

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
FACULDADE DE CIÊNCIAS ECONÔMICAS
PROGRAMA DE PÓS GRADUAÇÃO EM ECONOMIA

RODRIGO DE SÁ DA SILVA

BANCO CENTRAL E PREFERÊNCIAS ASSIMÉTRICAS: UMA
APLICAÇÃO DE SIEVE ESTIMATORS PARA OS EUA E O BRASIL

Porto Alegre

2011

RODRIGO DE SÁ DA SILVA

**BANCO CENTRAL E PREFERÊNCIAS ASSIMÉTRICAS: UMA
APLICAÇÃO DE SIEVE ESTIMATORS PARA OS EUA E O BRASIL**

Dissertação submetida ao Programa de Pós-Graduação em Economia da Faculdade de Ciências Econômicas da UFRGS, como quesito parcial para obtenção do título de Mestre em Economia, com ênfase em Economia Aplicada

Orientador: Prof. Dr. Marcelo Savino
Portugal

Porto Alegre

2011

DADOS INTERNACIONAIS DE CATALOGAÇÃO NA PUBLICAÇÃO (CIP)
Responsável: Biblioteca Gládis W. do Amaral, Faculdade de Ciências Econômicas da UFRGS

S586b Silva, Rodrigo de Sá da
Banco Central e preferências assimétricas : uma aplicação de sieve estimators para os EUA e o Brasil / Rodrigo de Sá da Silva. – Porto Alegre, 2011.
42 f. : il.

Orientador: Marcelo Savino Portugal.

Ênfase em Economia Aplicada.

Dissertação (Mestrado em Economia) - Universidade Federal do Rio Grande do Sul, Faculdade de Ciências Econômicas, Programa de Pós-Graduação em Economia, Porto Alegre, 2011.

1. Econometria : Estimação. 2. Política monetária : Modelo econométrico. I. Portugal, Marcelo Savino. II. Universidade Federal do Rio Grande do Sul. Faculdade de Ciências Econômicas. Programa de Pós-Graduação em Economia. III. Título.

CDU 330.115

RODRIGO DE SÁ DA SILVA

**BANCO CENTRAL E PREFERÊNCIAS ASSIMÉTRICAS: UMA
APLICAÇÃO DE SIEVE ESTIMATORS PARA OS EUA E O BRASIL**

Dissertação submetida ao Programa de Pós-Graduação em Economia da Faculdade de Ciências Econômicas da UFRGS, como quesito parcial para obtenção do título de Mestre em Economia, com ênfase em Economia Aplicada

Aprovada em Porto Alegre, 05 de Setembro de 2011

Banca Examinadora

Prof. Dr. Márcio Poletti Laurini

Ibmec-RJ

Prof. Dr. Edilean Kleber da Silva Arejano Aragón

PPGE/UFPB

Prof. Dr. Flávio Augusto Ziegelmann

PPGE/UFRGS

Prof. Dr. Marcelo Savino Portugal - Orientador

PPGE/UFRGS

Aos meus pais.

Agradecimentos

Ao professor Marcelo Portugal, pela oportunidade de trabalho que me foi concedida nesta dissertação, pela atenção e pela confiança ao longo de todo o trabalho, e por todos os ensinamentos transmitidos desde o início do Mestrado.

Aos professores do PPGE, por todos os conhecimentos adquiridos.

À Carina, pelo apoio nos momentos mais críticos deste trabalho.

À Secretaria do PPGE, pelo apoio sempre disponível.

Ao CNPq, pelo apoio financeiro.

À FEE, pelo apoio e consideração para a realização deste trabalho.

Mais uma vez aos meus pais, por terem sempre acreditado e me apoiado nesta realização.

Psico-História... Gaal Dornick, utilizando conceitos não matemáticos, definiu a psico-história como o ramo da matemática que trata das reações dos conglomerados humanos a estímulos sociais e econômicos fixos...

... Implícita em todas essas definições está a suposição de que o conglomerado humano que está em foco é suficientemente grande para um tratamento estatístico válido. O tamanho necessário de tal conglomerado pode ser determinado pelo Primeiro Teorema de Seldon, que... Uma suposição necessária posterior é que o conglomerado humano esteja ele próprio inconsciente da análise psico-histórica para que as suas reações sejam verdadeiramente aleatórias...

A base de toda a psico-história válida baseia-se no desenvolvimento das Funções Seldon, que exibem propriedades congruentes com as de forças sociais e econômicas como... (Enciclopédia Galática)

Resumo

Uma questão interessante na política monetária é se os Bancos Centrais dão pesos iguais para desvios positivos e negativos da inflação e do hiato do produto das suas respectivas metas. Para responder à esta questão, estimou-se a função perda da autoridade monetária não parametricamente através do método de *sieve estimator*, expandindo-a através de uma base composta de polinômios ortogonais. A economia foi modelada com agentes *forward-looking* e com comprometimento por parte da autoridade monetária. O método foi aplicado para a os Estados Unidos desde 1960 e para o Brasil a partir de 1999. Para a economia norte-americana não foram encontradas evidências de assimetria nas preferências da autoridade monetária. Já no Brasil o Banco Central mostrou ter preferências assimétricas quanto à inflação, auferindo uma maior perda de desvios negativos do que positivos em relação à meta.

Palavras-chave: Preferências do Banco Central. Função perda assimétrica. Sieve Estimators. Polinômios Ortogonais.

Abstract

An interesting question in monetary policy is whether the Central Bank gives equal weights to positive and negative deviations of inflation and output gap from their targets. Trying answering this question, we estimated the monetary authority's loss function nonparametrically, using the method of sieves, expanding it with orthogonal polynomials. The economy was model with forward-looking agents and with commitment of the monetary authority. We applied the method to U.S. monetary policy since 1960 and for Brazil since 1999. For the U.S. economy, it was not found evidence of asymmetry in the preferences of the monetary authority. In Brazil, the Central Bank proved to have asymmetric preferences about inflation, with a greater loss for negative deviations of inflation from the target rather for positive ones.

Palavras-chave: Central Banks' preferences. Asymmetric loss function. Sieve Estimators. Orthogonal polynomials.

Sumário

1	Introdução	9
2	A economia	13
2.1	Política monetária ótima	13
2.2	A função de reação e estática comparativa	15
3	<i>Sieve minimum distance estimate</i>	18
3.1	O estimador	18
3.2	Aproximação da função perda	22
3.3	Momentos condicionais e instrumentos	23
4	Resultados	25
4.1	Estados Unidos	25
4.2	Brasil	31
5	Conclusão	38
	Referências	40
A	Apêndice - Consistência do estimador	42

1 Introdução

A literatura dá grande atenção a condução da política monetária por parte dos Bancos Centrais. Eles agem para controlar a inflação e, dada a existência de fricções¹ na economia, afetam também o lado real no curto prazo. Com esse poder de afetar temporariamente o produto é possível que as autoridades monetárias tenham como objetivo, além da estabilidade de preços, a estabilidade do produto e do emprego.

No entanto, os Bancos Centrais nem sempre divulgam explicitamente seus objetivos. O que é mais facilmente observado é sua reação a mudanças nas variáveis econômicas - através das chamadas funções de reação. No trabalho seminal de Taylor (1993), a autoridade monetária ajusta a taxa nominal de juros respondendo a desvios da inflação em relação à sua meta e do hiato do produto.

A função de reação da autoridade monetária pode ser entendida como o resultado de um modelo onde o Banco Central busca minimizar uma função perda, condicionada pelas restrições impostas pela economia. Assim, observar a função de reação, o quanto a taxa de juros responde a desvios da inflação e do produto, por exemplo, não permite separar a parcela do efeito que vem das preferências da autoridade monetária daquela que é imposta pela estrutura da economia.

Existe na literatura uma vasta gama de trabalhos que buscam separar esses efeitos, estimando a preferências do Banco Central e não apenas a função de reação. A abordagem tradicional para este problema tem sido modelar a economia como equações lineares, *backward* ou *forward-looking*, e modelar a autoridade monetária como um agente que minimiza uma perda que é uma função quadrática da diferença entre a inflação e a sua meta e do hiato do produto,

$$L(\pi_t, x_t) = (\pi_t - \pi^*)^2 + \lambda x_t^2,$$

onde π é a inflação, π^* é a sua meta, x é o hiato do produto e λ é o seu peso relativo dado pela autoridade monetária. Uma grande vantagem dessa abordagem é a sua tratabilidade. Mesmo considerando o modelo sob comprometimento (isto é, as ações do Banco Central afetando as ações futuras dos agentes privados), ele resulta em um problema de programação dinâmica com função objetivo quadrática e funções de transição lineares, que pode ser resolvido analiticamente através da Equação de Riccati². Outra vantagem da função perda quadrática é que ela pode ser vista como uma aproximação de Taylor de segunda ordem da função de utilidade esperada de um agente representativo em um modelo de equilíbrio geral com expectativas racionais e fricções nos preços (Woodford, 2003, cap. 6).

Uma implicação de se assumir que a perda do Banco Central pode ser representada por uma função

¹Fricções ocorrem quando alterações no nível geral de preços da economia não são repassadas inteiramente aos preços e salários de uma determinada empresa ou setor, ocasionando alterações temporárias nos preços relativos.

²Ver, por exemplo, Ljungqvist e Sargent (2004, cap. 5).

quadrática é ele dar pesos iguais a desvios positivos e negativos de suas variáveis alvo. Por exemplo, a perda que a autoridade monetária teria com a inflação um ponto percentual acima da meta seria igual à obtida com a inflação um ponto abaixo da meta. Porém isso não é necessariamente verdade: para ganhar credibilidade é possível que o Banco Central tenha uma aversão maior a níveis de inflação superiores à meta, por motivos precaucionais. Um segundo ponto é que a abordagem linear-quadrática gera funções de reação lineares, com o efeito marginal sobre a taxa de juros da inflação sendo constante, tanto para taxas de inflação próximas ou distantes da sua meta, e o mesmo ocorrendo com o hiato do produto.

Uma primeira alternativa para se chegar a funções de reação que não são lineares é supor a não linearidade das funções estruturais da economia. Em um âmbito teórico, Nobay e Peel (2000) mostram que uma política discricionária sob uma curva de Phillips não linear gera um viés inflacionário se o produto for igual ao produto potencial. Schaling (2004), a partir de um modelo teórico onde a curva de Phillips é convexa (desvios positivos das variáveis são mais inflacionários do que os desvios negativos são desinflacionários), mostra que quanto maior a incerteza da economia, mais agressiva deveria ser a condução da política monetária. Utilizando dados norte-americanos para calibrar o seu modelo, Chiarella et al. (2003) modelam a economia como um sistema dinâmico de diversas equações, dentre as quais curvas de Phillips para preços e salários que apresentam uma maior rigidez para a queda das variáveis, e encontram que o equilíbrio desse sistema é instável. Por fim, Dolado et al. (2005), partindo de uma Curva de Phillips não linear (quadrática no hiato do produto), chegam a uma função de reação onde um dos termos é a interação (multiplicação) entra a expectativa de inflação e o hiato do produto. Utilizando dados mensais, os autores encontraram evidências de assimetria na condução da política monetária para Alemanha, França, Espanha e para a zona do Euro, porém não puderam rejeitar a hipótese de linearidade para o caso norte-americano.

A segunda alternativa para gerar funções de reação não lineares é modelar a função perda da autoridade monetária como uma função possivelmente assimétrica, diferentemente da função quadrática usual na literatura. Nessa linha, Orphanides e Wieland (2000) desenvolveram um modelo com não linearidades tanto na estrutura da economia quanto nas preferências da autoridade monetária. Eles consideram que a função perda da autoridade monetária tem um intervalo onde ela é linear, representando a banda compreendida entre os limites mínimo e máximo da meta de inflação. Além disso, também permitem que a curva de Phillips apresente uma porção linear. Como resultado, mostram que há um incentivo para que a autoridade monetária desvie de uma política linear. Cukierman e Muscatelli (2002) desenvolvem um modelo onde a especificação da função perda é genérica e, a partir da condição de primeira ordem da autoridade monetária, fazem um exercício de estática comparativa. Kim et al. (2005) estimam a função perda do Banco Central não-parametricamente através do método proposto por Hamilton (2001)³, encontrando assimetria da função para o período pré-Volcker mas linearidade para a amostra completa e

³Hamilton (2001) propôs estimar um modelo da forma $y = \mu(x) + \epsilon$, onde $\mu(\bullet)$ é uma função desconhecida, considerando μ , ela própria, como uma variável aleatória.

para o período Volcker-Greenspan.

Ainda nesta linha, Surico (2007), baseado em Nobay e Peel (2003), estimou a assimetria da função perda do Banco Central norte-americano, tanto para a inflação quanto para o hiato do produto, com a especificação *linex*⁴,

$$L(\pi_t, x_t) = \frac{e^{\alpha(\pi_t - \pi^*)} - \alpha(\pi_t - \pi^*) - 1}{\alpha^2} + \lambda \left(\frac{e^{\gamma x_t} - \gamma x_t - 1}{\gamma^2} \right),$$

onde os parâmetros α e γ captam a assimetria da função. Com $\alpha = \gamma = 0$, a especificação *linex* colapsa na função perda quadrática. Modelando o comportamento da autoridade monetária como discricionária, o autor encontrou evidências de assimetria na perda relativa ao hiato do produto para o período pré-Volcker, com o Banco Central mais preocupado com desvios negativos do que positivos do produto. Já para a inflação ele não rejeitou a hipótese de simetria da função perda. A metodologia de Surico (2007) foi aplicada ao Brasil por Aragón e Portugal (2010), que encontraram uma assimetria na perda do Banco Central Brasileiro com relação à inflação para o período 2000-2007, com a inflação abaixo da meta causando uma perda de bem estar maior do que a inflação acima da meta. Já para o subperíodo 2004-2007, ou para o hiato do produto na amostra completa ou nesse subperíodo, não encontraram evidências de assimetria.

A literatura relacionada às preferências à função de reação da autoridade monetária vem caminhando para um relaxamento das hipóteses do modelo, entre elas a que diz que as preferências são simétricas. A especificação *linex* utilizada por Nobay e Peel (2003), Surico (2007) e Aragón e Portugal (2010) generaliza a função perda, mas ainda assim é uma suposição paramétrica sobre o comportamento do Banco Central.

O caminho natural para a evolução desta literatura é reduzir ainda mais as restrições à forma funcional das preferências da autoridade monetária. Assim, o objetivo deste trabalho é aplicar um método de estimação da função perda do Banco Central totalmente não-paramétrico. Para tal é desenvolvido um modelo de otimização com comprometimento da autoridade monetária e, a partir das restrições resultantes sobre os momentos condicionais das variáveis, usar o método de *sieve* para estimar a função que representa as preferências do Banco Central. O método é empregado para dados norte-americanos e brasileiros. Assim será possível saber se as preferências das duas autoridades monetárias são assimétricas em relação à inflação e (ou) ao produto e qual é a forma exata das suas funções perda.

Estimar a forma de uma função implica minimizar a função critério do estimador sobre um espaço de dimensão infinita, o que apresenta diversas dificuldades. O método de *sieve*, proposto por Grenander (1981), consiste em minimizar a função critério sobre uma sequência de espaços mais simples, geralmente de dimensão finita, cuja dimensão aumenta com o tamanho da amostra.

Além desta introdução, este trabalho está dividido como segue: na seção seguinte é descrito como a economia é modelada; na terceira parte é apresentado o estimador utilizado para a estimação; os

⁴No modelo do autor, a função perda dependia ainda de um parâmetro quadrático na suavização da taxa de juros.

resultados para os EUA e para o Brasil são apresentados na quarta parte; uma breve conclusão é dada em seguida.

2 A economia

A economia é descrita pelo modelo de Svensson e Woodford (2005), caracterizadas por duas equações estruturais: uma curva de oferta agregada (AS) e uma curva IS aumentada pelas expectativas, ambas *forward-looking*. Estas duas equações representam a log-linearização do equilíbrio de um modelo micro-fundamentado, como em Clarida et al. (1999), diferindo apenas por incorporar um erro aleatório não previsível e não afetado pela política monetária. O modelo é apresentado como a seguinte hipótese:

Hipótese 1. *A economia é representada pelas equações*

$$\pi_{t+1} = \beta E_t \pi_{t+2} + \kappa E_t x_{t+1} + u_{t+1} \quad (2.1)$$

$$x_{t+1} = E_t x_{t+2} - \phi (E_t i_{t+1} - E_t \pi_{t+2} - r_{t+1}), \quad (2.2)$$

onde π_{t+1} é a inflação entre os períodos t e $t + 1$, x_{t+1} é o hiato do produto, i_t é a taxa nominal de juros de curto prazo (em desvio em relação ao seu valor de steady state) e u_{t+1} e r_{t+1} são termos de erro. Além disso, $0 \leq \beta < 1$ é o fator de desconto intertemporal e $\kappa > 0$ e $\phi > 0$ são constantes. E_t é o operador da esperança condicional à informação do instante t .

Note que, tanto na AS (equação (2.1)) quanto na IS (equação (2.2)) a esperança é tomada um período antes do que o usual na literatura, devido ao fato dos preços serem pré-determinados nesta economia. Para a análise que segue também é necessária uma suposição sobre a evolução dos termos de erro.

Hipótese 2. *Os termos de erro u_{t+1} e r_{t+1} seguem as seguintes equações de transição*

$$u_{t+1} = \rho u_t + \varepsilon_{t+1}$$

$$r_{t+1} = \bar{r} + \omega (r_t - \bar{r}) + \eta_{t+1},$$

onde ε_{t+1} e η_{t+1} são variáveis aleatórias independentes e identicamente distribuídas, $0 \leq \rho < 1$ e $0 \leq \omega < 1$ são constantes e \bar{r} é a média da taxa natural de juros real.

2.1 Política monetária ótima

Dadas as equações que descrevem a economia, a autoridade monetária busca minimizar a sua *perda*, que se supõe ser função do desvio da inflação em relação à sua meta, do hiato do produto e do desvio da taxa de juros em relação à sua meta. Os dois primeiros objetivos são usuais na literatura; já o último, que

representa uma preferência por uma menor variabilidade da taxa de juros, pode ser justificado por uma preocupação da autoridade em relação a fricções nas transações, como apresentado em Woodford (2003, 6.4.1, p. 420). Assim como a utilidade na teoria do consumidor, essa perda pode ser representada por uma função com certas características.

Hipótese 3. *A autoridade monetária pode ser descrita por a função perda separável*

$$L(\pi_t - \pi_t^*, x_t, i_t) = \tilde{L}(\pi_t - \pi_t^*, x_t) + \frac{\gamma}{2} (i_t - i_t^*)^2, \quad (2.3)$$

onde π_t^* e i_t^* são as metas de inflação e juros⁵, respectivamente, para o período t e $\gamma > 0$ é um parâmetro que mede o peso que a autoridade dá à suavização dos juros.

Hipótese 4. *A função $\tilde{L}(\bullet)$, que mede a perda com a inflação e com o hiato do produto,*

- (i) *tem como ponto mínimo $(0, 0)$;*
- (ii) *passa por $\tilde{L}(0, 0) = 0$;*
- (iii) *é limitada;*
- (iv) *é supermodular, isto é,*

$$\frac{\partial^2 \tilde{L}(\bullet)}{\partial \tilde{\pi}_t^2} \geq 0, \quad \frac{\partial^2 \tilde{L}(\bullet)}{\partial x_t^2} \geq 0, \quad \frac{\partial^2 \tilde{L}(\bullet)}{\partial \tilde{\pi}_t \partial x_t} \geq 0, \quad (2.4)$$

onde $\tilde{\pi}_t = \pi_t - \pi_t^*$ é o desvio da inflação em relação à sua meta;

(v) *pertence ao espaço das funções quadrado-integráveis, $\tilde{L}(\bullet) \in \mathcal{L}^2(\chi)$, definidas em um conjunto compacto χ .*

Supõe-se que o ponto mínimo da função perda da autoridade monetária se encontre no ponto $(0, 0)$ pois, caso contrário, a função implicaria uma meta (ponto que o Banco Central deseja alcançar) diferente do especificado anteriormente, que é o correntemente utilizado na literatura. Já as hipóteses de supermodularidade e de ser limitada garantirá a unicidade da função reação referente à política monetária ótima.

A hipótese (3) é a principal contribuição deste trabalho, por não se supor nenhuma função perda específica. Ela ser separável na taxa de juros é uma simplificação do presente modelo, porém ela é recorrente nos outros trabalhos da área, inclusive ao se supor uma função perda quadrática. Além disso, a função quadrática também atende a hipótese de supermodulariedade.

Dadas as restrições geradas pela estrutura da economia, a autoridade monetária minimiza o valor presente da sua função perda. Aqui se assume que ela o faz com comprometimento, isto é, ela leva em conta o efeito das suas ações sobre as expectativas do setor privado e os efeitos destas sobre as

⁵Note que a meta da autoridade monetária para o hiato do produto é assumida como igual a zero, como usualmente empregado na literatura. Como os juros são expressos em desvios em relação ao seu valor de *steady state* a meta de juros também é considerada como igual a zero.

suas possíveis ações futuras. O instrumento sob poder do Banco Central é a taxa nominal de juros de curto prazo, porém ela também afeta as expectativas sobre inflação e hiato do produto da iniciativa privada. Assim, o seu problema de otimização é minimizar o valor presente de (2.3) sujeito a (2.1) e (2.2), representado pelo seguinte Lagrangeano:

$$\begin{aligned} \min_{\{\pi_{t+1}, x_{t+1}, i_{t+1}\}} \mathfrak{L} &= E_{t_0} \sum_{t=t_0}^{\infty} E_t [\beta^{t+1-t_0} L(\tilde{\pi}_{t+1}, x_{t+1}, \tilde{i}_{t+1}) \\ &+ \lambda_{1,t+1} (\pi_{t+1} - \beta E_t \pi_{t+2} - \kappa E_t x_{t+1} - u_{t+1}) \\ &+ \lambda_{2,t+1} (x_{t+1} - E_t x_{t+2} + \phi (E_t i_{t+1} - E_t \pi_{t+2} - r_{t+1}))], \end{aligned} \quad (2.5)$$

onde $\tilde{i}_t = i_t - i_t^*$ é o desvio da taxa de juros em relação à sua meta.

As condições de primeira ordem do problema são

$$\frac{\partial \mathfrak{L}}{\partial \pi_{t+1}} = E_t [\beta^{t+1-t_0} \tilde{L}_{\pi,t+1} + \lambda_{1,t+1} - \beta \lambda_{1,t} - \lambda_{2,t}] = 0 \quad (2.6)$$

$$\frac{\partial \mathfrak{L}}{\partial x_{t+1}} = E_t [\beta^{t+1-t_0} \tilde{L}_{x,t+1} - \kappa \lambda_{1,t+1} + \lambda_{2,t+1} - \lambda_{2,t}] = 0 \quad (2.7)$$

$$\frac{\partial \mathfrak{L}}{\partial i_{t+1}} = E_t [\beta^{t+1-t_0} \gamma \tilde{i}_{t+1} + \phi \lambda_{2,t+1}] = 0, \quad (2.8)$$

$\forall t = t_0 + 1, \dots$, onde, por simplicidade de notação, $\tilde{L}_{\pi,t+1} = \partial \tilde{L}(\tilde{\pi}_{t+1}, x_{t+1}) / \partial \pi_{t+1}$, análogo para $\tilde{L}_{x,t+1}$. Isolando $\lambda_{2,t+1}$ em (2.8) e substituindo em (2.7), pode-se resolver essa equação para $\lambda_{1,t+1}$, de forma que

$$\lambda_{1,t+1} = \kappa^{-1} \left(\beta^{t+1-t_0} \tilde{L}_{x,t+1} - \beta^{t+1-t_0} \frac{\gamma}{\phi} \tilde{i}_{t+1} + \beta^{t-t_0} \frac{\gamma}{\phi} \tilde{i}_t \right).$$

Substituindo os multiplicadores em (2.6), chega-se a Equação de Euler da autoridade monetária sob comprometimento para a economia descrita nesta seção,

$$E_t \left[L_{\pi,t+1} + \frac{(L_{x,t+1} - L_{x,t})}{\kappa} - \frac{\gamma}{\kappa \phi} \tilde{i}_{t+1} + \left(\frac{\gamma}{\kappa \beta \phi} + \frac{\gamma}{\kappa \phi} + \frac{\gamma}{\beta} \right) \tilde{i}_t - \frac{\gamma}{\kappa \beta \phi} \tilde{i}_{t-1} \right] = 0. \quad (2.9)$$

Como a equação (2.9) deve ser satisfeita em todo período $t = t_0 + 1, \dots$ para que a autoridade monetária esteja otimizando a suas preferências, ela é uma candidata à regra de política ótima Woodford (2003, seção 7.5.1).

2.2 A função de reação e estática comparativa

Segundo Woodford (2003), o Banco Central deve escolher o seu instrumento, i_{t+1} , através de uma função de reação que garanta a validade da regra de política ótima. Isto é, i_{t+1} deve garantir que as projeções sobre a inflação e o hiato do produto, $E_t \pi_{t+1}$ e $E_t x_{t+1}$ sejam consistentes com (2.9). Uma forma é supor

que os agentes privados formem suas expectativas racionalmente, com as projeções acerca dessas duas variáveis sejam dadas pela esperança em t definida nas equações (2.1) e (2.2),

$$E_t x_{t+1} = E_t (x_{t+2} - \phi(i_{t+1} - \pi_{t+2} - r_{t+1})) \quad (2.10)$$

$$E_t \pi_{t+1} = E_t ((\beta + \kappa\phi)\pi_{t+2} + \kappa x_{t+2} - \kappa\phi(i_{t+1} - r_{t+1}) + u_{t+1}). \quad (2.11)$$

Assim, a autoridade monetária deve, em cada instante t , escolher i_{t+1} que satisfaça (2.9) com x_{t+1} e π_{t+1} substituídos por (2.10) e (2.11), respectivamente. Dada a natureza não paramétrica do problema, não se tem uma solução explícita para i_{t+1} e, além disso, não se pode garantir *a priori* que essa escolha caracterize uma função de reação implícita na regra de política ótima. A existência da função de reação é apresentada na proposição seguinte.

Proposição 1. *Existe uma função de reação $i_{t+1} = f(\bullet)$ implícita na regra de política ótima (2.9).*

Demonstração. Seja $\varphi(i_{t+1}|\bullet)$ o lado esquerdo da equação (2.9). A função de reação existe se e somente se existir i_{t+1} que soluciona $\varphi(i_{t+1}|\bullet) = 0$ for ele for único. Para que i_{t+1} exista e seja único é suficiente que a função $\varphi(i_{t+1}|\bullet)$ apresente uma troca de sinais e que o sinal de $\partial\varphi(i_{t+1}|\bullet)/\partial i_{t+1}$ seja igual ao sinal de $\lim_{i_{t+1} \rightarrow \infty} \varphi(\infty|\bullet)$. Como $\lim_{i_{t+1} \rightarrow -\infty} \varphi(-\infty|\bullet) = \infty > 0$ e $\lim_{i_{t+1} \rightarrow \infty} \varphi(\infty|\bullet) = -\infty < 0$ a função apresenta solução. Já $\partial\varphi(i_{t+1}|\bullet)/\partial i_{t+1} = -\kappa\phi\tilde{L}_{\pi\pi} - 2\phi\tilde{L}_{\pi x} - \phi/\kappa\tilde{L}_{xx} - \gamma/\kappa\phi < 0$ devido a todos os parâmetros serem positivos (hipóteses (1) e (3)) e a função perda ser supermodular (hipótese (4)), garantindo a unicidade da solução (onde $\tilde{L}_{\pi\pi} = \partial^2 \tilde{L}(\bullet)/\partial \pi^2$, sendo análogo para as outras derivadas parciais da equação). \square

Uma vez provado que a função de reação exista, ela dá a taxa nominal de juros de curto prazo ótima que o Banco Central deve escolher para satisfazer a sua regra de política monetária. No presente modelo a taxa de juros ótima em t é uma função $i_{t+1} = f(\pi_t, x_t, i_t, i_{t-1}, E_t \pi_{t+2}, E_t x_{t+2}, E_t u_{t+1}, E_t r_{t+1})$ do estado corrente da economia (inflação e hiato do produto em t e taxa de juros em t e $t-1$) e da expectativa em t quanto ao estado futuro da economia (inflação, $E_t \pi_{t+2}$, hiato do produto, $E_t x_{t+2}$ e choques da economia, $E_t u_{t+1}$ e $E_t r_{t+1}$). Diferenciando (2.9) implicitamente, pode-se obter a resposta marginal da função que define a taxa de juros em relação aos seus argumentos,

$$\begin{aligned}
\frac{\partial i_{t+1}}{\partial E_t \pi_{t+2}} &= \frac{(\beta + \kappa\phi) \tilde{L}_{\pi\pi} + (2\phi + \beta/\kappa) \tilde{L}_{\pi x} + \phi/\kappa \tilde{L}_{xx}}{\kappa\phi \tilde{L}_{\pi\pi} + 2\phi \tilde{L}_{\pi x} + \phi/\kappa \tilde{L}_{xx} + \gamma/\kappa\phi} > 0 \\
\frac{\partial i_{t+1}}{\partial E_t x_{t+2}} &= \frac{\kappa \tilde{L}_{\pi\pi} + 2\tilde{L}_{\pi x} + 1/\kappa \tilde{L}_{xx}}{\kappa\phi \tilde{L}_{\pi\pi} + 2\phi \tilde{L}_{\pi x} + \phi/\kappa \tilde{L}_{xx} + \gamma/\kappa\phi} > 0 \\
\frac{\partial i_{t+1}}{\partial E_t u_{t+1}} &= \frac{\tilde{L}_{\pi\pi} + 1/\kappa \tilde{L}_{\pi x}}{\kappa\phi \tilde{L}_{\pi\pi} + 2\phi \tilde{L}_{\pi x} + \phi/\kappa \tilde{L}_{xx} + \gamma/\kappa\phi} > 0 \\
\frac{\partial i_{t+1}}{\partial E_t r_{t+1}} &= \frac{\kappa\phi \tilde{L}_{\pi\pi} + 2\phi \tilde{L}_{\pi x} + \phi/\kappa \tilde{L}_{xx}}{\kappa\phi \tilde{L}_{\pi\pi} + 2\phi \tilde{L}_{\pi x} + \phi/\kappa \tilde{L}_{xx} + \gamma/\kappa\phi} \in (0, 1) \\
\frac{\partial i_{t+1}}{\partial \pi_t} &= -\frac{1/\kappa \tilde{L}_{\pi x}}{\kappa\phi \tilde{L}_{\pi\pi} + 2\phi \tilde{L}_{\pi x} + \phi/\kappa \tilde{L}_{xx} + \gamma/\kappa\phi} < 0 \\
\frac{\partial i_{t+1}}{\partial x_t} &= -\frac{1/\kappa \tilde{L}_{xx}}{\kappa\phi \tilde{L}_{\pi\pi} + 2\phi \tilde{L}_{\pi x} + \phi/\kappa \tilde{L}_{xx} + \gamma/\kappa\phi} < 0 \\
\frac{\partial i_{t+1}}{\partial i_t} &= \frac{\gamma/\kappa\beta\phi + \gamma/\kappa\phi + \gamma/\beta}{\kappa\phi \tilde{L}_{\pi\pi} + 2\phi \tilde{L}_{\pi x} + \phi/\kappa \tilde{L}_{xx} + \gamma/\kappa\phi} > 0 \\
\frac{\partial i_{t+1}}{\partial i_{t-1}} &= -\frac{\gamma/\kappa\beta\phi}{\kappa\phi \tilde{L}_{\pi\pi} + 2\phi \tilde{L}_{\pi x} + \phi/\kappa \tilde{L}_{xx} + \gamma/\kappa\phi} < 0.
\end{aligned}$$

Vê-se assim que a resposta dos juros no modelo apresentado aqui, que não restringe a forma funcional da preferência da autoridade monetária, apresenta os mesmos sinais que a tradicional, com o Banco Central aumentando os juros frente a um aumento das expectativas da inflação, do hiato do produto e dos choques. Além disso, ela tem uma resposta negativa quanto à variável de estado representada pela inflação e pelo hiato do produto do período anterior. Outro fato é que o efeito marginal das taxas de juros passadas sobre a função de reação oscila e decai com o tempo, $|\partial i_{t+1}/\partial i_t| > |\partial i_{t+1}/\partial i_{t-1}|$.

3 Sieve minimum distance estimate

Como exposto nas seções anteriores, o propósito deste trabalho é estimar a função perda do Banco Central. Estimar uma função significa estimar um parâmetro pertencente a um espaço de dimensão infinita, o que pode ser difícil de implementar, além de poder apresentar propriedades assintóticas não desejáveis, como inconsistência e uma baixa taxa de convergência (Chen, 2007, p. 1).

Grenander (1981) propôs estimar os parâmetros sobre uma sequência de espaços menos complexos, usualmente de dimensão finita, e deixar que a dimensão desses espaços cresça conforme cresce o tamanho da amostra, porém mais lentamente para que o estimador seja consistente.

Para garantir as propriedades desejadas do estimador, é necessário, além da taxa de crescimento da complexidade dos espaços utilizado na aproximação, que este aproxime bem o verdadeiro espaço onde o parâmetro populacional está contido. Para que ele seja um bom aproximador, é necessário que o espaço aproximado seja denso no espaço verdadeiro. Um conjunto \mathcal{M} ser denso em outro, \mathcal{N} , implica que, para qualquer ponto $n \in \mathcal{N}$, exista um ponto $m \in \mathcal{M}$ suficiente próximo, ou, equivalentemente, toda a bola aberta de \mathcal{N} inclua um ponto de \mathcal{M} . Assim, se o espaço aproximado for assintoticamente denso no espaço original, pode-se encontrar um estimador para o parâmetro populacional suficientemente próximo quando o tamanho da amostra cresce.

Como a função perda pertence a $\mathcal{L}_2(\chi)$, $\chi \subset \mathbb{R}^2$ compacto, o espaço que será usado como aproximação deve ser denso neste conjunto. Um espaço gerado por uma base composta de polinômios ortogonais que levam de χ em \mathbb{R} é denso em $\mathcal{L}_2(\chi)$.

A aproximação de uma função por uma base do seu espaço tem um caráter global, onde se encontra a função que seja mais próxima à função original, com o conceito de proximidade dado por uma norma do espaço que, em geral, leva em consideração todos os pontos do domínio. Isto difere do que ocorre com a aproximação por Taylor, que tem caráter local, isto é, se iguala o nível e a inclinação da função original em um dado ponto do domínio, mas o que não garante que essa aproximação seja boa em pontos fora dessa vizinhança.

3.1 O estimador

Nesta seção será mostrada a versão genérica do estimador, como apresentado em Ai e Chen (2003). Em modelos de expectativas racionais, é usual que, a partir do comportamento maximizador dos agentes, chegue-se a restrições nos momentos condicionais que as variáveis devem satisfazer.

Sejam (\mathbf{Y}, \mathbf{X}) as variáveis da economia, onde \mathbf{Y} é o vetor com os dados endógenos e \mathbf{X} o vetor com dados exógenos ao modelo ou pré-determinados. Então a condição de momento pode ser escrito como

$$E[\boldsymbol{\rho}(\mathbf{Z}, \boldsymbol{\theta}_0, \mathbf{h}_0(\bullet)) | \mathbf{X}] = \mathbf{0}, \quad (3.1)$$

onde $\mathbf{Z} = (\mathbf{Y}', \mathbf{X}'_z)$, $\mathbf{X}_z \subseteq \mathbf{X}$, $\boldsymbol{\rho}(\bullet)$ é um vetor de funções desconhecidas que dão as restrições sobre os momentos (resíduos, já que a esperança está igualada a zero). $E[\boldsymbol{\rho}(\mathbf{Z}, \boldsymbol{\theta}_0, \mathbf{h}_0(\bullet)) | \mathbf{X}]$ é a esperança condicional em \mathbf{X} de $\boldsymbol{\rho}(\mathbf{Z}, \boldsymbol{\theta}_0, \mathbf{h}_0(\bullet))$. $\boldsymbol{\theta}_0$ é um vetor com parâmetros contidos em um espaço de dimensão finita e $\mathbf{h}_0(\bullet) = (h_{01}(\bullet), \dots, h_{0q}(\bullet))$ é um vetor de funções desconhecidas. Deseja-se estimar $\boldsymbol{\alpha}_0 = (\boldsymbol{\theta}_0, \mathbf{h}_0(\bullet))$, o que torna o problema semi-paramétrico, no sentido ter-se parâmetros de interesse tanto de dimensão finita quanto infinita. Esta propriedade cabe perfeitamente ao problema do presente trabalho, onde além da função preferência da autoridade monetária, há parâmetros estruturais da economia que precisam ser estimados.

Como as variáveis apresentam dependência temporal é necessário estabelecer condições sob as quais, assintoticamente, possa-se utilizar o teorema do limite central. Neste caso, requer-se que os dados sejam β -mixing estacionários.

Definição 1. *Seja $\{Y_t, t = 0, \pm 1, \pm 2, \dots\}$ uma série temporal estritamente estacionária e defina*

$$\beta(n) = E \left[\sup_{B \in \mathcal{F}_n^\infty} |P(B) - P(B|Y_0, Y_{-1}, Y_{-2}, \dots)| \right],$$

onde \mathcal{F}_n^∞ é a σ -álgebra⁶ gerada por $\{Y_t, n \leq t < \infty\}$. Diz-se que o processo $\{Y_t\}$ é β -mixing se $\beta(n) \rightarrow 0$ quando $n \rightarrow \infty$. (Fan e Yao, 2005, p. 68)

Intuitivamente, um processo que seja β -mixing é assintoticamente independente. Isto porque $\beta(n) \rightarrow 0$ quando $n \rightarrow \infty$ implica que, para todo possível evento B (todo possível trajeto de Y_t entre n e ∞), as suas probabilidades incondicional e condicional (a Y_t passados) são iguais, o que caracteriza a independência assintótica entre $\{Y_t, n \leq t < \infty\}$ e $\{Y_0, Y_{-1}, Y_{-2}, \dots\}$.

Dispõe-se de uma amostra $\{(\mathbf{Y}_t, \mathbf{X}_t), t = 1, \dots, N\}$ para a estimação de (3.1), onde $(\mathbf{Y}_t, \mathbf{X}_t)$ é um processo β -mixing estacionário. As restrições sobre os momentos condicionais, $\boldsymbol{\rho} : \mathcal{Z} \times \mathcal{A} \rightarrow \mathbb{R}^{d_\rho}$, são conhecidas a menos dos parâmetros, $\boldsymbol{\alpha}_0 = (\boldsymbol{\theta}_0, \mathbf{h}_0(\bullet)) \in \mathcal{A} = \Theta \times \mathcal{H}$, onde Θ é o espaço dos parâmetros finito-dimensionais e $\mathcal{H} = \mathcal{H}_1 \times \dots \times \mathcal{H}_q$ é o espaço funcional onde \mathbf{h}_0 está contido. Se a função de distribuição de \mathbf{Y} dado \mathbf{X} , $F_{\mathbf{Y}|\mathbf{X}}$, fosse conhecida, poder-se-ia calcular cada um dos componentes de (3.1) como

$$E[\rho_i(\mathbf{Z}, \boldsymbol{\theta}_0, \mathbf{h}_0(\bullet)) | \mathbf{x}] \equiv m_i(\mathbf{x}, \boldsymbol{\alpha}) = \int \rho_i(\mathbf{y}, \mathbf{x}_z, \boldsymbol{\theta}, \mathbf{h}(\bullet)) dF_{\mathbf{Y}|\mathbf{X}=\mathbf{x}}(\mathbf{y}),$$

⁶Uma σ -álgebra em um conjunto Ω é uma coleção \mathcal{A} de subconjuntos de Ω tal que o conjunto vazio pertença a ela e que, dados $\mathcal{A}_n \in \mathcal{A}$, $(\bigcup_{n=1}^\infty \mathcal{A}_n) \in \mathcal{A}$ e $\mathcal{A}_n^c \in \mathcal{A}$ (Isnard, 2007, p. 38). Estas características da σ -álgebra são fundamentais na teoria da probabilidade pois garantem que um evento composto pela união de outros eventos que têm probabilidade definida também terá probabilidade e, além disso, se um evento tem uma certa probabilidade associada a ele, então também o tem o seu evento complementar.

$i = 1, \dots, d_\rho$. Seja $\mathbf{m}(\mathbf{x}, \boldsymbol{\alpha}) = (m_1(\mathbf{x}, \boldsymbol{\alpha}), \dots, m_{d_\rho}(\mathbf{x}, \boldsymbol{\alpha}))'$. Então o parâmetro verdadeiro $\boldsymbol{\alpha}_0$ minimiza

$$\inf_{\boldsymbol{\alpha}=(\theta, h) \in \Theta \times \mathcal{H}} E \left[\mathbf{m}(\mathbf{X}, \boldsymbol{\alpha})' [\boldsymbol{\Sigma}(\mathbf{X})]^{-1} \mathbf{m}(\mathbf{X}, \boldsymbol{\alpha}) \right], \quad (3.2)$$

onde $\boldsymbol{\Sigma}(\mathbf{X})$ é uma matriz positiva definida.

Um estimador natural para $\boldsymbol{\alpha}_0$ seria o argumento que minimizasse a versão amostral de (3.2). Porém, para viabilizar essa minimização, duas questões devem ser tratadas. Primeiramente, por não se conhecer a verdadeira forma funcional de $F_{Y|\mathbf{X}}$ e, conseqüentemente, de $\mathbf{m}(\mathbf{X}, \boldsymbol{\alpha})$, esta última função deve ser estimada por qualquer método não-paramétrico. Ai e Chen (2003) sugerem também estimá-la por *sieve*, mais precisamente por *linear sieve estimator*, que será discutido a seguir. Um segundo ponto é o fato de minimizar sobre o conjunto \mathcal{H} , de dimensão infinita, pode ser complexo e gerar propriedades indesejáveis nos estimadores. Como já discutido antes, uma alternativa é o método de *sieve*, substituindo \mathcal{H} por uma sequência de conjuntos mais simples, que sejam densos no original quando o tamanho da amostra N cresça. Assim, o processo de otimização se dá sobre o conjunto $\mathcal{H}_n = \mathcal{H}_n^1 \times \dots \times \mathcal{H}_n^q$, mais simples do que o original e, em geral, de dimensão finita. Com o tratamento apresentado para essas duas questões, tem-se o estimador *sieve minimum distance* (SMD),

$$\begin{aligned} \hat{\boldsymbol{\alpha}}_n &= (\hat{\boldsymbol{\theta}}_n, \hat{\mathbf{h}}_n) \\ &= \arg \min_{\boldsymbol{\alpha}=(\theta, h) \in \Theta \times \mathcal{H}_n} n^{-1} \sum_{t=1}^N \hat{\mathbf{m}}(\mathbf{X}_t, \boldsymbol{\alpha})' [\hat{\boldsymbol{\Sigma}}(\mathbf{X})]^{-1} \hat{\mathbf{m}}(\mathbf{X}_t, \boldsymbol{\alpha}), \end{aligned} \quad (3.3)$$

onde $\hat{\boldsymbol{\Sigma}}(\mathbf{X})$ é um estimador consistente para $\boldsymbol{\Sigma}(\mathbf{X})$, utilizado para tratar uma possível heterocedasticidade, e $\hat{\mathbf{m}}(\mathbf{X}, \boldsymbol{\alpha})$ é o *linear sieve estimator* para $\mathbf{m}(\mathbf{X}, \boldsymbol{\alpha})$.

O estimador $\hat{\mathbf{m}}(\mathbf{X}, \boldsymbol{\alpha})$ consiste em expandir a função original em uma base do seu espaço. Seja $\{p_j(\mathbf{X}), j = 1, 2, \dots\}$ uma base do \mathcal{L}^2 e $\mathbf{p}^{k_n}(\mathbf{X}) = (p_1(\mathbf{X}), \dots, p_{k_n}(\mathbf{X}))'$ o vetor que contém todas essas bases de ordem menor ou igual a k_n . Assim, essa base pode aproximar os momentos condicionais,

$$m_i(\mathbf{X}, \boldsymbol{\alpha}) = \sum_{j=0}^{k_n} c_{i,j} p_j(\mathbf{X}, \boldsymbol{\alpha}) = \mathbf{p}^{k_n}(\mathbf{X})' \mathbf{c}_i, \quad k_n \rightarrow \infty, \quad (3.4)$$

$i = 1, \dots, d_\rho$, onde $\mathbf{c}_i = (c_{i,1}, \dots, c_{i,k_n})'$ é o vetor com os coeficientes da i -ésima condição de momento. Os autores sugerem como *linear sieve estimator*, para uma dada ordem de aproximação k_n , estimar os coeficientes $\{c_{i,j}, i = 1, \dots, d_\rho, j = 1, \dots, k_n\}$ por mínimos quadrados ordinários⁷, regredindo $\rho_i(\mathbf{Z}, \boldsymbol{\alpha})$ contra $\mathbf{p}^{k_n}(\mathbf{X})$. Sejam $\boldsymbol{\rho}_i(\boldsymbol{\alpha}) = (\rho_i(\mathbf{Z}_1, \boldsymbol{\alpha}), \dots, \rho_i(\mathbf{Z}_N, \boldsymbol{\alpha}))'$ o vetor com a avaliação dessa condição de momento em cada ponto da amostra, definição análoga para $\mathbf{m}_i(\boldsymbol{\alpha})$, e \mathbf{P} a matriz $N \times k_n$ cujo elemento

⁷Para uma melhor descrição de *linear series estimators*, ver Newey (1997).

característico é $P_{t,j} = p_j^{k_n}(\mathbf{X}_t)$, ou ainda, $\mathbf{P} = (\mathbf{p}^{k_n}(\mathbf{X}_1), \dots, \mathbf{p}^{k_n}(\mathbf{X}_N))'$. Então, a regressão para cada restrição de momento, em notação matricial, é dada por

$$\hat{m}_i(\boldsymbol{\alpha}) = \hat{\rho}_i(\boldsymbol{\alpha}) = \mathbf{P}\hat{c}_i, \quad i = 1, \dots, d_\rho,$$

com o estimador $\hat{c}_i = (\mathbf{P}'\mathbf{P})^{-1}\mathbf{P}'\rho_i(\boldsymbol{\alpha})$. Substituindo-o em (3.4) chega-se a uma forma explícita para a esperança das restrições de momento⁸,

$$\hat{m}_i(\mathbf{X}, \boldsymbol{\alpha}) = \mathbf{p}^{k_n}(\mathbf{X}, \boldsymbol{\alpha})' (\mathbf{P}'\mathbf{P})^{-1} \mathbf{P}'\rho_i(\boldsymbol{\alpha}), \quad i = 1, \dots, d_\rho. \quad (3.5)$$

Em (3.5) tem-se uma equação para cada um dos k_n momentos. Elas podem ser agregadas em uma única equação,

$$\hat{\mathbf{M}}(\mathbf{x}, \boldsymbol{\alpha}) = \boldsymbol{\Xi}(\boldsymbol{\alpha}) \mathbf{P} (\mathbf{P}'\mathbf{P})^{-1} \mathbf{p}^{k_n}(\mathbf{x}), \quad (3.6)$$

onde $\hat{\mathbf{M}}(\mathbf{x}, \boldsymbol{\alpha}) = (\hat{m}_1(\mathbf{x}, \boldsymbol{\alpha}), \dots, \hat{m}_{d_\rho}(\mathbf{x}, \boldsymbol{\alpha}))'$ e $\boldsymbol{\Xi}(\boldsymbol{\alpha}) = (\rho_1(\mathbf{X}, \boldsymbol{\alpha}), \dots, \rho_{d_\rho}(\mathbf{X}, \boldsymbol{\alpha}))'$.

Note que a esperança dos momentos condicionais estimada, avaliada em um ponto \mathbf{x} qualquer, é uma função dos parâmetros desconhecidos do modelo, tanto dos de dimensão finita quanto dos infinito-dimensionais.

Substituindo essa esperança, (3.6), em (3.3) tem-se a função critério que precisa ser minimizada para se chegar ao SMD,

$$n^{-1} \sum_{t=1}^N \hat{\mathbf{M}}(\mathbf{X}_t, \boldsymbol{\alpha})' [\hat{\boldsymbol{\Sigma}}(\mathbf{X})]^{-1} \hat{\mathbf{M}}(\mathbf{X}_t, \boldsymbol{\alpha}).$$

O estimador por SMD é idêntico ao estimador pelo método generalizado dos momentos (GMM) usando a matriz identidade como matriz de pesos, $\hat{\boldsymbol{\Sigma}}(\mathbf{X}) = \mathbf{I}$. Este resultado é demonstrado por Ai e Chen (2003, p. 1799). Com $\hat{\boldsymbol{\Sigma}}(\mathbf{X}) = \mathbf{I}$, a função critério do SMD pode ser escrita como

$$\left(\sum_{t=1}^N \boldsymbol{\rho}(\mathbf{Z}_t, \boldsymbol{\alpha}) \otimes \mathbf{p}^{k_n}(\mathbf{X}_t) \right) (\mathbf{I} \otimes (\mathbf{P}'\mathbf{P})^{-1}) \left(\sum_{t=1}^N \boldsymbol{\rho}(\mathbf{Z}_t, \boldsymbol{\alpha}) \otimes \mathbf{p}^{k_n}(\mathbf{X}_t) \right)', \quad (3.7)$$

onde \mathbf{I} é a matriz identidade $d_\rho \times d_\rho$ e \otimes denota o produto de Kronecker. (3.7) é a função critério do estimador de GMM do sistema de restrições de momentos incondicionais

$$E[p_j(\mathbf{X})\boldsymbol{\rho}(\mathbf{Z}, \boldsymbol{\alpha}_o)] = 0, \quad j = 1, \dots, k_n$$

com $\mathbf{I} \otimes (\mathbf{P}'\mathbf{P})$ como matriz peso. Como ressaltam os autores, esta visão alternativa do estimador SMD

⁸Uma formulação equivalente para a equação é dada por Ai e Chen (2003, p. 1799): $\hat{m}_i(\mathbf{X}, \boldsymbol{\alpha}) = \sum_{j=1}^N \rho_i(\mathbf{Z}_j, \boldsymbol{\alpha}) \mathbf{p}^{k_n}(\mathbf{X}_j)' (\mathbf{P}'\mathbf{P})^{-1} \mathbf{p}^{k_n}(\mathbf{X})$.

dá um limite inferior para o grau da aproximação do *linear sieve estimator* k_n ,

$$d_\rho k_n \geq d_\theta + q_n,$$

onde d_θ é a dimensão do espaço dos parâmetros Θ e q_n é a dimensão apenas da aproximação do espaço dos parâmetros de dimensão infinita, \mathcal{H}_n .

O próximo passo na decisão é a escolha da sequencia que aproxima o espaço das funções \mathcal{H} . Neste trabalho usou-se o espaço gerado por polinômios ortogonais, mais precisamente pelos polinômios de Chebyshev. Para o problema aqui estudado, a conveniência desta base, ao invés de outras comumente utilizadas na literatura, como splines e redes neurais, se deve, primeiramente, aos polinômios apresentarem uma forma explícita e simples para as suas derivadas, que são como a função perda da autoridade monetária a ser estimada entra na função critério. Um segundo ponto é que, segundo as suposições do modelo econômico a ser estimado, a função perda do Banco Central é estritamente côncava para atender as condições de otimização deste, sendo assim não apresenta quinas, que seria a justificativa do uso de splines.

3.2 Aproximação da função perda

Uma base para um espaço de funções definidas em um conjunto bidimensional $\chi = \chi_1 \times \chi_2 \subset \mathbb{R}^2$ é um tensor das bases dos espaços das funções univariadas definidas em χ_1 e χ_2 . Seja B_1 uma base do espaço $\mathcal{L}_2(\chi_1)$ e B_2 uma base do espaço $\mathcal{L}_2(\chi_2)$. Então uma base para $\chi = \chi_1 \times \chi_2$ é o tensor (conjunto) $B_1 \otimes B_2 = \{f(\pi, x) = b_1(\pi)b_2(x) \mid b_1 \in B_1, b_2 \in B_2\}$. Porém, a quantidade de elementos em um tensor cresce exponencialmente com a dimensão, fazendo com que Judd (1998) argumente que é mais eficiente usar apenas os componentes do tensor $f(\pi, x) = b_1(\pi)b_2(x)$ que tenha grau menor ou igual a um k previamente escolhido.⁹ Neste trabalho, buscando uma maior eficiência computacional, também se usou esse método.

Como dito na subseção anterior, a base escolhida para aproximar a função perda foi a sequencia de polinômios de Chebyshev. Como é razoável que o par desvio da inflação em relação à sua meta e hiato do produto esteja contido em $[-1, 1]^2$, não precisou-se alterar o domínio dos polinômios. Seja U_n o polinômio de Chebyshev de grau n . A função perda da autoridade monetária é aproximada por

$$\begin{aligned} L(\tilde{\pi}_t, x_t, \tilde{i}_t) &= \tilde{L}(\tilde{\pi}_t, x_t) + \frac{\gamma}{2} (\tilde{i}_t)^2, \\ \tilde{L}(\tilde{\pi}_t, x_t) &= \sum_{i=1}^{K_\pi} \psi_{i,0} U_i(\tilde{\pi}_t) + \sum_{i=1}^{K_x} \psi_{0,i} U_i(x_t) + \sum_{i=1}^{k_\pi} \sum_{j=1}^{k_x} \psi_{i,j} U_i(\tilde{\pi}_t) U_j(x_t) + \psi_{0,0}. \end{aligned}$$

⁹Isto é chamado *complete polynomials*. Uma justificativa heurística dada por Judd (1998) é a aproximação de Taylor multivariada, na qual o termo de ordem k da expansão é composto pelos polinômios $x_1^{i_1} \dots x_n^{i_n}$ tal que $\sum_{l=1}^n i_l = k$.

Caso fosse utilizado o tensor completo ter-se-ia $K_\pi = K_x = k_\pi = k_x$. Por eficiência, aqui foi utilizado $K_\pi = K_x = 5$ e $k_\pi = k_x = 1$. Este grau da aproximação permite suficiente assimetria na função perda, além de ter um termo de interação entre a inflação e o hiato do produto.

Para que a função perda estimada atenda a hipótese (4), a otimização deve ser feita sob algumas restrições. Garantir que o mínimo da função se dê no ponto $(0, 0)$ permite diminuir dois parâmetros a serem estimados. Assim, resolvendo as equações que implicam que $(0, 0)$ seja o ponto de mínimo, $\partial \tilde{L}(0,0)/\partial \tilde{\pi}_t = 0$ e $\partial \tilde{L}(0,0)/\partial x_t = 0$, pode-se deixar $\psi_{1,0}$ e $\psi_{0,1}$ em função dos outros parâmetros,

$$\begin{aligned}\psi_{1,0} &= -(U_1(0))^{-1} \left(\sum_{i=2}^{K_\pi} \psi_{i,0} U'_1(0) + \sum_{i=1}^{k_\pi} \sum_{j=1}^{k_x} \psi_{i,j} U'_i(0) U_j(0) \right) \\ \psi_{0,1} &= -(U_1(0))^{-1} \left(\sum_{i=2}^{K_x} \psi_{0,i} U'_1(0) + \sum_{i=1}^{k_\pi} \sum_{j=1}^{k_x} \psi_{i,j} U_i(0) U'_j(0) \right),\end{aligned}$$

eliminando-os.

Já para garantir a supermodularidade da função perda, os parâmetros estimados devem satisfazer a equação (2.4) $\forall \tilde{\pi}, x \in (-1, 1)$ ¹⁰.

Como problemas de otimização são indiferentes à transformações afins¹¹ na função objetivo, por questões de identificabilidade deve-se normalizar a função perda a ser estimada. Neste trabalho normalizou-se a função fazendo com que $\psi_{2,0} = 1$, equivalente a dividi-la por $\|\psi_{2,0}\|$ ¹².

3.3 Momentos condicionais e instrumentos

O principal parâmetro de interesse neste trabalho é a função perda do Banco Central. Uma escolha natural para estimá-la e através da função de Euler (2.9). Porém, esta função não contempla toda a informação disponível no problema de otimização da autoridade monetária. Para incorporar essa informação restante, deve-se estimar conjuntamente as equações estruturais da economia, (2.1) e (2.2).

A equação de Euler já está no formato de momento condicional requerido pelo método. Já as outras duas equações precisam ser transformadas. Aplicando a esperança condicional ao conjunto de informação do período t em (2.1) chega-se a $E_t[\pi_{t+1} - \beta\pi_{t+2} - \kappa x_{t+1}] = E_t[u_{t+1}]$. Porém, ainda não se tem um momento condicional cuja esperança seja igual a zero. Para isso, é necessário subtrair de ambos os lados da equação a quantidade $E_t[u_{t+1}] = E_t[\varrho u_t] = E_t[\varrho(\pi_t - \beta\pi_{t+1} - \kappa x_t)]$, onde foi utilizada a hipótese (2) sobre a estrutura dos termos de erro. Assim, tem-se a restrição de momento condicional

¹⁰Na prática, como os valores da inflação e do hiato produto estão ao redor de zero e longe de -1 e 1 , exigiu-se a supermodularidade apenas em $(-0.9, 0.9)$. Isto permite que não se rejeite funções candidatas que não satisfaçam essa condição apenas nos extremos do seu suporte.

¹¹Transformação afim é um operador linear de forma $z \mapsto az + b$, $a > 0$.

¹²A partir de várias estimativas anteriores sabe-se que $\psi_{2,0} > 0$.

$$E_t [-\beta\pi_{t+2} + (1 + \rho\beta)\pi_{t+1} - \rho\pi_t - \kappa x_{t+1} + \rho\kappa x_t] = 0. \quad (3.8)$$

De forma análoga, pode-se chegar, a partir de (2.2) à terceira restrição de momento a ser estimada,

$$E_t [x_{t+2} + (1 + \omega x_{t+1}) - \omega x_t + \phi(i_{t+1} - \omega i_t) - \phi(\pi_{t+2} - \omega\pi_{t+1}) - \phi(1 + \omega)\bar{r}] = 0. \quad (3.9)$$

Dessa forma, o vetor com as restrições sobre os momentos condicionais que caracterizam a economia $E[\boldsymbol{\rho}(\mathbf{Z}, \boldsymbol{\theta}_0, \mathbf{h}_0(\bullet)) | \mathbf{X}] = \mathbf{0}$ é dado por as equações (2.9), (3.8) e (3.9).

No presente modelo, todas as variáveis utilizadas, $\{\pi_t, x_t, i_t, t = t_0, \dots, N\}$ são endógenas, isto é, todas são determinadas pelas equações que o descrevem.

Como descrito na seção anterior, o primeiro passo é estimar a esperança dos momentos condicionais, através de *linear sieve estimator*, como uma função das variáveis presentes no conjunto de informação. No instante t o conjunto de informação é dado por $I_t = \{\pi_\tau, x_\tau, i_\tau, \tau = t_0, \dots, t\}$. Porém, na prática não é necessário utilizar todas as variáveis presentes no conjunto de informação. Neste trabalho foram utilizadas apenas as variáveis do próprio instante t , além do período anterior, $t - 1$. Com estas variáveis foram montadas 14 componentes da base para a aproximação de $\mathbf{m}(\mathbf{X}, \boldsymbol{\alpha})$,

$$\mathbf{p} = (1, \pi_t, x_t, i_t, \pi_t^2, x_t^2, i_t^2, \pi_t x_t, \pi_t i_t, x_t i_t, \pi_t x_t i_t, \pi_{t-1}, x_{t-1}, i_{t-1}).$$

Uma vez que a função perda da autoridade monetária tenha sido estimada, pode-se testar se ela é ou não quadrática. Caso a função seja quadrática, a sua terceira derivada assume o valor zero em todos os pontos do seu domínio e, portanto, a esperança da terceira derivada é igual a zero.¹³ Assim, pode-se testar se a função perda não é quadrática testando se a esperança de sua terceira derivada é diferente de zero. Como a esperança é um funcional suave e a função perda é estimada por *sieve*, é possível utilizar o resultado encontrado por Chen e Shen (1998), que mostram que estes funcionais convergem para uma distribuição normal à taxa $N^{1/2}$.¹⁴ Dessa forma, a estatística de teste é $\mu_q = E(\partial^3 L / \partial q^3)$, $q = \{\pi, x\}$.

¹³Note que o oposto não é válido.

¹⁴O resultado é mostrado no teorema 2 de Chen e Shen (1998).

4 Resultados

Neste capítulo serão apresentados os dados e os resultados da estimação para a economia norte-americana e para a brasileira.

4.1 Estados Unidos

Para modelar as preferências do *Federal Reserve System* (Fed) foram utilizados dados trimestrais para o período compreendido entre o primeiro trimestre de 1960 e o último trimestre de 2006¹⁵, totalizando 188 observações. As variáveis utilizadas seguiram Surico (2007). Como medida de inflação usou-se a variação no logaritmo do deflator do *Personal Consumption Expenditure* (PCE), disponível no sistema FRED do Federal Reserve Bank of St. Louis. A taxa de juros utilizada foi o Effective Federal Funds Rate e a medida de produto foi o Industrial Production Index, ambos disponíveis no *site* do Fed. Todas as variáveis utilizadas foram calculadas como o seu valor acumulado nos últimos quatro trimestres. A meta de inflação foi assumida como constante e igual a 0,02. O hiato do produto e o desvio da taxa de juros com relação ao seu valor de *steady state* foram calculados utilizando o filtro Hodrick-Prescott.

Os histogramas dessas variáveis são mostrados na figura (1a). Nota-se que, considerando todo o período da amostra, na maior parte do tempo a inflação encontrou-se abaixo da sua meta. Já o hiato do produto teve tantos períodos positivos quanto negativos. As séries históricas dessas variáveis são apresentadas na figura (2)

Ao longo das quase cinco décadas compreendidas na amostra, é natural que tenha havido mudanças na condução da política econômica por parte do Fed. Clarida et al. (2000) encontraram evidências de que a função de reação do Banco Central mudou com a indicação de Paul Volcker como seu presidente em 1979. Segundo os autores, durante a era Volcker-Greenspan, a autoridade monetária americana passou a responder mais à expectativa de inflação, enquanto na era pré-Volcker a política tinha um caráter mais acomodativo. Com base nisso, a exemplo de Surico (2007), estimou-se aqui também a função perda do Fed para a subamostra correspondente a era Volcker-Greenspan, mais precisamente para o período compreendido entre o quarto trimestre de 1982 e o último de trimestre de 2006. O autor escolheu 1982 e não 1979 pois, segundo ele, entre 1979 e 1982 o instrumento de política do Banco Central dos Estados Unidos deixou de ser a taxa de juros de curto prazo, passando para os depósitos compulsórios.

¹⁵Não foram utilizados dados mais recentes para não contaminar a amostra com a Crise Financeira Internacional de 2008.

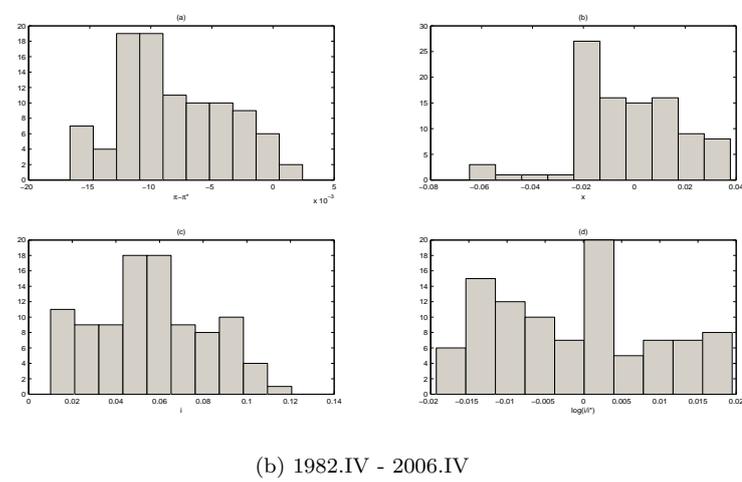
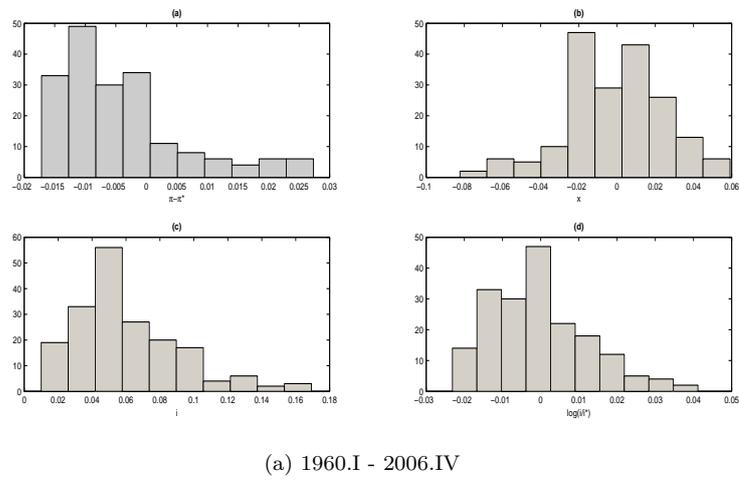


Figura 1: Histograma das variáveis utilizadas - EUA
 Fonte: elaboração própria.

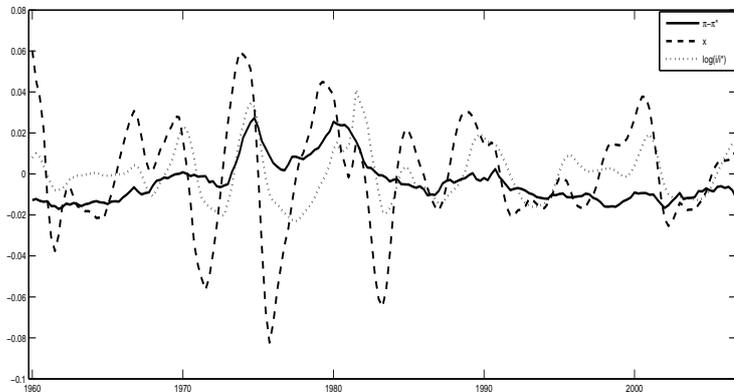
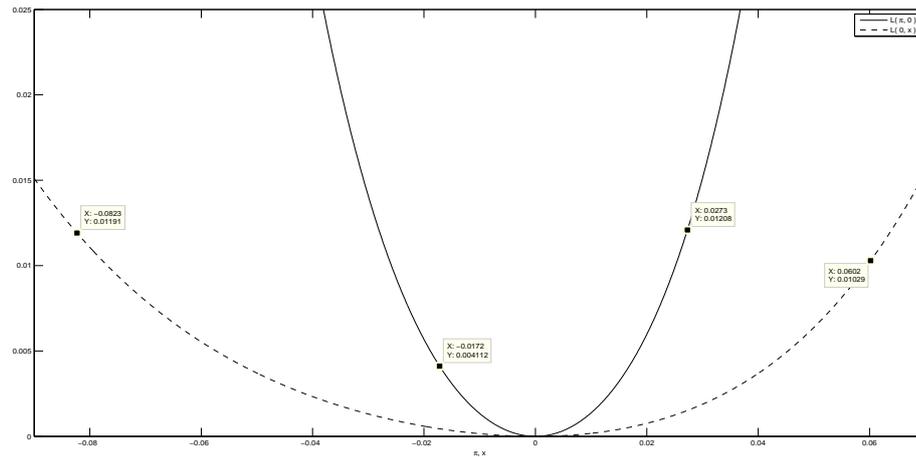


Figura 2: Série histórica das variáveis utilizadas - EUA - 1960.I-2006.IV
 Fonte: elaboração própria.

Tabela 1: Parâmetros estimados - EUA

Parâmetro	1960.I-2006.IV		1982.IV-2006.IV	
	Valor	SD	Valor	SD
$\psi_{1,0}$	0.4278	R	-0.0172	R
$\psi_{2,0}$	1.0000	N	1.0000	N
$\psi_{3,0}$	0.4670	0.4167	-0.0324	0.4089
$\psi_{4,0}$	0.4210	0.0030	0.4176	0.0029
$\psi_{5,0}$	0.1403	0.1262	-0.0167	0.1240
$\psi_{0,1}$	0.0240	R	-0.0118	R
$\psi_{0,2}$	0.0467	0.0332	0.0311	0.0556
$\psi_{0,3}$	0.0249	0.0167	-0.0130	0.0245
$\psi_{0,4}$	0.0195	0.0136	0.0109	0.0233
$\psi_{0,5}$	0.0068	0.0052	-0.0040	0.0075
$\psi_{1,1}$	0.0023	0.0061	0.0086	0.0072
γ	0.0001	0.0052	0.0028	0.0016
β	0.9228	0.0105	0.9158	0.0327
κ	0.0792	0.06074	0.1925	0.0794
δ	0.0158	0.04962	0.1070	0.0214
ϕ	0.0090	R	0.0258	R
\bar{r}	0.0468	0.00421	0.0410	0.0073
ϱ	0.8776	0.04720	0.8218	0.1991
ω	0.8133	0.07076	0.7870	0.1689

Nota: Valor é o valor estimado e SD o desvio padrão estimado. O desvio padrão refere-se a estimativa por processo de *bootstrap* com 500 reamostragens. Os parâmetros restritos (R) não são estimados diretamente, sendo uma função dada pelo modelo dos parâmetros estimados. N é o parâmetro normalizado na estimação.



(a) todo o domínio observado

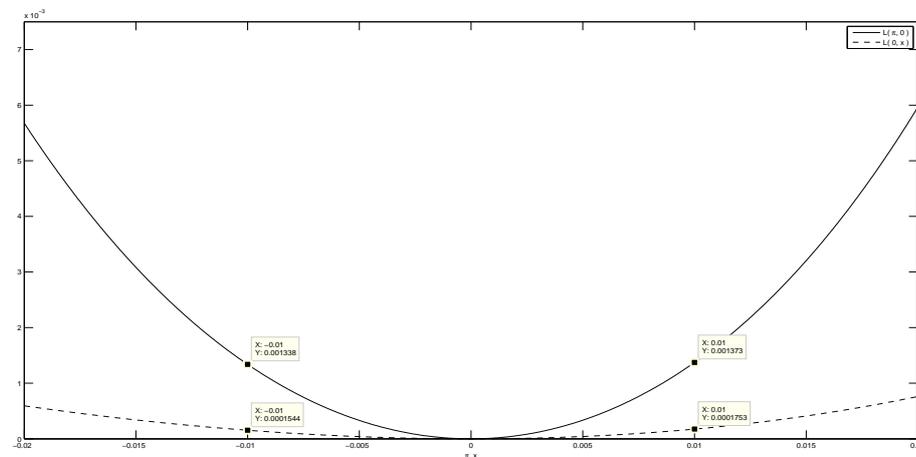
(b) domínio $(-0.02, 0.02)$

Figura 3: Função perda - EUA - 1960.I - 2006.IV

Nota: $L(\pi, 0)$ apresenta o valor da função perda do Banco Central para diversos valores da inflação, considerando o hiato do produto igual a zero. Análogo para $L(0, x)$. Os valores assinalados no gráfico superior representam o máximo e o mínimo da inflação e do hiato do produto. No gráfico inferior está se mostrando apenas o intervalo $(-0.02, 0.02)$.

Fonte: elaboração própria.

Tabela 2: Teste para função perda quadrática - EUA

(a) 1960.I - 2006.IV			
Estatística	Valor estimado	Desvio padrão	<i>t-value</i>
$\mu_\pi = E(\partial^3 L / \partial \pi^3)$	-0.201	0.322	-0.625
$\mu_x = E(\partial^3 L / \partial x^3)$	0.066	0.040	1.634
(b) 1982.IV - 2006.IV			
Estatística	Valor estimado	Desvio padrão	<i>t-value</i>
$\mu_\pi = E(\partial^3 L / \partial \pi^3)$	-0.051	0.168	-0.305
$\mu_x = E(\partial^3 L / \partial x^3)$	0.000	0.052	-0.008

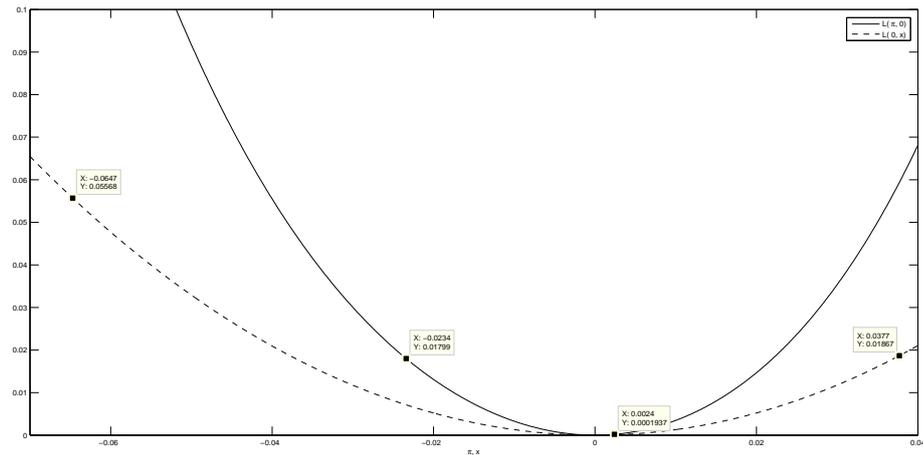
Nota: o *t-value* correspondente a 5% é 1.65.
O desvio padrão refere-se a estimativa por processo de *bootstrap* com 500 reamostragens.

A tabela (1) apresenta os resultados dos parâmetros estimados para a amostra inteira. Entre os parâmetros vale ressaltar o baixo valor do parâmetro que dá a sensibilidade do hiato do produto em relação à taxa de juros reais, $\phi = 0.009$.

A função perda estimada a partir dos parâmetros dados na tabela (1) é apresentada na figura (3). Como o parâmetro que trás a interação entre a inflação e o hiato do produto se mostrou baixo, $\psi_{1,1} = 0.0023$, a análise pode ser feita considerando as perdas provenientes de cada uma das variáveis em separado, sem perda de generalidade. Esse resultado suporta as abordagens correntes na literatura, onde a função perda da autoridade monetária é separável na inflação e no hiato do produto.

Segundo os resultados, o Banco Central norte-americano, no período 1960-2006, apresentou uma preocupação maior com a inflação do que com o hiato do produto, fato já encontrado em outros trabalhos. Deve-se ressaltar, no entanto, que, devido ao hiato do produto apresentar valores absolutos mais altos do que a inflação (figura (2a)), os seus valores extremos observados implicam uma perda à autoridade monetária de ordem semelhante aos valores extremos observados da inflação. Também se pode ver que as perdas associadas à inflação e ao hiato do produto são ambas levemente assimétricas à direita, com o Banco Central auferindo uma perda maior quando as variáveis estão acima das suas metas (a perda com um desvio da inflação de 1% foi 1,026 vezes maior do que a perda com um desvio de -1%; já a perda com um hiato do produto de 1% foi 1,135 vezes maior do que com um hiato de -1%). No entanto, para esse período não foi possível rejeitar, a um nível de 5% de significância, a hipótese que a verdadeira função perda seja quadrática em cada uma das variáveis, conforme o teste apresentado na tabela (2a).

Em relação à amostra completa, o Banco Central no período Volcker-Greenspan conseguiu manter a inflação um período ainda maior de tempo abaixo da sua meta, sendo que o desvio máximo da inflação registrado foi igual a 0.24% (figura (1b)). Os resultados para este período foram semelhantes ao da



(a) todo o domínio observado

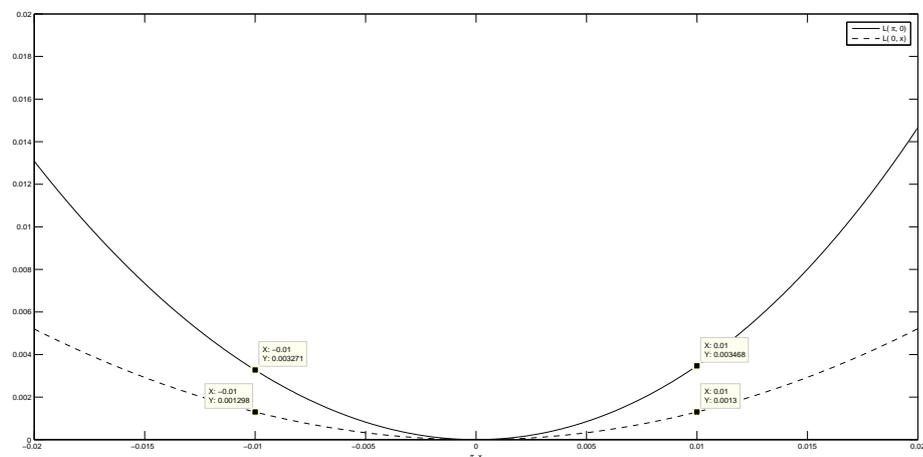
(b) domínio $(-0.02, 0.02)$

Figura 4: Função perda - EUA - 1982.IV - 2006.IV

Nota: $L(\pi, 0)$ apresenta o valor da função perda do Banco Central para diversos valores da inflação, considerando o hiato do produto igual a zero. Análogo para $L(0, x)$. Os valores assinalados no gráfico superior representam o máximo e o mínimo da inflação e do hiato do produto. No gráfico inferior está se mostrando apenas o intervalo $(-0.02, 0.02)$.

Fonte: elaboração própria.

amostra inteira, com a função perda estimada apresentada na figura (4). Assim como para o período completo, a função perda relativa à inflação se situa acima da relativa ao hiato do produto. Contudo, como no caso anterior, devido ao hiato do produto apresentar valores absolutos maiores do que a inflação, para um conjunto de observações a perda com aquele é maior do que com esta. Na era Volcker-Greenspan tanto a função perda relativa à inflação quanto a relativa ao hiato do produto apresentaram leve assimetria à direita. No entanto, não se pode rejeitar que nenhuma delas seja na verdade quadrática, como apresentado no teste da tabela (2b). Um motivo que dificulta a rejeição em relação à inflação é a quantidade de observações de desvios positivos da inflação ser bem menor do que a quantidade com desvios negativos, fazendo com que a variância da estimação para a porção positiva da reta seja grande.

Assim, tanto para a amostra completa quanto para a era Volcker-Greenspan, tanto a função perda relativa à inflação quanto a relativa ao hiato do produto se mostraram assimétricas a direita, isto é, com a autoridade monetária auferindo maiores perdas de desvios positivos das variáveis. Porém, não se pode rejeitar que a verdadeira função perda seja quadrática tanto para a inflação quanto para o hiato do produto, em nenhum dos períodos testados. A ausência de assimetrias nas preferências do Fed para esses períodos foi encontrada também por Surico (2007), modelando a assimetria restrita à equação *linear*.¹⁶

4.2 Brasil

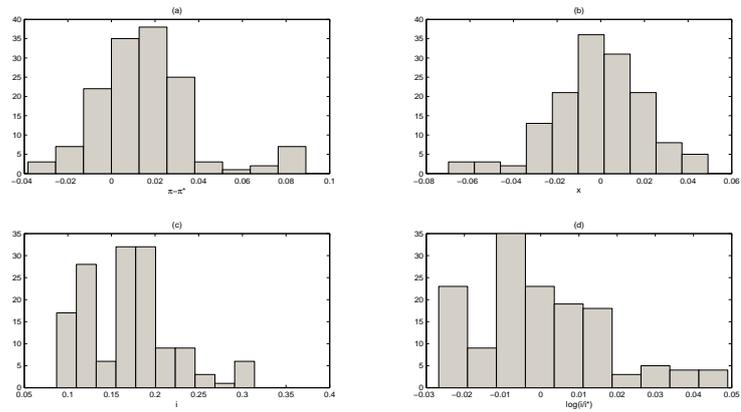
Como o sistema de metas de inflação brasileiro é mais recente, iniciado apenas em abril de 1999, usou-se dados mensais para estimar a função perda do Banco Central do Brasil (Bacen). A amostra tem 143 observações, indo de abril de 1999 até fevereiro de 2011. Para tentar separar da estimação o efeito da forte apreciação do dólar e do aumento da inflação, ocorridos entre 2002 e 2003, estimou-se também o modelo para o período compreendido entre janeiro de 2004 e fevereiro de 2011.

A medida de inflação utilizada foi o Índice de preço ao consumidor amplo (IPCA), índice este que baliza as ações do Bacen. No caso brasileiro, diferentemente do americano, a meta de inflação é explícita e muda ao longo do tempo. Como a meta é para a inflação anual, interpolou-se a meta assim como em Aragón e Portugal (2010)¹⁷. Para o produto foi utilizado o índice de produção industrial calculado pelo Instituto Brasileiro de Geografia e Estatística (IBGE), enquanto que para os juros foi utilizada a taxa *over-Selic*. Assim como na estimação anterior, todas as variáveis foram utilizadas com o seu valor acumulado ao longo dos doze meses anteriores e tanto o hiato do produto quanto o desvio dos juros em relação à sua meta foram calculados através do filtro Hodrick-Prescott.

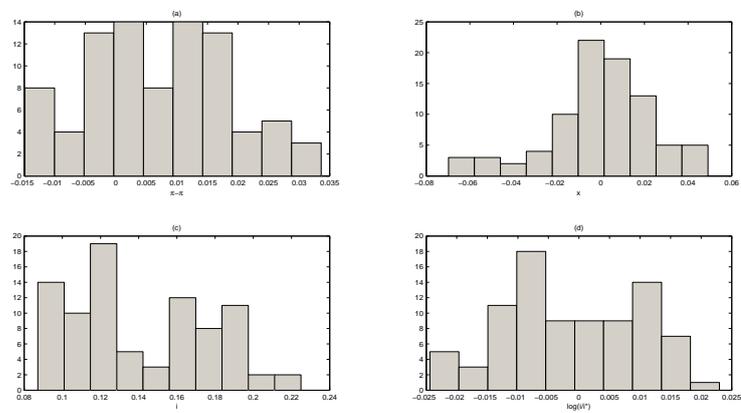
A figura (5a) apresenta o histograma das variáveis brasileiras. No Brasil, ao contrário dos EUA, a maioria das observações da inflação foi maior do que a meta, além da amplitude dos desvios da inflação ser maior do que a amplitude do hiato do produto. Ainda podem-se ver as altas taxas de juros apresentada pela economia brasileira, com a maior parte das observações estando acima de 10% a.a.. As trajetórias dessas variáveis estão plotadas na figura (6).

¹⁶O autor encontrou assimetria apenas relativa ao produto para o período pré-Volcker.

¹⁷Aragón e Portugal (2010) interpolam a meta de inflação de acordo com a equação $\pi_{m,t}^* = (12-m/12)\pi_{t-1}^* + (j/12)\pi_t^*$, onde $\pi_{m,t}^*$ é a meta de inflação interpolada para o mês m do ano t e π_t^* é a meta de inflação anual divulgada para o ano t . Para os anos de 2003 e 2004 foram utilizadas as metas ajustadas de 8,5% e 5,5%, respectivamente, estabelecidas pela Carta Aberta de janeiro de 2003.



(a) 1999.04 - 2011.02



(b) 2004.01 - 2011.02

Figura 5: Histograma das variáveis utilizadas - Brasil
 Fonte: elaboração própria.

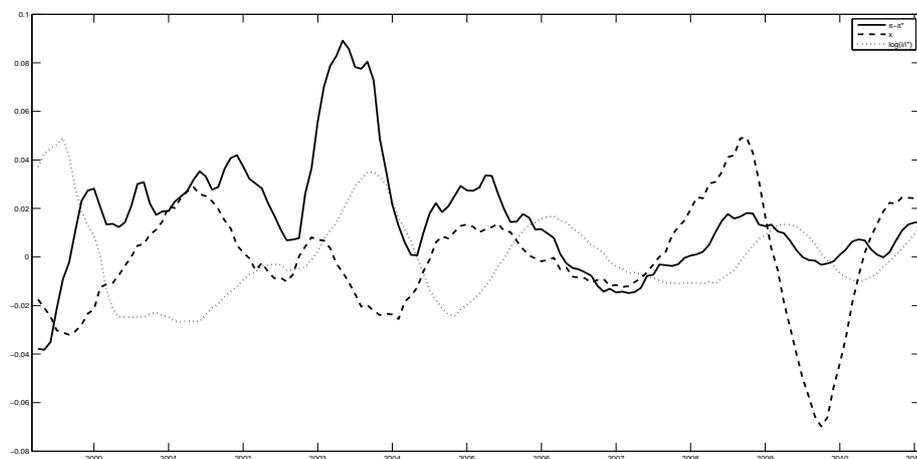
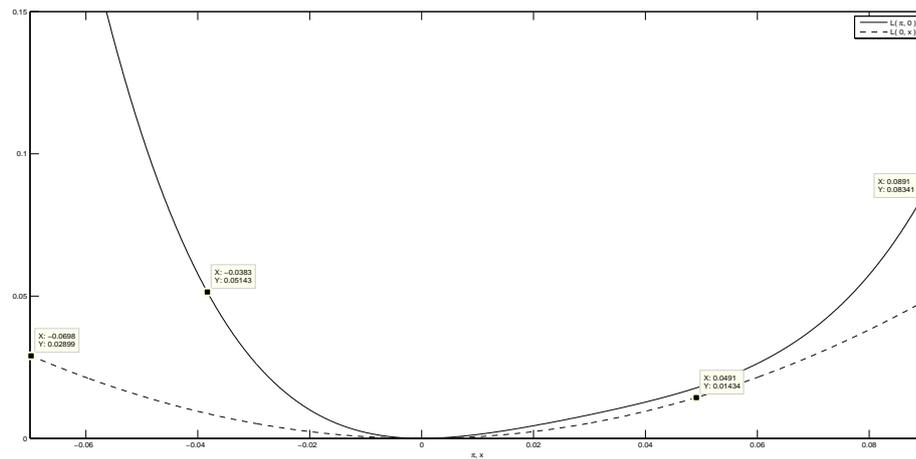


Figura 6: Série histórica das variáveis utilizadas - Brasil - 1999.04 - 2011.02
 Fonte: elaboração própria.

Tabela 3: Parâmetros estimados - Brasil

Parâmetro	1999.04-2011.02		2004.01-2011.02	
	Valor	SD	Valor	SD
$\psi_{1,0}$	-0.5475	R	-0.4064	R
$\psi_{2,0}$	1.0000	N	1.0000	N
$\psi_{3,0}$	-0.5529	0.2112	-0.4367	0.3154
$\psi_{4,0}$	0.4204	0.0019	0.4212	0.0008
$\psi_{5,0}$	-0.1434	0.0629	-0.1277	0.0958
$\psi_{0,1}$	0.0004	R	-0.0529	R
$\psi_{0,2}$	0.0035	0.8252	0.1280	0.1340
$\psi_{0,3}$	0.0004	0.4363	-0.0610	0.0713
$\psi_{0,4}$	0.0005	0.3490	0.0534	0.0566
$\psi_{0,5}$	0.0001	0.1297	-0.0200	0.0211
$\psi_{1,1}$	0.0044	0.0119	0.0032	0.0067
γ	0.0027	0.0045	0.0032	0.0009
β	0.9789	0.0509	0.9412	0.0230
κ	0.1494	0.1424	0.1282	0.1118
δ	0.0797	0.0365	0.1055	0.0130
ϕ	0.0335	R	0.0307	R
\bar{r}	0.0407	0.0091	0.0424	0.0122
ϱ	0.8622	0.1573	0.8756	0.1897
ω	0.9421	0.1300	0.9399	0.1794

Nota: Valor é o valor estimado e SD o desvio padrão estimado. O desvio padrão refere-se a estimativa por processo de *bootstrap* com 500 reamostragens. Os parâmetros restritos (R) não são estimados diretamente, sendo uma função dada pelo modelo dos parâmetros estimados. N é o parâmetro normalizado na estimação.



(a) todo o domínio observado

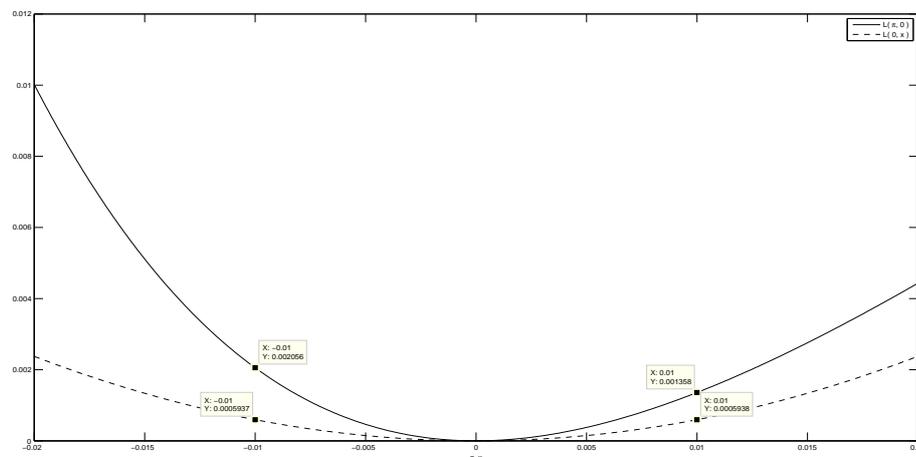
(b) domínio $(-0,02, 0,02)$

Figura 7: Função perda - Brasil - 1999.04 - 2011.02

Nota: $L(\pi, 0)$ apresenta o valor da função perda do Banco Central para diversos valores da inflação, considerando o hiato do produto igual a zero. Análogo para $L(0, x)$. Os valores assinalados no gráfico superior representam o máximo e o mínimo da inflação e do hiato do produto. No gráfico inferior está se mostrando apenas o intervalo $(-0,02, 0,02)$.

Fonte: elaboração própria.

Tabela 4: Teste para a função perda quadrática - Brasil

(a) 1999.04 - 2011.02			
Estadística	Valor estimado	Desvio padrão	<i>mathmor</i>
$\mu_\pi = E(\partial^3 L / \partial \pi^3)$	2.429	0.749	3.244*
$\mu_x = E(\partial^3 L / \partial x^3)$	0.000	0.949	0.000
(b) 2004.01 - 2011.02			
Estadística	Valor estimado	Desvio padrão	<i>t-value</i>
$\mu_\pi = E(\partial^3 L / \partial \pi^3)$	3.469	0.190	18.299*
$\mu_x = E(\partial^3 L / \partial x^3)$	0.122	0.128	0.956

Nota: o *t-value* correspondente a 5% é 1.65.

O desvio padrão refere-se a estimativa por processo de *bootstrap* com 500 reamostragens.

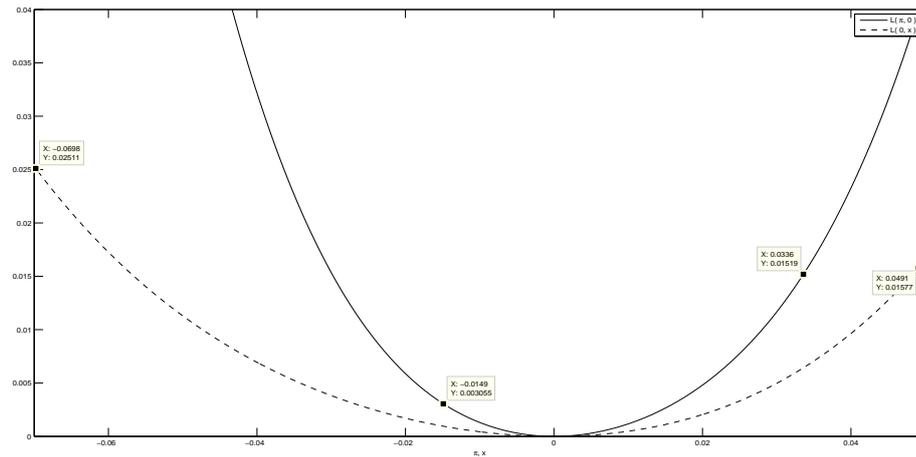
* significativo a 1%.

A tabela (3) apresenta o valor dos parâmetros estimados para a economia brasileira. Quanto à estrutura da economia pode-se ver que as inclinações da curva de Phillips e da curva IS (0.1494 e 0.0335, respectivamente) são maiores do que as da economia norte-americana.

A estimação da função perda da autoridade monetária durante o regime de metas é apresentada na figura (7). Novamente a perda relativa à inflação encontra-se acima da perda relativa ao hiato do produto, porém, no intervalo (0, 0.06) elas se encontram relativamente próximas. A perda associada à inflação se mostrou bastante assimétrica à esquerda, com desvios negativos da inflação em relação à sua meta auferindo uma maior desutilidade à autoridade monetária (a perda com um desvio da inflação de -1% foi 1.52 vezes maior do que a perda com um desvio de 1%). Esta leniência da autoridade monetária com desvios positivos da inflação pode resultar em dificuldades no controle da inflação. Já a perda com o hiato do produto estimada mostrou-se simétrica. O teste para a verdadeira forma da função perda permite rejeitar a forma quadrática da função perda relativa à inflação, confirmando a assimetria falada anteriormente, mas não permite rejeitá-la para a função perda relativa ao hiato do produto. O resultado do teste é apresentado na tabela (4a).

O subperíodo 2004-2011 apresentou estimativas para os parâmetros estruturais da economia semelhantes ao da amostra completa, como mostrado na tabela (3). Já a sua função perda estimada, mostrada na figura (8), apresenta um formato mais próximo do quadrático, contudo com ainda com uma assimetria à esquerda (um desvio de -1% da inflação representou uma perda à autoridade monetária 1,11 vezes maior do que a perda proveniente de um desvio de 1%). Enquanto isso a função perda em relação ao hiato do produto se mostrou levemente simétrica à direita. Testando esses resultados, pode-se, novamente, rejeitar que a verdadeira função perda da autoridade monetária brasileira em relação à inflação seja quadrática, o mesmo não ocorrendo com respeito ao hiato do produto. A tabela (4b) sumariza esses resultados.

Em resumo, para o caso brasileiro, rejeitou-se que a função perda relativa ao desvio da inflação seja



(a) todo o domínio observado

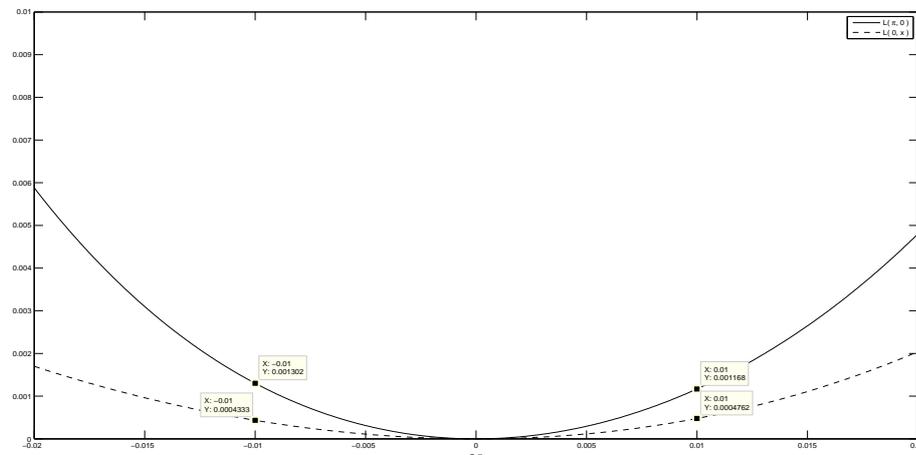
(b) domínio $(-0.02, 0.02)$

Figura 8: Função perda - Brasil - 2004.01 - 2011.02

Nota: $L(\pi, 0)$ apresenta o valor da função perda do Banco Central para diversos valores da inflação, considerando o hiato do produto igual a zero. Análogo para $L(0, x)$. Os valores assinalados no gráfico superior representam o máximo e o mínimo da inflação e do hiato do produto. No gráfico inferior está se mostrando apenas o intervalo $(-0.02, 0.02)$.

Fonte: elaboração própria.

quadrática, tanto para a amostra completa quanto para a subamostra 2004-2011. Em ambos os casos a função apresentou assimetria à esquerda. Essa assimetria negativa para a amostra completa já tinha sido encontrada por Aragón e Portugal (2010) , o mesmo não ocorrendo com a parte final da amostra¹⁸. Quanto a perda relativa ao hiato do produto, não se pode rejeitar que ela seja quadrática tanto para a amostra completa quanto para a subamostra, fato também encontrado por Aragón e Portugal (2010).

¹⁸Para a subamostra 2004-2007 Aragón e Portugal (2010) não rejeitaram a hipótese de que a perda seja simétrica. Porém, a subamostra utilizada no presente trabalho engloba mais quatro anos.

5 Conclusão

Neste trabalho aplicou-se a técnica não paramétrica de *sieve estimator* para a estimação das preferências das autoridades monetárias dos Estados Unidos e do Brasil. Dessa forma foi possível testar se a função perda dos Bancos Centrais desses países foi assimétrica durante o período analisado. Assim, pode-se testar se estas autoridades monetárias auferem uma maior perda de desvios positivos ou negativos da inflação e do hiato do produto.

A economia foi modelada com agentes *foward-looking* e a maximização da autoridade monetária feita sobre comprometimento, minimizando uma função perda genérica. A partir das condições de primeira ordem do problema, chegou-se a uma regra de política ótima na forma de uma restrição de momentos condicionais das variáveis da economia. Nessas restrições entram tanto os parâmetros de dimensão finita das equações estruturais da economia quanto a função perda do Banco Central a ser estimada. Dessa forma foi possível utilizar o método de *sieve* para a sua estimação, que consiste, de maneira informal, em estimar a função em um espaço menos complexo do que o qual ela está contida. O espaço aproximado utilizado aqui foi o gerado por polinômios de Chebyshev.

Não se encontrou evidências significativas de assimetria nas preferências do Fed tanto para o período completo da amostra (1960-2004) quanto para a era Volcker-Greenspan. No entanto, para ambos os períodos o Fed mostrou estar marginalmente mais preocupado com desvios positivos do que negativos, tanto da inflação quanto do hiato do produto.

Já o comportamento do Banco Central brasileiro mostrou-se assimétrico com relação à inflação tanto para a amostra completa (1999-2011) quanto para o subperíodo 2004-2011. E, em ambas as amostras, ele auferiu uma perda maior com a inflação abaixo da meta do que com a inflação acima da meta. Isto é o oposto do que se espera para um controle mais eficaz da inflação, o que pode levar a um viés inflacionário na economia brasileira. Essa forma das preferências do Banco Central brasileiro leva a uma maior velocidade da diminuição dos juros quando comparada aos momentos de elevação dos mesmos.

Esse resultado abre espaço para se pesquisar os motivos desse comportamento da autoridade monetária brasileira. Uma hipótese é a independência da autoridade monetária brasileira ser menor do que o esperado, fazendo com que a aversão do Governo por um desaquecimento da economia advindo de uma inflação inferior à meta fosse transferida, em parte, para o Banco Central sob a forma de uma rejeição à desvios negativos da inflação. Outra possibilidade é que o método tenha encontrado dificuldades em estimar a função perda da autoridade monetária em virtude das duas crises presentes na amostra (pré-eleição de 2002 e a crise financeira internacional do final da década de 2010). Isto porque em períodos de crise a autoridade monetária recorre a outros recursos além da taxa de juros para o combate a inflação, o que não é captado diretamente por a estrutura da economia adotada neste trabalho (onde o instrumento

do Banco Central é apenas a taxa de juros de curto prazo), enviesando a perda com uma inflação acima da meta para baixo.

Uma terceira possibilidade para a assimetria observada no caso brasileiro é o curto período da amostra e, particularmente, à maior observação de taxa de inflação acima da meta do que abaixo dela. A inflação esteve abaixo da sua meta em 23,1% das observações entre 1999-2011 e 31,4% entre 2004-2011. Esta diferença entre a proporção de desvios positivos e negativos da inflação pode ter levado a uma acurácia menor no intervalo negativo da função perda, dificultando comparações com o intervalo positivo. Contudo, vale lembrar que o teste apresentado no trabalho tem como hipótese nula a função perda ser quadrática, o que pode ser rejeitado mesmo que todas as observações fossem positivas.

Este trabalho suscita algumas questões que podem ser abordadas em outros trabalhos. Quanto à modelagem da economia, pode-se tentar incorporar novos elementos, como a taxa de câmbio, que busquem explicar a ação da autoridade monetária por meios outros que não a taxa de juros, como discutido para o caso brasileiro. Além disso, é necessário estudar as trajetórias das variáveis e o viés inflacionário gerados por esta abordagem não paramétrica. Quanto ao método do *sieve estimator*, seria interessante implementar uma simulação de Monte Carlo para testar sua performance em amostras finitas. Em um grau mais avançado, poder-se-ia tentar desenvolver um teste para a forma funcional da função perda que tenha como hipótese nula a função *linex*, utilizada em outros trabalhos empíricos.

Referências

- AI, C., E CHEN, X. Efficient estimation of models with conditional moment restrictions containing unknown functions. *Econometrica*, 71(6):1795–1843, 2003.
- ARAGÓN, E., E PORTUGAL, M. Nonlinearities in Central Bank of Brazil’s reaction function: the case of asymmetric preferences. *Estudos Econômicos*, 40(2):373–399, 2010.
- CHEN, X. Large sample sieve estimation of semi-nonparametric models. *Handbook of Econometrics*, 6: 5549–5632, 2007.
- CHEN, X., E LUDVIGSON, S. Land of addicts? An empirical investigation of habit-based asset pricing models. *Journal of Applied Econometrics*, 24(7):1057–1093, 2009.
- CHEN, X., E SHEN, X. Sieve extremum estimates for weakly dependent data. *Econometrica*, 66(2): 289–314, 1998.
- CHIARELLA, C., FLASCHEL, P., GONG, G., E SEMMLER, W. Nonlinear phillips curves, complex dynamics and monetary policy in a keynesian macro model. *Chaos, Solitons & Fractals*, 18(3):613–634, 2003.
- CLARIDA, R., GALÍ, J., E GERTLER, M. The science of monetary policy: A new keynesian perspective. *Journal of Economic Literature*, 37:1661–1707, 1999.
- CLARIDA, R., GALI, J., E GERTLER, M. Monetary policy rules and macroeconomic stability: Evidence and some theory. *Quarterly Journal of Economics*, 115(1):147–180, 2000.
- CUKIERMAN, A., E MUSCATELLI, V. Do central banks have precautionary demands for expansions and for price stability?: Theory and evidence. *CESiifo Working Paper Series*, 764, 2002.
- DOLADO, J., MARIA-DOLORES, R., E NAVEIRA, M. Are monetary-policy reaction functions asymmetric?: The role of nonlinearity in the phillips curve. *European Economic Review*, 49(2):485–503, 2005.
- FAN, J., E YAO, Q. *Nonlinear time series: nonparametric and parametric methods*. Springer, New York, 2005.
- GRENANDER, U. *Abstract inference*. John Wiley & Sons, 1981.
- HAMILTON, J. A parametric approach to flexible nonlinear inference. *Econometrica*, 69(3):537–573, 2001.
- ISNARD, C. *Introdução á Medida e Integração*. IMPA, Rio de Janeiro, 2007.
- JUDD, K. *Numerical methods in economics*. The MIT press, Cambridge, 1998.

- KIM, D. H., OSBORN, D. R., E SENSIER, M. Nonlinearity in the Fed's monetary policy rule. *Journal of Applied Econometrics*, 20(5):621–639, 2005.
- LJUNGQVIST, L., E SARGENT, T. *Recursive macroeconomic theory*. The MIT press, Cambridge, 2004.
- NEWKEY, W. Convergence rates and asymptotic normality for series estimators. *Journal of Econometrics*, 79(1):147–168, 1997.
- NEWKEY, W., E POWELL, J. Instrumental variable estimation of nonparametric models. *Econometrica*, 71(5):1565–1578, 2003.
- NOBAY, A. R., E PEEL, D. A. Optimal discretionary monetary policy in a model of asymmetric central bank preferences. *The Economic Journal*, 113(489):657–665, 2003.
- NOBAY, A., E PEEL, D. Optimal monetary policy with a nonlinear phillips curve. *Economics Letters*, 67(2):159–164, 2000.
- ORPHANIDES, A., E WIELAND, V. Inflation zone targeting. *European Economic Review*, 44(7):1351–1387, 2000.
- SCHALING, E. The nonlinear phillips curve and inflation forecast targeting: Symmetric versus asymmetric monetary policy rules. *Journal of Money, Credit, and Banking*, 36(3):361–386, 2004.
- SURICO, P. The Fed's monetary policy rule and U.S. inflation: The case of asymmetric preferences. *Journal of Economic Dynamics and Control*, 31(1):305 – 324, 2007.
- SVENSSON, L., E WOODFORD, M. Implementing optimal policy through inflation-forecast targeting. *The inflation-targeting debate*, 1:19 – 92, 2005.
- TAYLOR, J. Discretion versus policy rules in practice. In *Carnegie-Rochester conference series on public policy*, volume 39, pages 195–214. Elsevier, 1993.
- WOODFORD, M. *Interest and prices: Foundations of a theory of monetary policy*. Princeton Univ Press, Princeton, 2003.

A Apêndice - Consistência do estimador

Esta seção se baseia na abordagem de Chen e Ludvigson (2009) para estabelecer a consistência do estimador $\hat{\alpha}$ quando os dados não são independentes. Os autores o fazem aplicando o lema A1 de Newey e Powell (2003), enunciado a seguir.

Lema 2. *Suponha que (i) $Q(\alpha)$ tenha um único mínimo em \mathcal{A} , localizado em α_0 ; (ii) $\hat{Q}(\alpha)$ e $Q(\alpha)$ sejam contínuas, \mathcal{A} seja compacto e $\max_{\alpha \in \mathcal{A}} |\hat{Q}(\alpha) - Q(\alpha)| \xrightarrow{P} 0$; (iii) $\hat{\mathcal{A}}$ são subconjuntos compactos de \mathcal{A} tal que para qualquer $\alpha \in \mathcal{A}$ existe $\tilde{\alpha} \in \hat{\mathcal{A}}$ tal que $\tilde{\alpha} \xrightarrow{P} \alpha$. Então $\hat{\alpha} = \arg \min_{\alpha \in \hat{\mathcal{A}}} \hat{Q}(\alpha) \xrightarrow{P} \alpha_0$.*

Demonstração. Ver Newey e Powell (2003, p. 1576). □

A hipótese (i) é ligada à identificabilidade do modelo. Quanto à continuidade das funções critério amostral e populacional requerida em (ii), ela pode ser vista a partir das equações (3.2) e (3.3) (lembrando que as bases com as quais a função perda é aproximada são contínuas). Já a compacidade dos parâmetros finito-dimensionais é razoável pois não tem sentido econômico algum deles divergir para ∞ ou $-\infty$ e a compacidade da função perda do Banco Central é assumida como hipótese (4). A hipótese (iii) diz que o espaço aproximado (*sieve space*) utilizado é denso no espaço onde está contida a verdadeira função a ser estimada. O teorema de Stone-Weierstrass mostra que o espaço de todos os polinômios é denso em \mathcal{L}^2 . Seguindo Chen e Ludvigson (2009), mostrar a convergência uniforme solicitada em (iii) é equivalente a mostrar que valem as equações

$$\sup_{\mathbf{X}_t, \alpha} |\hat{m}(\mathbf{X}_t, \alpha) - m(\mathbf{X}_t, \alpha)| = o_p(1) \quad (\text{A.1})$$

e

$$\sup_{\mathcal{A} \times \Theta} \left| N^{-1} \sum_{t=1}^N m(\mathbf{X}_t, \alpha)' m(\mathbf{X}_t, \alpha) - E \left(m(\mathbf{X}_t, \alpha)' m(\mathbf{X}_t, \alpha) \right) \right| = o_p(1). \quad (\text{A.2})$$

O resultado (A.1) é obtido através do lema A1 de Ai e Chen (2003) modificado para permitir dados β -mixing estacionários. Já o resultado (A.2) pode ser obtido através de qualquer lei dos grandes números para dados β -mixing estacionários. Com as condições atendidas, tem-se $\|\alpha_N - \alpha_0\|_s = o_p(1)$. Chen e Ludvigson (2009, apêndice técnico, p. 7)