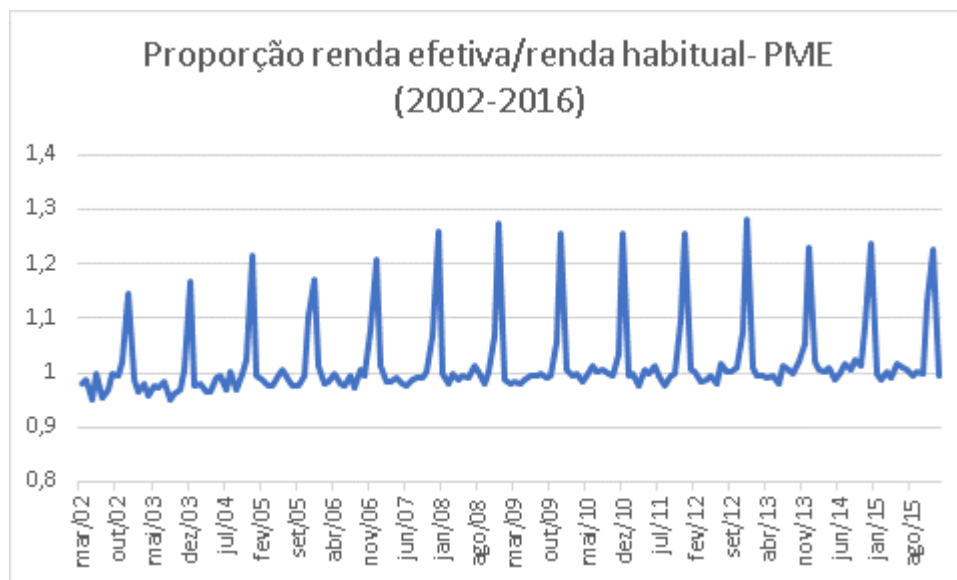
Fonte: PME e PNADC (IBGE)

O procedimento de ajuste da etapa ii, de eliminação dos picos sazonais de novembro-janeiro, decorre de diferenças na definição de rendimento para a PME entre o período de 1992-2001 e 2002-2016. O período 1992-2001 define rendimento como o total de ganhos salariais recebidos pelo trabalhador, incluindo tanto o salário mensalmente recebido quanto outros ganhos adicionais, decorrentes de bônus,

pagamentos extras e 13º salário. Em contraste, a pesquisa PME do período 2002-2016 realiza uma separação entre os rendimentos que excluem estes ganhos adicionais (rendimento habitualmente recebido) e os que incluem estes ganhos (rendimento efetivamente recebido). Convencionamos em adotar a definição de rendimento habitual como objeto de nossa retroposição, em função de ser uma série que apresenta uma maior suavidade nas tendências e por melhor acompanhar movimentações cíclicas de atividade econômica.

Podemos notar, assim, que o rendimento efetivo supera o habitual, em especial nos períodos de fim de ano, correspondente aos meses de pagamento de 13º salário. No gráfico abaixo, podemos notar esses descolamentos periódicos entre os rendimentos habituais e efetivos, concentrados em novembro-janeiro. No mesmo gráfico, também podemos notar que estes descolamentos entre o rendimento habitual e efetivo seguem a uma proporção relativamente constante ao longo dos anos. Como exemplo, nos meses de dezembro, o rendimento efetivo é superior ao habitual a um fator relativamente estável em torno de 1,1 a 1,2.

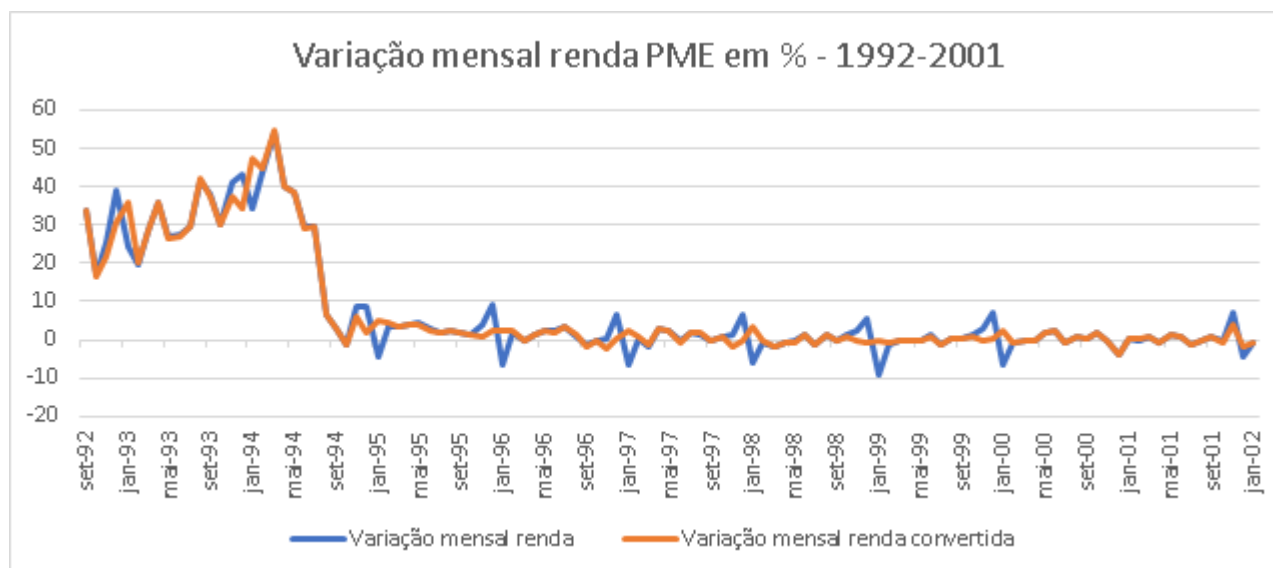
Gráfico 2A



Fonte: PME (IBGE)

Assumido o rendimento PME pré-2002 como o rendimento efetivo (por incluir ganhos de outras fontes como 13º salário), e as razões renda efetiva/habitual como relativamente estáveis, constantes e fixas para todo mês, podemos converter os níveis mensais dos rendimentos efetivos mensais da PME, existentes para o período 1991-2002, em níveis mensais de rendimentos habituais mensais. Conseguimos, deste modo, eliminar os principais picos de variação do rendimento PME, como podemos verificar no gráfico 3A abaixo.

Gráfico 3A



Fonte: PME (IBGE)

Para gerar séries coincidentes com as séries da PNAD Ajustada, nos meses de setembro, realizamos uma correção nas variações mensais da PME. Esse ajuste é necessário devido a existência de diferenças de tendências de crescimento entre as séries da PME e da PNAD Ajustada. Sem essa padronização das variações mensais da PME, à medida que tais variações vão sendo acumuladas mês a mês, poderia haver um descolamento crescente entre o nível da série retropolada e o nível da PNAD Ajustada para os meses de setembro.

A padronização da variação mensal da PME é feita em duas etapas: na primeira verifica-se a tendência de crescimento anual setembro a setembro da PNAD Ajustada em um determinado ano. Em seguida, busca-se calibrar os ruídos mensais da PME ao redor desta tendência anual, de modo que ambas as variações levem a um nível exatamente igual ao nível da PNAD Ajustada desejado em setembro. Assim, conseguimos construir uma série retropolada capaz de captar tanto a tendência anual presente na PNAD Ajustada quanto as variações mensais existentes na PME.

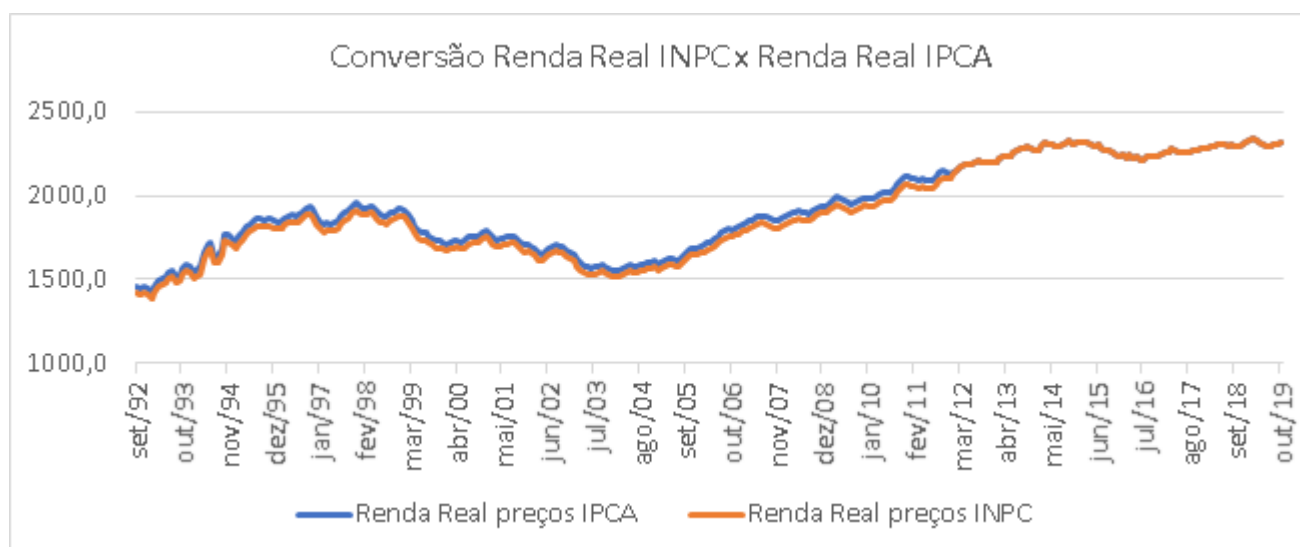
- 4) Os deflatores adotados para rendimentos da PNAD ajustada também necessitam de um tratamento de compatibilização, buscando a i) correção de erros decorrentes do deflacionamento do período inflacionário 1992-1994; ii) correção de erros decorrentes de desníveis de rendimentos entre a série deflacionada pós-2012 e a série deflacionada pré-2012.

A etapa i) é necessária de ser realizada, uma vez que notamos uma elevada volatilidade no período inflacionário brasileiro pré-Plano Real. Para contornar os erros da etapa i, optamos por utilizar uma

série de deflacionamento suavizada, construída a partir da metodologia de Corseuil e Foguel (2002). Este deflator suavizado, construído pelo método Corseuil e Foguel, baseia-se no indicador INPC e estende-se de 1980 a 2002.

Quanto à etapa ii), é preciso que esta se realize, uma vez que a série de rendimentos deflacionada pré-2012, que utiliza o deflator INPC Corseuil e Foguel, observa níveis de renda descolados em relação à série de rendimentos pós-2012, que utiliza o deflator IPCA. Este descolamento decorre do fato dos deflatores INPC e os deflatores IPCA serem variações de preços baseadas em unidades de medida de preços distintas, uma segundo o índice de preços INPC e outra segundo o índice de preços IPCA. Podemos harmonizar estas unidades de medida aplicando um fator fixo a toda a série de rendimentos pré-2012, convertendo os níveis de preços medidos segundo o índice INPC para níveis de preços medidos segundo o índice IPCA. Podemos pensar neste caso de forma análoga a conversões cambiais. Os índices de preços, assim como unidades de medida monetárias, podem ser convertidos de uma unidade monetária para outra unidade monetária a partir de uma proporção ou uma “taxa cambial”. Como exemplo, o índice de preços INPC, em outubro/2019, é definido, segundo o IBGE, pelo valor base de 5233 (para número índice fixado em dezembro 1993=100). Em contraste, o índice de preços IPCA, no mesmo período de outubro/2019, é definido como o valor base de 5355 (também para número índice fixado em dezembro 1993=100). A conversão da medida INPC de preços para a medida IPCA de preços se dará a partir de um fator ou “câmbio” de conversão, de $5355/5233=1,023344$. Aplicamos em seguida este fator para toda a série histórica de rendimentos pré-2012, convertendo reais INPC para reais IPCA. Os resultados obtidos podem ser notados no gráfico abaixo. O gráfico revela que este procedimento de conversão preços INPC- preços IPCA praticamente elimina os desníveis da renda pré-2012 e pós-2012.

Gráfico 4A



Fonte: PME (IBGE)