

Hudson Chaves Costa

Três Ensaaios em Comportamento dos
Preços na Economia Brasileira

Porto Alegre
2014

Hudson Chaves Costa

Três Ensaaios em Comportamento dos Preços na Economia Brasileira

Projeto de Tese apresentado ao Programa
de Pós-Graduação em Economia da Uni-
versidade Federal do Rio Grande do Sul na
Área de Economia Aplicada.

Orientador: Prof. Dr. Sabino Porto da Silva
Júnior

Porto Alegre
2014

Sumário

1	INTRODUÇÃO	1
2	REVISÃO BIBLIOGRÁFICA	5
3	METODOLOGIA	9
	Referências Bibliográficas	16
A	METODOLOGIA IBGE	18

Capítulo 1

INTRODUÇÃO

Sabe-se que a Teoria Econômica ortodoxa divide-se em dois grandes grupos: microeconomia e macroeconomia. Para [Bresser-Pereira \(1968\)](#), a primeira está focada na análise de funcionamento geral da economia por meio do comportamento dos agentes econômicos individuais (consumidores e produtores). Já a segunda, realiza essa mesma análise, partindo do estudo de agregados econômicos, como a renda, o consumo, e o investimento. Segundo [Ball and Mankiw \(1994\)](#) existem dois tipos de macroeconomistas: os que acreditam que a rigidez de preços tem um papel importante nas flutuações econômicas de curto prazo e os que atribuem a choques tecnológico tais flutuações.

A partir de Keynes, a teoria macroeconômica evoluiu em diversos caminhos. Em 1937 o modelo Hicks-Hansen, considerado uma das principais formulações matemáticas da Teoria Geral do Emprego, dos Juros e da Moeda, atribuiu importante valor à doutrina da preferência pela liquidez e sua interpretação dos efeitos do investimento sobre a renda (multiplicador) e dos efeitos da renda sobre o investimento (acelerador). Contudo, os keynesianos de Cambridge retomaram as ideias neoclássicas por meio de [Friedman and Kuttner \(1992\)](#) que introduziram a versão aceleracionista da curva de Phillips (taxa natural de desemprego e expectativas adaptativas). Além disso, o monetarismo defende que a taxa natural de desemprego é determinada pelas forças de oferta.

Não obstante, as críticas de [Lucas \(1972\)](#) que deram origem à macroeconomia novo-clássica, são consideradas como importantes para o pouco sucesso do monetarismo. Porém, devido às deficiências internas e pela falta de aderência dos modelos à realidade empírica, críticas surgiram sobre a macroeconomia novo-clássica ([McCallum \(1998\)](#)).

Nos anos 1980, vários autores neoclássicos questionaram fortemente a teoria Novo-Clássica e elaboraram um retorno ao keynesianismo neoclássico, com a teoria Novo-Keynesiana ([Blinder and Stiglitz \(1983\)](#)). Segundo [Dathe \(2000\)](#) em nível microeconômico, a teoria Novo-Keynesiana adota os fundamentos neoclássicos de agentes maximizadores fazendo o melhor uso das informações disponíveis. Em relação aos microfundamentos da macroeconomia, destacam-se as falhas ou barreiras de mercado. Dessa forma, constata-se a existência de falhas de coordenação de forma que decisões individuais podem gerar externalidades macroeconômicas. Por outro lado, no contexto econômico existe concorrência imperfeita, havendo agentes formadores, e não simples-

mente tomadores, de preços e salários. Então, para os agentes econômicos, existe rigidez de preços e salários originada por assimetrias de mercado, heterogeneidade de bens e fatores e informações imperfeitas. A competição imperfeita, por outro lado, cria externalidades de demanda agregada que levam a que os custos sociais da rigidez de preços sejam maiores que os custos privados.

Como é possível observar, há décadas diversas pesquisas teóricas têm focado na fundamentação microeconômica da rigidez de preços, um elemento chave nas explicações dos efeitos reais da política monetária. Por outro lado, a literatura empírica sobre rigidez de preços é pouco encontrada. [Bils and Klenow \(2004\)](#) fizeram uma importante contribuição ao estudar de forma desagregada dados do *Consumer Price Index* (CPI) dos EUA a partir de 1990 e mostrar que o preço médio mudou uma vez a cada 4,3 meses. Embora outras significantes contribuições seguiram este trabalho, importantes questões empíricas permanecem em grande parte sem resposta.

[Cavallo \(2010\)](#) fez questionamentos sobre os fundamentos da rigidez de preços. Entre eles, se as decisões de preço são temporalmente dependentes ou relacionada ao estado econômico subjacente, se a rigidez de preços é realmente conduzida por custo de menu e assimetria de informação, o papel da competição e sincronização dos preços e como o ambiente econômico, experiências passadas de inflação e quadros institucionais influenciam na forma como os preços se ajustam.

A principal restrição para pesquisa empírica é que os dados em nível de produto são limitados em termos de frequência, países e contextos em que eles são coletados. Dados do CPI dos EUA e Europa têm se tornado viáveis para pesquisadores em uma base limitada. Embora esses bancos de dados cubram uma vasta gama de produtos, eles são tipicamente viáveis para países desenvolvidos com ambiente macroeconômico estável, onde choques agregados são leves e os mecanismos relevantes em nível micro com objetivos macro são difíceis de identificar.

Por fim, dados tradicionais de pesquisas de instituições públicas e privadas não possuem tais características e assim, são limitantes para pesquisas empíricas sobre o impacto de políticas monetárias e rigidez de preços.

O presente projeto de tese propõe a utilização da tecnologia de *web scraping* assim como [Cavallo \(2010\)](#) para coletar preços de sites de diversos setores da economia brasileira. Assim, pretende-se contribuir para a pesquisa empírica no mercado brasileiro no que tange à avaliação empírica da rigidez de preços e também gerar um índice de inflação que possa ser considerado proxy para os índices divulgados pelo governo.

JUSTIFICATIVA

Os preços são fundamentais para o funcionamento da economia. São diversas fontes de preços em áreas distintas: bens e serviços, trabalho (salário), dinheiro e tempo (juros), empresas (ações), uma moeda em relação a outra (câmbio). Historicamente, o Brasil sofreu com altas taxas de inflação que prejudicavam o poder de compra da população nos anos 80 e início dos anos 90. A partir da estabilização dos preços com o Plano Real, fez-se necessário o acompanhamento do seu comportamento para evitar o retorno às

taxas anteriores de inflação. Um dos principais preços da economia é a taxa de juros básicos utilizada pelo Banco Central para o controle da inflação. Tal taxa é aumentada quando há sinal de pressão inflacionária e ciclo de atividade em alta.

Assim, políticas monetárias e fiscais são usadas como forma de manter os preços sob controle e por conseguinte, influenciar o desempenho da economia. A qualidade das informações sobre os preços é um fator importante para que os órgãos públicos possam tomar decisões eficientes no que tange aos objetivos de controle de inflação. Estudos empíricos já mostraram a importância de dados coletados da *web* na avaliação dos pressupostos de rigidez de preços e proposição de medida de inflação oriunda de informações *online*.

O presente projeto de tese apresenta o seguinte problema de pesquisa: preços coletados diretamente da web produzem índices de inflação que são condizentes com os publicados pelos órgãos públicos brasileiros? Empiricamente, a rigidez de preços pode ser mensurada e avaliada tanto qualitativamente quanto estatisticamente? É possível avaliar o impacto das políticas públicas sobre os preços em diferentes regiões do país ou até mesmo setores da economia?

OBJETIVOS

O objetivo geral deste projeto é avaliar empiricamente a rigidez de preços na economia brasileira, bem como, propor um índice de inflação oriundo de preços coletados da *web*.

Dentro deste escopo, os objetivos específicos são: melhor compreensão do comportamento microeconômico dos preços e suas implicações para a economia. Tal ponto pode ser um fator vital para a construção de políticas públicas mais eficientes tanto em âmbito regional quanto em setores diversos da economia.

Capítulo 2

REVISÃO BIBLIOGRÁFICA

Muitos mecanismos microeconômicos têm sido propostos para explicar porque os preços são rígidos. A maior parte deles podem ser amplamente classificados em dois tipos de modelos: tempo-dependente e estado-dependente. Nos modelos tempo-dependentes, o momento de ajustamento nos preços é determinado exogenamente. Uma firma é capaz de definir o preço ótimo depois de um dado número de períodos (conforme [Taylor \(1980\)](#)), ou aleatoriamente em cada período (conforme [Calvo \(1983\)](#)). Com ajustamentos aleatórios, mudanças nos preços de qualquer tamanho são possíveis e existem uma fração estável de firmas ajustando em cada período, independentemente de quanto tempo já passou desde a última alteração no preço.

Modelos estado-dependentes, começam com os modelos de custo de menu de [Barro \(1972\)](#) e [Sheshinski and Weiss \(1977\)](#) e mais recentemente [Dotsey et al. \(1999\)](#) e [Golosov and Lucas Jr \(2007\)](#), tendem a ter maior fundamentação microeconômica. Eles são baseados na hipótese de que firmas são capazes de mudar seus preços em qualquer momento, mas devem enfrentar custos de ajustamento para fazer isso. Tais custos podem incluir custos trabalhistas para fazer as alterações, custos gerenciais para tomar a decisão, ou mesmo custos de “ira do cliente” como reação do cliente após o ajustamento nos preços. Em modelos estado-dependentes existem algumas pequenas alterações (dado que não vale a pena pagar o custo de menu) e os ajustem tendem a ocorrer mais frequentemente em preços “velhos” (onde o desvio do ótimo é maior).

Como já mencionado anteriormente, pesquisas empíricas que busquem avaliar a rigidez de preços tem pouca audiência em função da dificuldade de acesso a dados de preços de produtos e serviços em uma frequência considerável. O presente projeto de tese tem como objetivo avaliar se dados coletados da nuvem de diversos sites (supermercados, postos de gasolina, farmácias, ...) são capazes de contribuir para o estudo empírico da rigidez de preços. Além disso, propor um índice de inflação que possa ser utilizado como *proxy* para os índices divulgados pelos órgãos públicos e comumente usado no mercado.

O tema ainda é pouco explorado na academia, mas a difusão da internet e a facilidade de acesso via telefones celulares, por exemplo, são fatores que sinalizam para a importância que as informações armazenadas nas páginas de sites têm para estudos econômicos.

O *Billion Prices Project* (BPP) no *Massachusetts Institute of Technology* (MIT) é uma iniciativa acadêmica que usa preços coletados diariamente de varejistas com lojas online em todo o mundo para conduzir pesquisa econômica. O projeto tem sido uma fonte de dados para diversas pesquisas acadêmicas no que tange à rigidez de preços, impacto de união monetária nos preços relativos internacionais, taxas de câmbios reais, lei do preço único, inflação online versus inflação oficial de países emergentes onde as medidas divulgadas para esta variável são questionadas pelo mercado, comportamento dos preços e oferta de produtos em desastres naturais como os terremotos do Chile em 2010 e Japão em 2011.

Cavallo (2010) apresenta de forma inovadora a utilização de dados diários de preços individuais coletados de páginas de supermercados na Argentina, Brasil, Chile e Colômbia entre outubro de 2007 a outubro de 2008 para avaliar o comportamento microeconômico dos preços e sua implicação para modelos macroeconômicos. O autor apresenta fatos estilizados sobre a rigidez dos preços nestes quatro países e mostram que a distribuição do tamanho da mudança nos preços é bimodal (com algumas alterações próximas a zero por cento). As funções de risco agregado são inclinadas para cima ou em forma côncava, e existe sincronização de mudanças nos preços para marcas concorrentes. Esses fatores desafiam visões comumente vistas na literatura de rigidez de preços que tem sido influenciada por trabalhos teóricos anteriores.

Além disso, Cavallo (2010) providencia um índice de inflação alternativo para a Argentina onde estatísticas oficiais tem se tornado irreal para muitas pessoas. Eles mostraram que a inflação anual era consistentemente duas ou três vezes maior do que o índice oficial divulgado.

Cavallo et al. (2012) utilizaram um conjunto de dados de preços online de bens idênticos vendidos em dezenas de países por quatro dos maiores varejistas globais (Apple, IKEA, H&M, ZARA) para estudar a taxa de câmbio real e seus comportamentos agregados. Em contraste com a literatura, demonstraram que a lei do preço único (LPU) se matem dentro da zona do euro para dez dos milhares de bens vendidos pelos quatro varejistas em três indústrias não relacionadas, implicando taxa de juro real igual a um.

Não obstante, Cavallo et al. (2012) mostraram que desvios da LPU são significativamente maiores para esses mesmos produtos em países com diferentes moedas, mesmo se sua taxa de câmbio real é indexada. Por exemplo, preços na zona do euro diferem tipicamente daqueles na Suécia, que tem uma taxa de câmbio flutuante e também daqueles na Dinamarca, que vincula sua moeda ao euro. Isto esclarece que a moeda comum por si e não simplesmente a falta de volatilidade nominal, é importante na redução da dispersão dos preços entre países. Além disso, a LPU com os EUA se mantem mais para países dolarizados como Equador e El Salvador do que para países como Hong Kong ou Jordânia, que tem suas próprias moedas, mas vinculadas ao dólar americano.

Cavallo et al. (2013) estudaram o comportamento dos preços diários de supermercados e disponibilidade de produtos seguindo dois desastres naturais recentes: terremoto no Chile em 2010 e terremoto no Japão em 2011. Em ambos os casos existiu um efeito imediato e persistente sobre a disponibilidade de produtos. O número de bens disponíveis para venda caiu 32% no Chile e 17% no Japão a partir do dia do desastre

para seus pontos mais baixos, que ocorreram 61 e 18 dias depois dos terremotos, respectivamente. A disponibilidade de produtos recuperou lentamente e uma significativa parte dos bens permaneceram sem estoque depois de seis meses. Ao contrário, os preços ficaram estáveis por meses até mesmo para os bens que estavam experimentando grave carência. Essas tendências estão presentes para todos os níveis de agregação, mas existe heterogeneidade entre categorias. Ainda, [Cavallo et al. \(2013\)](#) para a frequência e magnitude nas mudanças de preços em ambos os países e encontraram que os resultados no Chile são consistentes com modelos de precificação onde varejistas tem medo da “ira do cliente”. No Japão a evidência sugere um maior papel para a ruptura de distribuição que restringe a capacidade dos varejistas em distribuir após o terremoto.

Recentemente, [Cavallo et al. \(2014\)](#) refletiram sobre as implicações de um país aderir a uma área de moeda comum nos preços relativos. Os autores consideraram o caso da Letónia, que recentemente deixou sua taxa de câmbio indexada e se juntou à zona do euro em janeiro de 2014. O artigo é o primeiro a usar dados de alta frequência para demonstrar que aderir a uma união monetária tem implicações econômicas para os preços dos países e a taxa de câmbio real. Tal experimento é particularmente útil porque a entrada da Letónia à zona do euro não carregou consigo profundas mudanças nas políticas governamentais.

Utilizando preços de milhares de bens vendidos pela Zara (coletados via web scraper), a maior varejista de roupas do mundo, os autores observaram que a dispersão dos preços na Letónia desmoronou rapidamente após a adoção do euro. O percentual de bens com preços quase idênticos em Letónia e Alemanha aumentou de 6% para 89%. Por outro lado, a tamanho mediano do preço declinou de 7% para 0%. Se um número maior de firmas também se comporta desta maneira, esses resultados sugerem que se tornar membro de uma união monetária tem significante implicações para a taxa de câmbio real de países.

Capítulo 3

METODOLOGIA

WEB SCRAPING

O surgimento da internet, em particular a Web (*Web Wide Word*) trouxe um crescimento exponencial nas disposições de informações. Embora muitas dessas informações sejam úteis, elas raramente estão de uma forma que podemos utilizá-las, pois ainda é comum pessoas gastarem horas na coleta manual de dados de páginas da web. Especificamente, apesar de disponibilidade, poucos trabalhos acadêmicos em Economia utilizam desta fonte de dados para análise econômica empírica. Tal característica pode ser dada pela dificuldade dos pesquisadores na área em lidar com linguagem de programação que demandam maior conhecimento de computação.

Em recente publicação, [Varian \(2014\)](#) salienta que as técnicas utilizadas na Ciência da Computação e outras áreas correlatas para manipular e analisar dados, têm muito a oferecer. [Varian \(2014\)](#) defende que economistas deveriam conhecer melhor esses métodos e usá-los em seus trabalhos. Além disso, [Varian \(2014\)](#) cita a comumente colaboração entre os departamentos de Ciência da Computação e Estatística nas universidades dos EUA. Porém, o autor espera que em um futuro próximo os estudantes de econometria tenham maior colaboração com esses perfis e assim, contribuir para a pesquisa econômica empírica.

Uma metodologia que facilita o processo de coleta de dados da web é conhecida como *web scraping* que envolve escrever algoritmos que executam automaticamente o que nós fazemos manualmente quando navegamos por uma página de um site de e-commerce, por exemplo. Além disso, necessita-se de pouco conhecimento em programação para iniciar o processo de coleta de dados da internet. Segundo [Manning et al. \(2008\)](#) *web scraping* é o processo de tirar informações desestruturadas de páginas da web e transformá-las em informações estruturadas que podem ser usadas para análise.

A maior parte das páginas de sites são construídas usando uma linguagem de codificação estruturada chamada de *HyperText Markup Language* (HTML). Este código tem “tags”, tais como `< center >` e `< bold >`, que determinam o estilo e localização do texto em uma página. Estas tags tendem a permanecer constantes ao longo do tempo, uma vez que proporcionam um “look and feel” distinto para cada página. Por contraste,

a informação dentro dessas *tags*, tais como preço de produtos, mudam ao longo do tempo. O software de *scraping* pode ser ensinado a utilizar as *tags* em HTML para localizar informações relevantes sobre um produto e guardá-las em um banco de dados. A repetição desse processo todos os dias produz um banco de dados em formato de painel com um registro por produto por dia. Em adição, o endereço da página (URL) onde cada produto é localizado pode ser usado para classificar produtos em categorias padronizadas.

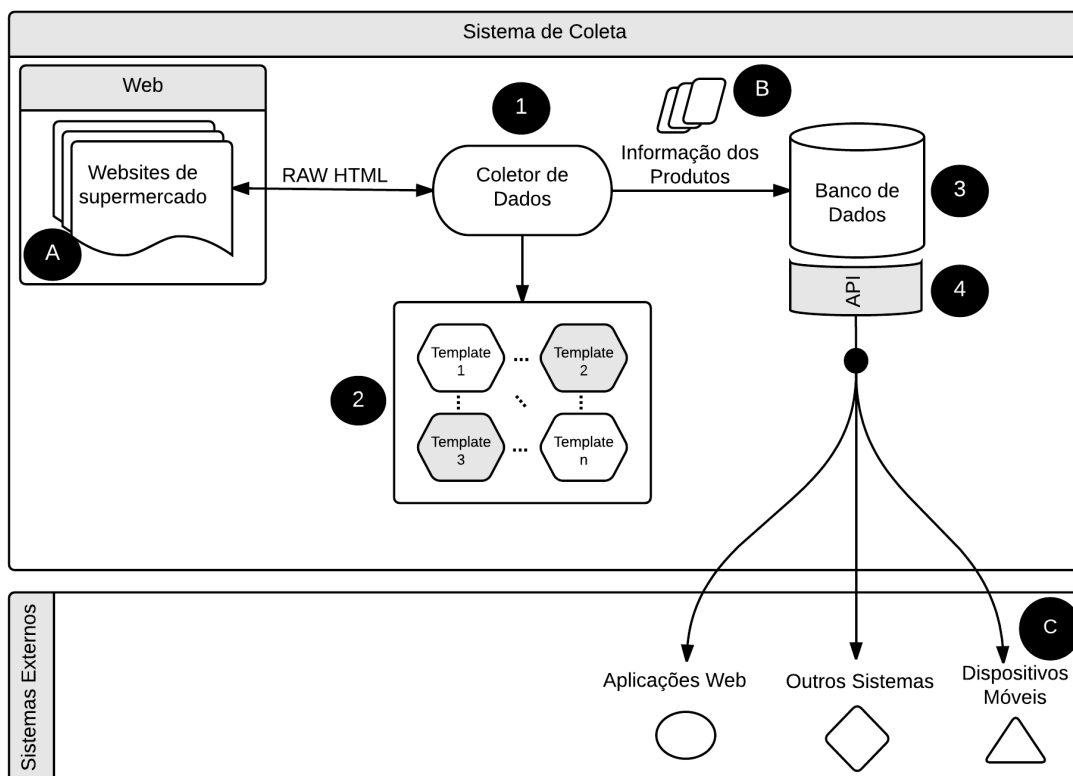


Figura 3.1: Arquitetura do Sistema de Coleta e Disponibilização dos Dados

Através de um coletor é possível arquitetar e executar de forma lógica e escalável todo esse processo. Para que um coletor seja funcional é necessário que o mesmo seja capaz de interagir com páginas da Web, extrair a informação de interesse e estruturar e armazenar os dados para futuras consultas. Em geral, exemplos corriqueiros de coletores podem ser citados como os desenvolvidos pelo Google e Microsoft para atuar na procura por páginas da internet ou outros mais específicos para coleta de preços de produtos como os portais de agregadores como Bom de Faro, Buscapé dentre outros. Assim, como [Cavallo \(2010\)](#), o presente projeto de tese busca de forma inovadora para a economia brasileira, explorar preços coletados de sites de supermercados, farmácias, companhia de energia elétrica, lojas de varejo online, lojas de roupas e calçados, entre outros, e propõe o uso de um sistema de coleta como o apresentado a seguir.

O sistema de coleta de preços está apresentado na 3.1. O coletor recebe como entrada os templates dos sites que se deseja coletar e produz como resposta informações estruturadas com os atributos dos produtos. Para que o coletor seja capaz de realizar a tarefa de extração de informação, o mesmo deve apresentar os seguintes componentes: um componente centralizado capaz de ler instruções e aplicar regras para extração (1 - módulo coletor de dados); ter um conjunto de regras que descreva de forma não ambígua como realizar a coleta dos dados e os atributos de interesse (2 - templates dos webistes de supermercado, por exemplo); ter um banco de dados capaz de lidar com as características dos dados armazenados (3 - módulo de banco de dados); ter uma interface para facilitar o acesso aos dados por meio de outras aplicações ou sistemas web (4 - módulo de disponibilização da informação). Dessa forma, o coletor (1) é o centralizador do processo de coleta de dados, fazendo a interação com os templates (2), os webistes (A) e o módulo de banco de dados (3). Em resumo, o coletor através de um algoritmo inicia o processo de coleta carregando em uma lista os templates de coleta dos websites (2) e em seguida através de um processo iterativo visita o website, coleta os documentos de interesse e realiza a extração das informações indicadas pelo template. Ao final da coleta os dados são estruturados em um formato de documento denominado JSON e armazena-os em um banco de dados NoSQL adequado a essa estrutura de dados. O processo se repete para cada template até que todos os templates sejam avaliados. O algoritmo que descreve esse processo é apresentado em Algoritmo 1. Por fim, além da coleta em si, há um módulo para disponibilizar o acesso a informação coletada. Esse módulo (4) é responsável por permitir de forma segura e racional o uso dos dados coletados por diferentes sistemas e aplicações web existentes (C).

Segundo Cavallo (2010), preços coletados da internet possuem duas desvantagens: Primeiro, percentual menor de empresas disponibiliza seus produtos e preços na internet em comparação com as lojas físicas. Tal limitação pode ser minimizada ao longo do tempo com uma maior oferta de produtos e serviços na internet. Segundo, os preços coletados da internet não incluem informações sobre as quantidades vendidas o que impede de obter market share e estimativas de elasticidade.

Avaliações futuras precisarão ser feitas de forma que seja capaz explorar se os preços online e off-line se comportam similarmente. Preocupado com tal validação, Cavallo (2010) fez pesquisa de preços nas lojas físicas dos supermercados utilizados para coleta de dados da internet. Desta forma, o autor examinou se os preços dos produtos nas lojas físicas eram similares aos preços nos sites. Uma importante característica é que um alto percentual de produtos vendidos nas lojas físicas também era comercializado nos sites em todos os países. O autor comparou os preços tanto em termo de nível quanto em tamanho e intervalo de tempo de alterações nos preços. Tal comportamento é muito importante para a avaliação de rigidez nos preços. Para tanto, o autor criou uma série de mudança de preços para cada produto que recebe valor 1 se o preço aumentou, 0 se o preço permaneceu constante e -1, caso contrário. Assim, foi possível avaliar se os preços dos produtos nas lojas físicas são semelhantes em nível e em direção de mudança para cada produto e supermercado dos países avaliados por Cavallo (2010).

Não obstante, Cavallo (2010) apresenta algumas vantagens dos preços coletados da

internet que os fazem uma fonte única de informação para análise de rigidez nos preços. Primeiro, pode-se obter preços diários para os produtos e serviços e por conseguinte, reduzir medidas de erro em relação à frequência de cálculo da inflação, analisar promoções de produtos, controles e sincronização nos preços. Segundo, os dados estão disponíveis para vários países, com maior facilidade de acesso e possibilidade de comparação entre países. Terceiro, existem informações detalhadas sobre cada produto e não há substituições forçada de itens como ocorre em estatísticas oficiais de inflação. Por fim, preços coletados da internet estão viáveis em tempo real, sem qualquer atraso para acessá-los. Isto pode ser usado para providenciar estimativas de rigidez nos preços em tempo real.

ÍNDICE DE PREÇOS ONLINE

Para calcular o índice de preços online que será comparado com o Índice Nacional de Preços ao Consumidor Amplo (IPCA) e Índice Nacional de Preços ao Consumidor (INPC) divulgados pelo Instituto Brasileiro de Geografia e Estatística (IBGE), utilizaremos a abordagem proposta por [Cavallo \(2010\)](#).

Assim, o índice de preços usa a combinação de dados online e as estruturas de ponderação oficiais do IBGE para as categorias da “cesta de mercadorias”¹ de cada índice de inflação. Maiores detalhes sobre a metodologia de coleta e cálculo do IPCA e INPC podem ser obtidas no [A](#). Dados diários serão utilizados para construir o índice de preços online o que é útil para observar padrões de curto prazo nos dados que ajudam a validar as informações online.

O índice de preço online será calculado utilizando os preços de todos os produtos disponíveis para compra em cada site. Isto implica que a cesta de bens muda dinamicamente ao longo do tempo podendo um produto aparecer ou desaparecer da cesta a qualquer momento devido à disponibilidade ou indisponibilidade no site. Além disso, o número de preços por produto tende a ser muito maior o que os coletados usualmente pelos órgãos governamentais. Para construir o índice, mudanças de preço são calculadas em nível de produto, então as médias dentro das categorias usando média geométrica ponderada e finalmente agregado entre categorias com uma média aritmética ponderada. Em particular, o primeiro passo é obter a média geométrica ponderada das mudanças nos preços na categoria j para cada dia t :

$$R_{t,t-1}^j = \prod_i \left(\frac{p_t^i}{p_{t-1}^i} \right)^{1/n_{j,t}} \quad (3.1)$$

onde p_t^i é o preço do bem i no tempo t , $n_{j,t}$ é o número de produtos na categoria j que estão presentes na amostra neste dia.

O segundo passo é computar o índice em nível de categoria em t :

¹Segundo o [IBGE \(2013\)](#) os índices constituem uma medida síntese de movimento de preços de um conjunto de bens e serviços, chamado “cesta de mercadorias”, representativo de um determinado grupo populacional, em certo período de tempo

$$I_t^j = R_{1,0}^j * R_{2,1}^j * \dots * R_{t,t-1}^j \quad (3.2)$$

Finalmente, o índice de preços no tempo t é a média aritmética ponderada de todos os índices das categorias:

$$IPO_t = \sum_j \frac{w_j}{W} I_t^j \quad (3.3)$$

onde w^j é o peso oficial utilizado pelo IBGE para tal categoria e W a soma de todos os pesos incluídos na amostra.

A classificação de produtos e pesos de categorias é uma das partes mais complexas deste processo. Nos dados originais, cada produto é atrelado à um endereço de web (URL) que corresponde à página onde o produto é localizado.

RIGIDEZ DE PREÇOS

Diversas estatísticas poderão ser utilizadas para a análise da rigidez de preços coletados da internet, como por exemplo: frequência de produtos com alterações diárias e frequências de alta e baixa em relação ao total de alterações nos preços em um dia. Assim, teremos um parâmetro que reflete a probabilidade incondicional de mudança no preço de uma firma ao longo de um dado período de tempo. Porém, a análise da frequência apresenta uma visão parcial do comportamento dos preços e faz-se necessário avaliar o tamanho da mudança dos preços por meio do valor absoluto da alteração no preço de um determinado produto (também o tamanho das mudanças positivas e negativas) e avaliação da sua distribuição de probabilidade. Desta forma, poder-se-á comparar o comportamento da inflação em determinadas regiões, municípios e estados em relação ao tamanho da mudança nos preços nestes locais.

[Cavallo \(2010\)](#) encontrou uma característica bimodal na distribuição do tamanho das alterações nos preços e uma forte queda da densidade das alterações próximo a 0% em alguns países, o que é consistente com os modelos de custo de menu que mudanças muito pequenas não são ótimas na presença de custo de ajuste. Outro tipo de análise é a avaliação da assimetria na densidade das alterações que pode refletir maior quantidade de preços crescendo que diminuindo.

Análise de Sobrevida

A análise de frequência ajudará na avaliação da rigidez de preços, mas ela sugere que a probabilidade de um preço se alterar é independente do tempo que uma mudança ocorre em relação à última alteração no preço. Ainda, a taxa de risco do preço se alterar é constante ao longo do tempo durante todo o período amostral. Embora esse método seja simples e efetivo para a comparação do grau de rigidez entre setores, regiões, cidades e países, um importante ponto reside sobre a forma da função risco.

Para avaliar a função risco utilizaremos a Análise de Sobrevida assim como [Cavallo \(2010\)](#). Conforme [Colosimo and Giolo \(2006\)](#), em análise de sobrevivência, a variável

resposta é, geralmente, o tempo até a ocorrência de um evento de interesse. Tal tempo é comumente conhecido como tempo de falha. Em medicina é comum o uso do método para a avaliação do tempo até a morte, transplante, doença, cura entre outros. No contexto de preços, estamos interessados no tempo até o ajuste do preço. Assim, tanto o aparecimento do risco e o evento de falha ocorrem quando uma firma muda seus preços.

A principal característica de dados de sobrevivência é a presença de censura que é a observação parcial da resposta. Isto se refere a situações em que por alguma razão, o acompanhamento do preço foi interrompido, seja porque a firma não vende mais um produto ou este não é produzido. A variável aleatória não-negativa T , usualmente contínua, que representa o tempo de falha, é geralmente especificada em análise de sobrevivência pela sua função de sobrevivência ou pela função de risco (tempo de falha). Estas duas funções são extensivamente usadas na análise de dados de sobrevivência.

Segundo Colosimo and Giolo (2006), a função de sobrevivência é definida como a probabilidade de uma observação não falhar até um certo tempo t , ou seja, a probabilidade de uma observação sobreviver (preço não se alterar) ao tempo t . Por outro lado, se T é a variável aleatória que mede a duração do preço, com função densidade $f(t)$ e densidade acumulada $F(t)$, o risco $h(t)$ é a probabilidade limite de que a mudança no preço ocorra em t , condicional ao preço não se alterar até este momento.

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{Pr(t < T < t + \Delta t | t < T)}{\Delta t} = \frac{f(t)}{1 - F(t)} \quad (3.4)$$

Esta função risco mede o risco instantâneo de um preço se alterar, condicionado à sobrevida. Podemos adicionar todas as taxas de risco ao longo do tempo e obter o risco total de um preço alterar acumulado até o tempo t . Isto é representado pelo função risco acumulado, $H(t)$:

$$H(t) = \int_0^t h(u) du = -\ln(1 - F(t)) \quad (3.5)$$

$H(t)$ é um aumento, função ilimitada de t , que acumula a probabilidade condicional do preço mudar ao longo do tempo. No contexto de repetidas "falhas"(preço se alterar), ela pode ser interpretada com o número esperado de ajustamento nos preços de 0 à t . O risco acumulado recebe grande atenção na Análise de Sobrevida porque ele é mais fácil de estimar do que a função risco sózinha.

Para estimar $H(t)$ e $h(t)$ empiricamente, usaremos conforme Cavallo (2010) uma abordagem não paramétrica dada por Nelson (1972) e Aalen (1978), que não requer hipóteses de distribuição de probabilidade. Métodos semi-paramétricos como o modelo Cox podem ser utilizados futuramente uma vez que permitem a incorporação de variáveis explicativas e a consideração da heterogeneidade não observável em nível de categoria de preços. Uma estimativa simples da função risco acumulado, $H(t)$, é dado por:

$$\hat{H}(t) = \sum_{j|t_j \leq t} \frac{c_j}{n_j} \quad (3.6)$$

onde c_j é o número de preços que mudaram em t_j e n_j é o número de preços sob risco em t_j , ou seja, os preços que não alteraram e não foram censurados até o instante imediatamente anterior a t_j . O passo incremental $\frac{c_j}{n_j}$ é uma estimativa para a probabilidade do preço mudar em t_j , levando em consideração apenas aqueles preços que sobreviveram até este ponto no tempo.

Para obter a função risco suavizada $\hat{h}(t)$, pode-se usar a seguinte equação:

$$\hat{h}(t) = \frac{1}{b} \sum_{j \in D} K\left(\frac{t - t_j}{b}\right) \Delta \hat{H}(t_j) \quad (3.7)$$

onde K é um kernel com densidade simétrica, b é a *bandwidth* de suavização e D é o conjunto de vezes com mudança de preços.

Referências Bibliográficas

- Ball, L. and Mankiw, N. G. (1994). A sticky-price manifesto.
- Barro, R. J. (1972). A theory of monopolistic price adjustment. *The Review of Economic Studies*, pages 17–26.
- Bils, M. and Klenow, P. J. (2004). Some evidence on the importance of sticky prices. *Journal of Political Economy*, 112(5):947–985.
- Blinder, A. S. and Stiglitz, J. E. (1983). Money, credit constraints, and economic activity.
- Bresser-Pereira, L. C. (1968). Da macroeconomia clássica à keynesiana.
- Calvo, G. A. (1983). Staggered prices in a utility-maximizing framework. *Journal of monetary Economics*, 12(3):383–398.
- Cavallo, A., Cavallo, E., and Rigobon, R. (2013). Prices and supply disruptions during natural disasters. Technical report, National Bureau of Economic Research.
- Cavallo, A., Neiman, B., and Rigobon, R. (2012). Currency unions, product introductions, and the real exchange rate. Technical report, National Bureau of Economic Research.
- Cavallo, A., Neiman, B., and Rigobon, R. (2014). The price impact of joining a currency union: evidence from latvia. Technical report, National Bureau of Economic Research.
- Cavallo, A. F. (2010). *Scraped data and prices in macroeconomics*. Harvard University.
- Colosimo, E. A. and Giolo, S. R. (2006). Análise de sobrevivência aplicada. In *ABE-Projeto Fisher*. Edgard Blücher.
- Dathein, R. (2000). O crescimento do desemprego nos países desenvolvidos e sua interpretação pela teoria econômica: as abordagens neoclássica, keynesiana e schumpeteriana. campinas. *Campinas, SP: IE. UNICAMP, 2000.(Tese de Doutorado)*.
- Dotsey, M., King, R. G., and Wolman, A. L. (1999). State-dependent pricing and the general equilibrium dynamics of money and output. *Quarterly journal of Economics*, pages 655–690.
- Friedman, B. M. and Kuttner, K. N. (1992). Money, income, prices, and interest rates. *The American Economic Review*, pages 472–492.

- Golosov, M. and Lucas Jr, R. E. (2007). Menu costs and phillips curves. *Journal of Political Economy*, 115(2).
- IBGE (2013). *Sistema Nacional de Índices de Preços ao Consumidor - Métodos de Cálculo*. Diretoria de Pesquisas - Instituto Brasileiro de Geografia e Estatística (IBGE).
- Lucas, R. E. (1972). Expectations and the neutrality of money. *Journal of economic theory*, 4(2):103–124.
- Manning, C. D., Raghavan, P., and Schütze, H. (2008). *Introduction to information retrieval*, volume 1. Cambridge university press Cambridge.
- McCallum, B. (1998). Stickiness: A comment. In *Carnegie-Rochester Conference Series on Public Policy*, volume 49, pages 357–363. Elsevier.
- Sheshinski, E. and Weiss, Y. (1977). Inflation and costs of price adjustment. *The Review of Economic Studies*, pages 287–303.
- Taylor, J. B. (1980). Aggregate dynamics and staggered contracts. *The Journal of Political Economy*, pages 1–23.
- Varian, H. R. (2014). Big data: New tricks for econometrics. *The Journal of Economic Perspectives*, 28(2):3–27.

Apêndice A

METODOLOGIA IBGE

O Sistema Nacional de Preços ao Consumidor (SNIPC) efetua a produção e sistemática de índices de preços ao consumidor tendo como unidade de coleta estabelecimentos comerciais e de prestação de serviços, concessionária de serviços públicos e domicílios (para levantamento de aluguel e condomínio). O sistema abrange as regiões metropolitanas do Rio de Janeiro, Porto Alegre, Belo Horizonte, Recife, São Paulo, Belém, Fortaleza, Salvador e Curitiba, além do Distrito Federal e do município de Goiânia. A partir de janeiro de 2014, o SNIPC passou a incorporar a Região Metropolitana de Vitória/ES e o município de Campo Grande/MS.

As motivações para criação do Índice Nacional de Preços ao Consumidor Amplo (IPCA) e Índice Nacional de Preços ao Consumidor (INPC) foram a obtenção de medida geral de inflação e a indexação salarial, respectivamente.

As etapas para a construção dos índices de preços são elencadas abaixo. Para maiores detalhes, consultar [IBGE \(2013\)](#).

1. Definição da população objetivo:
 - (a) Para o INPC são as famílias residentes nas áreas urbanas das regiões de abrangência do SNIPC com rendimentos de 1 a 6 salários-mínimos e cujos chefes são assalariados;
 - (b) Para o IPCA, as famílias residentes nas áreas urbanas das regiões de abrangência do SNIPC com rendimentos de 1 a 40 salários-mínimos, qualquer que seja a fonte de rendimentos.
2. Obter estruturas de ponderação: O conjunto de bens e serviços representativos do consumo dos grupos e os valores de despesa que lhes são associados.
 - (a) Pode ser diferente para uma determinada população-objetivo;
 - (b) São resultado da consolidação dos orçamentos familiares levantados pela POF;

(c) São montadas de forma que categorias de consumo de mesma natureza fiquem juntas. Hierarquicamente ²:

- i. Grupo
- ii. Subgrupo
- iii. Item
- iv. Subitem

3. Calcular os Pesos:

- (a) Anualisar os valores de despesa com consumo oriúndas da POF que são coletados em diferentes períodos de referência;
- (b) Colocar as despesas anuais em preços constantes de 15 de janeiro de 2009;
- (c) Somar para cada subitem, despesas das famílias pertencentes à população-objetivo;
- (d) A razão da soma anterior e a despesa total de todas as famílias da região em questão gera o índice.

4. Definir estruturas de consumo: A partir da participação dos subitens, define-se quais permanecerão para o cálculo do índice. Para tanto, utiliza-se o seguinte critério:

- (a) subitens com participação igual ou superior a 0,07% fazem parte das estruturas;
- (b) subitens com participação inferior a 0,01% em hipótese alguma fazem parte das estruturas;
- (c) subitens com ponderação igual ou superior a 0,01% e inferior a 0,07% podem fazer parte para assegurar que o item do qual fazem parte tenha cobertura de 70% dos gastos realizados com os componentes do item.

5. Cadastrar informantes: Por meio da Pesquisa de Locais de Compra (PLC) faz-se o cadastro dos estabelecimentos.

6. Cadastrar produtos: Por meio da Pesquisa de Especificação de Produtos e Serviços (PEPS) obtém-se os produtos.

7. Coletar preços: Tarefa contínua, realizada mensalmente, nas áreas de cobertura da pesquisa, ao longo do mês. Para viabilizá-la existem pesquisadores de campo dedicados à coleta de informações necessárias à produção dos índices. Questionário eletrônico de coleta instalado em computador de mão, no qual estão descritas as características dos produtos ou serviços nele investigados.

²Por exemplo, Laranja-pera é um subitem do item "Frutas" que conjuntamente com outros itens formam o subgrupo "Alimentação no Domicílio", o qual, unido ao subgrupo "Alimentação Fora do Domicílio" compõe o grupo "Alimentação e Bebidas". Retirado de [IBGE \(2013\)](#)

A tabela A.1, apresenta um resumo das fontes de informações relevantes para os índices de preços: Pesquisa de Orçamentos Familiares (POF), Pesquisa de Locais de Compra (PLC) e Pesquisa de Especificação de Produtos e Serviços (PEPS).

PESQUISAS BÁSICAS	
POF	Fornece as estruturas de ponderação para cada grupo de bens e serviços).
PLC	Fornece o cadastro de informantes da pesquisa que tem manutenção constante.
PEPS	Fornece o cadastro de produtos e serviços a serem pesquisados.

Tabela A.1: Principais Pesquisas Utilizadas na Metodologia

Por fim, temos a metodologia de cálculo dos índices de preços. Sinteticamente, partindo-se de milhares de preços coletados mensalmente, obtêm-se no primeiro processo-síntese, as estimativas dos movimentos de preços referentes a cada produto pesquisado. Estes resultados são agregados por uma fórmula elementar de cálculo e geram a estimativa para variação de preços de cada subitem; essas estimativas, por sua vez, por outro processo agregativo, produzem os índices referentes a itens, que, por fim, geram os índices regionais e nacional mensais de cada população-objetivo.

Cálculo no nível de produto

Primeiro, calcula-se mensalmente o relativo de preços referentes a dois meses e temos a estimativa da variação mensal dos preços do produto j , ou relativo do produto j , conforme:

$$R_{t-1,t}^j = \frac{\bar{P}_t^j}{\bar{P}_{t-1}^j} = \frac{\frac{1}{n_t} \sum_{L=1}^{n_t} p_t^{j,L}}{\frac{1}{n_{t-1}} \sum_{L=1}^{n_{t-1}} p_{t-1}^{j,L}} \quad (\text{A.1})$$

onde $R_{t-1,t}^j$ é a medida da variação de preços do produto j entre os meses $t-1$ e t , \bar{P}_t^j e \bar{P}_{t-1}^j os preços médio do produto j nos meses t e $t-1$, respectivamente, assim como $p_t^{j,L}$ e $p_{t-1}^{j,L}$ são os preços do produto j no local L nos meses t e $t-1$. Por fim, n_t e n_{t-1} são os números de locais que compõem a amostra do produto nos meses t e $t-1$.

Por conseguinte, o próximo passo é a agregação no nível de subitem. Para tanto, calcula-se a média geométrica dos resultados obtidos para cada produto que compõe o subitem. Assim,

$$R_{t-1,t}^k = \sqrt[m_k]{\prod_{j=1}^{m_k} R_{t-1,t}^j} \quad (\text{A.2})$$

onde $R_{t-1,t}^k$ é a variação média de preços entre os meses $t-1$ e t dos produtos que compõem o subitem k , $R_{t-1,t}^j$ é a variação do preço do produto j entre os meses $t-1$ e t (fórmula

No que diz respeito aos resultados ao longo do tempo, evidencia-se a importância de manter-se o painel de produtos fixos, a exemplo do que ocorre com o painel de locais, sob pena de incorporar falsas variações de preços. Desta forma, o IBGE imputa o preço de um produto para determinado local ou subitem. Para maiores informações de como é feito esse processo, consultar [IBGE \(2013\)](#).

Cálculo no nível de item

Usa-se a fórmula de Laspeyres que expressa a razão entre o gasto efetuado no momento t , necessário para consumir as mesmas quantidades do momento 0, e o gasto efetuado no momento 0. A fórmula [A.3](#) representa o índice:

$$L_{0,t} = \sum_{i=1}^n \left(\frac{p_0^i q_0^i}{\sum_{j=1}^n p_0^j q_0^j} \right) \left(\frac{p_t^i}{p_0^i} \right) \quad (\text{A.3})$$

onde $\frac{p_t^i}{p_0^i} = R_{0,t}^i$ é o estimador da variação de preços do subitem i entre os momentos 0 e t e $\frac{p_0^i q_0^i}{\sum_{j=1}^n p_0^j q_0^j} = W_0^i$ é o peso do subitem i obtido a partir da POF. Observe-se que tanto $R_{0,t}^i$ como W_0^i referem-se, na prática, a pequenos agregados de produtos. Para se conhecer a variação de preços do item m para uma determinada área e faixa de rendimento em ciclos mensais utiliza-se a fórmula [A.4](#).

$$I_{t-1,t}^m = \frac{\sum_{i=1}^n W_{t-1}^i R_{t-1,t}^i}{\sum_{i=1}^n W_{t-1}^i} \quad (\text{A.4})$$

onde $I_{t-1,t}^m$ é o índice do item m entre os meses $t-1$ e t , W_{t-1}^i é o peso do subitem i em $t-1$ e $R_{t-1,t}^i$ é o relativo do subitem i entre $t-1$ e t . Além disso, o peso W_{t-1}^i a partir de $t=2$ é dado por:

$$W_{t-1}^i = W_0^i \prod_{j=0}^{t-2} \frac{R_{j,j+1}^i}{I_{j,j+1}} \quad (\text{A.5})$$

onde W_0^i é o peso do subitem i obtido a partir da POF, $R_{j,j+1}^i$ é o relativo do subitem i entre os meses j e $j+1$ e $I_{j,j+1}$ é o resultado do índice geral entre os meses j e $j+1$.

Cálculo dos índices regionais

O resultado mensal para a área A e população-objetivo F é dado por:

$$IPC_{t-1,t}^{A,F} = \sum_m^M W_{t-1}^m I_{t-1,t}^m \quad (\text{A.6})$$

onde $I_{t-1,t}^m$ é o índice do item m obtido conforme a equação [A.4](#) e W_{t-1}^m corresponde ao peso de cada item e é obtido somando-se os pesos no período $t-1$ por meio da equação [A.5](#) utilizando todos os subitens que compõem o respectivo item m .

Cálculo dos índices nacionais

Os índices nacionais são obtidos a partir dos índices regionais. O método empregado para obtenção dos índices nacionais consiste no cálculo de uma média aritmética ponderada dos índices regionais mensais, conforme:

$$INPC_{t-1,t} = \sum_{A=1}^{11} W^{A,F} IPC_{t-1,t}^{A,F} \quad (\text{A.7})$$

onde $INPC_{t-1,t}$ é o índice nacional referente à variação de preços entre os meses $t-1$ e t , $IPC_{t-1,t}^{A,F}$ é o índice da área A , população-objetivo F , obtido via [A.6](#). Além disso, $W^{A,F}$ é o peso da área A , população-objetivo F . Na mais recente atualização, tendo como fonte a POF 2008-2009, os pesos das regiões foram obtidos com base nas estimativas da população urbana para os estados, Grandes Regiões e Brasil. A tabela [A.2](#), apresenta os índices regionais antes e após a alteração.

Regiões	IPCA	INPC
Belém	4,65	7,03
Fortaleza	3,49	6,61
Recife	5,05	7,17
Salvador	7,35	10,67
Belo Horizonte	10,86	10,60
Vitória	1,78	1,83
Rio de Janeiro	12,06	9,51
São Paulo	30,67	24,24
Curitiba	7,79	7,29
Porto Alegre	8,40	7,38
Campo Grande	1,51	1,64
Goiânia	3,59	4,15
Brasília	2,80	1,88

Tabela A.2: Participação do índice regional no agregado nacional