

# MICROESTRUTURA EMPÍRICA DE MERCADO - UMA ANÁLISE PARA A TAXA DE CÂMBIO BRL/US\$ USANDO DADOS DE ALTA FREQUÊNCIA.

MÁRCIO POLETTI LAURINI  
LUIZ GUSTAVO CASSILATTI FURLANI  
MARCELO SAVINO PORTUGAL

RESUMO. Neste artigo realizamos uma análise de microestrutura de mercado empírica para o mercado de câmbio BRL/US\$ usando dados de cotações de bid e ask de alta frequência (tick by tick). Os objetivos do artigo são verificar a importância da presença de informação assimétrica na dinâmica dos preços, elaborar um modelo para o processo de descoberta de preços e analisar os determinantes empíricos do spread entre bid e ask através de um modelo condicional que captura uma resposta assimétrica do spread em relação ao conjunto de informações passadas.

A verificação da presença de informação assimétrica é realizada através de um teste não-paramétrico de independência condicional para a propriedade Markoviana. O modelo de descoberta de preços é construído usando um vetor de correção de erros entre bid e o ask, controlando para efeitos de duração e volatilidade. Como resultado deste vetor, construímos uma série de desvios do spread de equilíbrio, e mostramos que a distribuição condicional dos desvios do spread de equilíbrio responde de forma assimétrica aos spreads passados e durações e volatilidades esperadas. Esta decomposição é realizada usando como ferramentas o quantilograma e um modelo de autoregressão quantílica. Nós relacionamos os efeitos encontrados a diversos fatos estilizados encontrados na literatura teórica de microestrutura de mercado.

RESUMO. This paper provides an analysis of empirical microstructure for the exchange rate market in Brazil (BRL/US\$) using the high frequency bid and ask quotations data (tick by tick). The aims of the article are to verify the importance of the presence of asymmetric information in the dynamics of prices, to build a model for the prices discovery process and to analyze the empirical determinatives of spread between bid and ask through a conditional model that captures an asymmetric reply for the spread regarding past information. The asymmetric information hypotheses is tested through a non-parametric test of conditional independence for the Markovian property. The model of prices discovery is built using an error correction vector between bid and ask, controlling for the duration and volatility. As a result of this vector, we build a series of the deviation from the equilibrium spread, and we show that conditional distribution of spreads deviations from equilibrium answers asymmetrically to the spreads changes and expected conditional volatilities and durations. This is made by using the quantilogram and a quantile autoregression model as tools. We relate the findings to some facts presented in the theoretical literature of market microstructure.

## 1. INTRODUÇÃO

A análise dos efeitos empíricos relacionados à microestrutura nos mercados de taxas de câmbio é uma área de pesquisa de grande desenvolvimento recente<sup>1</sup>. Existe um reconhecimento que os determinantes de curto prazo na precificação de ativos podem estar mais relacionados às estruturas de mercado do que a fatores ligados aos fundamentos dos ativos, conforme apontado por [Flood & Taylor, 1996]. A literatura de microestrutura de mercado indica que, no curto prazo, fatores como custos de transação, balanceamento de estoques e prêmios por liquidez podem ser mais importantes nos preços dos ativos do que fatores ligados a fundamentos macroeconômicos.

A literatura de taxas de câmbio fornece uma motivação fundamental para o estudo dos efeitos da microestrutura de mercado, ao mostrar uma série de evidências de que a abordagem macroeconômica tradicional de determinação das taxa de câmbio só tem poder explicativo sobre movimentos

---

<sup>1</sup>Duas obras de referência na literatura de microestrutura do mercado de câmbio são [Frenkel *et al.*, 1996] e [Sarno & Taylor, 2002]. A monografia de [Lyons, 2001] é um estudo profundo da microestrutura de mercado baseada no fluxo de ordens.

de longo prazo e em situações extremas como hiperinflações e crises cambiais. Em situações de funcionamento normal do mercado de câmbio os movimentos são determinados pela microestrutura do mercado (e.g. [Flood & Taylor, 1996], [Taylor, 1995] e [Frankel & Rose, 1995]).

O outro fator que possibilita o estudo dos efeitos de microestrutura de câmbio é a maior disponibilidade de informações estruturadas sobre as operações no mercado intradiário de câmbio, permitida pela existência de sistemas proprietários de transações, como os sistemas Reuters 2000-2 Dealing System, o Electronic Broking System (EBS), Spot Dealing System e no Brasil o sistema Sisbex da BM&F. Informações destes sistemas são coletadas de forma sistemática e tornadas publicamente disponíveis através produtos como o sistema Reuters e o sistema Bloomberg Data License (utilizado neste estudo), permitindo a realização de estudos abrangentes de microestrutura de mercado usando dados de operações em alta frequência (operação por operação) nestes mercados.

A disponibilidade destas informações permite estudar algumas características do mercado de câmbio que não podem ser sistematicamente explicadas pelos modelos macroeconômicos usuais. Entre os efeitos não explicados está a grande persistência nos retornos de câmbio nos dados intradiários, relacionada a desvios da propriedade martingale dos retornos que caracterizaria eficiência de mercado e ausência de condições de arbitragem. Outros efeitos não explicados pela análise macro são os determinantes do spread entre o bid e o ask; a importância do conteúdo informacional presente no fluxo de ordens e seu poder preditivo sobre taxas futuras; os impactos de análise charlista no mercado de câmbio; influência do volume transacionado, localização espacial dos agentes e volatilidade na determinação dos preços e a importância da informação privada da determinação de preços e spreads<sup>2</sup>

Uma diferença fundamental dos modelos de microestrutura de mercado em relação aos modelos de determinação macroeconômica são as restrições teóricas assumidas. Modelos de determinação macroeconômica normalmente são baseados em estruturas de agentes representativos, informação simétrica, expectativas racionais e ausência de custos de transação. Os modelos de microestrutura de mercado, em oposição a estes pressupostos, são usualmente caracterizados por estruturas de assimetria e heterogeneidade<sup>3</sup>. Existem diferenças no conjunto de informação dos agentes e diversos tipos agentes - neste mercado coexistem traders, market-makers e clientes com objetivos estratégicos e conjuntos de informação diferentes.

Outra característica importante dada pela presença de efeitos de microestrutura é relacionada a eficiência do funcionamento do mercado, que pode ser afetada por características institucionais. Como o mercado de câmbio é descentralizado<sup>4</sup> e os operadores estão fisicamente separados, os conjuntos de informações entre os agentes são distintos, tornando informações privadas relevantes na determinação dos preços. Este diferencial de informações não apenas permite, mas torna comum a existência de arbitragem, podendo afetar o grau de eficiência no funcionamento do mercado, como relatado em [Flood, 1994].

A riqueza do conteúdo informacional presente nas operações intradiárias permite analisar questões não possíveis de serem tratadas através de dados de frequência mais baixa. Além dos preços, as operações intradiárias carregam outras fontes de informação relevantes, como o tempo decorrido entre duas ordens no mercado (durações de ordens) e a possibilidade de identificação parcial da fonte de informação de cada ordem através do fornecedor da cotação. O tempo decorrido entre duas ordens está ligado à chegada de novas informações no mercado e também é uma medida inerente de liquidez<sup>5</sup>. Estas informações são relevantes nos modelos de microestrutura de mercado em função da possibilidade de que o preço seja impactado pelas operações mais recentes

<sup>2</sup>Veja [Sarno & Taylor, 2002], Cap. 9, para uma extensiva revisão bibliográfica sobre microestrutura do mercado de câmbio.

<sup>3</sup>Veja [O'Hara, 1995] para uma revisão dos modelos teóricos de assimetria de informação no contexto de microestrutura de mercado e [Hasbrouck, 2007] para implicações empíricas.

<sup>4</sup>[Sarno & Taylor, 2002] contém uma descrição das estruturas e agentes nos mercados de câmbio, e para uma descrição minuciosa do funcionamento do mercado de câmbio BRL/US\$ veja [Garcio & Urban, 2004].

<sup>5</sup>Para uma revisão da literatura de modelagem da duração e a análise do conteúdo informacional da duração condicional veja [Engle & Russel, 1998], [Engle, 2000]. [Fernandes & Grammig, 2005a] e [Fernandes & Grammig, 2005b]. são referências sobre os modelos econométricos de duração condicional e testes de especificação nestes modelos.

([Hasbrouck, 1991] e [Dufour & Engle, 2000]), fazendo com que os preços e o spread presentes na próxima operação sejam afetados não só pelos preços passados, mas também pelo volume, spread e o tempo decorrido nas operações passadas.

O objetivo do nosso artigo é analisar efeitos empíricos de microestrutura de mercado revelados pelas cotações intradiárias de bid e ask no mercado da taxa de câmbio R\$/US\$. Analisamos neste artigo a importância de informações privadas no mercado através de um teste da propriedade Markoviana (Seção 4). O resultado deste teste motiva um modelo empírico de descoberta de preços usando um vetor de correção de erros entre o bid e o ask tendo como variáveis explicativas as durações passadas de preços e a volatilidade presente nas ordens. Com este modelo conseguimos também verificar o impacto nos preços das operações mais recentes. Usando os resultados desse modelo, analisamos os determinantes dos desvios do spread de equilíbrio entre bid e ask, elaborando uma modelagem que permite resposta assimétrica do spread ao conjunto de informação passada.

As principais contribuições do artigo na literatura de microestrutura empírica de mercado são: a análise da dinâmica das cotações de bid e ask no mercado intradiário de câmbio BRL/US\$, o uso do teste da propriedade Markoviana dos preços derivado por [Fernandes & Amaro de Matos, 2007] para dados de mercado de câmbio e a análise da assimetria na distribuição condicional do spread através do uso do quantilograma de [Linton & Whang, 2007] e de métodos de autoregressão quantílica ([Koenker & Xiao, 2006]).

O artigo está estruturado da seguinte forma: as Seções 2 e 3 descrevem os dados utilizados, mostram algumas características destas séries e comentam a relação com trabalhos anteriores utilizando dados de câmbio BRL/US\$; a Seção 4 verifica a presença de informação assimétrica usando o teste da propriedade Markoviana; Seção 5 descreve o modelo vetorial de correção de erros usado no processo de descoberta de preços e a Seção 6 analisa os padrões periódicos no spread e os efeitos de assimetria na distribuição condicional do spread. As conclusões finais estão na Seção 7.

## 2. FONTES DE DADOS E TRABALHOS RELACIONADOS

**2.1. Bases de Dados e Trabalhos Anteriores.** Embora exista uma grande literatura sobre microestrutura de mercado de câmbio, ainda são poucos os trabalhos analisando a microestrutura do mercado de câmbio BRL/US\$. Os trabalhos mais importantes sobre microestrutura do mercado de câmbio BRL/US\$ são [Garcio & Urban, 2004] e [Wu, 2007]. O primeiro trabalho descreve de forma minuciosa as instituições e o funcionamento do mercado interbancário de câmbio no Brasil, mostrando os agentes, instituições e os mecanismos de operação existentes. Além desta descrição do funcionamento do mercado, o artigo apresenta uma evidência econométrica da existência de uma direção de causalidade de Granger dos retornos do mercado futuro para o mercado spot, utilizando dados diários.

O trabalho de [Wu, 2007] é um estudo completo das microestruturas do mercado de câmbio no Brasil usando dados agregados diariamente do sistema Sisbacen do Banco Central do Brasil, que contém todas as operações consolidadas de câmbio realizados no Brasil, permitindo inclusive identificar todas as partes envolvidas nesta operação. O completo banco de dados do fluxo de ordens utilizado por [Wu, 2007] é o único estudo na literatura mundial que utiliza uma base de dados que compreende 100% das operações oficiais de câmbio em um país, permitindo isolar o efeitos dos fluxos de câmbio relacionados a operações comerciais, financeiras, intervenções do Banco Central no mercado de câmbio e as consequências destes fluxos sobre a taxa de câmbio.

São duas as possíveis críticas ao banco de dados utilizados por [Wu, 2007]. A primeira é que o conjunto de dados utilizado não corresponde a informação publicamente disponível no tempo de decisão dos agentes. Parte das informações utilizadas não são tornadas públicas pelo Banco Central, e parte das informações é disponibilizada com atraso, não sendo assim equivalente ao conjunto de informação disponível utilizada pelos agentes na tomada de decisões no mercado intradiário.

Nossa análise difere desses estudos anteriores pelo uso das cotações de bid e ask observadas em alta frequência, contendo todas as cotações, operação a operação, existentes no mercado. Esta

TABLE 1. Formato dos Dados

BRL   2003-09-30 12:51:35   B   2.9   BNPL     A   2.905   PION
BRL   2003-09-30 12:51:28   B   2.9   ALLI     A   2.904   BNPL
BRL   2003-09-30 12:51:08   B   2.9   ALLI     A   2.902   PION
BRL   2003-09-30 12:51:06   B   2.9   ALLI     A   2.903   ALLI
BRL   2003-09-30 12:50:38   B   2.9   ALLI     A   2.902
BRL   2003-09-30 12:50:12   B   2.9       A   2.902
BRL   2003-09-30 12:50:10   B   2.899   BNPL     A   2.902
BRL   2003-09-30 12:49:16   B   2.901   PION     A   2.903
BRL   2003-09-30 12:46:17   B   2.899   PION     A   2.9025   BNPL
BRL   2003-09-30 12:46:12   B   2.899   PION     A   2.902
BRL   2003-09-30 12:46:11   B   2.899   PION     A   2.902   ALLI

metodologia é análoga àquelas utilizadas em [Goodhart, 1989] e [Bollerslev & Domowitz, 1993], e permite analisar a influência dos efeitos relacionados a microestrutura no funcionamento do mercado intradiário de câmbio<sup>6</sup> presentes nas cotações intradiárias.

Nossa fonte de dados consiste no livro de cotações de operações de câmbio spot fornecido pelo sistema de dados Bloomberg Data License. Este formato de dados é conhecido como FXX DATA. Este sistema coleta informações sobre operações realizadas em vários mercados, como o Sisbex (Sistema de negociação eletrônica de dólares da Bolsa de Mercadorias & Futuros) e operações over the counter, contando com informações fornecidas por vários participantes de mercado. A amostra utilizada neste estudo contém todas as ordens gravadas neste sistema iniciando em 28/05/2006 e finalizando em 30/11/2006. Uma amostra com o formato dos dados utilizados está exemplificado na Tabela 1.

O arquivo de dados contém 8 colunas. A primeira coluna identifica o ativo (moeda) transacionado. A segunda coluna contém o momento de registro da operação no sistema, com precisão de segundos. A terceira e a quarta colunas identificam o bid e o preço de bid; a quinta coluna identifica o agente responsável pelo fornecimento do valor de bid. Colunas 6, 7 e 8 identificam o ask, o preço de ask e o operador responsável pelo valor do ask. Note que nem sempre os dados identificam qual o operador responsável pelo bid e pelo ask, já que é possível entrar ordens anônimas no mercado.

Os dados de alta frequência do mercado de câmbio apresentam muitas limitações em relação a outros bancos de dados utilizados em microestrutura empírica. O TAQ (Transaction and Quotes) de ações negociadas na NYSE, por exemplo, que contém informações sobre as transações realizadas, o preço efetivamente utilizado na transação e volume de cada operação. Os dados de câmbio contém apenas a evolução das cotações, mas não contém o valor efetivamente transacionado. Outras limitações são a ausência de informação de volume e a impossibilidade de descobrir se a ordem foi iniciada por um comprador ou por um vendedor, informações existentes na base de dados TORQ (Transaction Orders and Quotes) compilada em [Hasbrouck, 1992]. A não existência desses dados (preços, volumes e se a transação foi iniciada por um comprador ou por um vendedor) sobre as transações efetivamente realizadas decorre da inexistência de uma cláusula de *disclosure* para operações de câmbio.

Embora o escopo das análises de microestrutura de mercado sobre este conjunto de dados seja limitado pela condição de *no disclosure*, é importante notar que este é o conjunto de informação disponível publicamente para os operadores do mercado spot, e corresponde as indicative quotes como analisadas em [Goodhart, 1989] e [Bollerslev & Domowitz, 1993].

Existem algumas evidências de que a omissão das transações não afeta o resultado das estimativas, como discutido em [Agmon & Barnea, 1977] e [Glassman, 1987], mas também são apontadas

<sup>6</sup>[Lyons, 2001] contém uma descrição completa dos diferentes bancos de dados utilizados em microestrutura de mercados de câmbio.

na literatura alguns problemas no uso das cotações, já que são apenas indicações e não transações. [Lyons, 1996] mostra que os spreads interdealers são menores do que os existentes nas cotações indicativas<sup>7</sup>. Contra estas críticas pode-se argumentar que estas cotações correspondem a informação publicamente disponível no mercado spot, e assim sobre o qual os agentes decidem. Estudos como o de [Lyons, 1996] que utilizam dados de um dealers particular podem ser criticados em dois aspectos - o intervalo de tempo estudado é muito mais curto (semanas no máximo) e eles capturam o comportamento das transações de um dealer particular e não necessariamente podem ser representativos do comportamento do mercado, dada a grande heterogeneidade dos agentes no mercado de câmbio.

Um ponto importante é notar que os dados de alta frequência utilizados nesse trabalho contém duas informações relevantes não contidas nos dados utilizados em outros estudos para o mercado de câmbio como [Wu, 2007]. A informação do horário da transação nos permite construir a variável de duração das ordens, dada pelo tempo decorrido entre a chegada de duas ordens no mercado. Esta variável carrega consigo informações de liquidez e volatilidade de mercado, e sua evolução representa a chegada de novas informações ao mercado ([Engle & Russel, 1998], [Engle, 2000], [Fernandes & Grammig, 2005a]). Outra variável construída é a volatilidade condicional derivada de um modelo GARCH para a média entre o bid e o ask como uma proxy do preço transacionado. O uso do *mid quote* como preço de transação é justificado na literatura de microestrutura pois em alguns modelos o mid quote está relacionado com o preço “correto” do ativo<sup>8</sup>.

**2.2. Construção e Filtragem dos Dados.** Os dados utilizados no artigo são as séries temporais de informações de preços de bid e ask dadas pela Tabela 1. Duas variáveis adicionais são construídas: a variável de duração, através do número de segundos entre a chegada de duas ordens, e a variável de mid quote, dada pela média entre o bid e o ask, que será utilizada na construção da proxy de volatilidade através de um modelo GARCH para esta variável.

O armazenamento em tempo real destes dados gera um número relativamente grande de operações com informações incorretas<sup>9</sup>. A correção foi realizada retirando-se observações claramente aberrantes dadas por erros de grafia (por exemplo um bid gravado como 0.219 ao invés de 2.19), observações com spreads negativos ou não compatíveis com o comportamento local do spread. Estes outliers foram filtrados usando uma regra de filtragem que trata como outliers operações cujo spread entre bid e ask esteja acima de 10 desvios padrões do spread. Esta regra captura principalmente erros de grafia no banco de dados.

Outra verificação de consistência realizada no horário da transação permite analisar se esse foi gravado corretamente e está na ordem temporal correta. Note que, seguindo o padrão da literatura de microestrutura de câmbio, não limitamos o horário das transações, já que a cotação nos mercados de câmbio é realizada 24 horas por dia, pois o mercado está em funcionamento contínuo nos 3 grandes centros (EUA, Europa e Japão). A limitação de informações levaria a exclusão de informações que servem de referência para a formação do preço, pois informações de outros mercados fora do horário de transação do Brasil podem afetar a determinação do câmbio no mercado doméstico. Após essas filtrações destas observações nosso banco de dados consiste em 279737 observações operação por operação no intervalo de tempo entre 28/5/2006 e 30/11/2006.

A construção da variável Volatilidade é através de um modelo GARCH(1,1) para a série de mid quotes entre o bid e o ask. Esta variável é construída de forma a ser um proxy da volatilidade do preço transacionado, que não é observado diretamente. Esta variável é obtida a partir de um modelo GARCH(1,1)<sup>10</sup> estimado por máxima verossimilhança com parâmetros ( $\omega = 1.51E09, \alpha = .0.220275, \beta = .7773732$ ). A série de durações é obtida como o número de segundos entre duas transações. Note que o mercado de câmbio é um mercado de funcionamento 24 horas por dia, e assim existem cotações durante todo o dia. Nós não excluimos as transações fora dos períodos de transação normal do mercado brasileiro, já que as cotações nestes períodos podem ser indicadores

<sup>7</sup>Para obter esta conclusão [Lyons, 1996] utiliza o fluxo de ordens de um dealer particular.

<sup>8</sup>Veja [Hasbrouck, 2007] para uma revisão dos procedimentos e modelos utilizados em modelagem empírica de microestrutura de mercado.

<sup>9</sup>Veja [Falkenberg, 2002] para mais detalhes sobre os problemas no tratamento de dados de alta frequência.

<sup>10</sup>O modelo não é colocado no texto por restrições de espaço, mas pode ser obtido por solicitação aos autores.

FIGURA 1. Bid-Ask e Spread

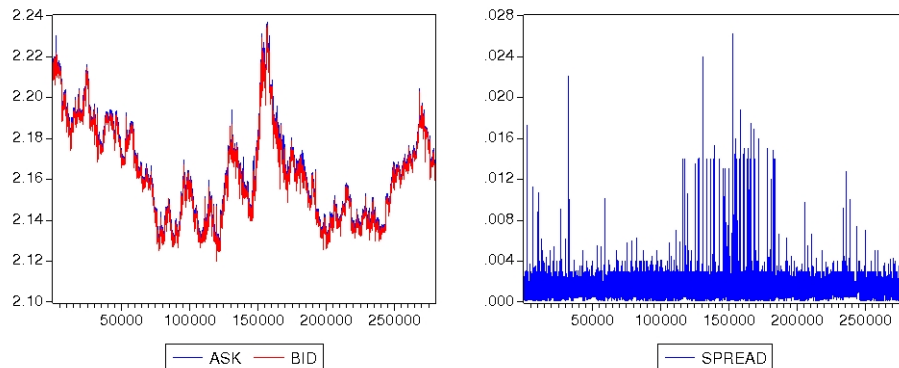
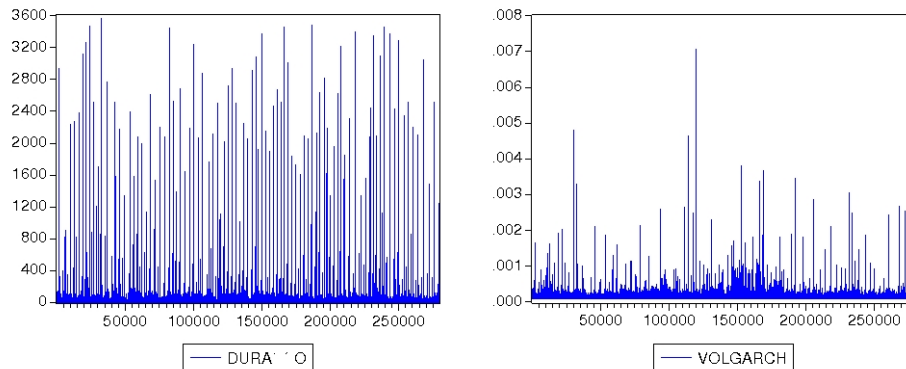


FIGURA 2. Durações e Volatilidades



de preço em mercados externos. Existem cerca de 4% de cotações fora do horário de transação normal do sistema Sisbex da BM&F.

### 3. ESTATÍSTICAS DESCRITIVAS E PADRÕES DE PERIODICIDADE (SPREADS, VOLATILIDADE E DURAÇÕES)

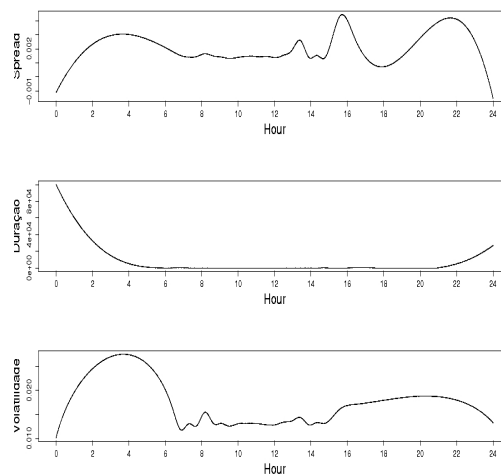
**3.1. Estatísticas Descritivas e Testes de Raiz Unitária.** As Figuras 1 e 2 mostram os gráficos das séries de bid e ask, spread, durações e a variável Volatilidade. No Apêndice temos uma datação de eventos relevantes na série de câmbio no período em estudo. As figuras apontam que o período de elevação nos spreads correspondem aos maiores valores alcançados pelas séries de bid e ask. A Tabela 2 mostra as estatísticas descritivas destas variáveis, bem como testes Phillips-Perron de raízes unitárias para as séries de bid e ask, e log retornos de bid e ask. Os resultados dos testes indicam que não podemos rejeitar a hipótese nula de raiz unitária para bid e ask em nível, mas as séries de log retornos são estacionárias.

**3.2. Padrões Periódicos.** Um fato estilizado em séries financeiras intradiárias é a presença de padrões de periodicidade dentro do dia. [Zivot & Yan, 2003] é uma referência na análise destes padrões de periodicidade. Para verificar os padrões de periodicidade, realizamos um ajuste não paramétrico que captura os padrões periódicos dentro do dia, rodando o modelo de smoothing spline tendo como variáveis dependentes spreads, durações e volatilidades. Um smoothing spline ([Green & Silverman, 1994]) pode ser definido com a solução do problema de se minimizar o seguinte funcional:

TABELA 2. Estatísticas Descritivas

	Ask	Bid	Log Retornos Ask	Log Retornos Bid	Duração	Variância
Média	2.163719	2.162170	-2.67E-07	-2.07E-07	12.75741	4.17E-08
Mediana	2.162000	2.160500	0.000000	0.000000	1.999999	2.47E-08
Máximo	2.236900	2.355400	0.010915	0.0100925	3563.100	0.000509
Mínimo	2.122600	2.12000	-0.009199	-0.008751	0.000000	1.54E-08
Desvio Padrão	0.022946	0.02887	0.000187	0.000203	72.92399	1.76E-06
Assimetria	0.502882	0.494570	1.111605	-0.13605	24.9654	54.49727
Curtose	2.515426	2.490172	247.0023	176.0372	814.3413	4827.558
Jarque-Bera	14518.0	14433.53	6.93E+08	8.20E+11	7.96e+09	7.46E+14
Probabilidade	0.000000	0.000000	0.00000	0.000000	0.000000	0.000000
pvalor raiz unitária pp	0.0372	0.0348	0.0001	0.0001	-	-
Observações	279737	279737	2797376	279736	279736	279736

FIGURA 3. Padrões Periódicos



(a) Smoothing Spline - 24 horas

$$(1) \quad S_\lambda(g) = \sum_{i=1}^n (Y_i - g(X_i))^2 + \lambda \int (g''(x))^2 dx$$

onde  $g$  é permitido ser qualquer curva,  $x$  é o conjunto de dados e  $\lambda$  é o que controla a suavidade do ajuste, controlando a troca entre minimização do resíduo e a aspereza (roughness) do ajuste. Para obter os padrões, ajustamos smoothing splines utilizando 24 nós para cada série. Não foram encontrados padrões de periodicidade relevantes nos log retornos de bid e ask. A Figura3 mostra os padrões obtidos para o spread, a duração e a volatilidade.

Os padrões observados mostram que o spread tende a ser bem maior fora dos horários de transação no mercado brasileiro (intervalo de 2-6 horas). Mas a figura b mostra que existe uma tendência de aumento do spread nos horários de abertura e fechamento do mercado de câmbio no Brasil, e são análogos ao efeito de um padrão “U” obtido por [Bollerslev & Domowitz, 1993] e consistentes com o modelo teórico de determinação do spread de [Admati & Pfleiderer, 1988]. Um efeito relevante é que o spread tende a aumentar próximo as 17 horas, que é horário limite de registro de transações de câmbio no Sisbacen. Este efeito pode ser racionalizado pelo fato de que transações não registradas no Sisbacen tem que ser desfeitas, levando a custos de ajuste e desbalanceamento de inventário e margens de risco.

TABELA 3. Distribuição do último dígito do Spread

último Dígito	Número	Proporção
zero	16833	6.017438%
um	135928	48.59136%
dois	111837	39.97934%
três	11260	4.025409%
quatro	1259	0.4500656%
cinco	434	0.1551457%
seis	445	0.159078%
sete	893	0.3192284%
oito	657	0.2348635%
nove	191	0.06827842%
Teste $\chi^2$	Estat.	pvalor.
$H_0$ – probabilidades iguais	842649.8	< 2.2e-16

O padrão periódico para a série de duração mostra que fora dos horários de pregão no Brasil a duração de tempo entre duas cotações é bem longa, mas nos horários compatíveis com funcionamento do mercado no Brasil as durações são mais curtas, tendo um ligeiro aumento próximo das 17 horas. Os padrões periódicos para a série de volatilidade mostram que as volatilidades são grandes fora dos horários de transação (devido ao menor número de cotações os saltos nos preços serão maiores), e que dentro do horário de funcionamento do mercado existe uma tendência de aumento na volatilidade, possivelmente associada com o fechamento de posições. Note que estas figuras indicam uma possível correlação entre o spread, a duração e a volatilidade. Esta associação será testada nas seções 5 e 6.

Uma hipótese adicional sobre a distribuição nos spreads a ser testada é a existência de efeitos de aglomeração no spread, o que pode ser consistente com um comportamento de colusão de preços ([Christie *et al.*, 1994], [Hasbrouck, 1999]). A hipótese de colusão de preços pode ser sintetizada como uma tendência das cotações do spread permanecerem em múltiplos da variação mínima permitida. No caso da séries de câmbio, por exemplo, o spread mínimo é 0.001, mas as distribuições do spread tenderiam a permanecer concentradas em um certo valor múltiplo do valor mínimo. Uma forma simples de testar este efeito é verificar a distribuição do último dígito do spread. Na hipótese nula de não aglomeração de preços, a distribuição do último dígito deveria ser uma distribuição uniforme e a proporção de valores em cada dígito deveria ser estatisticamente igual. Esta é uma das abordagens seguidas em [McGroarty *et al.*, 2006], que contém uma análise extensa do efeito aglomeração para taxas de câmbios.

A Tabela 3 mostra a distribuição do último dígito do spread. Nesta tabela vemos que o último dígito do spread se concentra nos valores 1, 2 e 3, indicando que o spread tende a se concentrar na faixa de valores de 0.001 a 0.003. O teste de igualdade de proporções é realizado utilizando um teste Qui-Quadrado, que rejeita a hipótese nula de proporções iguais. Note que embora o teste rejeite a hipótese nula de não aglomeração, ele não indica qual a possível causa da aglomeração de preços. Os trabalhos de [Hasbrouck, 1999] e [Hasbrouck, 1999] discutem algumas causas possíveis para esse efeito e [Hasbrouck, 1999] complementa esta análise com um modelo dinâmico. Nestes modelos a aglomeração de preços pode estar relacionada com um custo estocástico de provisão de liquidez enfrentado pelo market maker. Na seção 5 nós mostraremos que o spread de equilíbrio é de cerca de 0.0025, e que a concentração de valores logo abaixo 0.001 e 0.003 unidades está relacionada com desvios de curto prazo do spread de equilíbrio.

#### 4. PROPRIEDADE MARKOVIANA

Uma propriedade do conceito de eficiência de mercado<sup>11</sup>, obtida pelo pressuposto de racionalidade dos agentes e processamento eficiente da informação disponível, é a de que preços de ativos

<sup>11</sup>Veja [Campbell *et al.*, 1997] para uma revisão dos testes empíricos de eficiência de mercado.



devem ser compatíveis com um processo de Markov de primeira ordem. Desta forma, o preço da data t só deve depender da informação mais recente em t-1 mais um processo de inovações:

$$(2) \quad P_t = P_{t-1} + \varepsilon_t$$

Se a informação é processada de forma eficiente, toda a informação disponível<sup>12</sup> até o período t-1 já deve estar contida no preço  $P_{t-1}$ , e assim a variação de preço  $P_t - P_{t-1}$  deve corresponder a um processo de erro não sistemático<sup>13</sup>. Outra característica derivada é que a esperança condicional  $E_Q[P_t|I_t] = P_{t-1}$  na medida neutra ao risco deve ser um processo *martingale*, levando ao conceito de medida martingale equivalente.

Entretanto, uma predição dos modelos de microestrutura baseados em informação assimétrica como [Glosten & Milgrom, 1985] e [Easley & O'Hara, 1987], é a de que a existência de conjuntos de informação diferenciada entre os agentes afetaria a propriedade Markoviana nos preços de bid e ask. Nestes modelos a existência de informação assimétrica faz com que os preços no tempo t dependam de toda a história de transações e não somente da informação mais recente, invalidando a propriedade Markoviana nos preços e, de forma indireta, caracterizando uma forma de ineficiência de mercado. Uma discussão relacionada está em [Flood, 1994], que mostra que a descentralização dos agentes no mercado de câmbio representaria uma forma informacionalmente menos eficiente do que os mercados de ações, cujas transações são centralizadas em bolsas. A descentralização tornaria a disseminação das informações mais lenta, e desta forma os preços seriam correlacionados não só com o preço mais recente mais com um conjunto longo de preços passados.

Dessa forma, podemos testar a presença de informação assimétrica através de testes da propriedade Markoviana. Um teste da propriedade Markoviana que leva em conta o espaçamento irregular das cotações de preços no tempo existente em dados financeiros de alta frequência é o teste proposto por [Fernandes & Amaro de Matos, 2007]. Este teste é um teste não-paramétrico de independência condicional, baseado na hipótese nula de que se propriedade Markoviana é válida, a duração entre duas operações deve ser independente da realização da variável ligada aos preços dos ativos, no caso o spread entre bid e ask.

A hipótese nula do teste derivado por [Fernandes & Amaro de Matos, 2007] é dada por:

$$(3) \quad H_0 : f_{iX_j}(d_i, x, d_j) = f_{i|X}(d_i)f_{x_j}(x, d_j)$$

onde  $f_{iX_j}(d_i, x, d_j)$  representa a densidade conjunta da duração  $d_i$ , do spread  $x$  e duração  $d_j$ ,  $f_{i|X}(d_i)$  a densidade condicional da duração  $i$  e  $f_{x_j}(x, d_j)$  a densidade da distribuição conjunta entre o spread  $x$  e a duração  $j$ , para  $i > j$ . Se a propriedade de independência condicional é válida, a hipótese nula dada pela equação 3 é equivalente a validade da propriedade Markoviana. O teste derivado por [Fernandes & Amaro de Matos, 2007] é baseado na distância quadrática ponderada entre  $f_{iX_j}(d_i, x, d_j) - f_{i|X}(d_i)f_{x_j}(x, d_j)$  quando as densidades são substituídas por estimadores não paramétricos de densidade. A estatística de teste é dada por:

$$(4) \quad \Lambda_{\hat{f}} = \frac{1}{n} w_i(d_{k+j}, X_k, d_k) \left[ \hat{f}_i(d_{k+j}, X_k, d_k) - \frac{\hat{f}_{iX}(d_{k+j}, X_k)}{\hat{f}_X(X_k)} \hat{f}_{X_j}(X_k, d_k) \right]$$

onde os estimadores  $\hat{f}_{(\cdot)}$  são estimadores de densidade por kernel para as densidades conjuntas, marginais e condicionais e  $w_i$  é uma função de ponderação. [Fernandes & Amaro de Matos, 2007] mostram que a estatística dada pela equação 5 tem distribuição assintótica normal padrão :

<sup>12</sup>Formalmente: o processo de informação é uma Filtração  $\mathcal{F}_t$  dada por uma seqüência crescente de sub-sigma álgebras  $B_t \subset B_u \subset B$  para  $0 \leq t \leq T$  definidas em um espaço de probabilidade  $(\Omega, \mathcal{B}, P)$ , e esta filtração representa o conjunto de informação dado pelo passado das séries. Assumimos a filtração usual neste artigo e por isso simplificamos a notação nos casos usuais.

<sup>13</sup>Diferentes categorizações sobre eficiência de mercado se relacionam aos pressupostos sobre o processo  $\varepsilon_t$ . Eficiência tipo III seria associada a um processo  $\varepsilon_t$  não correlacionado, enquanto que eficiência tipo II corresponderia a hipótese mais restritiva de um processo independente  $\varepsilon_t$  e a hipótese mais restritiva de todas (Eficiência tipo I) que o processo  $\varepsilon_t$  seja independente e identicamente distribuído. Testes de eficiência de mercado são baseados em testes para estas propriedades sobre o processo de inovações. ([Campbell *et al.* , 1997] ).

TABELA 4. Teste Não-Paramétrico da Propriedade Markoviana

Série	Estatística de Teste	p-valor
Spread/Duração	135.10	0.000
Spread/Duração Ajustados	102.74	0.000
Desvios dos Spread de Equilíbrio	84.78	0.000

$$(5) \quad \hat{\lambda}_n = \frac{nb_n^{3/2}\Lambda\hat{f} - b_n^{-3/2}\hat{\delta}_\Lambda}{\hat{\sigma}_\Lambda} \xrightarrow{d} N(0, 1)..$$

onde nesta equação  $n$  é o tamanho da amostra,  $b_n$  é o valor de bandwidth,

$$(6) \quad \begin{aligned} \hat{\delta}_\Lambda &= \frac{e_k}{n} \sum_{i=1}^n w_i X_j(d_{k+j}, X_k, d_k) \hat{f}_i(d_{k+j}, X_k, d_k) \\ \hat{\sigma}_\Lambda &= \frac{v_k}{n} \sum_{i=1}^n w_i^2 X_j(d_{k+j}, X_k, d_k) \hat{f}_i^3(d_{k+j}, X_k, d_k) \end{aligned}$$

e  $e_k$  e  $v_k$  são constantes que dependem da função de kernel escolhida. Para realizar o teste da propriedade Markoviana nos preços, seguimos a metodologia proposta por [Fernandes & Amaro de Matos, 2007] e utilizamos as séries de spreads no tempo  $t$ , e de durações nos períodos  $t$  e  $t-1$ , e criamos as séries de log-spreads e log-durações padronizadas segundo a mesma metodologia. Nós computamos os estimadores das densidades marginais, condicionais e conjuntas usando um kernel quártico e as mesmas regras de [Fernandes & Amaro de Matos, 2007] para a escolha do bandwidth,  $e_k$  e  $v_k$ . A Tabela 4 mostra o resultado dos testes da propriedade markoviana para os spreads e durações brutas e também para os spreads e durações ajustados para o padrão de periodicidade encontrado na seção 3. Os resultados mostram que rejeitamos a hipótese nula da validade da propriedade Markoviana de primeira ordem para qualquer nível de significância nos dois conjuntos de dados. Esta evidência apoia a conclusão dada por Richard Lyons - 'Contrary to the asset approach - exchange rate determinations is not wholly a function of public news'<sup>14</sup>.

A rejeição da hipótese nula da propriedade Markoviana corrobora para o mercado de câmbio as predições da existência de efeitos de informação assimétrica apontados por [Glosten & Milgrom, 1985] e [Easley & O'Hara, 1987]. Obtemos resultados análogos aos encontrados por [Flood, 1994], mostrando que a informação diferenciada entre os agentes é importante da determinação dos spreads e que a existência de informação diferenciada afeta o processo de descoberta de preços pelos agentes. Note que é possível explicar a violação da propriedade Markoviana pela estrutura de funcionamento do mercado de câmbio, baseada em uma estrutura de múltiplos dealers com localizações distintas<sup>15</sup>.

A propriedade mais simples é que o processo de descoberta de preço não será baseado apenas no preço dado no instante imediatamente anterior e sim todo o conjunto de informação passada. Veremos que esta propriedade afetará diretamente o modelo vetorial de correção de erros utilizado na descoberta de preço, que terá um longo número de lags para o processo de ajuste da dinâmica de curto prazo.

## 5. MODELO VECM DE DESCOBERTA DE PREÇOS

Uma das predições da literatura de microestrutura de mercado na presença de informações assimétricas é que os agentes devem descobrir o verdadeiro preço de equilíbrio do ativo num processo que foi denominado descoberta de preços<sup>16</sup>. Neste processo os agentes tentam encontrar o preço fundamental do ativo, que é observado nas cotações correntes misturado com os efeitos de microestrutura.

<sup>14</sup>[Lyons, 2001], pg 9.

<sup>15</sup>Veja [Lyons, 2001] para um discussão sobre os efeitos de uma estrutura de múltiplos dealers sobre o mercado de câmbio.

<sup>16</sup>Para maiores referências e modelos econométricos de descoberta de preços veja [Hasbrouck, 1996][Hasbrouck, 1988, Hasbrouck, 1996, Hasbrouck, 2007, Hasbrouck, 1991].

Quando o processo de preços está baseado na existência de diversos preços cotados para o mesmo ativo, como, por exemplo, ações transacionadas em várias bolsas ou então existência de preços de bid e ask, a descoberta do preço fundamental do ativo é relacionada com o mecanismo de busca de um preço de equilíbrio entre as diversas cotações do mesmo ativo. Isso é equivalente a existência de mecanismos de correção dos desvios do preço de equilíbrio em relação aos preços existentes em cada cotação. Essas cotações contêm o preço de equilíbrio mais componentes de microestrutura e choques não previsíveis de forma sistemática.

Uma forma de representar este processo multivariado de descoberta de preços é a partir de um modelo vetorial de correção de erros formulado em tempo de eventos<sup>17</sup>. Assumindo um vetor bivariado de preços  $P_t = [p_{1t} p_{2t}]$ , o modelo de vetorial de correção de erros é representado da seguinte forma:

$$(7) \quad \begin{aligned} \Delta P_t &= \mu_0 + A_1 \Delta P_{t-1} + A_2 \Delta P_{t-2} + \dots + A_k \Delta P_{t-k} + \gamma(Z_{t-1} - \mu_1) + \lambda_1 X_{t-1} + \dots + \lambda_j X_{t-k} \\ Z_{t-1} &= [p_{1t-1} - B_1 p_{2t-1}] \end{aligned}$$

Neste modelo o vetor  $Z_{t-1}$  dá os desvios dos valores de equilíbrio de longo prazo entre  $p_1$  e  $p_2$  e  $B_1$  são os vetores de coeficientes na relação de equilíbrio;  $\gamma$  é o vetor de coeficientes que controla o mecanismo da correção de erros; os coeficientes  $A_i$  representam os coeficientes de ajuste de curto prazo;  $X_{t-k}$  é um vetor de possíveis variáveis explicativas que não sejam cointegradas com  $p_1$  e  $p_2$  e  $\lambda$  um vetor de coeficientes capturando a influência de  $X_{t-k}$  na dinâmica de curto prazo de dada por  $\Delta P_t$ .

A idéia do modelo VECM (*Vector Error Correction Model*) é decompor a dinâmica do ajuste dos preços em dois componentes - um ligado a dinâmica de curto prazo dado pelos componentes  $A_i \Delta P_{t-k}$  e  $\lambda_j X_{t-k}$  e outro ligado a dinâmica de desvios do equilíbrio de longo prazo dada pelo vetor de cointegração  $Z_{t-1}$ . No contexto do modelo de descoberta de preço, o vetor de cointegração representa o equilíbrio entre duas medidas de preço do ativo, e nesse equilíbrio reside uma medida do preço fundamental do ativo, dado por  $Z_{t-1}$  no modelo VECM.

Em nosso trabalho o vetor de preços é dado pelos preços de bid e ask  $P_t = [bid_t ask_t]$  e desta forma o vetor de cointegração captura a relação do valor do spread de equilíbrio dado por  $ask - B * bid$ . O modelo VECM formulado neste artigo tem o objetivo de decompor as variações no bid e no ask no período t em dois componentes: um componente de curto prazo dado pelas variações nos bids e asks passados; outro componente corresponde ao efeito de um conjunto de variáveis explicativas e a um mecanismo ligado ao equilíbrio de longo prazo, dado pela correção dos desvios da relação de longo prazo entre o bid e o ask.

A violação da propriedade Markoviana mostra que o processo de descoberta de preços não pode ser baseado somente no conjunto de informação da observação imediatamente anterior ( $\Delta P_{t-1}$  e  $X_{t-1}$ ), já que as cotações no período t-1 não contêm toda a informação disponível no mercado (a informação privada revelada nas transações anteriores não é incorporada de forma instântanea aos preços, levando a violação da propriedade Markoviana estudada na Seção 4). A consequência é que o processo de descoberta de preço não vai depender somente do preço imediatamente anterior nas de todo o passado, propriedade compatível com alguns modelos de informação assimétrica como [Glosten & Milgrom, 1985] e [Easley & O'Hara, 1987].

Em nosso modelo de equilíbrio também colocamos como variáveis explicativas os valores das durações passadas e das volatilidades condicionais do mid quote<sup>18</sup> passadas. Essas duas variáveis são incluídas como forma de verificar se os bids e asks respondem a efeitos de liquidez e chegada de

<sup>17</sup>Para uma revisão dos modelos multivariados de descoberta de preços veja [Hasbrouck, 2007], e também [Engle, 2000] para modelos de correção de erros em tempo de eventos. Tempo de eventos representa o uso do índice t como um indexador de transações, substituindo o tempo calendário como índice do processo estocástico. [Hasbrouck, 2007] discute as vantagens do uso do tempo de eventos em estudos de microestrutura, que estão relacionadas a processos de deformação temporal que contêm informações significantes sobre o conjunto informacional de cada transação.

<sup>18</sup>A variável de volatilidade condicional é construída estimando um modelo GARCH(1,1) para a série de mid quotes entre bid e ask. O objetivo desta variável é ser um proxy da volatilidade do preço transacionado nas operações passadas. Esta metodologia é análoga a abordagem seguida em [Jorion, 1996].

TABELA 5. Teste de Cointegração

Teste do Traço de Cointegração				
Número de Eq. de Cointegração	Traço	Estat.	Valor Crítico 0.05	Prob.**
Nenhum*	0.002669	940.6628	12.32090	0.0001
No máximo 1	1.86E-06	0.653874	4.129906	0.4789
Teste de Cointegração do Posto (Máximo Autovalor)				
Número de Eq. de Cointegração	Max. Autovalor	Estat.	Valor Crítico 0.05	Prob.**
Nenhum *	0.002846	940.0089	11.22480	0.0000
No máximo 1	0.002669	0.653874	4.129906	0.4789
**MacKinnon-Haug-Michelis (1999) p-valores				
*Indica 1 Eq. de cointegração no nível de 0.05				

novas informações contidas nestas duas variáveis, o que poderia representar impactos nos preços de efeitos de risco, medido pela volatilidade do ativo e liquidez, usando como proxy as durações de preços.

A especificação do modelo VECM é válida na existência de um vetor de equilíbrio entre bid e ask, hipótese que pode ser testada através de um teste de cointegração<sup>19</sup>. Para testar a existência da relação de equilíbrio, nós usamos o teste de cointegração de Johansen, cujos resultados estão colocados na Tabela 5, para as séries de logaritmos de bid e ask. A especificação do teste de cointegração é o modelo de correção de erros baseada uso de 120 lags para variações nos logs dos bids e asks, 24 lags para as durações passadas e 20 lags para as volatilidades passadas, aonde esta especificação foi decidida com o critério de informação de Schwartz. Como discutido na Seção 4, esse grande número de lags é relacionado com a violação da propriedade Markoviana. Os resultados do teste mostram que rejeitamos a hipótese nula de não existência de um vetor de cointegração com p-valor de 0.001 pelos dois testes (posto e traço) obtidos pelo procedimento de Johansen, e não rejeitamos a existência de um vetor pelo mesmo teste, indicando a existência de um mecanismo de equilíbrio de longo prazo entre os logs do bid e ask.

Note que o teste de cointegração de Johansen é baseado em pressupostos de distribuições normais nos resíduos e ausência de quebras estruturais. Estas duas situações poderiam alterar o poder do teste. A hipótese de distribuições normais não é válida para nosso conjunto de dados, já que a curtose e a assimetria indicam distribuições não gaussianas, Desta forma os valores críticos usados no teste devem ser afetados por esta violação. Mas note que existe uma justificativa econômica a priori para a existência de um vetor de cointegração entre o bid e o ask, já que uma relação de desequilíbrio entre bid e o ask, representando um spread não estacionário, levaria a oportunidades sistemáticas de arbitragem. Desta forma não rejeitamos a evidência em favor da hipótese de cointegração, mesmo com os possíveis problemas de distorção no poder do teste.

O modelo vetorial de correção de erros está apresentado parcialmente na Tabela 6, aonde apresentamos o vetor de cointegração estimado (Equação de Cointegração), a matriz de cargas do ajuste dos desvios do longo prazo (Correção de Erros) e os dois primeiros lags do mecanismo de curto prazo<sup>20</sup>. A equação de cointegração mostra que o vetor normalizado de cointegração para o log bid ask é [1 -1.000949], o que representa um spread de equilíbrio de .00259 (a série de ask é 1.00259) a de bid).

Este valor do spread de equilíbrio pode ser explicado por 3 fatores básicos: custos relacionados às funções dos dealers de mercado, custos de carregamento de inventários de moeda e um último fator relacionado à informação assimétrica dado por seleção adversa. Os custos dos dealers são relacionados à provisão de liquidez imediata. Os custos de inventários são relacionados à provisão de liquidez e a possibilidade de que os dealers estejam operando com agentes com informação privilegiada (insiders). O problema de seleção adversa é dado pelo fato de que os dealers não conseguem separar agentes com demandas de liquidez e hedge e os insiders, e desta forma aumentam

<sup>19</sup>Veja Hamilton (1994) para uma descrição de procedimentos de estimação e inferência sobre testes de cointegração e modelos vetoriais de correção de erro.

<sup>20</sup>A Tabela é truncada por restrições de espaço, mas é disponível em contato com os autores.

TABELA 6. Vetor de Correção de Erros

Dinâmica	de	Longo Prazo
Equação	de	Cointegração
LOG(ASK(-1))	1.000000	
LOG(BID(-1))	-1.000949	
	(1.7E-05)	
	[-575570.]	
Correção	de	Erros
	D(LOG(ASK))	D(LOG(BID))
CointEq1	-0.015364	0.014691
	(0.00089)	(0.00094)
	[-17.2781]	[ 15.9325]
Dinâmica	de	Curto Prazo
	D(LOG(ASK))	D(LOG(BID))
D(LOG(ASK(-1)))	-0.209626	0.094811
	(0.00202)	(0.00214)
	[-103.647]	[ 44.3896]
D(LOG(BID(-1)))	0.087291	-0.238502
	(0.00200)	(0.00211)
	[ 32.5363]	[-70.4570]
MDUR(-1)	-2.93E-06	-4.17E-06
	(1.4E-07)	(1.5E-07)
	[-16.5732]	[-27.4075]
MDUR(-2)	-1.38E-07	-9.12E-07
	(7.2E-07)	(1.5E-07)
	[ -16.5732]	[-5.61120]
GARCH(-1)	77.68290	89.27822
	(3.29714)	(3.48201)
	[23.5607]	[25.6398]
GARCH(-2)	-69.63691	-174.6838
	(4.73599)	(5.00153)
	[-14.7038]	[-34.9260]

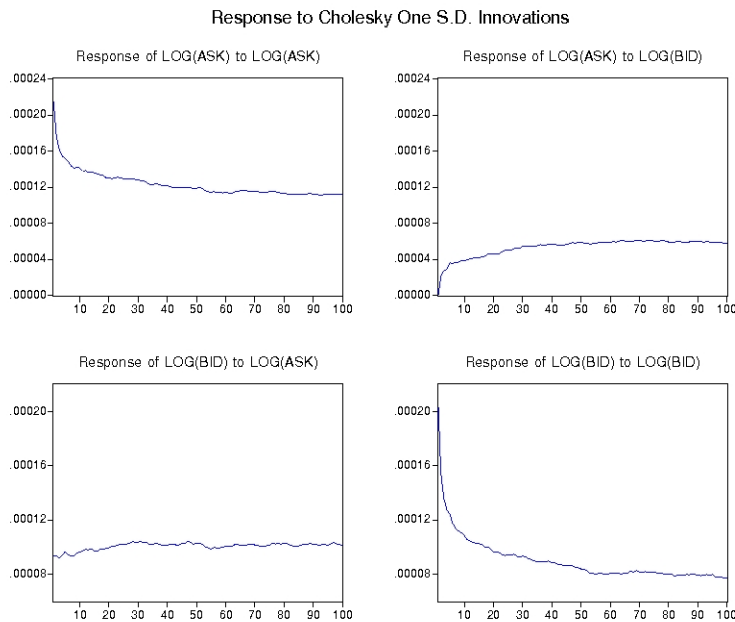
o spread para as duas classes de agentes<sup>21</sup>. Nós analisaremos fatores adicionais na determinação do spread na Seção 6.

A matriz de cargas  $\gamma$  estimada é dada por valor de -0.015364 para a variação do log do ask e 0.014691 para a variação do log do bid. Podemos interpretar estes valores da seguinte forma - desvios positivos do spread de equilíbrio são ajustados com uma variação negativa no ask no ask e variações positivas do bid, e o ajuste é mais rápido no ask, para spreads abaixo do valor de equilíbrio, o mecanismo é inverso, dado pelo aumento do ask e redução no bid. Os mecanismos de curto prazo são mais difíceis de serem interpretados, dado o grande número de lags e a troca de sinal nos coeficientes. Uma evidência é que 120 lags correspondem em média a 20 minutos no tempo calendário, mostrando que o tempo de médio de incorporação de informações pode ser aproximado por este valor.

A figura 4 mostra as funções de impulso resposta generalizadas (independentes da ordenação das séries) obtidas pelo modelo VECM. A figura indica que choques convergem em cerca de 50 observações para o seu valor permanente, e mostram um comportamento estável de convergência aos valores de longo prazo.

<sup>21</sup>Novamente [Frenkel *et al.*, 1996] e [Sarno & Taylor, 2002] para referências sobre fatores determinantes do spread.

FIGURA 4. Impulso-Resposta



## 6. DETERMINAÇÃO DO SPREAD

O modelo VECM estimado na seção 5 permitiu encontrar o valor do spread de equilíbrio e o mecanismo de correção dos desvios do spread de equilíbrio, mas não permite diretamente identificar os fatores que influenciam e impactam o spread. Para analisar os determinantes do spread, começamos por investigar as características empíricas da série de desvios do spread. A partir das características encontradas, formulamos um modelo de resposta assimétrica para a distribuição condicional do spread baseado nos resultados da estimação do quantilograma ([Linton & Whang, 2007]) e um modelo de distribuição condicional assimétrico baseado em uma formulação usando autoregressão quantílica.

Note que a existência de um vetor de longo prazo entre o log bid e o log ask é equivalente à existência de um processo estacionário para os desvios do spread. Para descrever esse processo, nós formulamos um primeiro modelo linear para os desvios do spread. Esta metodologia é baseada em regressões sobre o spread como estudadas em [Jorion, 1996]. Seguindo esta metodologia, formulamos um modelo linear autoregressivo para o spread com a adição de variáveis representando os valores esperados para a volatilidade e para a duração. A formulação desse modelo procura controlar e verificar os impactos estocásticos dos custos dos dealers e o impacto de valores esperados do risco e liquidez na determinação dos desvios do spread de equilíbrio (que podem ser vistos como os desvios sobre os custos médios embutidos no spread de equilíbrio) discutidos na seção 5<sup>22</sup>.

O modelo é baseado em um modelo autoregressivo de terceira ordem para os desvios do spread de equilíbrio, com a adição de previsões um passo a frente para a volatilidade e para a duração. As previsões para a volatilidade são obtidas através do mesmo modelo GARCH utilizado na construção da variável de volatilidade. Para as previsões para a duração, estimamos um modelo de duração condicional autoregressivo<sup>23</sup> ([Engle & Russel, 1998]) e utilizamos como variável no modelo para o spread as previsões um passo a frente obtidas por este modelo. A incorporação das previsões

<sup>22</sup>Uma forma alternativa de olhar para a determinação dos bids e asks é utilizando decomposições de efeitos permanentes-transitórios em séries temporais ([Hasbrouck, 2007]). Os efeitos permanentes estariam ligados aos fundamentos do ativo e os efeitos transitórios a microestrutura de mercado. Como esta abordagem é baseada nos preços efetivamente transacionados nós não utilizamos em nossa análise esta abordagem.

<sup>23</sup>O modelo estimado é um modelo ACD estimado como  $\psi_t = 1.37e-05 + 0.196830x_{t-1} + .884851\psi_t$ , onde  $x_t$  são as durações de preços. Classes mais gerais de modelos ACD poderiam ser utilizadas - veja [Bauwens & Giot, 2000],

TABELA 7. Modelo Linear Para o Spread

DESVIOS DO SPREAD		SPREAD		Method: L	
Variable	Coefficient	Std. Error	t stat	Prob.	
C	-1.63E-06	3.27E-06	-4.996740	0.0005	
PREVISÃO MDURAÇÃO	-7.25E-06	4.88E-06	-1.487233	0.1370	
PREVISÃO VOLATILIDADE	0.087317	0.01883	4.624163	0.0000	
DESVIOS DO SPREAD(-1)	0.673913	0.013154	52.23440	0.0000	
DESVIOS DO SPREAD(-2)	0.136297	0.014759	9.235058	0.0000	
DESVIOS DO SPREAD(-3)	0.130158	0.008060	16.14869	0.0000	
R-squared 0.849712	Mean dependent var -1.54E-05	S.E. of regression 0.000234	Akaike info criterion -13.88596		
Adjusted R-squared 0.778203	S.D.dependent var 0.000593	Sum squared resid 0.231517	Schwarz criterion -13.88573		
Log likelihood 1942215	F-statistic 316321.	Included obs.: 297737 after adjustments	Newey-West HAC (lag trunc=23)		
Durbin-Watson stat 2.01857	Prob(F-statistic) 0.000000				

um passo a frente é incorporar no modelo de determinação do spread o efeito das expectativas dos agentes sobre as volatilidades e durações esperadas para a próxima transação.

Os resultados obtidos nesta decomposição do spread (Tabela 7) mostram que existe uma dependência alta do spread em relação aos spreads passados (a persistência medida pela soma dos coeficientes autoregressivos é aproximadamente .93). Outros efeitos importantes são os sinais obtidos para os coeficientes relacionados as previsões da volatilidade e da duração. Obtemos um sinais positivo para a volatilidade e um sinal negativo, mas não significativo para a duração. Isso pode ser interpretado como um prêmio adicionais no spread para a incerteza, medida pelo valor esperado da volatilidade. A não significância da duração pode ser interpretada como uma medida de liquidez neste mercado, aonde os agentes não precisam pagar um prêmio pela urgência nas transações. Estes efeitos podem ser interpretados como proteção dos dealers em relação a incerteza (volatilidade pode estar relacionada com a chegada de insiders com informações privilegiadas e proteção contra um maior custo de carregamento no aumento da volatilidade). Efeitos semelhantes são obtidos em [Glassman, 1987] que encontra correlação positivas entre spread e volatilidade.

Uma possível deficiência deste modelo é o tratamento simétrico dado aos desvios do spread - spreads abaixo do valor de equilíbrio são tratados de forma igual a valores acima do spread de equilíbrio. Note que estas situações intuitivamente são relacionadas a situações distintas de mercado. Logo, a imposição de mesmo comportamento nas duas situações pode ser uma restrição inválida.

**6.1. Quantilograma.** Para verificar uma possível resposta assimétrica nos desvios do spread, nós utilizamos uma ferramenta conhecida como quantilograma, derivada por [Linton & Whang, 2007]. O quantilograma é uma extensão do correlograma, baseado na média condicional, para a modelagem da dependência existente nos quantiles condicionais da distribuição da série temporal. O quantilograma é também é uma medida de previsibilidade direcional, conforme discutido em [Linton & Whang, 2007] e está inserida na literatura de testes mais gerais de previsibilidade e eficiência de mercado.

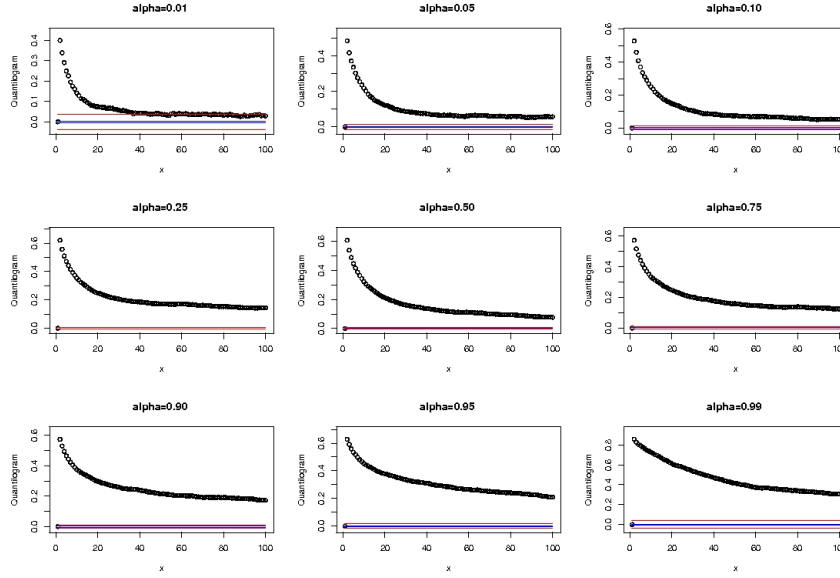
Seja  $y_1, y_2, \dots$ <sup>24</sup> um processo estacionário cuja distribuição marginal tem quantiles  $\mu_\alpha$  para  $\alpha \in (0, 1)$ . Na hipótese nula de não previsibilidade direcional condicionada ao quantile  $\alpha$ :

$$(8) \quad E[\psi_\alpha(y_t - \mu_\alpha) | \mathcal{F}_{t-1}] = 0$$

[Bauwens *et al.*, 2002][Fernandes & Grammig, 2005a] e [Fernandes & Grammig, 2005b] para maiores referências sobre modelos de duração condicional autoregressiva.

<sup>24</sup>A série  $y_i$  analisada pode ser um processo diretamente observado ou resíduos de um modelo estimado em um primeiro estágio, como em nosso caso. [Linton & Whang, 2007] derivam distribuições assintóticas que são válidas nas duas situações. A inferência no quantilograma também é válida na presença de efeitos gerais de heterocedasticidade condicional como processos GARCH estacionários.

FIGURA 5. Quantilograma



(a) Smoothing Spline - 24 horas

onde  $\psi_\alpha(x) = \alpha - 1(x < 0)$  é uma função indicadora que mede se a variável atingiu o quantile  $\alpha$  e  $\mathcal{F}_{t-1} = y_{t-1}, y_{t-2}, \dots$  é uma filtração usual. Na hipótese nula de não previsibilidade direcional, se a variável no período  $t$  está abaixo do quantile  $\alpha$ , a chance é não mais que  $\alpha$  que a série  $y$  atingirá novamente este quantile no período  $t-1$ . Violações desta hipótese são evidências de previsibilidade para condicional à este quantile. Note que uma forma tradicional de teste de eficiência fraca de mercado pode ser colocado neste contexto, utilizando a média condicional:

$$(9) \quad E[y_t - \mu | \mathcal{F}_{t-1}] = 0$$

O quantilograma apresenta duas vantagens em relação às medidas usuais de previsibilidade direcional: estimação de quantiles condicionais é robusta a presença de outliers e certos quantiles da distribuição de retornos de ativos tem uma interpretação direta em gestão de risco, sendo ligadas a medidas como Value At Risk e Expected Shortfall. Note que existe uma interpretação similar na análise do spread - se existe persistência nos quantiles mais elevados da distribuição do spread isso implica em uma persistência dos spread maiores, o que afeta as rendas dos dealers e os custos de transação neste mercado. Para medir a dependência nestes quantiles condicionais, o quantilograma derivado por [Linton & Whang, 2007] é dado pela seguinte expressão:

$$(10) \quad \rho_{\alpha\kappa} = \frac{E[\psi_\alpha(y_t - \mu_\alpha)\psi_\alpha(y_{t+\kappa} - \mu_\alpha)]}{E[\psi_\alpha^2(y_t - \mu_\alpha)]}$$

o estimador amostral é dado por:

$$(11) \quad \hat{\rho}_{\alpha\kappa} = \frac{\sum_{t=1}^{T-\kappa} [\psi_\alpha(y_t - \hat{\mu}_\alpha)\psi_\alpha(y_{t+\kappa} - \hat{\mu}_\alpha)]}{\sqrt{\sum_{t=1}^{T-\kappa} [\psi_\alpha^2(y_t - \hat{\mu}_\alpha)]} \sqrt{\sum_{t=1}^{T-\kappa} [\psi_\alpha^2(y_{t+\kappa} - \hat{\mu}_\alpha)']}}$$

aonde o estimador para  $\hat{\mu}_\alpha$  é dado pelo quantile amostral, que é um estimador do processo:

$$(12) \quad \hat{\mu}_\alpha = \arg \min_{\mu \in R} \sum_{i=1}^T \rho_\alpha(y_i - \mu)$$



A hipótese nula de não previsibilidade direcional é dada por

$$(13) \quad H_0 E[\psi_\alpha(y_t - \mu_\alpha)\psi_\alpha(y_{t+k} - \mu_\alpha)] = 0$$

para todo quantile  $\alpha$ . A Figura 5 mostra o quantilograma estimado para os quantiles (0.01, 0.05, 0.10, 0.25, 0.50, 0.75, 0.90, 0.95, 0.99). O quantilograma estimado para estes quantiles mostra que o mesmo padrão de dependência autoregressiva encontrada na média se mantém nos quantiles, mas note que o padrão de dependência é diferente em cada quantile, mostrando uma correlação nos primeiros lags próximas a .4 para o quantile .01 e uma tendência de aumento da persistência nos quantiles mais altos; no quantile de 0.99 (os spreads mais altos) a persistência é próxima de .95. Esse efeito de assimetria mostra que os percentis mais baixos são caracterizados por uma baixa persistência e um reversão rápido ao quantile não condicional, enquanto que para os pontos aonde o spread está muito acima dos valores de equilíbrio (percentis acima de .90) existe uma grande persistência. Note que esse efeito de assimetria tem uma grande importância financeira, já que mostra que spreads elevados tendem a serem mais persistentes do que spreads baixos. Novamente podemos interpretar este efeito como uma resposta dos dealers a choques não antecipados, como situações de aumento de incerteza e maiores custos de manutenção de inventários de moeda.

**6.2. Autoregressão Quantílica.** O quantilograma nos indica que existem padrões diferenciados de dependência temporal para cada quantile condicional, mas não representa um modelo completo para este explicar este comportamento. Isto indica a necessidade de montar um modelo para a distribuição condicional do spread para cada quantile usando a estrutura autoregressiva, mas também controlando para os efeitos de volatilidade e durações esperadas, análogo ao modelo linear estimado para o spread. Uma ferramenta possível para este tipo de análise é o modelo de autoregressão quantílica ([Koenker & Xiao, 2006]), formulando o modelo de regressão quantílica utilizando como variáveis explicativas defasagens da própria variável dependente. Em uma regressão quantílica ([Koenker & Basset, 1978]) a função objetivo é formulada diretamente em função do quantile condicional de interesse  $\alpha$ , minimizando a função:

$$(14) \quad \min_{\beta \in R^p} \sum_{i=1}^n \rho_\alpha(y_i - x'_i \beta(\alpha))$$

que corresponde a uma função  $\rho_\alpha$  de perda condicional ao quantile  $\alpha$ , onde  $\alpha \in (0, 1)$ . Definimos a função perda como  $\rho_\alpha(u) = u(\alpha - I(u < 0))$ , onde  $I(\cdot)$  é uma função indicadora, e  $u$  é a diferença entre o valor observado  $y_i$  e o previsto por  $x'_i \beta(\alpha)$ . Estimadores para  $\hat{\beta}(\alpha)$  são obtidos pela minimização da função perda dada por 14, obtidos pelo valor esperado de  $\rho_\alpha(y_i - x'_i \beta(\alpha))$  em relação a cada  $\beta(\alpha)$ .

Note que o modelo regressão quantílica pode ser estendido para estruturas autoregressivas, através de uma regressão quantílica para o processo autoregressivo<sup>25</sup>:

$$(15) \quad Q_{y_t}(\alpha|y_{t-1}, \dots, y_{t-p}) = \beta_0(\alpha) + \beta_1(\alpha)y_{t-1} + \dots + \beta_p(\alpha)y_{t-p}$$

Nesta forma um modelo quantílico autoregressivo de primeira ordem (denominado QAR(1)) pode ser escrito como:

$$(16) \quad Q_{y_t}(\alpha|y_{t-1}) = \beta_0(\alpha) + \beta_1(\alpha)y_{t-1}$$

Note que podemos representar este modelo na forma  $y_t = \beta_0(U_t) + \beta_1(U_t)y_{t-1}$  e onde  $U_t$  é distribuído de forma uniforme entre (0,1). Nesta formulação o modelo autoregressivo AR(1) tradicional é obtido quando  $\beta_0(u) = \sigma\Phi^{-1}$  e  $\beta_1(u) = \beta_1$ . Para modelar a possível estrutura de assimetria na resposta dos desvios do spread nós formulamos o seguinte modelo:

<sup>25</sup>[Engle & Manganelli, 2004] utiliza uma formulação semelhante para o cálculo de value at risk condicional, e as regressões sobre o quantile neste artigo são análogas ao modelo de autoregressão quantílica.

TABELA 8. Regressões Quantílicas

$\alpha$	intercepto	Vol Forecast	Dur Forecast	$\phi_1$	$\phi_2$	$\phi_3$	$\phi_1+\phi_2+\phi_3$
0.01	-0.000305	-1.5471	1.7146e-007	0.33812	0.06216	0.08322	0.48350
dv. padrão	0.00048006	1.3000	0.0096843	0.25598	0.29264	0.14321	
t estat	-0.63534	1.1900	1.7705e-005	1.1554	0.24286	0.58113	
0.05	2.7483e-005	-1.9203	-7.2699e-006	0.47112	0.087408	0.094346	0.652878
dv. padrão	1.0019e-005	0.053672	2.9878e-007	0.0055919	0.0033144	0.0021805	
t estat	2.7430	-35.778	24.332	84.251	26.372	43.268	
0.10	0.00011128	-1.8572	7.2020e-006	0.55422	0.10253	0.11588	0.772628
dv. padrão	9.0151e-006	0.050679	4.1718e-007	0.0041372	0.0031668	0.0038092	
t estat	12.344	-36.645	17.264	133.96	32.378	30.421	
0.25	9.5605e-005	-0.95422	1.8677e-005	0.73684	0.10723	0.097129	0.941204
dv. padrão	4.9570e-006	0.024695	8.2370e-005	0.0035932	0.0024669	0.0020315	
t estat	19.287	38.639	0.22675	205.07	43.469	47.812	
0.50	-5.5749e-007	-0.0030460	5.8815e-006	0.92536	0.050826	0.019937	0.61295
dv. padrão	2.3941e-007	0.00067876	2.7470e-006	0.018844	0.013394	0.0056270	
t estat	-2.3286	-4.4877	2.1410	49.108	3.7948	3.5430	
0.75	-0.00011275	0.90265	1.2775e-005	0.79667	0.10759	0.088515	0.99277
dv. padrão	2.0740e-006	0.016636	4.3110e-006	0.0042570	0.0040564	0.0023968	
t estat	-54.363	54.260	2.9634	187.15	26.523	36.930	
0.90	-0.00031379	2.9543	3.3944e-005	0.65856	0.093343	0.20397	0.955868
dv. padrão	9.8263e-006	0.057821	2.0210e-005	0.0032553	0.0058786	0.0045187	
t estat	-31.934	51.093	1.6796	202.30	15.878	45.138	
.95	-0.00034529	3.8288	0.00012853	0.64144	0.086261	0.22052	0.948226
dv. padrão	2.1127e-005	0.12358	4.1586e-006	0.0061701	0.0087383	0.0066088	
t estat	-16.344	30.981	30.907	103.96	9.8717	33.368	
0.99	-0.00015489	5.1649	0.00047318	0.79704	0.087104	0.073036	0.957176
dv. padrão	0.00041945	5.1930	0.021389	0.73670	0.072878	0.54358	
t estat	-0.36928	0.99457	0.022123	1.0819	1.1952	0.13436	

(17)

$$Q_{y_t}(\alpha|y_{t-1}, \dots, y_{t-p}) = \beta_0(\tau) + \nu(\alpha)E[Vol|\mathcal{F}_{t-1}] + \delta(\alpha)E[Dur|\mathcal{F}_{t-1}] + \beta_1(\alpha)y_{t-1} + \dots\beta_p(\alpha)y_{t-p}$$

que é uma versão para o quantile condicional do modelo estimado na Tabela 7.

Estimamos<sup>26</sup> estes modelos para os mesmos quantiles (0.01, 0.05, 0.10, 0.25, 0.50, 0.75, 0.90, 0.95, 0.99) utilizados no quantilograma (Tabela 8). O modelo de regressão quantílica (Eq. 17) incorpora o efeito de assimetria mostrado pelo quantilograma - a persistência dos choques é mais baixa para os quantiles abaixo da mediana condicional e se torna mais alta para os quantiles acima da mediana condicional, ficando próximo de 1 nesses quantiles, evidenciando a resposta assimétrica dos desvios do spread. Um outro efeito interessante é que o sinal das parâmetros relacionados a volatilidade esperada - os sinais dos coeficientes volatilidade sempre são negativos para os quantiles abaixo da mediana e positivos para os percentis incluindo e acima da mediana, e neste modelo condicional os coeficientes relacionados a duração se tornam significativos e positivos para os quantiles acima da mediana.

Esse efeito de assimetria indica que a volatilidade e duração aumentam os spreads para os desvios do spread acima da mediana condicional, e esse efeito é amplificado nos quantiles mais altos. Podemos interpretar este efeito de assimetria como relação assimétrica do spread com a volatilidade e a duração - spreads acima do spread de equilíbrio tem persistência alta e são influenciados positivamente pela volatilidade e pelas durações esperadas, enquanto que spreads abaixo do spread

<sup>26</sup>Para referências sobre a estimação de modelos de regressão quantílica veja a monografia de [Koenker, 2005]. Nós utilizamos o método de rank inversion para o cálculo da matriz de variância-covariância dos parâmetros.

de equilíbrio são pouco persistentes e relacionados negativamente com as volatilidades e as durações esperadas<sup>27,28</sup>, indicando um mecanismo assimétrico de reversão ao spread de equilíbrio.

## 7. CONCLUSÕES

Neste artigo analisamos algumas propriedades empíricas relacionadas à microestrutura de mercado utilizando os dados de alta frequência de cotações para a série de dados de bid e ask para o mercado de câmbio R\$/US\$. O artigo obtém a replicação de alguns fatos estilizados observados na literatura internacional sobre microestrutura do mercado de câmbio para a série R\$/US\$ e introduz novas ferramentas na análise de efeitos de microestrutura empírica.

Entre os efeitos estudados neste artigo, observamos que a violação da propriedade Markoviana implica que neste mercado não existe incorporação imediata das novas informações nos preços, levando uma estrutura de longa dependência temporal nos retornos do bid e ask. Para capturar esta estrutura de longa dependência, formulamos um modelo de descoberta de preços utilizando um modelo vetorial de correção de erros, parametrizando este processo de incorporação de informações e obtendo uma estimativa para o spread de equilíbrio, que é interpretado na literatura de microestrutura como uma medida dos custos médios de provisão de liquidez e carregamento de estoques pelos dealers que operam neste mercado.

A modelagem do spread mostra que existe um mecanismo de resposta assimétrica nesta variável, aonde valores do spread acima e abaixo do valor do spread de equilíbrio reagem de forma distinta as informações passadas do próprios spread, volatilidade e durações. Valores do spread acima do valor de equilíbrio tem persistência alta e reagem de forma positiva e proporcional ao quantile em relação á volatilidade e a duração condicional esperadas, enquanto que encontramos uma relação inversa para os quantiles inferiores da distribuição dos spreads, com uma relação negativa dos spreads com as volatilidades e durações esperadas e uma baixa persistência, caracterizando uma estrutura não linear de reversão a média nos spreads.

Esta análise de assimetria, usando ferramentas como o quantilograma para a identificação da assimetria nos quantiles condicionais e a modelagem desta estrutura usando modelos de autoregressão quantílica, é original na literatura de microestrutura do mercado de câmbio. Estas evidências empíricas apontam para para novos fatos estilizados a serem incorporados nos modelos teóricos de microestrutura de mercado.

## REFERÊNCIAS

- [Admati & Pfleiderer, 1988] Admati, A., & Pfleiderer, P. 1988. A theory of intraday trading patterns: Volume and trading variability. *Review of Financial Studies*, **1**, 3–41.
- [Agmon & Barnea, 1977] Agmon, T., & Barnea, A. 1977. Transaction Costs and Marketability Services in the Eurocurrency Money Market. *Journal of Monetary Economics*, **3**, 359–366.
- [Bauwens & Giot, 2000] Bauwens, L., & Giot, P. 2000. The Logarithmic ACD Model: An Application to the Bid-Ask Quote Process of Three NYSE Stocks. *Annales d'Économie et de Statistique*, **60**, 117–149.
- [Bauwens *et al.*, 2002] Bauwens, L., Giot, P., Grammig, J., & Veredas, D. 2002. A comparison of financial duration models through density forecasts. *International Journal of Forecasting*, **20**, 589–609.
- [Bollerslev & Domowitz, 1993] Bollerslev, T., & Domowitz, I. 1993. Trading Patterns and Prices in the Interbank Foreign Exchange Market. *Journal of Finance*, **48**, 1421–1424.
- [Campbell *et al.*, 1997] Campbell, John Y, Lo, Andrew W, & MacKinlay, A Craig. 1997. *The econometrics of financial markets*. Princeton University Press.
- [Christie *et al.*, 1994] Christie, W., Harris, J, H., & Schultz, P. 1994. Why did NASDAQ market makers stop avoiding odd-eights quotes? *Journal of Finance*, **49**, 1841–1890.
- [Dufour & Engle, 2000] Dufour, A., & Engle, R. 2000. Time and the Price Impact of a Trade. *Journal of Finance*, **55**(6), 2467–2498.
- [Easley & O'Hara, 1987] Easley, D, & O'Hara, M. 1987. Price, Trade Size, and information in securities markets. *Journal of Financial Economics*, **19**, 69–90.
- [Engle, 2000] Engle, R. 2000. The Econometrics of Ultra High Frequency Data. *Econometrica*, **68**, 1–22.

<sup>27</sup>. Efeitos semelhantes de resposta assimétrica são obtidos pelo modelo de autoregressão quantílica para as séries de desemprego, preços gasolina no varejo e taxas de juros de curto prazo nos EUA ([Koenker & Xiao, 2006])

<sup>28</sup>Uma forma alternativa de especificar a modelagem de assimetria seria através de modelos de parâmetros variantes no tempo como modelos de *threshold autoregression*, modelos de parâmetros aleatórios ou então modelos de mudanças de regimes. Veja Cap 18 em [Zivot, 2006] para uma revisão desta literatura.

## MICROESTRUTURA EMPÍRICA DE MERCADO - UMA ANÁLISE PARA A TAXA DE CÂMBIO BRL/US\$ USA

- [Engle & Manganelli, 2004] Engle, R., & Manganelli. 2004. CAViaR: Conditional autoregressive value at risk by regression quantiles. *Journal of Business and Economic Statistics*, **22**(4), 367–381.
- [Engle & Russel, 1998] Engle, R., & Russel, J. 1998. Autoregressive Conditional Duration: A New Model for Irregularly Transaction Data. *Econometrica*, **66**, 1127–1162.
- [Falkenberry, 2002] Falkenberry, T. N. 2002. *High Frequency Data Filtering*. Tech. rept. Tick Data.
- [Fernandes & Amaro de Matos, 2007] Fernandes, M., & Amaro de Matos, J. 2007. Testing the Markov property with high frequency data. *Journal Of Econometrics*, **forthcoming**.
- [Fernandes & Grammig, 2005a] Fernandes, M., & Grammig, J. 2005a. A family of autoregressive conditional duration models. *Journal of Econometrics*.
- [Fernandes & Grammig, 2005b] Fernandes, M., & Grammig, J. 2005b. Nonparametric specification tests for conditional duration models. *Journal Of Econometrics*, **127**(1), 35–68.
- [Flood, 1994] Flood, M.D. 1994. Market Structure and Inefficiency in the Foreign Exchange Market. *Journal Of International Money and Finance*, **13**, 52–70.
- [Flood & Taylor, 1996] Flood, R. P., & Taylor, M. P. 1996. *The Microstructure of Exchange Rate Markets*. National Bureau of Economic Research. Chap. Exchange Rate Economics: What’s Wrong with the Conventional Macro Approach?, pages 261–294.
- [Frankel & Rose, 1995] Frankel, J. A., & Rose, A. K. 1995. *Handbook of International Economics*. North-Holland. Chap. Empirical Research in Nominal Exchange Rates, pages 1698–1729.
- [Frenkel *et al.*, 1996] Frenkel, J. A., Galli, G., & Giovannini, A. (eds). 1996. *The Microstructure of Foreign Exchange Market*. National Bureau of Economic Research.
- [Garcio & Urban, 2004] Garcio, M., & Urban, F. 2004. *O Mercado Interbancário de Câmbio no Brasil*. Working Paper não publicado.
- [Glassman, 1987] Glassman, D. 1987. Exchange rate risk and transactions costs: evidence from bid-ask spreads. *Journal Of International Money and Finance*, **6**, 481–490.
- [Glosten & Milgrom, 1985] Glosten, L., & Milgrom, P. 1985. Bid, ask and transaction prices in a specialist market with heterogenously informed traders. *Journal of Financial Economics*, **14**(Mar.), 71–100.
- [Goodhart, 1989] Goodhart, C. 1989. News and the Foreign Exchange Market. *In: Manchester Statistical Society*.
- [Green & Silverman, 1994] Green, P. J., & Silverman, B. W. 1994. *Nonparametric Regression and Generalized Linear Models: A Roughness Penalty Approach*. Chapman and Hall.
- [Hasbrouck, 1991] Hasbrouck, J. 1991. Measuring the information content of stock trades. *Journal of Finance*, **46**, 179–207.
- [Hasbrouck, 1992] Hasbrouck, J. 1992. *Using the TORQ Database*. Tech. rept. NYSE Working Paper 92-05.
- [Hasbrouck, 1999] Hasbrouck, J. 1999. Security Bid/Ask Dynamics with Discreteness and Clustering. *Journal of Financial Markets*, **2**, 1–28.
- [Hasbrouck, 2007] Hasbrouck, J. 2007. *Empirical Market Microstructure*. Oxford Univerirsity Press.
- [Hasbrouck, 1988] Hasbrouck, Joel. 1988. Trades, Quotes, Inventories and Information. *Journal of Financial Economics*, **22**, 229–252.
- [Hasbrouck, 1996] Hasbrouck, Joel. 1996. Modelling Market Microstructure Time Series. *Chap. 22 of: Maddala, G S, & Rao, C R (eds), Handbook of Statistics*, vol. 14 Statistical Methods in Finance. North-Holland.
- [Jorion, 1996] Jorion, J. 1996. *The Microstructure of Foreign Exchange Markets*. National Bureau of Economic Research. Chap. Risk and Turnover in The Foreign Exchange Market, pages 19–40.
- [Koenker, 2005] Koenker, R. 2005. *Quantile Regression*. Cambridge University Press.
- [Koenker & Basset, 1978] Koenker, R., & Basset, G. 1978. Regression Quantiles. *Econometrica*, **46**, 33–50.
- [Koenker & Xiao, 2006] Koenker, R., & Xiao, Z. 2006. Quantile Autoregression. *Journal of the American Statistical Association*, **475**, 980–1006.
- [Linton & Whang, 2007] Linton, O., & Whang, J. 2007. A Quantilogram Approach to Evaluating Directional Predictability. *Journal Of Econometrics*, **forthcoming**.
- [Lyons, 1996] Lyons, R. 1996. *The Microstructure of Foreign Exchange Markets*. National Bureau of Economic Research. Chap. Foreign Exchange Volume: Sound and Fury Signifying Nothing?, pages 183–208.
- [Lyons, 2001] Lyons, Richard K. 2001. *The Microstructure Approach to Exchange Rates*. MIT Press.
- [McGroarty *et al.*, 2006] McGroarty, F., Gwiylm, O., & Thomas, S.H. 2006. Microstructure effects, bid-ask spreads and volatility in the spot foreign exchange market pre and post-EMU. *Global Finance Journal*, **17:1**, 23–49.
- [O’Hara, 1995] O’Hara, M. 1995. *Market Microstrcuture Theory*. Blackwell.
- [Sarno & Taylor, 2002] Sarno, L., & Taylor, M. 2002. *The Economics of Exchange Rates*. Cambridge University Pres.
- [Taylor, 1995] Taylor, M. P. 1995. The Economics of Exchange Rates. *Journal of Economic Literature*, **83**, 19–47.
- [Wu, 2007] Wu, T. 2007. *Order Flow in the South: Anatomy of the Brazilian FX Market*. University of California, Santa Cruz, Unpublished Working Paper.
- [Zivot, 2006] Zivot, E. 2006. *Modeling Financial Time Series with S-PLUS, second edition*. Springer-Verlag.
- [Zivot & Yan, 2003] Zivot, E., & Yan, B. 2003. *"Analysis of High-Frequency Financial Data with S-PLUS"*