

## Estimativas da Curva de Phillips Híbrida Novo Keynesiana para o Brasil: um modelo de autorregressão univariado não causal

Reinaldo Luiz Bedim Junior e Elcyon Caiado Rocha Lima

Universidade do Estado do Rio de Janeiro

### RESUMO

Neste trabalho estima-se uma Curva de Phillips Híbrida, Novo Keynesiana, para o Brasil. Como em Lanne e Luoto (2013) é possível provar, adotando-se a curva de Phillips Híbrida de Galí e Gertler (1999) e hipótese genéricas sobre o hiato do produto, que a taxa de inflação segue um processo autorregressivo não causal univariado. Um modelo autorregressivo não causal univariado, AR(r,s), estabelece a relação entre o valor corrente da variável e seus valores defasados (até “r” defasagens) e futuros (até “s” períodos à frente). Admitindo-se que a taxa de inflação siga tal processo pode-se testar a importância relativa das expectativas sobre a taxa de inflação futura e passada na determinação da taxa de inflação corrente.

Neste artigo foi utilizado o mesmo método de estimação de modelos autoregressivos univariados não causais descritos em Lanne e Luoto (2013). Utilizando-se dados para a taxa de inflação brasileira, medida pelo IPCA, entre 1996 (1º trimestre) e 2014 (3º trimestre) não foi recusada a hipótese de que o seu processo gerador possa ser representado por um modelo autorregressivo não causal univariado. Detectou-se também um componente fortemente antecipativo (*forward-looking*) na dinâmica da inflação brasileira. Tal resultado tem relevantes implicações sobre o papel da autoridade monetária no controle inflacionário, sugerindo que as expectativas são o principal determinante do processo de formação de preços da economia brasileira no período analisado.

Palavras-chave: Curva de Phillips. Modelo autorregressivo não causal. Comportamento *Forward-looking*.

### ABSTRACT

This article estimates a New Keynesian Hybrid Phillips Curve for Brazil. As in Lanne e Luoto (2013) it is possible to prove, adopting the Hybrid Phillips Curve of Galí e Gertler (1999) and very generic assumptions about the output gap's generation process, that the inflation rate follows a univariate noncausal autoregressive process. A Noncausal Autoregressive Model, AR(r,s), establishes a relationship between the current value of the variable and its “r” lags and “s” leads. Admitting that the inflation rate follows such a process allows for the testing of the magnitudes of the forward-looking and backward-looking components in inflation dynamics.

This paper uses the same method of estimation of noncausal autoregressions described in Lanne e Luoto (2013). The hypothesis that the Brazilian inflation rate, measured by the IPCA index, follows a univariate noncausal autoregressive process was not rejected by quarterly data from 1996-Q1 to 2014-Q3. The estimated model also found a strong forward-looking component in Brazilian inflation rate dynamics in this period. This latter result has important implications for the role of monetary authority on inflation targeting in Brazil, suggesting that inflation expectations were the main factor impacting the Brazilian prices dynamics during the analyzed period.

Keywords: Phillips Curve. Noncausal Autoregressive Model. Forward-looking behavior.

JEL: C13; C22; E31

# 1. INTRODUÇÃO

Neste trabalho estima-se uma Curva de Phillips Híbrida Novo Keynesiana para o Brasil. Como em Lanne e Luoto (2013) é possível provar, adotando-se a curva de Phillips Híbrida de Galí e Gertler (1999) e hipótese genéricas sobre o hiato do produto, que a taxa de inflação segue um processo autorregressivo não causal univariado. Um modelo autorregressivo não causal univariado, AR(r,s), estabelece a relação entre o valor corrente da variável e seus valores defasados (até “r” defasagens) e futuros (até “s” períodos à frente). Admitindo-se que a taxa de inflação siga tal processo pode-se testar a importância relativa das expectativas sobre a taxa de inflação futura e passada na determinação da taxa de inflação corrente.

A obtenção de estimadores de máxima verossimilhança (MV) para os parâmetros dos modelos autorregressivos não causais foi explorada inicialmente por Breidt et al (1991) e depois por Andrews et al (2006) e Lanne e Saikkonen (2010). Neste artigo adotamos os mesmos procedimentos de estimação descritos em Lanne e Luoto (2013). Foram utilizados dados para a taxa de inflação brasileira, medida pelo IPCA, entre 1996 (1º trimestre) e 2014 (3º trimestre).

Este artigo está organizado da seguinte forma. Além desta introdução, o trabalho está dividido em mais quatro seções. Na segunda seção, é realizada uma revisão da literatura existente sobre a Curva de Phillips Novo Keynesiana. São explorados os aspectos da evolução teórica dos temas bem como alguns dos principais resultados empíricos existentes. Na terceira seção, são apresentados os detalhes dos modelos autorregressivos não causais. É realizada, também, a dedução da Função de Verossimilhança que será utilizada nas estimativas. A quarta seção apresenta os principais resultados estimados. Por fim, concluímos na última seção.

## 2. A CURVA DE PHILLIPS

### 2.1 A Curva de Phillips Novo Keynesiana

A Curva de Phillips Novo Keynesiana (CPNK), adota expectativas racionais e um comportamento *forward-looking* dos agentes. Entretanto, é introduzido um grau de rigidez de preços. Maka (2013), em um bom resumo deste tema, descreve as equações provenientes de um modelo *Dynamic Stochastic General Equilibrium* – DSGE com rigidez de preços, desenvolvido por Calvo (1983). Nesse modelo, é considerada a seguinte regra de ajuste de preços:

$$\rho_t = \omega \rho_{t-1} + (1 - \omega) z_t \quad (2.1)$$

onde:  $\omega$  é a proporção de firmas que não reajustam preços em função de alguma rigidez nominal existente e, portanto, cobram em  $t$  valores de  $t - 1$ ; e,  $z_t$  é a proporção de firmas que reajustam preços e  $\rho_t$  é o nível geral dos preços. Admitindo um mercado de competição monopolística sem fricções, as firmas estabeleceriam seus preços a partir de um *mark-up* fixo, acima dos seus respectivos custos marginais. Essas firmas utilizariam a seguinte regra de decisão para o reajuste de preços:

$$z_t = (1 - \omega\beta) \sum_{j=0}^{\infty} (\omega\beta)^j E_t (MC_{t+j} + \mu) \quad (2.2)$$

onde:  $\mu$  é o *mark-up* ótimo;  $MC_t$  é o custo marginal nominal; e,  $\beta$  é o fator de desconto das firmas ( $0 < \beta < 1$ ). As equações (2.1) e (2.2) podem ser combinadas e, após algumas manipulações algébricas, formam a versão da Curva de Phillips Novo Keynesiana – CPNK.

$$\pi_t = \beta E_t \pi_{t+1} + \frac{(1-\omega)(1-\omega\beta)}{\omega} (mc_t + \mu) \quad (2.3)$$

A equação (2.3) relaciona, então, a inflação no período corrente, a inflação no período posterior e os desvios dos custos marginais reais acrescidos do *mark-up* ótimo em relação ao seu valor potencial. Ainda, com apontado por Maka (2013), sob condições relativamente gerais, os custos marginais agregados são proporcionais ao hiato do produto.

Assumindo essas condições, é possível reescrever a CPNK como:

$$\pi_t = \beta E_t \pi_{t+1} + \gamma x_t \quad (2.4)$$

onde:  $x_t$  é o hiato do produto. Nesse caso, o hiato de produto está substituindo o hiato de emprego das formulações anteriores e  $x_t$  pode ser interpretado como uma taxa de crescimento do produto que não acelera a inflação. Essa formulação é muito semelhante à utilizada por Lanne e Luoto (2013), que será a base, uma vez acrescida da transformação realizada na seção 4.4, da modelagem da inflação deste trabalho.

O modelo da CPNK, postula que somente as expectativas sobre a inflação futura afetariam a inflação presente. Entretanto, segundo Maka (2013), muitas estimativas indicaram influência, também, da inflação passada. Ou seja, haveria uma inércia inflacionária no comportamento dos preços.

Considerando essa percepção, Galí e Gertler (1999) desenvolveram um modelo, também deduzido a partir de um DSGE, que permite que uma fração das firmas adote um comportamento *backward-looking* para ajustar os seus preços, deixando a hipótese de um comportamento totalmente *forward-looking* como um caso particular. A partir do modelo obtém-se uma curva de Phillips novo keynesiana que também apresenta um comportamento *backward-looking*, o que gerou um modelo híbrido.

$$\pi_t = \omega \pi_{t-1} + (1 - \omega) \beta E_t \pi_{t+1} + \gamma x_t \quad (2.5)$$

A equação (2.5) é o modelo básico que será testado para a economia brasileira neste trabalho. Um ponto importante a ser destacado é que os modelos novo keynesianos, híbridos ou não, deduzidos a partir de DSGEs, apontam o custo marginal como variável teoricamente apropriada para estimativas de Curvas de Phillips. Entretanto, o custo marginal agregado da economia é de difícil mensuração, ensejando a utilização de variáveis aproximadas. Assim, é comum que diferentes estimativas utilizem diferentes medidas de hiato do produto em seus modelos, o que parece provocar discrepâncias nos resultados obtidos.

## 2.2 Principais evidências empíricas sobre a Curva de Phillips Híbrida (CPH)

Um dos primeiros trabalhos que buscou estimar uma CPH para o caso dos EUA foi Galí e Gertler (1999). Nesse estudo seminal, os autores estimaram, primeiramente, uma versão novo keynesiana, conforme a equação (2.4) e, depois, uma versão híbrida, conforme a equação (2.5). Em ambos os casos, foi utilizada a método dos momentos generalizados, *Generalized Method of Moments* – GMM. Os resultados apontaram para uma predominância do comportamento *forward-looking*, embora o componente *backward-looking* tenha sido estatisticamente relevante.

Esse trabalho possui ainda um resultado intermediário importante para a discussão ora proposta. Quando da estimação da CPNK, os autores testaram tanto o hiato de produto<sup>1</sup> diretamente quanto uma medida aproximada do custo marginal<sup>2</sup>. Enquanto a estimativa utilizando o custo marginal apresentou-se estatisticamente relevante e teoricamente consistente, a estimativa que utilizou o hiato do produto se mostrou contraditória, apresentando um sinal oposto ao que seria esperado pela teoria. Motivo pelo qual os autores concluíram ser o custo marginal (aproximado pela participação dos salários na economia) uma medida mais apropriada para estimativas dessa natureza.

Rudd e Whelan (2005), entretanto, criticaram os pressupostos adotados nas estimativas realizadas por Galí e Gertler (2000), bem como a aplicação do método GMM. Primeiramente, os autores rejeitaram a adoção da participação dos salários na renda (*labour share*) como uma *proxy* para os custos marginais. Criticaram, ainda, a Teoria das Expectativas Racionais como regra para formação dos preços

<sup>1</sup> O hiato do produto foi calculado como uma diferença entre o valor da variável (PIB) e sua tendência.

<sup>2</sup> *Since real marginal cost is not directly observable, we use restrictions from theory to derive a measure based on observables.* (GALÍ & GERTLER, 2000, p.8). Os autores introduzem como proxy do custo marginal a participação dos salários na economia (*labour share*).

pelos agentes. Por fim, os autores concluem que os modelos CPNK que utilizam a participação dos salários na renda como *proxy* dos custos marginais oferecem pouca explicação sobre a dinâmica dos preços. Ainda, os modelos híbridos forneceriam pouca evidência a respeito da importância do componente *forward-looking* no comportamento da inflação.

Galí et al (2005), no entanto, reafirmaram as conclusões de Galí e Gertler (2000). Os autores voltaram a testar o hiato do produto diretamente nos modelos em alternativa a participação dos salários e, novamente, concluíram pela melhor explicação desse. Além disso, Galí et al (2005) ainda testaram métodos de estimação alternativos<sup>3</sup> e aferiram resultados aderentes aos estimados por GMM, o que atestaria a robustez desse método.

Outro enfoque de investigação foi desenvolvido por Monacelli (2005). Nesse trabalho, o autor estima um DSGE para uma economia pequena e introduz o comportamento cambial em uma CPNK. Monacelli (2005) concluiu que a taxa de câmbio tem um papel significativo na dinâmica do modelo estimado.

Outras técnicas econométricas foram utilizadas para estudos nessa área. Fanelli (2005), por exemplo, testou uma CPH desenvolvida por Galí e Gertler (2000), para a Zona do Euro, utilizando uma modelagem por Vetores Autorregressivos (VAR) e concluiu que este modelo híbrido não apresentou uma boa aproximação para o comportamento da inflação naquela região.

De modo similar, existe uma série de trabalhos que buscaram testar, para o caso do Brasil, o arcabouço contido na CPH. Areosa e Medeiros (2007) testaram a importância do componente cambial para uma economia pequena e aberta (conforme Monacelli (2005)) *vis-à-vis* uma economia fechada. Concluíram que o componente cambial é estatisticamente significativo para o caso do Brasil e que a introdução do setor externo aumenta a importância do componente *forward-looking* nas estimativas.

Seguindo uma linha de investigação similar, Holland e Mori (2010) estimaram uma CPH para o Brasil com o intuito de testar o impacto da globalização dos mercados na dinâmica interna da inflação. Os autores concluíram que a introdução dos “fatores globais” reduziria a importância do hiato de produto utilizado nos modelos.

Divino e Mazali (2010) estimaram uma CPH, baseados em um modelo com rigidez de salários reais, buscando explicar o *trade-off* de curto prazo entre inflação e atividade econômica. Concluíram haver elevada importância do componente *backward-looking* e, consequentemente, um custo deflacionário (“custo social”) significativo para o caso do Brasil.

Fica evidente, portanto, que há relativa variedade de técnicas utilizadas para cercar este assunto. Em todos os estudos apresentados, no entanto, a seleção da variável de hiato de produto e da técnica de estimação que serão utilizadas parece influenciar sobremaneira nos resultados auferidos. Ainda, o custo marginal da economia, que seria a variável teoricamente mais apropriada, não pode ser observado diretamente e nem aproximado de maneira simples.

No caso brasileiro, o imbróglio se torna ainda mais complexo em função dos resultados estimados por Monacelli (2005). Nesse estudo, o autor afirma que, em economias abertas e pequenas, que é a realidade brasileira, a taxa de câmbio se torna bastante influente na dinâmica inflacionária. Tal resultado parece ser corroborado tanto por Areosa e Medeiros (2007) quanto por Holland e Mori (2010).

Ainda, conforme exemplificado pelo trabalho de Divino e Mazali (2010), existem estudos divergentes sobre o peso dos componentes *backward-looking* e *forward-looking* para o caso do Brasil. Tal divergência pode ser originária tanto das técnicas de estimação quanto das medidas de hiato de produto utilizadas nas modelagens.

---

<sup>3</sup> Os autores testam os seguintes métodos: Mínimos Quadrados não Lineares, Mínimos Quadrados não Lineares com variáveis instrumentais e de Maximização de Verossimilhança.

Neste trabalho, foi adotada uma técnica de autorregressão não causal. Essa metodologia, como será visto nas próximas seções, vem sendo crescentemente utilizada para estudos econômicos e ainda não foi aplicada à realidade brasileira. Uma das principais vantagens dessa técnica é a dispensa da escolha, *a priori*, da variável hiato do produto para a estimação do peso dos componentes *backward-looking* e *forward-looking*.

### 3 OS MODELOS AUTOREGRESSIVOS

#### 3.1 Modelos Autorregressivos não Causais

Os modelos autorregressivos são utilizados, usualmente, para descrever a dinâmica de séries de tempo em economia. Aplicações empíricas típicas incluem de mensuração de persistência de séries a modelagens macroeconômicas complexas. Entretanto, a maior parte dessas aplicações se restringe aos modelos autorregressivos causais, nos quais a variável independente é função de valores defasados dela mesma (LANNE & SAIKKONEN, 2010).

Neste trabalho, será utilizada uma abordagem na qual é permitida a não causalidade de um modelo autorregressivo. Ou seja, além de estabelecer uma relação com valores passados da variável, será estabelecida, também, uma relação com os valores futuros dela própria.

*We argue that this is a particularly useful feature in economic applications where expectations play a central role. Another interesting feature of the noncausal autoregressive model is that its errors are predictable by past values of the series which may be made use of in improving forecast accuracy if noncausality is detected. An interpretation of such predictability is that the errors contain effects of omitted variables that are predictable by the considered series* (LANNE & SAIKKONEN, 2010, p.2).

Em outras palavras, a introdução da não causalidade pode corrigir erros de especificação do modelo selecionado, uma vez que esse pode omitir variáveis determinantes para o processo e que o econometrista não possui capacidade de avaliar corretamente. Embora a aplicação dessa técnica seja relativamente recente e ainda pouco utilizada na área econômica, existem diversos trabalhos que desenvolveram o tema estatisticamente.

Breidt et al (1991), em artigo seminal, propôs um método estatístico para a construção de uma função de verossimilhança aproximada para modelos autorregressivos não causais, além de realizar simulações para um processo  $AR(0,2)$ . Breidt et al (2001) utilizou essa mesma metodologia aplicada ao comportamento dos preços das ações da *Microsoft* em um período de três anos. De maneira semelhante, Lii e Roseblatt (1996) estudaram a função de verossimilhança para modelos ARMAs não causais, simulando um  $ARMA(1,1)$ . Andrews et al (2006) estudou o mesmo método sugerido por Breidt et al (1991), fazendo algumas alterações, que serão detalhadas ao longo deste trabalho. Andrews et al (2009) estimaram, ainda, um processo não causal para a série de preços das ações do Wal-Mart.

No entanto, a interpretação, como será dada neste trabalho, da solução não causal como uma possível mensuração das expectativas econômicas, foi sugerida por Hansen e Sargent (1991). Ainda, a aplicação dessa modelagem especificamente para a estimação de uma CPH foi explorada por Lanne e Saikkonen (2010) e Lanne e Luoto (2013). Ambos os trabalhos abordaram a realidade americana e produziram resultados que reforçaram a percepção da importância do componente *forward-looking* na dinâmica da inflação dos Estados Unidos.

Ressalte-se, ainda, que a literatura sobre modelos não causais aplicados à economia, embora ainda escassa, está ganhando volume recentemente. Leeper et al (2012) modelaram um VARMA não causal para estimar efeitos de políticas fiscais antecipadas pelos agentes.

Gourieroux e Jasiak (2015) propuseram uma metodologia de estimação distinta da proposta por Breidt et al (1991) por meio da construção de um VAR não causal semi paramétrico. Nesse trabalho, os autores sugeriram, inclusive, a possibilidade de identificação entre defasagens e valores à frente mesmo

em modelos cujos resíduos sejam gaussianos. Esse ponto será mais detalhadamente abordado nas seções que se seguem. Por ora, registre-se, apenas, que a metodologia desenvolvida por Breidt et al (1991), que será, em grande medida, utilizada neste trabalho, carece da não normalidade dos resíduos para a identificação entre ordem de defasagens e valores futuros. Lanne e Saikkonen (2010) estimaram um modelo puramente não causal para a inflação americana, enquanto que Lanne e Luoto (2013) estimaram uma prevalência do componente *forward-looking* de uma CPH utilizando uma estimativa não causal também para a economia americana.

Este trabalho, até onde alcançou a pesquisa realizada, é o primeiro a utilizar um modelo não causal para estimar uma CPH para a realidade do Brasil. Utilizaremos, em acordo com a bibliografia, a função de verossimilhança aproximada introduzida por Breidt et al (1991) e alterada por Andrews et al (2006) e Lanne e Saikkonen (2010). Para a modelagem da inflação, será utilizado, como base, Lanne e Luoto (2013). As próximas seções detalham ambos os procedimentos e os resultados aferidos.

### 3.2 O modelo Autorregressivo Não Causal – ARNC

Esta seção foi retirada, em grande medida, de Lanne e Saikkonen (2010). Considere um processo estocástico de  $y_t$  ( $t = 0, \pm 1, \pm 2, \dots$ ) gerado pela seguinte equação:

$$\phi(B)\varphi(B^{-1})y_t = \varepsilon_t \quad (3.1)$$

onde:  $\phi(B) = 1 - \phi_1 B - \dots - \phi_r B^r$ ;  $\varphi(B^{-1}) = 1 - \varphi_1 B^{-1} - \dots - \varphi_s B^{-s}$ ; e,  $\varepsilon_t$  é uma sequência de resíduos independentes e identicamente distribuídos – iid, contínuos, com média igual a 0 e variância igual a  $\sigma^2$ . Ou, de outra forma  $\varepsilon_t \sim \text{iid}(0, \sigma^2)$ .

Em complemento,  $B$  é o operador de defasagens usual. Ou seja,  $B^k y_t = y_{t-k}$  ( $k = 0, \pm 1, \pm 2, \dots$ ) e  $B^{-k} y_t = y_{t+k}$  ( $k = 0, \pm 1, \pm 2, \dots$ ). Além disso, os polinômios  $\phi(z)$  e  $\varphi(z)$  possuem raízes fora do círculo unitário.

$$\phi(z) \neq 0 \text{ para } |z| \leq 1 \text{ e } \varphi(z) \neq 0 \text{ para } |z| \leq 1 \quad (3.2)$$

Se  $\varphi(z) \neq 0$  para algum  $j \in \{1, \dots, s\}$  e  $\phi_1 = \phi_2 = \dots = \phi_r = 0$ , o modelo será puramente não causal. Por outro lado, se  $\phi(z) \neq 0$  para algum  $j \in \{1, \dots, s\}$  e  $\varphi_1 = \varphi_2 = \dots = \varphi_s = 0$ , o modelo será puramente causal.

Partindo da equação (3.1) e fazendo  $\varphi(B^{-1})y_t = \mu_t$ , será obtida a seguinte representação em médias móveis (MA):

$$\begin{aligned} \phi(B)\varphi(B^{-1})y_t &= \varepsilon_t \\ \phi(B)u_t &= \varepsilon_t \\ u_t &= \varepsilon_t / (1 - \phi_1 B - \dots - \phi_r B^r) \\ u_t &= \sum_{j=0}^{\infty} \alpha_j \varepsilon_{t-j} \end{aligned} \quad (3.3)$$

onde:  $\alpha_0 = 1$  e  $\alpha_j$  decresce geometricamente quanto  $j$  vai para o infinito. De forma semelhante, pode-se fazer  $\phi(B)y_t = v_t$  e obter a seguinte equação:

$$\begin{aligned} \phi(B)\varphi(B^{-1})y_t &= \varepsilon_t \\ \varphi(B^{-1})v_t &= \varepsilon_t \\ v_t &= \varepsilon_t / (1 - \varphi_1 B^{-1} - \dots - \varphi_s B^{-s}) \\ v_t &= \sum_{j=s}^{\infty} \beta_j \varepsilon_{t+j} \end{aligned} \quad (3.4)$$

onde:  $\beta_0 = 1$  e  $\beta_j$  decresce geometricamente quanto  $j$  vai para o infinito. Em ambos os casos, nas equações (3.3) e (3.4), a estacionaridade descrita é garantida somente se as condições assinaladas em

(3.2) forem atendidas. Do mesmo modo, conforme Lanne e Saikkonen (2010), o próprio processo descrito pela equação (3.1) pode ser escrito como um processo de médias móveis em duas direções.

$$\phi(B)\varphi(B^{-1})y_t = \varepsilon_t \quad (3.1)$$

$$y_t = \frac{\varepsilon_t}{\phi(B)\varphi(B^{-1})}$$

$$y_t = \sum_{j=-\infty}^{\infty} \psi_j \varepsilon_{t-j} \quad (3.5)$$

onde:  $\psi_j$  é o coeficiente de  $z^j$  em uma expansão da série de Laurent do polinômio  $[\phi(z)^{-1}\varphi(z^{-1})^{-1}] = \psi_j z^j$  (Breidt et al, 1991) e  $\phi(z)^{-1}\varphi(z^{-1})^{-1} \stackrel{\text{def}}{=} \psi(z)$ . Especificamente, pela condição de (2.2), obtemos:

$$\psi(z) = \sum_{j=-\infty}^{\infty} \psi_j z^j$$

existente para todo  $b \neq 0$  tal que  $b < |z| < b^{-1}$  e  $b < 1$ . As equações (3.3) (puramente causal) e (3.4) (puramente não causal) são, portanto, casos particulares da equação (3.5).

É utilizada a abreviação  $AR(r, s)$ , onde  $r$  é o número de defasagens e  $s$  de valores à frente, para representar simplificada o modelo definido em (3.1). Note-se que, para  $s = 0$ , temos o modelo  $AR(r)$ , que é o modelo autorregressivo causal no qual há dependência apenas dos valores passados. No caso contrário, para  $r = 0$ , temos o  $AR(0, s)$ , que é um modelo puramente não causal, no qual os valores da variável dependem apenas dos valores dos períodos posteriores.

Dada a definição do processo  $v_t$  e tomando a esperança não condicional em ambos os lados da equação (3.1), fica evidente que os valores presentes da variável dependente ( $y_t$ ) dependem dos erros futuros esperados:

$$y_t = \phi_1 y_{t-1} + \dots + \phi_r y_{t-r} + \sum_{j=s}^{\infty} \beta_j E_t(\varepsilon_{t+j}) \quad (3.6)$$

No caso de um modelo puramente causal,  $\beta_j = 0$ , para todo  $j > 0$ , o último termo da equação seria nulo, o que significaria que erros futuros de previsão não afetariam o valor presente de  $y_t$ . Nesse caso, o modelo poderia ser estimado por Mínimos Quadrados Ordinários (MQO). Entretanto, considerando  $\beta_j \neq 0$ , o último termo não é nulo em modelos com componente não causal e, portanto, a utilização de MQO não é adequada para estimar a equação (3.6).

Conforme salientado por Lanne e Saikkonen (2010), um problema prático dos processos não causais é o de que eles são não identificáveis por propriedades de segunda ordem ou por funções de verossimilhança gaussianas. Dessa forma, assim como em trabalhos já realizados, será assumido que  $\varepsilon_t$  possui uma densidade de *Lebesgue*, tal que  $f_{\sigma}(\sigma, \omega) = \sigma^{-1} f(\sigma^{-1}x; \omega)$ , dependente do vetor  $\omega(dx1)$  e do parâmetro de escala  $\sigma$ . Essa formulação, presente em Andrews et al (2006) e Lanne e Saikkonen (2010), difere da inicialmente formulada por Breidt et al (1991) ao introduzir  $\omega(dx1)$ .

### 3.3 A equivalência entre as representações causal e não causal

Considere a seguinte representação causal do modelo não causal:

$$y_t - a_1 y_t - \dots - a_p y_{t-p} = \varepsilon_t \quad (3.7)$$

$$a(\beta)y_t = \varepsilon_t \quad (3.8)$$

onde:  $a(\beta) = 1 - a_1\beta - \dots - a_p\beta^p$  e  $\varepsilon_t \sim \text{iid}(0, \sigma^2)$ , possivelmente não gaussiano. O polinômio  $a(\beta)$  pode ser reescrito da seguinte forma:

$$a(z) = \varphi^*(z)\phi(z) \quad (3.9)$$

onde:  $\varphi^*(z) = 1 - \varphi_1^* z - \dots - \varphi_s^* z^s$ ;  $\varphi^*(z) \neq 0$ ,  $|z| \geq 1$ ; e,  $s > 0$  implica  $|\varphi_s^*| > 1$ ;  $\phi(z) = 1 - \phi_1 z - \dots - \phi_r z^r$ ,  $\phi(z) \neq 0$ ;  $|z| \leq 1$ ; e,  $r > 0$  implica  $|\phi_r| < 1$ . As condições de estacionaridade, neste caso, são:

$$\varphi^*(z) \neq 0 \text{ para } |z| \geq 1 \text{ e } \phi(z) \neq 0 \text{ para } |z| \leq 1 \quad (3.10)$$

Quando o processo é estacionário, a condição (3.10) será válida e será possível, então, estabelecer uma relação entre os coeficientes das representações causal e não causal do processo não causal. Esse procedimento será realizado a seguir.

O polinômio  $\varphi^*(z)$  pode ser expresso da seguinte forma:

$$\varphi^*(z) = -\varphi_s^* z^s \left( 1 + \frac{\varphi_{s-1}^*}{\varphi_s^*} z^{-1} + \dots + \frac{\varphi_1^*}{\varphi_s^*} z^{1-s} + \frac{1}{\varphi_s^*} z^{-s} \right)$$

$$\varphi^*(z) = -\varphi_s^* z^s \varphi(z^{-1})$$

onde:  $\varphi(z^{-1}) = 1 - \phi_1 z^{-1} - \dots - \phi_s z^{-s}$  e  $\varphi(z^{-1}) \neq 0$  para  $|z| \leq 1$ .

Então,  $\varphi_{s-j}^*/\varphi_s^* = -\phi_j$  para  $j = 1, \dots, s-1$  e  $1/\varphi_s^* = \phi_s$ . Como as raízes de  $\varphi^*(z)$  estão dentro do círculo unitário, as raízes de  $\varphi(z^{-1})$  estarão fora do círculo unitário. Dessa maneira, a representação (3.8) (causal), fazendo-se  $\varepsilon_t = -(1/\varphi_s^*)\varepsilon_{t+s}$ , torna-se  $\phi(B)\varphi(B^{-1})y_t = \varepsilon_t$ , que é a representação apresentada na equação (3.1) (não causal).

Similarmente, se as condições de estacionaridade mencionadas anteriormente forem satisfeitas, será possível obter a representação causal a partir da representação não causal. Como  $\phi_s \neq 0$ , pondo-se  $\varepsilon_t = -(1/\phi_s)\varepsilon_{t-s}$  e  $\varphi_j^* = -\phi_j/\phi_s$ , para  $j = 1, \dots, s-1$ , e  $\varphi_s^* = 1/\phi_s$ , chega-se na equação (3.8) (causal) a partir da equação (3.1) (não causal).

### 3.4 Deduzindo a Função de Verossimilhança aproximada

O método de estimação por maximização de verossimilhança (MV) para modelos autorregressivos não causais foi estudado inicialmente por Breidt et al (1991) e depois por Andrews et al (2006) e Lanne e Saikkonen (2010)<sup>4</sup>. Em todos eles, a representação básica do modelo utilizada foi a expressada pela equação (3.7).

Além disso, como já mencionado, foi utilizada uma distribuição de densidade  $f(x, \omega)$  para a qual foram assumidas todas as características descritas em Lanne e Saikkonen (2010) e Andrews et al (2006), a saber:  $f(x, \omega)$  é duas vezes diferenciável em relação a  $x$  e a  $\omega$ , não gaussiana e positiva para todo  $x \in \mathbb{R}$  e para todo  $\omega$  admissível (LANNE & LUOTO, 2012). Ainda, o espaço de parâmetros possíveis de  $\omega(dx1)$ , chamado de  $\Lambda$ , é algum subespaço de  $\mathbb{R}^d$ , os valores admissíveis dos parâmetros  $\phi$  e  $\varphi$  são definidos pela condição de estacionaridade (3.2) e  $\sigma > 0$  (LANNE & SAIKKONEN, 2010).

Conforme subseção anterior, considere-se a seguinte representação causal do modelo não causal de ordem  $p$  –  $AR(p)$ :

$$y_t - a_1 y_{t-1} - \dots - a_p y_{t-p} = \varepsilon_t \quad (3.7)$$

onde:  $a(z) = 1 - a_1 z - \dots - a_p z^p \neq 0$  para  $|z| = 1$ ;  $a_p \neq 0$ ; e,  $\varepsilon_t$  é uma sequência iid. tal que  $\varepsilon_t \sim iid(0, \sigma^2)$ , cuja função de densidade de probabilidade é dada por  $f_\sigma(\sigma, \omega) = \sigma^{-1} f(\sigma^{-1} x; \omega)$ , na qual  $\sigma$  é um fator de escala e  $\omega$  é um parâmetro vetorial.

Note-se que, pelas equações (3.3) e (3.4),  $u_t$  e  $v_{t-s+1}$  são iguais a:

$$u_t = \sum_{j=0}^{\infty} \alpha_j \varepsilon_{t-j} = \varepsilon_t + \alpha_1 \varepsilon_{t-1} + \dots + \alpha_j \varepsilon_{t-j} \quad (3.11)$$

<sup>4</sup> Breidt et al utiliza  $f_x = \sigma^{-1} f(\sigma^{-1} x)$ , onde  $\sigma$  é o parâmetro de escala. Andrews et al (2006) e Lanne e Saikkonen (2010) utilizam uma função do tipo  $f_\sigma(\sigma, \omega) = \sigma^{-1} f(\sigma^{-1} x; \omega)$ , conforme descrito no final do item anterior, em que, além do parâmetro de escala  $\sigma$ , também existe o parâmetro vetorial  $\omega(dx1)$ .



$$v_{t-s+1} = \sum_{j=s}^{\infty} \beta_j \varepsilon_{(t-s+1)+j} = \beta_s \varepsilon_{t+1} + \beta_{s+1} \varepsilon_{t+2} + \dots + \beta_{s+j} \varepsilon_{t+j} \quad (3.12)$$

Então,  $u_t$  é independente de  $v_{t-s+1}$  e a função de densidade de probabilidade do vetor estocástico  $(u_1, \dots, u_{T-s}, v_{T-s+1}, \dots, v_T)$  é  $h_U(u_1, \dots, u_r) f_{\sigma}(u_{r+1} - \phi_1 u_{r+2} - \dots - \phi_{T-s+1} u_{T-s}; \omega) h_V(v_{T-s+1}, \dots, v_T)$ .

Utilizando as definições  $u_t = \varphi(B^{-1})y_t$  e  $v_t = \phi(B)y_t$  a matriz (3.13) pode ser escrita:

$$\begin{bmatrix} u_t \\ \vdots \\ u_{T-s} \\ v_{T-s+1} \\ \vdots \\ v_T \end{bmatrix} = \begin{bmatrix} y_1 - \phi_1 y_2 - \dots - \phi_s y_{s+1} \\ \vdots \\ y_{T-s} - \phi_1 y_{T-s+1} - \dots - \phi_s y_T \\ y_{T-s+1} - \phi_1 y_{T-s} - \dots - \phi_r y_{T-s+1-r} \\ \vdots \\ y_T - \phi_1 y_{T-1} - \dots - \phi_r y_{T-r} \end{bmatrix} = A \begin{bmatrix} y_1 \\ \vdots \\ y_{T-s} \\ y_{T-s+1} \\ \vdots \\ y_T \end{bmatrix} \quad (3.13)$$

Ou, resumidamente:

$$x = Ay$$

Se o vetor-coluna  $x$  for expandido, a matriz (3.13) poderá ser reescrita da seguinte maneira:

$$\begin{bmatrix} u_1 \\ \vdots \\ u_r \\ \varepsilon_{r+1} \\ \vdots \\ \varepsilon_{T-s} \\ v_{T-s+1} \\ \vdots \\ v_T \end{bmatrix} = \begin{bmatrix} u_1 \\ \vdots \\ u_r \\ u_{r+1} - \phi_1 u_r - \dots - \phi_r u_1 \\ \vdots \\ u_{T-s} - \phi_1 u_{T-s-1} - \dots - \phi_r u_{T-s-r} \\ v_{T-s+1} \\ \vdots \\ v_T \end{bmatrix} = C \begin{bmatrix} u_1 \\ \vdots \\ u_r \\ u_{r+1} \\ \vdots \\ u_{T-s} \\ v_{T-s+1} \\ \vdots \\ v_T \end{bmatrix} \quad (3.14)$$

Ou, de forma compacta:

$$z = Cx$$

Assim, os vetores  $z$  e  $y$  são relacionados pela seguinte equação:

$$z = CAy$$

A função de densidade conjunta desses parâmetros pode, então, ser escrita:

$$h_U(u_1, \dots, u_r) \left( \prod_{t=r+1}^{T-s} f_{\sigma}(\varepsilon_t, \omega) \right) h_V(v_{T-s+1}, \dots, v_T) \quad (3.15)$$

onde:  $h_U$  e  $h_V$  são as funções de densidade conjunta de  $(u_1, \dots, u_r)$  e de  $(v_{T-s+1}, \dots, v_T)$ , respectivamente. A equação (3.15) somente pode ser escrita como um produto porque os componentes de  $z$ , dados por  $(u, \dots, u_r)$ ,  $(\varepsilon_{r+1}, \dots, \varepsilon_{T-s})$  e  $(v_{T-s+1}, \dots, v_T)$  são independentes (LANNE & SAIKKONEN, 2010). Esse fato pode ser concluído pela observação das equações (3.11) e (3.12).

Ambas as matrizes  $A$  e  $C$  são não singulares e o determinante de  $C$  é um número real. Desse modo, é possível realizar uma transformação de variáveis e expressar a função de densidade conjunta do vetor  $y$  da seguinte maneira:

$$h_U(\varphi(B^{-1})y_1, \dots, \varphi(B^{-1})y_r) \left( \prod_{t=r+1}^{T-s} f_{\sigma}(\varphi(B^{-1})\phi(B)y_t; \omega) \right) \times h_V(\phi(B)y_{T-s+1}, \dots, \phi(B)y_T) |det(A)| \quad (3.16)$$

A partir da equação (3.11), fica evidente que os valores  $(u_1, \dots, u_r)$  devem ser tratados como fixos, uma vez que os  $r$  primeiros valores de  $u_t$  não podem ser estimados dentro da amostra. Da mesma forma, a equação (3.12) torna claro que os valores  $(v_{T-s+1}, \dots, v_T)$  também devem ser tratados como fixos, uma vez que os  $s$  últimos valores de  $v_t$  também não podem ser estimados dentro da amostra. Se esses valores são fixos, tanto  $h_U(u_1, \dots, u_r)$  quanto  $h_V(v_{T-s+1}, \dots, v_T)$  não alteram as estimativas dos parâmetros e podem ser retirados da função de verossimilhança. Assim sendo, a função de densidade conjunta de  $y$  pode ser aproximada pelo segundo termo da equação (3.16), o que dá origem a uma função aproximada, condicional, do logaritmo da função de verossimilhança, como segue:

$$l_T(\theta) = \sum_{t=r+1}^{T-s} g_t(\theta) \quad (3.17)$$

onde  $\theta = (\phi, \varphi, \sigma, \omega)$  e

$$g_t = \log f(\sigma^{-1}(u_t(\varphi) - \phi_1 u_{t-1}(\varphi) - \dots - \phi_r u_{t-r}(\varphi)); \omega) - \log \sigma$$

$$g_t = \log f(\sigma^{-1}(v_t(\phi) - \varphi_1 v_{t+1}(\phi) - \dots - \varphi_s v_{t+s}(\phi)); \omega) - \log \sigma$$

onde:  $u_t(\varphi)$  e  $v_t(\phi)$  são as séries  $u_t$  e  $v_t$  tratadas como funções dos parâmetros  $\varphi$  e  $\phi$ , respectivamente. Maximizando  $l_T(\theta)$  para cada valor permitido de  $\theta$ , é possível, então, aferir um valor aproximado do logaritmo da máxima verossimilhança e dos parâmetros  $\theta$ . Importante destacar que, para aplicar esse procedimento, a ordem do modelo  $p = r + s$  é assumida como dada.

Em complemento, como demonstrado por Lanne e Saikkonen (2010) e Andrews et al (2006), os estimadores apresentados são consistentes e eficientes. Essa prova não será apresentada neste trabalho. Por fim, os autores demonstram que um estimador da matriz hessiana de  $g_t$  pode ser calculado da seguinte maneira:

$$\hat{Q} \stackrel{\text{def}}{=} -(T - p)^{-1} \sum_{t=r+1}^{T-s} \frac{\partial^2}{\partial \theta \partial \theta'} g_t(\hat{\theta}) \quad (3.18)$$

onde:  $\hat{Q}$  é o estimador da matriz hessiana;  $\hat{\theta}$  é o vetor que contém os parâmetros estimados; e,  $\frac{\partial^2}{\partial \theta \partial \theta'} g_t(\hat{\theta})$  é o vetor que contém as derivadas de segunda ordem em relação aos parâmetros. Assim, a partir da equação (3.18), o desvio padrão dos componentes do vetor de parâmetros estimados ( $\hat{\theta}$ ) pode ser estimado pelo cálculo da raiz quadrada dos elementos da diagonal da matriz  $(T - p)\hat{Q}^{-1}$  (LANNE & SAIKKONEN, 2010).

## 4 RESULTADOS

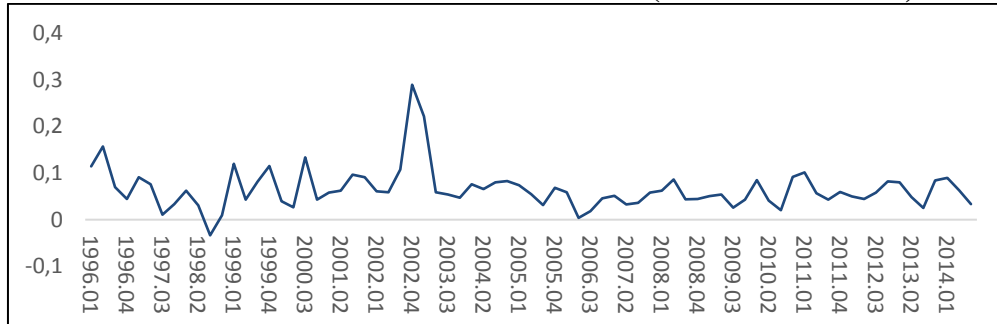
### 4.1 Os dados utilizados

Foram utilizados os dados mensais, de janeiro de 1996 a setembro de 2014, do Índice de Preços ao Consumidor Amplo – IPCA coletados junto ao Instituto Brasileiro de Geografia e Estatística – IBGE. Os dados, com frequência mensal, foram trimestralizados e, depois, anualizados da seguinte forma:

$$\pi_t = [\prod_{m=1}^3 (1 + \pi_m)]^4 - 1$$

onde:  $\pi_m$  é a inflação no mês  $m$  do trimestre em questão. Esse procedimento gerou uma série com 75 observações do primeiro trimestre de 1996 até o terceiro trimestre de 2014.<sup>5</sup>

**Gráfico 1 – IPCA trimestral anualizado (1996-01 a 2014-03)**



### 4.2 A não causalidade do processo inflacionário brasileiro

O objetivo desta seção é estimar o processo autorregressivo,  $AR(r, s)$ , que melhor se ajusta à série de preços apresentada no item anterior. Lanne e Luoto (2013) sugerem, primeiramente, definir a

<sup>5</sup> A série ainda foi testada no software Demetra e foi recusada a hipótese de sazonalidade. Versão 1.4.1. O pacote utilizado foi o X11 com as seguintes especificações: retiradas as opções “trading days”, “easter” e “sazonal adjustment”.

ordem ( $p = r + s$ ) do modelo e, depois, distribuir essa ordem entre defasagens ( $r$ ) e valores à frente da variável ( $s$ ).

Para a definição da ordem do modelo, Lanne e Saikkonen (2010) utilizam tanto o *Akaike Information Criterion* (AIC) quanto o *Schwarz Information Criterion* (SIC)<sup>6</sup>. Entretanto, Ivanov e Killian (2005) afirmam que, para dados trimestrais e com uma amostra inferior a 120 observações, que é o caso da série utilizada neste trabalho, o SIC apresenta os melhores resultados.

A ordem do modelo é definida pela estimação de um  $AR(p)$  gaussiano por Mínimo Quadrados Ordinários (MQO). A tabela 1 apresenta os resultados das estimativas de modelos auto regressivos causais –  $AR(p)$ , no software *EViews* (versão 7), e os valores estimados dos SICs correspondentes. O melhor modelo, por esse critério, foi o  $AR(3)$ .

**Tabela 1 – Número de defasagens do IPCA\***

Variável	Nº de defasagens	SIC
$\pi_t$	1	-3,280448
	2	-3,236687
	<b>3</b>	<b>-3,450648</b>
	4	-3,382797
	5	-3,324755
	6	-3,268573
	7	-3,222692
	8	-3,143206
	9	-3,070670
	10	-2,993276

\*Embora a tabela apresente somente o SIC, vale destacar que o AIC também selecionou três defasagens para o modelo.

Destaca-se, ainda, que os resíduos do  $AR(3)$  estimado por MQO não apresentaram correlação nos testes realizados<sup>7</sup>, que é a condição apontada por tanto por Lanne e Luoto (2013) quanto por Lanne e Saikkonen (2010) para a aceitação da ordem ( $p$ ) selecionada em estimativas por MQO de processos não causais.

Uma vez escolhido um número de defasagens, tal que  $p = r + s = 3$ , é necessário, então, definir a distribuição entre defasagens ( $r$ ) e valores à frente ( $s$ ) da variável. É importante ressaltar que, segundo Lanne e Luoto (2013) e Lanne e Saikkonen (2010), a distribuição entre defasagens e valores à frente só pode ser identificada caso o processo seja não Gaussiano. Com efeito, após a estimação do modelo  $AR(p)$ , no qual  $p = 3$ , foi realizado o teste de Jarque-Bera, que recusou a normalidade dos resíduos, conforme tabela a seguir<sup>8</sup>.

Uma vez determinada a ordem do modelo  $AR(p)$  e rejeitada a hipótese de normalidade dos resíduos, adota-se a distribuição entre defasagens e valores à frente que maximiza a Função de Verossimilhança ( $l_T(\theta)$ ), admitida como uma *t-student*, deduzida no item 3.4 deste trabalho. Foi utilizado o pacote “R” para esses cálculos.

<sup>6</sup> Tanto o AIC quanto o SIC realizam ponderações entre os valores de verossimilhança estimados e a quantidade de parâmetros utilizados. No caso do SIC, a ponderação é realizada da seguinte forma:  $SIC = -2\ln(\hat{L}) + k\ln(n)$ , onde  $\hat{L}$  é o valor estimado da verossimilhança,  $k$  é o número de parâmetros estimados e  $n$  é o tamanho da amostra. O melhor modelo é o que minimiza o valor do SIC.

<sup>7</sup> A inexistência de correlação foi verificada por meio do correlograma dos resíduos e pelo teste Breusch-Godfrey no *software* *EViews*. Ambos indicaram ausência de correlação.

<sup>8</sup> A estatística de teste foi 166,43, que gerou um p-valor de 0,0000.

**Tabela 2 – Composição entre lags e leads**

Modelo	p	r	s	LogL
AR(3,0)	3	3	0	-178,0854
AR(2,1)	3	2	1	-180,6082
AR(1,2)	3	1	2	-177,2128
<b>AR(0,3)</b>	<b>3</b>	<b>0</b>	<b>3</b>	<b>-173,2179</b>

O modelo selecionado, portanto, foi o  $AR(0,3)$ , puramente não causal. Esse modelo pode ser representado genericamente, seguindo a equação (3.1), da seguinte maneira:

$$\phi(B)\varphi(B^{-1})y_t = \varepsilon_t \quad (3.1)$$

$$\pi_t = \varphi_1\pi_{t+1} + \varphi_2\pi_{t+2} + \varphi_3\pi_{t+3} + \varepsilon_t$$

Os parâmetros estimados<sup>9</sup> da equação acima estão indicados na tabela 4.

**Tabela 3 – Processo  $AR(0, 3)$  estimado para o IPCA**

Modelo	AR(0,3)	
Parâmetro	Valor estimado	Desvio Padrão
$\varphi_1$	0,5420	0,0936
$\varphi_2$	-0,4396	0,0907
$\varphi_3$	0,4293	0,1008
$v$	94,10	12,950
$\sigma$	2,00	$0,217 \times 10^{-11}$

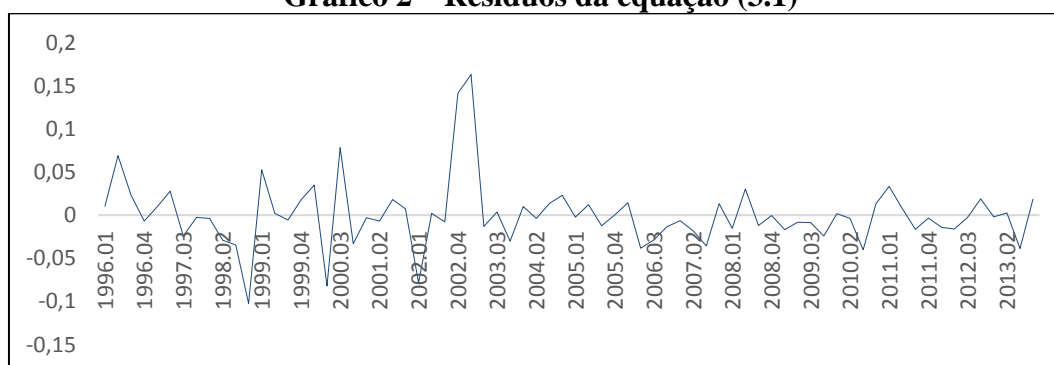
As variáveis  $v$  e  $\sigma$  são, respectivamente, as estimativas de graus de liberdade e de escala da distribuição  $t$ -student dos resíduos do modelo. As estimativas dos desvios padrão foram realizadas a partir da matriz hessiana da função de verossimilhança  $l_T(\theta)$  utilizada. Nesse caso específico, sabe-se que  $l_T(\theta) = l(\varphi_1, \varphi_2, \varphi_3, v, \sigma)$ .

O processo autorregressivo pode, então, ser escrito conforme a equação (4.1).

$$\pi_t = \underset{(0,0936)}{0,5420} \pi_{t+1} - \underset{(0,0906)}{0,4396} \pi_{t+2} + \underset{(0,1008)}{0,4293} \pi_{t+3} + \varepsilon_t \quad (4.1)$$

onde:  $\varepsilon_t$  independente e identicamente distribuído (iid), com média nula e distribuição  $t$ -student parametrizada de acordo com  $v$  e  $\sigma$ . O gráfico 2 apresenta os resíduos do modelo estimado.

**Gráfico 2 – Resíduos da equação (3.1)**



<sup>9</sup> Utilizado o pacote “R” para essas estimativas.

### 4.3 As representações do Processo Autorregressivo Estimado

Conforme deduzido na seção 3.3 deste trabalho, o modelo estimado possui duas representações: uma não causal e outra causal. A primeira representação do modelo, puramente não causal, é dada pela equação (4.1):

$$\pi_t = 0,5420\pi_{t+1} - 0,4396\pi_{t+2} + 0,4293\pi_{t+3} + \varepsilon_t \quad (4.1)$$

$$(1 - 0,5420B^{-1} + 0,4396B^{-2} - 0,4293B^{-3})\pi_t = \varepsilon_t \quad (4.2)$$

onde:  $B$  é o operador de defasagens, tal que  $B^{-r}y_t = y_{t+r}$ . A equação característica deste processo pode ser escrita da seguinte forma:

$$0,4293z^3 - 0,4396z^2 + 0,5420z - 1 = 0 \quad (4.3)$$

As raízes da equação (4.3) são:  $z_1 = 1,3577$ ,  $z_2 = -0,1699 + 1,2992i$  e  $z_3 = -0,1699 - 1,2992i$ . Assim sendo,  $|z| > 1$  e a condição de estacionaridade (3.2) está respeitada. Conforme apresentado na seção 3.3, partindo da equação (4.3), chega-se à seguinte representação causal do processo puramente não causal estimado:

$$-2,394z^3 + 1,2625z^2 - 1,024z + 1 = 0 \quad (4.4)$$

A equação (4.4) relaciona-se ao seguinte processo causal:

$$\pi_t = 1,024\pi_{t-1} - 1,2625\pi_{t-2} + 2,3294\pi_{t-3} + \varepsilon_t \quad (4.5)$$

Como as raízes de (4.4) são o inverso das calculadas em (4.3) e estas estão fora do círculo unitário, aquelas estão dentro. Entretanto, qualquer teste de estacionaridade que utilize a representação causal não é adequado. Ainda, o  $AR(p)$ , tal que  $p = 3$ , estimado na seção anterior, não pode ser estimado pelo método de mínimos quadrados ordinários (MQO), pois os resíduos estarão correlacionados com os valores passados da variável, conforme demonstrado na seção 3.2. A representação causal, com resíduos independentes e identicamente distribuídos (iid), conforme a equação (4.5), somente pode ser obtida utilizando-se a representação não causal do processo (equação 4.1), pelo estimador de máxima verossimilhança descrito na seção 3.

### 4.4 A CPH como um processo autorregressivo não causal

O objetivo desta seção é apresentar a Curva de Phillips Híbrida (CPH), descrita na seção 2, como um processo autorregressivo possivelmente não causal. A partir disso, será possível estimar o peso da inflação esperada na CPH brasileira. Formulou-se, então, uma curva novo keynesiana com um componente *backward-looking*, gerando um modelo híbrido. A CPH pode ser expressa conforme a equação a seguir:

$$\pi_t = \omega\pi_{t-1} + (1 - \omega)\beta E_t \pi_{t+1} + \gamma x_t \quad (2.5)$$

onde:  $\omega$  é a proporção de firmas que seguem uma regra de ajuste de preços do tipo *backward-looking*,  $\beta$  é o fator de desconto das firmas ( $0 < \beta < 1$ ) e  $x_t$  é uma medida de hiato do produto. A estimativa do hiato de produto ( $x_t$ ), conforme abordado na seção 2, é altamente controversa na literatura. A escolha da medida para estimar essa variável afeta decisivamente os resultados nesse tipo estudo. Uma vantagem adicional da metodologia ora adotada é a de que ela prescinde de uma definição prévia de medida de hiato do produto para a modelagem da inflação.

De forma alternativa, para facilitar as manipulações algébricas, é possível reescrever a equação (2.5), em acordo com Lanne e Luoto (2013):

$$\pi_t = \lambda_f E_t \pi_{t+1} + \lambda_b \pi_{t-1} + \gamma x_t \quad (4.6)$$

Reescrevendo a equação (4.6):

$$\pi_t = \lambda_f \pi_{t+1} + \lambda_b \pi_{t-1} + \eta_{t+1} \quad (4.7)$$

onde:  $\eta_{t+1} = \lambda_f (E_t \pi_{t+1} - \pi_{t+1}) + \gamma x_t \equiv \xi_{t+1} + \gamma x_t$ . Nos modelos baseados na teoria das expectativas racionais (TER), o erro esperado de  $\xi_{t+1}$  é assumido como independente e identicamente distribuído (iid) e com esperança nula. Por meio de algumas manipulações algébricas, é possível obter a equação (4.8).

$$(1 - B\lambda_f^{-1} + B^2\lambda_b\lambda_f^{-1})\pi_t = -\lambda_f^{-1}\eta_t \quad (4.8)$$

Agora, considerando que o polinômio  $a(z) \equiv 1 - \lambda_f^{-1}z + \lambda_b\lambda_f^{-1}z^2$  pode ser reescrito como  $a(z) \equiv (1 - \phi z)(1 - \varphi^* z)$ , onde  $\phi$  e  $\varphi^*$  são as raízes da equação (4.8), sendo:

$$\phi = \frac{1}{2} \left( \lambda_f^{-1} - \sqrt{\lambda_f^{-2} - 4\lambda_f^{-1}\lambda_b} \right) \text{ e } \varphi^* = \frac{1}{2} \left( \lambda_f^{-1} + \sqrt{\lambda_f^{-2} - 4\lambda_f^{-1}\lambda_b} \right) \quad (4.9)$$

Com o conjunto de valores possíveis de  $\lambda_f$  e  $\lambda_b$ , os parâmetros serão tal que  $\phi < 1$  e  $\varphi^* > 1$ . Torna-se, então, conveniente reescrever  $a(z)$ :

$$a(z) \equiv (1 - \phi z)(1 - \varphi^* z) = -(1 - \phi z)\varphi^* \left( 1 - \frac{1}{\varphi^*} z^{-1} \right) z = -\varphi^*(1 - \phi z)z(1 - \varphi z^{-1})$$

onde  $\varphi = \frac{1}{\varphi^*}$ . Por analogia, substituindo esses resultados na equação (4.8):

$$(1 - \phi B)(1 - \varphi B^{-1})\pi_t = \epsilon_t \quad (4.10)$$

sendo:

$$\epsilon_t = (\varphi^* \lambda_f)^{-1} \eta_t. \quad (4.11)$$

Se  $\eta_t$  fosse iid, a equação (4.10) se tornaria um AR(1,1) e estimadores consistentes com os parâmetros da equação (4.6) poderiam ser obtidos estimando um modelo AR(1,1) para a inflação pelo método de maximização de verossimilhança (LANNE & LUOTO, 2013).

Entretanto,  $\eta_{t+1} = \xi_{t+1} + \gamma x_t$  e  $x_t$  provavelmente não é iid. Logo, é necessário adequar a modelagem de  $\eta_t$  para permitir auto-correlação da série (LANNE & LUOTO, 2013). Considere, então, o processo autorregressivo causal  $AR(r-1)$  a seguir:

$$\rho(B)\theta(B^{-1})\eta_t = \zeta_t \quad (4.12)$$

onde:  $\rho(B) = 1 - \rho_1 B - \dots - \rho_{r-1} B^{r-1}$ ,  $\theta(B^{-1}) = 1 - \theta_1 B^{-1} - \dots - \theta_{s-1} B^{-s+1}$  e  $\zeta_t$  é o erro iid.

Substituindo a equação (4.12) na equação (4.10):

$$\rho(B)(1 - \phi B)\theta(B^{-1})(1 - \varphi B^{-1})\pi_t = \epsilon_t \quad (4.13)$$

$$\phi(B)\varphi(B^{-1})\pi_t = \epsilon_t \quad (4.14)$$

onde:  $\phi(B) \equiv \rho(B)(1 - \phi B)$ ,  $\varphi(B^{-1}) \equiv \theta(B^{-1})(1 - \varphi B^{-1})$  e  $\epsilon_t \equiv (\varphi^* \lambda_f)^{-1} \zeta_{t+1}$ . A equação (4.14) será o modelo utilizado para o comportamento da inflação. Os valores para  $\phi$  e  $\varphi$  não são, necessariamente, únicos. Entretanto, qualquer par de raízes reais encontradas pode ser utilizado para resolver as equações (4.9) e aferir valores para  $\lambda_f$  e  $\lambda_b$ . Restrições teóricas impostas a  $\lambda_f$  e  $\lambda_b$  podem eliminar a multiplicidade dos valores encontrados.

No caso de um modelo puramente não causal, que foi o caso estimado neste trabalho para a economia brasileira,  $\lambda_b = 0$ . Assim, podemos simplificar a equação (4.8).

$$(1 - B\lambda_f^{-1} + B^2\lambda_b\lambda_f^{-1})\pi_t = -\lambda_f^{-1}\eta_t \quad (4.8)$$

$$(1 + B\lambda_f^{-1})\pi_t = -\lambda_f^{-1}\eta_t$$

Logo, o polinômio  $a(z)$  também pode ser simplificado e reescrito da seguinte forma:

$$a(z) = (1 - \varphi^*z) = (1 + \lambda_f^{-1}z)$$

De onde, extraem-se as seguintes relações:

$$\frac{1}{\varphi^*} = \varphi = \lambda_f \quad (4.15)$$

Destaque-se, ainda, que a equação (3.14) é uma aplicação da equação (2.1) para o caso particular de  $y_t = \pi_t$ .

$$\phi(B)\varphi(B^{-1})y_t = \varepsilon_t \quad (3.1)$$

$$\phi(B)\varphi(B^{-1})\pi_t = \varepsilon_t \quad (4.14)$$

#### 4.5 A estimativa da Curva de Phillips para o caso brasileiro

Na seção anterior, provou-se que a CPH, expressa nos termos da equação (4.6), pode ser escrita de acordo com a equação (4.13).

$$\pi_t = \lambda_f E_t \pi_{t+1} + \lambda_b \pi_{t-1} + \gamma x_t \quad (4.6)$$

$$\rho(B)(1 - \phi B)\theta(B^{-1})(1 - \varphi B^{-1})\pi_t = \varepsilon_t \quad (4.13)$$

Considerando que o processo estimado na seção 4.2 foi puramente não causal, podemos reescrever a equações (4.13) de maneira mais simplificada, sem os componentes causais.

$$\theta(B^{-1})(1 - \varphi B^{-1})\pi_t = \varepsilon_t \quad (4.16)$$

onde:  $\theta(B^{-1})$  e  $(1 - \varphi B^{-1})$  são os polinômios referentes ao hiato do produto e à inflação, respectivamente. Note-se, ainda, que é permitido que  $\theta(B^{-1})$  tenha mais de um elemento, enquanto  $(1 - \varphi B^{-1})$  é, obrigatoriamente, de primeira ordem. Esse fato está coerente com a CPH utilizada (equação 4.6), que permite apenas um *lead* da inflação, enquanto que o hiato do produto foi modelado permitindo a ocorrência de diversos *leads* e *lags* (equação 4.12).

Repetindo o processo realizado na seção 4.3 para representar a equação característica do processo inflacionário estimado na seção 4.2, a equação (4.16) aparece parcialmente estimada<sup>10</sup>:

$$\pi_t = 0,5420\pi_{t+1} - 0,4396\pi_{t+2} + 0,4293\pi_{t+3} + \varepsilon_t \quad (4.1)$$

$$(1 - 0,5420B^{-1} + 0,4396B^{-2} - 0,4293B^{-3})\pi_t = \varepsilon_t \quad (4.2)$$

O polinômio característico do processo  $(1 - 0,4293z^3 + 0,4396z^2 - 0,5420z = 0)$  pode ser decomposto em dois polinômios de ordem inferior, tornando possível reescrever a equação (4.2) da seguinte maneira:

$$[(1 + 0,1946B^{-1} + 0,5829B^{-2})(1 - 0,7365B^{-1})]\pi_t = \varepsilon_t \quad (4.17)$$

A equação (4.17) é a equação (3.16) estimada, onde  $\theta(B^{-1}) = (1 + 0,1946B^{-1} + 0,5829B^{-2})$  e  $(1 - \varphi B^{-1}) = (1 - 0,7365B^{-1})$ . Logo,  $\theta_1 = -0,1946$ ,  $\theta_2 = -0,5829$  e  $\varphi = 0,7365$ .

Dessa maneira, conforme já deduzido na seção 3.4, para modelos puramente não causais, vale a seguinte relação:

<sup>10</sup> A equação (3.2) é uma estimativa completa da equação (4.16).

$$\frac{1}{\varphi^*} = \varphi = \lambda_f \quad (4.15)$$

Considerando então as estimativas realizadas,  $\varphi = \lambda_f = 0,7365$  e  $\lambda_b = 0$ . Logo, a CPH brasileira se torna puramente não causal. Assim sendo, o modelo estimado para a Curva de Phillips brasileira é compatível com a formulação original novo keynesiana elaborada por Calvo (1983), descrita na seção 2 deste trabalho e representada pela equação (2.4):

$$\pi_t = \beta E_t \pi_{t+1} + \gamma x_t \quad (2.4)$$

$$\pi_t = 0,7365 E_t \pi_{t+1} + \gamma x_t \quad (4.18)$$

Essa CPH pode ser reescrita nos termos da equação (4.7):

$$\pi_t = \lambda_f \pi_{t+1} + \lambda_b \pi_{t-1} + \eta_{t+1} \quad (4.7)$$

$$\pi_t = 0,7365 \pi_{t+1} + \gamma x_t + \xi_{t+1} \quad (4.19)$$

onde:  $\eta_{t+1} = \gamma x_t + \xi_{t+1}$ ;  $\lambda_f(E_t \pi_{t+1} - \pi_{t+1}) = \xi_{t+1}$ ; e  $\xi_{t+1}$  é, de acordo com a teoria das expectativas racionais, um resíduo iid com esperança nula. Ainda podemos estimar os desvios padrão das variáveis, utilizando o hessiano da função de verossimilhança, seguindo o mesmo procedimento da seção 4.2. A tabela 4 sumariza os resultados estimados para a Curva de Phillips estimada neste trabalho.

**Tabela 4 – Processo para a CP**

Modelo	$\pi_t = \lambda_f \pi_{t+1} + \lambda_b \pi_{t-1} + \gamma x_t + \xi_{t+1}$	
Parâmetro	Valor estimado	Desvio Padrão
$\varphi = \lambda_f$	0,7365	0,0682
$\lambda_b$	-	-
$v$	94,10	12,950
$\sigma$	2,00	$0,217 \times 10^{-11}$

A equação (4.19) estimada para a economia brasileira aponta para um comportamento fortemente *forward-looking*. Esse comportamento é coerente com estudos para a economia brasileira, como Areosa e Medeiros (2007), principalmente quando o fator cambial é adicionado em acordo com a metodologia introduzida por Monacelli (2005). No entanto, o comportamento antecipativo estimado dos agentes, no presente artigo, é ainda mais dominante.

Quando comparado aos resultados obtidos para a realidade americana, as estimativas explicitadas na tabela 4 também se mostram similares aos resultados obtidos por Lanne e Luoto (2013). Neste trabalho, os autores estimaram dois modelos utilizando índices distintos para captar o comportamento dos preços na economia americana, o deflator implícito do PIB e o *Consumer Price Index* – CPI. Em ambos os casos, o comportamento *forward-looking* foi, em larga medida, dominante, embora o comportamento *backward-looking* não tenha sido descartado em nenhum deles.

Tal resultado fornece relevantes implicações para o papel da autoridade monetária no controle inflacionário, sugerindo que as expectativas inflacionárias foram o principal fator no processo de formação de preços da economia brasileira no período analisado.

## 5. CONCLUSÃO

Neste trabalho estima-se uma Curva de Phillips Híbrida, Novo Keynesiana, para o Brasil. Como em Lanne e Luoto (2013) é possível provar, adotando-se a curva de Phillips Híbrida de Galí e Gertler (1999) e hipóteses genéricas sobre o hiato do produto, que a taxa de inflação segue um processo autorregressivo não causal univariado. Admitindo-se que a taxa de inflação siga tal processo pode-se



testar a importância relativa das expectativas sobre a taxa de inflação futura e passada na determinação da taxa de inflação corrente.

A obtenção de estimadores de máxima verossimilhança (MV) para os parâmetros dos modelos autorregressivos não causais foi explorada inicialmente por Breidt et al (1991) e depois por Andrews et al (2006) e Lanne e Saikkonen (2010). Neste artigo adotamos os mesmos procedimentos de estimação por máxima verossimilhança descritos em Lanne e Luoto (2013). Foram utilizados dados para a taxa de inflação brasileira, medida pelo IPCA, entre 1996 (1º trimestre) e 2014 (3º trimestre).

Neste artigo, foi detectado um importante componente *forward-looking* na curva de Phillips Nova Keynesiana estimada para o Brasil. Esse resultado é coerente com outros estudos para a economia brasileira, como Areosa e Medeiros (2007), principalmente quando o fator cambial é adicionado por eles seguindo a metodologia introduzida por Monacelli (2005). Quando comparado aos resultados obtidos para a realidade americana, as estimativas se mostraram semelhantes às aquelas obtidas por Lanne e Luoto (2013), que estimaram dois modelos utilizando índices distintos para captar o comportamento dos preços da economia americana: o deflator implícito do PIB; e, o *Consumer Price Index* – CPI. Em ambos os casos, o comportamento *forward-looking* foi, em larga medida, dominante, embora o componente *backward-looking* não tenha sido descartado em nenhum deles.

Os resultados obtidos neste trabalho fornecem relevantes implicações para a atuação da autoridade monetária no controle inflacionário, sugerindo que o papel das expectativas inflacionárias foi o principal fator no processo de formação de preços da economia brasileira no período analisado.

## 6. REFERÊNCIAS BIBLIOGRÁFICAS

- ANDREWS, B., M. CALDER, e R.A. DAVIS (2009). "Maximum likelihood estimation for stable autoregressive processes". *Annals of Statistics* 37, 1946-1982.
- ANDREWS, B., R.A. DAVIS, e F.J. BREIDT (2006). "Maximum likelihood estimation for all-pass time series models". *Journal of Multivariate Analysis* 97, 1638-1659.
- AREOSA, W. e MEDEIROS, M. (2007). "Inflation Dynamics in Brazil: The Case of a Small Open Economy". *Brazilian Review of Econometrics* 27, 131-166.
- BREIDT, J., R.A. DAVIS, K.S. LI, e M. ROSENBLATT (1991). "Maximum likelihood estimation for noncausal autoregressive processes". *Journal of Multivariate Analysis* 36, 175-198.
- BREIDT, J., R.A. DAVIS, e A.A. TRINDADE (2001). "Least absolute deviation estimation for all-pass time series models". *The Annals of Statistics* 29, 919-946.
- CALVO, G. A. (1983). "Staggered Prices in a Utility-Maximizing Framework". *Journal of Monetary Economics*, 12, 383-398.
- FANELLI, L. (2005). "Testing the new keynesian phillips curve through vector autoregressive models: results from the euro area". *Universita Ca Foscari, Venezia*, January 24-25, <http://www.cide.info/conf/congress.php>.
- FASOLO, A. e PORTUGAL, M. (2004). "Imperfect Rationality and Inflationary Inertia: a New Estimation of the Phillips Curve for Brazil". *Estudos Econômicos*, 34(4), 725-776.
- FIALHO, M. e PORTUGAL, M. (2005). "Monetary Policy and Fiscal Policy Interactions in Brazil: an Application of the Fiscal Theory of the Price Level". *Estudos Econômicos* 35: 657-685.
- FRIEDMAN, M. (1968). "The Role of Monetary Policy". *American Economic Review*, 58, 1-17.
- GALÍ, J. e GERTLER, M. (1999). "Inflation Dynamics: A Structural Econometric Analysis". *Journal of Monetary Economics*, 44, 195-222.
- GALÍ, J., GERTLER, M., LOPEZ-SALIDO, J.D. (2005). "Robustness of the estimates of the hybrid New Keynesian Phillips curve". *Journal of Monetary Economics* 52, 1107-1118.
- GOURIEROUX, C. e JASIAK, J. (2015). "Semi-Parametric Estimation of Noncausal Vector Autoregression". *CHEST DP*.
- HOLLAND, M. e MORI, R. (2010). "Dinâmica da Inflação no Brasil e os Efeitos Globais". *Revista Economia*, 13(3), 649-670.

- IVANOV, V., e L., KILIAN (2005), "A Practitioner's Guide to Lag-Order Selection for VAR Impulse Response Analysis", *Studies in Nonlinear Dynamics and Econometrics* 9(1).
- LANNE, M., SAIKKONEN, P. (2010). "Noncausal autoregressions for economic time series". Munich Personal RePEc Archive.
- LANNE, M., e J. LUOTO (2013). "Autoregression-based estimation of the new Keynesian Phillips curve". *Journal of Economic Dynamics and Control* 37, 561-570.
- LEEPER, E., WALKER, T., e S., YANG (2012): "Fiscal Foresight and Information Flows", *Econometrica*, 81, 1115-1145.
- LII, K. S. E, M. ROSENBLATT (1996). "Maximum likelihood estimation for non-Gaussian nonminimum phase ARMA sequences". *Statistica Sinica* 6, 1.22.
- LIMA, E. (2003). "The NAIRU, Unemployment and the Rate of Inflation in Brazil". *Revista Brasileira de Economia*, 57, 899-930.
- LANNE, M. e LUOTO, J. (2013). "Autoregression-based estimation of the new Keynesian Phillips curve". *Journal of Economic Dynamics & Control* 37, (561-572).
- MAKA, A (2013). "On Testing the Phillips Curves, the IS Curves, and the Interaction between Fiscal and Monetary Policies". Fundação Getúlio Vargas – FGV.
- MAZALI, J. e DIVINO, A. (2010). "Real Wage Rigidity and the New Phillips Curve: The Brazilian Case". *Revista Brasileira de Economia*, 64, 291-306.
- MONACELLI, T. (2005). "Monetary Policy in a Low Pass-Through Environment". *Journal of Money Credit and Banking*, 37, 1047-1066.
- PHELPS, E., S. (1967). "Phillips Curves, Expectations of Inflation and Optimal Unemployment over Time". *Economica* 34 (135): 254-281.
- PHELPS, E., S. (1968). "Money-Wage Dynamics and Labor-Market Equilibrium". *Journal of Political Economy* 76 (4): 678-711.
- PHILLIPS, A.W. (1958), "The Relation between Unemployment and the Rate of Change of Money Wage Rates in the United Kingdom, 1861- 1957", *Economica* 45, November, pp.283-299.
- RUDD, J. e WHELAN, K. (2005). "Modelling Inflation Dynamics: A Critical Survey of Recent Research". *Journal of Money Credit and Banking*, 39, 155-170.