

O valor dos dados: uma estimativa do custo de (não) atualizar os índices de preços no Brasil

Autores: Ricardo Vale, Otavio Conceição e Gabriel Lara Ibarra.

- As pesquisas de orçamento familiar são a base para ajustar os ponderadores dos índices de preços ao consumidor. No caso brasileiro, tanto a estrutura de ponderação do INPC quanto a do IPCA são ajustadas em decorrência dos resultados encontrados na POF, sendo de suma importância para o orçamento do governo, bem-estar das famílias, política monetária e reajuste de contratos do setor privado.
- No caso do INPC no Brasil, a baixa periodicidade com que os dados de consumo das famílias são incorporados no índice pode gerar grandes diferenças na magnitude da inflação mensurada. Estima-se que um adiamento de um ano na atualização da estrutura de ponderação do índice em 2020 levaria a uma diferença de 0,8 pontos percentuais na inflação anual medida. A diferença da inflação anual poderia chegar a 1,1 pontos percentuais em 2022.
- O INPC é o índice de referência para o reajuste do salário mínimo e para os benefícios previdenciários. Uma simples extrapolação dos resultados contrafactuais sugere que os gastos com benefícios previdenciários em 2021 teriam sido R\$ 5,5 bilhões maiores caso a estrutura de ponderação do INPC não tivesse sido atualizada em 2020.
- Embora esse resultado seja contexto-específico para o Brasil e num período de tempo particular, ele demonstra que a POF é um investimento importante que contribui para a produção de diversas estatísticas oficiais do país, incluindo na elaboração do orçamento público e na formulação da política monetária. Qualquer análise de custo-benefício dessa pesquisa-chave do sistema estatístico nacional deve, portanto, levar em conta as consequências de interrupções na sua frequência de realização.

Introdução

Os índices de preços ao consumidor são centrais para monitorar, orientar e definir o caminho de desenvolvimento econômico de um país. Ao capturar como os preços variam ao longo do tempo, um índice de preços ao consumidor fornece uma medida da evolução do custo de vida das famílias. O impacto dos índices de preços na economia é muito amplo, afetando desde o reajuste das aposentadorias e pensões à política monetária do Banco Central, da transferência de renda dos programas sociais aos contratos do setor privado. A mensuração da inflação tem, portanto, consequências reais na evolução do país.

Um fato importante sobre os índices de preços é que eles são calculados tendo como referência a cesta de consumo da população – informação que vem de pesquisas nacionais domiciliares de orçamento das famílias. Tais pesquisas coletam informações detalhadas sobre o padrão de consumo das famílias em um determinado momento do tempo. Em particular, elas permitem aferir a importância relativa dos diferentes itens de consumo através de sua participação nas despesas totais da família. Dessa forma, os indicadores sobre os itens de consumo podem ser agregados (em nível local ou nacional) em um índice de preços.

A relação dos índices de preço com as pesquisas domiciliares de orçamento demonstra a importância da disponibilidade de informação de qualidade e atualizada sobre os padrões de consumo das famílias. De fato, os índices de preço têm suas estruturas de ponderação ajustadas com certa frequência para refletir a realidade

da alocação de despesas de suas populações-alvo. Isso é fundamental, pois defasagens na captura de mudanças nos padrões de consumo, como a substituição de câmeras fotográficas, videocassetes e telefones fixos por smartphones e serviços de streaming podem levar a estimativas de custo de vida muito diferentes. Além disso, ajustes do consumo das famílias ao seu nível de renda, o qual é correlacionado com o crescimento econômico e seus ciclos, também são um fator primordial na determinação real dos custos de vida. Não sem razão, a ONU (2009, parágrafo 4.45) sugere que, idealmente, informações sobre a alocação das despesas familiares sejam coletadas todos os anos, enquanto o FMI et al. (2020, parágrafo 3.43) recomenda um intervalo de cinco anos.

Nesta nota, estudamos o caso brasileiro e focamos nas potenciais implicações fiscais da indisponibilidade de uma pesquisa domiciliar de orçamento das famílias em tempo hábil. A nota apresenta dois exercícios hipotéticos que variam o momento no qual o INPC incorpora novas informações de uma POF. A variação no tempo de adoção das informações de despesas nos permite criar um índice de preços contrafactual que pode ser comparado com o verdadeiro INPC em diferentes momentos do tempo. A diferença entre estes índices fornece uma estimativa, ainda que imperfeita, do efeito de se incorporar a informação sobre a despesa das famílias. Assim, buscamos responder à seguinte questão: qual teria sido o impacto sobre as despesas do governo caso a atualização do INPC tivesse sido adiada? A nota foca nas despesas com aposentadorias devido à disponibilidade de dados. Dado que existem muitas outras políticas públicas que dependem de estimativas de inflação, os valores apresentados podem ser interpretados como um limite inferior para este impacto.

O caso brasileiro

Índices de preços ao consumidor são definidos a partir de uma cesta representativa do consumo da população-alvo. No caso do Brasil, o Índice Nacional de Preços ao Consumidor (INPC) tem a composição de sua cesta definida e atualizada pelo consumo médio das famílias com renda entre 1 e 5 salários mínimos medido nos dados da Pesquisa de Orçamento Familiar (POF) do Instituto Brasileiro de Geografia e Estatística (IBGE). A POF tem ocorrido com uma periodicidade média de 6,25 anos nas últimas três décadas (isto é, desde a POF 1987/1988). A atualização da cesta de consumo de referência para o INPC só ocorre depois que a POF é coletada e processada. A sequência das duas últimas atualizações das informações da POF no cálculo do INPC foi a seguinte: (i) a POF realizada entre julho de 2008 e julho de 2009 foi incorporada em janeiro de 2012; e (ii) a POF realizada entre julho de 2017 e julho de 2018 foi incorporada em janeiro de 2020.

No Brasil, mudanças na mensuração da variação de preços geram grandes impactos econômicos. Os índices de preços norteiam aumentos dados ao salário mínimo e demais negociações salariais, aposentadorias e pensões, transferências de renda, reajustes de preços em contratos de serviços e fornecimentos, atualização de dívidas, definição de juros e prêmios em ativos financeiros. Em especial, o INPC é o índice de preços oficial utilizado como referência para o reajuste do salário mínimo e como mínimo legal para o reajuste das aposentadorias e pensões pagas pelo governo federal (Art. 41-A, Lei nº 8.213, 1991). O Índice Nacional de Preços ao Consumidor Amplo (IPCA), por sua vez, é o índice de referência para o regime de metas de inflação brasileiro. Assim, toda a política monetária é desenhada em torno do objetivo de trazer o IPCA para o centro da meta. Portanto, divergências na estimativa da cesta típica do consumidor brasileiro têm o potencial de afetar também a taxa básica de juros e, conseqüentemente, o nível de atividade econômica e o emprego.

A não realização frequente da POF tem o potencial de distorcer a mensuração do custo de vida dos consumidores. Passaram-se nove anos entre a conclusão da última pesquisa, de 2017/2018, e a POF anterior, realizada em 2008/2009. Se, em 2008, assistiam-se filmes em DVD e a comunicação ainda se dava bastante pelas chamadas telefônicas tradicionais ou por SMS, em 2017 os serviços de *streaming* (e.g., Netflix) e os softwares de troca de mensagens (e.g., Whatsapp) já eram ferramentas amplamente utilizadas em muitos domicílios do país. Entre as mudanças, podemos destacar o caso da categoria “Alimentação e Bebidas”, que representava 24,5% da cesta de consumo de referência do INPC em dezembro de 2019 e cujo peso caiu para

19,3% com a atualização dos ponderadores em janeiro de 2020, o que representa uma diferença de 5,2 pontos percentuais. Em contraste, a categoria “Comunicação” subiu de 3,4% para 5,7% no mesmo período.

Esta nota estima os impactos que a não atualização da cesta com a periodicidade habitual teria no INPC.

Nosso interesse é compreender as consequências da realização da POF com diferentes periodicidades. Em particular, apresentamos os resultados relativos a dois exercícios contrafactuais diferentes. O primeiro estima impactos de curto prazo em um cenário hipotético que supõe que a atualização da cesta de referência do INPC não tivesse ocorrido em janeiro de 2020 e em nenhum período posterior. Notoriamente, 2020, 2021 e 2022 foram anos atípicos em termos de condições econômicas que afetaram os preços ao consumidor devido à pandemia, à invasão russa na Ucrânia e às eleições presidenciais no Brasil. Para testar o exercício em diferentes circunstâncias, conduzimos um segundo cenário também hipotético que supõe que a atualização da cesta não tivesse ocorrido em janeiro de 2012 e em nenhum período posterior.

A priori não é possível prever quais seriam os impactos de uma mudança na frequência de incorporação das informações da POF.

Existem vários fatores que afetam o consumo de bens e serviços das famílias e o impacto desses fatores pode mudar ao longo do tempo. Se o preço de um bem varia, é preciso considerar o efeito substituição e o efeito renda sobre a elasticidade-preço da demanda. Não só a elasticidade-preço do próprio bem afeta sua demanda, como também a elasticidade cruzada entre bens. Além disso, não existe apenas a dinâmica de preços agindo sobre a demanda de um bem, mas vários outros fatores. Se as famílias estão ficando mais ricas ou mais pobres em média, a elasticidade-renda da demanda é outro fator a ser considerado. Soma-se a isso o fato de que as dinâmicas tecnológicas e econômicas determinam a demanda das famílias e que as preferências culturais e hábitos não são estáticos. Vale lembrar que as famílias podem consumir milhares de bens e serviços diferentes, o que torna desafiador, ao menos com o conhecimento existente hoje, atualizar de forma precisa a ponderação dos itens consumidos sem que haja uma pesquisa de campo que gere dados atualizados.

É importante destacar que os resultados apresentados são específicos ao contexto em que estão inseridos.

Como previsto pela teoria econômica, existem vários fatores que afetam as mudanças na demanda de bens e serviços, e, portanto, não existe um padrão de impacto sobre o índice que pode ser estabelecido *ex-ante* somente com o conhecimento de um adiamento ou edição extra da pesquisa. Os resultados dos exercícios aqui apresentados podem não ser em outro ponto do tempo, o que não diminui as deduções que fazem transparecer. No caso em que o índice de inflação está sendo superestimado, contratos e preços estão sendo sobre-ajustados e recursos podem ser desperdiçados. Por outro lado, caso a inflação esteja sendo subestimada, o trabalhador assalariado que tem seu salário ajustado pelo índice tem o seu poder de compra reduzido em taxa equivalente à da subestimação. O ponto principal é que é necessário ter dados atualizados para minimizar potenciais erros de cálculo.

Dados e metodologia

Dados

Para fazer os cálculos deste estudo, utilizam-se os dados mensais do INPC publicados pelo IBGE.

Especificamente, utilizamos as planilhas de variação de preços e dos pesos de cada categoria que compõe a cesta de consumo do INPC. A fonte dos dados é o Sistema IBGE de Recuperação Automática (SIDRA), que consolida os dados do INPC a nível de Subitem¹. É importante mencionar que os dados do SIDRA contêm os

¹ Existem atualmente duas fontes de dados para o INPC a nível de Subitem: (i) os dados “brutos” do FTP IBGE, que são publicados na forma de planilhas no formato .xls, cada uma relativa a um determinado ano-mês do período entre janeiro de 2008 e maio de 2023; e (ii) os dados consolidados do Sistema IBGE de Recuperação Automática das Tabelas

códigos identificadores das categorias, o que permite a identificação segura das categorias que formam os diferentes níveis de agregação do INPC ao longo do tempo.

A análise é conduzida ao nível mais granular do INPC, o nível de subitem, o que gera desafios específicos para o cálculo dos contrafactuais. O INPC tem quatro níveis de agregação, na seguinte ordem do mais agregado para o menos agregado: (i) Grupo, (ii) Subgrupo, (iii) Item, e (iv) Subitem. É importante ressaltar que dada a forma como a variação do INPC é calculada de um período para o outro podemos computar tanto o INPC factual quanto o contrafactual em qualquer nível de agregação das categorias que formam a cesta de consumo. A principal dificuldade, contudo, associada à realização de exercícios contrafactuais em níveis mais desagregados é que novas categorias podem surgir e outras podem desaparecer ao longo do tempo, o que implica que o número de elementos das cestas comparadas não é necessariamente o mesmo. Além disso, não observamos a variação de preços dos bens que foram retirados da cesta e, portanto, não temos todos os insumos necessários para o cálculo do INPC contrafactual de início. É importante ressaltar que, conforme mostramos no Apêndice A, se os cálculos contrafactuais fossem realizados a níveis mais agregados, perderíamos a variação dentro das categorias e encontraríamos resultados potencialmente enganosos.

Metodologia

Imputação de uma variação de preços para os bens não observados nos cenários contrafactuais

Para contornar a ausência de dados relativos à variação de preços no exercício contrafactual, implementou-se uma imputação de média que preserva os pesos das categorias agregadas de consumo. Uma abordagem simples para lidar com a falta de variação de preços de alguns subitens seria abandoná-los e reponderar a cesta com base nos subitens restantes, ou de forma equivalente, imputar a variação média de preços da categoria correspondente a nível de Item². Nesse caso, contudo, poderíamos produzir um contrafactual errôneo em relação à cesta original. Em especial, se a maioria dos subitens que foram removidos na nova estrutura de ponderação pertencessem a uma categoria específica de item, por exemplo, “Comunicação”, o nosso contrafactual se tornaria a representação de uma cesta em que a categoria Comunicação teria uma importância menor do que realmente teve no passado. Assim, para aumentar a semelhança entre a cesta utilizada para as estimativas contrafactuais e aquela em vigor antes da atualização na estrutura de ponderação, imputamos a variação de preço médio dos subitens remanescentes dentro da mesma categoria a nível de Item aos que subitens que foram removidos (Figura 1). Se todos os elementos no nível de Item estiverem faltando, imputamos a eles a média do nível de Subgrupo e assim por diante. Implicitamente, nossa hipótese é que a variação de preço esperada do subitem é correlacionada com a variação de subitens similares.

O índice de preços

O INPC é um índice que busca se aproximar de um índice de custos de vida (IBGE, 2005, p.13). O índice se baseia em uma cesta de consumo típica que foi observada em algum período de tempo e no qual foram calculados os pesos de cada categoria sobre o custo total da cesta. A partir da observação desses pesos no primeiro período, monitora-se apenas a variação dos preços ao longo do tempo. A variação total do índice de preços pode ser calculada como a média das variações de preço de cada categoria ponderada pelo peso da categoria na cesta de consumo do período base. Este índice é equivalente a um Índice de Laspeyres, no qual as quantidades consumidas no período base são mantidas fixas e o custo da cesta é atualizado pelos novos preços ao longo do tempo.

2938 (relativa ao período entre julho de 2006 e dezembro de 2011), 1419 (relativa ao período entre janeiro de 2012 e dezembro de 2019) e 7060 (relativa ao período entre janeiro de 2020 e o último ano-mês antes de uma atualização do INPC em razão de uma nova POF (i.e., posterior à POF 2017/2018).

² Uma outra abordagem para contornar esta ausência de informações seria estimar o cenário contrafactual ao nível do item. No Apêndice A, contudo, apresentamos exemplos de como isso poderia levar a erros.

Iterativamente, o peso de cada categoria é ajustado ao longo do tempo levando-se em conta apenas a variação de preço da categoria. O peso de uma categoria é atualizado multiplicando o peso base por um fator de correção que considera a razão entre a variação de preço da categoria e a variação de preço da cesta de consumo do índice. O IBGE segue essa dinâmica de estimação da variação do índice e atualização dos pesos até que seja implementada uma nova cesta típica com base em uma nova POF. Desde maio de 2018, o índice é calculado para 16 regiões e o INPC a nível nacional é uma média ponderada dos índices específicos das regiões. O peso regional é determinado pela população da região. Formalmente, o índice pode ser definido da seguinte forma.

Seja i uma determinada categoria, r uma dada região, e t um determinado ano-mês, $p_{i,r,t}$ o preço da categoria i na região r no período t , e $q_{i,r,t}$ a quantidade consumida de i na mesma região e período. O custo total da cesta na região r é, portanto, a soma do dispêndio associado a cada categoria j pertencente à cesta e pode ser definido como $\sum_j p_{j,r,t} q_{j,r,t}$, sendo o peso de cada categoria i na região definido por $w_{i,r,t} = \frac{p_{i,r,t} q_{i,r,t}}{\sum_j p_{j,r,t} q_{j,r,t}}$. Portanto, a variação de preço da categoria i na região r no período t em relação ao período anterior $t - 1$ é dada por $pvar_{i,r,t} = \frac{p_{i,r,t} - p_{i,r,t-1}}{p_{i,r,t-1}}$. Assim, se o índice é chamado *INPC*, e consiste em um número-índice base, pode-se definir a variação do INPC entre t e $t - 1$ como:

$$(1) \quad INPCvar_{r,t} = \frac{\sum_{i \in I_{r,t}} w_{i,r,t-1} * pvar_{i,r,t}}{\sum_{i \in I_{r,t}} w_{i,r,t-1}}$$

onde $INPCvar_{r,t}$ é a variação no INPC na região r , e $I_{r,t}$ é o conjunto de categorias que compõem o índice na região r no período t . A nível nacional, o INPC é calculado da seguinte maneira:

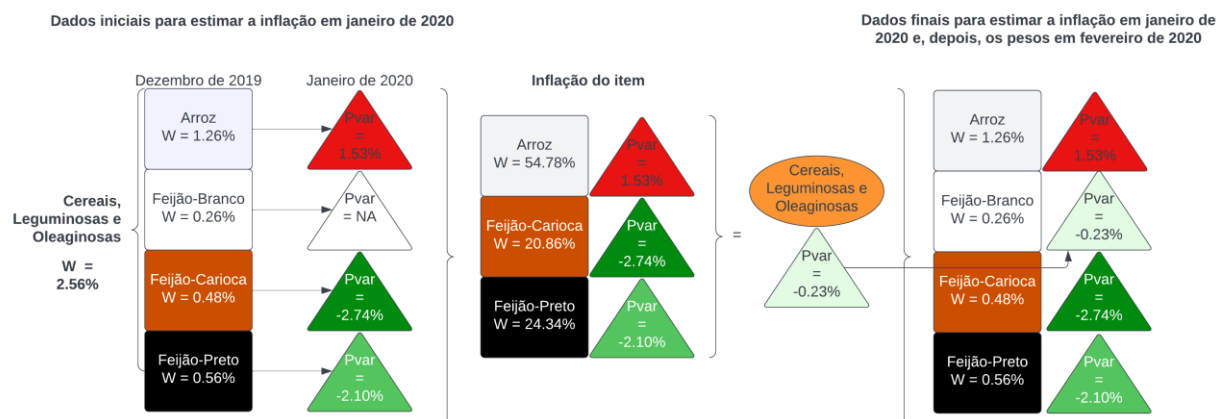
$$(2) \quad INPCvar_{Brasil,t} = \frac{\sum_{r \in R_t} \alpha_{r,t} * INPCvar_{r,t}}{\sum_{r \in R_t} \alpha_{r,t}}$$

onde $INPCvar_{Brasil,t}$ é a variação do INPC a nível Brasil, $\alpha_{r,t}$ é o peso da região r no índice nacional no período t e R_t é o conjunto das regiões que compõem o índice no período. O peso de cada categoria i na região r , por sua vez, é atualizado do período $t - 1$ para o período t pela seguinte fórmula:

$$(3) \quad w_{i,r,t} = \frac{w_{i,r,t-1} * (1 + pvar_{i,r,t})}{(1 + INPCvar_{r,t})}$$

(4)

Figura 1. Imputando a variação de preços nos subitens não mais observados



Fonte: elaboração própria.

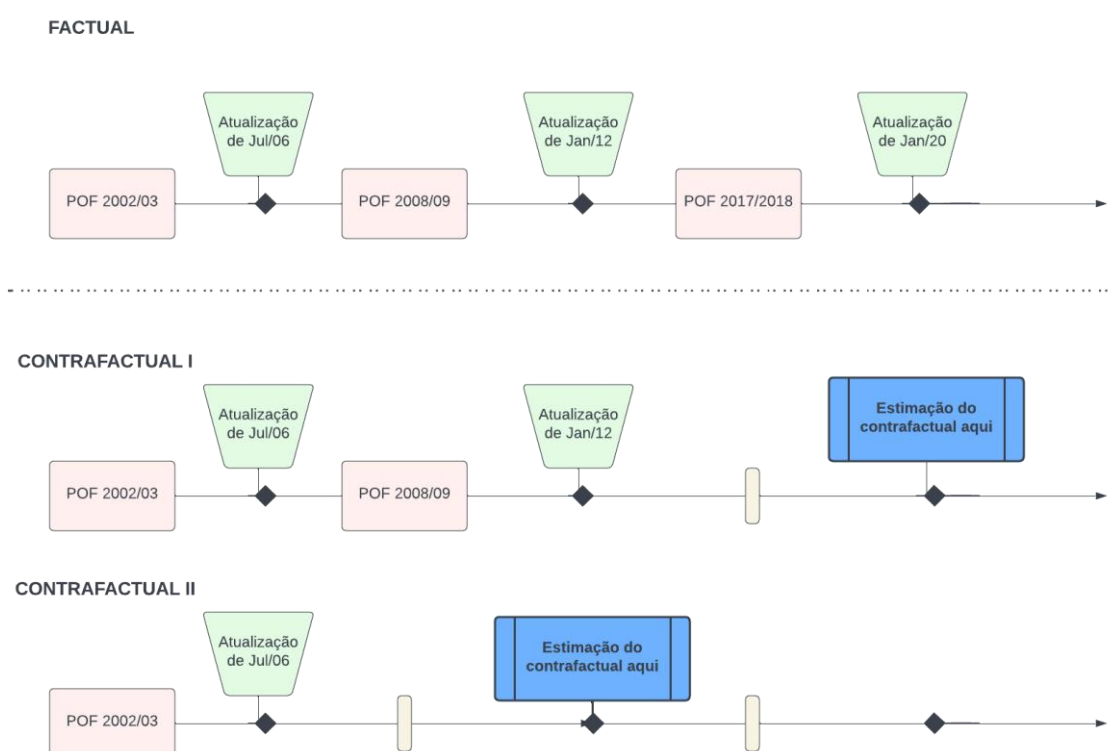
Os Contrafactuais

Nos últimos 20 anos, o IBGE atualizou a estrutura de ponderação do INPC três vezes. A primeira em julho de 2006 em decorrência da POF 2002/2003 (IBGE, Nota Técnica 01/2006), a segunda em janeiro de 2012 em decorrência da POF 2008/2009 (IBGE, 2012), e a terceira em janeiro de 2020 após a POF 2017/2018 (IBGE, 2020). No intervalo entre as atualizações, os pesos são atualizados conforme a equação (3) acima.

Esta nota apresenta dois exercícios contrafactuais cujos pressupostos são ilustrados abaixo (Figura 2). No primeiro, a situação hipotética é que a atualização da estrutura de ponderação de janeiro de 2020 não ocorreu. Ou seja, ao invés de se empregar os pesos publicados, utiliza-se a equação (3) para atualizar os pesos entre os dados mensais publicados em dezembro de 2019 e janeiro de 2020. Em seguida, recalcula-se a variação nacional do INPC em janeiro de 2020 a partir da equação (2). Iterativamente, é possível reestimar toda a série do índice. O segundo contrafactual parte do mesmo método descrito anteriormente, porém para o mês de janeiro de 2012. Ao invés de se estimar a inflação com os pesos publicados naquele mês, calculam-se os pesos contrafactuais a partir dos dados mensais publicados em dezembro de 2011. A partir daí, uma nova série de inflação é estimada como se nenhuma atualização da estrutura de ponderadores tivesse ocorrido desde então.

Para levar em consideração as mudanças na cobertura geográfica do INPC ao longo do tempo, estimamos duas versões do primeiro exercício. Quando o IBGE implementou a atualização dos pesos em janeiro de 2012, coletavam-se dados de preços em 11 regiões para calcular o INPC nacional. Contudo, antes da atualização seguinte da cesta de consumo, o IBGE adicionou cinco regiões ao indicador nacional. Assim, para verificar a robustez dos resultados, construímos um cenário contrafactual que mantém os pesos regionais originais de janeiro de 2012 constantes ao longo do tempo, e outro que pondera a inflação contrafactual nas 16 regiões consideradas no índice desde maio de 2018. Não há diferenças significativas entre as duas abordagens, conforme mostrado no Apêndice B. Para simplificação, apresentamos nos resultados a seguir apenas a versão que considera o conjunto completo de regiões.

Figura 2. A POF, a atualização da estrutura de ponderadores do INPC ao longo do tempo e os desenhos contrafactuais



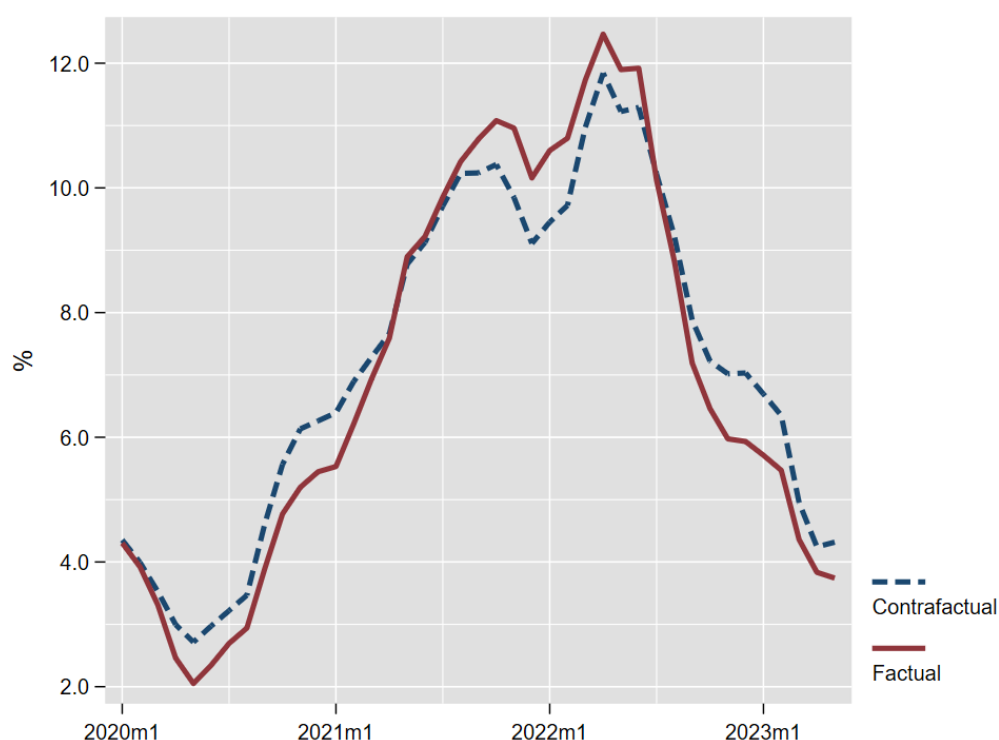
Fonte: elaboração própria.

Resultados

Exercício contrafactual 1: análise de curto prazo (sem atualização recente)

A não atualização dos pesos do INPC em 2020 teria levado a uma taxa de inflação anual mais elevada nos primeiros dezesseis meses, mas menor entre maio de 2021 e junho de 2022.³ A Figura 3 mostra a evolução da variação do INPC em 12 meses comparando o real com o contrafactual no período entre janeiro de 2020 e maio de 2023. Observa-se que a taxa contrafactual permaneceu superior à real entre janeiro de 2020 e abril de 2021, mas depois a situação se inverte até junho de 2022. Este resultado tem uma consequência importante para o debate sobre a periodicidade da POF: a taxa de inflação contrafactual nem sempre é maior ou menor que a observada.

Figura 3: Taxas de inflação anual factual e contrafactual entre janeiro de 2020 e maio de 2023 - Exercício 1



Fonte: elaboração própria a partir dos dados do SIDRA IBGE.

Caso não tivesse ocorrido a atualização da estrutura de ponderação do INPC em 2020 com base na POF 2017/2018, a taxa de inflação medida ao final daquele ano teria sido 0,8 ponto percentual maior. A Tabela 1 apresenta a taxa de inflação anual observada e a contrafactual em dezembro de cada ano entre 2020 e 2022. Como se pode observar, ao final do primeiro ano (dezembro de 2020), a diferença entre o INPC contrafactual e o factual ficou em torno de 0,8 pontos percentuais. Este resultado mostra que mesmo num período relativamente curto de um ano, as diferenças entre o valor observado e o contrafactual podem ser significativas.

³ A taxa de inflação anual é a variação do INPC acumulada nos últimos 12 meses.

Tabela 1: Taxas de inflação entre 2019 e 2022 – Exercício 1

Ano-mês	Factual (%)	Contrafactual (%)	Diferença (p.p.)
2019m12	4.48	4.48	0.00
2020m12	5.45	6.26	0.82
2021m12	10.16	9.11	-1.05
2022m12	5.93	7.03	1.10

Fonte: elaboração própria a partir dos dados do SIDRA IBGE.

Uma análise detalhada sugere que o resultado líquido entre a inflação real e a contrafactual depende da combinação entre a diferença de pesos e a variação de preços observada para cada categoria em cada mês específico. A Tabela 2 mostra a comparação do INPC factual com o contrafactual em cada mês de 2022, destacando aquele em que vemos a maior diferença entre os índices, e quando a taxa de inflação contrafactual anual se torna maior que a factual. Esta mudança de rumo foi no mês em que a isenção de impostos sobre combustíveis e serviços de energia foi implantada. Como a parcela dos gastos com gasolina e outros bens isentos é menor na cesta desatualizada do que na atualizada, o cenário contrafactual subestima a queda dos preços nesse período. No Apêndice C apresentamos uma análise mais detalhada das diferenças relativas a este período.

Tabela 2: INPC factual e contrafactual entre janeiro e dezembro de 2022 – Exercício 1

Ano-mês	Variação mensal INPC			Variação acumulada INPC 12 meses		
	Factual (%)	Contrafactual (%)	Dif. (p.p)	Factual (%)	Contrafactual (%)	Dif. (p.p)
2022m1	0.67	0.67	0.00	10.60	9.45	-1.15
2022m2	1	0.91	-0.09	10.80	9.72	-1.08
2022m3	1.71	1.83	0.12	11.73	10.97	-0.76
2022m4	1.04	1.22	0.18	12.47	11.84	-0.62
2022m5	0.45	0.36	-0.09	11.90	11.23	-0.67
2022m6	0.62	0.64	0.02	11.92	11.29	-0.63
2022m7	-0.6	-0.09	0.51	10.12	10.22	0.10
2022m8	-0.31	-0.07	0.24	8.83	9.23	0.40
2022m9	-0.32	-0.19	0.13	7.19	7.88	0.69
2022m10	0.47	0.46	-0.01	6.46	7.22	0.76
2022m11	0.38	0.41	0.03	5.97	7.02	1.04
2022m12	0.69	0.68	-0.01	5.93	7.03	1.10

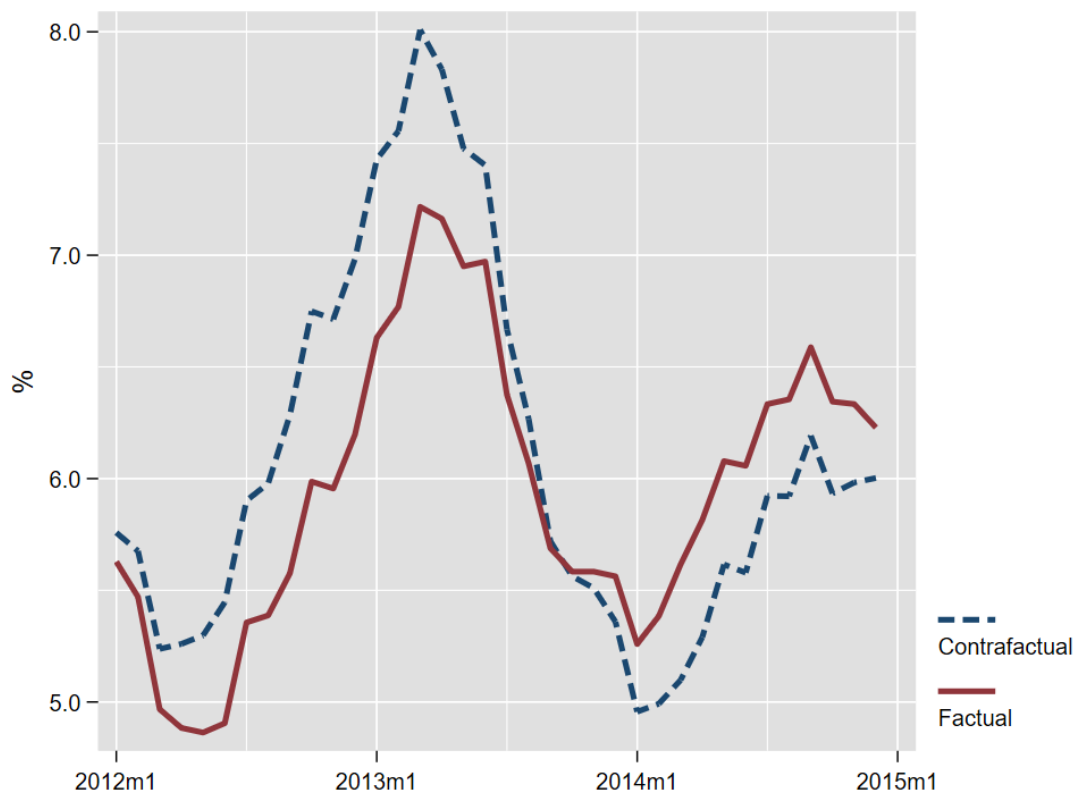
Fonte: elaboração própria a partir dos dados do SIDRA IBGE.

Exercício contrafactual 2: análise histórica

No cenário hipotético em que o país tivesse adiado a POF na década anterior, a inflação mensurada também teria sido significativamente diferente. A Figura 4 apresenta a comparação entre a taxa de inflação anual observada com a contrafactual no período entre janeiro de 2012 e dezembro de 2014 considerando nosso segundo exercício. Tal como no primeiro, a inflação contrafactual em 12 meses é superior no primeiro período de análise, mas tal relação se inverte no último trimestre de 2013. A diferença é de 0,8 pontos percentuais na taxa de inflação de 2012 (Tabela 3), semelhante em magnitude ao primeiro ano de análise no Exercício 1. Conforme discutido anteriormente, os resultados são determinados pela combinação entre a diferença de pesos e a variação de preços associada aos bens que perderam ou ganharam importância. Ou seja, nada garante 0,8 p.p. de diferença em cada primeiro ano do contrafactual. No entanto, este segundo exercício

reforça que o atraso na atualização da estrutura de ponderação do INPC devido à ausência de uma POF tem consequências importantes nas taxas de inflação mensuradas.

Figura 4: Taxas de inflação anual factual e contrafactual entre janeiro de 2012 e dezembro de 2014 - Exercício 2



Fonte: elaboração própria a partir dos dados do SIDRA IBGE.

Tabela 3: Taxas de inflação entre 2011 e 2014 – Exercício 2

Ano-mês	Factual (%)	Contrafactual (%)	Diferença (p.p.)
2011m12	6.08	6.08	0.00
2012m12	6.20	6.98	0.79
2013m12	5.56	5.36	-0.20
2014m12	6.23	6.00	-0.23

Fonte: elaboração própria a partir dos dados do SIDRA IBGE.

Consequências orçamentárias da não realização frequente da POF

As consequências da não realização frequente da POF sobre a inflação oficial medida pelos índices de preços ao consumidor podem ser de enorme impacto para o orçamento público. Notadamente, muitas rubricas do orçamento são diretamente afetadas pela inflação, assim como a própria política monetária através da definição da taxa básica de juros no Brasil, a SELIC. Isso implica que os impactos sobre a economia brasileira podem ser substanciais. Por exemplo, a parcela das aposentadorias e salários dos servidores representava, respectivamente, 30% e 13% das despesas primárias do governo central do Brasil em 2022 (STN, 2023). Estas rubricas são afetadas diretamente pelo reajuste do salário mínimo

derivado da variação do INPC. Além disso, alguns títulos da dívida pública do governo federal são indexados pelo IPCA, o que implica que o próprio custo da dívida também é impactado pelo cálculo oficial da inflação.

Caso a atualização da estrutura de ponderação do INPC não tivesse ocorrido em janeiro de 2020, as despesas do governo federal com benefícios previdenciários em 2021 teriam sido R\$ 5,45 bilhões maiores. A fim de mostrar uma estimativa dos custos econômicos da não realização da POF sobre o cálculo do INPC, apresentamos os efeitos sobre a atualização monetária da despesa do governo federal com benefícios previdenciários. Segundo dados do Resultado Primário do Governo Central, estas despesas no Brasil em 2021 foram da ordem de R\$ 709,5 bilhões. Por sua vez, os resultados do nosso primeiro exercício contrafactual indicam que a taxa de inflação anual factual foi de 5,45% em dezembro de 2020, enquanto a contrafactual foi de 6,26% (veja a Tabela 1) – uma diferença de 0,81 pontos percentuais. Portanto, se considerarmos que as despesas previdenciárias de 2021 tivessem sido atualizadas pelo INPC acumulado em 12 meses contrafactual de dezembro de 2020, o valor passaria de R\$ 709,5 bilhões para R\$ 715,0 bilhões.

Como se observa, no caso em que a inflação está sendo sobreestimada, o governo está perdendo recursos que poderiam ser usados para financiar outros programas governamentais. Por exemplo, a economia com a atualização da estrutura de ponderação do INPC em 2020 representa 8,8% do custo do Programa de Prestação Continuada (BPC) em 2021 (Tabela 4). A estimativa também representa 9,6% do custo do programa Auxílio Emergencial em 2021, e 17,5% do custo do principal programa de transferência de renda, o Bolsa Família, em 2019. Tais resultados demonstram que a diferença entre a atualização das despesas previdenciárias pelo INPC acumulado em 12 meses contrafactual e o factual é substancial mesmo em apenas um ano.

Tabela 4: Comparação da estimativa como proporção do custo de alguns programas sociais

Programa	Custo (em bilhões de reais)	Diferença como proporção do custo dos programas
BPC (2021)	61.8	8.8%
Auxílio Emergencial (2021)	56.8	9.6%
Bolsa Família (2019)	31.1	17.5%
Seguro-desemprego (2021)	30.8	17.7%
Bolsa Família (2021)	27.2	20.0%
Auxílio Brasil (2021)	6.4	85.2%

Fonte: elaboração própria a partir dos dados do Portal de Dados Abertos do Governo Federal.

Conclusões

Esta nota apresenta o resultado de dois exercícios contrafactuais que simulam situações hipotéticas nas quais é possível observar a variação do INPC caso a POF não tivesse sido realizada com determinada periodicidade. O primeiro exercício simula uma situação em que a cesta de consumo e os ponderadores do INPC não foram atualizados em janeiro de 2020 e em nenhum período posterior, enquanto o segundo supõe que estes não foram atualizados em janeiro de 2012 e em nenhum período subsequente. Os resultados mostram que o INPC contrafactual pode ser significativamente diferente do factual nos dois cenários hipotéticos, especialmente se preços específicos sofrerem choque, como no caso das isenções fiscais para combustíveis em 2022.

Esses resultados sugerem que a não atualização frequente da cesta de consumo e dos ponderadores do INPC no Brasil pode ter grandes consequências econômicas, não apenas sobre o bem-estar das famílias, mas também sobre as despesas públicas. Atualmente, a atualização da estrutura de ponderação do INPC é feita somente quando é realizada uma nova POF, de forma que a não realização frequente desta implica a potencial desatualização da estrutura do INPC. Portanto, a evidência destaca a importância da realização frequente da POF em razão de sua vinculação aos índices de preço oficiais⁴.

É preciso ressaltar que os resultados são contexto-específicos e refletem mudanças particulares nos padrões de consumo no Brasil durante o período de estudo, isto é, janeiro de 2020 a maio de 2023 no primeiro exercício e janeiro de 2012 a dezembro de 2014 no segundo. Uma destas mudanças foi a eclosão da pandemia de COVID-19 e o período de alta inflação no país, que impacta diretamente as estimativas do nosso primeiro exercício contrafactual. Não há, dessa forma, garantia de que esses resultados serão os mesmos no futuro, o que reforça a necessidade de dados frequentes e representativos do padrão de consumo das famílias ao longo do tempo.

O caso dos índices de preço no Brasil é ilustrativo da realidade de potencialmente vários países nos quais as pesquisas domiciliares de orçamento são a principal fonte de dados para a atualização da cesta de consumo dos índices de preço ao consumidor e nos quais elas são realizadas com baixa periodicidade. Nossos resultados mostram que diante do custo estimado de tais pesquisas, os custos econômicos podem ser substantivamente maiores na ausência de dados atualizados sobre os padrões de consumo das populações.

⁴ Poderia se argumentar que a estrutura de ponderação do INPC pode ser atualizada usando dados mais frequentes das Contas Nacionais. Contudo, as estimativas de despesas das Contas Nacionais também se baseiam parcialmente em estimações da POF, dependendo também da atualização frequente desta (ONU, 2009, parágrafo 4.40).

Referências

- Hill, Peter (2010). Lowe indices. In: 2008 World Congress on National Accounts and Economic Performance Measures for Nations. Washington, DC.
- IBGE (2005). Sistema Nacional de Índices de Preços ao Consumidor: estruturas de ponderação a partir da Pesquisa de Orçamentos Familiares 2002-2003. Séries Relatórios Metodológicos v.34, Rio de Janeiro.
- IBGE (2020). Sistema Nacional de Índices de Preços ao Consumidor: métodos de cálculo. Séries Relatórios Metodológicos v.14, 8ª edição, Rio de Janeiro.
- IMF, ILO, OECD, EU, UN, and WB. (2020). Consumer price index manual: concepts and methods. (International Monetary Fund, International Labour Office, Organisation for Economic Co-operation and Development, European Union, United Nations, World Bank). Washington, DC : International Monetary Fund
- STN (2023). Estatísticas Fiscais do Governo Central Orçamentário. Secretaria do Tesouro Nacional. Brasília-DF, Brasil.
- ONU (2009). Practical Guide to Producing Consumer Price Indices. United Nations, Geneva, Switzerland.

Apêndice A. Como a análise a níveis mais agregados pode ser enganosa

O IBGE disponibiliza dados da variação de preços de bens de consumo para diferentes níveis de agregação. O IBGE publica dados no nível mais desagregado (subitem), mas também em outros níveis que agregam subitens por similaridade (itens, subgrupos e grupos). Por exemplo, o arroz é um subitem, que está agrupado no item “cereais, leguminosas e oleaginosas”, que está agregado no subgrupo “alimentação no domicílio”, que por sua vez pertence ao grupo mais agregado “Alimentação e Bebidas”. Quando as cestas de consumo dos índices de preços ao consumidor são atualizadas, vários subitens podem sair da cesta enquanto outros podem ingressar. Estas substituições refletem a mudança dos hábitos de consumo ao longo do tempo, a qual pode ser explicada pelas mudanças tecnológicas e pelas dinâmicas culturais. Vale ressaltar, porém, que as categorias de item, subgrupo e grupo raramente mudam.

Uma alternativa para realizar os exercícios contrafactuais propostos nesta nota seria trabalhar com níveis mais agregados do INPC. Conforme mencionado no corpo do texto, uma dificuldade nos cálculos contrafactuais é o fato de não observarmos a variação de preços dos subitens que foram retirados da cesta do INPC. Uma abordagem mais simples seria trabalhar com os dados de variação de preços e pesos ao nível do item, ignorando as mudanças relativas que acontecem dentro desta categoria.

A utilização de dados mais agregados (por exemplo, itens), contudo, tem consequências importantes para nossos resultados. Um dos motivos, por exemplo, é que não capturamos os efeitos de substituição entre categorias pertencentes ao mesmo nível de agregação (item), como é o caso do feijão preto e do feijão marrom (*feijão-carioca*). Assim, caso o feijão marrom tivesse uma participação mais elevada nas despesas das famílias e por alguma razão o seu preço aumentasse relativamente mais do que o do feijão preto (por exemplo, devido a uma praga específica que afeta apenas o *feijão-carioca*), subestimaríamos os aumentos de preços no cenário contrafactual porque a variação de preços disponibilizada ao nível do item já leva em consideração a ponderação factual dos subitens.

A utilização de níveis mais agregados induz a análise contrafactual ao erro porque não considera o rebalanceamento dos subitens pela variação média do preço. Fornecemos um exemplo numérico e uma ilustração desse fato utilizando as categorias a seguir (Figura A.1). Antes de janeiro de 2020, o arroz representava 0,72% das despesas do consumidor típico de São Paulo, enquanto essa participação era de 0,49% para o *feijão-carioca*. Isto significa que o arroz teve 60% do total de 1,21% gasto no item relativo aos cereais (“cereais, leguminosas e oleaginosas”), enquanto o feijão marrom ficou com 40%. Após a atualização da estrutura de ponderação, as participações do arroz e do feijão marrom passaram a ser de 68% e 32% nos cereais, respectivamente, enquanto os cereais obtiveram uma parcela inferior das despesas totais (0,91%). O arroz teve variação de preço maior que o feijão marrom no período recente. Como consequência, na análise com dados ao nível do subitem, os cereais têm uma participação mais elevada nas despesas (1,21% vs. 0,91%), mas a sua variação de preços é suavizada por uma participação relativamente menor do arroz no nível do item (60% vs. 68%). Contudo, se tivéssemos trabalhado com dados ao nível de item, obteríamos a mesma percentagem mais elevada de cereais sem compensar a variação de preços.

De modo geral, teríamos superestimado o INPC contrafactual se tivéssemos trabalhado com níveis mais agregados. A Figura A.2 mostra a trajetória das taxas de inflação factual a nível de item e contrafactuais a nível de item, subgrupo e grupo.

Figura A.1. Como a agregação ao nível de item pode levar a uma variação contrafactual errônea

a. Dados factuais e contrafactuais para “Cereais, leguminosas e oleaginosas” em diferentes níveis de agregação do INPC

Factual			
Subitem	W	var_p	
1101002.Arroz	0.62		
		2.62	
1101051.Feijão - mulatinho	-	-	
1101052.Feijão - preto	-	-	
1101053.Feijão - macáçar (fradinho)	-	-	
1101073.Feijão - carioca (rajado)	0.29	-1.4	

Item	W	var_p
1101.Cereais, leguminosas e oleaginosas	0.91	1.33

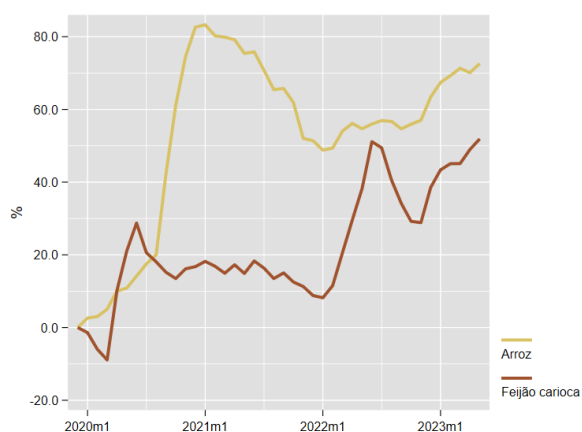
Dados contrafactuais a nível de item			
Item	W	var_p	
1101.Cereais, leguminosas e oleaginosas	1.21	1.33	

Item	W	var_p
1101.Cereais, leguminosas e oleaginosas	1.21	1.33

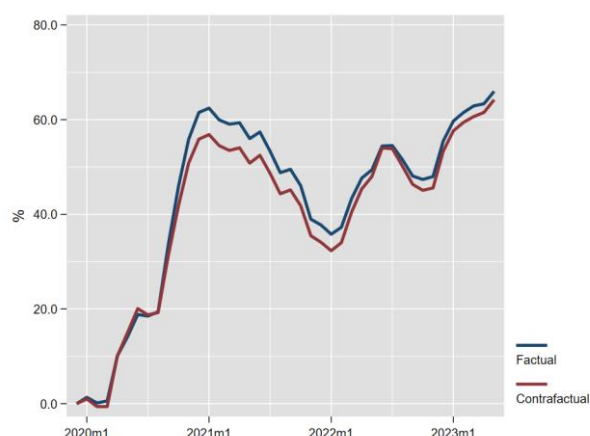
Dados contrafactuais a nível de subitem			
Subitem	W	var_p	
1101002.Arroz	0.72	2.62	
1101051.Feijão - mulatinho	-	-	
1101052.Feijão - preto	-	-	
1101053.Feijão - macáçar (fradinho)	-	-	
1101073.Feijão - carioca (rajado)	0.49	-1.4	
1101075.Feijão - branco	-	-	
1101079. Milho em grão	-	-	

Item	W	var_p
1101.Cereais, leguminosas e oleaginosas	1.21	0.99

b. Variação do preço do arroz e do feijão carioca

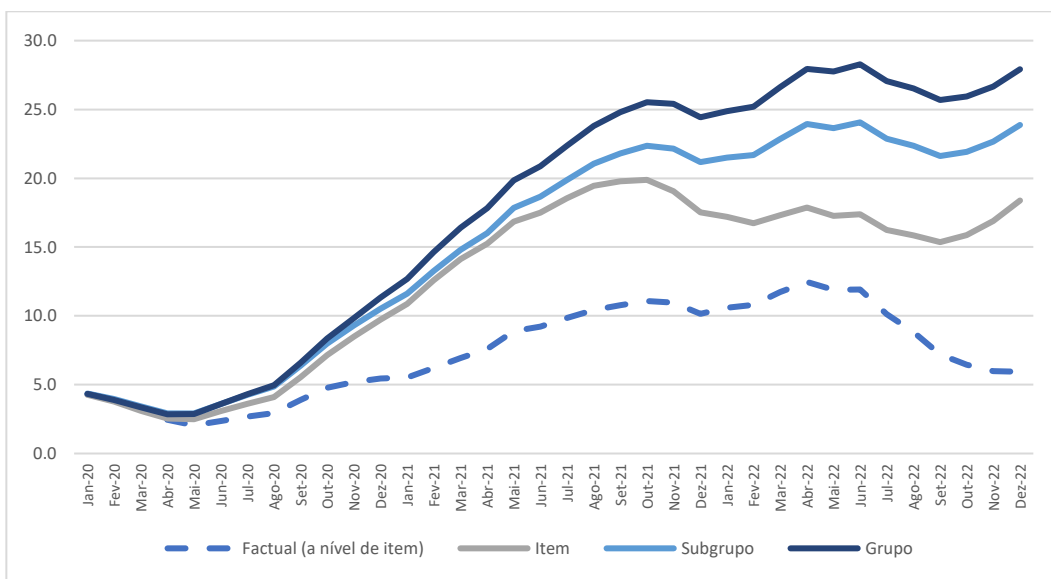


c. Variação acumulada de preço dos cereais: factual versus contrafactual a nível de item



Fonte: elaboração própria a partir dos dados do SIDRA IBGE.

Figura A.2. Comparação do INPC acumulado em 12 meses factual a nível de item com os contrafactuais a nível de Item, Subgrupo e Grupo – Exercício 1



Fonte: elaboração própria a partir dos dados do SIDRA IBGE.

Apêndice B. Diferentes ponderações regionais para o INPC nacional

O número de regiões consideradas no cálculo do INPC a nível nacional mudou ao longo do tempo. Em 1979, quando o índice começou a ser calculado, os preços eram coletados apenas no Rio de Janeiro-RJ, Porto Alegre-RS, Belo Horizonte-MG e Recife-PE. Em 1980, foram incluídos São Paulo-SP, Brasília-DF e Belém-PA. Além disso, Campo Grande-MS e Vitória-ES foram adicionados em 2014, enquanto Rio Branco-AC, São Luís-MA e Aracaju-SE foram adicionados em maio de 2018.

Isto implica que, teoricamente, poderíamos considerar ou a situação contrafactual em que nada mudou desde a atualização anterior da estrutura de ponderação (regiões e cesta de consumo), ou aquela em que novas regiões foram introduzidas, mas a cesta de consumo de referência permaneceu a mesma. A fim de testar a sensibilidade do INPC contrafactual a diferentes especificações de pesos regionais, estimamos o INPC contrafactual a nível nacional calculando a média das estimativas regionais em ambas as situações contrafactuais consideradas acima (Tabela B.1).

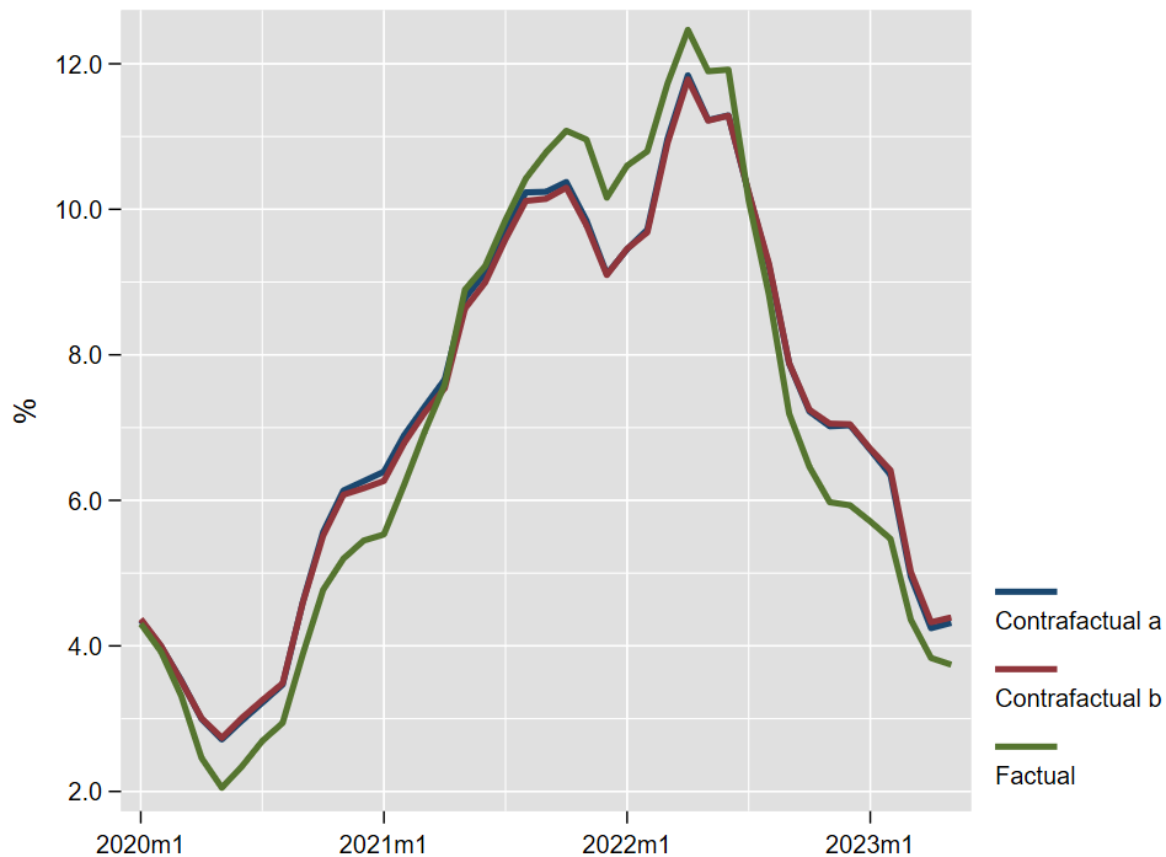
Como pode ser visto, as diferenças entre os dois cenários hipotéticos não são substantivas. Em particular, as regiões que foram introduzidas ao longo do tempo têm uma participação menor no índice nacional, não contribuindo assim para alterações significativas nas estimativas contrafactuais (Figura B.1). Conforme explicado no texto principal, os pesos regionais são determinados pelas populações das regiões. O mesmo foi replicado para o segundo exercício proposto, estimando a inflação contrafactual utilizando i) a estrutura de ponderação original, e ii) considerando a introdução de novas regiões em 2014, e as diferenças se mostraram pequenas.

Table B.1. Diferentes estruturas de ponderação regional para o INPC contrafactual a nível nacional

a. 16 regiões		b. 11 regiões	
Região	Peso no INPC (%)	Região	Peso no INPC (%)
Brasil	100	Brasil	100
Rio de Janeiro	9.38	Rio de Janeiro	9.91
Porto Alegre	7.15	Porto Alegre	7.38
Belo Horizonte	10.35	Belo Horizonte	11.04
Recife	5.6	Recife	7.17
São Paulo	24.6	São Paulo	25.24
Brasília	1.97	Brasília	2.39
Belém	6.95	Belém	7.03
Fortaleza	5.16	Fortaleza	6.61
Salvador	7.92	Salvador	10.67
Curitiba	7.37	Curitiba	7.29
Goiânia	4.43	Goiânia	5.27
Vitória	1.91		
Campo Grande	1.73		
Rio Branco	0.72		
São Luís	3.47		
Aracaju	1.29		

Source: IBGE (2020).

Figura B.1. Comparação entre a inflação anual contrafactual e factual entre janeiro de 2020 e maio de 2023 - Exercício 1

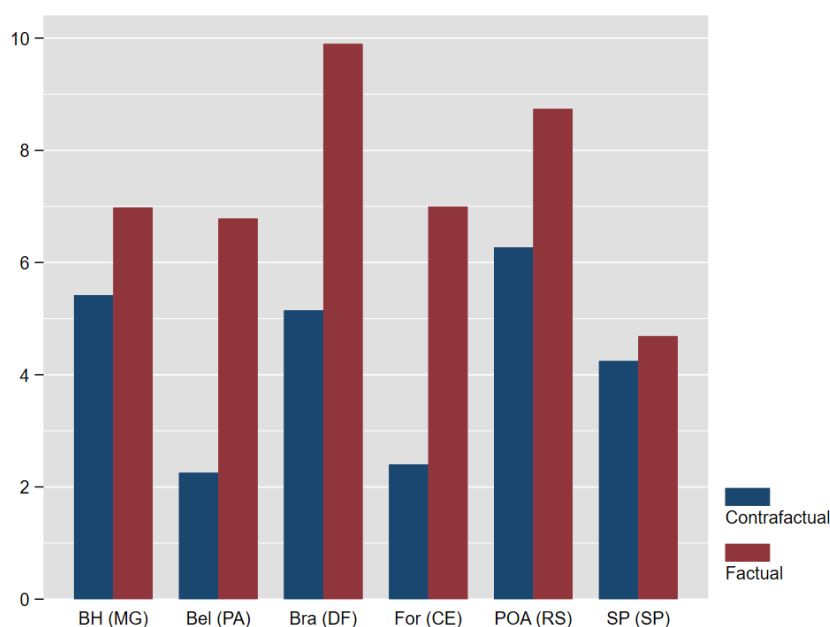


Fonte: elaboração própria a partir dos dados do SIDRA IBGE. Nota: “Contrafactual a”=primeira forma de se calcular a inflação contrafactual descrita neste apêndice, “Contrafactual b”= segunda forma de se calcular a inflação contrafactual descrita neste apêndice. A inflação de 12 meses no ano-mês t é calculada como a inflação acumulada no ano-mês t e nos onze anos-meses anteriores a t.

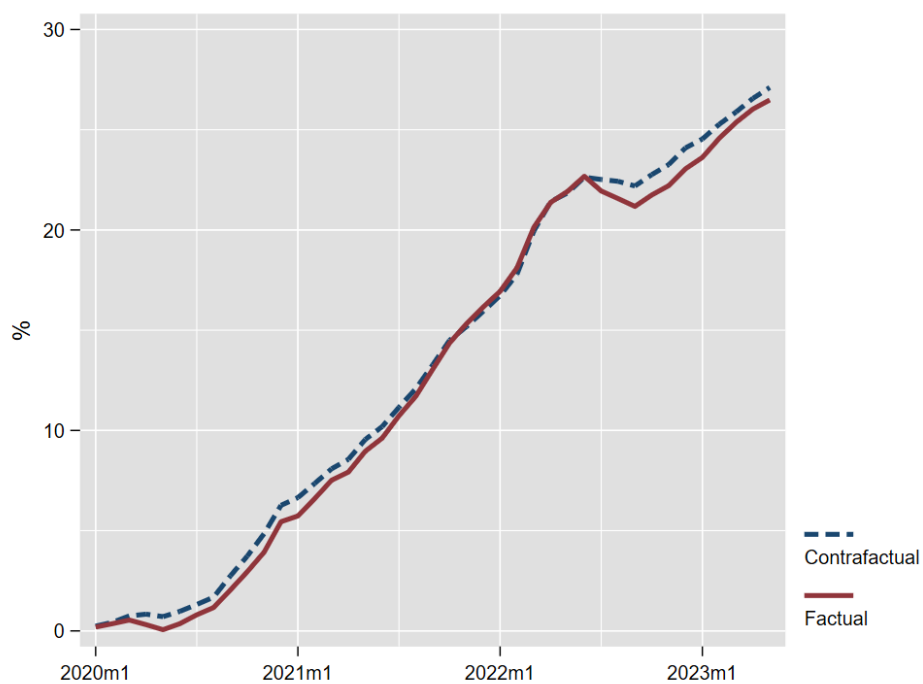
Apêndice C. Estudo de caso da isenção fiscal da gasolina em 2022

A isenção fiscal para a gasolina e alguns serviços selecionados aprovada em 2022 esclarece as diferenças entre as estruturas de ponderação factuais e contrafactuais. Ao comparar o contrafactual com os dados registrados, observamos que a maior diferença se materializa em julho de 2022. Não por acaso, este é o mês logo após a aprovação da Lei Complementar nº 194/2022, que isentou impostos sobre combustíveis, energia elétrica, serviços de telecomunicação e transporte público. A isenção resultou em uma redução de aproximadamente 15% nos preços da gasolina em todas as regiões onde são observados os preços do INPC. É importante ressaltar que a gasolina tem atualmente a maior participação na cesta do INPC na maioria das regiões. Contudo, antes da última atualização da estrutura de ponderação, tinha uma participação menor (Figura C.1). Assim, não sem razão, julho de 2022 é o ponto de inflexão em que a inflação contrafactual acumulada desde o início dos cálculos se torna irreversivelmente superior à observada (Figura C.2).

Figura C.1. Participação da gasolina nas cestas de consumo de referência do INPC (Contrafactual vs. Factual)



Fonte: elaboração própria a partir dos dados do SIDRA IBGE. Notas: BH (MG)=Belo Horizonte-MG, Bel (PA)=Belém-PA, Bra (DF)=Brasília-DF, For (CE)=Fortaleza-CE, PoA=Porto Alegre-RS, and SP (SP)=São Paulo-SP. As barras azuis representam as participações contrafactuais, enquanto as vermelhas representam as factuais. Algumas regiões selecionadas para simplificação.

Figure C.2. Comparação entre a inflação acumulada factual e a contrafactual desde janeiro de 2020

Fonte: elaboração própria a partir dos dados do SIDRA IBGE. Nota: Variação de preço acumulada (%), calculada desde o início do exercício contrafactual 1 em janeiro de 2020.