

Universidade Federal do Espírito Santo
Centro de Ciências Jurídicas e Econômicas
Mestrado em Economia

Eduardo Reis Araújo

**INFLUÊNCIA DAS TAXAS DE JUROS INTERNACIONAIS SOBRE VOLUME DE
OPERAÇÕES DE CRÉDITO EXTERNO NO BRASIL: APLICAÇÃO DO MÉTODO
DE REGRESSÃO NÃO-PARAMÉTRICA COM ESTIMADORES DE NÚCLEO**

Vitória - ES

2009

Eduardo Reis Araújo

**INFLUÊNCIA DAS TAXAS DE JUROS INTERNACIONAIS SOBRE VOLUME DE
OPERAÇÕES DE CRÉDITO EXTERNO NO BRASIL: APLICAÇÃO DO MÉTODO
DE REGRESSÃO NÃO-PARAMÉTRICA COM ESTIMADORES DE NÚCLEO**

Dissertação de mestrado apresentada ao Curso de Mestrado em Economia, para obtenção do grau de mestre em teoria econômica, sob a orientação do Prof. Valdério A. Reisen.

Vitória - ES

2009

Dados Internacionais de Catalogação-na-publicação (CIP)
(Biblioteca Central da Universidade Federal do Espírito Santo, ES, Brasil)

A663i Araujo, Eduardo Reis, 1981-
Influência das taxas de juros internacionais sobre volume de operações de crédito externo no Brasil : aplicação do método de regressão não-paramétrica com estimadores de núcleo / Eduardo Reis Araujo. – 2009.
50 f. : il.

Orientador: Valdério Anselmo Reisen.
Dissertação (mestrado) – Universidade Federal do Espírito Santo, Centro de Ciências Jurídicas e Econômicas.

1. Taxas de juros internacionais. 2. Crédito. 3. Econometria. 4. Macroeconomia. I. Reisen, Valdério Anselmo. II. Universidade Federal do Espírito Santo. Centro de Ciências Jurídicas e Econômicas. III. Título.


CDU: 330

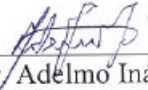
“Influência das taxas de juros internacionais sobre o volume de operações de crédito externo no Brasil: aplicação da regressão não-paramétrica com estimador de núcleo”

Eduardo Reis Araujo

Dissertação submetida ao Programa de Pós-Graduação em Economia da Universidade Federal do Espírito Santo como requisito parcial para a obtenção do título de Mestre em Economia.

Aprovada em 05 de fevereiro de 2009 por:


Prof. Dr. Valdério Anselmo Reisen – Orientador -UFES


Prof. Dr. Adélmo Inácio Bertoldi - UFES


Profa. Dra. Clélia Maria de Castro Tolo – USP

À Lauriana, à memória de meu pai e irmão.

AGRADECIMENTOS

Agradeço a Deus por me dar vida, saúde e disposição; ao Prof. Valdério A. Reisen pela sagacidade, interesse e envolvimento em todas as etapas desta dissertação; à Prof^a. Clélia M. C. Tolo e Prof. Adelmo Bertolde por terem aceitado o convite para composição da Banca; ao doutorando Fabio Fajardo pelas análises críticas e sugestões na etapa final do trabalho; aos Professores do Programa de Mestrado em Economia por todos ensinamentos, em especial aos Professores Alain Herscovici, Robson Grassi, Sônia Dalcomuni e Rogério Arthman pela contribuição prestada nos preparativos; aos colegas do Núcleo de Modelagem Estatística da UFES; aos colegas da Cesan, em especial, Wagner, Zenaide, Sérgio, Nilo, Marcelo e Celso que me deram todo suporte necessário; aos colegas Adriano Carmo, Alan do Pará, Alex Hoffman, Alexandre Bongestab, Breno Barreto, Celso Bissoli, Claudeci Neto, Hélio Gonçalves, Diego Corrêa, Fabiano Venturim, Guilherme Lucas, Igor Trazzi, Silvia Varejão e Ricardo Paixão pela parceria no estudo.

Nonparametric regression is a simple and useful tool for obtaining insight into the structure of data.

Chu, C. K. e Marron, J. S. (1991, p. 404).

RESUMO

A influência dos juros sobre o volume de crédito é um tema relevante na área econômica. O assunto já despertou interesse de diversos autores, que fizeram uso de modelos estatísticos paramétricos para discutir o assunto (Topel, 1988; Vivacqua, 2007; entre outros). Esta dissertação apresenta a análise da influência das taxas de juros internacionais sobre o volume de operações de crédito externo no Brasil por meio de modelos de regressão não-paramétrica com estimadores de núcleo. A motivação para uso dos métodos não-paramétricos está associada ao fato de que esses métodos permitem ajustar curvas de regressão num espaço vetorial mais amplo que o método regressão clássico (Härdle *et al.*, 2004). Por utilizar regressão entre séries observadas ao longo do tempo, o trabalho também verificou a dependência de longo prazo das variáveis no contexto de séries temporais. O teste ADF (Dickey e Fuller, 1979) e os estimadores de parâmetro de memória longa do modelo ARFIMA (Geweke e Porter-Hudak (1983); Reisen, (1994)) foram aplicados analisar a estacionariedade dos resíduos das regressões. Os resultados das análises evidenciaram que as variáveis do estudo são co-integradas de ordem fracionária, isto é, possuem dependência de longo prazo. A trajetória da curva de regressão não-paramétrica mostrou empiricamente que, no intervalo de Taxas de Juros correspondente a 1% e 2%, o aumento dessa Taxa não implica necessariamente em redução no volume de operação de crédito externo.

Palavras-chave: regressão não-paramétrica, estimador de núcleo e *bandwidth*.

ABSTRACT

The influence of interest on the amount of credit is an important topic in economic. In fact, it has attracted interest of several authors who have made use of parametric statistical models (Topel, 1988; Vivacqua, 2007, among others). This dissertation presented to the analysis of the influence of international interest rates on the volume of foreign credit operations in Brazil by means of regression models with non-parametric estimators of core. The motivation for the use of non-parametric methods is linked to the fact that these methods can adjust the regression curves in a vector space wider than the classical regression method (Härdle *et al*, 2004). By using regression between series observed over time, this work also found a long-term dependence of the variables in the context of time series. The ADF test (Dickey and Fuller, 1979) and the estimators of long memory parameter of the ARFIMA model (Reisen, 1994) were applied to analyze the stationarity of waste from regressions. The results of this analysis showed that the variables are co-integrated fractional order, ie, have long-term dependency. The trajectory of the curve of non-parametric regression showed empirically that, in the interest rate corresponding to 1% and 2%, the rate of increase does not necessarily imply a reduction in volume of external credit operation.

Keywords: nonparametric regression, kernel functions, bandwidth.

LISTA DE FIGURAS

Figura 1. Curvas das funções de núcleo para $u \in [-1,1]$: (1) Epanechnikov, (2) Tricubica, (3) Triangular, (4) Normal ou Gaussiana, (5) Uniforme ou Retangular, (6) Quartica.	21
Figura 2. Aplicação da Regressão com estimadores de núcleo para diferentes valores <i>bandwidth</i> : (a) $h=0,05$; (b) $h=0,015$; e (c) $h=0,45$	24
Figura 3. <i>Trade-off</i> entre viés e variância para diferentes valores h : (a) $h=0,05$; (b) $h=0,015$; e (c) $h=0,45$	25
Figura 4. Procedimentos para estimação de modelos de regressão paramétrica.	27
Figura 5. Procedimentos para ajuste de curvas de regressão não-paramétrica com estimadores de núcleo.	27
Figura 6. Séries temporais das taxas de juros internacionais e do volume de operações de crédito externo.	34
Figura 7. Gráfico de dispersão juros contra crédito.....	35
Figura 8. FAC das variáveis em estudo: (a) FAC Crédito; (b) FAC Juros.	36
Figura 9. FAC dos resíduos de regressão paramétrica: (a) Modelo 1; (b) Modelo 2.	38
Figura 10. FAC dos resíduos da regressão não-paramétrica: (a) Modelo 3; (b) Modelo 4.....	39
Figura 11. Ajustes de curvas de regressão método paramétrico versus não-paramétrico	42
Figura 12. Intervalo de Confiança de 95% para Regressão <i>Kernel</i>	43
Figura 13. Gráfico de dispersão dos Resíduos contra tempo: (a) Modelo 1; (b) Modelo 2; (c) Modelo 3; e (d) Modelo 4.	44
Figura 14. Gráfico de dispersão dos Resíduos contra variável Taxa de Juros: (a) Modelo 1; (b) Modelo 2; (c) Modelo 3; e (d) Modelo 4.....	45

LISTA DE TABELAS

Tabela 1. Tipos de funções Kernel.	20
Tabela 2. Teste de Dickey-Fuller Aumentado aplicado às Variáveis.....	36
Tabela 3. Ordem de integração das variáveis obtida a partir do estimador GPH com <i>bandwidth n^α</i>	37
Tabela 4. Tipos de modelos utilizados na aplicação.....	37
Tabela 5. Teste de Dickey-Fuller Aumentado aplicado aos resíduos.....	39
Tabela 6. Ordem de integração dos resíduos obtida a partir do estimador GPH com <i>bandwidth n^α</i>	40

SUMÁRIO

1	INTRODUÇÃO.....	13
2	METODOLOGIA.....	17
2.1	MODELO DE REGRESSÃO	17
2.2	ESTIMADOR DE NADARAYA WATSON	18
2.3	TIPOS DE FUNÇÕES KERNEL	20
2.4	SELEÇÃO DO BANDWIDTH.....	22
2.5	SÍNTESE DOS PROCEDIMENTOS DE AJUSTE DE MODELOS.....	26
2.6	REGRESSÃO ESPÚRIA E CO-INTEGRAÇÃO	28
3	ANÁLISES E RESULTADOS.....	34
3.1	ANÁLISE DESCRITIVA DAS VARIÁVEIS	34
3.2	ANÁLISES DE RELAÇÕES DE LONGO PRAZO	35
3.3	COMPARAÇÃO DOS AJUSTES DE REGRESSÃO	41
4	CONCLUSÕES E RECOMENDAÇÕES.....	46
5	REFERÊNCIAS.....	48

1 INTRODUÇÃO

As operações de crédito externo possuem um papel estratégico no desenvolvimento econômico de um país. As fontes externas de financiamento permitem ampliar as possibilidades de investimento das empresas quando o custo de captação em bancos nacionais é superior ao de bancos estrangeiros.

Greumad, Vasconcellos e Toneto (2002) afirmam que quando ocorre uma crise de liquidez no mercado financeiro internacional, os bancos podem adotar uma maior parcimônia na concessão de empréstimos. Com isso, as taxas de juros podem se elevar e o volume das operações de crédito se reduzir.

Estudos realizados por Topel (1988) e Vivacqua (2007) ressaltaram a importância de se analisar a influência dos juros sobre volume de operações de crédito, através de metodologias estatísticas de regressão paramétrica. Topel (1988) utilizou modelos de regressão dinâmica para analisar a sensibilidade do investimento habitacional dos Estados Unidos quando ocorrem mudanças das taxas de juros, destacando a existência de uma associação negativa entre taxas de juros e investimento. Vivacqua (2007) estudou o impacto da política monetária sobre o investimento privado no Brasil através de um experimento empírico com base em dados em painel, e demonstrou que o investimento privado é afetado negativamente pelos juros¹.

Em 2005, a Federação das Indústrias do Estado de São Paulo publicou um estudo para avaliar a influência da variação das taxas de juros de longo prazo sobre o volume de ope-

¹ Esses trabalhos seguem a linha da Teoria Pós-Keynesiana, que defendem a não-neutralidade da moeda como ponto focal da crítica à teoria ortodoxa. Para esclarecimentos ver Carvalho (1992).

rações de crédito direcionado pelo Banco Nacional de Desenvolvimento Econômico e Social. O estudo utilizou a técnica de regressão paramétrica por Mínimos Quadrados Ordinários (MQO) e demonstrou que a redução de 1% nos juros produz acréscimo de R\$ 1,1 bilhão em operações de crédito.

Os trabalhos supracitados, por serem desenvolvidos com base nas técnicas de modelagem paramétrica, requerem pressupostos teóricos para inferência estatística. Isto é, eles pressupõem uma pré-especificação do relacionamento funcional das variáveis. Essa característica é considerada um aspecto de rigidez do modelo (Härdle, 1990). Por essa razão, abordagens alternativas à modelagem paramétrica foram propostas na literatura, o que inclui os modelos semiparamétricos e não-paramétricos.

Aplicações dos modelos semiparamétricos podem ser encontradas nos trabalhos de Härdle e Stoker (1989), Ghysels e Serena (1996), Ferreira e Tusell (1996), Hastie e Tibshirani (1991), Yatchew (2001), entre outros.

No contexto não-paramétrico, Härdle e Kirman (1995) ajustaram uma curva de regressão não-paramétrica com estimadores de núcleo (ou *Kernel*) para estudar a influência de quantidades comercializadas sobre os preços no mercado de piscicultura Marseille (França). A justificativa para aplicação dessa técnica foi motivada em três aspectos: em primeiro lugar, porque os autores não possuíam pressupostos teóricos que justificassem o uso de uma determinada especificação prévia do modelo; segundo, porque a utilização de modelos paramétricos certamente impossibilitaria a identificação de características dos dados que somente poderiam ser observadas com uma técnica mais flexível; e por fim, porque os estimadores de núcleo apresentam simplicidade matemática e maior compreensão intuitiva. Outros estudos envolvendo uso de técnica de regressão não-paramétrica podem ser encontrados em Azzalini,

Bowman e Härdle (1989); Ghysels e Serena (1996); Ferreira e Tusell (1996); Banks, Blundell e Lewbel (1997); Ziegelmann (2000); Shimakura *et al* (2001); Simonassi e Cândido Junior (2005) Racine (2008); entre outros.

Seguindo a linha do trabalho de Härdle e Kirman (1995), essa dissertação propõe a aplicação do método de regressão não-paramétrica com uso do estimador de núcleo para analisar a relação existente entre as taxas de juros internacionais e o volume de operações de crédito externo no Brasil, no período de dezembro de 2002 a agosto de 2008.

Neste trabalho, compara-se a metodologia não-paramétrica com uso do estimador de núcleo e o método de regressão paramétrica utilizando estimadores MQO. A trajetória da curva de regressão não-paramétrica demonstrou que, no intervalo de Taxas de Juros correspondente a 1% e 2%, a elevação dos juros não implica necessariamente numa redução do volume de crédito externo.

Por se tratar de um estudo envolvendo regressão com variáveis observadas ao longo do tempo, a dissertação tangenciou temas relativos a séries temporais e co-integração. O teste ADF (Dickey e Fuller, 1979) e os estimadores de parâmetro de memória longa do modelo ARFIMA (Geweke e Porter-Hudak (1983); Reisen, 1994) foram aplicados para analisar a estacionariedade dos resíduos das regressões. Essa análise de não-estacionariedade aplicada aos resíduos da regressão não-paramétrica com estimadores de núcleo permitiu identificar uma relação de longo prazo entre as variáveis estudadas.

Esta dissertação está estruturada da seguinte forma. No Capítulo 2 é voltado à discussão dos principais aspectos metodológicos associados à regressão não-paramétrica com estimador núcleo, bem como alguns conceitos básicos da análise de séries temporais. O Capítulo 3 corresponde a aplicação dos métodos de regressão não-paramétricos com estimadores

de núcleo ao conjunto de dados de Taxas de Juros Internacionais expresso pelas Taxa *Libor* ao ano e o Volume de Operações de Crédito realizados por pessoas empresas no Brasil, no período de dezembro de 2002 a agosto de 2008. Finalmente, conclusões e indicações para pesquisas futuras são apresentadas no Capítulo 4.

2 METODOLOGIA

A seguir são apresentadas algumas definições consideradas no desenvolvimento deste trabalho. A maior parte das informações foram obtidas em Härdle (1990), Wand e Jones (1995), Eubank (1999), Yachew (2003), Morettin e Toloi (2004) e Racine (2008).

2.1 MODELO DE REGRESSÃO

Definição 1: Seja $(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$ uma amostra aleatória bivariada, independente, e x_i um valor observado de X_i . Considere o modelo

$$Y_i = m(x_i) + \varepsilon_i, \quad i = 1, 2, \dots, n, \quad (2.1)$$

onde $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ são variáveis aleatórias não-correlacionadas que representam os termos de erro, tal que $E(\varepsilon_i | X_i = x_i) = 0$ e $Var(\varepsilon_i | X_i = x_i) = \sigma^2(x_i)$, $m(g)$ é chamada de função (ou curva) de regressão e é dada por $m(x_i) = E(Y_i | X_i = x_i)$.

Note que o modelo (2.1) remove as restrições paramétricas para $m(x_i)$ e permite que estruturas de várias formas sejam consideradas. Em particular, se $m(x_i) = \beta_0 + \beta_1 X_i$ tem-se um modelo de regressão linear simples, $m(x_i) = \beta_0 + \beta_1 \ln X_i$ representa o modelo de Lin-Log (Gujarati, 2004).

Definição 2: Um estimador $\hat{m}(g)$ de $m(g)$ é conhecido como suavizador linear, se para cada $x \in \mathcal{R}$, existe uma seqüência de pesos $w_1(x), \dots, w_n(x)$, com $\sum_{i=1}^n w_i(x) = 1$ e

$$\hat{m}(x) = \sum_{i=1}^n w_i(x) Y_i. \quad (2.2)$$

Para estimar a função $m(x)$ várias propostas podem ser encontradas na literatura. No contexto de regressão não-paramétrica com estimadores de núcleo, o estimador mais conhecido foi desenvolvido por Nadaraya (1964) e Watson (1964). Outras estimadores alternativos foram propostos por Gasser (1979) e Müller (1979), Cheng e Ling (1981), Priestley e Chao (1962), entre outros.

De acordo com Geenens (2008) a popularidade do estimador de Nadaraya-Watson deve-se à simplicidade de derivação, interpretação e programação. A seguir serão apresentadas a forma e a propriedade de consistência estatística desse estimador².

2.2 ESTIMADOR DE NADARAYA WATSON

Definição 3: (Estimador de Nadaraya-Watson) Seja $h \in \mathbb{R}_+$, $h > 0$. O estimador de núcleo de Nadaraya-Watson (NW) é definido como a equação (2.2) com

² O estimador de “Nadaraya-Watson também é conhecido como estimador “constante local”. A explicação pra esta denominação é descrita, geralmente, quando se introduz o assunto de regressão polinomial local (Fan (1992)).

$$w_i(x) = \frac{K\left(\frac{x-x_i}{h}\right)}{\sum_{j=1}^n K\left(\frac{x-x_j}{h}\right)}, \quad (2.3)$$

onde h é chamado de *bandwidth*, $K(g)$ é uma função par, contínua, limitada, e satisfaz as seguintes propriedades:

$$(i) \quad \int |K(u)| du < \infty,$$

$$(ii) \quad \lim_{|u| \rightarrow \infty} |uK(u)| = 0,$$

$$(iii) \quad \int K(u) du = 1.$$

A Proposição 1 apresenta a consistência do estimador NW, provada para amostras independentes por Roseblatt (1969) e Robinson (1983) para amostras dependentes.

Proposição 1. *Suponha que a função $K(g)$ satisfaz as condições (i) e (ii) acima,*

$EY_i^2 < \infty$, para todo $i = 1, 2, \dots, n$, então, $NW \xrightarrow{p} m(x)$, quando $n \rightarrow \infty$.

Observação: \xrightarrow{p} significa convergência em probabilidade.

Na equação (2.3), $K(g)$ é uma função conhecida como núcleo (ou *Kernel*). Em regressão não-paramétrica, com estimadores de núcleo, diferentes tipos dessas funções podem ser utilizados. Na seção seguinte serão apresentados alguns tipos.

2.3 TIPOS DE FUNÇÕES KERNEL

O tipo mais simples de função núcleo é conhecido como uniforme (ou retangular) que assume valor $K(u) = 1$ para o intervalo $u \in [-1, 1]$. Na Tabela 1 mostra as funções mais populares. As ilustrações gráficas podem auxiliar entender como os pesos são atribuídos, conforme a Figura 1.

Tabela 1. Tipos de funções Kernel.

#	Nome da função Kernel	$K(u)$	Domínio
1	<i>Epanechnikov</i>	$(3/4)(1 - u^2)$	$u \in [-1, 1]$
2	Tricubica	$(70/81)(1 - u ^3)^3$	$u \in [-1, 1]$
3	Triangular	$(2)(1 - 2 u)$	$u \in [-1, 1]$
4	Normal ou Gaussiana	$(1/\sqrt{2\pi})\exp(u^2/2)$	$u \in (-\infty, +\infty)$
5	Uniforme ou retangular	1	$u \in [-1, 1]$
6	Quartica	$(15/16)(1 - u^2)^2$	$u \in [-1, 1]$

Com exceção da função de núcleo uniforme, as funções de núcleo atribuem pesos menores às observações que se afastam do centro da janela selecionada (ver Figura 1). Assim, a forma como os pesos são atribuídos dependem do tipo de função escolhidas.

De acordo com Pagan e Ullah (1999) e Yatchew (2003), a escolha do tipo de função núcleo produz pouca interferência prática no ajuste da curva de regressão. O critério mais relevante é a seleção adequada do *bandwidth*, que será objeto de discussão na seção seguinte.

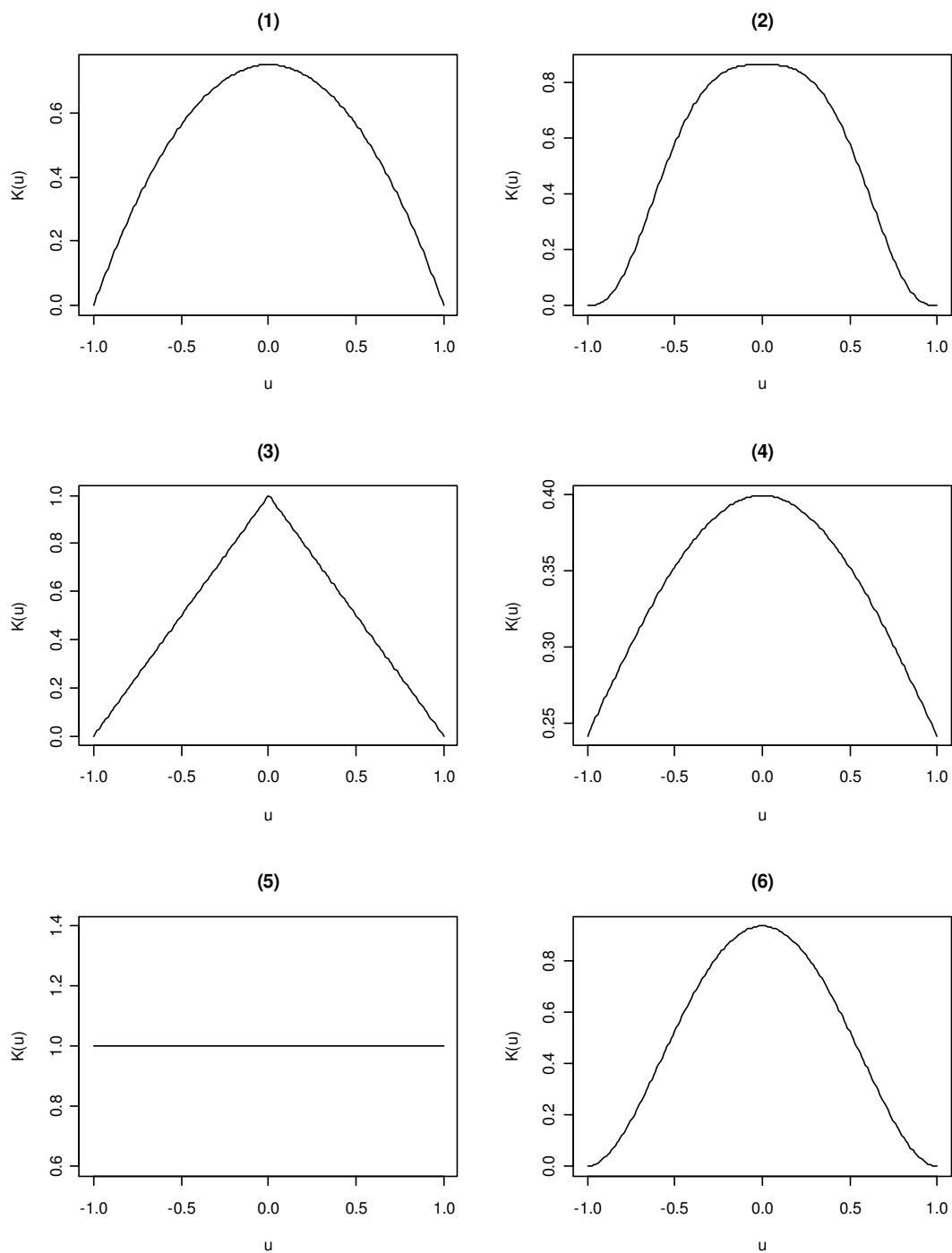


Figura 1. Curvas das funções de núcleo para $u \in [-1, 1]$: (1) Epanechnikov, (2) Tricubica, (3) Triangular, (4) Normal ou Gaussiana, (5) Uniforme ou Retangular, (6) Quartica.

2.4 SELEÇÃO DO BANDWIDTH

Esta seção é motivada pelas seguintes questões: qual é a importância da escolha do *bandwidth* no contexto da regressão não-paramétrica com funções de núcleo? O que acontece com a curva de regressão quando esses valores são alterados? Que técnica pode ser empregada para seleção adequada do valor do *bandwidth*?

De acordo com Yatchew (1998) e Wand (1994), a seleção do *bandwidth* é o critério de maior relevância no ajuste de curvas não-paramétricas, pois determina o grau de suavidade da curva de regressão.

Para exemplificar como o ajuste de regressão não-paramétrica com estimadores de núcleo é influenciado pela escolha do *bandwidth*, foi realizada a simulação da função (2.1) com $m(x_i) = x_i \cos(4\pi x_i)$, $\{\varepsilon_i\}$ independente e identicamente distribuído (i.i.d) com distribuição normal de média 0 e variância 0,09, $x_i \in [0,1]$ e $i = 1, 2, \dots, 100$ para diferentes valores h : (a) $h=0,05$; (b) 0,15 ; (c) 0,45. Todos os cálculos foram obtidos no *Software* estatístico R versão 2.7.1, subrotinas *ksmooth* e *locpoly*, disponível no site [HTTP://www.r-project.org](http://www.r-project.org).

Na Figura 2, a curva pontilhada refere-se à regressão ajustada e a curva sólida à verdadeira distribuição. Note que se o valor do *bandwidth* é pequeno (por exemplo, quando $h=0,05$) a curva de regressão ajustada tende a reproduzir os valores observados. Porém, se o valor do *bandwidth* aumentar (por exemplo, quando $h=0,15$ e $h=0,45$) a curva de regressão fica mais suave ou alisada.

Do ponto de vista teórico Härdle (1990) mostra que, quando $h \rightarrow \infty$, e considerando o estimador NW definido em (2.3), tem-se $K\left(\frac{x-X_i}{h}\right) \rightarrow K(0)$ e então

$$\hat{m}(x) \rightarrow \frac{\sum_{i=1}^n K(0)Y_i}{\sum_{i=1}^n K(0)} = n^{-1} \sum_{i=1}^n Y_i.$$

Segundo Gasser (1984), se por um lado a redução do valor do *bandwidth* aproxima melhor o ajuste da curva de regressão aos dados, por outro a elevação do valor do *bandwidth* ocasiona decréscimos na variância do suavizador linear ao estimar a função $m(x)$. Esse dilema é conhecido como *trade-off* entre viés e variância (Härdle (1990) e Yatchew (1998)), ilustrado na Figura 3. Note que na medida em que o valor do *bandwidth* aumenta, o desvio-padrão $\sigma = [Var(\hat{m}(x))]^{1/2}$ do estimador se reduz e, como conseqüência, o intervalo de confiança fica menor.

Portanto, a escolha de um valor ótimo do *bandwidth* é um ponto crucial na obtenção de uma boa estimativa da curva e regressão. Existem vários métodos estatísticos para obtenção de h , dentre os quais destacam-se Validação Cruzada (CV), Validação Cruzada Generalizada (GVC), o método *plug-in* originalmente proposto por Woodroffe (1970) e de suas respectivas variações propostas por Chiu (1991), entre outros. A Validação Cruzada, definida abaixo, é utilizada neste trabalho, por conta da popularidade e simplicidade de aplicação do método.

Definição 4: (*Critério de Validação Cruzada ou Cross Validation*). Seja

$$CV(h) = n^{-1} \sum_{i=1}^n \{Y_i - \hat{m}_{-i}(x_i)\}^2 w(x_i), \text{ onde } \hat{m}_{-i}(x_i) \text{ corresponde ao estimador definido em}$$

(2.2) com a retirada da i -ésima observação. O valor do *bandwidth* ótimo é obtido a partir da

expressão $\hat{h} = \arg \min_h CV(h)$.

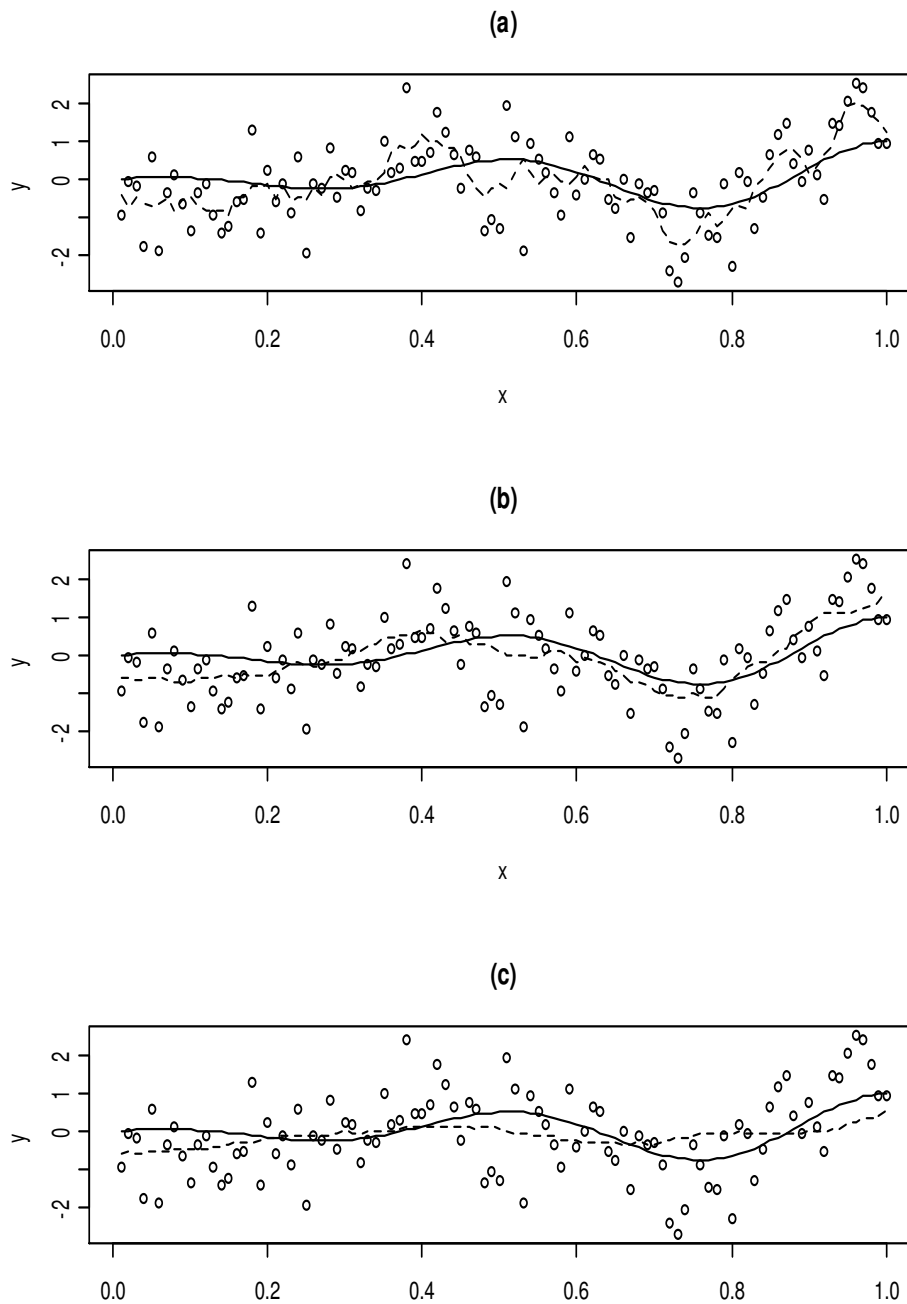


Figura 2. Aplicação da Regressão com estimadores de núcleo para diferentes valores *bandwidth*: (a) $h=0,05$; (b) $h=0,015$; e (c) $h=0,45$

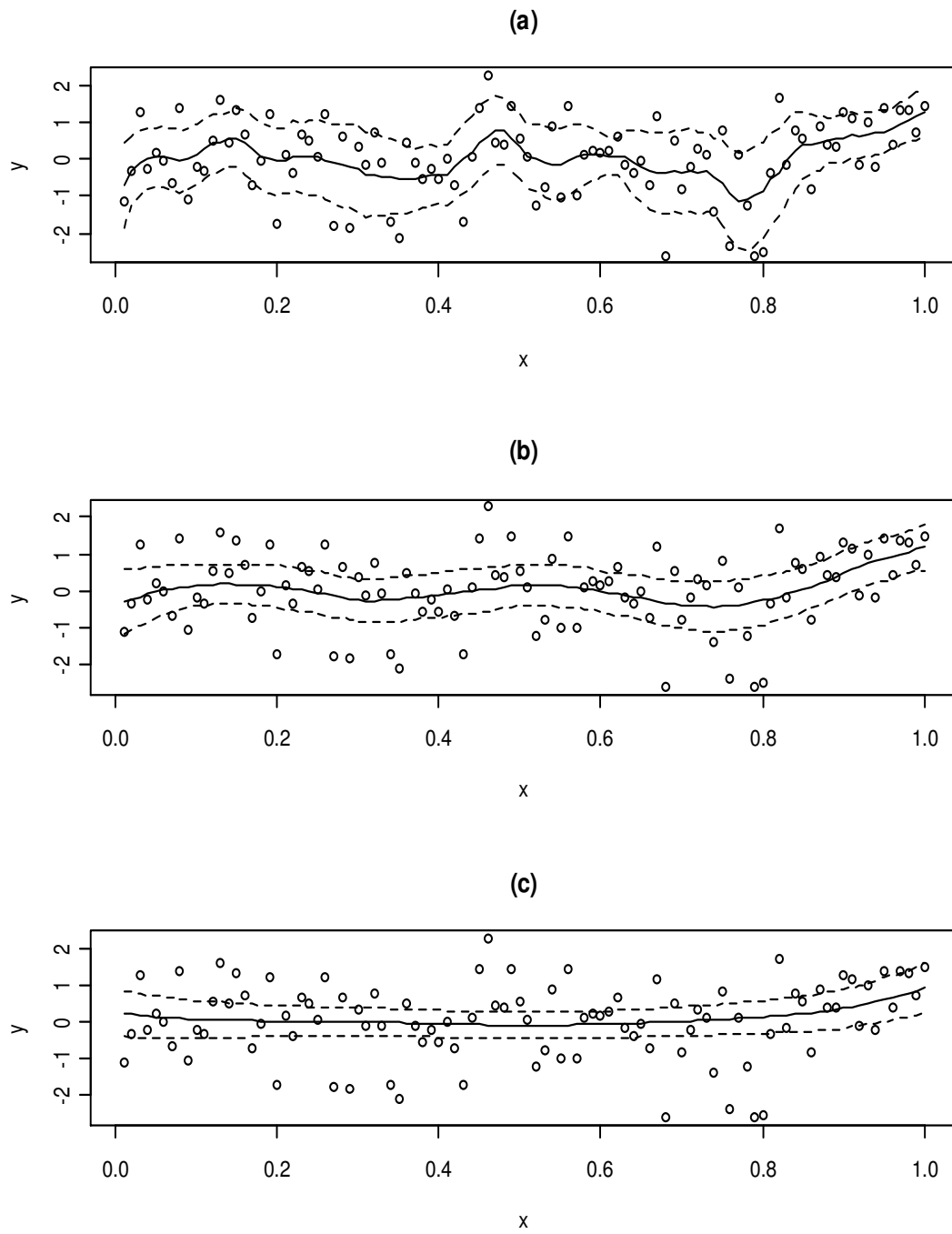


Figura 3. Trade-off entre viés e variância para diferentes valores h : (a) $h=0,05$; (b) $h=0,015$; e (c) $h=0,45$

2.5 SÍNTESE DOS PROCEDIMENTOS DE AJUSTE DE MODELOS

As seções anteriores se dedicaram na apresentação das principais definições e conceitos relativos à regressão não-paramétrica com estimadores de núcleo. Agora serão discutidos os procedimentos ou etapas a serem seguidos para utilização desses modelos.

De acordo Yatchew (1998), a pouca popularização das técnicas de regressão não-paramétricas devem-se basicamente a três razões: 1) maior complexidade teórica em relação à abordagem paramétrica; 2) ao fato de que, para validar as propriedades assintóticas dos estimadores de regressão não-paramétricos, são necessários conjuntos de dados significativamente maiores que no caso da regressão tradicional; 3) por fim, em razão da pequena popularidade de softwares ou pacotes estatísticos para ajuste dos modelos.

De fato, as técnicas de regressão não-paramétricas são menos difundidas na área da econometria aplicada. Entretanto, aos poucos essa metodologia está sendo incorporada aos manuais de econometria, e novos pacotes estatísticos são disponibilizados³.

As etapas de estimação de modelos não-paramétricos envolvem procedimentos diferentes daqueles utilizados no contexto de regressão paramétrica (ver Figuras 4 e 5). O esquema da Figura 4 foi desenvolvido com base em informações provenientes de um livro-texto clássico da área de econometria (Gujarati, 2006).

³ Neste trabalho, o ajuste de curvas de regressão não-paramétrica com uso de estimadores de núcleo foram realizado com base em pacotes disponíveis nos *softwares* estatísticos livres: *R-Program* (www.r-project.org) e *XploRe* (<http://fedc.wiwi.hu-berlin.de/xplore.php>).

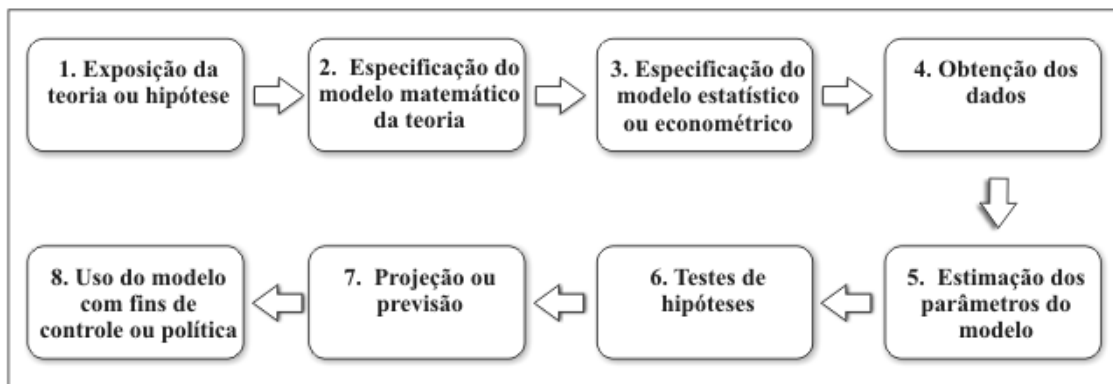


Figura 4. Procedimentos para estimação de modelos de regressão paramétrica.

A Figura 5 apresenta as etapas para ajuste de curvas pelo método de regressão não-paramétrico com estimadores de núcleo. Em síntese, os procedimentos desse esquema são os mesmos utilizados na Seção 3.3 deste trabalho.

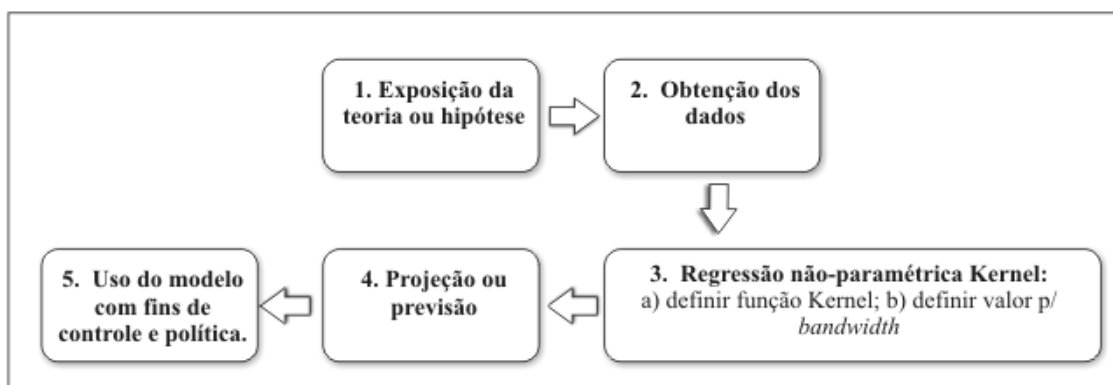


Figura 5. Procedimentos para ajuste de curvas de regressão não-paramétrica com estimadores de núcleo.

Na comparação entre a Figura 4 e Figura 5 destacam-se algumas diferenças nos procedimentos de ajuste de curva paramétrica e não-paramétrica⁴. De maneira geral, as diferenças estão associadas ao fato de que na abordagem não-paramétrica não é necessário especi-

⁴ O que se quer destacar com os Esquemas 1 e 2 é apenas uma noção geral dos procedimentos adotados para ajuste de regressão paramétrico e não-paramétrico. Por uma questão de simplificação, os esquemas ocultam algumas etapas como, por exemplo, a análise de resíduos.

ficar o modelo previamente, tão pouco estimar parâmetros e testar hipóteses⁵. A questão mais importante, como discutido na Seção 2.4, refere-se à seleção adequada do *bandwidth* ou parâmetro de suavização.

O estudo de regressão envolvendo variáveis observadas ao longo do tempo pressupõe a utilização de alguns conceitos relativos a séries temporais e co-integração, tendo em vista os riscos de se obter regressão espúria entre as variáveis (Yule, 1926). Diante desse contexto, a seguir serão apresentados conceitos básicos para diagnóstico de relações de longo prazo entre as variáveis.

2.6 REGRESSÃO ESPÚRIA E CO-INTEGRAÇÃO

A regressão espúria (ou regressão sem sentido) é um problema reportado com frequência na literatura econométrica. O assunto está relacionado à tendência de se aceitar a existência de uma relação entre duas séries, quando a relação não existe. Tecnicamente, esse problema pode ser verificado quando as equações de regressão de duas séries temporais apresentam um grau de ajuste R^2 *aparentemente* alto, mas com valor de estatística de Durbin-Watson pequeno. O problema é usualmente encontrado quando as series consideradas são não-estacionárias (ver Granger e Newbold (1974)).

Um processo estocástico é definido como uma família indexada de variáveis aleatórias $Z(\omega, t)$, onde ω pertence a um espaço amostral e t pertence a um conjunto indexado T . Para um t fixo, $Z(\omega, t)$ é uma variável aleatória. . A população em séries temporais consis-

⁵ De fato, enquanto na abordagem paramétrica os modelos são expressos na forma de equações algébricas (por exemplo, $\hat{y} = \hat{\alpha} + \hat{\beta}_x$), os métodos “totalmente” não-paramétricos utilizam ferramenta gráfica como mecanismo para analisar a curva de regressão ajustada.

te em todas as possibilidades de realizações de $Z(\omega, t)$ e é definida como “ensemble”. Para um dado ω , $Z(\omega, t)$ é função de t , e denominada “função amostral” ou “realização” ou “série temporal”. Portanto, série temporal é uma realização ou função amostral de certo processo estocástico (para detalhes ver Morettin e Toloi (2004) e Wei, 2006). A seguir será apresentado o conceito de processos estocástico. Será assumido que T é um conjunto de números inteiros, onde $\{Z_t\}$ é uma variável aleatória do processo estocástico $\{Z(\omega, t) : t = 0, \pm 1, \pm 2, \dots\}$.

Definição 5. (*Estacionariedade*⁶) Um processo estocástico $\{Z_t\}$ é chamado fracamente estacionário de ordem 2 se, para qualquer $t \in \mathbb{Z}$, então $E(Z_t) = \mu$, $\gamma_0 = E(Z_t - \mu)^2$ são constantes que não dependem t e a covariância $\gamma(t_1, t_2) = Cov\{Z_{t_1}, Z_{t_2}\}$ é uma função $|t_1 - t_2|$.

Definição 6. A função de autocorrelação (FAC) de um processo estacionário $\{Z_t\}$ é definida como $\rho_k = \frac{E(Z_t - \mu)(Z_{t+k} - \mu)}{\sqrt{E(Z_t - \mu)^2} \sqrt{E(Z_{t+k} - \mu)^2}} = \frac{\gamma_k}{\gamma_0}$, γ_0 corresponde à variância de Z_t , onde $k \in \mathbb{Z}$.

Para um processo estacionário (Definição 5), a FAC apresenta as seguintes propriedades:

- (i) $\rho_0 = 1$,
- (ii) $|\rho_k| \leq 1$, para todo k ,
- (iii) $\rho_k = \rho_{-k}$, para todo k , isto é, a FAC é uma função par e, portanto, simétrica, em torno da origem.

⁶ Neste trabalho, será utilizada a expressão *estacionário* para evitar o uso constante da expressão *estacionário de ordem 2*.

Definição 7. A função de autocorrelação parcial (FACP) de um processo estacionário de $\{Z_t\}$ defini-se como a correlação entre as variáveis Z_t e Z_{t+k} depois que a dependência linear das variáveis $Z_{t+1}, Z_{t+2}, \dots, Z_{t+k-1}$ tenha sido removida.

A seguir será apresentada a definição de um dos modelos mais populares na área de séries temporais.

Definição 8. (Processo ARFIMA) Seja $\{Z_t\}$ um processo estocástico com a seguinte representação

$$\Phi(B)(1-B)^d Z_t = \Theta(B)\varepsilon_t, \quad (2.4)$$

onde $d \in \mathbb{R}$, $(1-B)^d$ é o operador de diferenciação fracionária dado por

$$(1-B)^d = \sum_{j=0}^{\infty} \binom{d}{j} (-B)^j = 1 - dB - \frac{d}{2}(1-d)B^2 - \dots,$$

B é o operador de retardo, isto é, $B^m Z_t = Z_{t-m}$. O processo $\{\varepsilon_t\}$ é ruído branco com $E(\varepsilon_t) = 0$ e $\text{var}(\varepsilon_t) = \sigma^2$, para $t \in \mathbb{Z}$. Os polinômios $\Phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$ e $\Theta(B) = 1 - \theta_1 B - \dots - \theta_q B^q$ são os polinômios autorregressivo e de médias móveis, respectivamente. O processo $\{Z_t\}$ com representação (2.4) é chamado de *processo autorregressivo fracionário e de médias móveis*, ARFIMA (p,d,q). Para detalhes ver Reisen (1994) e Morettin e Tolo (2004).

Hosking (1981) mostrou que se $d < 0,5$ e todas as raízes do polinômio $\Phi(z) = 0$ ficam fora do círculo unitário, então $\{Z_t\}$ é estacionário. O processo ARFIMA (p,d,q) tem sido bastante explorado na literatura. Diferentes aplicações podem ser encontradas em todas as áreas de conhecimento, devido à capacidade de modelar séries com diferentes tipos de

memória: memória curta ($d = 0$), memória intermediária ($d < 0$) e memória longa ($d > 0$) (ver Reisen (1994)).

Vários métodos de estimação do parâmetro d foram propostos na literatura e podem ser divididos, basicamente, em dois grupos: métodos paramétricos e métodos semi-paramétricos. Dentre os métodos paramétricos podem-se citar os trabalhos de Fox e Taquq (1986), Sowell (1992), entre outros. Dentre os métodos semi-paramétricos, o mais popular é conhecido como GPH, proposto por Geweke e Porter-Hudak (1983). Outras variações do método foram propostas por Reisen (1994), Robinson (1994), entre outros. Neste trabalho foram considerados os métodos GPH e Reisen (1994), obtidos com funções do *software* estatístico R versão 2.7.1, subrotina *fracdiff*.

A não-estacionariedade é uma característica freqüentemente observada no estudo de séries temporais econômicas. Um procedimento comum adotado antes de realizar a análise desse tipo de séries é a diferenciação. A diferenciação torna as séries estacionárias e, a partir de então, é possível aplicar as metodologias econométricas para estudar a dinâmica das séries. O conceito de integração, introduzido por Engle e Granger (1987), está relacionado com a não-estacionariedade das séries. Por essa razão, será apresentado abaixo.

Definição 9. (Integração) Seja $\{Z_t\}$ um processo representado por (2.4). $\{Z_t\}$ é chamado integrado de ordem d , denotado por $Z_t \sim I(d)$, se são requeridas d diferenças para torná-lo estacionário.

Quando $d = 1$, diz-se que o processo é integrado de ordem 1, ou o processo apresenta uma raiz unitária. Outras definições de integração podem ser encontradas em Abadir e Taylor (1999).

O conceito de co-integração é amplamente conhecido na área de econometria. A teoria introduzida por Granger (1981) e Engle e Granger (1987) tem sido um dos ícones da pesquisa na área. Esses autores receberam o prêmio Nobel de Economia em função da relevante contribuição científica no campo de modelagem de séries financeiras. Abaixo é apresentada a definição formal de co-integração.

Definição 10. (Co-integração) Seja $\mathbf{Z}_t = (Z_{1t}, Z_{2t}, \dots, Z_{kt})$ um vetor de séries temporais. As componentes de \mathbf{Z}_t são ditas co-integradas de ordem d , b , denotado por $\mathbf{Z}_t \sim CI(d, b)$ se:

- (i) Todas as componentes de \mathbf{Z}_t são $I(d)$;
- (ii) existe um vetor β tal que $\beta' \mathbf{Z}_t \sim I(d-b), b > 0$. O vetor β é chamado de *vetor de co-integração*.

A idéia básica de co-integração é estudar as relações de longo prazo entre duas ou mais variáveis. Vários estudos empíricos têm investigado a possibilidade de co-integração entre os elementos de \mathbf{Z}_t . Os testes de hipóteses do tipo $H_0: \beta' \mathbf{Z}_t \sim I(1)$ contra $H_1: \beta' \mathbf{Z}_t \sim I(0)$ são conhecidos como “*testes de raiz unitária*” (Engle e Grenger (1987); Engle e Yoo (1987); Philips e Ouliaris (1990)).

Entre os testes de raízes unitárias mais utilizados nos estudos de co-integração estão: o testes de Dickey-Fuller (1979), Phillips-Perron (1988), Dickey-Fuller Aumentado (ADF)⁷ proposto por Said e Dickey (1984), entre outros.

⁷ O teste ADF é uma extensão do teste de Dickey-Fuller (1979) e consiste em testar a nulidade do parâmetro ρ da regressão na regressão $\Delta Z_t = \rho Z_{t-1} + \sum_{j=1}^k \delta_j \Delta Z_{t-j} + \varepsilon_t$, onde $\Delta = 1 - B$ e $k = n^{1/4}$. Para maiores detalhes ver Morettin e Toloí (2004).

Uma desvantagem dos testes de raiz unitária tradicionais no estudo da co-integração é que não são apropriados quando as séries em estudo são fracionalmente integradas. Uma alternativa a esses métodos é utilizar o estimador GPH para processos fracionalmente co-integrados.

Definição 12. Seja $\mathbf{Z}_t = (Z_{1t}, Z_{2t}, \dots, Z_{kt})$ um vetor de séries $I(d)$, as componentes de \mathbf{Z}_t são ditas fracionalmente co-integradas se o erro de co-integração é fracionalmente integrado de ordem d^* , onde $d^* < d$.

3 ANÁLISES E RESULTADOS

3.1 ANÁLISE DESCRITIVA DAS VARIÁVEIS

Para avaliar a relação existente entre a Taxa de Juros Internacional (X_t) e o Volume de Operações Crédito Externo (Y_t) utilizaram-se as seguintes *proxies*: Taxa de Juros *Libor* por dólar americano (% a.a.) e o Saldo das Operações de Crédito Externo realizado por pessoas jurídicas referenciadas em taxas de juros externas (% do PIB). O conjunto de dados possui 68 observações coletadas, mensalmente, no período dezembro/2002-agosto/2008⁸ (ver Figura 6). Todos os cálculos foram realizados no *Software* estatístico R versão 2.7.1 e disponível no site [HTTP://www.r-project.org](http://www.r-project.org).

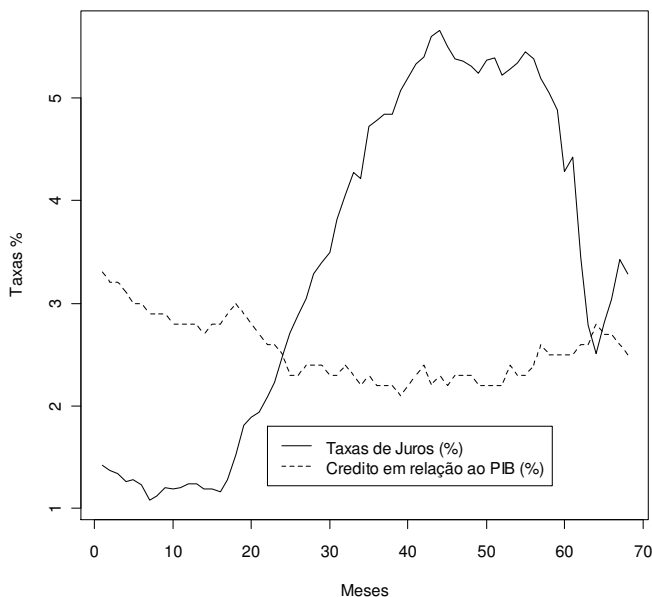


Figura 6. Séries temporais das taxas de juros internacionais e do volume de operações de crédito externo.
Fonte: Banco Central do Brasil.

⁸ Será utilizado “juros” e “crédito” para se referir à taxas de juros internacional e volume de operações de crédito externo no restante deste documento.

A Figura 6 sugere a existência de uma relação inversa entre os *juros* e *crédito*. Sendo assim, quando as Taxas de Juros internacionais se elevam, o Saldo das Operações de Crédito Externo tende a se reduzir. Isto ocorre porque, o aumento nos *juros* produz elevação dos custos para contratar operações de *crédito*, como mostra o gráfico de dispersão apresentado na Figura 7. Esse comportamento sugere relações de longo prazo entre as variáveis, que serão analisadas na seção seguinte.

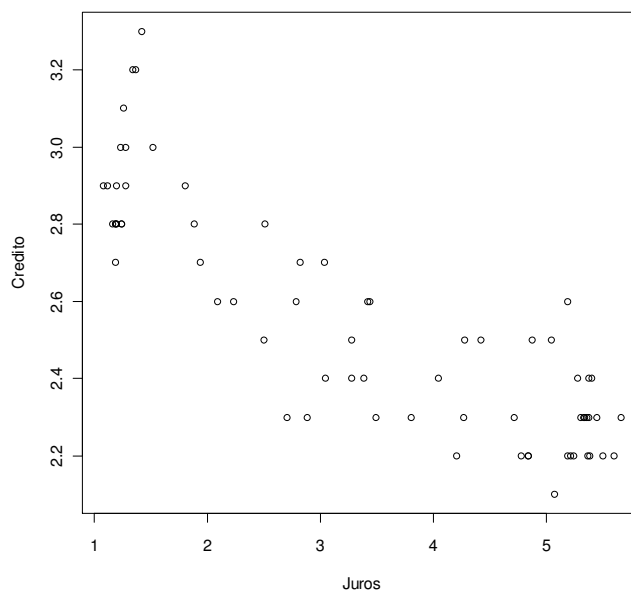


Figura 7. Gráfico de dispersão juros contra crédito.

3.2 ANÁLISES DE RELAÇÕES DE LONGO PRAZO

Neste trabalho, o procedimento adotado para verificar a existência de relação de longo prazo das variáveis compreendeu duas etapas: 1º) verificar a igualdade da ordem de integração das variáveis *Juros* e *Crédito*; 2º) caso as variáveis sejam da mesma ordem de integração, então, verificar a estacionariedade do vetor de resíduos.

Os correlogramas das séries *Juros* e *Crédito* mostram um decaimento lento que sugere comportamento não-estacionário na série (ver Figura 8). A confirmação desse diagnóstico é possível ser feita por meio da aplicação dos testes ADF (ver Tabela 2) e o teste de GPH (ver Tabela 3).

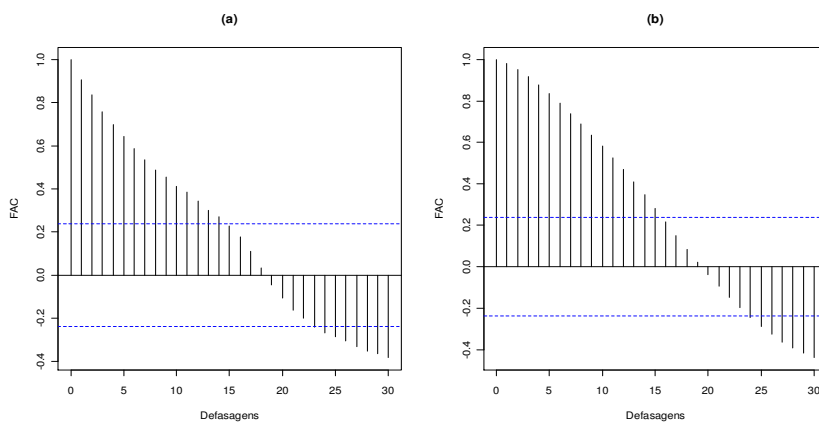


Figura 8. FAC das variáveis em estudo: (a) FAC Crédito; (b) FAC Juros.

Os resultados dos Testes ADF, na Tabela 2, indicam que as variáveis Crédito (Y_t) e Juros (X_t) são $I(1)$, portanto, não-estacionárias. A confirmação desse diagnóstico pode ser feita com base no estimador GPH na Tabela 3.

Tabela 2. Teste de Dickey-Fuller Aumentado aplicado às Variáveis.

Variável	Defasagens	Estatística t_τ	Valor p	Integração
Y_t	4	-1,2029	0,8989	I(1)
X_t	4	-0,9828	0,9340	I(1)

Na Tabela 3, o valor da estimativa d confirma a não-estacionariedade das variáveis Y_t e X_t para diferentes valores do *bandwidth* dessa estatística. Todos os cálculos foram feitos pelo *Software* estatístico R versão 2.7.1 com uso da sub-rotina *fracdiff*. Lembre-se que, de acordo com a Definição 8, quando $d \geq 0,5$ a variável é não estacionária.

Tabela 3. Ordem de integração das variáveis obtida a partir do estimador GPH com $bandwidth\ n^\alpha$

Variável	Estatísticas	$\alpha = 0,5$	$\alpha = 0,7$	$\alpha = 0,9$
Y_t	Estimativa d	1,2543	1,1566	1,1804
	Erro-Padrão	0,3480	0,1955	0,1386
X_t	Estimativa d	1,2235	0,7874	0,7256
	Erro-Padrão	0,3480	0,1955	0,1386

* Observação: o conceito de $bandwidth$ utilizado no teste GPH é diferente daquele definido, na seção 2.4, para regressão não-paramétrica. Para saber mais sobre a escolha ótima do $bandwidth$ no teste GPH ver Robinson (1994b).

A ordem de integração das variáveis indica o risco de obter regressões espúrias entre as variáveis, como enunciado na Seção 2.6. Uma forma de contornar esse problema é verificar a existência de um vetor de co-integração entre as variáveis, que possibilitará o ajuste da curva de regressão com séries em níveis⁹.

Então, como visto acima, X_t e Y_t são $I(1)$, agora será preciso verificar se o termo de erro (ε_t) da regressão entre essas variáveis é $I(0)$. Esse tipo de análise consistirá na utilização de resíduos provenientes de quatro tipos de modelos de regressão, sintetizados na Tabela 4.

Tabela 4. Tipos de modelos utilizados na aplicação.

Modelos Paramétricos	Modelos Não-Paramétricos
Modelo 1: $Y_t = \beta_1 + \beta_2 X_t + \varepsilon_t$	Modelo 3: $Y_t = m_k(X_t) + \varepsilon_t$
Modelo 2: $Y_t = \beta_1 + \beta_2 \ln X_t + \varepsilon_t$	Modelo 4: $Y_t = m_s(X_t) + \varepsilon_t$

* Observação: Os termos m_k e m_s são utilizados aqui apenas para diferenciar os modelos 3 (Kernel ou estimador núcleo) 4 (Splines).

A justificativa para escolha dos modelos na Tabela 4 é com intuito de comparar dois tipos de modelos paramétricos e não-paramétricos de maior popularidade: o Modelo 1 representa o modelo de regressão linear simples; o Modelo 2 corresponde a uma do modelo conhecido como Log-Linear, que é utilizado em situações que envolvem uso da variável taxas

⁹ O termo “níveis” refere-se ao uso das séries sem diferenciação.

de juros (GUJARATI, 2004); o Modelos 3 corresponde ao modelo de regressão não paramétrica com estimador de núcleo; e, por fim, Modelo 4 refere-se ao modelo de regressão não-paramétrica com uso de *Splines*¹⁰.

Após obter os resíduos dos modelos da Tabela 4, será verificada a estacionariedade dos resíduos. Para ilustrar a estrutura de autocorrelação dos resíduos de cada um dos modelos. A Figura 9 apresenta os correlogramas dos resíduos dos Modelos 1 e 2 (paramétricos). O diagnóstico da FAC sugere que esses resíduos são $I(0)$. Os testes estatísticos ADF (ver Tabela 5) e GPH (1993) (ver Tabela 6) serão utilizados para confirmar essa análise.

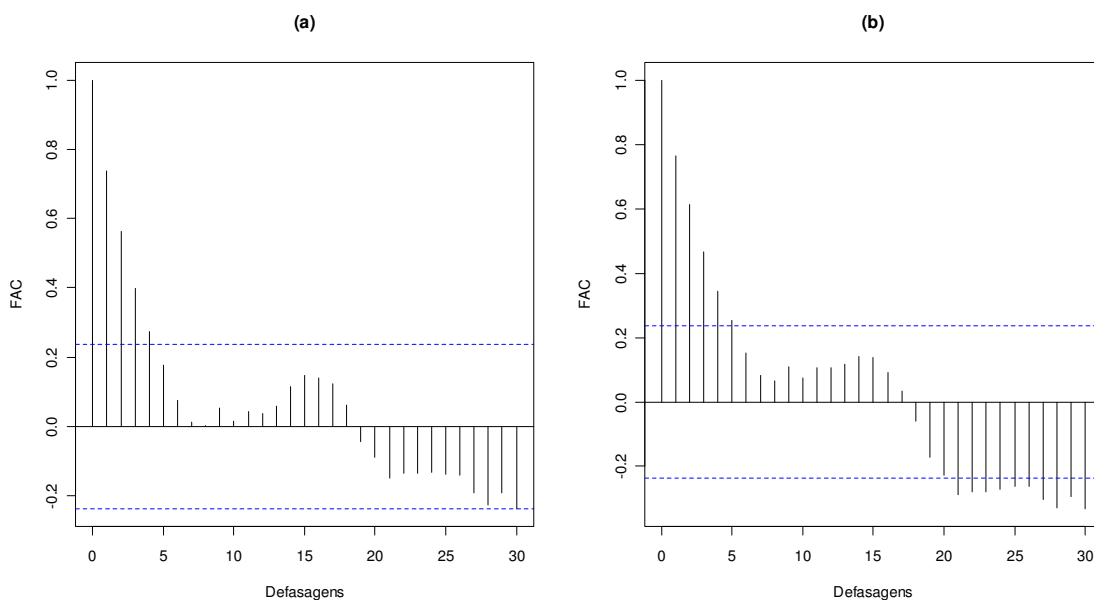


Figura 9. FAC dos resíduos de regressão paramétrica: (a) Modelo 1; (b) Modelo 2.

O que mostra a análise de estacionariedade dos resíduos da regressão não-paramétrica com estimadores de núcleo? A Figura 10 mostra que os resíduos Modelo 3 é $I(0)$ e o Modelo 4, por sua vez, é $I(1)$. Sendo assim, os resíduos da regressão não-paramétrica

¹⁰ Para detalhes do método não-paramétrico *Splines*, ver Eubank (1999), Wahba (1990) e Wegman (2003).

com uso de estimadores de núcleo indicam a existência de relações de longo prazo entre as variáveis em estudo. A confirmação dessas análises é feita por meio da aplicação de testes estatísticos, conforme Tabelas 5 e 6.

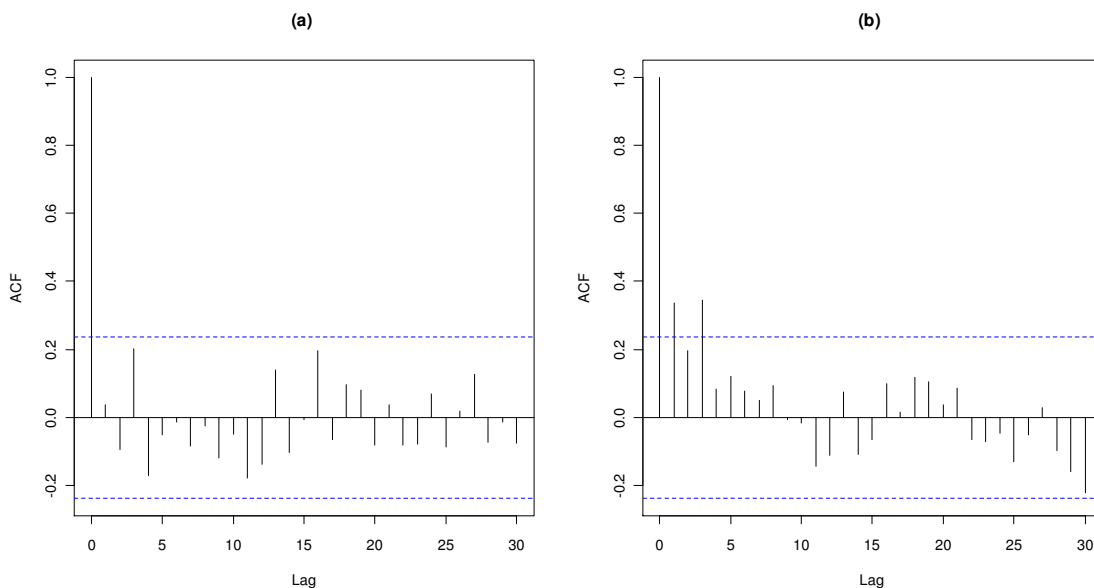


Figura 10. FAC dos resíduos da regressão não-paramétrica: (a) Modelo 3; (b) Modelo 4.

A Tabela 5 mostra os resultados da aplicação do teste ADF para os resíduos das regressões paramétrica e não-paramétricas. Segundo esse teste, os resíduos obtidos a partir do Modelo 2 e 3 são $I(0)$ e os resíduos dos Modelos 1 e 4 são $I(1)$.

Tabela 5. Teste de Dickey-Fuller Aumentado aplicado aos resíduos

Resíduos das regressões	Defasagens	Valor p	Integração
Modelo 1: Linear	4	0,1810	$I(1)$
Modelo 2: Log Linear	4	0,0452	$I(0)$
Modelo 3: <i>Kernel</i>	4	0,04356	$I(0)$
Modelo 4: <i>Splines</i>	4	0,2544	$I(1)$

Mas, será que as conclusões baseadas no Teste ADF são realmente corretas? Qual é a interpretação possível com uso de estimadores fracionários? A presença de autocorrelações estatisticamente significativas nos resíduos do Modelo 2, conforme Figura 9 (b), sugere aplicação do teste GPH.

A Tabela 6 indica que o Modelo 3 é o único com resíduos estacionários. Note que a ordem de integração obtida a partir do estimador GPH oferece uma interpretação distinta do teste ADF, no tocante a estacionariedade dos resíduos do Modelo 2. O que sugere que teste ADF não é adequado para aplicações em processos com ordem de integração fracionária (Franco e Reisen (2007); Santander, Reisen e Abraham (2003); Toscano, Reisen e Tojeiro (2006); Reisen, Cribari e Jensen (2005)).

Tabela 6. Ordem de integração dos resíduos obtida a partir do estimador GPH com $bandwidth\ n^\alpha$

Variável ou Resíduo da regressão	Estatísticas	$\alpha = 0,5$	$\alpha=0,7$	$\alpha=0,9$
Modelo 1	Estimativa d	0,6588	0,7271	0,5761
	Erro-padrão	0,3392	0,2133	0,1389
Modelo 2	Estimativa d	0,7113	0,7840	0,6713
	Erro-padrão	0,3480	0,1955	0,1385
Modelo 3	Estimativa d	-0,1667	0,0539	0,1189
	Erro-padrão	0,2215	0,1492	0,1199
Modelo 4	Estimativa d	0,6628	0,5612	0,3602
	Erro-padrão	0,2403	0,2051	0,1285

As estimativas do parâmetro d para os resíduos dos modelos acima também foram obtidas pelo método de Reisen (1994). Os resultados reforçaram que $d < 1$, e na estimativa de alguns modelos, os valores de \hat{d} foram menores que 0,5, isto é, o que indicam resíduos estacionários.

De acordo com exposto, é possível concluir que existe uma relação de dependência de longo prazo entre as variáveis *Juros* e *Crédito* no período de dezembro/2002 e agosto/2008, sendo possível a análise de regressão pelo método não-paramétrico com estimadores de núcleo, sem o risco de obter relações espúrias entre as mesmas.

3.3 COMPARAÇÃO DOS AJUSTES DE REGRESSÃO

Após constatar a relação de dependência de longo prazo entre as variáveis, pergunta-se: qual é a influência das Taxas de Juros externos sobre o volume de operações de crédito externo contratado por empresas no Brasil no período de dezembro/2002 e agosto/2008? Essas questões serão discutidas a seguir.

Apesar de não ter sido evidenciada a estacionariedade nos resíduos dos Modelos 1, 2 e 4, as curvas de regressão desses modelos foram comparadas à curva de regressão não-paramétrica com estimador de núcleo (Modelo 3), como demonstrado na Figura 11.

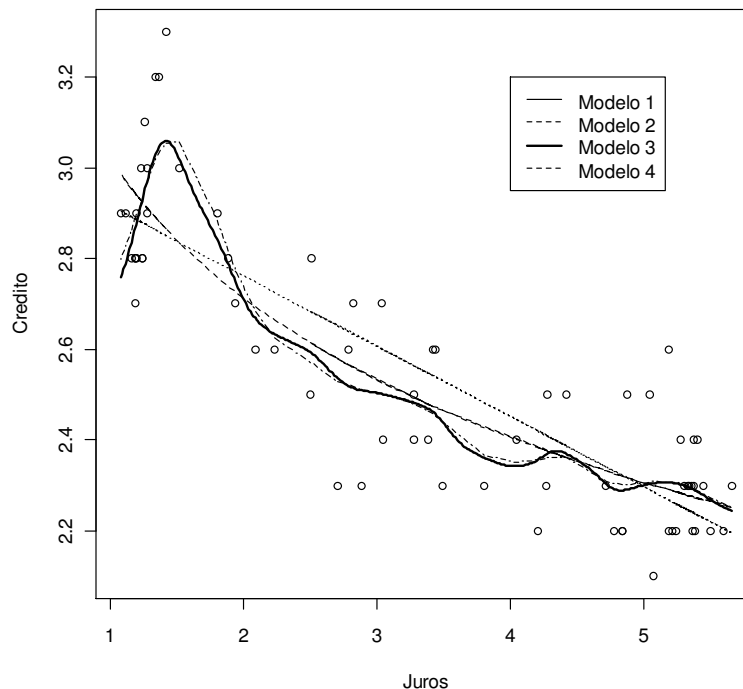


Figura 11. Ajustes de curvas de regressão método paramétrico versus não-paramétrico

A figura revela algumas limitações dos modelos paramétricos (Modelo 1 e 2) na

tentativa de explicar a dinâmica das variáveis em estudo, por não captar as oscilações apresentadas ao longo da série. O ajuste da curva de regressão não-paramétrica *Splines*¹¹ (Modelo 4) apresentou resultados próximos à curva de regressão não-paramétrica com estimadores de núcleo (Modelo 3)¹².

A curva do modelo de regressão não-paramétrica com estimador de núcleo revela que, no intervalo de Taxas de Juros Internacionais correspondente a 1% e 2%, as variáveis *Juros* e *Crédito* possuem uma associação não necessariamente negativa¹³. A interpretação

¹¹ Os cálculos do parâmetro de suavização ou *bandwidth* (λ) e o ajuste da curva de regressão não-paramétrica *Splines* (Modelo 4) foram feitos com uso da sub-rotina *smooth.splines* do *Software* estatístico R versão 2.7.1. O valor do *bandwidth* pelo critério de Validação Cruzada Generalizada (GVC) $\lambda = 0,0206$.

¹² O *bandwidth* (h) utilizado para ajuste do Modelo 3 foi calculado através do critério de Validação Cruzada (Definição 4), o valor de $h=0,73$. A função de núcleo utilizada foi a *Epanechnikov*. Os cálculos foram feitos por meio da sub-rotina *locpoly* do *Software* estatístico R versão 2.7.1.

¹³ Os testes estatísticos utilizados para verificar a significância dos parâmetros dos Modelos 1 e 2 (paramétricos) revelam que os mesmos são significativos. A interpretação das estimativas dos parâmetros mostra que a elevação de 1 ponto percentual na Taxa Anual de Juros Internacionais causa a redução média de 0,15 pontos

econômica possível para esse comportamento da curva é que, ao patamar de taxas de juros internacionais reduzidos, os agentes tendem ser indiferentes à elevação dessas taxas. O intervalo de confiança de 95% do Modelo 3 apresentado na Figura 12.

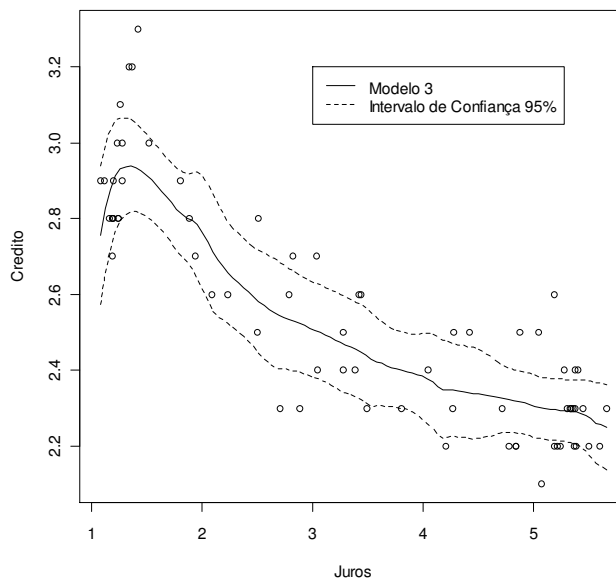


Figura 12. Intervalo de Confiança de 95% para Regressão com estimador de núcleo

A Figura 13 exibe gráficos de dispersão dos resíduos contra o tempo para os diferentes tipos modelos ajustados. Com base na análise da trajetória dos resíduos dos Modelos 1 e 2 (Figuras 13 (a) e (b), respectivamente) é possível identificar existência de um padrão. Nas Figuras 13 (c) e (d) os resíduos das regressões dos Modelos 3 e 4 são apresentados, onde se observa que os resíduos do Modelo 4 apresenta um padrão quadrático, enquanto o Modelo 3 não parece ter padrão algum que mostre correlação serial nos resíduos.

percentuais no Volume de Crédito Externo em relação ao PIB. O coeficiente angular do Modelo 2 significa que um aumento de 1%, em média, na magnitude das Taxa de Juros Internacionais será acompanhada de uma redução de 0,44 no montante de Credito de Externo em relação ao PIB, no período estudado.

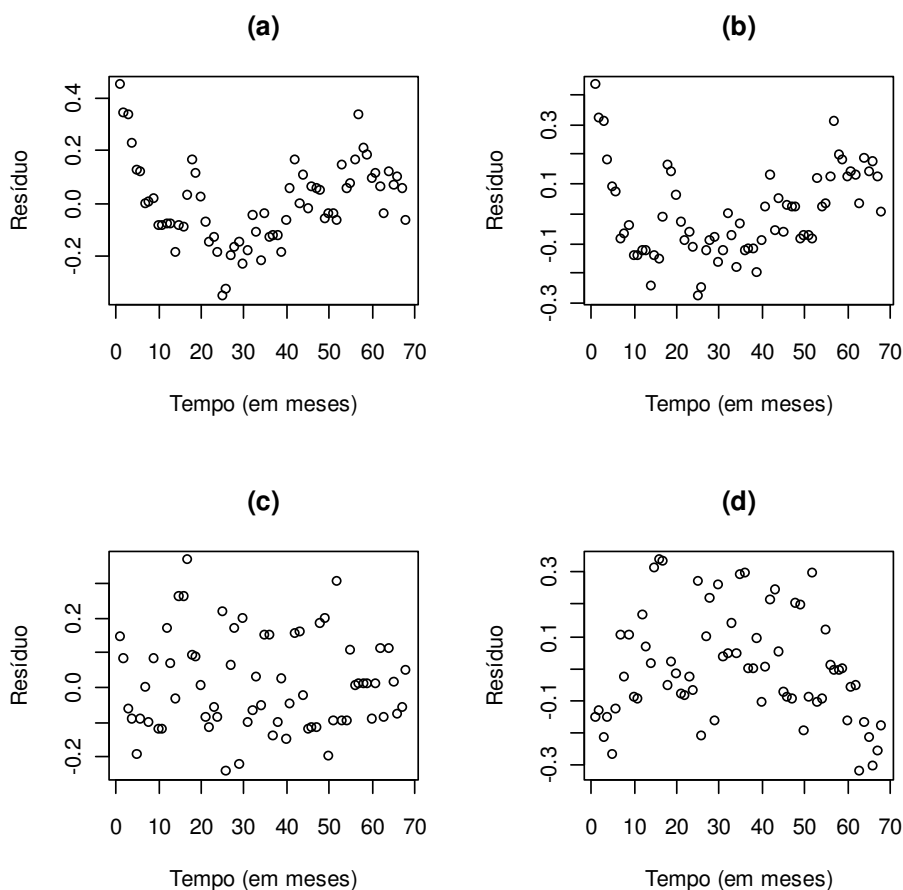


Figura 13. Gráfico de dispersão dos Resíduos contra tempo: (a) Modelo 1; (b) Modelo 2; (c) Modelo 3; e (d) Modelo 4.

A Figura 14 apresenta os resíduos dos Modelos 1, 2, 3 e 4 observados contra a variável *Juros*. Com base nos gráficos, parece haver ausência de heterocedasticidade. No caso dos modelos paramétricos, o teste de Jarque-Bera não mostrou violação da suposição de normalidade nos vetores de resíduos¹⁴.

¹⁴ Os métodos de regressão não-paramétricos não requerem análise de normalidade dos resíduos, porque esses estimadores não fazem nenhuma suposição sobre o tipo de distribuição dos dados.

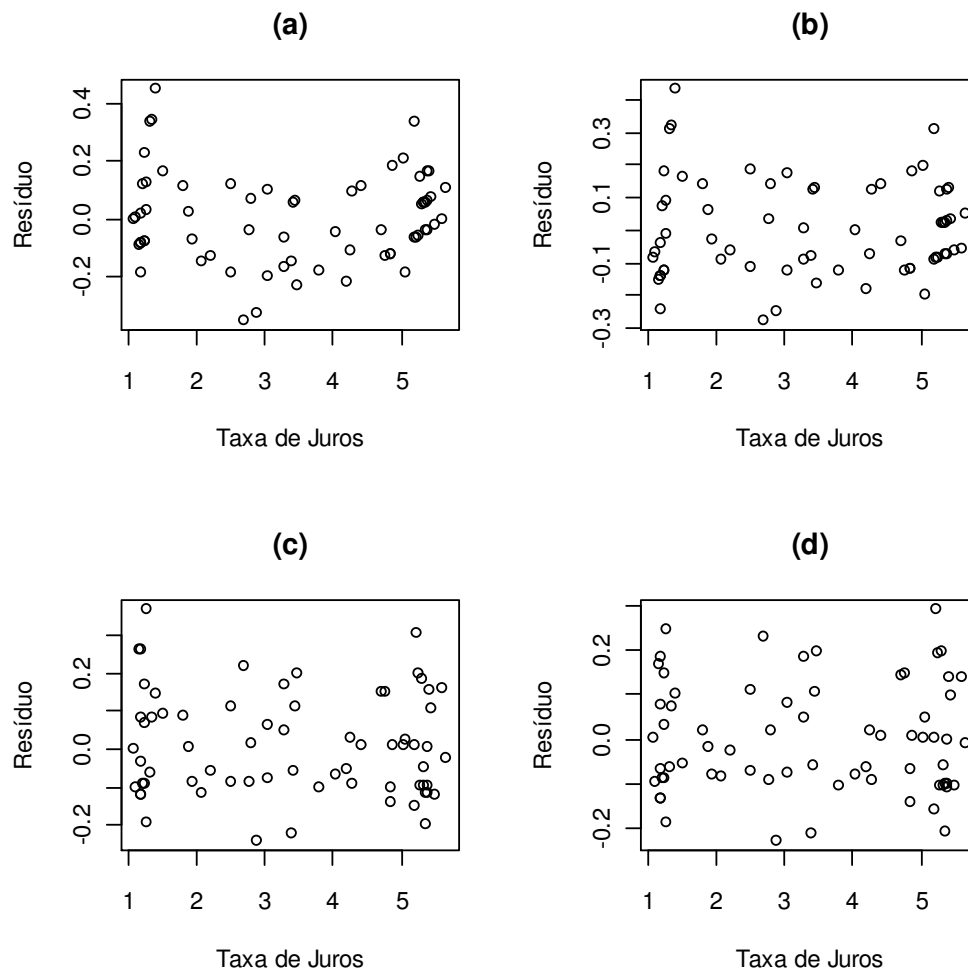


Figura 14. Gráfico de dispersão dos Resíduos contra variável Taxa de Juros: (a) Modelo 1; (b) Modelo 2; (c) Modelo 3; e (d) Modelo 4.

4 CONCLUSÕES E RECOMENDAÇÕES

A aplicação do método não-paramétrico com estimadores de núcleo demonstrou que as variáveis taxas de juros internacionais e o volume de operações de crédito externo realizado por empresas brasileiras entre o período de dezembro/2002 e agosto/2002 possuem uma associação negativa, no intervalo de taxas de juros entre 2% e 6%.

A curva ajustada pela técnica de regressão não-paramétrica com estimador de núcleo captou algumas oscilações na curva de regressão dos *juros* contra *crédito*, que não foram observadas com base em duas especificações paramétricas testadas. A trajetória da curva de regressão não-paramétrica com estimador de núcleo demonstrou que, no intervalo de Taxas de Juros correspondente a 1% e 2%, os aumentos na Taxa de Juros não implicam necessariamente numa redução do volume de operações de crédito externo realizadas por empresas brasileiras. A interpretação econômica possível para este comportamento é que, ao patamar de taxas reduzidas, os agentes tendem a ser indiferentes às oscilações dessas taxas.

Os resultados da regressão não-paramétrica com estimador de núcleo, associadas a métodos de estimação de parâmetros de memória longa do modelo ARFIMA (Geweke e Porter-Hudak (1983); Reisen, 1994), evidenciaram a existência de relações de longo prazo entre as variáveis estudadas.

O trabalho demonstrou a viabilidade de se utilizar técnicas de cointegração fracionária combinadas com regressão não paramétrica. Como recomendação para trabalhos futuros sugere-se considerar regressão não paramétrica na presença de erros correlacionados, conforme proposições de Chiu (1989), Chu (1991), Hart (1991), Chu e Marron (1991), Altman (1990), Opsomer, Wang e Yang (2001), entre outros. O uso de técnicas de *bootstrap* pode ser uma alternativa interessante para escolha do *bandwidth* do estimador de núcleo. No

que tange a co-integração fracionária, uma sugestão é avançar, teórica e empiricamente, no estudo desse assunto. Finalmente, sugere-se testar a inclusão de outras variáveis explicativas (como taxas de juros internas, taxas de cambio, taxas doméstica de crescimento econômico), partindo assim para o desenvolvimento de modelos de regressão multivariados com base em técnicas semiparamétricas (por exemplo, modelos aditivos generalizados).

5 REFERÊNCIAS

ABADIR, K. M.; TAYLOR A.M. (1999) On the definitions of (co-)integration. **Journal of Time Series Analysis**, **20**, 129-137.

ALTMAN, N. S. (1990) Kernel smoothing of data with correlated errors. **Journal of the American Statistical Association**, **85**, 749–759.

CHIU, S. T. (1989) Bandwidth selection for kernel estimate with correlated noise. **Statistic Probability Letters**, 347-354.

CHU, C. K. (1991). Comparison of two bandwidth selectors with dependent errors. **Annals of Statistic**, **19**, 1906–1918.

CHU, C. K.; MARRON, J.S. (1991). Choosing a Kernel Regression Estimator 1991. **Statistic Science**, **6**, 404-436.

DICKEY, D. A.; FULLER, W. A. (1979). Distribution of the Estimators for Autorregressive Time Series With Unit Root. **Journal of the American Statistical Association**, v. 74, 427-431.

EUBANK, E. G. (1999). **Nonparametric Regression and Spline Smoothing**. Texas: Textbooks and Monographs.

FOX, R., TAQQ, M.S. (1986). Large-sample properties of parameter estimates for strongly dependent stationary Gaussian time series. **The Annals of Statistics**, v.14, 517-532.

FRANCO, G. C., REISEN, V. A. (2007). Bootstrap approaches and confidence intervals for stationary and non-stationary long-range dependence processes. **Physica**, v. 375, 546 - 562.

GASSER, T. H. (1984). Nonparametric regression analysis of growth curves. **Annals of Statistics**, v. 12.

GEWEKE, J.; PORTER-HUDAK, S. (1983). The estimation and application of long memory time series. **Journal of Time Series Analysis**, v. 4, 221-238.

GHYSELS, E. S.; SERENA, N.. (1996). A Semi-Parametric Factor Model for Interest Rates. **Scientific Series Cirano**.

GRANGER, C. W. J.; NEWBOLD, P. (1974). Spurious regression in econometrics. **Journal of Econometrics**, v. 2, 111-120.

- GREMAUD, A. P.; VASCONCELLOS, M. A.; TONETTO JUNIOR, R.. (2002). **Economia Brasileira Contemporânea**. 4ª Edição. Editora Atlas. Rio de Janeiro.
- HÄRDLE, W. (1990). **Applied Nonparametric Regression**. Cambridge University Press, England: v.19, 261 p.
- HÄRDLE, W. (1993). Comparing nonparametric versus parametric regression fits. **Annals of Statistics**, 21, 1926-1947.
- HÄRDLE, W.; MÜLLER, M.; SPERLICH, S. ; WERWATZ, A. (2004). **Nonparametric and Semiparametric Models**. Berlin: Springer Series in Statistics.
- HART, J. D. (1991). Kernel regression estimation with time series errors. **Journal Royal Statistical Association** , 53 , 173–187.
- HART, J. D.; WEHRLY. T.E. (1986). Kernel regression estimation using repeated measurements data. **Journal of the American Statistic Association**, 81, 1008-1088.
- HOSKING, J. (1981). Fractional differencing. **Biometrika**, 165-176.
- LEMONTE, A. J.; REISEN, V. A.; TOSCANO, E. M. M. O Poder do teste de Dickey-Fuller aumentado contra processos ARFIMA(p,d,q). **Revista Brasileira de Estatística**, v. 65, n. 223, 2005.
- MAUNIS, M. L. (2004). **O investimento fixo das empresas registradas para negociação em Bolsa de Valores: 1987-2002**. Rio de Janeiro: PUC-RIO.
- MORETTIN, P.; TOLOI, C. M. C. (2004). **Análise de séries temporais**. São Paulo: Edgar Blücher.
- NADARAYA, E. A. (1964). On Estimating Regression. **Theory of Probability and Its Applications**, 9.
- NADARAYA, E. A. (1965). On nonparametric estimates of density functions. **Theory of Applied Probability** (10), 186–190.
- OPSOMER, J.; WANG, Y; YANG, Y. (2001). Nonparametric Regression With Correlated Errors. **Statistical Science**, 16, 134-153.
- PAGAN, A.; Ullah, A. (1999). **Nonparametric Econometrics**. Cambridge University Press. Nova York.

R DEVELOPMENT CORE TEAM (2007). A Language and Environment. **R Foundation for Statistical Computing**. Vienna, Austria.

RACINE, J. (2008). Nonparametric Econometrics: A primer. **Econometrics**, 3, 1-88.

REISEN, V. A. (1994). Long memory time series models. **Journal of Time Series Analysis**, 335-350.

REISEN, V. A.; CRIBARI NETO, F.; JENSEN, M. Long Memory inflationary dynamics; The case of Brazil. **The Studies For Nonlinear Dynamics And Econometrics**, v. 7, n. 3, p. 1-15, 2003.

ROBINSON, P. M. (1983). Nonparametric estimators for time series. **Journal of Time Series Analysis**, 185-207.

ROBINSON, P. M. (1994a). Semiparametric analysis of long-memory time series. **Annals of Statistics**, 22, 515-539.

ROBINSON, P. M. (1994b). Rates of convergence and optimal spectral bandwidth for long range dependence. **Probab. Thoery Related Fields**, 99, 443-473.

SANTANDER, L. A. M.; REISEN, V. A.; ABRAHAM, B. Non-cointegration tests and a fractional ARFIMA process. **Statistical Methods Journal**, v. 5, n. 1, p. 1-22, 2003.

SOWELL, G. (1992). Tests for unit roots: a monte carlo investigation. **Journal of Business and Economic statistics**, v.53, 147-159.

TOPEL, R.; ROSEN, S. (1988). Housing Investment in the United States. The University of **Chicago Press**, v. 96, n. 4, 718-740.

TOSCANO, E. M., REISEN, V. A., TOJEIRO, C. A. Short, long or unit root process? **Statistical Methods**, V. 8, 226-242. 2006.

VIVACQUA, M. V. (2007). **Política Monetária e Investimento no Brasil**. Dissertação (Mestrado em Economia). Programa de Pós-Graduação em Economia, Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro.

WAHBA, G. (1990). **Spline models for observational data**. Philadelphia: Siam.

WAND, M. P; JONES, M. C (1995). **Kernel Smoothing**. New York: Chapman and Hall.

WATSON, G. S. (1964). Smooth Regression Annalysis. **Sankhya**, 26, 359-372.

WEGMAN, E. J. (1983). Splines in statistics. *Journal of American Statistical Association* , 78, 351-365.

YATCHEW, A. (2003). *Semiparametric Regression for the Applied Econometrician*. Nova York: Cambridge University Press.

YULE, G. U. (1926). "Why do We Sometimes Get Nonsense Correlations Between Time Series? A Study in Sampling and the Nature of Time Series". **Journal of the Royal Statistical Society**, 1-64.

ZIEGELMANN, F. (2000). Nonparametric Estimation of Volatility Functions: The Local Logistic Estimator. **Finance Lab Working Paper**, 15.