**UNIVERSIDADE DE SÃO PAULO
FACULDADE DE ECONOMIA, ADMINISTRAÇÃO E CONTABILIDADE
DEPARTAMENTO DE ECONOMIA - RP
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA**

**NÚCLEO DA INFLAÇÃO COMO FATOR COMUM DO IPCA: UMA ABORDAGEM DO
MODELO DE FATOR DINÂMICO GENERALIZADO**

Ana Paula de Almeida Alves

Orientador: Prof. Dr. Jaylson Jair da Silveira

2009

Dissertação defendida e aprovada no Departamento de Economia da Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo/RP Programa de Pós-Graduação em Economia Aplicada, pela seguinte banca examinadora:

Alves, Ana Paula de Almeida

Núcleo da inflação com fator comum do IPCA: uma abordagem do modelo de fator dinâmico generalizado / Ana Paula de Almeida Alves - Ribeirão Preto , 2009
60 p.

Dissertação (Mestrado) - Universidade de São Paulo, 2009

Bibliografia

1. Núcleos de inflação - Brasil 2. Modelo de Fatores I. Universidade de São Paulo. Faculdade de Economia, Administração e Contabilidade II. Título

ERRATA

Alves, A.P.A. **Núcleo da inflação como fator comum do IPCA: uma abordagem do modelo de fator dinâmico generalizado**. 2009. 60 pp. Dissertação (Mestrado em Economia Aplicada) - Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, 2009.

página	equação	onde se lê	leia-se
24	13		<p>A projeção de χ_{it} terá a forma $\chi_{it} = K_i^n(L)X_t$, em que:</p> $K_i^n(L) := \sum_{j=1}^q p_{ji}(L)p_j(L)$

página	linha	onde se lê	leia-se
13	24	Bryan e Cecchetti (1996)	Bryan e Cecchetti (1993)
15	21	Fiorencio e Moreira (2002)	Fiorencio e Moreira (2001)
21	23	Ou seja, são quadrado somáveis, ou seja são quadrado somáveis.	Ou seja, são quadrado somáveis
22	22	$\Gamma_{\chi}^n = (o)$	$\Gamma_{\chi}^n = (\theta)$
23	10	$\Gamma_{n,-M}^T, \dots, \Gamma_{n0}^T, \dots, \Gamma_{nM}^T$	$\Gamma_{n,-M}^T, \dots, \Gamma_{n,0}^T, \dots, \Gamma_{n,M}^T$
39	2	Referência [31]	Picchetti, P. e Toledo, C. (2000) "How Much to Trim ? A Methodology for Calculating Core Inflation, With an Application for Brazil", Economia Aplicada, Vol.4 – N. 4
39	3	Referência [32]	Chamberlain, G. e Rothschild, M (1983), "Arbitrage, Factor Structure and Mean-Variance Analysis in Large Asset Markets", Vol.51, 1305-1324
89	10	Referência [34]	Forni e Reichlin (2001), "Federal

Policies and Local Economies: Europe and the US", European Economic Reviews, No.45, pp. 109-134.

página

figura

onde se lê

leia-se

35

12

1.27***

-1.27***

35

12

1.56***

-1.56***

Núcleo da inflação como fator comum do IPCA: uma abordagem do modelo de fator dinâmico generalizado

Ana Paula de Almeida Alves *

2009/Janeiro

*Dissertação de Mestrado em Economia Aplicada/FEARP/USP-2009

Agradeço o apoio incondicional de meus pais, meus irmãos e Andres.

Agradeço a todos os meus amigos pela paciência e compreensão de minha ausência, mas principalmente por sempre estarem ao meu lado. Em especial aos meus amigos de muitos anos Fabiana, Karen, Bruno, Daniela e Preta; que nunca desistiram de mim!

Agradeço aos professores do departamento com quem tive a chance de ter aulas e muito me ensinaram, e as contribuições de Roseli Silva e Alexandre Chibebé Nicolella.

Agradeço aos meus colegas de mestrado pelas longas conversas e apoio mútuo em momentos decisivos, Ludmila Sá Fonseca, Rodrigo Takeuchi, Caio Casella, Leandro Stocco, Michael Martins e Beatriz Selan.

Agradeço a Octavio de Barros por todo apoio e confiança depositados desde o início deste projeto, e pela oportunidade de trabalhar em um ambiente de grande inquietude intelectual, onde todos de alguma forma contribuíram para a execução deste trabalho, em especial, minha amiga e colega de trabalho, Andrea Bastos Damico. As contribuições pontuais e essenciais de Leandro Camara e Myriã Bast.

Agradeço especialmente a Fernando Freitas, Thais Ortega e Bruno Rocha pelas críticas e sugestões sempre fundamentais.

Agradeço a amiga Clara Setoguchi pela ajuda na correção final.

Agradecimento especial ao meu orientador, Jaylson Silveira, pela paciência, comentários e apoio nas horas mais decisivas.

Muito obrigada!

Resumo: Sob o regime de metas de inflação cabe à autoridade monetária balisar seus instrumentos de política de forma a manter a estabilidade do nível geral de preços. Neste aspecto, pelo caráter volátil dos índices de inflação cheia os bancos centrais de todo o mundo utilizam o conceito de núcleo da inflação para tentar capturar com maior acurácia a tendência subjacente da taxa de inflação. Muitas vezes os índices de preços ao consumidor estão altamente sujeitos a volatilidades decorrentes de fatores temporários e muitas vezes localizados. E já que o objetivo da autoridade monetária está em zelar pela estabilidade "real"(ou de fato) do nível geral de preços, mudanças temporárias ou localizadas não afetam as taxas de inflação no longo prazo e, conseqüentemente, não cabe à autoridade monetária responder a tais mudanças, pois isso poderia gerar uma volatilidade desnecessária à política monetária com consequência sobre as flutuações da atividade econômica no período. Dessa forma, Bancos Centrais do mundo inteiro fazem uso de núcleos de inflação. Este trabalho aplica uma nova metodologia de cálculo de núcleo para a inflação brasileira, utilizando o modelo de fatores dinâmicos generalizados. Esta abordagem permite diferenciar fatores localizados (idiossincráticos) dos choques comuns (generalizados) em um grande conjunto de dados. Usamos o IPCA em seu nível mais desagregado e geramos o choque comum entre este conjunto. E a este choque chamamos de núcleo da inflação. Sua eficiência em termos de antecedência à inflação cheia no curto prazo foi testada por meio de uma cointegração, VEC, tais resultados foram comparados com o desempenho do núcleo por Exclusão, mostrando uma maior eficiência do núcleo aqui encontrado

Palavras-chave: componente comum; modelo dinâmico; análise espectral; núcleo da inflação.

Abstract: Under the inflation target system lies to the monetary authority the evaluation of the best tools to keep general price stability. In this context, due to the volatile character of the inflation, central banks around the world use the concept of the inflation core in attempt to capture in a more accurately way the prices trends. Several times, consumer prices indexes are subjected to very volatile prices, due to temporary or localized factors. As the vigilance of the monetary authority relies on the real stability of the general prices level, temporary or localized changes doesn't affect the inflation indexes in the long run and, therefore, it's not an issue to the central bank to respond to this variations, this could indeed create an unnecessary volatility to the monetary politics with consequences to the economic activity in the period. This way, central banks around the world calculate and use inflation core. This paper applies a new methodology to calculate the inflation core to the Brazilian inflation, using the generalized dynamic factor model. With this approach it's possible to differentiate the localized factors from the common (generalized) shocks in a great data set. We use IPCA on its more disaggregated level and create a common shock in the data set, and we name this shock the inflation core. We test the advance of this core to the inflation in the short run using a VEC, and compare with the results of the Exclusion core, we show that your core by using dynamic factor model is more efficient then Exclusion core.

Keywords: common components; dynamic model; spectral analysis; core inflation

Sumário

1	Introdução	9
2	A Literatura sobre núcleo da Inflação	13
3	Modelo de fatores dinâmicos: análise de séries de tempo no domínio da frequência	17
3.1	Literatura de modelo de fatores	17
3.2	Procedimento de estimação do núcleo	19
4	Núcleo da Inflação Brasileira	25
4.1	Conjunto de Dados	25
4.1.1	IPCA <i>reponderado</i>	25
4.1.2	Pré-Tratamento do Dados	26
4.2	Determinação do número de fatores para o núcleo FHLR	27
5	Resultados: teste de cointegração	31
6	Conclusões	36
7	Referências Bibliográficas	37
8	Apêndices	40
8.1	Análise espectral	40
8.2	Apêndice: testes e resultados para a estimação do modelo FHLR	47
8.3	Rotinas para o Matlab	58

Lista de Figuras

1	IPCA IBGE <i>versus</i> IPCA <i>reponderado</i>	26
2	Variâncias explicadas dos primeiro 8 autovalores dinâmicos	28
3	Média dos autovalores dinâmicos para as diferentes frequências	29
4	Evolução dos componentes idiossincráticos e comuns	30
5	IPCA <i>reponderado</i> e núcleo FHLR	30
6	IPCA <i>reponderado</i> e núcleo Exclusão	31
7	Núcleo FHLR e núcleo por Exclusão	32
8	Comparação dos ciclos de política monetária e desvios da inflação de longo prazo pelo modelo FHLR	33
9	Índices base 100=Ago/1999 do IPCA <i>reponderado</i> , Exclusão e FHLR . . .	34
10	Testes de Cointegração - Estatísticas do Traço e Auto-Valor	34
11	Vetor de Correção de Erros	35
12	Coeficientes de Longo Prazo	35
13	Teste de Raiz Unitária no logaritmo - probabilidade	56
14	Critério de Informação - Modelo: IPCA <i>reponderado</i> e Núcleo Exclusão . .	57
15	Critério de Informação - Modelo: IPCA <i>reponderado</i> e Núcleo FHLR . . .	57

1 Introdução

No início da década de noventa o regime de metas de inflação surgiu como alternativa de política monetária em detrimento ao controle de agregados monetários e da taxa nominal de câmbio¹. Sua disseminação a partir de então, deveu-se em grande medida à atribuição em nível teórico da sua capacidade de aumentar o bem-estar econômico, uma vez que reduz os custos de desinflação da economia².

Sob a hipótese quase que consensual de que a redução da inflação traz um maior bem-estar, mas não sem alguma perda de produto no horizonte de curto prazo, pode-se dizer que há um claro *trade-off* entre produto e inflação presente no regime de metas, tal como uma curva de Phillips de curto prazo. Entretanto, a minimização deste sacrifício³ em termos de PIB (ou desemprego) foi importante para que o regime fosse amplamente adotado por diversos países como estratégia de manutenção do nível geral de preços desde seu advento⁴.

Diante do exposto acima, existe um alto custo, em termos de produto, para a autoridade monetária ao cometer equívocos na interpretação dos dados correntes de inflação. Não obstante, o estudo de formas de cálculo e a análise dos índices de inflação ganharam cada vez mais destaque na literatura econômica após a implantação do regime de metas, sendo as principais delas: o estudo de núcleos de inflação e as formas de cálculo de índices de preços. Desta forma, o debate sobre núcleos⁵ de inflação cresceu em todo o mundo. Diversos Bancos Centrais⁶ usam estimações de medidas de inflação subjacente.

No Brasil, a implantação do regime de metas se deu em junho de 1999 e, a partir de então, o Banco Central brasileiro passou a ter como objetivo o cumprimento de metas pré-

¹A dificuldade do controle dos agregados monetários e consequentemente a ineficiência do regime no controle de preços determinou a quase estição do regime de agregados monetários, nas palavras de Alan Blider (1999): *Não foi a política monetária que abandonou os agregados monetários e sim os agregados monetários que abandonaram a política monetária*. Dificuldades no controle do fluxo de capitais também tornou bastante problemático o regime de taxa nominal de câmbio, principalmente entre os países emergentes.

²Para evidências empíricas ver Carvalho, 2006.

³A credibilidade do Banco Central é tomada pela literatura como ponto-chave na minimização deste sacrifício (Bernanke et al., 1999). Entre as vantagens do regime de metas estão: a transparência da política monetária, diminuição da variância e possibilidade de acomodação parcial de choques.

⁴A Nova Zelândia, no início da década de 1990, foi o primeiro país a utilizá-lo, entretanto, já em 2008 vinte e quatro países o utilizavam como diretriz de política monetária. Entre os países emergentes que o adotaram neste período inclui-se Brasil, Chile, Hungria, Coreia do Sul, África do Sul, México e Turquia mais recentemente.

⁵Em 1999 o BIS promoveu um seminário abordando o tema.

⁶Nos EUA, Reino Unido e Nova Zelandia, União Européia, entre outros.

determinadas pelo Conselho Monetário Nacional⁷(CMN). Sob este regime, a autoridade monetária deve estar atenta à desenvolver ferramentas que minimizem os ruídos de curto prazo presentes nos cálculos de índices de preços e que não representam, desta forma, tendências de longo prazo para o nível geral de preços e, conseqüentemente, não requerem respostas de política monetária. Tal fato motivou o desenvolvimento deste trabalho.

A partir da adoção do referido regime, uma série de eventos conjunturais provocou desvios da inflação com relação às metas, culminando em uma lenta trajetória de convergência. Crises externas⁸ e internas⁹ entre 2001 e 2003 aumentaram a aversão ao risco e provocaram forte depreciação cambial com conseqüência sobre as taxas de inflação, que permaneceram acima da meta estipulada neste período. Tais eventos motivaram a autoridade monetária a introduzir, em algumas ocasiões, metas ajustadas com o objetivo de acomodar parte de choques primários sobre o nível geral de preços. A utilização mais recente deste mecanismo foi em 2003 quando o Banco Central alterou a meta de 4,0% para 8,5% e, conseqüentemente, para acomodar a inércia mais alta vinda do ano em questão, alterou a meta do ano subsequente (2004), de 3,75% para 5,5%. Ainda em relação a 2004, era necessário acomodar também parte do choque de oferta originário da alta do preço internacional de algumas *commodities*¹⁰.

Ainda assim, a autoridade monetária conseguiu manter as taxas de inflação em níveis relativamente baixos e, apesar de lenta, as taxas convergiram para a meta. No período entre 2002 e 2006 a trajetória do Índice de Preços ao Consumidor Amplo (IPCA) foi de franca desaceleração, passando de 12,5% no início do período e terminando em 3,14% em 2006¹¹. Durante este período a taxa de juros (instrumento do Banco Central) aumentou substancialmente sua eficácia sobre a demanda agregada e, conseqüentemente, sobre o nível geral de preços (Minella, 2001).

Neste contexto, cabe ao Banco Central compreender cada vez melhor os mecanismos de transmissão da política monetária. O acompanhamento de tais mecanismos deve ultrapassar a análise agregada de suas defasagens e magnitudes e incluir, também, o entendimento dos impactos diferenciados que a política monetária exerce sobre a composição final dos diversos preços da economia. Para tanto, ele deve considerar e diferenciar características idiossincráticas de cada setor produtivo.

Sob este aspecto, cabe à autoridade monetária diferenciar entre movimentos generalizados e movimentos localizados de preços, além de fazer distinção entre os duradouros e temporários. Entretanto, os índices de inflação são regras deterministas (Fava, 2007)

⁷O Conselho Monetário Nacional é composto pelo ministro da fazenda, presidente do Banco Central e o ministro do planejamento.

⁸Atentado terrorista nos Estados Unidos e crise Argentina.

⁹Sucessão presidencial de 2002.

¹⁰As *commodities* metálicas subiram cerca de 64% em dólar no período segundo dados da Bloomberg-LMX (London Metal Exchange).

¹¹Em 2003 o IPCA variou 9,3%, em 2004 7,6% e 5,7% em 2005.

sujeitas a erros amostrais, não obstante, possuem flutuações altamente voláteis que dificultam a diferenciação de tais movimentos.

Por conta disto, um regime de metas necessita de uma outra medida de inflação, que não o índice cheio, para melhor nortear as decisões de política monetária. Tal medida é conhecida como núcleo (*core*), e tem como objetivo principal medir a tendência subjacente da inflação corrente. Aqui dito, por tendência subjacente, deve-se entender que o núcleo será capaz de diferenciar entre movimentos generalizados e idiossincráticos, permanentes e temporários, e ainda ser um bom antecedente da inflação no curto prazo - definição também utilizada por Fiorenco e Moreira (2002)¹².

Isto posto, a definição de um núcleo consistente é de suma importância para a condução da política monetária em um contexto de metas de inflação. Caso contrário, a política monetária poderia se tornar bastante volátil ao reagir a movimentos temporários, sem identificar sua tendência. Vale aqui uma ressalva, o debate sobre a utilização dos *headline* dos índices de inflação ou do núcleo como meta ou objetivo de política monetária propriamente dito também é vasto, entretanto, não é escopo deste trabalho discutir as eficiências em termos de bem-estar econômico da utilização de um ou outro método, caberá apenas encontrar uma ferramenta adicional na gestão da política monetária e não uma meta.

Há uma vasta literatura sobre núcleos de inflação. A metodologia mais tradicional e simples - bastante usada por diversos bancos centrais¹³ - é o núcleo que exclui itens pré-determinados, em geral alimentos e energia. A exclusão de tais itens deve-se ao fato de que são produtos altamente sujeitos a sazonalidades, como safra e entressafra no caso dos alimentos, e fatores geopolíticos, no caso do petróleo (energia). O Banco Central brasileiro usa uma versão adaptada às peculiaridades domésticas ao excluir também o grupo de produtos ditos monitorados. Tal grupo inclui preços cujos reajustes seguem uma metodologia quase que contratual¹⁴ ou estão atreladas a eventos exógenos - como o caso da energia proveniente do petróleo. No próximo capítulo detalharemos os demais métodos utilizados na construção de núcleos.

Considerando o exposto anteriormente, este trabalho tem como objetivo estimar um núcleo de inflação com base no modelo de fatores dinâmicos, proposto por Forni *et al.* (2002), que é capaz de diferenciar choques generalizados e localizados, bem como choques duradouros e temporários. O modelo de fatores pressupõe que o comportamento de um grande conjunto de dados é estabelecido por poucas fontes (fatores) comuns, somadas

¹²A diferenciação entre choques de oferta e demanda também é requerida pela literatura na definição de núcleos, mas isto não é necessário pois o objetivo do trabalho é encontrar e diferenciar os movimentos constantes nas taxas de inflação, não importando suas causas.

¹³Como por exemplo os bancos centrais americano, japonês, grego, francês e australiano.

¹⁴Energia elétrica que é reajustada em geral com base na inflação passada mais a variação cambial; telefonia fixa também reajustada pela inflação passada ponderada pelos custos do setor de telecomunicações; tarifa de água e esgoto; ônibus urbano; entre outros.

aos choques idiossincráticos. Desta forma, chamaremos de núcleo da inflação os fatores comuns entre os diversos preços constantes do IPCA.

Este trabalho está organizado como segue. Na próxima seção faremos uma resenha da literatura sobre núcleos de inflação. Na seção 3 apresentaremos o Modelo de Fatores Dinâmicos Generalizados. Na seção 4 detalharemos o conjunto de dados utilizados e os tratamentos aplicados a ele, juntamente com os resultados encontrados. Na seção 5 faremos uma comparação do núcleo aqui encontrado e o núcleo por exclusão atualmente calculado pelo Banco Central brasileiro.

2 A Literatura sobre núcleo da Inflação

Do ponto de vista de política monetária, a literatura sobre índices de preços pode ser dividida em dois grupos diferenciados com base no foco das análises, a saber: o primeiro diz respeito ao cálculo estrito dos índices de preços e formas de diminuição de vieses amostrais e o segundo engloba trabalhos sobre as formas de cálculos de inflação subjacentes derivadas dos índices de preços.

No primeiro enfoque toma-se como referência um índice de inflação perfeito. Em teoria este índice seria aquele capaz de medir as variações de preços necessárias para se manter um determinado nível de utilidade (custo de vida). Entretanto, é bastante problemático medir variações em termos de utilidade, assim, a metodologia mais utilizada em todo o mundo para se medir a inflação é o índice de preços do tipo Laspeyres Modificado, que compara os preços de uma cesta de bens e serviços em um determinado período-base ao longo do tempo. Sendo assim, ao utilizar a metodologia Laspeyres, o IPCA está sujeito a uma série de vieses metodológicos (diferença entre o custo de vida e o Laspeyres). Nos Estados Unidos há uma série de textos nos quais são discutidas as falhas metodológicas no cálculo da inflação do CPI - *Consumer Price Index*¹⁵, com metodologia semelhante ao IPCA. Há poucos estudos aplicados ao caso brasileiro, apenas Rojas (2008) encontrou um viés de 0,31 ponto percentual para o IPCA de cada ano somente pelo viés de substituição de produtos e serviços no IPCA.

No segundo tipo de enfoque incluem-se os cálculos de núcleos de inflação bastante utilizados por diversos bancos centrais em todo o mundo. É neste último tipo de abordagem que está inserido o presente trabalho. Mais precisamente, analisaremos um modelo alternativo de cálculo de núcleo de inflação para o caso brasileiro.

A literatura não apresenta uma definição clara para núcleo da inflação (Bryan e Cecchetti, 1996), mas de maneira geral há dois tipos de abordagem que contam com diversas formas de mensuração para o núcleo. A primeira, e mais antiga, tende a diferenciar dentro do núcleo as causas dos desvios da inflação no curto prazo com relação à sua taxa no longo prazo. Sendo assim, esta abordagem tende a diferenciar choques de oferta de choques de demanda. Já o segundo, busca apenas separar (sem identificar causas) os movimentos generalizados dos movimentos localizados/idiossincráticos, bem como os choques temporários dos choques permanentes.

A literatura mais recente também atribui a um núcleo eficaz a capacidade de ser um antecedente da inflação no curto prazo. Fiorenco e Moreira (2001) definem as seguintes qualidades de um bom núcleo ainda no segundo campo de abordagem: (i) generalidade, (ii) persistência e (iii) antecipação.

¹⁵Boskin Report (1996), encontrou para o caso do CPI norte americano um viés total de 1,1 ponto percentual ao ano.

Sobre a primeira abordagem de núcleos, Eckstein (1981) foi precursor no uso de uma definição mais detalhada de núcleo da inflação, fazendo um estudo para a inflação americana de 1957 a 1979. Para ele, a inflação pode ser decomposta em três componentes: (i) uma porção atribuída à oferta agregada; (ii) uma porção atribuída à demanda agregada, e (iii) fatores idiossincráticos dos preços. Em suma: $\pi = c + d + s$, em que c são os preços dos fatores de produção, e d é a parcela atrelada à demanda e s são os chamados "choques de oferta"¹⁶ ou fatores específicos (exógenos) de cada setor.

Eckstein em seu livro também reserva um capítulo teórico sobre núcleo da inflação, decompondo as muitas causas da inflação e chamando de "surpresas" o caráter volátil dos índices de preços. Decompor as causas da inflação é o foco central de seu trabalho e, para ele, o núcleo da inflação é definido como a taxa que prevalecerá no longo prazo (Eckstein, 1981, p. 25):

A taxa do núcleo da inflação pode ser vista como a taxa que irá ocorrer em uma economia na trajetória de crescimento de longo prazo; e desta trajetória excluem-se os choques, e o impacto da demanda será neutro pela percepção de que os mercados estão em equilíbrio de longo prazo.(Tradução Própria)

Neste sentido, a contribuição da "inflação de demanda" no longo prazo deve ser nula, pois não haverá desvios da taxa de desemprego efetiva sobre a taxa natural da economia. Assim, o autor considera como núcleo da inflação (no longo prazo) os custos dos fatores de produção¹⁷, pois no longo prazo o impacto da demanda é nulo¹⁸.

Alan Blinder (1982), subsequentemente discutiu o tema dos núcleos utilizando uma linha mais simples que permanece até hoje. Usou uma definição de núcleo pautada na segunda forma de abordagem de núcleos (separação de termos idiossincráticos e movimentos temporários), excluindo alimentos, energia e preços de moradia (*mortgage interest rate*). Entretanto, Blinder observa explicitamente que, mesmo com abordagens distintas, seu núcleo por exclusão pode ser comparado aos "choques" exógenos (" s ") propostos por Eckstein e, portanto, a única diferença aparente de seu núcleo seria os preços das moradias¹⁹ e a não-consideração do efeito da demanda agregada.

¹⁶Aqui definido como restrição temporária em algum determinado setor produtivo.

¹⁷O autor coloca custo unitário do trabalho menos produtividade somado ao custo do capital como sendo a variação dos preços dos fatores.

¹⁸Como lembrado por Blinder (1982), trabalhos anteriores apontavam como sendo o "núcleo da inflação" ou "inflação subjacente" no longo prazo, apenas mudanças nos salários nominais menos a produtividade, diferentemente de Eckstein que (i) inclui o custo do capital e (ii) de certa forma a série de salários são ciclicamente correlacionadas ao considerar os desvios da taxa de desemprego efetiva com a taxa natural.

¹⁹Preços presentes na metodologia utilizada pelo CPI americano

Em suma, está é a literatura de núcleos anterior ao surgimento do regime de metas de inflação, que suscitou novos debates sobre o tema.

Quah e Vahey (1995), supõem que o núcleo da inflação seria *mais do que "mensurar erros"*²⁰, definindo como núcleo, a inflação que no médio e longo prazo não possui impacto no produto (consistente com a curva de Phillips de longo prazo) da economia usando um modelo VAR (vetor autoregressivo). Fazendo uma crítica explícita para mensurações *ad hoc* de núcleos de inflação.

Entretanto, uma das abordagens mais disseminadas na última década foi o uso das médias aparadas com estimadores truncados (simétricas ou não). Proposta por Bryan e Cecchetti (1993b), esta abordagem explora a distribuição das variações dos preços pesquisados pelos índices ao consumidor ao longo da *cross-section* (mensal ou trimestral). Empiricamente, os autores mostram que estas possuem alto grau de assimetria e lepturtose. O estimador (núcleo) de médias aparadas descarta parte das caudas da distribuição. A principal crítica a este método é a dificuldade em especificar o grau ótimo de corte nas caudas.

No Brasil, a discussão sobre núcleos de inflação só ganhou força a partir da implantação do regime de metas. A primeira publicação de um núcleo de inflação foi em 2000, pela Fundação Getúlio Vargas, do Índice de Preços ao Consumidor usando o método das médias aparadas. O Instituto de Pesquisa Econômica Aplicada (IPEA) apresentou resultados semelhantes usando o método das médias aparadas suavizado. Entretanto, há poucos trabalhos aplicados ao caso brasileiro, Fiorenco e Moreira (2002) calculam um núcleo com base na tendência comum dos preços usando um estimador de médias truncadas para a inflação, Picchetti e Toledo (2000) usam o filtro de Kalman (com componentes comuns estáticos) e Figueiredo (2001) faz uma resenha da literatura e estima diversas medidas de núcleos. E Bryan e Cecchetti (2001) propuseram um núcleo de médias aparadas assimétricas para o Brasil. Os autores encontraram um corte ótimo em cerca de 14% da cauda inferior e 10% da cauda superior da distribuição. Entretanto, o núcleo que prevaleceu para a análise da autoridade monetária brasileira foi o por médias aparadas simétricas proposto por Figueiredo (2001), com um corte de 20% nas duas caudas.

A busca por uma medida apropriada para o núcleo da inflação persiste e é uma questão a ser discutida, e como exposto acima, há diversas abordagens distintas de estimação. Mas de forma geral, do ponto de vista teórico, um bom núcleo de inflação deve (i) separar choques localizados dos generalizados, bem como os temporários dos permanentes, e (ii) ainda ser um bom predecessor da inflação.

Como lembrado por Quah e Vahey (1995), grande parte da literatura ²¹ acaba por definir critérios *ad hoc* para a exclusão dos "ruídos" da taxa de inflação. Como é o caso do núcleo por exclusão de alimentos e energia, dado que são preços classificados como tendo

²⁰Tradução Própria.

²¹Brauer e Wu (1991) e Ganley et al. (1991).

um alto grau de componentes sazonais e localizados.

Nossa proposta é a determinação de um núcleo que não seja *ad hoc*, como definido por Quah e Shaun, mas expurgue os componente idiossincráticos implícitos em cada preço do IPCA. Na próxima seção exporemos a metodologia usada pelo Modelo de Fatores Dinâmicos Generalizados. Tal modelo foi aplicado para a estimação de um núcleo de inflação para a Zona do Euro por Cristadoro *et al.* (2003). No trabalho, os autores utilizaram dados desagregados de diversos países pertencentes à citada região. A metodologia aqui apresentada se assemelha às utilizadas pelos autores, embora haja algumas diferenças em relação ao banco de dados usado nos dois casos. Para o Brasil usaremos o IPCA com séries de preços desagregadas, no caso europeu os autores utilizaram dados de preços, atividade, agregados monetários e outros²².

²²Para detalhes ver Cristadoro *et al.* (2003).

3 Modelo de fatores dinâmicos: análise de séries de tempo no domínio da frequência

Há dois tipos de enfoques na análise de séries temporais utilizados de acordo com o objetivo da análise e mostram, cada um, aspectos diferentes das séries temporais. O primeiro enfoque é feito com base na análise no *domínio do tempo* e, neste caso, o interesse está na relação entre as observações em diferentes instantes do tempo e suas magnitudes. A base para esta análise está nas suas **autocorrelações** e funções delas derivadas, sendo assim, neste caso nos deparamos com modelos paramétricos. O segundo tipo de abordagem é feito no *domínio da frequência* e o interesse está na frequência com que certos eventos ocorrem em determinado período de tempo²³. Para tanto, a ferramenta usada é o **espectro** e, desta forma, estaremos diante de um modelo não paramétrico como veremos à frente.

Há dois teoremas de representação de uma série temporal, que basicamente justificam tanto a abordagem no domínio do tempo quanto a do domínio da frequência, são eles o de Wold e de Cramér (para mais detalhes ver Apêndice 1). De maneira resumida e intuitiva, a análise clássica de séries de tempo trata da decomposição de uma série temporal em quatro componentes: ciclo, tendências, aleatório e sazonalidade. Desta forma, mudanças cíclicas são caracterizadas por alterações periódicas, tal que este período sofra apenas pequenas mudanças ao longo do tempo, em contraste com a sazonalidade cujo período é constante. As tendências são os efeitos a longo prazo na média. Assim, a sequência em que certos eventos ocorrem é fundamental na análise de séries de tempo e é sobre este arcabouço teórico de séries de tempo que constituiremos o núcleo nesta seção.

Nesta seção apresentaremos a literatura mais recente de modelos de fatores, destacando algumas de suas aplicações econômicas e algumas de suas variações (seja no domínio do tempo ou da frequência). Na seção seguinte faremos a estimação do Modelo de Fator Dinâmico Generalizado e sua aplicação ao núcleo da inflação brasileira.

3.1 Literatura de modelo de fatores

Não há interesse nas causas das mudanças nas taxas de inflação, pretende-se apenas diferenciar variações generalizadas (macroeconômicas) das localizadas e as variações

²³Ciclos, por exemplo.

temporárias das permanentes. Para tanto, usaremos o modelo de fatores dinâmicos generalizados que, dito de forma intuitiva, divide um grande conjunto de dados em dois componentes distintos (a) os choques comuns (fatores) e (b) os idiossincráticos. Como veremos mais à frente, ao isolarmos estes dois componente e excluirmos também fatores sazonais, chegamos à inflação subjacente, ou seja, aos fatores comuns. Faremos agora uma breve exposição da resenha da literatura do modelo de fatores.

Como destaca Forni *et al.* (2000), para identificar apropriadamente as peculiaridades da economia, um modelo de fatores deve apresentar duas características: (i) ele deve ser dinâmico, ou seja, deve acomodar o fato de que os choques podem afetar as diversas variáveis em períodos diferentes; e (ii) ele deve permitir correlação ao longo dos componentes idiossincráticos, porque a ortogonalidade é uma hipótese não realista para a maioria das aplicações econômicas. O modelo construído por Forni *et al.*, (2000), daqui em diante denominado modelo FHLR, atende a estas duas características.

Um modelo de fator dinâmico foi proposto primeiramente por Sargent e Sims (1977) e Geweke (1977). Forni *et al.* (2000) propõem um modelo de fator com componentes dinâmicos e não ortogonais, de *cross-section* infinita, generalizando os modelos de Chamberlain e Rothschild (1983) e de Sargent e Sim (1977). O modelo FHLR fornece uma condição de identificação propondo um estimador dos componentes comuns que leva a uma convergência, ao mesmo tempo, da *cross-section* e do tempo; com a *cross-section* indo para o infinito. O modelo FHLR se apresenta como uma alternativa interessante à aplicação de um VAR ou VARMA para grandes conjuntos de dados, pois ele comporta a hipótese de que n (o número da *cross-section*) é muito grande e maior que o número T de observações ao longo do tempo.

No quadro adiante temos um resumo do marco teórico do modelo de fatores. De maneira geral, há três hipóteses básicas que podem aparecer diferenciadas em cada modelo. A primeira diz respeito ao choques comuns (fatores). Em um modelo pode-se supor que estes são estáticos ou dinâmicos - se dinâmicos, admitimos que o mesmo choque pode afetar cada série com defasagens distintas; no caso estático, supomos que todas as séries são afetadas pelos choques comuns no mesmo instante. No primeiro caso tratamos a série de tempo no domínio do tempo e, no segundo caso, no domínio da frequência. As segunda e terceira hipóteses que estão sujeitas a adaptações - de acordo com as aplicações teóricas de cada modelo - dizem respeito aos termos idiossincráticos. No caso da segunda hipótese, podemos supor ou não a ortogonalidade entre eles. Já na terceira podemos, ou não, supor a convergência infinita da *cross-section*, ou seja, o modelo acomoda a análise de grandes conjuntos de dados.

Quadro 1: Literatura recente sobre modelos de fatores

Autores	Hipóteses	Tipo de Fator	Hipótese do termo idiossincrático	Hipótese da cross-section
Sargent e Sims (1977) e Geweke (1977)	Primeira proposta do modelo de fatores dinâmicos (domínio da frequência)	dinâmico	ortogonalidade	finita
Chamberlain e Rotschild (1983)	Utiliza uma versão estática do modelo de Sargent e Sim (1977) para séries financeiras	estático	ortogonalidade	infinita
Stock e Watson (1998)	Correlação entre os termos idiossincráticos , mas com componentes principais estáticos (domínio do tempo) para a estimação dos fatores	estático	há correlação entre os termos idiossincráticos	infinita
Forni <i>et al.</i> (2000)	Generalização dos dois primeiros, permitindo a correlação dos termos idiossincráticos com fatores dinâmicos	dinâmico	há correlação entre os termos idiossincráticos	infinita

Fonte: Elaboração própria

O uso do modelo de fatores foi bastante difundido na última década com aplicações destinadas à análise macroeconômica em bases desagregadas (setores ou regiões), como Quah e Sargent (1993), Forni e Reichlin (1996,1997,1998) e Forni e Lippi (1997). Para o Brasil, Rocha (2007) apresenta uma análise dos efeitos diferenciados da política monetária sobre as regiões do país. Cristadoro *et al.* (2003) propuseram a aplicação do modelo FHLR para a estimação de um núcleo de inflação para a União Européia. Usaremos a mesma técnica para o caso brasileiro. Na próxima seção exporemos os procedimentos de estimação.

3.2 Procedimento de estimação do núcleo

Para nossa estimação, considere um vetor de n séries de tempo $x_t = (x_{1t}, x_{2t}, \dots, x_{nt})$. A i -ésima série x_{it} pode ser apresentada como a soma de dois componentes não observáveis²⁴:

$$x_{it} = x_{it}^* + \epsilon_{it} \quad (1)$$

O objetivo é estimar x_{it}^* , por meio dos valores passados e presentes de todos os componentes do referido vetor. Para efeitos do cálculo de nosso núcleo, contudo, não nos interessa o componente ϵ_{it} pois ele retrata somente os erros e choques idiossincráticos de x_{it} e não o componente de longo prazo. O principal objetivo de nossa estimação é garantir que as variáveis do painel de dados possam ser representadas pelo modelo FHLR, ou seja, que cada variável x_{it} , $i = 1, \dots, n$, possa ser representada como a soma de dois

²⁴Esta exposição está baseada em Forni *et al.* (2002) e Cristadoro *et al.* (2003).

componentes não observáveis mutuamente ortogonais: o componente comum, designado por χ_{it} , e o componente idiossincrático, denotado por ξ_{it} . O componente comum é guiado por um pequeno número de fatores (choques) comuns u_{jt} , $j = 1, \dots, q$ (sendo q o número de fatores), que são os mesmos para toda a *cross-section*, mas impactam com defasagens distintas cada variável. Assim, pelo modelo de FHLR teremos que a i -ésima série poderá ser representada como:

$$x_{it} = \chi_{it} + \xi_{it}, \quad (2)$$

em que $\chi_{it} = \sum_{j=1}^q b_{ij}(L)u_{jt}$, $i \in \mathbb{N}$, $t \in \mathbb{Z}$, e L é o operador defasagem.

Seja χ_{1t} uma variável estacionária, de acordo com o teorema de Wold e Cramer, podemos representá-la como uma "soma" de ondas de diversas periodicidades (decomposição espectral). Em particular, podemos decompor χ_{it} em dois componentes, um de curto prazo, χ_{it}^S , e outro de longo prazo, χ_{it}^L , tal que (2) represente a desagregação de suas respectivas ondas periódicas, ou seja:

$$x_{it} = \chi_{it}^S + \chi_{it}^L + \xi_{it}. \quad (3)$$

Considerando (1) e (3) segue que:

$$x_{it} = x_{it}^* + \epsilon_{it} = \chi_{it}^S + \chi_{it}^L + \xi_{it}. \quad (4)$$

O objetivo central da análise é estimar:

$$x_{it}^* = \chi_{it}^L, \quad (5)$$

que nada mais é do que o componente comum de longo prazo da inflação após a eliminação

de $\epsilon_{it} = \chi_{it}^S + \xi_{it}$. Considerando a equação (3), a eliminação de χ_{it}^S por meio da extração de fatores sazonais²⁵ presentes em cada séries permite expressar o componente de longo prazo como $\chi_{it}^L = x_{it} - \xi_{it}$. A idéia de retirar os componentes sazonais e os choques de curto prazo está no fato de que a origem da volatilidade dos dados mensais de inflação, em grande medida, é atribuída à este tipo de ruído. Desse modo, em linha com a literatura de núcleos, nossa idéia de construir um indicador da inflação no médio e longo prazo.

²⁵O procedimento de dessazonalizações será detalhado na próxima seção.

Como dito anteriormente, para a estimação de χ_{1t}^L temos que garantir que o conjunto de dados x_t possa ser representado pelo modelo de FHLR, tal como na equação (2). Forni *et al.* (2002) mostram que isto é possível dado um conjunto de quatro hipóteses, as quais apresentamos, por conveniência, adiante.

Por simplificação notacional, vamos considerar a partir de agora que da variável x_{it} já excluimos a sazonalidades χ_{1t}^S , e a esta altura podemos escrever simplesmente que nosso objetivo será achar:

$$\chi_{it}^L = \chi_{it} = x_{it}^*. \quad (6)$$

Assim de acordo com o modelo de FHLR:

$$x_{it} = \chi_{it} + \xi_{it} = \sum_{j=1}^q b_{ij}(L)u_{jt} + \xi_{it}, \quad (7)$$

ou

$$x_{it} = b_{i1}(L)u_{1t} + b_{i2}(L)u_{2t} + \dots + b_{iq}(L)u_{qt} + \xi_{it}, \quad (8)$$

na qual as variáveis u_{jt} ; com $j = 1, \dots, q$; são os *choques comuns* do modelo especificado por (7) e (8), desta forma χ_{it} é o *componente comum* no qual estamos interessados como descrito em (6), ou seja $x_{it}^* = \chi_{it}^L = \chi_{it}$, que nada mais será do que uma combinação linear dinâmica de q choques comuns não observados do modelo acima descrito. O modelo de FHLR está fundamentado nas quatro premissas que seguem.

HIPÓTESE I

- (1) $u_t = (u_{1t}, u_{2t}, \dots, u_{qt})$ é um ruído branco ortogonal com $E(u_{jt}) = 0$, $Var(u_{jt}) = 1$, para qualquer j e t , $u_{jt} \perp u_{jt-k}$ para qualquer j, t e $k \neq 0$, $u_{jt} \perp u_{st-k}$ para qualquer $s \neq j, t$ e k ;
- (2) $\xi_t^{(i)} = (\xi_{it}, \xi_{2t}, \dots, \xi_{nt})$ é um processo estacionário com $E(\xi_t^{(i)}) = 0$ e $\xi_{it} \perp u_{jt-k}$ para qualquer i, j e k ;
- (3) Os coeficientes dos filtros $b_{ij}(L) = \sum_{k=-\infty}^{\infty} b_{ij}(L)^k$, ou seja, são quadrado somáveis, ou seja, são quadrado somáveis.

Uma consequência da *Hipótese I* é a de que as variáveis do painel serão estacionárias e, portanto, pelos teoremas de Wold e Crámer, $\Sigma^{(n)}(\theta)$, será a matriz de densidade espectral, e pode ser decomposta como a soma da densidade espectral do *componente comum*, $\Sigma_{\chi}^{(n)}(\theta)$, com a densidade espectral do *componente idiossincrático*, $\Sigma_{\xi}^{(i)}(\theta)$.

HIPÓTESE II: sendo $\Sigma^{(n)}(\theta)$ a matriz de densidade espectral do vetor x_t para a frequência θ e $\sigma_{nj}(\theta)$ suas entradas, segue que para qualquer $i \in \mathbb{N}$, existe $c_i > 0$ real tal que $\sigma_{ii}(\theta) \leq c_i$ para qualquer $\theta \in [-\pi, \pi]$. Esta hipótese implica que todas as entradas da matriz de densidade espectral são limitadas em módulo.

Denota-se por λ_{nj} o *autovalor dinâmico* da matriz de densidade espectral, que associa a qualquer $\theta \in [-\pi, \pi]$ o j -ésimo autovalor real não negativo de $\Sigma^{(n)}(\theta)$ em ordem decrescente de magnitude. Da mesma forma, λ_{nj}^{χ} e λ_{nj}^{ξ} são os autovalores dinâmicos de $\Sigma_{\chi}^{(n)}$ e $\Sigma_{\xi}^{(n)}$, respectivamente.

HIPÓTESE III: O primeiro autovalor dinâmico idiossincrático λ_{n1}^{ξ} é uniformemente limitado, isto é, existe um real Λ tal que $\theta \in [-\pi, \pi]$ para qualquer $n \in \mathbb{N}$.

HIPÓTESE IV: Os q primeiros autovalores dinâmicos comuns divergem em $[-\pi, \pi]$, isto é, $\liminf_{\theta \rightarrow \infty} \lambda_{i1}^{\chi} = \infty$, para $j \leq q$ em $[-\pi, \pi]$.

Sob as hipóteses de 1 a 4, o modelo e são chamados de Modelo de Fator Dinâmico Generalizado e Forni et al. (2000) demonstram que é possível obter estimativas consistentes de χ_{it} e ξ_{it} . Feitas as considerações sobre a representação do modelo, vamos passar à questão da identificação dos componente comum e idiossincrático.

Recuperando os Componentes Comuns

Chamaremos de $\Gamma_{\chi}^n(o)$ a matriz de variância e covariância de x_{it} , e $\mu_{\chi j}^n$ seu o j -ésimo autovalor e, dada a matriz densidade espectral $\Sigma^{(n)}(\theta)$ de x_{it} , serão designados como $\lambda_{nj}(\theta)$ seus autovalores, e como $\lambda_{nj}^{\chi}(\theta)$ e $\lambda_{nj}^{\xi}(\theta)$ os j -ésimos autovalores da matriz de densidade espectral dos componentes comuns e idiossincráticos, respectivamente. Vamos assumir que:

- (i) $p_{nj}(\theta)$ é o autovetor de $\Sigma^{(n)}(\theta)$ correspondente ao autovalor $\lambda_{nj}(\theta)$, isto é, $p_{nj}(\theta) \Sigma^{(n)}(\theta) = \lambda_{nj}(\theta) p_{nj}(\theta)$ para qualquer $(\theta) \in [-\pi, \pi]$;
- (ii) $|p_{nj}(\theta)|^2 = 1$ para qualquer j e $\theta \in [-\pi, \pi]$;
- (iii) $\lim_{n \rightarrow \infty} \mu_{\chi r}^n = \infty$, em que $r = q(s + 1)$;
- (iv) $p_{nj}(\theta)$ é mensurável em $[-\pi, \pi]$;
- (v) existe um real Λ no qual $\lambda_{\chi 1}^n(\theta) \leq \Lambda$ para qualquer frequência $\theta \in [-\pi, \pi]$ e qualquer n .

Se as propriedades de (i) a (v) são atendidas, $p_{nj}(\theta)$ é denominado autovetor dinâmico

de $\Sigma^{(n)}(\theta)$, sendo autovetores dinâmicos os autovetores da matriz densidade espectral como função da frequência θ . Dessa forma, a matriz $\Sigma^{(n)}(\theta)$ poderá ser decomposta em termos de seu autovalores e autovetores dinâmicos, ou seja, tais hipóteses são necessárias para que possamos distinguir os termos idiossincráticos dos componentes comuns.

Estimação das covariâncias

Forni *et al.*(2000) propõem o seguinte procedimento de estimação. O primeiro passo é estimar a matriz de densidade espectral $\Sigma^{(n)}(\theta)$ de x_{it} e a covariância dos componentes comuns, a estimação será concluída utilizando-se uma janela de retardamento espectral de Bartlett²⁶. Para um dado inteiro $M=M(T)$, estimaremos a matriz de covariância Γ_{nk}^T de x_{nt} e $x_{n,t-k}$ para $k = 0, 1, \dots, M$ e os $(2M+1)$ pontos da transformada de Fourier para as sequências $\Gamma_{n,-M}^T, \dots, \Gamma_{n0}^T, \dots, \Gamma_{nM}^T$, em que $\Gamma_{n,-k} = \Gamma_{nk}'$. Mais precisamente, computaremos:

$$\Sigma^{(n)}(\theta) = \frac{1}{2\pi} \sum_{k=-M}^M w_k \cdot \Gamma(k) \cdot e^{-i\theta k}, \quad (9)$$

em que:

$$\theta_h = \frac{2\pi h}{(2M+1)} \quad (10)$$

e $h = 0, 1, \dots, 2M$.

Vale ressaltar que no caso de $M = 0$ estaríamos estimando simplesmente os autovalores e autovetores da matriz de variância e covariância de x_{it} , o que significa que o componente principal dinâmico se reduziria ao componente principal do caso estático. A equação (10) nada mais é do que os pesos correspondentes a uma janela de Bartlett de tamanho $M = (\sqrt{T}/4)^{27}$ sugerido pelos autores do modelo FHRL, e os pesos $w_k = 1 - [\frac{|k|}{(M+1)}]$, $k = -M, \dots, M$ para a janela de tamanho M .

O espectro será calculado em $(1 + 2M)$ frequências de espaço no intervalo de $[-\pi, \pi]$, até a frequência $\theta_h = \frac{2\pi h}{2M+1}$, $h = 0, 1, \dots, 2M$ e, então, procedemos à decomposição do componente principal dinâmico. Para cada frequência da rede vamos computar o autovalor e autovetor de $\Sigma^{(n)}(\theta)$. As funções dos autovalores e autovetores - $\lambda_j^\xi(\theta)$ e $U_j(\theta)$, $j = 1, \dots, n$,

²⁶Para detalhes ver Apêndice sobre análise espectral e janelas de retardamento.

²⁷Para detalhes sobre janela espectral ver Apêndice. De forma resumida a janela espectral garante a estimação consistente da matriz de densidade espectral reduzindo a "janela" de estimação dos parâmetros da transformada de Fourier.

respectivamente - são obtidos pela ordenação dos autovalores - em ordem decrescente - de cada frequência. A função $\lambda_{nj}(\theta)$ pode ser interpretada como a densidade espectral do j-ésimo componente principal, e em analogia como componente principal estático teremos que a taxa:

$$p_j = \frac{\int_{-\pi}^{\pi} \lambda_j(\theta) d\theta}{\sum_{j=1}^n \int_{-\pi}^{\pi} \lambda_j(\theta) d\theta} \quad (11)$$

e representa a contribuição do j-ésimo componente principal para a variância total do sistema, a qual chamaremos de *commonality* mais à diante. Seja $\Lambda_q(\theta)$ a matriz diagonal com os autovalores de $\lambda_1(\theta), \dots, \lambda_q(\theta)$ e $U(\theta)$ será a matriz $(n \times q)$ de $(U_1(\theta) \dots U_q(\theta))$, a estimação da densidade espectral do componente comum χ_{it} será dada por:

$$\Sigma_{(n)}^{(\chi)}(\theta) = U(\theta)\Lambda(\theta)U(\theta)'. \quad (12)$$

Dada uma determinada escolha apropriada de q , resultados consistentes serão encontrados pela matriz definida na equação (12) quando n e T vão para o infinito²⁸. Segundo Forni *et al.* (2000) a escolha ótima de q dever ser a quantidade necessária para explicar a variância do conjunto de dados²⁹.

²⁸Detalhes Forni *et al.* (2000).

²⁹A estimação da matriz de densidade espectral dos termos idiossincrático pode ser encontrada para diferença $\Sigma_{(n)}^{(\varepsilon)}(\theta) = \Sigma^{(n)}(\theta) - \Sigma_{(n)}^{(\chi)}(\theta)$.

4 Núcleo da Inflação Brasileira

A maior parte dos trabalhos que envolvem núcleos de inflação utilizam dados (em geral de forma desagregada) de índices de preços ao consumidor. Neste trabalho também utilizaremos como base o Índice de Preços ao Consumidor Amplo (IPCA), pesquisado pelo Instituto Brasileiro de Geografia e Estatística (IBGE).

4.1 Conjunto de Dados

Antes de procedermos a estimação do modelo FHRL, fez-se necessário o tratamento do conjunto de dados. O primeiro "obstáculo" que encontramos foi a descontinuidade dos dados desagregados do IPCA quando encadeados pelas POF (Pesquisa de Orçamento Familiar) de 1995/1996 e 2002/2003, depois tivemos que garantir que o conjunto pudesse ser representado pelo modelo de fatores dinâmicos.

4.1.1 IPCA *reponderado*

O IPCA é composto de 9 grandes grupos de dispêndio, a saber: alimentação, habitação, artigos de residência, vestuário, transportes, saúde e cuidados pessoais, despesas pessoais, educação e comunicação. A maior desagregação do IPCA é o subitem, que é representado pela variação dos preços de 384 produtos. É com este nível máximo de desagregação que iremos trabalhar neste estudo. O grupo com maior número de itens pesquisados é o de alimentação que engloba 159 produtos pesquisados mensalmente pelo IBGE, quase metade dos subitens. Os demais estão divididos entre os outros 8 grupos de dispêndio.

O índice calculado pelo IBGE utiliza a metodologia de Laspeyres Modificado, que supõe uma cesta de bens fixa ao longo do tempo. Para compor a cesta de bens coletados, o IBGE tem como referência a POF, também pesquisada pelo mesmo órgão. A POF pesquisa famílias de 9 regiões metropolitanas³⁰, além de Goiânia e Distrito Federal, para compor a cesta básica de consumo das famílias com rendimentos mensais de até 40 salários mínimos. A última atualização da POF (2002/2003) foi incorporada à pesquisa do IPCA em julho de 2006, antes disto vigorava a cesta de consumo calculada pela POF de 1995/1996.

Assim, se fossemos considerar apenas os dados divulgados pelo IBGE com base na POF de 2002/2003 teríamos apenas dois anos para análise. Dessa forma, optamos por usar uma

³⁰Belém, Belo Horizonte, Curitiba, Fortaleza, Porto Alegre, Recife, Rio de Janeiro, São Paulo e Salvador.

Figura 1: IPCA IBGE *versus* IPCA *reponderado*

versão compatibilizada dos dados divulgados a partir das duas POFs. A compatibilização foi feita usando o tradutor de dados das POFs, disponibilizado pelo IBGE. Com esta compatibilização foi possível estender os dados analisados por mais sete anos, desde agosto de 1999. Enfim, os dados analisados compreendem o período de agosto de 1999 até outubro de 2008.

Com a compatibilização das duas estruturas das POFs, algumas séries de novos produtos que foram incorporados à nova estrutura de ponderação tiveram que ser excluídas do banco de dados. Por exemplo, o item *mão-de-obra* não existia na POF de 1995/1996 e, portanto, não temos sua série de preços desde 1999. Do total de 384 produtos pesquisados pelo IPCA, com o corte feito das séries para as quais não tínhamos os valores pesquisados pela POF de 1995/1996, ficamos com 334 produtos que correspondem a 95% do peso do IPCA. A este "novo IPCA" chamaremos de *IPCA reponderado*, para que não haja confusão com o IPCA geral divulgado oficialmente. A Figura 1 mostra que o nosso *IPCA reponderado* não perde muita informação para o IPCA geral oficial divulgado pelo IBGE. Tendo isto em mente, sem perda de generalidades, as comparações e exercícios posteriores serão feitos com nosso *IPCA reponderado*.

4.1.2 Pré-Tratamento do Dados

Como observado na seção 3 deste trabalho, o modelo de fatores dinâmicos demanda a estimação da matriz de densidade espectral do conjunto. Para tanto é requerido um

pré-tratamento dos dados para garantir a estacionariedade das séries. Ou seja, para chegarmos à especificação da equação (6); que nada mais é do que o componente comum de longo prazo χ_{it}^L da inflação após a eliminação de χ_{it}^S e dos termos idiossincráticos ξ_{it} do modelo FHRL; temos que adotar alguns procedimentos. Em nosso caso o pré-tratamento envolveu quatro passos.

Primeiro, detectamos e removemos *outliers* usando as especificações do Tramo-Seats³¹. De forma geral, a análise gráfica das séries mostra que este procedimento apenas tratou os movimentos mais intensos de preços decorrentes da desvalorização de 2002.

O segundo passo foi retirar, quando presente, o componente sazonal determinístico de cada série. Para isto, utilizou-se também o Tramo-Seats. Assim, concluímos a eliminação do componente χ_{it}^S dos dados. Não utilizamos nenhum procedimento de dessazonalização no conjunto dito preços monitorados, pois estes possuem uma lógica de reajustes bastante particular que não reflete movimentos sazonais. Por exemplo, o reajuste de tarifas de transporte público não possui data nem fórmula de cálculo pré-estabelecida, assim faz sentido que não haja sazonalidade. E mesmo para os que possuem datas específicas (como é o caso de energia elétrica) a forte presença de atrasos na vigoração de reajustes fez com que a sazonalidade ficasse prejudicada, assim tais preços também não foram dessazonalizados.

O terceiro passo foi testar a estacionariedade das 334 séries envolvidas em nossa análise através do teste ADF (*Augmented Dickey-Fuller*) para avaliar a existência, ou não, de raiz unitária. Concluímos que as séries não possuem raiz unitária³², assim, podemos passar para a estimação do Modelo de Fatores Dinâmicos Generalizados.

O quarto passo foi a centralização dos dados na média e a padronização pelo desvio padrão. Naturalmente todo o processo foi feito com as variações de preços.

4.2 Determinação do número de fatores para o núcleo FHLR

Calculamos a matriz de densidade espectral $\Sigma^{(n)}(\theta)$ dos dados x_{it} , após o tratamento descrito anteriormente e, a partir dela, estimamos seus autovalores $\lambda_{nj}(\theta)$ e autovetores de $\Sigma^{(n)}(\theta)$. Uma forma de ver o grau de colinearidade dos n elementos através do chamado grau de "*commonality*" de cada variável como resto do painel. Para tanto, precisamos apenas olhar para os autovalores dinâmicos (q) da densidade espectral de x_{it} e ver quantos são necessários para explicar grande parte da variância total do conjunto. A Figura 2 mostra a variância, explicada pelos $j = 1, \dots, 8$ primeiros autovalores dinâmicos

³¹O método Tramo-Seats também foi utilizado por Cristadoro *et al.*(2003).

³²A metodologia e os resultados dos testes podem ser vistos no Apêndice 2 deste trabalho. Em todos os testes observa-se a não-existência de raiz unitária.

Figura 2: Variâncias explicadas dos primeiro 8 autovalores dinâmicos

Número de autovalores dinâmicos	Variância explicada <i>Commonality</i>
1	0,09
2	0,14
3	0,21
4	0,30
5	0,38
6	0,49
7	0,58
8	0,67

de forma acumulada em percentual da variância total. Já a Figura 3 abaixo, mostra no eixo horizontal o número de séries r utilizadas até $n = 334$. Um dos critérios de seleção propostos por Forni *et al.* (2000) para a definição do número ótimo de autovalores é a observação da inclinação da média dos autovalores para as diferentes frequências apontados pela Figura 3, os primeiros 8 autovalores possuem uma velocidade maior, ou seja, uma inclinação mais positiva, enquanto os demais são quase horizontais. Verificou-se que os primeiros 8 autovalores explicam de forma acumulada 67% da variância total. O resultado para as 334 séries da razão entre a variância do componente comum e a variância total, medindo o grau de *commonality* de cada variável no sistema, está no Apêndice 2, neste caso para $q = 8$. Quase a totalidade das séries apresentou *commonality* acima de 50%, ou seja, grande parte das séries tem componente idiossincrático reduzido. De fato, a média do componente comum de todas as variáveis é 68%, enquanto a do componente idiossincrático é 32%, assim, escolhemos $q = 8$.

Vamos passar para a estimação de nosso núcleo com base no modelo FHLR. Como descrito na seção 3, nosso núcleo deve excluir $\epsilon_{it} = \chi_{it}^S + \xi_{it}$, sendo que a eliminação de χ_{1t}^S foi feita através do procedimento de dessazonalização. Assim nos resta encontrar $x_{it}^* = \chi_{it}^L$, como dito na equação (5), com a eliminação de ξ_{it} .

Pelo modelo FHLR encontramos ξ_{it} (componentes idiossincráticos), o que nos leva a $\epsilon_{it} = \chi_{it}^S + \xi_{it}$. A obtenção de $x_{1t}^* = \chi_{1t}^L$; como descrito na equação (6); foi feita pela diferença vetor de variáveis, $x_t = (x_{1t}, x_{2t}, \dots, x_{nt})$ menos os componentes de curto prazo (ϵ_{it}).

Após o procedimento de eliminação de $\epsilon_{it} = \chi_{it}^S + \xi_{it}$, já podemos reponderar x_{1t}^* ; que este nada mais é do que $x_{it}^* = \chi_{it}^L = x_{it} - \epsilon_{it}$, através da multiplicação dele pela matriz de pesos usada para a construção do IPCA *reponderado* e, assim, encontramos o núcleo

Figura 3: Média dos autovalores dinâmicos para as diferentes frequências

FHLR, ou seja:

$$núcleoFHLR = \sum_{i=1}^n q_t^i x_{it}^* \quad (13)$$

em que: q_t^i é o peso do bem i no período t na cesta do IPCA *reponderado*.

Na Figura 4 observa-se a evolução, tal como extraído do modelo FHLR³³, dos componentes comuns χ_{it}^L e dos termos idiossincráticos ξ_{it} do conjunto de dados, ambos ponderados pela matriz de pesos do IPCA *reponderado*.

Em linha com a teoria, notamos graficamente na comparação entre a evolução do núcleo FHLR e do IPCA *reponderado* uma menor volatilidade do núcleo em relação à inflação cheia, Figura 5. Por outro lado se compararmos o IPCA com o núcleo por exclusão, notamos uma forte semelhança entre os dois (Figura 6). Em termos de desvio padrão, um *ranking* em ordem crescente de volatilidade³⁴ seria: FHLR com a menor, 0,24; em seguida o núcleo por Exclusão com 0,35; e o com maior volatilidade seria o IPCA *reponderado*, com 0,47. Na próxima seção, testaremos à luz da teoria de núcleos a eficiência do núcleo com relação ao núcleo atualmente utilizado pelo Banco Central.

³³Naturalmente, como descrito na seção de pré-tratamento, os dados na geração do modelo FHLR são centralizados e padronização pelo desvio padrão. E, por isso, a evolução são desvios em relação ao ponto zero.

³⁴Aqui mensurada pelo desvio padrão.

Figura 4: Evolução dos componentes idiossincráticos e comuns

Figura 5: IPCA *reponderado* e núcleo FHRL

Figura 6: IPCA *reponderado* e núcleo Exclusão

5 Resultados: teste de cointegração

Nesta seção foi testada a eficiência do núcleo FHLR contra o núcleo por exclusão atualmente calculado pelo Banco Central. A Figura 7 mostra a comparação entre a evolução dos dois núcleos. Como exposto na seção 2 um núcleo, além de diferenciar movimentos idiossincráticos e temporários dos índices de inflação³⁵ deve ser um bom antecedente para a inflação cheia. Testamos a eficiência na antecedência de núcleos por meio de uma cointegração. Será com este instrumental que desenvolveremos esta seção.

A partir do regime de metas, o Banco Central brasileiro tem como objetivo explícito zelar pela estabilidade do nível de preços. Se tomarmos as ações de política monetária como respostas aos desvios da inflação com relação à trajetória das metas, de forma a analisá-las como uma regra de Taylor simples, podemos ter uma idéia da eficiência do núcleo FHLR. A Figura 8 compara a evolução dos ciclos de política monetária do Brasil e os desvios da inflação com relação à trajetória de longo prazo (χ^L) do modelo FHLR, ou seja, os choques comuns que afetaram os preços, tendo eles como consequência uma inflação abaixo ou acima da trajetória de longo prazo. As áreas hachuradas apontam os períodos em que a autoridade monetária promoveu alterações na política econômica por meio de mudanças na taxa de juros (instrumento de política). As áreas vermelhas apontam períodos de contração (elevação da Selic) e as áreas mais claras indicam períodos de expansão monetária (queda da Selic) e, naturalmente as áreas brancas são períodos de Selic constante, tais períodos coincidem com os de inflação acima da trajetória de longo prazo e abaixo, respectivamente.

No último ciclo de afrouxamento monetário brasileiro que durou 25 meses, de outubro

³⁵Procedimento realizado na seção 4.

Figura 7: Núcleo FHLR e núcleo por Exclusão

de 2005 até outubro de 2007, segundo o modelo FHLR, a inflação corrente ficou em média 0,20 ponto percentual abaixo da trajetória de longo prazo. Já no ciclo de aperto monetário anterior, de outubro de 2004 até junho de 2005, a inflação começou a subir antes do início do ciclo. No período entre junho de 2004 e junho de 2005 a inflação ficou em média 0,11 ponto percentual acima da trajetória de longo prazo. Se considerarmos que a trajetória de longo prazo da inflação implícita no modelo de fatores dinâmicos é igual a meta de política para a inflação, ou seja, 4,5%, no primeiro ciclo em questão a inflação estava caminhando em termos anualizados próxima a 2%, ou seja, 0,50 ponto percentual abaixo da banda inferior ao centro da meta³⁶. Este resultado de inflação bem baixa, pode ter justificado o maior ciclo de política monetária desde a implantação do regime de metas. Já no ciclo de alta de 2004/2005, a inflação em termos anualizados estava muito perto de 5,9%, valor bem acima do número central.

A Figura 8 é apenas um indício da eficácia do núcleo FHLR, mas nesta seção foi obtido um teste através da relação de longo prazo (cointegração) entre os índices do IPCA *reponderado* e os dois núcleos considerados neste estudo - a Figura 9 mostra a evolução dos índices.

Em um contexto de séries de tempo, a análise de cointegração é utilizada quando o objetivo é identificar uma relação de equilíbrio estável de longo prazo entre um conjunto de variáveis. A verificação da existência de uma relação de cointegração entre as duas séries de preços e, conseqüentemente, de uma relação estável entre elas, significa que eventuais desvios desta relação de longo prazo possuem caráter transitório. Este procedimento é adequado quando o objetivo é identificar um conjunto de parâmetros na presença de variáveis não estacionárias, isto é, cujas médias não são constantes ao longo do tempo. Mais

³⁶O Banco Central brasileiro possui uma banda de 2% para cima ou para baixo do centro da meta de 4,5% estipulada pelo CMN.

Figura 8: Comparação dos ciclos de política monetária e desvios da inflação de longo prazo pelo modelo FHLR

do que isto, a cointegração permite identificar precedências temporais entre as variáveis cointegradas. Esta última informação será valiosa para corroborar a teoria sobre núcleos de inflação, a qual atribui a eles caráter precedente à inflação corrente.

Para a cointegração usamos as séries dos índices do IPCA *reponderado*, núcleo por Exclusão e FHLR com base em agosto de 1999. O primeiro passo para avaliação da possibilidade de existência de uma tendência comum de longo prazo entre as séries, aferida por um modelo VEC, consiste na realização de testes de raiz unitária. A análise do correlograma das séries e os testes de raiz unitária ADF (Augmented Dickey-Fuller) apontam que as séries são todas não-estacionárias, possuindo uma única raiz unitária³⁷. Em seguida, são escolhidas a partir de alguns critérios as características do modelo VAR a ser estimado. São determinados o número de defasagens e a inclusão ou não de termos deterministas e/ou constantes, dentro e fora do vetor de cointegração. No caso, três dos critérios de seleção³⁸ mostraram que a segunda defasagem é a mais adequada para identificar o processo gerador de dados bivariado entre os dois modelos estimados, (i) IPCA *reponderado* e Núcleo Exclusão e (ii) IPCA *reponderado* e Núcleo FHLR.

A Figura 10 apresenta os resultados dos testes de cointegração³⁹. Ambos os modelos

³⁷O Apêndice 3 com os testes para a cointegração, apresenta os resultados para a identificação de raiz unitária para as séries aqui analisadas.

³⁸Foram aplicados os testes LM e a análise do círculo unitários. Ambos mostraram estabilidade dos parâmetros analisados. O resultado para o teste LM pode ser encontrado no Apêndice 3 sobre os testes de cointegração.

³⁹A hipótese nula do teste do traço é a existência de no máximo r vetores de cointegração, ou seja, dos n autovalores estimados, apenas os r maiores são significantes: $H_0 : \lambda_{r+1} = \lambda_{r+2} = \dots = \lambda_n = 0$. A estatística do teste é dada por $traco = -T \sum_{i=r+1}^n \ln(1 - \hat{\lambda}_i)$, onde T é o tamanho da amostra, r é o número de vetores de cointegração e $\lambda_1, \dots, \lambda_n$ são os n autovalores. Se essa soma for pequena, então

Figura 9: Índices base 100=Ago/1999 do IPCA *reponderado*, Exclusão e FHLR

Figura 10: Testes de Cointegração - Estatísticas do Traço e Auto-Valor

		IPCA e Núcleo Exclusão		IPCA e Núcleo FHRL	
		Estatística	Probabilidade	Estatística	Probabilidade
Estatística do Traço					
Nenhum	14.6	0.02	27.07	0.01	
No máximo 1	2.64	0.12	3.47	0.08	
Estatística do Auto-Valor					
Nenhum	11.96	0.04	23.6	0.01	
No máximo 1	2.64	0.12	3.47	0.08	

selecionados não incluem intercepto dentro do vetor de cointegração uma vez que as séries são índices com o mesmo período-base. Em relação ao número de defasagens, para ambos os modelos foram considerados uma defasada. Os resultados do teste indicam que a hipótese de não-existência de um vetor de cointegração entre as séries foi rejeitada a um nível de significância de 5%. Por outro lado, a hipótese de existência de apenas um vetor de cointegração não foi rejeitada. Assim, com 95% de certeza, podemos dizer que existe uma relação estável de longo prazo entre o índice IPCA e os dois núcleos considerados.

Este é um importante indicativo de que o índice IPCA guarda estreita relação com os núcleos. De acordo com os modelos estimados, as relações entre as séries são positivas e

os autovalores $r+1, r+2, \dots, n$ serão próximos de zero. Este teste é conhecido como teste do traço. O teste do máximo autovalor testa a significância de λ_{r+1} . Este último é um teste que possui maior poder estatístico, pois é um teste conjunto, e o seu resultado foi utilizado, prioritariamente, para a determinação do número de vetores de cointegração.

Figura 11: Vetor de Correção de Erros

Modelo	Núcleo Exclusão		Núcleo FHRL	
	D(IPCA)	D(Núcleo)	D(IPCA)	D(Núcleo)
Coef. de Ajustamento	0.011***	-0.007***	-0.004***	-0.005***
D(IPCA(-1))	0.365*	0.187*	0.572***	0.189***
D(Núcleo(-1))	0.442***	0.519***	0.166	0.312***

Fonte: IBGE. * Significante a 10%; ** a 5%; *** a 1%.

Figura 12: Coeficientes de Longo Prazo

	Exclusão	FHRL
IPCA	1.00	1.00
Núcleo	1.27***	1.56***

Fonte: IBGE. * Significante a 10%; ** a 5%; *** a 1%.

estatisticamente significantes, isto é, ambas variam no mesmo sentido ao longo do tempo. Como observado no início desta seção, a existência de uma relação de cointegração implica que desvios desta relação são transitórios e que há uma dinâmica de ajustamento em direção a esta relação estável. Mais do que isto, os testes mostraram que na hipótese de desvios da trajetória de longo prazo é o índice de inflação também se ajusta para manter a relação de equilíbrio, significando que há uma relação de antecendência do núcleos com relação à inflação cheia.

O coeficientes de longo prazo apresentados na Figura 12 mostram que o núcleo FHRL possui uma maior influência para o IPCA no longo prazo, 1,56, enquanto o núcleo por Exclusão apresentou um coeficiente de 1,27. Entretanto, os coeficientes de ajustamento (curto prazo) mostram uma maior resposta do IPCA *reponderado* ao núcleo por exclusão (Figura 11).

6 Conclusões

Bancos Centrais se deparam com a grande volatilidade nos índices de preços, seja por deficiências metodológicas ou por características idiossincráticas e temporárias de alguns preços. A utilização de núcleos de inflação para uma melhor avaliação dos dados correntes de inflação é fundamental para evitar que a política monetária sob o regime de metas de inflação seja desnecessariamente volátil. A aplicação do método de componentes principais dinâmicos (proposto por Forni *et al.* (2000)) para extrair os componentes comuns do índice de inflação brasileiro (IPCA) mostrou bom desempenho.

O núcleo com base no modelo de fatores foi capaz de extrair movimentos temporários e idiossincráticos. Além disso, há uma relação de equilíbrio de longo prazo identificada pela cointegração do índice de inflação e o núcleo FHLR. Esta relação mostrou uma precedência do FHLR em relação à inflação cheia, como prescreve a teoria. A comparação da relação de longo prazo do núcleo aqui encontrado com o núcleo por Exclusão, mostrou uma maior correlação do FHLR com a inflação cheia.

Os fatores comuns detectados e usados na estimação do núcleo não são identificados neste trabalho, pois não era nosso objetivo identificar os fatores comuns e dar uma interpretação econômica a eles. Assim, uma extensão deste trabalho seria relacionar os fatores com a economia. Poderia-se estimar fatores para cada categoria de preços (serviços, bens duráveis, não duráveis e semiduráveis), por exemplo. Outra interpretação seria analisar os fatores conjuntamente como choques que atingem a economia, como variações cambiais, atividade econômica, etc. A identificação destes choques é fundamental para que a análise de política monetária seja flexível em relação a origens dos choques de forma a atingir as metas pré-estabelecidas.

7 Referências Bibliográficas

Referências

- [1] BERNANKE, B. e MISHKIN F. S. (1997) *Inflation Targeting: a new framework for monetary policy?*, The Journal of Economic Perspective, vol11, n 2 pp. 97-116
- [2] BLINDER, ALLAN (1998) *Central Banking in theory and practice*, Massachusetts Institute of Technology
- [3] BOSKIN, M.J.; DULBERGER, E.R.; GORDON, R.J.; GRILICHES, Z.; JORGENSON, D.W. (1998) *Consumer Prices, the Consumer Price Index, and the Cost of Living. Journal of Economic Perspective*. Vol. 12: p. 3-26,1998.
- [4] BRAUER, G. E WU L. (1991) *In brief economic capsules: an overview of inflation measurement*, Federal Reserve of New York, Summer
- [5] BRYAN, M. E CECCHETTI, G. (1993) *The CPI as a measure of inflation*, Economic Review, Federal Reserve of Cleveland, 29, 15-24
- [6] BRYAN, M. E CECCHETTI, G. (1993) *Measuring core inflation*, National Bureau of Economic Research, Cambridge, MA, Working Paper n 4303
- [7] BRYAN, M. E CECCHETTI g. (2001) *A note on the efficient estimation of inflation in Brazil*, Banco Central do Brasil, Working Papers Series 11
- [8] CECCHETTI, G. (1995). *Inflation Indicators and Inflation Policy*, NBER Macroeconomics Annual, 10, p. 189-219
- [9] CRISTADORO, R., FORNI, Reichlin, L., VERONESE, G. (2003). *A core inflation indicator for the Euro Area*, BCE, 17, p. 1113-41
- [10] ECKSTEIN, O. (1981) *Core inflation*, The Journal of Political Economy, vol. 90. n 6,
- [11] FAVA V,(2007). *A precisão dos índices de preços*, Revista Economia, Brasília(DF), v.8 n.1, p. 39-63
- [12] FUHRER C. JEFFREY,(1997). *Central Bank Independence and Inflation Targeting: Monetary Policy Paradigms for the Next Millenium?*, New England Economic Review.
- [13] FIGUEIREDO, F.M.R. (2001). *Evaluating core inflation measures for Brazil*, Banco Central do Brasil, Working Papers Series, 14
- [14] FIORENCIO, A., MOREIRA, A. (2001). *O núcleo da inflação como a tendência comum dos preços*, Revista Brasileira de Economia, Abr/Jun, p. 175-98

- [15] FORNI, M., HALLIN, M., LIPPI, M. e Reichlin, L. (2000). *The generalized factor model: representation theory*, Econometric Theory, 17, p. 1113-41
- [16] FORNI, M. e REICHLIN, L. (1998). *Let's get real: a factor analytical approach to disaggregated business cycle dynamics*, Review of Economic Studies 65, p. 453-73
- [17] GANLEY, J., HENRY, S. G. B. E ROBINSON, G. N. (1991) '*Measuring inflation*', Bank of England.
- [18] GEWEKE, J. (1977). *The dynamic factor analysis of economic time series*, In Dennis J. Aigner and Arthur S. Goldberger (eds.). Latent Variables in Socio-Economic Models, North-Holland, Amsterdam.
- [19] MANKIW, G. e REIS, R (2002). *Sticky Information versus Sticky Prices: a proposal to replace the Phillips Curve*, The Quarterly Journal of Economics. vol 117, n 4, pp 1295-1328.
- [20] MINELLA, A. (2001). *Monetary policy and inflation in Brazil (1975 - 2000): a VAR estimation*, Revista Brasileira de Economia, 57, p.605-635.
- [21] MORETTIN, P. A. (1979). *Análise Harmônica de Processos Estocásticos*, Rio de Janeiro - Instituto de Matemática Pura e Aplicada.
- [22] QUAH, D. e VAHEY S.P.(1995). *Measuring core inflation*, Economic Journal, vol. 105, n. 432 p. 1130-44
- [23] ORTEGA, T. (2004). *Grandes conjuntos de dados, modelo de fatores e a condução da política monetária*. Dissertação de Mestrado Universidade de São Paulo.
- [24] PAIS, M. B. Z.; PEREIRA, B. B. e SALES P. R. H. (1986) *Análise Espectral de Séries de Tempo: uma introdução para Engenharia, Economia e Estatística*, I Edição - Rio de Janeiro, Arte Final Leasing Editorial Ltda.
- [25] ROCHA, R. (2006). *O mecanismo de transmissão monetária nos Estados brasileiros: uma abordagem do modelo de fatores dinâmicos*. Dissertação de Mestrado Universidade de São Paulo.
- [26] SARGENT, T.J. e SIMS, C.A. (1977). *Business cycle modeling without pretending to have too much a priori economic theory*, In Christopher A. Sims (ed.), New Methods in Business Research, Federal Reserve Bank of Minneapolis.
- [27] SIMS, C. (1992). *Interpreting the macroeconomic time series facts: the effects of monetary policy*. European Economic Review, 36, p. 975-1000.
- [28] SIMS, C., STOCK, J.H e WATSON, M. W. (1990). *Interpreting the macroeconomic time series facts: the effects of monetary policy*. European Economic Review, 36, p. 975-1000.
- [29] NAKANE, M.;ROCHA, F.;TAKEDA,T. (2005). *The reaction of bank lending to monetary policy in Brazil*, Revista Brasileira de Economia, vol.59, n.1.

- [30] Rojas A., (2008) *Viés de substituição de preços do IPCA*, Dissertação de Mestrado, FEA-USP

8 Apêndices

Os apêndices matemáticos expostos a seguir foram retirados de Morettin (1979) e Pais *et al.* (1986). A inclusão destes tem o intuito de facilitar a compreensão do modelo de série de tempo no domínio da frequência (ou análise espectral).

8.1 Análise espectral

Teoremas de representação

Faremos a seguir uma breve exposição dos dois teoremas básicos que fundamentam a análise de séries de tempo, tanto no domínio da frequência quanto no domínio do tempo.

I) O teorema de *Wold* é a base teórica da análise de séries de tempo e, de acordo com ele, todo processo estacionários Y_t admite a seguinte representação:

$$Y_t = D_t + E_t = T_t + S_t + E_t, \quad (14)$$

em que T_t é a componente determinística representada pela tendência, S_t é a componente determinística representada pela sazonalidade e E_t é a componente estocástica definida como um ruído branco, onde:

- i) D_t e E_t são não correlacionados
- ii) D_t é determinístico, pois valores futuros de D_t são previstos com precisão baseados em função linear de valores passados $D_{t-1}, D_{t-2}, \dots$
- iii) E_t é um processo linear discreto geral (ou média móvel infinito, $MA(\infty)$), isto é $E_t = \epsilon_t + \psi_1 \epsilon_{t-1} + \dots$, onde $E(\epsilon_t) = 0$ e $E(\epsilon_t \epsilon_{t+k}) = 0$ para qualquer t e k , ou seja ϵ é ruído branco.

A decomposição⁴⁰ de Wold mostra que é razoável esperar que o verdadeiro processo gerador dos dados possa ser representado adequadamente por um mecanismo que gera um tipo simples de processo determinístico mais um ruído E_t .

⁴⁰A decomposição de Wold foi popularizada em 1970 por Box e Jenkins.

A análise de uma série temporal no domínio da frequência é feita com base em dois teoremas, o de Bochner-Klitchin e o principal é o de Crámer, sejam eles representados por:

II) Teorema de Bochner-Klitchin - a função de autocorrelação $\rho(\tau)$ ⁴¹ de um processo estacionário discreto admite a seguinte representação:

$$\rho(\tau) = \int_{-\pi}^{\pi} \pi e^{i w \tau} dF(w) = \int_0^{\pi} [\text{sen} w \tau dQ(w) + \text{cos} w \tau dC(w)], \quad (15)$$

em que $F(w)$ tem as propriedades de uma função de distribuição acumulada no intervalo $[-\pi, \pi]$, isto é, $F(-\pi)=0$ e $F(\pi)=1$. A notação $dF(w)$ significa integração ou soma conforme $F(w)$ seja contínua ou discreta. A função $F(w)$ é denominada função espectral.

III) Teorema de Crámer: todo processo estacionário discreto no tempo, com função de distribuição espectral $F(w)$, admite a representação:

$$X_t = \int_{-\pi}^{\pi} e^{i w t} dF(w) = \int_0^{\pi} [\text{sen} w t \alpha(w) dw + \text{cos} w t \beta(w) dw], \quad (16)$$

em que $\alpha(w)$ e $\beta(w)$ são funções estocásticas com propriedades análogas a α_i e β_i do processo harmônico (ou série de Fourier) da equação (18) abaixo designada, da mesma podemos expressar a variância de X_t pode ser decomposta:

$$\int_{-\pi}^{\pi} dF(w) = F(\pi). \quad (17)$$

Na realidade, a representação acima é a extensão do processo harmônico ou série de Fourier:

$$X_t = \sum_{i=0}^k [\alpha_i \text{sen} w_i \tau + \beta_i \text{cos} w_i \tau]. \quad (18)$$

⁴¹Seria a autocovariâncias sobre a variância.

Desta forma, a análise espectral trata do estudo da estrutura da série X_t , identificando componentes (senóides) importantes para X_t através da análise e estimação de $F(w)$.

As decomposições de Crámer e Wold usam conceitos de espaço de Hilbert, onde tenta-se expressar qualquer processo estocástico em termos de combinações lineares de variáveis aleatórias não correlacionadas. Nesta abordagem uma série temporal é tratada como um vetor de um espaço infinito-dimensional e pode ser convenientemente estudada pela sua decomposição ortogonal ou canônica, expressa como soma ou integral. No domínio do tempo (paramétrico), os eixos são variáveis aleatórias não correlacionadas, enquanto os coeficientes formam uma função determinista. Já no domínio da frequência (não paramétrico), os eixos são funções deterministas ortogonais, tais como $e_{i\lambda_t}$ e os coeficientes são variáveis aleatórias.

De forma resumida acabamos de expor sem demonstrações o arcabouço teórico da análise de séries de tempo. Primeiramente tem-se o conceito de processo estocástico⁴², desta forma, a primeira suposição que fizemos no teorema de Wold foi a estacionariedade onde, intuitivamente, um processo deste tipo se desenvolve no tempo de modo que sua origem não é importante e não existe mudanças sistemáticas na média e na variância. A partir do teorema de Wold usaremos o arcabouço de processos estacionários para analisar séries de tempo.

Análise de Fourier

A análise de Fourier⁴³ é a aproximação de qualquer função por uma soma de senos e cossenos. Ou seja, na análise de séries temporais no domínio da frequência⁴⁴, o objetivo é aproximar uma função de tempo por uma combinação linear de harmônicos (série de Fourier). De maneira geral podemos dizer que a Análise de Fourier ou Harmônica de uma série temporal consiste em decompor esta série em uma soma de harmônicos cossenos e senos, a partir de registros de frequência (comprimento finito).

Assim, pelo teorema de Fourier qualquer $f(x)$, periódica - com p períodos - pode ser expressa com uma série de Fourier, sendo:

⁴²A introdução do conceito de processo estocástico se dá pelo fato prático de em séries temporais não se pode observar outras realizações, ou seja, não se pode repetir o experimento. Assim, a cada instante t do tempo em geral tem-se somente uma observação $x(t)$ de $X(t)$, Morettin (1979)

⁴³Análogo às séries de Taylor no sentido de que ambas fornecem uma forma simples de expressar funções complexas.

⁴⁴derivadas de processos estocásticos

$$f(x) = \sum_{j=0}^{\infty} \left\{ \alpha_j \cos\left(\frac{2\pi j}{p}x\right) + \beta_j \left(\frac{2\pi j}{p}x\right) \right\}, \quad (19)$$

em que $\alpha_0, \beta_0, \alpha_1, \beta_1, \alpha_2, \beta_2, \dots$ são constantes que podem ser determinadas a partir de $f(x)$.

Diz-se ainda que $x(t)$ é uma série de Fourier de frequência w e amplitude A (para w e A positivos) se:

$$x(t) = A \cos wt \quad (20)$$

ou

$$x(t) = A \sin wt. \quad (21)$$

Como para todo inteiro k :

$$\cos[w(t + (2\pi/w)k)] = \cos(wt + 2\pi k) = \cos wt. \quad (22)$$

A frequência dá o número de ciclos complementares em 2π unidades de tempo; e é chamada frequência angular. Com o mesmo ocorrendo com o seno, podemos notar que a transformada de Fourier de frequência w tem período $T = 2\pi/w$.

Podemos considerar também a frequência em ciclos por unidade de tempo, v onde $v = w/2\pi$ e deste modo $v = 1/T$ (frequência ou número de repetições do ciclo por unidade de tempo). Assim, uma onda cosenoide de período $T = 5$ segundos e $w = 2\pi v = 0,4$ radianos por segundo. Dizer que $v = 0,2$ significa dizer que em um segundo, $1/5$ de um ciclo tem lugar. Outro elemento básico de uma função senoidal é a chamada fase, que dá o deslocamento da primeira onda (seja ela o pico ou o vale) relativamente à origem, se denotarmos ϕ pela fase uma função harmônica pode ser expressa por:

$$x(t) = A \cos(wt + \phi), \quad (23)$$

,

seja $-\infty < t < +\infty$, e como $\cos(wt + \phi) = \cos wt \cos \phi - \sin wt \sin \phi$, se chamarmos $b = -A \sin \phi$, temos:

$$x(t) = a \cos wt + b \sin wt. \quad (24)$$

De modo geral, uma série temporal $x(t)$ na forma acima mencionada pode ser superposicionada na forma de harmônicos com frequências e fases variáveis, ou seja:

$$equation x(t) = \sum_{j=1}^k A_j \cos[w_j t + \phi(w)], \quad (25)$$

em que, $t = 1, 2, \dots, T$ e $k = 2, 3, \dots, [T/2]$, que representa um processo harmônico simples.

Da mesma maneira usando regras de cálculo de seno e cosseno, podemos reescrevê-la como:

$$x(t) = \sum_{j=1}^k (\alpha_j \cos w_j t + \beta_j \sin w_j t), \quad (26)$$

,

Desta forma as representações acima são chamadas representações espectrais de $x(t)$.

Coeficientes de Fourier - Aproximação por mínimos quadrados

Se considerarmos uma função $x(t)$ com período T e aproximarmos por uma combinação linear da forma com um número finito de termos, teremos:

$$x_n(t) = \sum_{n=0}^N \alpha_n \cos w_n + \beta_n \sin w_n t, \quad (27)$$

com $w_n = \frac{2\pi}{T}n$, $n = 0, 1, 2, \dots, N$, é chamado de polinômio harmônico de grau N e período T . O critério que será utilizado é a aproximação de mínimos quadrados, consiste em minimizar quadrados, sendo:

$$\int_0^T [x(t) - x_n(t)]^2 dt \quad (28)$$

Isto significa que devemos encontrar os valores dos coeficientes α_n, β_n , $0 \leq n \leq N$, sendo assim:

$$S(\alpha_n, \beta_n) = \int_0^T [x(t) - \sum_{n=0}^N \alpha_n \cos W_n t + \beta_n \sin w_n t]^2 dt, \quad (29)$$

seja minimizada, considerando-se as derivadas, $n = 0, \dots, N$, chegamos que α e β serão dados por:

$$\alpha_n = 2/T \int_0^T x(t) \cos(n \frac{2\pi}{T} t) dt \quad (30)$$

$$\beta_n = 2/T \int_0^T x(t) \sin(n \frac{2\pi}{T} t) dt \quad (31)$$

Tais expressões são chamadas de coeficientes de Fourier da função $x(t)$, para que as expressões acima valham é necessário que $\int_0^T |x(t)| dt < \infty$. Os coeficientes de Fourier são os mesmos da estimação por mínimos quadrados.

Séries de Fourier

Usando (27) verificamos que:

$$\int_0^T [x(t) - x_N(t)]^2 dt = \int_0^T x^2(t) dt - \sum_{n=0}^N (\alpha_n^2 + \beta_n^2) \quad (32)$$

quando $N \longrightarrow \infty$, podemos escrever:

$$x(t) = \sum_n^{\infty} \alpha_n \cos w_n t + \beta_n \sin w_n t \quad (33)$$

o segundo termo diz-se expansão em série de Fourier de $x(t)$, onde $w_n = 2 \pi / T .n$ e α_n e β_n são dados por (30) e (31).

Forma complexa da série de Fourier

Se usarmos a fórmula de Euler, podemos escrever:

$$x(t) \sim \sum_{n=-\infty}^{\infty} C_n e^{i n \frac{2\pi}{T} t} \quad (34)$$

Janela espectral

É um procedimento para manter a consistência da estimação, considere a seguinte transformada de Fourier da função de autocovariância:

$$\rho(w) = \frac{1}{\pi} [c(0) + 2 \sum_{k=1}^{T-1} c(k) \cos w k], \quad (35)$$

mas a precisão de $c(k)$ diminui quando k cresce, temos que atribuir menor peso a estes valores na medida em que k cresce. Para isto a literatura utiliza as chamadas "janelas espectrais" de retardo e $M(< T)$ é chamado ponto de truncamento. Há uma série de janelas utilizadas, dentre elas as de Tukey, Parzen, mas a sugerida por Forni *et al.* (2000) é a janela de Bartlett. A expressão dada por (35) ficará:

$$\rho(w) = \frac{1}{\pi} [\lambda_0 c(0) + 2 \sum_{k=1}^M \lambda_k c(k) \cos wk], \quad (36)$$

em que λ_k é o peso dado pela janela espectral de retardo e $\lambda_k = 1 + \frac{k}{M}$, $k = 0, 1, \dots, M$.

8.2 Apêndice: testes e resultados para a estimação do modelo FHLR

Teste para a existência de raiz unitária e variância explicada de cada série pelo modelo FHLR

A aplicação do teste Dickey-Fuller Aumentado (ADF) requer dois tipos de especificação acerca do processo gerador de dados. A primeira especificação está relacionada à presença, ou não, de uma tendência determinística e/ou uma constante. A segunda diz respeito ao número de defasagens no modelo. Determinado o modelo, calcula-se a estatística de teste e a compara com o respectivo valor crítico. Rejeita-se a hipótese nula de raiz unitária caso a estatística do teste seja inferior ao valor crítico de Dickey-Fuller:

H_0 : uma raiz unitária

H_1 : série estacionária

Se a estatística do teste for menor que o valor crítico, a presença de uma raiz unitária é rejeitada, e conclui-se que a série é estacionária. Caso contrário, se a estatística do teste for maior que o valor crítico, a hipótese nula não é rejeitada e a série é não estacionária. As tabelas abaixo mostram o resultado do teste de Dickey Fuller para as séries de nosso banco de dados, de acordo com a especificação: com intercepto e com tendência. Os testes levam à rejeição da hipótese nula, isto quer dizer que as séries são estacionárias. A última coluna mostra quanto da variância de cada série é comum às demais pelo modelo FHLR.

Número.	Nomenclatura dada pelo IPCA .	V. crítico ADF	Var explicada FHLR(q=8)
1	Arroz	-4.91	0.742
2	Feijão-mulatinho	-6	0.559
3	Feijão-preto	-4.66	0.635
4	Feijão-macassar	-6.92	0.565
5	Feijão-carioca(rajado)	-6.17	0.566
6	Fava	-8.19	0.555

7	Farinha de arroz	-5.88	0.655
8	Macarrão	-4.35	0.811
9	Fubá de milho	-4.59	0.82
10	Farinha de trigo	-4.46	0.749
11	Farinha vitaminada	-9.78	0.63
12	Farinha de mandioca	-4.68	0.74
13	Massa Semi Preparada	-4.61	0.698
14	Batata-inglesa	-6.69	0.522
15	Inhame	-10.17	0.567
16	Mandioca(aipim)	-4.57	0.63
17	Abóbora	-9.45	0.557
18	Chuchu	-11.13	0.655
19	Pimentão	-8.46	0.544
20	Quiabo	-7.91	0.667
21	Tomate	-8.9	0.624
22	Cebola	-6.82	0.551
23	Cenoura	-7.63	0.606
24	Açúcar refinado	-6.44	0.643
25	Açúcar cristal	-6.15	0.597
26	Chocolate em barra	-4.03	0.747
27	Gelatina de frutas em pó	-8.61	0.639
28	Sorvete	-7.56	0.655
29	Achocolatado em pó	-3.27	0.807
30	Doce de frutas	-10.86	0.605
31	Doce de leite	-5.56	0.624
32	Alface	-10.16	0.653
33	Coentro	-14.61	0.643
34	Couve	-11.58	0.631
35	Couve-flor	-13.33	0.721
36	Repolho	-7.28	0.57
37	Cheiro-verde	-13.01	0.592
38	Brócolis	-15.02	0.693
39	Banana-da-terra	-8.99	0.659
40	Abacaxi	-11.16	0.611
41	Banana-d'agua	-5.28	0.622
42	Banana-maçã	-12.64	0.633
43	Banana-prata	-9.71	0.574
44	Limão	-7.51	0.583
45	Maçã	-6.94	0.536
46	Mamão	-14.15	0.562
47	Manga	-8.77	0.498
48	Maracujá	-8.93	0.57
49	Melância	-8.44	0.53
50	Melão	-10.23	0.491
51	Pera	-9.36	0.613

52	Tangerina	-8.42	0.589
53	Uva	-9.33	0.564
54	Ameixa	-9.23	0.514
55	Laranja-pera	-6.6	0.576
56	Goiaba	-10.83	0.513
57	Fígado	-7.43	0.604
58	Outras Vísceras	-8.22	0.452
59	Carne de porco	-6.06	0.644
60	Contrafilé	-8.39	0.705
61	Filé mignon	-8.18	0.738
62	Chã-de-dentro	-8.01	0.742
63	Alcatra	-8.95	0.703
64	Patinho	-7.39	0.745
65	Lagarto redondo	-8.22	0.647
66	Lagarto plano	-8.97	0.718
67	Músculo	-7.23	0.693
68	Pá	-8.34	0.633
69	Acém	-7.25	0.674
70	Peito	-3.09	0.678
71	Capa de Filé	-11.04	0.555
72	Costela	-8.09	0.666
73	Peixe-anchova	-8.73	0.46
74	Peixe-corvina	-11.43	0.574
75	Peixe	-13.09	0.556
76	Peixe-pescadinha	-12.27	0.591
77	Peixe-tainha	-11.71	0.619
78	Peixe-sardinha	-11.3	0.513
79	Camarão	-13.09	0.574
80	Peixe-vermelho	-13.74	0.614
81	Peixe-cavala	-14.44	0.622
82	Peixe-Linguado	-9.06	0.555
83	Peixe-Cação	-10.95	0.65
84	Peixe-Merluza	-4.99	0.65
85	Peixe-serra	-10.34	0.584
86	Peixe-pargo	-7.92	0.588
87	Peixe-pescada	-15.21	0.546
88	Caranguejo	-13.04	0.61
89	Peixe-Piramutaba	-8.9	0.679
90	Peixe-acará	-9.05	0.545
91	Peixe-dourada	-11.69	0.623
92	Presunto	-3.02	0.755
93	Salsicha e salsichão	-3.94	0.78
94	Lingüiça	-3.79	0.664
95	Mortadela	-4.42	0.715
96	Bacalhau	-4.83	0.7

97	Carne-seca	-8.39	0.617
98	Carne de hamburguer	-8.47	0.64
99	Frango Inteiro	-8.27	0.58
100	Frango em pedaços	-6.33	0.616
101	Ovo de galinha	-5.79	0.743
102	Leite pasteurizado	-7.88	0.63
103	Leite condensado	-4.82	0.653
104	Leite em pó	-4.24	0.686
105	Queijo	-5	0.63
106	Crema de leite	-7.51	0.624
107	Iogurte	-4.63	0.696
108	Manteiga	-4.36	0.781
109	Leite Fermentado	-9.68	0.555
110	Leite com Sabor	-4.91	0.667
111	Biscoito	-3.8	0.839
112	Pão francês	-4.77	0.73
113	Pão doce	-9.03	0.642
114	Pão de forma	-6.8	0.641
115	Bolo	-6.9	0.705
116	Pão de queijo	-11.32	0.582
117	Óleo de soja	-4.82	0.645
118	Azeite de oliva	852	0.66
119	Margarina vegetal	-4.22	0.792
120	Suco de frutas	-4.56	0.708
121	Polpa de açaí	-9.09	0.562
122	Café moído	-5.1	0.719
123	Café solúvel	-3.53	0.681
124	Chá	-9.91	0.608
125	Refrigerante e Água	-7.65	0.703
126	Cerveja	-7.58	0.619
127	Outras bebidas alcoólicas	-10.32	0.573
128	Ervilha em lata	-9.74	0.718
129	Feijoadada em lata	-10.23	0.553
130	Palmito em conserva	-10.22	0.488
131	Sardinha em Conserva	-4.19	0.72
132	Salsicha em Conserva	-6.78	0.557
133	Carne de boi em lata	-6.58	0.692
134	Patê	-4.79	0.483
135	Sopa desidratada	-7.79	0.663
136	Azeitona	-4.29	0.626
137	Milho verde em lata	-3.58	0.72
138	Cogumelo em Conserva	-10.82	0.467
139	Atum em lata	-4.41	0.639
140	Leite de coco	-7.95	0.654
141	Atomatado	-3.89	0.711

142	Alho	-5.91	0.538
143	Sal refinado	-3.66	0.725
144	Colorau	-3.24	0.637
145	Fermento	-7.43	0.632
146	Maionese	-5.19	0.775
147	Vinagre	-6.96	0.669
148	Caldo concentrado	-9.14	0.675
149	Tempero misto	-4.92	0.71
150	Molho de Soja	-4.15	0.679
151	Refeição	-7.72	0.696
152	Lanche	-6.49	0.707
153	Café da manhã	-10.41	0.657
154	Refrigerante	-10.23	0.71
155	Cafezinho	-11.16	0.649
156	Cerveja	-7.14	0.671
157	Chopp	-10.13	0.602
158	Outras bebidas alcoólicas	-5.78	0.711
159	Doces	-9.89	0.596
160	Aluguel residencial	-4.45	0.707
161	Condomínio	-7.18	0.605
162	Taxa de água e esgoto	-8.21	0.569
163	Mudança	-8.78	0.512
164	Ferragens	-4.58	0.664
165	Material de eletricidade	-4.53	0.576
166	Material de pintura	-4.2	0.711
167	Tinta	-3.79	0.735
168	Ferramentas	-12.09	0.591
169	Vassoura	-9.45	0.633
170	Água sanitária	-4.56	0.665
171	Detergente	-7.57	0.716
172	Sabão em pó	-3.78	0.758
173	Desinfetante	-4.98	0.688
174	Inseticida	-12.34	0.616
175	Cera para assoalho	-5.36	0.671
176	Sabão em barra	-4	0.823
177	Esponja de limpeza	-4.58	0.61
178	Limpador com amoníaco	-11.81	0.701
179	Amaciante, alvejante	-5.91	0.711
180	Carvão vegetal	-11.8	0.591
181	Gás de bujão	-10.39	0.635
182	Gás encanado	-10.33	0.629
183	Energia elétrica	-1.33	0.683
184	Móvel para sala	-4.47	0.718
185	Móvel para quarto	-7.4	0.728
186	Móvel para copa e cozinha	-3	0.748

187	Colchão	-5.88	0.752
188	Tapete	-10.88	0.675
189	Cortina	-9.45	0.621
190	Utensílios p/ copa de metal	-9.9	0.639
191	Utensílios p/ copa de louça	-7.56	0.622
192	Flores naturais	-14.24	0.673
193	Roupa de cama	-7.1	0.661
194	Roupa de banho	-5.57	0.674
195	Refrigerador	-5.44	0.73
196	Condicionador de ar	-3.22	0.71
197	Máquina de costura	-10.08	0.555
198	Máquina de lavar roupa	-5.74	0.691
199	Liquidificador	-9.06	0.664
200	Ventilador	-4.86	0.619
201	Fogão	-5.86	0.724
202	Lâmpada	-8.81	0.533
203	Forno de microondas	-5.41	0.776
204	Televisor	-3.67	0.684
205	Aparelho de som	-4.41	0.628
206	Microcomputador	-3.12	0.682
207	Conserto de geladeira	-10.77	0.6
208	Conserto de som	-10.19	0.617
209	Conserto de máq de lavar	-10.95	0.651
210	Reforma de estofado	-10.55	0.648
211	Conserto bomba d'água	-10.48	0.633
212	Calça masculina	-5.25	0.566
213	Terno	-11.5	0.598
214	Agasalho masculino	-12.21	0.552
215	Short/bermuda masculina	-10.4	0.58
216	Cueca	-10.6	0.591
217	Camisa Masculina	-10.3	0.609
218	Calça feminina	-9.49	0.64
219	Agasalho feminino	-9.47	0.49
220	Saia	-7.16	0.642
221	Vestido	-7.48	0.673
222	Blusa	-7.64	0.653
223	Lingerie	-9.43	0.584
224	Roupa de banho feminina	-12.15	0.573
225	Bermuda e short feminino	-10.7	0.631
226	Uniforme	-9.22	0.631
227	Calça comprida infantil	-11.51	0.618
228	Agasalho infantil	-11.67	0.525
229	Vestido infantil	-8.45	0.638
230	Short e bermuda infantl	-9.75	0.536
231	Camisa infantil	-10.84	0.626

232	Fralda	-2.99	0.687
233	Conjunto esportivo infantil	-10.19	0.441
234	Sapato masculino	-9.98	0.52
235	Sapato feminino	-10.19	0.604
236	Sapato infantil	-5.31	0.537
237	Sandália/chinelo masculino	-9.55	0.598
238	Sandália/chinelo feminino	-8.47	0.611
239	Sandália/chinelo infantil	-10.36	0.637
240	Bolsa	-9.7	0.636
241	Tenis	-8.89	0.548
242	Bijuteria	-11.07	0.623
243	Jóia	-5.25	0.644
244	Relógio de pulso	-3.32	0.659
245	Tecido	-4.11	0.683
246	Artigo de armarinho	-9.05	0.617
247	Ônibus urbano	-7.32	0.591
248	Táxi	-9.06	0.573
249	Trem	-9.74	0.689
250	Ônibus intermunicipal	-7.13	0.625
251	Ônibus interestadual	-10.56	0.488
252	Ferry-boat	-8.74	0.53
253	Avião	-7.63	0.614
254	Metrô	-7.31	0.558
255	Barco	-9.11	0.584
256	Transporte escolar	-9	0.657
257	Automóvel novo	-6.97	0.538
258	Emplacamento e licença	-9.39	0.54
259	Seguro de veículo	-5.71	0.61
260	Óleo	-7.04	0.737
261	Acessórios e peças	-7.16	0.694
262	Pneu e câmara-de-ar	-3.71	0.788
263	Conserto de automóvel	-12.11	0.567
264	Estacionamento	-9.06	0.543
265	Pedágio	-9.59	0.558
266	Lubrificação e lavagem	-11.4	0.632
267	Automóvel usado	-4.11	0.605
268	Pintura de veículo	-9.35	0.676
269	Motocicleta	-8.99	0.736
270	Gasolina	-4.43	0.63
271	Álcool	-8.57	0.668
272	Óleo diesel	-6.7	0.722
273	Antiinfecioso e antibiótico	-7.45	0.641
274	Analgésico e antitérmico	-7.12	0.749
275	Antiinflamatório/reumático	-7.36	0.708
276	Antigripal/tussígeno	-4.71	0.777

277	Antimicótico	-8.72	0.739
278	Antialérgico/dilatador	-4.01	0.719
279	Gastroprotetor	-4.67	0.712
280	Vitamina e fortificante	-8.23	0.724
281	Lente de grau	-8.62	0.713
282	Armação de óculos	-12.81	0.609
283	Óculos sem grau	-10.3	0.617
284	Lente de contato	-11.16	0.634
285	Médico	-8.69	0.489
286	Dentista	-9.53	0.588
287	Aparelho dentário	-3.79	0.579
288	Artigos ortopédicos	-12.31	0.619
289	Tratamento psicológico	-10.53	0.599
290	Exame de laboratório	-9.23	0.708
291	Hospitalização e cirurgia	-7.33	0.586
292	Eletrodiagnóstico	-10.34	0.573
293	Radiografia	-3.11	0.561
294	Plano de saúde	-2.89	0.648
295	Produto para Cabelo	-3.63	0.664
296	Produto para Pele	-4.64	0.669
297	Produto Higiene Bucal	-2.89	0.714
298	Produto para unhas	-4.45	0.682
299	Perfume	-8.82	0.553
300	Desodorante	-3.2	0.689
301	Absorvente higiênico	-10.78	0.528
302	Sabonete	-3.1	0.762
303	Papel higiênico	-8.66	0.582
304	Artigo de maquiagem	-9.82	0.557
305	Costureira	-9.47	0.609
306	Manicure e pedicure	-11.42	0.609
307	Barbeiro	-9.33	0.552
308	Cabeleireiro	-9.67	0.679
309	Empregado doméstico	-6.99	0.583
310	Depilação	-11.79	0.607
311	Cartório	-7.47	0.74
312	Despachante	-9.4	0.535
313	Serviço bancário	-9.31	0.477
314	Conselho de classe	-3.4	0.715
315	Cinema	-10.58	0.584
316	Ingresso para jogo	-9.87	0.479
317	Clube	-11.07	0.572
318	Disco Laser	-8.56	0.588
319	Instrumento musical	-9.55	0.622
320	Bicicleta	-8.74	0.652
321	Alimento para cães	-10.1	0.612

322	Brinquedos	-9.56	0.595
323	Aluguel de fita de vídeo	-11.31	0.548
324	Boite, danceteria	-11.44	0.57
325	Jogos lotéricos	-8.87	0.375
326	Motel	-10.47	0.64
327	Material esportivo	-9.89	0.6
328	Hotel	-11.69	0.647
329	Excursões	-10.49	0.553
330	Cigarro	-8.28	0.604
331	Máquina fotográfica	-8.71	0.57
332	Filme e flash descartável	-4.5	0.726
333	Revelação e cópia	-3.59	0.676
334	Creche	-8.16	0.547
335	Educação Infantil	-6.7	0.597
336	Ensino Fundamental	-12	0.606
337	Ensino Médio	-2.6	0.608
338	Ensino Superior	-12.27	0.583
339	Jornal diário	-10.31	0.648
340	Assinatura de jornal	-7.78	0.692
341	Revista não técnica	-6.52	0.7
342	Livro não didático	-6.59	0.687
343	Caderno	-11.66	0.578
344	Fotocópia	-4.54	0.713
345	Artigo de papelaria	-11.7	0.547
346	Correio	-8.34	0.61
347	Telefone fixo	-8.4	0.536
348	Telefone público	-9.22	0.625
349	Telefone celular	-10.25	0.536
350	TV a cabo	-9.23	0.644
351	Aparelho telefônico	-7.32	0.712

Testes para cointegração

Testes de raiz unitária para os índices do IPCA *reponderado*, núcleo Exclusão e núcleo FHRL

As tabelas abaixo mostram o resultado do teste de Dickey Fuller, de acordo com as várias especificações: com intercepto e sem tendência, com intercepto e com tendência e sem tendência e sem intercepto. O critério de defasagem ótima foi obtido através do Critério de Informação de Akaike Modificado (MAIC). Os testes levam a não rejeição da hipótese nula, isto quer dizer que as séries são não-estacionárias.

Figura 13: Teste de Raiz Unitária no logaritmo - probabilidade

	Modelos		
	Intercepto	Intercepto e Tendência	Sem Intercepto e Sem Tendência
IPCA	0.89	0.72	1.00
Núcleo Exclusão	0.93	0.45	1.00
Núcleo FHRL	0.97	0.48	1.00

Fonte: IBGE.

Estatísticas do teste de cointegração

Uma das hipóteses principais assumidas para a realização de testes de cointegração é a de que os erros do modelo são do tipo ruído branco (white noise). Deste modo, para a seleção do número de defasagens para a realização dos testes de cointegração, leva-se em consideração, conjuntamente, critérios de determinação estatística da ordem do modelo auto-regressivo vetorial (VAR) e a inexistência de autocorrelação dos resíduos dos modelos VEC estimados. A conjugação de critérios de informação para seleção das defasagens do VAR (Akaike, Schwarz, Final Prediction Error e Hannan-Quinn) com testes de razão de multiplicador de Lagrange (LM) para a existência de autocorrelação de resíduos fornece um modelo parcimonioso em que os resíduos não são correlacionados, compatíveis com os requisitos da teoria econométrica. A escolha da inclusão ou não de termos deterministas e/ou constantes, dentro e fora do vetor de cointegração, foi inicialmente conduzida por meio de avaliação gráfica, seguida de análise da significância dos coeficientes estimados (testes t e F).

Analisando os critérios para a determinação da ordem do VAR nos modelos que relacionam o índice do IPCA com os núcleos, observa-se que para o modelo com o núcleo de exclusão quatro dos cinco critérios de informação indicam a utilização de 2 defasagens no VAR enquanto para o modelo com o núcleo FHRL três critérios também recomendam a utilização de 2 defasagens no VAR. Deste modo, foi adotado para os modelos VEC uma defasagem.

Figura 14: Critério de Informação - Modelo: IPCA *reponderado* e Núcleo Exclusão

Lag	LogL	LR	FPE	AIC	SC	HQ
0	-560.7272	NA	296.6261	11.36823	11.42065	11.38944
1	-85.52529	921.6036	0.021781	1.848996	2.006276	1.912632
2	-47.02295	73.11556	0.010850*	1.151979*	1.414112*	1.258038*
3	-43.75948	6.065438	0.011016	1.166858	1.533845	1.315342
4	-42.02319	3.156889	0.011537	1.212590	1.684430	1.403497
5	-38.17456	6.842017	0.011582	1.215648	1.792341	1.448979
6	-33.69435	7.783798	0.011484	1.205946	1.887493	1.481701
7	-33.02801	1.130747	0.012305	1.273293	2.059693	1.591472
8	-28.84847	6.923682	0.012287	1.269666	2.160919	1.630269
9	-21.12585	12.48101	0.011428	1.194462	2.190568	1.597488
10	-18.87757	3.542746	0.011881	1.229850	2.330810	1.675300
11	-15.58019	5.062642	0.012101	1.244044	2.449857	1.731918
12	-8.799146	10.13732*	0.011498	1.187862	2.498528	1.718159

* indicates lag order selected by the criterion
 LR: sequential modified LR test statistic (each test at 5% level)
 FPE: Final prediction error
 AIC: Akaike information criterion
 SC: Schwarz information criterion
 HQ: Hannan-Quinn information criterion

Figura 15: Critério de Informação - Modelo: IPCA *reponderado* e Núcleo FHLR

Lag	LogL	LR	FPE	AIC	SC	HQ
0	-643.7182	NA	1392.160	12.91436	12.96647	12.93545
1	-94.38316	1065.710	0.025526	2.007663	2.163973	2.070925
2	-64.70710	56.38452*	0.015276*	1.494142*	1.754659*	1.599578*
3	-64.06204	1.199816	0.016341	1.561241	1.925965	1.708851
4	-63.19590	1.576362	0.017407	1.623918	2.092849	1.813703
5	-62.43239	1.359059	0.018587	1.688648	2.261785	1.920607
6	-58.18199	7.395689	0.018515	1.683640	2.360984	1.957773
7	-57.40664	1.318097	0.019780	1.748133	2.529684	2.064441
8	-52.81348	7.624643	0.019588	1.736270	2.622028	2.094752
9	-50.83516	3.204884	0.020451	1.776703	2.766668	2.177360
10	-49.09615	2.747631	0.021469	1.821923	2.916095	2.264754
11	-48.08077	1.563687	0.022883	1.881615	3.079994	2.366621
12	-43.29988	7.171340	0.022638	1.865998	3.168583	2.393177

* indicates lag order selected by the criterion
 LR: sequential modified LR test statistic (each test at 5% level)
 FPE: Final prediction error
 AIC: Akaike information criterion
 SC: Schwarz information criterion
 HQ: Hannan-Quinn information criterion

8.3 Rotinas para o Matlab

Rotinas utilizadas no MatLab para estimação dos componetes

1- Esta rotina computa a autocovariância da matriz X (dados centralizados e normalizados)⁴⁵

```
function [lag0lead] = autocov(X,maxlag);
ciao=[];
[n,k]=size(X);
centered=center(X);
for i=0:maxlag
    gammina=0;
    for j=1:n-i
        ciao=(centered(j+i,:))'*(centered(j,:));
        gammina=gammina+ciao;
    clear ciao;
    end
    if i==0
        colonna=reshape(gammina,k^2,1);
        lag0lead=(1/n)*colonna;
        clear colonna
    else
        colonnalag=reshape(gammina,k^2,1);
        colonnalead=reshape(gammina',k^2,1);
        lag0lead=[(1/n)*colonnalag lag0lead (1/n)*colonnalead];
        clear gammina;
        clear colonnalag;
        clear colonnalead;
    end
end
```

2- Estima o autovalor e o autovetor dinâmico da matriz espectral de X

```
function [values2,freq2] = dyneig2(X,maxlag);
[spettro2,freq2] = spectra2(X,maxlag);
[n s]=size(X);
vectors2=[];
values2=[];
for j=1:n
    uno=reshape(spettro2(:,j),s,s);
    [eigvect2,eigvalues2]=eig(uno);
    due=diag(eigvalues2);
```

⁴⁵rotinas extraídas do site de Lucrezia Reichlin

```

[tre I]=sort(due);
quattro=flipud(tre);
values2=[values2 quattro];
clear uno egei egeo eigvect2 eigvalues2 due tre quattro
end

```

3- Estima a matrix de densidade espectral dos dados, matriz X

```

function [spettro2,freq2] = spectra2(X,maxlag);
[n k] = size(X);
[lag0lead] = autocov(X,maxlag);
lag0lead = n*lag0lead/(n-1);
freq2=[];
for j=-((n-1)/2):n/2
freq2=[freq2 2*pi*j/n];
end
weight=[];
for k=-maxlag:maxlag
w=1-((abs(k))/(maxlag+1));
weight=[weight w];
clear w;
end
for w=1:n
for k=1:2*maxlag+1
frequency(k,w)=exp(i*(k-maxlag-1)*freq2(w));
end
end
weights=diag(weight);
spettro2=lag0lead*weights*frequency/(2*pi);

```

4- Rotina principal que utiliza as 3 rotinas acima detalhadas

```

dados1=data(:,:);
[W,N] = size(dados1);
maxnlags = round(sqrt(W)/4);
dados2=center(dados1);
dados3=dados2*inv(diag(std(dados2)));
X=dados3;
[values2,freq2] = dyneig2(X,3);
[c v]=size(X);
for d=1:v
integr=2*(pi/c.*((values2(d,:)+conj(values2(d,:)))./2));
somma(d)=sum(integr);
integral(d)=sum(integr)/v;
end

```

```

plot(integral,'*'),grid;
for d=2:v
fac(d-1)=integral(d)-integral(d-1);
end
eig=zeros(351,351);
for y=1:v
[values2,freq3] = dyneig2(X(:,1:y),2);
eig(1:y,y)=mean(values2,2);
clear values2 freq3
end
for u=1:351
hold on
plot(1:351,eig(u,:));
end
hold off

```

5- Rotina para gerar os componentes comuns dado a escolha de q

```

nlaggedcov=q;
nfac=q;
e = kestimat(dados3,nfac,nlaggedcov);
[covc,covi] = covestim(dados3,nfac,nlaggedcov);
commonality(:,1)=diag(covc(:,,nlaggedcov+1))./diag(cov(dados3));
idiosincrco=mean(diag(covi(:,,nlaggedcov+1)));
commonco=sum(diag(covc(:,,nlaggedcov+1)))/351;
commonality2=diag(cov(e))./diag(cov(dados3));

```