



imp^a Instituto Nacional de Matemática Pura e Aplicada

Toxicidade no Mercado Brasileiro

Autor: **Diego Andrés de Barros Lima Barbosa**

Orientador:
Milton Jara

Rio de Janeiro
Abril de 2014

Dedico este trabalho aos meus pais

Agradecimentos

Agradeço a todos aqueles que me ajudaram e contribuíram para o desenvolvimento deste trabalho.

Faço menção especial ao meu pai, Marcos, e à minha mãe, Marialzira, que sempre me estimularam a estudar e me ensinaram todos os valores que acredito hoje. Sou muito grato a eles por tudo. Agradeço também ao meu irmão, Marcos, às minhas duas irmãs, Marília e Mariana, e à minha namorada, Bianca, por sempre me ajudarem a evoluir pessoalmente, academicamente e profissionalmente.

Agradeço ao professor Jorge Zubelli por todo o esforço e dedicação tanto por este trabalho, quanto pela organização deste curso de mestrado. Também agradeço aos professores Milton Jara e Luca Mertens por todo apoio e disposição no esclarecimentos de dúvidas.

Finalmente agradeço a todos os professores e monitores que tive no IMPA.

Resumo

A teoria de microestrutura de mercado divide os negociadores, de acordo com o seu nível de informação, em dois grupos: negociadores informados e negociadores não-informados. Um mercado é definido como tóxico, quando existe uma alta probabilidade dos provedores de liquidez estarem transacionando com negociadores informados. Um negócio é dito informado, quando um dos negociadores dele é informado. Nesse contexto, esse trabalho explana duas formas para mensurar a toxicidade de mercado: a probabilidade de o negócio ser informado (PIN) e a probabilidade de o negócio ser informado sincronizado por volume (VPIN). Foram utilizados diferentes tipos de classificações de negócios para calcular o VPIN no mercado futuro de índice bovespa, no mercado futuro de taxa de câmbio real-dólar e no mercado futuro de taxa DI de um dia. Nos momentos em que o VPIN atinge valores históricos máximos, foi feita uma análise qualitativa e quantitativa do mercado associando cada evento desse tipo a algum fato relevante do mercado. No mercado futuro de taxa de câmbio real-dólar, foi mostrado evidências de um aumento de toxicidade no período em que o governo passou a intervir diariamente no câmbio.

Palavras Chave: Microestrutura de mercado, toxicidade, liquidez, PIN, VPIN, classificação de negócios.

Abstract

The market microstructure theory divides negotiators, according to their level of information, into two groups : informed and non-informed traders. A market is defined as toxic when there is a high probability of liquidity providers negotiate with informed traders. A trade is said to be informed when one of negotiators are informed. In this context, this work explains two models to measure toxicity in the market: the probability of informed trading (PIN) and the volume synchronized probability of informed trading (VPIN). Trade classification models was used to calculate the VPIN in bovespa index future, real-dollar currency future and one day DI interest rate future. When the VPIN reaches the historical high level, a qualitative and quantitative analysis of the market was taken by associating each such event to any relevant fact in the market. In real-dollar future contract, was shown to increase toxicity in the period when the government started to intervene daily.

Key words: Market microstructure, toxicity, liquidity, PIN, VPIN, trade classification.

Sumário

Índice	viii
1 Introdução	1
2 Revisão Bibliográfica	3
2.1 Classificação de Negócios	3
2.2 Medidas de Precisão	4
3 Medidas de Toxicidade	7
3.1 Probabilidade de Negócio Informado - PIN	7
3.2 Probabilidade de Negócio Informado Sincronizado Por Volume - VPIN	10
4 Considerações Gerais e Metodologia	13
4.1 Tratamento de Dados	13
4.2 Mercados Analisados	13
4.3 Período de Amostragem	15
5 Considerações Gerais e Metodologia	17
5.1 Tratamento de Dados	17
5.2 Mercados Analisados	17
5.3 Período de Amostragem	19
6 Toxicidade no Brasil	21
6.1 Dólar-Real	21
6.2 Índice Bovespa	26
6.3 Taxa Juros DI	30
7 Toxicidade em Eventos	35
7.1 Programa de Leilões de Swap Cambial - Banco Central do Brasil	35
8 Considerações Finais	37
8.1 Síntese Final	37
8.2 Conclusões	37
A Resultados Técnicos e Notações	41
A.1 Códigos Classificação de Negócios e Cálculo do VPIN	41
Referências Bibliográficas	57

Capítulo 1

Introdução

Sob o ponto de vista econômico, o preço de equilíbrio de um bem é o preço do encontro de sua curva de oferta e de sua curva de demanda. Ele representa o equilíbrio dos interesses dos agentes do mercado. A partir disso duas perguntas interessantes podem ser feitas: Qual é a eficiência do mecanismo para um agente expressar suas intenções de preço? De que forma esse preço de equilíbrio foi alcançado? A teoria da microestrutura de mercado tenta propor modelos que respondam essas e outras perguntas relacionadas, mais precisamente, ela está interessada em estudar como o mecanismo de negociação afeta o processo de formação do preço [21].

De acordo com o nível de informação disponível para os negociadores do mercado, classificaremos eles como: *negociadores informados* e *negociadores não-informados*. Um negociador informado é aquele que possui uma informação relevante sobre o ativo negociado, fazendo seus negócios, na média, causarem perdas em suas contrapartes (nas pessoas que fornecem liquidez a ele). O negociador informado possui uma vantagem em comparação ao não-informado. Essa classificação é essencial para os modelos de microestrutura de mercado, conforme pode ser visto em [20] e em [21]. Dessa forma definiremos um negócio como *negócio informado*, se um dos negociadores envolvidos é informado. Vale notar que na definição de negociador informado, não necessariamente essa vantagem possuída tem relação com algo ilícito, pode ser algo como uma detecção de um padrão, modelo proprietário de precificação ou qualquer outro tipo de vantagem. Quando existir num mercado uma grande proporção dos negócios sendo informados, o chamaremos de *mercado tóxico*.

O trabalho a ser apresentado é da área de microestrutura de mercado e explanará algumas medidas e metodologias para se medir a toxicidade do mercado. Será estudado o momento em que o mercado fica altamente tóxico, pois a alta proporção de negócios informados fará os provedores de liquidez sair momentaneamente do mercado devido aos seguidos prejuízos. Esse movimento pode causar uma queda imediata da liquidez do mercado e ocasionar um aumento na volatilidade de curto prazo. Além disso, serão feitas análises de como a toxicidade se comportou antes, durante e depois de grandes eventos relevantes.

Capítulo 2

Revisão Bibliográfica

2.1 Classificação de Negócios

Os negócios realizados no mercado unem duas partes (de compra e de venda) que negociam num específico preço. Uma das partes, que age de maneira ativa, está retirando a liquidez do mercado e a outra parte, que age de maneira passiva, está fornecendo liquidez ao mercado.

A parte ativa, que fez acontecer o negócio naquele específico momento, é chamado de *agressor do negócio*. Quando o agressor do negócio for um comprador (vendedor), chamaremos o negócio de *agressão de compra (venda)*. Os negócios ocorridos no leilão de abertura ou leilão de fechamento do mercado não possuem agressores e serão chamados de *negócios do leilão*. Além disso, serão denominados *negócios do balcão*, os negócios que são registrados no histórico de negociação do ativo, mas que não foram realizados por ordens que entraram a mercado (negócios realizados no mercado de balcão).

Os modelos de toxicidade, que serão estudados nesse trabalho, usarão a classificação de negócios como ferramenta primordial em seus algoritmos. Por conta disso, uma atenção especial será dado posteriormente aos negócios do leilão e aos negócios do balcão, pois dependendo de sua proporção no mercado em estudo, ele poderá causar grandes distorções nas medidas de toxicidade. Nesta seção concentraremos em alguns métodos para tentar identificar o agressor de um negócio.

Classificação da Bolsa

BM&F Bovespa disponibiliza dados de negociação *tick-by-tick* com uma precisão de milissegundos no seu *ftp* [8]. Os dados a partir de novembro de 2011 possuem gravados o lado agressor. Assim, nós temos uma classificação exata dada pela bolsa brasileira. Essa classificação é chamada de *Exchange Classification* (Classificação da Bolsa) ou abreviadamente EC, que será a sigla utilizada no durante o trabalho.

A Regra do Tick

Essa regra classifica um negócio como agressão de compra, se o seu preço foi maior do que o preço do último negócio (negócio de tick para cima). Analogamente, classificamos um negócio como agressão de venda, se o seu preço foi menor do que o preço do último negócio (negócio de tick para baixo). Se o preço do negócio for o mesmo valor do preço do último negócio, iremos olhar o o negócio mais próximo já realizado que não seja nesse mesmo preço. Caso esse preço anterior seja maior (menor), classificaremos como agressão de compra (venda).

Formalmente, seja $(P_i)_{i \in \mathbb{N}}$ a série de preços dos negócios ocorridos. Para cada $i \in \mathbb{N}$, seja $j = \sup\{k \in \mathbb{N}; k < i \wedge P_k \neq P_i\}$.

- $P_i > P_j$: Classificamos o negócio como agressão de compra
- $P_i < P_j$: Classificamos o negócio como agressão de venda

A regra do tick será referenciada pela sigla TR (letras iniciais de *Tick Rule*), conforme [13].

Classificação Maciça por Volume

No artigo [16] foi introduzido o conceito de fazer uma classificação contínua dos negócios. Para isso, iremos amostrar os negócios em barras iguais de tempo que denotaremos por $(I_i)_{i \in \mathbb{N}}$, onde chamaremos P_i como o preço de fechamento da barra I_i . O volume negociado em cada barra I_i será chamado por V_i . Da série das diferenças, $\Delta P_i = P_i - P_{i-1}$, calcularemos $\sigma_{\Delta P}$ como sendo a estimação de seu desvio padrão. Sendo Z a função distribuição acumulada da normal padrão, definiremos:

- Volume de agressão de compra:

$$V_i^c = Z\left(\frac{\Delta P_i}{\sigma_{\Delta P}}\right)V_i$$

- Volume de agressão de venda :

$$V_i^v = \left(1 - Z\left(\frac{\Delta P_i}{\sigma_{\Delta P}}\right)\right)V_i$$

Como pode ser algebricamente notado: $V_i = V_i^c + V_i^v$. Fracionamos o volume de uma barra em dois termos (volume de agressão de compra e agressão de venda). Variações e detalhes desse tipo de classificação pode ser visto em [13] e [16]. Esta classificação é chamada de *Bulk Volume Classification* (Classificação Maciça por Volume) ou abreviadamente BVC, que será a sigla utilizada no presente trabalho.

2.2 Medidas de Precisão

Tendo em vista que a bolsa brasileira fornece a classificação de negócios, vamos considerar essa classificação como a exata (EC com precisão de 100%). Iremos considerar a letra grega η para representar a precisão de uma classificação calculada pela comparação negócio por negócio dado por uma específica classificação e a dada por EC. Usaremos a letra $\underline{\eta}$ para representar a precisão de uma classificação calculada pela a fração de volume classificado correto no interior das barras de volume.

Usaremos a seguinte denotação: dado uma série *tick-by-tick*, cada negócio i terá o preço P_i e quantidade Q_i . Os negócios serão amostrados em barras iguais de tempo que denotaremos pela série I_j , onde chamaremos V_j o volume da barra I_j . Cada V_j foi dividido pela classificação EC por volume de agressão de compra V_j^c e volume por agressão de venda V_j^v , onde $V_j = V_j^c + V_j^v$. Para qualquer detalhe e esclarecimento sobre precisão de uma classificação de negócio, consultar [14].

Precisão TR

$$\eta^{TR} = \frac{\sum_{i=1}^N (Q_i \mathbf{1}_i)}{\sum_{i=1}^N Q_i}$$

, onde a função $\mathbf{1}_i$ (função característica) representa 1 quando o negócio foi classificado corretamente e 0 quando o negócio foi classificado de forma errada e N representa o número total de negócios analisados.

Para a precisão da Regra do Tick nos mercados futuros de dolar, índice bovespa e DI de um dia; conseguimos achar as seguintes precisões:

Mercado	DOL	IND	DI
η^{TR}	79.57%	75.84%	80.50%

Precisão BVC

Consideraremos que o V_i foi dividido pela classificação BVC nos seguintes volumes \underline{V}_i^c e \underline{V}_i^v , cada um representando a soma das frações das quantidades dos negócios do intervalo que foram classificados como agressão de compra e de venda, respectivamente. Assim, podemos definir:

$$\eta^{BVC} = \frac{\sum_{j=1}^K (\min\{\underline{V}_j^c, V_j^c\} + \min\{\underline{V}_j^v, V_j^v\})}{\sum_{j=1}^K V_j}$$

, onde K representa o número total de barras de tempo analisado.

Para a precisão BVC nos mercados futuros de dolar, índice bovespa e DI de um dia; conseguimos achar as seguintes precisões:

Mercado	DOL	IND	DI
η^{BVC}	62.34%	71.68%	55.86%

Capítulo 3

Medidas de Toxicidade

Para mensurar a toxicidade do mercado, serão apresentados os dois principais modelos existentes na teoria: Probabilidade de Negócio Informado (PIN) e Probabilidade de Negócio Informado Sincronizado Por Volume (VPIN).

3.1 Probabilidade de Negócio Informado - PIN

Modelo

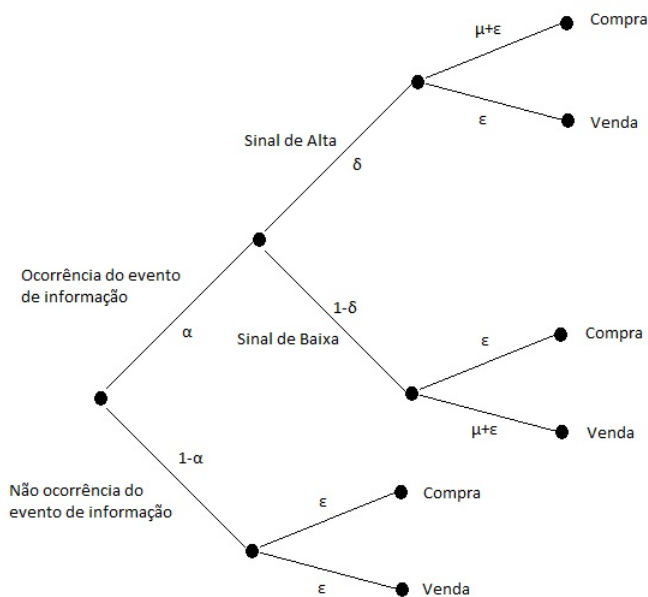


Figura 3.1: Modelo PIN

A sigla PIN vem da abreviação das palavras *probability of informed trading*, que é a probabilidade de negociação informada. Vimos no Capítulo 1 que a definição de mercado tóxico está associada a proporção de negociadores informados e, por isso, o PIN é uma medida de toxicidade do mercado.

Dividiremos o tempo em períodos de negociação enumerados. Assim, denotaremos por $(P_i)_{i \in \mathbb{N}}$ o preço de um ativo no fim de cada período de negociação i (Esses períodos podem ser

segundos, minutos, dias ou qualquer outra medida de tempo). O tempo no interior de cada um desses períodos é uma variável contínua e seu intervalo será indexado por $t \in (a_i, b_i)$.

O modelo considera, antes de iniciar cada período de negócio, a possibilidade de ocorrer um evento relevante sobre o ativo naquele período. Esses eventos são considerados independentemente distribuídos e ocorrem com probabilidade α_t . É considerado também que cada fato relevante tem uma probabilidade δ_t de fazer o ativo subir de preço e tem probabilidade $1 - \delta_t$ de fazer o ativo cair de preço. Condicionado a ocorrer um fato relevante no período i , P_i^+ será chamado o preço do ativo no fim do período caso a notícia seja um sinal de alta, e de modo análogo, P_i^- será chamado o preço do ativo no fim do período caso a notícia seja um sinal de baixa. Caso não ocorra fatos relevantes, denotaremos o preço do ativo no fim do período por P_i^* . Considera-se $P_i^+ > P_i^* > P_i^-$, $\forall i \in \mathbb{N}$.

Por facilidade de notação, denotaremos $1 - \alpha_t$ como $p_n(t)$ (a letra "n" representa que nenhuma informação ocorreu), $\alpha_t \delta_t$ como $p_a(t)$ (a letra "a" representa uma alta no preço do ativo) e $\alpha_t(1 - \delta_t)$ como $p_b(t)$ (a letra "b" representa uma baixa no preço do ativo). Assim, para cada período i e com $t \in (a_i, b_i)$ temos:

$$\mathbb{E}[P_i|t] = p_n(t)P_i^* + p_a(t)P_i^+ + p_b(t)P_i^- \quad (3.1.1)$$

Conforme mostrado no Capítulo 1, os negociadores são divididos em duas classes: os informados e os não informados. Será assumido que a chegada de ordens de compra e de venda para um negociador não informado são independentes e seguem um processo de Poisson de taxa ϵ e que a chegada de ordens de compra e de venda para um negociador informado são independentes e seguem um processo de Poisson de taxa μ . A Figura 3.1 representa uma árvore que sintetiza o que é assumido acima para o modelo PIN.

Vamos considerar o vetor tri-dimensional $p(t) = (p_n(t), p_b(t), p_a(t))$, com seus componentes representando a probabilidade a priori de não ocorrer notícia, de ocorrer notícia de baixa e de ocorrer notícia de alta, respectivamente.

Denotaremos v_t (c_t) o evento de chegar uma ordem de venda (compra) no tempo t e considere a probabilidade condicional $p(t|v_t)$ como a atualização das componentes de probabilidade dado que chega uma ordem de venda no tempo t (dado v_t). Assim, Pela regra de Bayes ($P(A|B) = \frac{P(B|A)P(A)}{P(B)}$) temos:

$$p(t|v_t) = (p_n(t|v_t), p_b(t|v_t), p_a(t|v_t))$$

\Rightarrow

$$p_n(t|v_t) = \frac{p(v_t | \text{"Sem - Notcia"}) p_n(t)}{p(v_t)}$$

$$p_b(t|v_t) = \frac{p(v_t | \text{"Notcia - Baixa"}) p_b(t)}{p(v_t)}$$

$$p_a(t|v_t) = \frac{p(v_t | \text{"Notcia - Alta"}) p_a(t)}{p(v_t)}$$

\Rightarrow

$$p(t|v_t) = \left(\frac{\epsilon p_n(t)}{\epsilon + p_b(t)\mu}, \frac{(\epsilon + \mu)p_b(t)}{\epsilon + p_b(t)\mu}, \frac{\epsilon p_a(t)}{\epsilon + p_b(t)\mu} \right) \quad (3.1.2)$$

Portanto, para qualquer tempo $t \in (a_i, b_i)$ o preço justo para o valor de uma oferta de compra livre de arbitragem será calcular o $\mathbb{E}[P_i|v_t]$:

$$OC(t) = \mathbb{E}[P_i|v_t] = p_n(t|v_t)P_i^* + p_b(t|v_t)P_i^- + p_a(t|v_t)P_i^+ = \frac{p_n(t)\epsilon + p_b(t)(\epsilon + \mu) + p_a(t)\epsilon}{\epsilon + p_b(t)\mu} \quad (3.1.3)$$

De forma análoga, o preço justo para o valor de uma oferta de compra livre de arbitragem será calcular o $\mathbb{E}[P_i|c_t]$:

$$OV(t) = \mathbb{E}[P_i|c_t] = p_n(t|c_t)P_i^* + p_b(t|c_t)P_i^- + p_a(t|c_t)P_i^+ = \frac{p_n(t)\epsilon + p_b(t)\epsilon + p_a(t)(\epsilon + \mu)}{\epsilon + p_a(t)\mu} \quad (3.1.4)$$

Subtraindo a Equação (3.1.4) da Equação (3.1.3) e substituindo a Equação (3.1.1), chega-se na seguinte equação para o valor esperado do *spread* (melhor oferta de venda menos melhor oferta de compra) de um ativo num tempo t :

$$Spread(t) = OV(t) - OC(t) = \frac{\mu p_a(t)}{\epsilon + \mu p_a(t)}(P_i^+ - \mathbb{E}[P_i|t]) + \frac{\mu p_b(t)}{\epsilon + \mu p_b(t)}(\mathbb{E}[P_i|t] - P_i^-) \quad (3.1.5)$$

Nota-se que o valor do *spread* decresce com o aumento da taxa de negociadores não informados (ϵ), atingindo o valor máximo: $spread_{max} = P_i^+ - P_i^-$, para $\epsilon = 0$. Também deve-se notar que o *spread* cresce com o aumento da taxa de negociadores informados (μ), atingindo o valor mínimo: $spread_{min} = 0$, para $\mu = 0$. Quando a probabilidade de chegar um sinal de alta é igual a probabilidade de chegar um sinal de baixa, temos que:

$$1 - \delta_t = \delta_t \Rightarrow \delta_t = 1/2 \Rightarrow p_a(t) = \alpha_t(1 - 1/2) = \frac{\alpha_t}{2} = \frac{1 - p_n(t)}{2} \quad (3.1.6)$$

Substituindo a Equação (3.1.6) na Equação (3.1.5):

$$Spread(t) = \frac{p_a(t)\mu}{p_a(t)\mu + \epsilon}(P_i^+ - P_i^-) = \frac{(1 - p_n(t))\mu}{(1 - p_n(t))\mu + 2\epsilon}(P_i^+ - P_i^-) \quad (3.1.7)$$

A partir da Equação (3.1.7), naturalmente será definido o PIN como o primeiro termo da multiplicação:

$$PIN(t) = \frac{(1 - p_n(t))\mu}{(1 - p_n(t))\mu + 2\epsilon} = \frac{\alpha_t\mu}{\alpha_t\mu + 2\epsilon} \quad (3.1.8)$$

Como pode ser observado, $PIN(t) \in [0, 1]$. Esse termo fará o *spread* diminuir ou aumentar de acordo com a proporção de negociadores informados no mercado. Para uma alta probabilidade de negócio informado (PIN alto), teremos pela Equação (3.1.8) um *spread* livre de arbitragem grande e limitado a $P_i^+ - P_i^-$ pelo modelo. Esse comportamento mostrado faz sentido e concorda com o racional do mercado, pois no momento em que o mercado estiver com um PIN alto, ninguém vai querer dar liquidez e isso ocasionará o aumento do *spread*. Da mesma forma que quando ocorrer um PIN baixo, os provedores de liquidez vão querer dar liquidez aquele mercado, ocasionando um menor *spread*.

Estimação

Estimar o vetor $\theta = (\alpha, \delta, \epsilon, \mu)$ é um problema difícil, pois os parâmetros não são diretamente observáveis. Olhando os negócios, conseguimos observar a chegada de ordens de compra e de

venda, mas não conseguimos observar quais negociadores são informados e quais são não informados. Por conta disso, precisamos de um modelo que extrairá a informação dessas variáveis que são observáveis (chegada de ordens de compra e venda). Conforme o modelo do PIN: períodos com notícias de alta ocorrerão mais compras (processo de Poisson com taxa $\mu + \epsilon$), períodos com notícias de baixa ocorrerão mais vendas (processo de Poisson com taxa $\mu + \epsilon$) e períodos sem notícias teremos as ordens de compra e de venda como um processo de Poisson com taxa ϵ .

Considere um período com notícia de baixa, assim as ordens de venda serão um processo de Poisson com taxa $\mu + \epsilon$ e as ordens de compra serão um processo de Poisson com taxa ϵ . Considerando uma amostragem sequencial de negócios, interna ao período em análise, com duração T e que tenha no total C compras e V vendas, temos que a função de máxima verossimilhança será:

$$L_1(\theta|(C, V)) = e^{-\epsilon T} \frac{(\epsilon T)^C}{C!} e^{-(\epsilon + \mu)T} \frac{((\epsilon + \mu)T)^V}{V!} \quad (3.1.9)$$

De modo análogo em um período de notícia de alta:

$$L_2(\theta|(C, V)) = e^{-\epsilon T} \frac{(\epsilon T)^V}{V!} e^{-(\epsilon + \mu)T} \frac{((\epsilon + \mu)T)^C}{C!} \quad (3.1.10)$$

Da mesma forma num período em que não ocorrer notícia temos:

$$L_3(\theta|(C, V)) = (e^{-\epsilon T} \frac{(\epsilon T)^C}{C!} e^{-(\epsilon)T} \frac{((\epsilon)T)^V}{V!}) \quad (3.1.11)$$

Assim a função máxima verossimilhança observando C e V em um dia desconhecido será as equações (3.1.9), (3.1.10) e (3.1.11) multiplicadas pelas suas respectivas probabilidades:

$$L(\theta|(C, V)) = \alpha(1 - \delta)L_1(\theta|(C, V)) + \alpha\delta L_2(\theta|(C, V)) + (1 - \alpha)L_3(\theta|(C, V)) \quad (3.1.12)$$

Com apenas 1 período de estimação, o estimador $\bar{\theta}$ calculado teria valores de α e δ sendo apenas 0 ou 1. Assim, dado K dias com parâmetros $M = (C_i, V_i), \forall i \in \{1, 2, \dots, k\}$, temos que maximizar a função:

$$L(\theta|M) = \prod_{i=1}^K L(\theta|(C_i, V_i)) \quad (3.1.13)$$

Para uma análise mais ampla do problema de otimização associado a (3.1.13), ver as referências [19] e [15].

3.2 Probabilidade de Negócio Informado Sincronizado Por Volume - VPIN

Modelo

A sigla VPIN vem da abreviação das palavras *v*olume *s*ynchronized *p*robability of *i*nformed *t*rading, que é a probabilidade de negociação informada sincronizada pelo volume. Para cada período i , será denotado o volume das agressões de venda por V_i^v e o volume das agressões de compra por V_i^c . Os períodos são definidos como o tempo necessário de ser negociado um

determinado volume V_i (valor conhecido a priori do período), definido antes do início do período de negociação. Observe que os períodos não necessariamente possuem a mesma duração de tempo.

Podemos agora dar a definição de VPIN [16]:

$$VPIN = \mathbb{E}\left[\frac{|V_i^v - V_i^c|}{V_i}\right] \quad (3.2.1)$$

Foi mostrado na referência [18] que temos a seguinte aproximação:

$$\mathbb{E}[|V_i^v - V_i^c|] \approx \alpha\mu \quad (3.2.2)$$

Observe que o lado direito da aproximação acima é exatamente o numerador da formula do PIN em (3.1.7). Assim, em [16] mostrou-se:

$$VPIN \approx \frac{\alpha\mu}{\alpha\mu + 2\epsilon} = PIN \quad (3.2.3)$$

Estimação

A sincronia com o volume permite que os negociadores consigam atualizar o valor do VPIN imediatamente após o negócio que completa o valor do volume esperado e, além disso, trabalhar com buckets de volume faz com que se recupere parcialmente a normalidade dos retornos. Uma maior abordagem sobre consequências da mudança da variável de tempo pela variável volume pode ser encontrada em [17].

Em todo o trabalho será assumido, para o cálculo do número de contratos de cada bucket de volume (V), o valor esperado de 50 buckets para um dia. Assim, V será exatamente o volume estimado para aquele dia dividido por 50 ($m = 50$). Vale ressaltar que para cada dia teremos um V . Na estimação do volume diário, usamos o média de volume dos 10 dias anteriores ($q = 10$).

Assumindo que estamos querendo calcular o valor do volume diário, $Volume_K^{Dia}$, e o volume do bucket, V_K , do dia K ($K > q$), temos:

$$Volume_K^{Dia} = \frac{\sum_{i=K-q}^{K-1} Volume_i^{Dia}}{q} = \frac{\sum_{i=K-10}^{K-1} Volume_i}{10} \quad (3.2.4)$$

$$V_K = Volume_K^{Dia}/m = Volume_K^{Dia}/50; \quad (3.2.5)$$

Usaremos em todo calculo de VPIN a média das últimas 50 amostras ($n = 50$) do módulo do desbalancamento entre agressões de compra e de venda dividido pelo o volume do bucket (Equação (3.2.1)). Dessa forma, para um bucket j ($j > n$) teremos sempre que o VPIN será:

$$VPIN_j = \mathbb{E}\left[\frac{|V_j^v - V_j^c|}{V_j}\right] = \frac{\sum_{i=j-n}^{j-1} \frac{|V_i^v - V_i^c|}{V_i}}{n} = \frac{\sum_{i=j-50}^{j-1} \frac{|V_i^v - V_i^c|}{V}}{50} \quad (3.2.6)$$

Sintetizando acima: $(q, m, n) = (10, 50, 50)$. Esses valores foram escolhidos com base nas análises colocadas em [22] e [16].

Exemplo Formador de Mercado

O problema matemático formal do formador de mercado pode ser encontrado na referência [23]. Em linhas gerais, um formador de mercado é um negociador que simultaneamente coloca ordens limitadas de compra e de venda (ele sempre é um provedor de liquidez). Seu objetivo é fazer diversos negócios administrando seu inventário e controlando o risco de mercado para ganhar, na média, o *spread* do ativo. O formador de mercado vai querer exatamente evitar o negociador informado, pois o valor esperado desse tipo de negócio é negativo. A proximidade do preço de suas ordens de compra (venda) com o preço da melhor ordem de compra (venda) pode ser controlado através de alguma métrica que use a variável VPIN como uma de suas variáveis de controle.

Capítulo 4

Considerações Gerais e Metodologia

4.1 Tratamento de Dados

Todos os dados presente nesse trabalho foram obtidos diretamente do *ftp* da bolsa brasileira BM&F Bovespa [8]. Os dados se encontram gravados em arquivos de texto (ver figura 5.2) num formato próprio da bolsa (ver Figura 5.1). Foi necessário tratar esses arquivos brutos seguindo seus padrões e criar séries históricas *tick-by-tick* de cada ativo.

Identificação	1	2	RH Cabeçalho - RT Rodapé
Nome Arquivo	4	20	
Data Inicial	25	10	Data Inicial desse arquivo
Data Final	36	10	Data Final desse arquivo
Total de Linhas	47	9	No caso do rodapé apresenta o total de linhas do arquivo
Detalhes			

Coluna	Posição Inicial	Tamanho	Descrição

Data Sessão	1	10	Data de sessão
Símbolo do Instrumento	12	50	Símbolo do Instrumento para Código de Bolsa "BMF"
Nr.Negócio	63	10	Número do negócio
Preço Negócio	74	20	Preço do negócio
Quantidade	95	18	Quantidade negociada
Hora	114	15	Horário da negociação (formato HH:MM:SS.NNNNNN)
Ind.Anulação	130	1	Indicador de Anulação
Data Oferta Compra	132	10	Data da oferta de compra
Seq.Oferta Compra	143	10	Número sequencial da oferta de compra
GenerationID - Of.Compra	154	10	Número de geração (GenerationID) da Oferta de Compra.
Condição Oferta de Compra	165	1	Código que identifica a condição da oferta de compra.
Data Oferta Venda	167	10	Data da oferta de venda
Seq.Oferta Venda	178	10	Número sequencial da oferta de venda
GenerationID - Of.Venda	189	10	Número de geração (GenerationID) da Oferta de Venda.
Condição Oferta de Venda	200	1	Código que identifica a condição da oferta de venda.
Obs: Delimitador das colunas de detalhes ';'			

Figura 4.1: Exemplo de formato de arquivo da bolsa.

4.2 Mercados Analisados

O estudo é focado no mercado de contratos futuros. Os movimentos nesse mercado são normalmente mais relacionados a riscos sistemáticos ou a mudanças em variáveis de cobertura de risco [16]. Dessa forma, focaremos nos principais futuros negociados no Brasil: futuros de dólar comercial, de índice Bovespa e de taxa DI de um dia. Na Figura 5.3 temos o volume mensal dos principais contratos futuros da BM&F Bovespa [11].


```

2012-10-18;DOLX12 ;0000010; 000000002037.000000;00000000000000000005;09:00:43.641;1;2012-10-18;7449758433;0000498414;2;2012-10-18;7449758
2012-10-18;DOLX12 ;0000070; 000000002037.000000;0000000000000000010;09:00:45.899;1;2012-10-18;7449758455;0000498464;1;2012-10-18;7449758
2012-10-18;DOLX12 ;0000080; 000000002037.000000;00000000000000000135;09:00:47.090;1;2012-10-18;7449758456;0000498468;2;2012-10-18;7449758
2012-10-18;DOLX12 ;0000090; 000000002037.500000;00000000000000000015;09:00:58.514;1;2012-10-18;7449758482;0000498506;1;2012-10-18;7449758
2012-10-18;DOLX12 ;0000100; 000000002037.500000;00000000000000000005;09:00:59.204;1;2012-10-18;7449758482;0000498513;2;2012-10-18;7449758
2012-10-18;DOLX12 ;0000120; 000000002037.500000;0000000000000000010;09:01:07.371;1;2012-10-18;7449758423;0000498534;1;2012-10-18;7449758
2012-10-18;DOLX12 ;0000110; 000000002037.500000;00000000000000000005;09:01:07.371;1;2012-10-18;7449758423;0000498532;1;2012-10-18;7449758
2012-10-18;DOLX12 ;0000130; 000000002037.500000;00000000000000000005;09:01:07.595;1;2012-10-18;7449758423;0000498538;2;2012-10-18;7449758
2012-10-18;DOLX12 ;0000170; 000000002037.000000;00000000000000000005;09:01:08.738;1;2012-10-18;7449758480;0000498555;2;2012-10-18;7449758
2012-10-18;DOLX12 ;0000180; 000000002037.000000;0000000000000000010;09:01:08.738;1;2012-10-18;7449758481;0000498557;2;2012-10-18;7449758
2012-10-18;DOLX12 ;0000190; 000000002037.000000;00000000000000000005;09:01:08.738;1;2012-10-18;7449758485;0000498559;2;2012-10-18;7449758
2012-10-18;DOLX12 ;0000150; 000000002037.000000;0000000000000000010;09:01:08.738;1;2012-10-18;7449758468;0000498551;2;2012-10-18;7449758
2012-10-18;DOLX12 ;0000210; 000000002037.000000;00000000000000000040;09:01:08.738;1;2012-10-18;7449758488;0000498563;2;2012-10-18;7449758
2012-10-18;DOLX12 ;0000220; 000000002037.000000;00000000000000000025;09:01:08.738;1;2012-10-18;7449758489;0000498565;2;2012-10-18;7449758
2012-10-18;DOLX12 ;0000140; 000000002037.500000;0000000000000000030;09:01:08.738;1;2012-10-18;7449758423;0000498549;2;2012-10-18;7449758
2012-10-18;DOLX12 ;0000160; 000000002037.000000;0000000000000000010;09:01:08.738;1;2012-10-18;7449758470;0000498553;2;2012-10-18;7449758
2012-10-18;DOLX12 ;0000200; 000000002037.000000;0000000000000000010;09:01:08.738;1;2012-10-18;7449758486;0000498561;2;2012-10-18;7449758
2012-10-18;DOLX12 ;0000230; 000000002037.000000;0000000000000000020;09:01:09.347;1;2012-10-18;7449758505;0000498571;1;2012-10-18;7449758
2012-10-18;DOLX12 ;0000240; 000000002037.000000;0000000000000000025;09:01:10.464;1;2012-10-18;7449758512;0000498581;1;2012-10-18;7449758
2012-10-18;DOLX12 ;0000250; 000000002037.000000;0000000000000000040;09:01:10.469;1;2012-10-18;7449758513;0000498584;1;2012-10-18;7449758
2012-10-18;DOLX12 ;0000260; 000000002037.000000;0000000000000000005;09:01:11.856;1;2012-10-18;7449758517;0000498593;2;2012-10-18;7449758
2012-10-18;DOLX12 ;0000280; 000000002037.000000;0000000000000000005;09:01:14.036;1;2012-10-18;7449758519;0000498605;2;2012-10-18;7449758
2012-10-18;DOLX12 ;0000270; 000000002037.000000;0000000000000000045;09:01:14.036;1;2012-10-18;7449758517;0000498603;2;2012-10-18;7449758
2012-10-18;DOLX12 ;0000290; 000000002037.000000;0000000000000000010;09:01:14.036;1;2012-10-18;7449758523;0000498607;2;2012-10-18;7449758
2012-10-18;DOLX12 ;0000300; 000000002037.000000;0000000000000000040;09:01:14.389;1;2012-10-18;7449758526;0000498610;1;2012-10-18;7449758
2012-10-18;DOLX12 ;0000310; 000000002037.000000;0000000000000000010;09:01:15.442;1;2012-10-18;7449758526;0000498615;2;2012-10-18;7449758
2012-10-18;DOLX12 ;0000320; 000000002037.000000;0000000000000000005;09:01:16.587;1;2012-10-18;7449758533;0000498622;1;2012-10-18;7449758
2012-10-18;DOLX12 ;0000330; 000000002037.000000;0000000000000000020;09:01:21.277;1;2012-10-18;7449758541;0000498634;1;2012-10-18;7449758
2012-10-18;DOLX12 ;0000340; 000000002037.000000;0000000000000000005;09:01:21.277;1;2012-10-18;7449758541;0000498636;1;2012-10-18;7449758
2012-10-18;DOLX12 ;0000350; 000000002037.000000;0000000000000000010;09:01:21.769;1;2012-10-18;7449758541;0000498640;2;2012-10-18;7449758
2012-10-18;DOLX12 ;0000360; 000000002037.000000;0000000000000000015;09:01:28.235;1;2012-10-18;7449758552;0000498655;1;2012-10-18;7449758
    
```

Figura 4.2: Exemplo de arquivo de texto com dados bruto.

Volume Geral

referente: outubro/2013 :: Seleccione o mês

Mercado	Nº de Negócios	Contratos Negociados	Volume Financeiro		Contratos em aberto	Contratos Negociados jan-out
			R\$/mil	US\$/mil		
Pregão (negociação)						
Ibovespa futuro	226.281	1.585.595	85.112.824	38.894.407	201.600	18.044.142
IBRX-50 futuro	370	4.366	395.074	180.966	16.557	57.614
Contrato Futuro S&P 500	3.418	17.316	3.243.087	1.481.877	3.816	214.681
DI de um dia futuro	254.045	23.616.174	2.038.978.736	931.384.577	13.543.900	349.509.872
Setic de um dia futuro	83	204.165	19.368.860	8.864.043	203.930	205.320
Cupom cambial futuro	70	42.553	4.332.414	1.982.304	2.399.035	1.228.272
Dólar comercial futuro	388.356	6.068.525	667.441.019	304.976.367	1.189.673	71.606.094
Euro futuro	454	26.745	4.036.751	1.844.100	12.667	297.301
Iene futuro	86	5.042	566.804	259.813	3.763	135.658
Boi gordo (em reais)	24.320	138.598	4.953.025	2.260.906	20.851	754.111
Café arábica futuro	3.631	10.515	318.121	145.474	8.118	121.847

Figura 4.3: Volumes mensais de alguns contratos futuros da BM&F Bovespa.

Para cada mercado a ser abordado existe uma figura representando o padrão médio de volume intradiário. Dividiu-se cada dia em barras de 10 minutos e calculou-se a proporção do volume diário que está em cada barra. Logo após, retirou-se a média (252 dias de negociação - 21/12/2012 até 30/12/2013) da proporção de volume do dia que cada barra possui.

Futuro de Dólar Comercial

O mercado brasileiro de futuro de taxa de câmbio de reais por dólar comercial da BM&F Bovespa funciona no horário de 09:00 até 18:00. Cada contrato desse mercado, equivale a

USD 50.000,00. No Brasil, a taxa de real para dólar à vista que é usada como referência para a maioria dos contratos cambiais chama-se PTAX. A formação de preço da PTAX é feita consultando os grandes bancos, em janelas aleatórias de dois minutos, quatro vezes ao dia nos horários: consulta 1 - 09:50 às 10:10, consulta 2 - 10:50 às 11:10, consulta 3 - 11:50 às 12:10 e consulta 4 - 12:50 às 13:10, conforme pode ser visto na circular 3506 do Banco Central do Brasil [9].

Os principais fluxos de volume nesse mercado aparecem durante: a abertura do mercado, as consultas da PTAX, o ajuste e o fechamento. Antes da abertura do mercado ocorre um leilão que resulta no primeiro preço negociado do dia. O cálculo do ajuste é feito como a média ponderada de negócios de 15:50 até 16:00, ver [4]. Para mais informações desse contrato futuro, ver [5]. Qualquer citação de contrato de dólar futuro, estaremos implicitamente nos referindo a esse contrato.

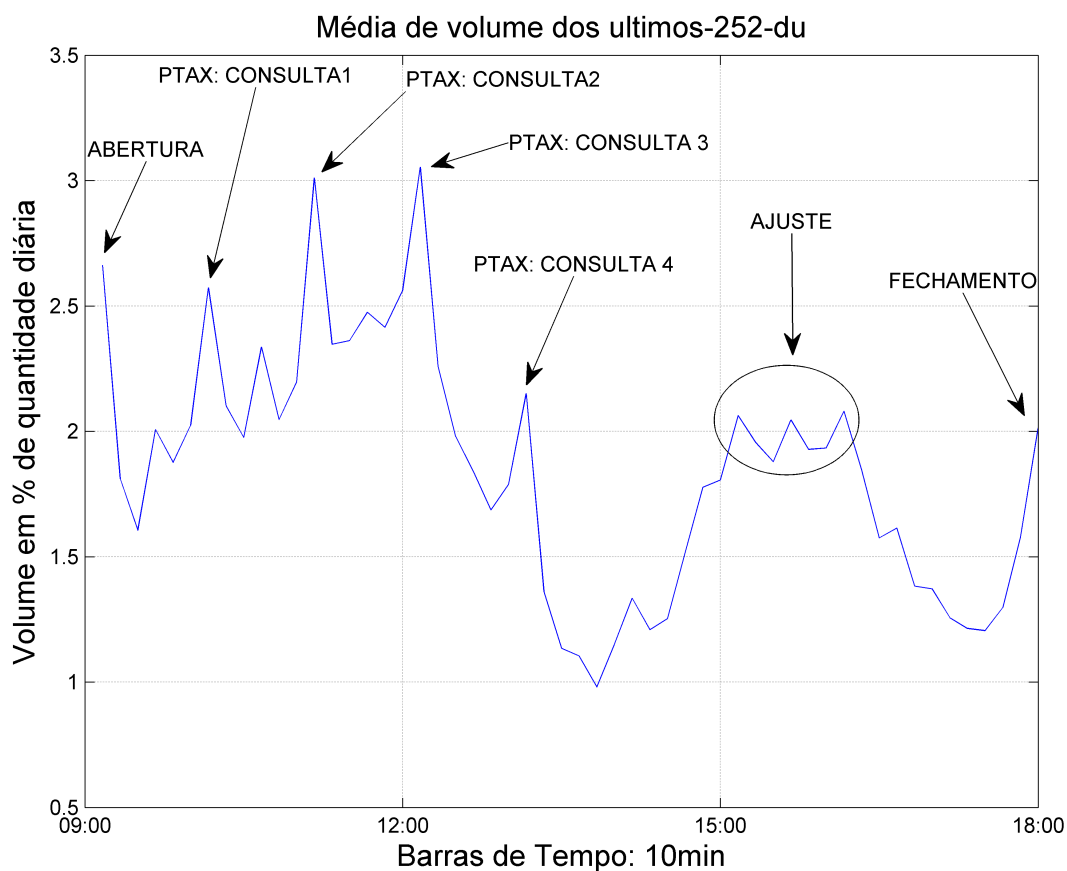


Figura 4.4: Gráfico de concentração de volume médio do futuro de dólar - 252 dias de negociações - 21/12/2012 até 30/12/2013.

A Figura 5.4 mostra a média da proporção do volume diário em cada parte do tempo intradiário. Conforme pode ser visto, os eventos citados ficam extremamente em evidência no gráfico.

Futuro de Taxa DI de Um Dia

O mercado brasileiro de futuro de taxa de DI de um dia da BM&F Bovespa funciona no horário de 09:00 até 18:00. Cada contrato desse mercado, equivale a R\$ 100.000,00 trazido a valor

presente pela taxa negociada (esse valor presente chamado de preço unitário - PU). Antes da abertura do mercado, ocorre um leilão que resulta no primeiro preço negociado do dia. A Figura 5.5 mostra a média da proporção do volume diário em cada parte do tempo intradiário. Conforme pode ser visto, o leilão do ajuste representa o mais significativo nível de volume do dia. Todos os negócios desse leilão ocorrem em um único específico preço que maximiza a quantidade de contratos negociados tendo em vista as ofertas de compra e de venda do leilão. Esse preço será o ajuste do dia, conforme [4]. A partir das 16:00 o mercado é interrompido para a ocorrência do leilão. Após o leilão, o mercado volta a ser negociado normalmente. Qualquer citação de contrato de taxa DI de um dia, estaremos implicitamente nos referindo a esse contrato. Para mais informações desse contrato futuro, ver [6].

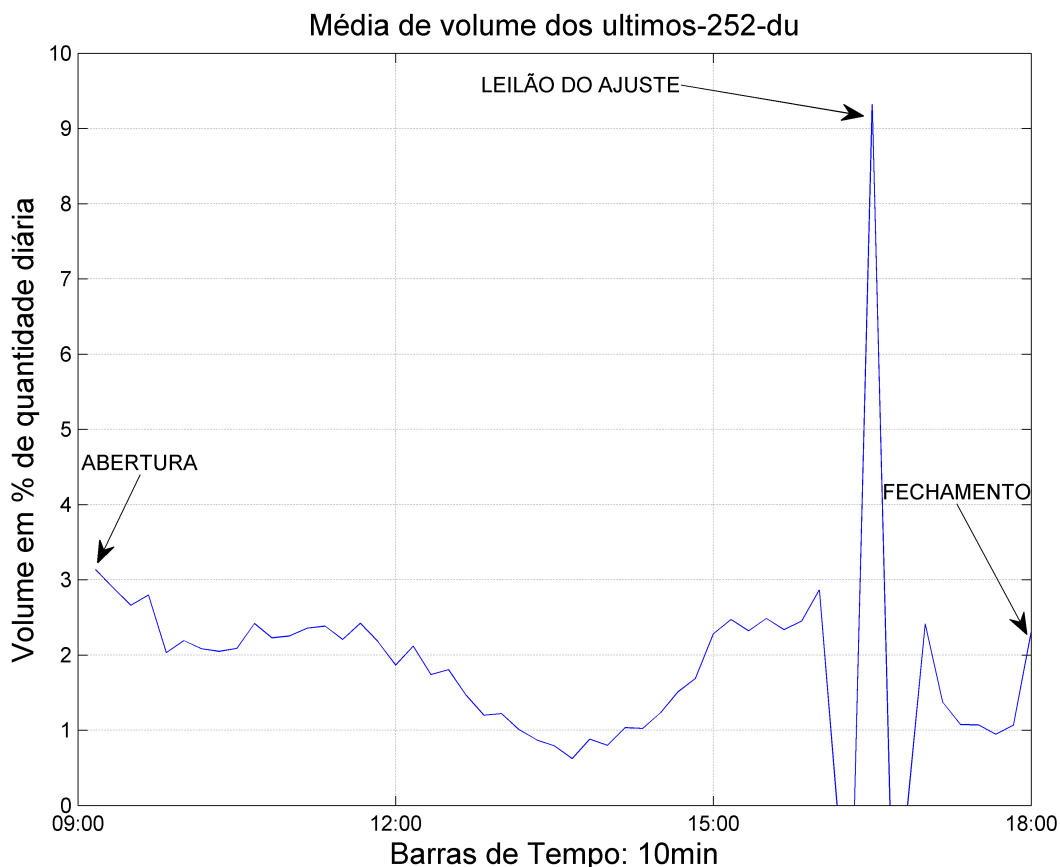


Figura 4.5: Gráfico de concentração de volume médio do futuro de DI de um dia - 252 dias de negociações - 21/12/2012 até 30/12/2013.

Futuro de Índice Bovespa

O mercado brasileiro de futuro de índice bovespa da BM&F Bovespa funciona no horário de 09:00 até 18:00, podendo variar no horário de verão e também de acordo com a liquidez do mercado de à vista. Cada contrato desse mercado, equivale ao valor do mercado futuro convertendo 1 ponto para R\$ 1,00. Sua abertura e fechamento possuem leilões para formação do preço de abertura e do preço de fechamento. A Figura 5.6 mostra a média da proporção do volume diário em cada parte do tempo intradiário. Conforme pode ser visto, a abertura e o fechamento do mercado à vista influencia bastante o padrão da curva de volume intradiário.

Qualquer citação do futuro de índice bovespa, estaremos implicitamente nos referindo a esse contrato. Para mais informações desse contrato futuro, ver [7].

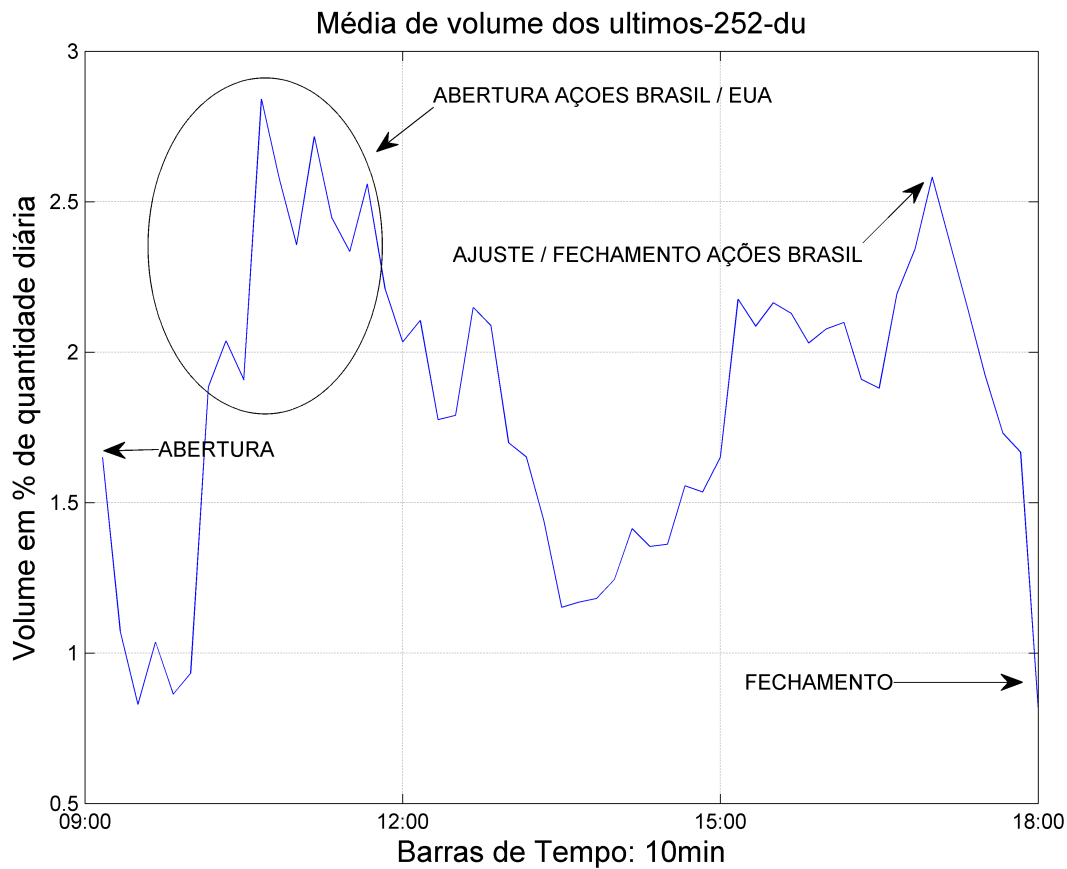


Figura 4.6: Gráfico de concentração de volume médio do futuro de índice bovespa - 252 dias de negociações - 21/12/2012 até 30/12/2013.

4.3 Período de Amostragem

O período escolhido para o estudo do presente trabalho foi de 01/11/2011 até 30/12/2013 que corresponde a um pouco mais de 2 anos de dados. Entre esse período ocorreram 535 dias de negociação. Conforme foi falado no Capítulo 2, desde novembro de 2011 temos o lado agressor (agressão de compra e de venda) sendo gravado pela bolsa nos 3 mercados estudados. Mais precisamente, esses lados de agressões no futuro do índice bovespa começaram a serem gravadas em 24/10/2011, no futuro do dólar comercial em 01/10/2011 e no futuro de taxa DI de um dia em 01/10/2011.

Capítulo 5

Considerações Gerais e Metodologia

5.1 Tratamento de Dados

Todos os dados presente nesse trabalho foram obtidos diretamente do *ftp* da bolsa brasileira BM&F Bovespa [8]. Os dados se encontram gravados em arquivos de texto (ver figura 5.2) num formato próprio da bolsa (ver Figura 5.1). Foi necessário tratar esses arquivos brutos seguindo seus padrões e criar séries históricas *tick-by-tick* de cada ativo.

Identificação	1	2	RH Cabeçalho - RT Rodapé
Nome Arquivo	4	20	
Data Inicial	25	10	Data Inicial desse arquivo
Data Final	36	10	Data Final desse arquivo
Total de Linhas	47	9	No caso do rodapé apresenta o total de linhas do arquivo
Detalhes			

Coluna	Posição Inicial	Tamanho	Descrição

Data Sessão	1	10	Data de sessão
Símbolo do Instrumento	12	50	Símbolo do Instrumento para Código de Bolsa "BMF"
Nr.Negócio	63	10	Número do negócio
Preço Negócio	74	20	Preço do negócio
Quantidade	95	18	Quantidade negociada
Hora	114	15	Horário da negociação (formato HH:MM:SS.NNNNNN)
Ind.Anulação	130	1	Indicador de Anulação
Data Oferta Compra	132	10	Data da oferta de compra
Seq.Oferta Compra	143	10	Número sequencial da oferta de compra
GenerationID - Of.Compra	154	10	Número de geração (GenerationID) da Oferta de Compra.
Condição Oferta de Compra	165	1	Código que identifica a condição da oferta de compra.
Data Oferta Venda	167	10	Data da oferta de venda
Seq.Oferta Venda	178	10	Número sequencial da oferta de venda
GenerationID - Of.Venda	189	10	Número de geração (GenerationID) da Oferta de Venda.
Condição Oferta de Venda	200	1	Código que identifica a condição da oferta de venda.
Obs: Delimitador das colunas de detalhes ';'			

Figura 5.1: Exemplo de formato de arquivo da bolsa.

5.2 Mercados Analisados

O estudo é focado no mercado de contratos futuros. Os movimentos nesse mercado são normalmente mais relacionados a riscos sistemáticos ou a mudanças em variáveis de cobertura de risco [16]. Dessa forma, focaremos nos principais futuros negociados no Brasil: futuros de dólar comercial, de índice Bovespa e de taxa DI de um dia. Na Figura 5.3 temos o volume mensal dos principais contratos futuros da BM&F Bovespa [11].

```

2012-10-18;DOLX12 ;0000010; 000000002037.000000;00000000000000000005;09:00:43.641;1;2012-10-18;7449758433;0000498414;2;2012-10-18;7449758
2012-10-18;DOLX12 ;0000070; 000000002037.000000;0000000000000000010;09:00:45.899;1;2012-10-18;7449758455;0000498464;1;2012-10-18;7449758
2012-10-18;DOLX12 ;0000080; 000000002037.000000;00000000000000000135;09:00:47.090;1;2012-10-18;7449758456;0000498468;2;2012-10-18;7449758
2012-10-18;DOLX12 ;0000090; 000000002037.500000;0000000000000000015;09:00:58.514;1;2012-10-18;7449758482;0000498506;1;2012-10-18;7449758
2012-10-18;DOLX12 ;0000100; 000000002037.500000;0000000000000000005;09:00:59.204;1;2012-10-18;7449758482;0000498513;2;2012-10-18;7449758
2012-10-18;DOLX12 ;0000120; 000000002037.500000;0000000000000000010;09:01:07.371;1;2012-10-18;7449758423;0000498534;1;2012-10-18;7449758
2012-10-18;DOLX12 ;0000110; 000000002037.500000;0000000000000000005;09:01:07.371;1;2012-10-18;7449758423;0000498532;1;2012-10-18;7449758
2012-10-18;DOLX12 ;0000130; 000000002037.500000;0000000000000000005;09:01:07.595;1;2012-10-18;7449758423;0000498538;2;2012-10-18;7449758
2012-10-18;DOLX12 ;0000170; 000000002037.000000;0000000000000000005;09:01:08.738;1;2012-10-18;7449758480;0000498555;2;2012-10-18;7449758
2012-10-18;DOLX12 ;0000180; 000000002037.000000;0000000000000000010;09:01:08.738;1;2012-10-18;7449758481;0000498557;2;2012-10-18;7449758
2012-10-18;DOLX12 ;0000190; 000000002037.000000;0000000000000000005;09:01:08.738;1;2012-10-18;7449758485;0000498559;2;2012-10-18;7449758
2012-10-18;DOLX12 ;0000150; 000000002037.000000;0000000000000000010;09:01:08.738;1;2012-10-18;7449758468;0000498551;2;2012-10-18;7449758
2012-10-18;DOLX12 ;0000210; 000000002037.000000;0000000000000000040;09:01:08.738;1;2012-10-18;7449758488;0000498563;2;2012-10-18;7449758
2012-10-18;DOLX12 ;0000220; 000000002037.000000;0000000000000000025;09:01:08.738;1;2012-10-18;7449758489;0000498565;2;2012-10-18;7449758
2012-10-18;DOLX12 ;0000140; 000000002037.500000;0000000000000000030;09:01:08.738;1;2012-10-18;7449758423;0000498549;2;2012-10-18;7449758
2012-10-18;DOLX12 ;0000160; 000000002037.000000;0000000000000000010;09:01:08.738;1;2012-10-18;7449758470;0000498553;2;2012-10-18;7449758
2012-10-18;DOLX12 ;0000200; 000000002037.000000;0000000000000000010;09:01:08.738;1;2012-10-18;7449758486;0000498561;2;2012-10-18;7449758
2012-10-18;DOLX12 ;0000230; 000000002037.000000;0000000000000000020;09:01:09.347;1;2012-10-18;7449758505;0000498571;1;2012-10-18;7449758
2012-10-18;DOLX12 ;0000240; 000000002037.000000;0000000000000000025;09:01:10.464;1;2012-10-18;7449758512;0000498581;1;2012-10-18;7449758
2012-10-18;DOLX12 ;0000250; 000000002037.000000;0000000000000000040;09:01:10.469;1;2012-10-18;7449758513;0000498584;1;2012-10-18;7449758
2012-10-18;DOLX12 ;0000260; 000000002037.000000;0000000000000000005;09:01:11.856;1;2012-10-18;7449758517;0000498593;2;2012-10-18;7449758
2012-10-18;DOLX12 ;0000280; 000000002037.000000;0000000000000000005;09:01:14.036;1;2012-10-18;7449758519;0000498605;2;2012-10-18;7449758
2012-10-18;DOLX12 ;0000270; 000000002037.000000;0000000000000000045;09:01:14.036;1;2012-10-18;7449758517;0000498603;2;2012-10-18;7449758
2012-10-18;DOLX12 ;0000290; 000000002037.000000;0000000000000000010;09:01:14.036;1;2012-10-18;7449758523;0000498607;2;2012-10-18;7449758
2012-10-18;DOLX12 ;0000300; 000000002037.000000;0000000000000000040;09:01:14.389;1;2012-10-18;7449758526;0000498610;1;2012-10-18;7449758
2012-10-18;DOLX12 ;0000310; 000000002037.000000;0000000000000000010;09:01:15.442;1;2012-10-18;7449758526;0000498615;2;2012-10-18;7449758
2012-10-18;DOLX12 ;0000320; 000000002037.000000;0000000000000000005;09:01:16.587;1;2012-10-18;7449758533;0000498622;1;2012-10-18;7449758
2012-10-18;DOLX12 ;0000330; 000000002037.000000;0000000000000000020;09:01:21.277;1;2012-10-18;7449758541;0000498634;1;2012-10-18;7449758
2012-10-18;DOLX12 ;0000340; 000000002037.000000;0000000000000000005;09:01:21.277;1;2012-10-18;7449758541;0000498636;1;2012-10-18;7449758
2012-10-18;DOLX12 ;0000350; 000000002037.000000;0000000000000000010;09:01:21.765;1;2012-10-18;7449758541;0000498640;2;2012-10-18;7449758
2012-10-18;DOLX12 ;0000360; 000000002037.000000;0000000000000000015;09:01:28.235;1;2012-10-18;7449758552;0000498655;1;2012-10-18;7449758
    
```

Figura 5.2: Exemplo de arquivo de texto com dados bruto.

Volume Geral

referente: outubro/2013 :: Seleccione o mês

Mercado	Nº de Negócios	Contratos Negociados	Volume Financeiro		Contratos em aberto	Contratos Negociados jan-out
			R\$/mil	US\$/mil		
Pregão (negociação)						
Ibovespa futuro	226.281	1.585.595	85.112.824	38.894.407	201.600	18.044.142
IBRX-50 futuro	370	4.366	395.074	180.966	16.557	57.614
Contrato Futuro S&P 500	3.418	17.316	3.243.087	1.481.877	3.816	214.681
DI de um dia futuro	254.045	23.616.174	2.038.978.736	931.384.577	13.543.900	349.509.872
Setic de um dia futuro	83	204.165	19.368.860	8.864.043	203.930	205.320
Cupom cambial futuro	70	42.553	4.332.414	1.982.304	2.399.035	1.228.272
Dólar comercial futuro	388.356	6.068.525	667.441.019	304.976.367	1.189.673	71.606.094
Euro futuro	454	26.745	4.036.751	1.844.100	12.667	297.301
Iene futuro	86	5.042	566.804	259.813	3.763	135.658
Boi gordo (em reais)	24.320	138.598	4.953.025	2.260.906	20.851	754.111
Café arábica futuro	3.631	10.515	318.121	145.474	8.118	121.847

Figura 5.3: Volumes mensais de alguns contratos futuros da BM&F Bovespa.

Para cada mercado a ser abordado existe uma figura representando o padrão médio de volume intradiário. Dividiu-se cada dia em barras de 10 minutos e calculou-se a proporção do volume diário que está em cada barra. Logo após, retirou-se a média (252 dias de negociação - 21/12/2012 até 30/12/2013) da proporção de volume do dia que cada barra possui.

Futuro de Dólar Comercial

O mercado brasileiro de futuro de taxa de câmbio de reais por dólar comercial da BM&F Bovespa funciona no horário de 09:00 até 18:00. Cada contrato desse mercado, equivale a

USD 50.000,00. No Brasil, a taxa de real para dólar à vista que é usada como referência para a maioria dos contratos cambiais chama-se PTAX. A formação de preço da PTAX é feita consultando os grandes bancos, em janelas aleatórias de dois minutos, quatro vezes ao dia nos horários: consulta 1 - 09:50 às 10:10, consulta 2 - 10:50 às 11:10, consulta 3 - 11:50 às 12:10 e consulta 4 - 12:50 às 13:10, conforme pode ser visto na circular 3506 do Banco Central do Brasil [9].

Os principais fluxos de volume nesse mercado aparecem durante: a abertura do mercado, as consultas da PTAX, o ajuste e o fechamento. Antes da abertura do mercado ocorre um leilão que resulta no primeiro preço negociado do dia. O cálculo do ajuste é feito como a média ponderada de negócios de 15:50 até 16:00, ver [4]. Para mais informações desse contrato futuro, ver [5]. Qualquer citação de contrato de dólar futuro, estaremos implicitamente nos referindo a esse contrato.

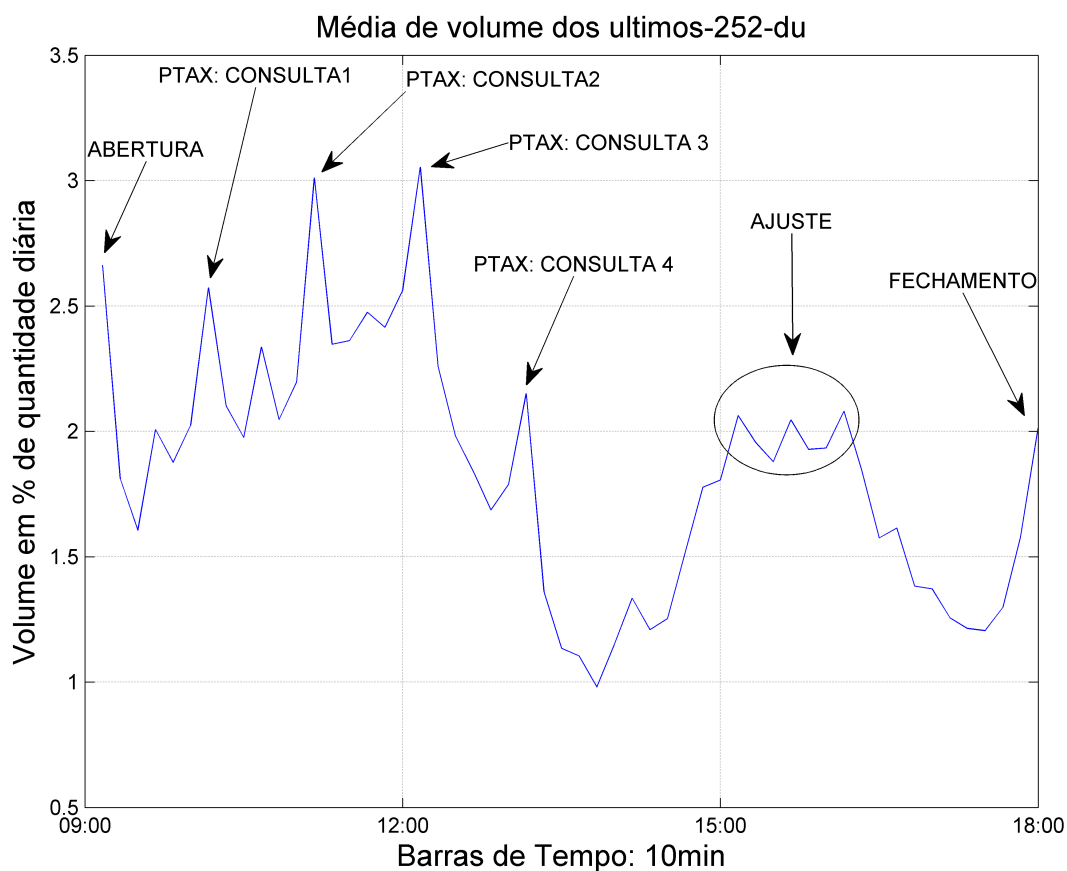


Figura 5.4: Gráfico de concentração de volume médio do futuro de dólar - 252 dias de negociações - 21/12/2012 até 30/12/2013.

A Figura 5.4 mostra a média da proporção do volume diário em cada parte do tempo intradiário. Conforme pode ser visto, os eventos citados ficam extremamente em evidência no gráfico.

Futuro de Taxa DI de Um Dia

O mercado brasileiro de futuro de taxa de DI de um dia da BM&F Bovespa funciona no horário de 09:00 até 18:00. Cada contrato desse mercado, equivale a R\$ 100.000,00 trazido a valor

presente pela taxa negociada (esse valor presente chamado de preço unitário - PU). Antes da abertura do mercado, ocorre um leilão que resulta no primeiro preço negociado do dia. A Figura 5.5 mostra a média da proporção do volume diário em cada parte do tempo intradiário. Conforme pode ser visto, o leilão do ajuste representa o mais significativo nível de volume do dia. Todos os negócios desse leilão ocorrem em um único específico preço que maximiza a quantidade de contratos negociados tendo em vista as ofertas de compra e de venda do leilão. Esse preço será o ajuste do dia, conforme [4]. A partir das 16:00 o mercado é interrompido para a ocorrência do leilão. Após o leilão, o mercado volta a ser negociado normalmente. Qualquer citação de contrato de taxa DI de um dia, estaremos implicitamente nos referindo a esse contrato. Para mais informações desse contrato futuro, ver [6].

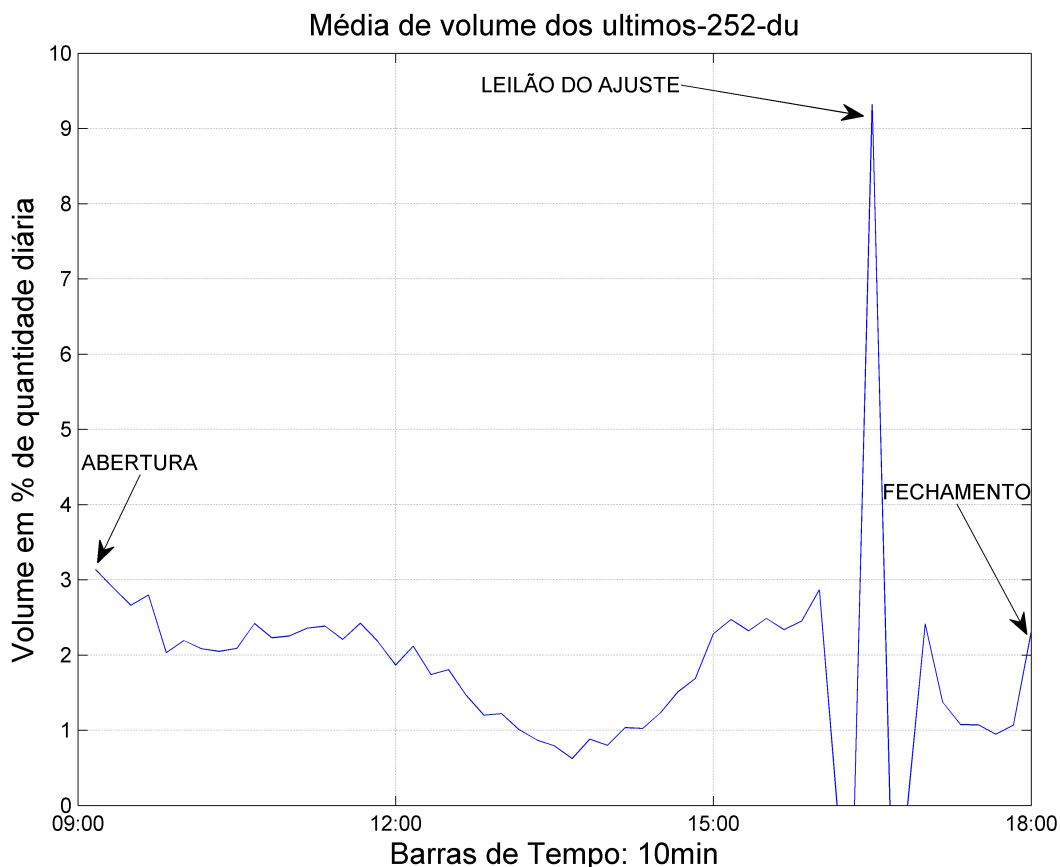


Figura 5.5: Gráfico de concentração de volume médio do futuro de DI de um dia - 252 dias de negociações - 21/12/2012 até 30/12/2013.

Futuro de Índice Bovespa

O mercado brasileiro de futuro de índice bovespa da BM&F Bovespa funciona no horário de 09:00 até 18:00, podendo variar no horário de verão e também de acordo com a liquidez do mercado de à vista. Cada contrato desse mercado, equivale ao valor do mercado futuro convertendo 1 ponto para R\$ 1,00. Sua abertura e fechamento possuem leilões para formação do preço de abertura e do preço de fechamento. A Figura 5.6 mostra a média da proporção do volume diário em cada parte do tempo intradiário. Conforme pode ser visto, a abertura e o fechamento do mercado à vista influencia bastante o padrão da curva de volume intradiário.

Qualquer citação do futuro de índice bovespa, estaremos implicitamente nos referindo a esse contrato. Para mais informações desse contrato futuro, ver [7].

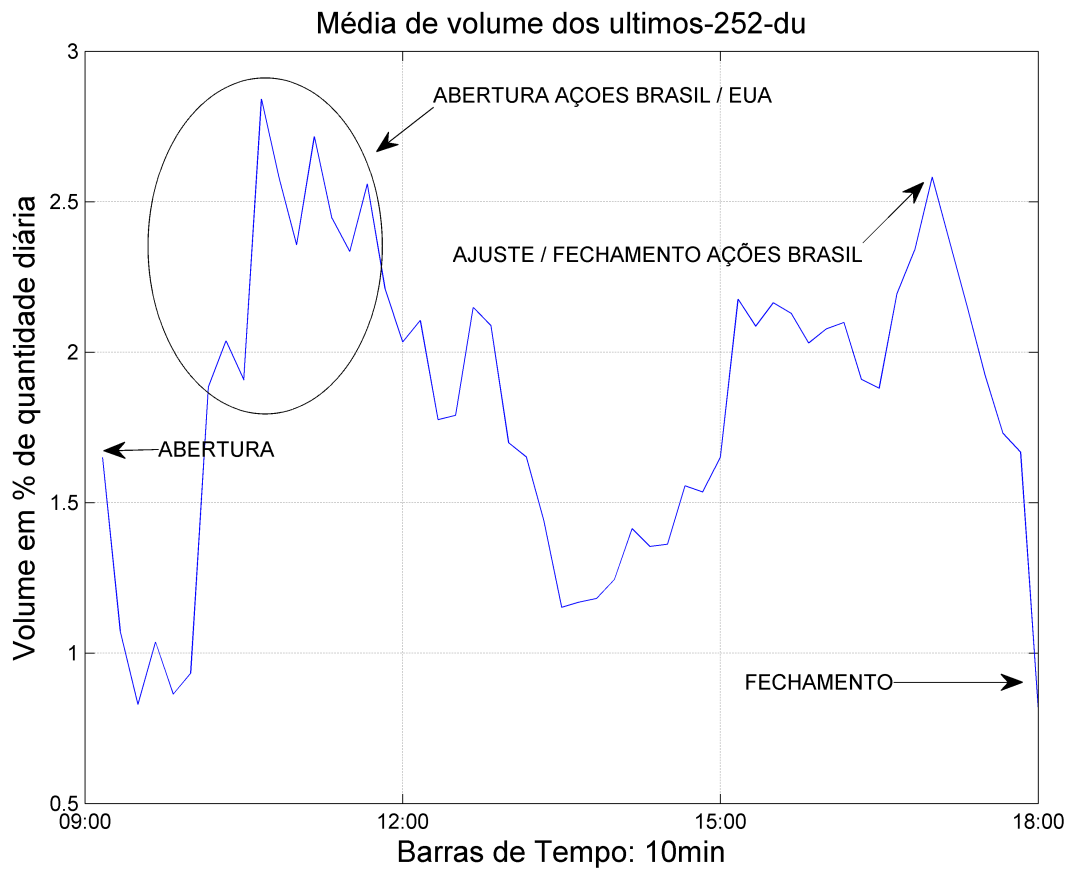


Figura 5.6: Gráfico de concentração de volume médio do futuro de índice bovespa - 252 dias de negociações - 21/12/2012 até 30/12/2013.

5.3 Período de Amostragem

O período escolhido para o estudo do presente trabalho foi de 01/11/2011 até 30/12/2013 que corresponde a um pouco mais de 2 anos de dados. Entre esse período ocorreram 535 dias de negociação. Conforme foi falado no Capítulo 2, desde novembro de 2011 temos o lado agressor (agressão de compra e de venda) sendo gravado pela bolsa nos 3 mercados estudados. Mais precisamente, esses lados de agressões no futuro do índice bovespa começaram a serem gravadas em 24/10/2011, no futuro do dólar comercial em 01/10/2011 e no futuro de taxa DI de um dia em 01/10/2011.

Capítulo 6

Toxicidade no Brasil

Dividiremos o estudo pelos seguintes mercados: futuro de dólar, futuro de índice bovespa e futuro de DI de 1 dia (ver explanação desses mercados no Capítulo 5). Para cada mercado, será analisado as 3 formas de classificação de negócio estudadas no Capítulo 2. Vale lembrar que o período amostral foi do dia 01/11/2011 até o dia 30/12/2013 (pouco mais de 2 anos de mercado).

Seguimos o esquema representado na Figura 6.1. Inicialmente, os negócios são extraídos dos dados brutos e colocados na forma *tick-by-tick*, conforme procedimento citado no Capítulo 5. A partir disso, serão usados os algoritmos de classificação de negócios para serem gerados três diferentes tipos de VPIN: TR-VPIN, BVC-VPIN e EC-VPIN. A implementação dos algoritmos foram na linguagem matlab e os códigos estão disponibilizados no apêndice.

Em todas os nossos cálculos de VPIN foram usados $(q, m, n) = (10, 50, 50)$ conforme foi detalhado no Capítulo 3.

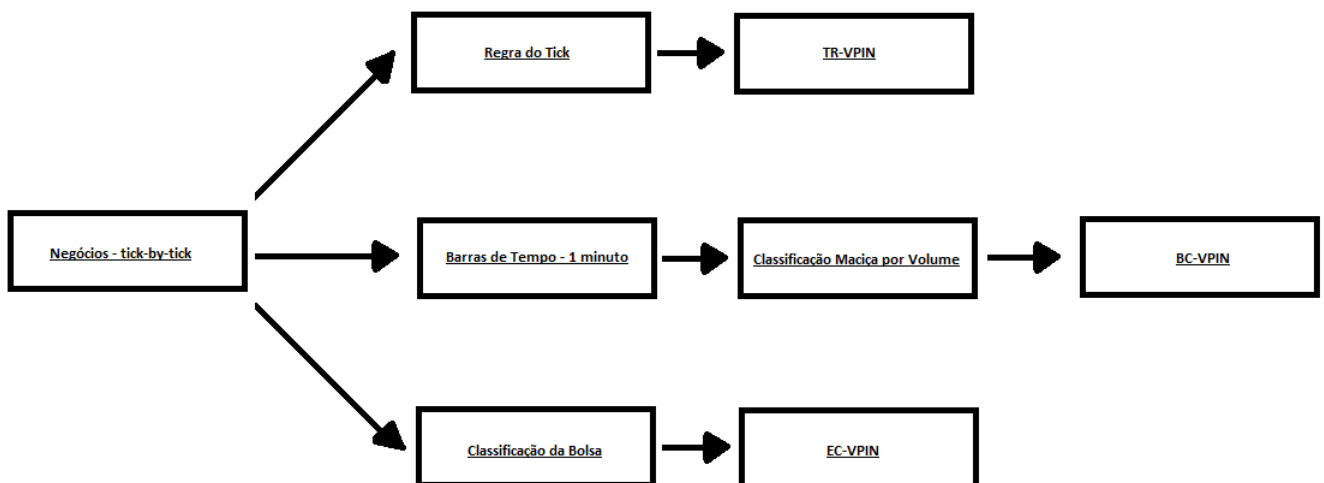


Figura 6.1: Divisão para o cálculo do VPIN por diferentes metodologias de classificação de negócio.

Será mostrado o histograma e a evolução do VPIN durante todo o período analisado. Além disso, serão mostrados mais dois gráficos em cada análise. Para fugir de subjetividade e de viés na escolha dos períodos de cada gráfico, iremos padronizar os períodos mostrados nesses dois gráficos: um será o gráfico intradiário do dia em que o VPIN máximo foi atingindo e um outro gráfico numa janela de 6 dias, composta pelo dia em que atingiu o máximo, mais 4 dias anteriores e 1 dia posterior. Nesse gráfico diário, cada dia é delimitado por círculos. Quanto

ao conteúdo, em cada um desses gráficos, serão mostradas as evoluções do preço do contrato futuro, do VPIN e da função de distribuição acumulada (FDA) do VPIN. Iremos analisar os períodos próximos ao máximo do VPIN no qual o FDA VPIN atinge um valor alto (90% é usado em diversas análises em [22]).

6.1 Dólar-Real

Foi usado o contrato de dolar comercial futuro negociado na BM&F Bovespa.

Mercado	dólar futuro
Início da Amostra	01/11/2011
Fim da Amostra	30/12/2013
Número de dias de negociação	535
Número Total de <i>Buckets</i>	27231

A Figura 6.2 mostra o histograma da quantidade de 1 *bucket* e a Figura 6.3 mostra o valor da quantidade de 1 *bucket* no decorrer dos 535 dias amostrados. Em períodos de muita liquidez, o estimador de volume diário (a média de volume dos ultimos 10 dias) acaba aumentando e gera um aumento nessa quantidade. Todavia, em momentos de pouca liquidez o estimador de volume diário acaba diminuindo e, conseqüentemente, faz a quantidade de 1 *bucket* diminuir.

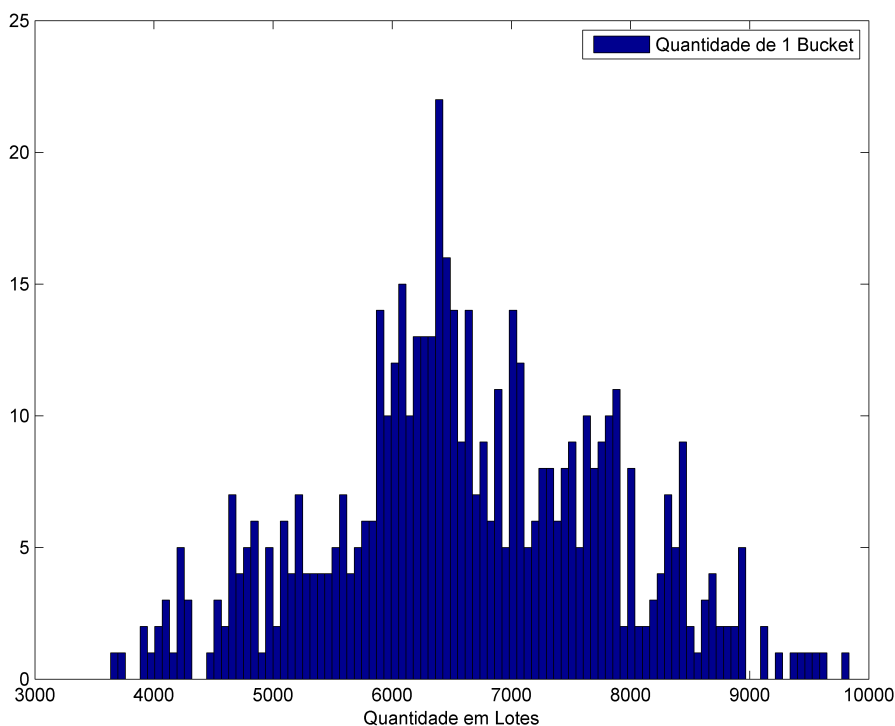


Figura 6.2: Histograma da quantidade de 1 *bucket* no dólar futuro.

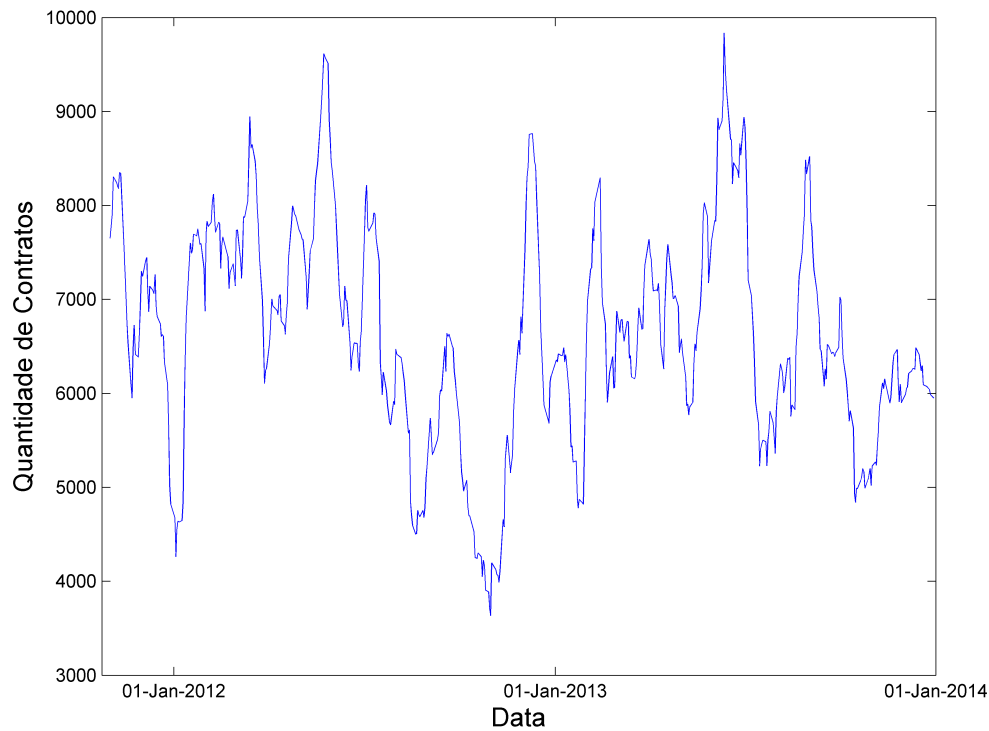


Figura 6.3: Gráfico da quantidade de 1 *bucket* no dólar futuro.

BVC-VPIN

Teste de Jarque-Bera ($\alpha = 5\%$)	Rejeitado
Média	0.2451
Mediana	0.2398
Desvio Padrão	0.0441
Simetria	0.7424
Curtose	4.2326
Máximo	0.4849

Vemos na Figura 6.5 um indicação mostrando que o BVC-VPIN atinge o valor máximo do dia 09/11/2012. Mais precisamente no *bucket* que inicia: 09/11/2012 12:04:00.000 e termina : 09/11/2012 12:09:00.000; Nesse dia o dólar chegou a subir um pouco mais de 1% e a PTAX, que é a principal referência para o dólar à vista, atingiu o seu maior valor por um pouco mais de 2 meses (desde o dia 30/08/2012). O gráfico da PTAX durante esse intervalo de tempo pode ser visto na ser visto na Figura 6.6.

Na Figura 6.8 o FDA BVC-VPIN ultrapassou 95% no fechamento do dia 08/11/2012 podendo dar indícios que o dia 09/11/2012 seria incomum.

O coeficiente de correlação de Pearson do VPIN e do logaritmo do VPIN com o módulo do retorno do *bucket* subsequente foram calculados:

$$\rho(VPIN_t, |\frac{P_{t+1}}{P_t} - 1|) = 0.0310$$

$$\rho(\ln(VPIN_t), |\frac{P_{t+1}}{P_t} - 1|) = 0.0369$$

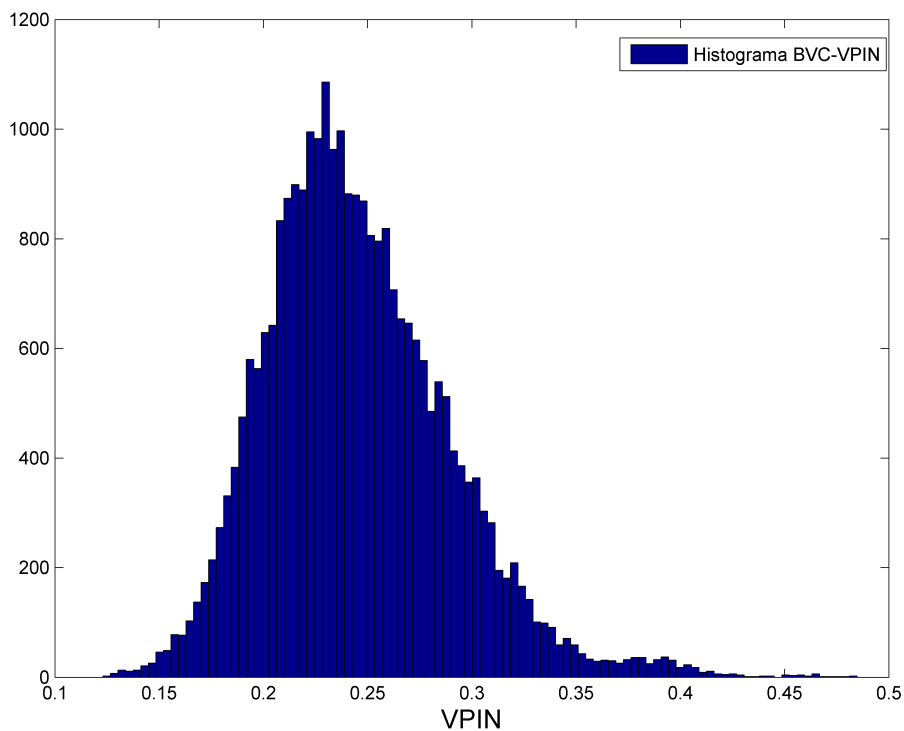


Figura 6.4: Histograma do BVC-VPIN no dólar futuro.

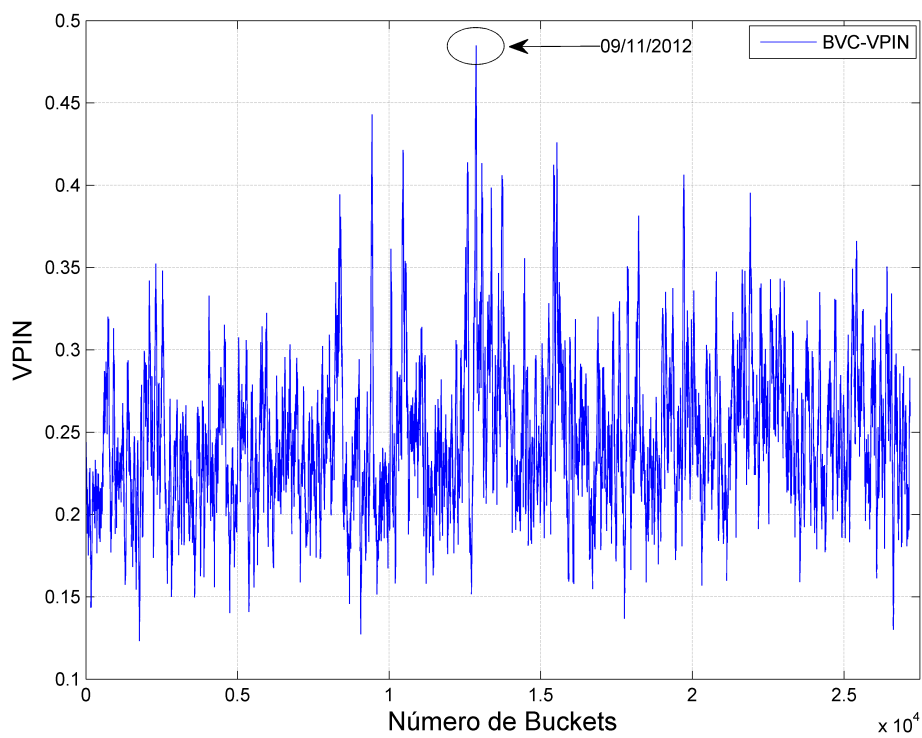


Figura 6.5: Máximo atingido pelo BVC-VPIN no dólar futuro.

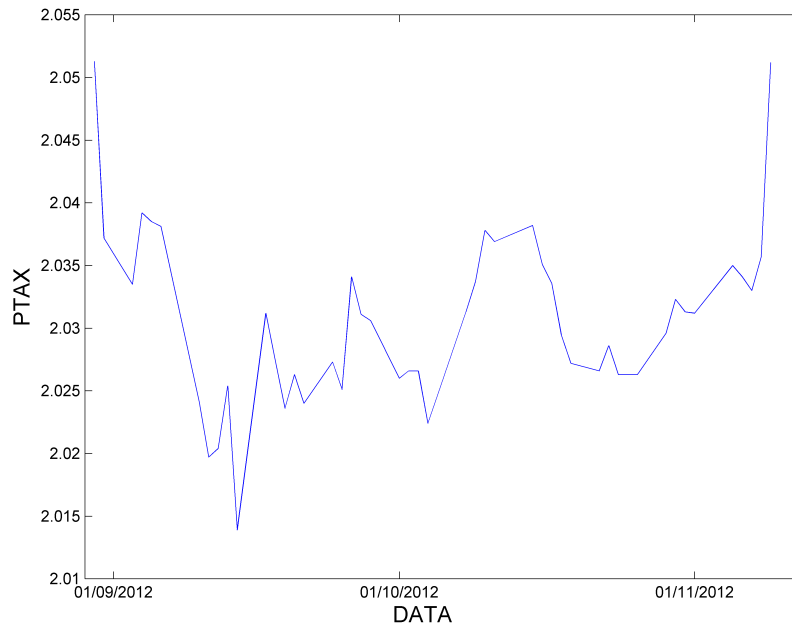


Figura 6.6: PTAX - do dia 30/08/2012 até o dia 09/11/2012.

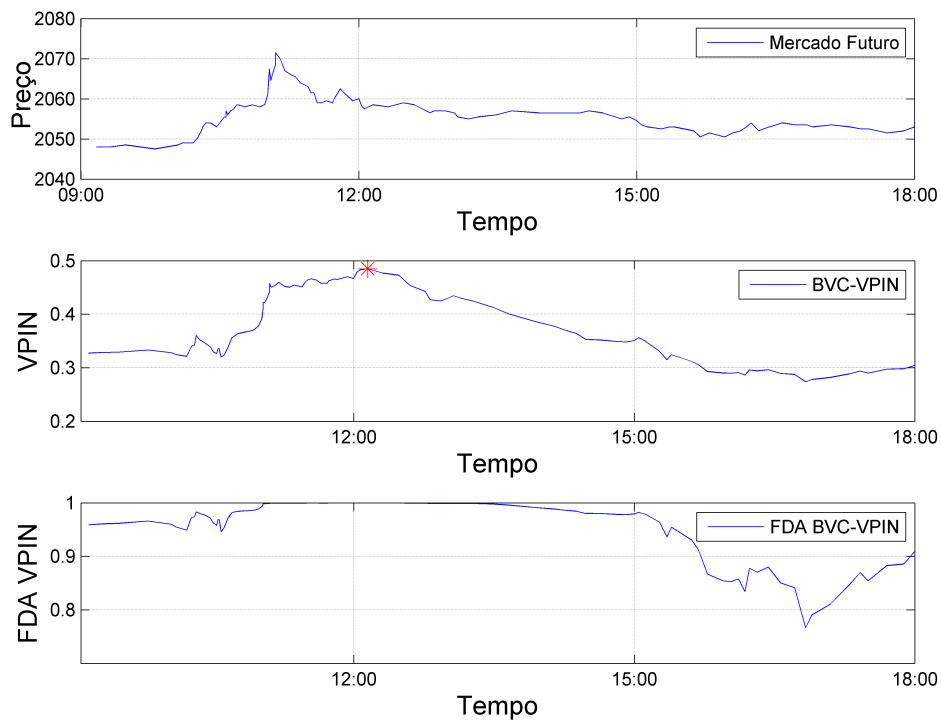


Figura 6.7: Comportamento intradiário do BVC-VPIN do dia 09/11/2012 no dólar futuro.

EC-VPIN

Teste de Jarque-Bera ($\alpha = 5\%$)	Rejeitado
Média	0.2459
Mediana	0.2350
Desvio Padrão	0.0611
Simetria	1.1907
Curtose	5.0611
Máximo	0.5656

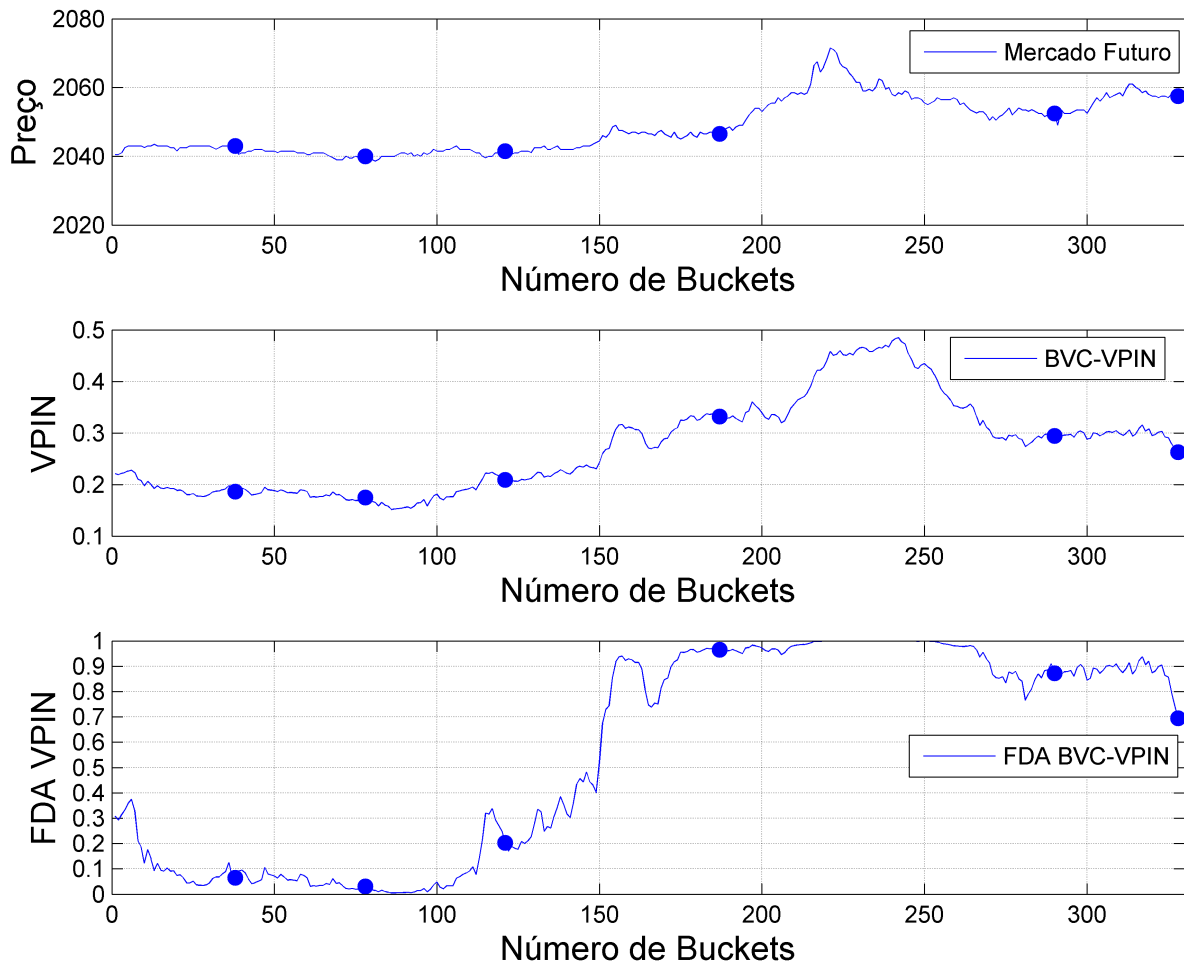


Figura 6.8: Comportamento do BVC-VPIN - 6 dias úteis - dia 05/11/2012 até o dia 12/11/2012 no dólar futuro.

Vemos na Figura 6.10 um indicação mostrando que o EC-VPIN atinge o valor máximo do dia 31/10/2012. Mais precisamente no *bucket* que inicia: 31/10/2012 13:00:30.960 e termina : 31/10/2012 13:00:33.984;

O coeficiente de correlação de Pearson do VPIN e do logaritmo do VPIN com o módulo do retorno do *bucket* subsequente foram calculados:

$$\rho(VPIN_t, |\frac{P_{t+1}}{P_t} - 1|) = -0.2245$$

$$\rho(\ln(VPIN_t), |\frac{P_{t+1}}{P_t} - 1|) = -0.2283$$

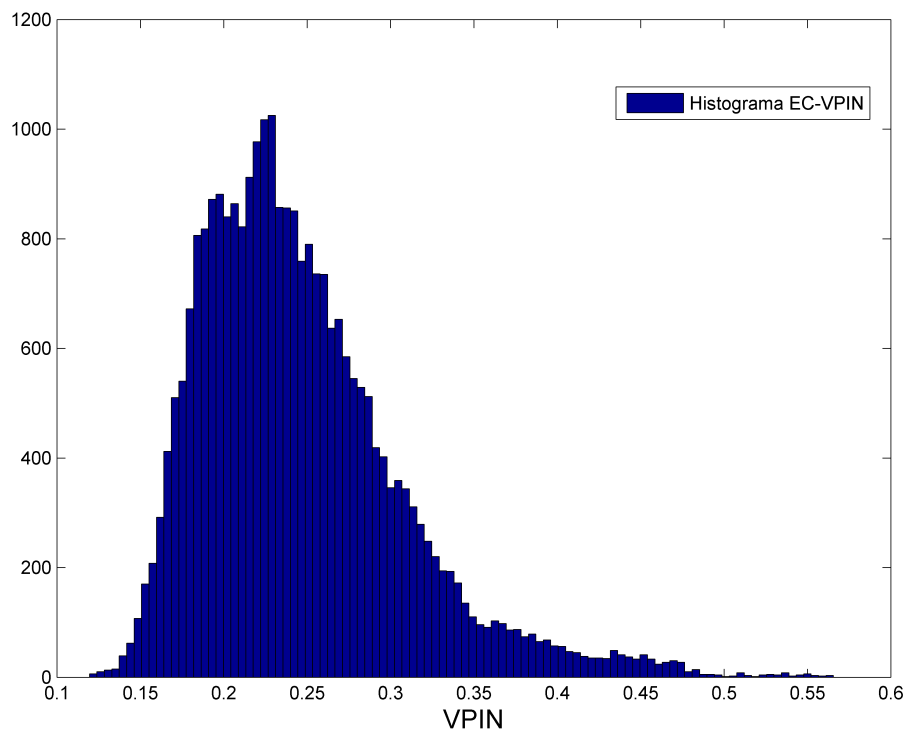


Figura 6.9: Histograma do EC-VPIN no dólar futuro.

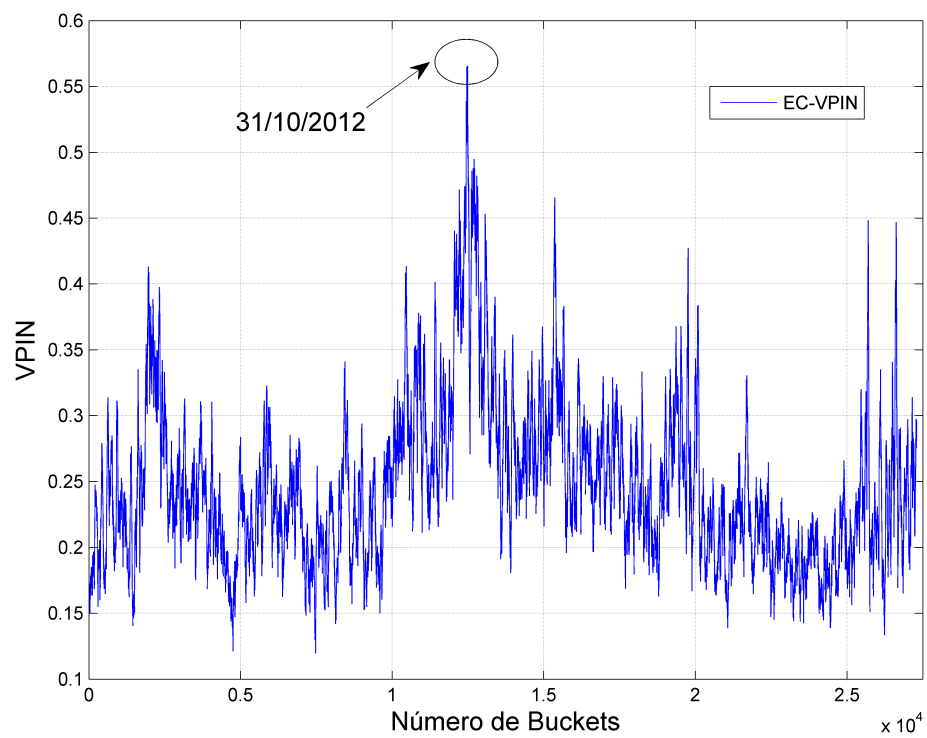


Figura 6.10: Máximo atingido pelo EC-VPIN no dólar futuro.

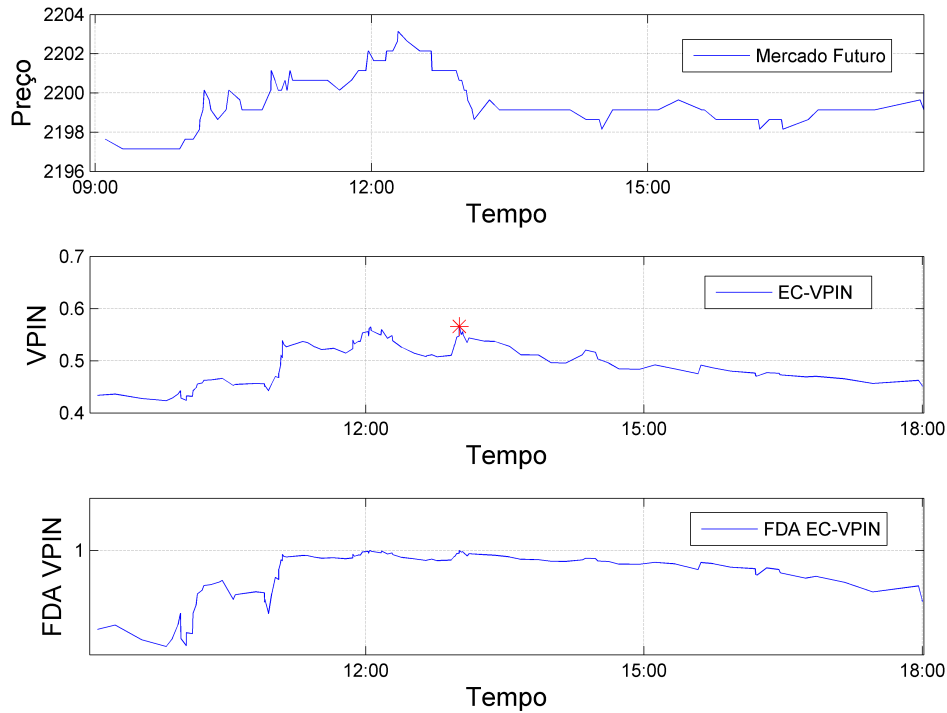


Figura 6.11: Comportamento intradiário do EC-VPIN do dia 31/10/2012 no dólar futuro.

TR-VPIN

Teste de Jarque-Bera ($\alpha = 5\%$)	Rejeitado
Média	0.3197
Mediana	0.3095
Desvio Padrão	0.0661
Simetria	1.0043
Curtose	4.5946
Máximo	0.6753

Vemos na Figura 6.14 um indicação mostrando que o TR-VPIN atinge o valor máximo do dia 01/11/2012. Mais precisamente no *bucket* que inicia: 01/11/2012 09:18:31.018 e termina : 01/11/2012 09:23:17.002;

O coeficiente de correlação de Pearson do VPIN e do logaritmo do VPIN com o módulo do retorno do *bucket* subsequente foram calculados:

$$\rho(VPIN_t, |\frac{P_{t+1}}{P_t} - 1|) = -0.2207$$

$$\rho(\ln(VPIN_t), |\frac{P_{t+1}}{P_t} - 1|) = -0.2236$$

6.2 Índice Bovespa

Foi usado o contrato de índice bovespa futuro negociado na BM&F Bovespa.

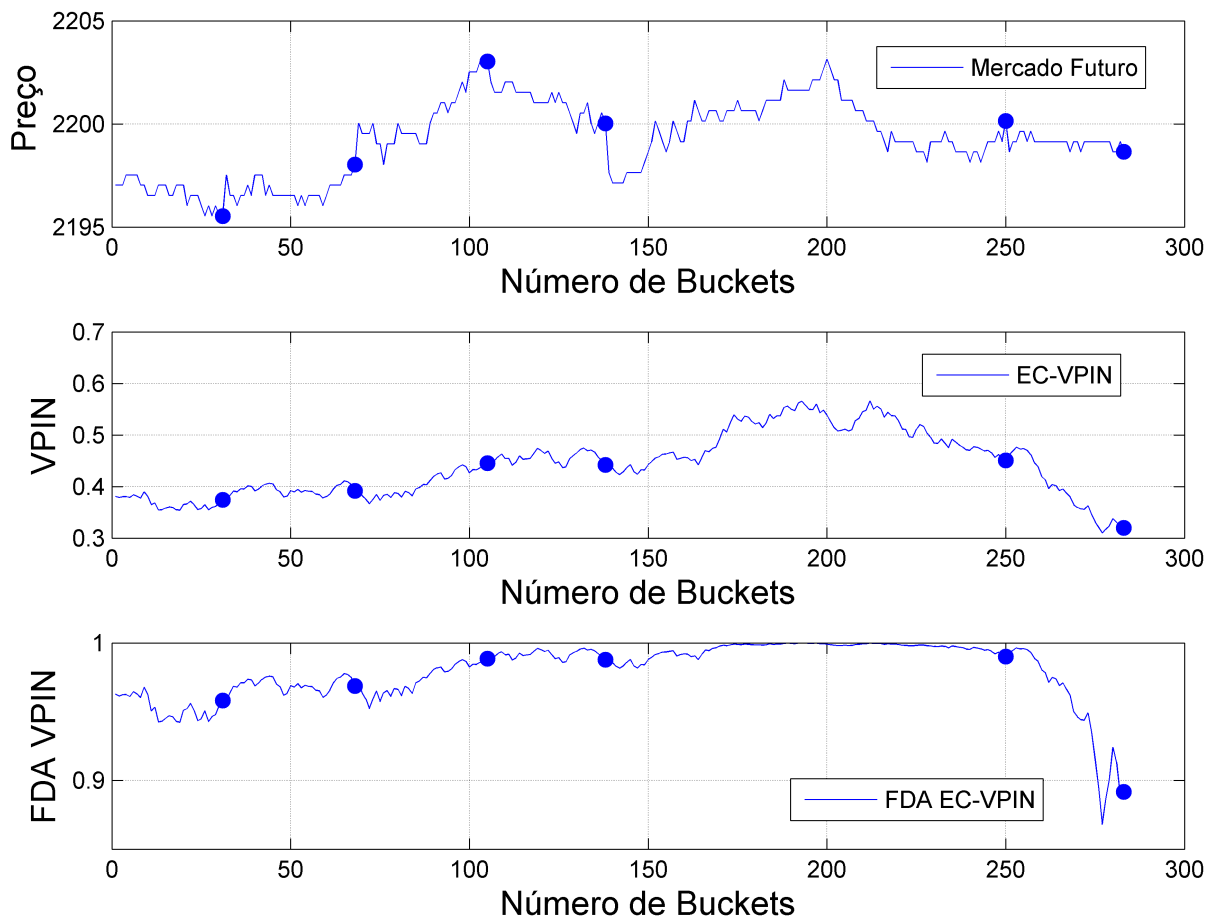


Figura 6.12: Comportamento do EC-VPIN - 6 dias úteis - dia 25/09/2012 até o dia 01/11/2012 no dólar futuro.

Mercado	índice futuro
Início da Amostra	01/11/2011
Fim da Amostra	30/12/2013
Número de dias de negociação	535
Número Total de <i>Buckets</i>	27079

A Figura 6.17 mostra o histograma da quantidade de 1 *bucket* e a Figura 6.18 mostra o valor da quantidade de 1 *bucket* no decorrer dos 535 dias amostrados. Em períodos de muita liquidez, o estimador de volume diário (a média de volume dos últimos 10 dias) acaba aumentando e gera um aumento nessa quantidade. Todavia, em momentos de pouca liquidez o estimador de volume diário acaba diminuindo e, conseqüentemente, faz a quantidade de 1 *bucket* diminuir.

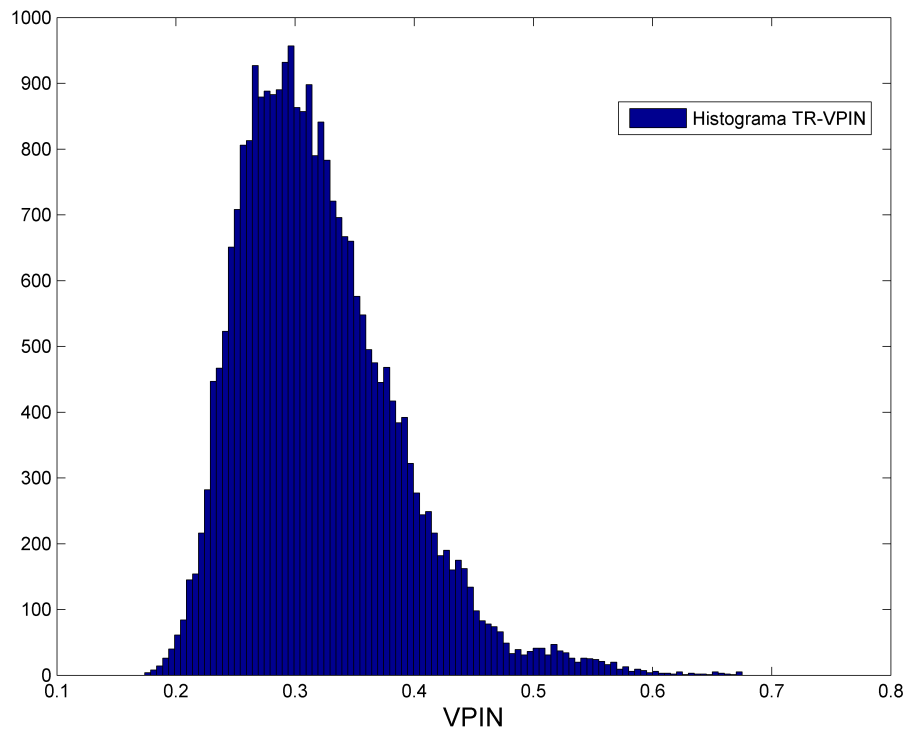


Figura 6.13: Histograma do TR-VPIN no dólar futuro.

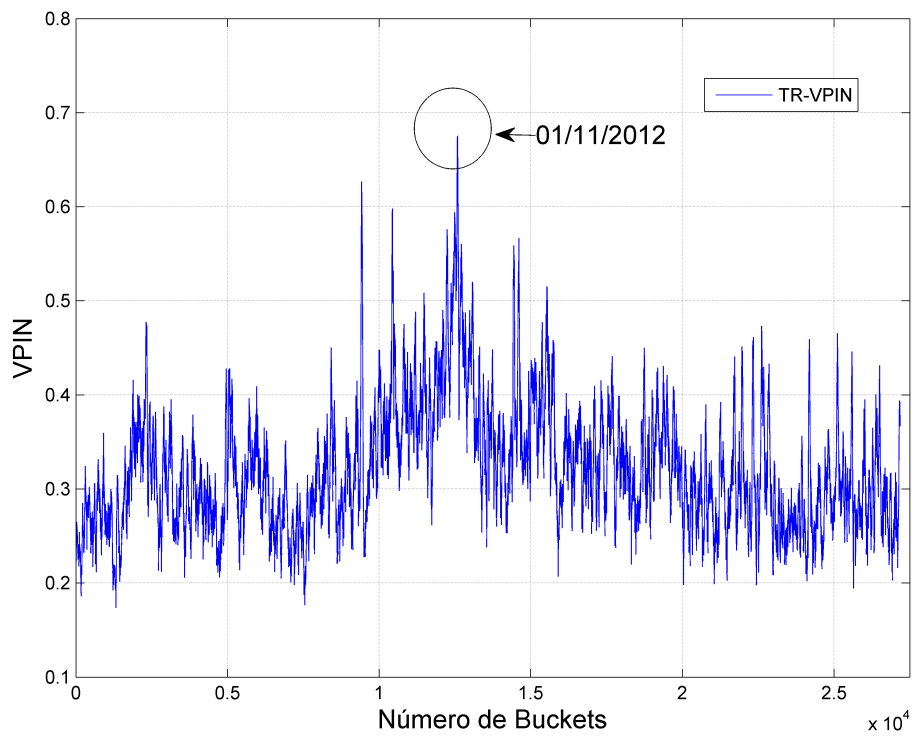


Figura 6.14: Máximo atingido pelo TR-VPIN no dólar futuro.

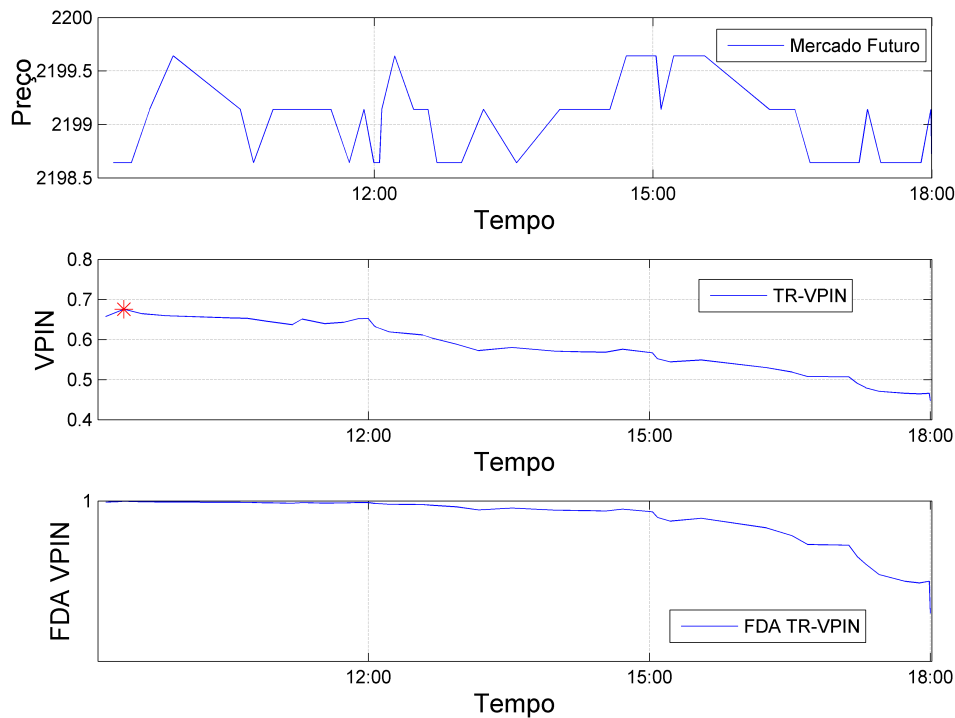


Figura 6.15: Comportamento intradiário do TR-VPIN do dia 01/11/2012 no dólar futuro.

BVC-VPIN

Teste de Jarque-Bera ($\alpha = 5\%$)	Rejeitado
Média	0.2686
Mediana	0.2656
Desvio Padrão	0.0373
Simetria	0.7688
Curtose	4.8891
Máximo	0.4881

Vemos na Figura 6.20 aparece uma indicação mostrando que o BVC-VPIN atinge o valor máximo do dia 27/07/2012. Mais precisamente no *bucket* que inicia: 27/07/2012 15:36:00.000 e termina : 27/07/2012 15:37:00.000; nesse dia o índice bovespa fechou subindo 4.72% (conforme [2]) e atingiu a maior alta intradiária de quase um ano (no dia 09/08/2011 o índice subiu 5.10%, [1]). Também nesse dia ocorreu uma quebra de recorde de negócios de ações com 1.110.540 negócios (conforme nota em [3]).

Olhando a Figura 6.21 vemos que seria muito difícil antecipar esse movimento inteiro. Entretanto, um pouco antes das 15:00 o FDA do BVC-VPIN já havia atingido mais de 95% e desse momento em diante, o mercado ainda andou um pouco mais de 1000 pontos. Olhando a Figura 6.22 não existe indicativo algum que aconteceria isso nesse dia.

O coeficiente de correlação de Pearson do VPIN e do logaritmo do VPIN com o módulo do retorno do *bucket* subsequente foram calculados:

$$\rho(VPIN_t, |\frac{P_{t+1}}{P_t} - 1|) = -0.0337$$

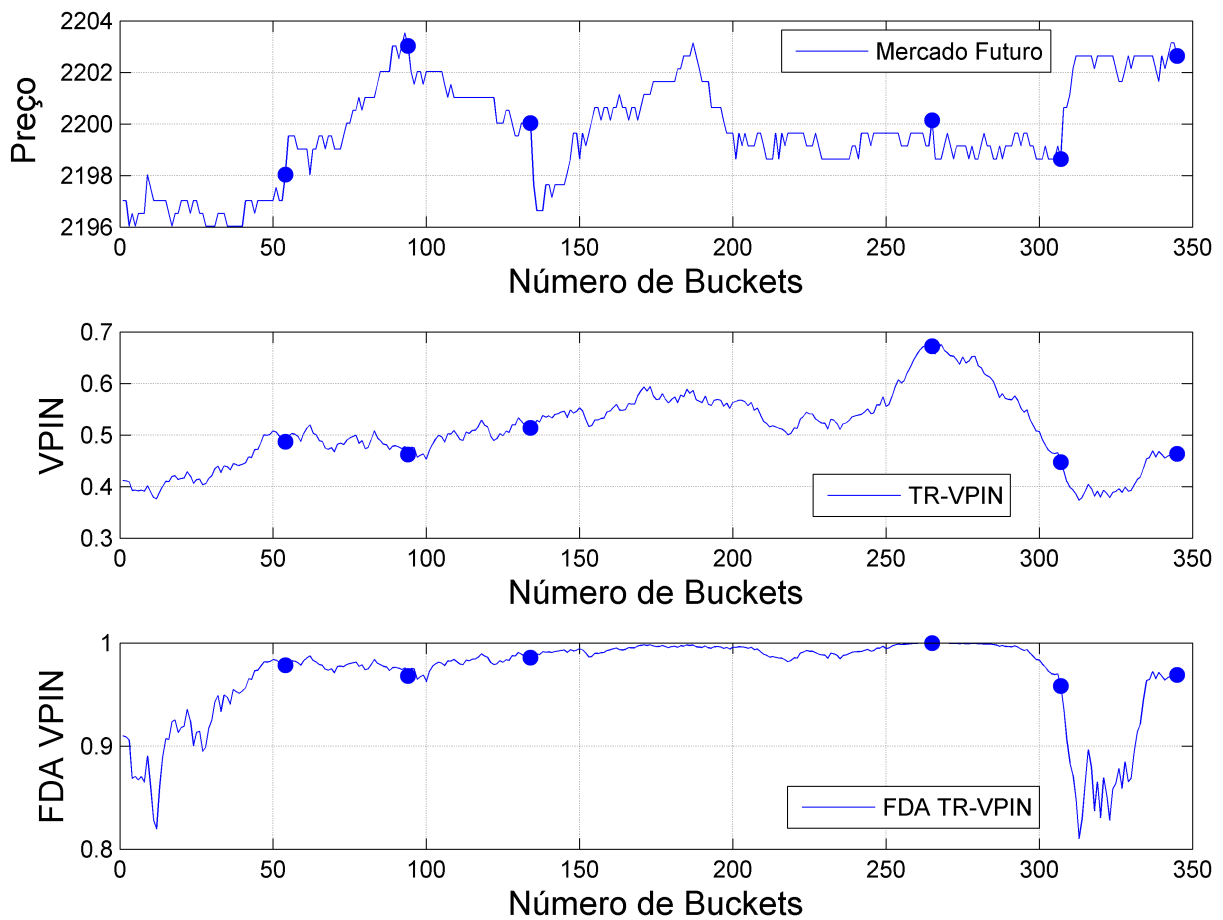


Figura 6.16: Comportamento do TR-VPIN - 6 dias úteis - dia 26/09/2012 até o dia 02/11/2012 no dólar futuro.

$$\rho(\ln(VPIN_t), |\frac{P_{t+1}}{P_t} - 1|) = -0.0329$$

EC-VPIN

Teste de Jarque-Bera ($\alpha = 5\%$)	Rejeitado
Média	0.1668
Mediana	0.1642
Desvio Padrão	0.0281
Simetria	0.5176
Curtose	3.2444
Máximo	0.2870

Vemos na Figura 6.24 um indicação mostrando que o EC-VPIN atinge o valor máximo do dia 02/03/2012. Mais precisamente no *bucket* que inicia: 02/03/2012 16:06:32.976 e termina : 02/03/2012 16:12:01.037;

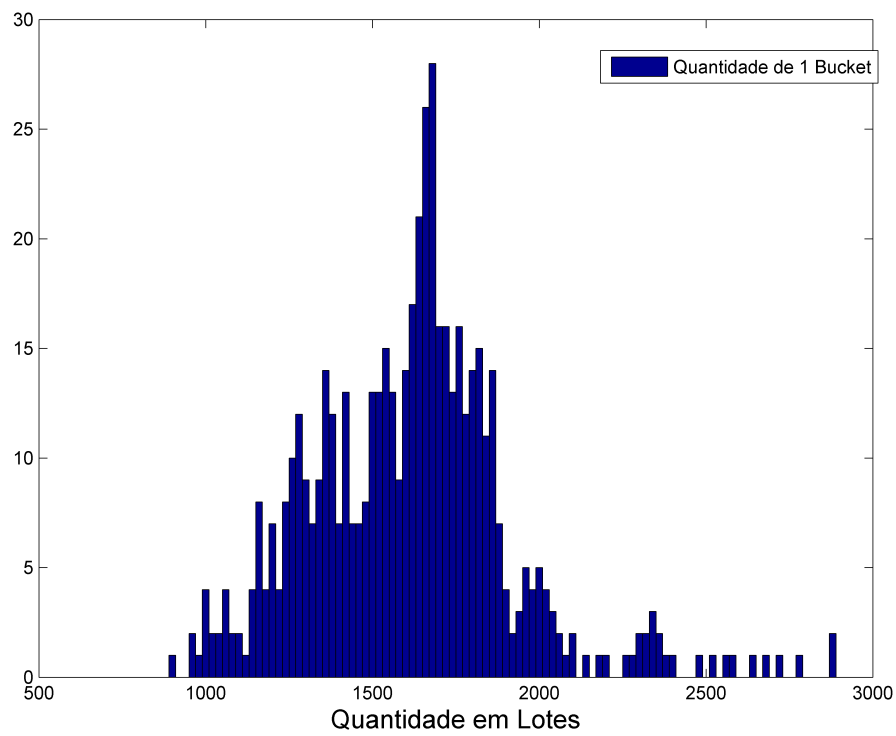


Figura 6.17: Histograma da quantidade de 1 *bucket* no índice bovespa futuro.

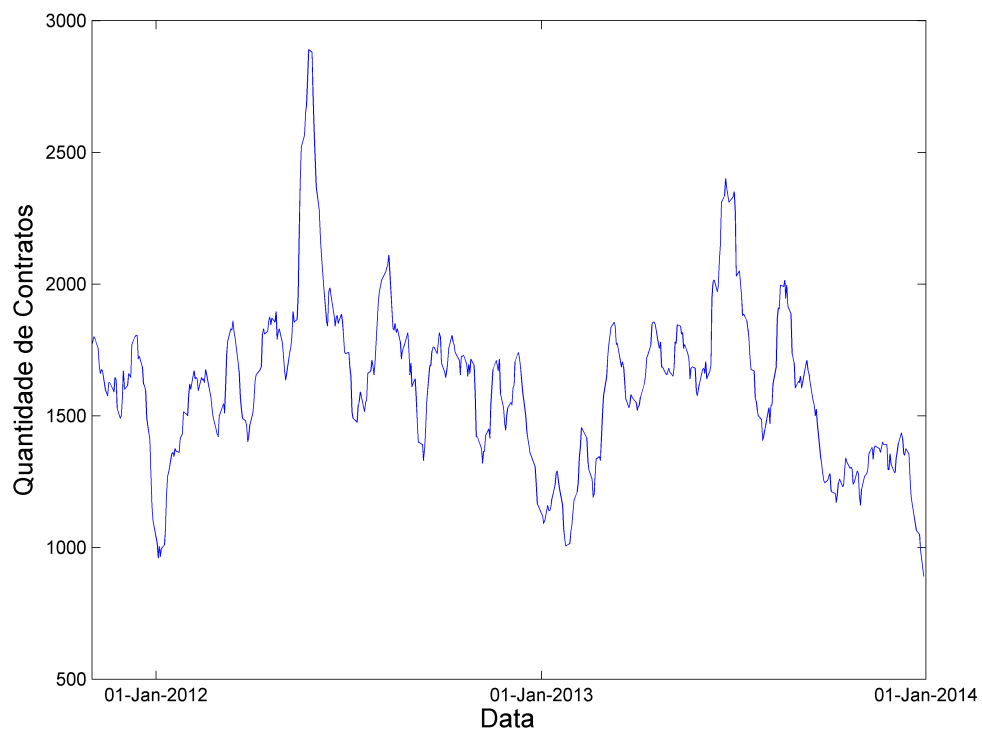


Figura 6.18: Gráfico da quantidade de 1 *bucket* no índice bovespa futuro.

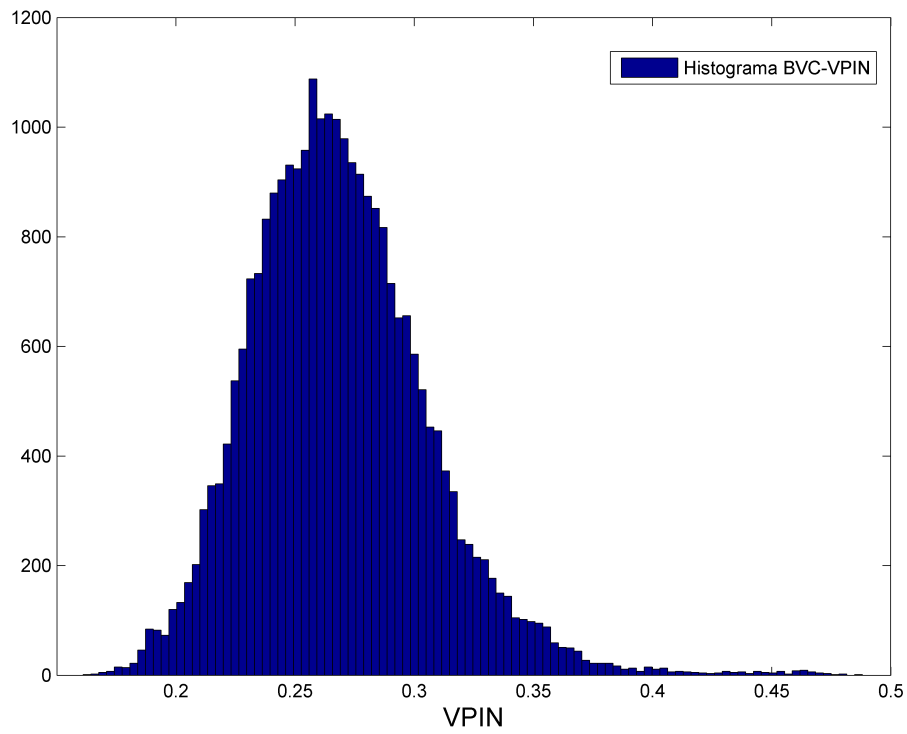


Figura 6.19: histograma do BVC-VPIN no índice bovespa futuro.

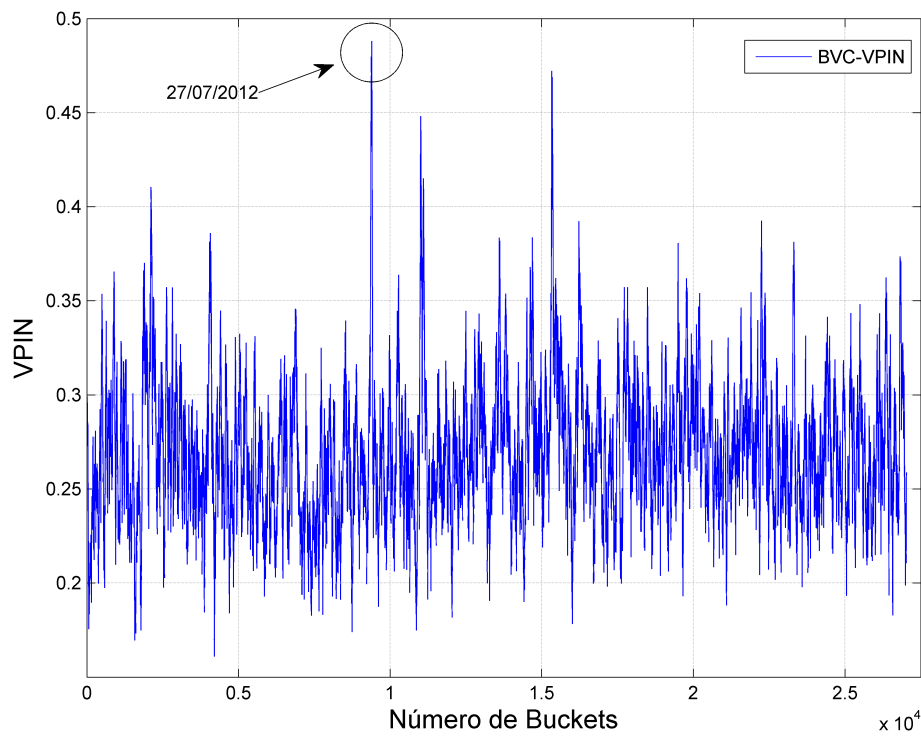


Figura 6.20: Máximo atingido pelo BVC-VPIN no índice bovespa futuro.

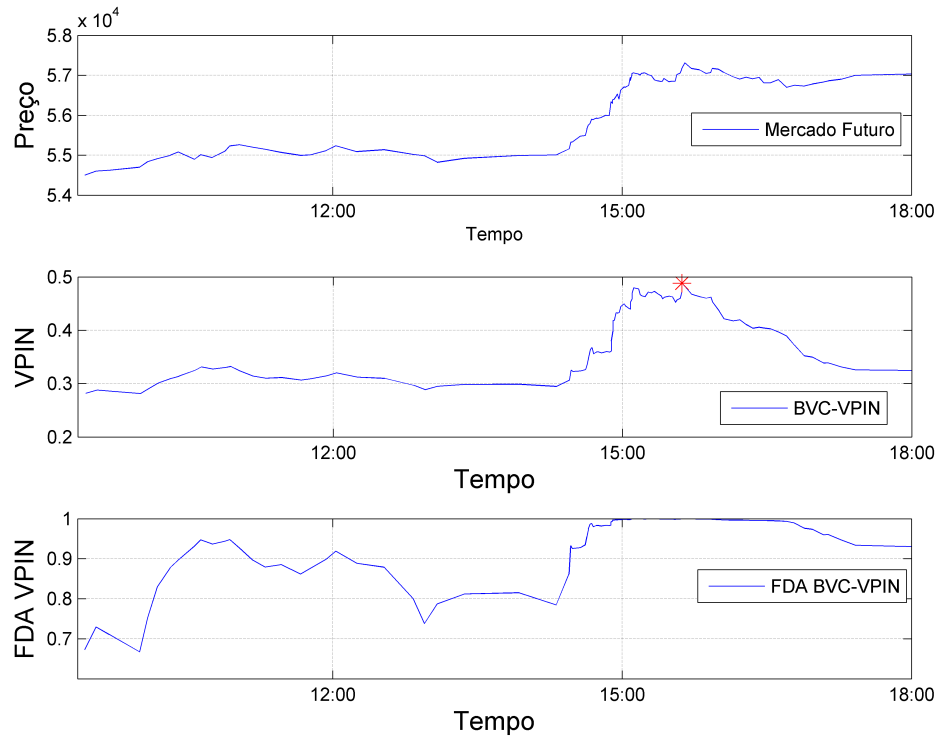


Figura 6.21: Comportamento intradiário do BVC-VPIN do dia 27/07/2012 no índice bovespa futuro.

O coeficiente de correlação de Pearson do VPIN e do logaritmo do VPIN com o módulo do retorno do *bucket* subsequente foram calculados:

$$\rho(VPIN_t, |\frac{P_{t+1}}{P_t} - 1|) = -0.0857$$

$$\rho(\ln(VPIN_t), |\frac{P_{t+1}}{P_t} - 1|) = -0.1023$$

TR-VPIN

Teste de Jarque-Bera ($\alpha = 5\%$)	Rejeitado
Média	0.2144
Mediana	0.2128
Desvio Padrão	0.0292
Simetria	0.5125
Curtose	3.6511
Máximo	0.3777

Vemos na Figura 6.28 um indicação mostrando que o TR-VPIN atinge o valor máximo do dia 29/12/2011. Mais precisamente no *bucket* que inicia: 29/12/2011 15:52:08.976 e termina : 29/12/2011 15:59:17.002;

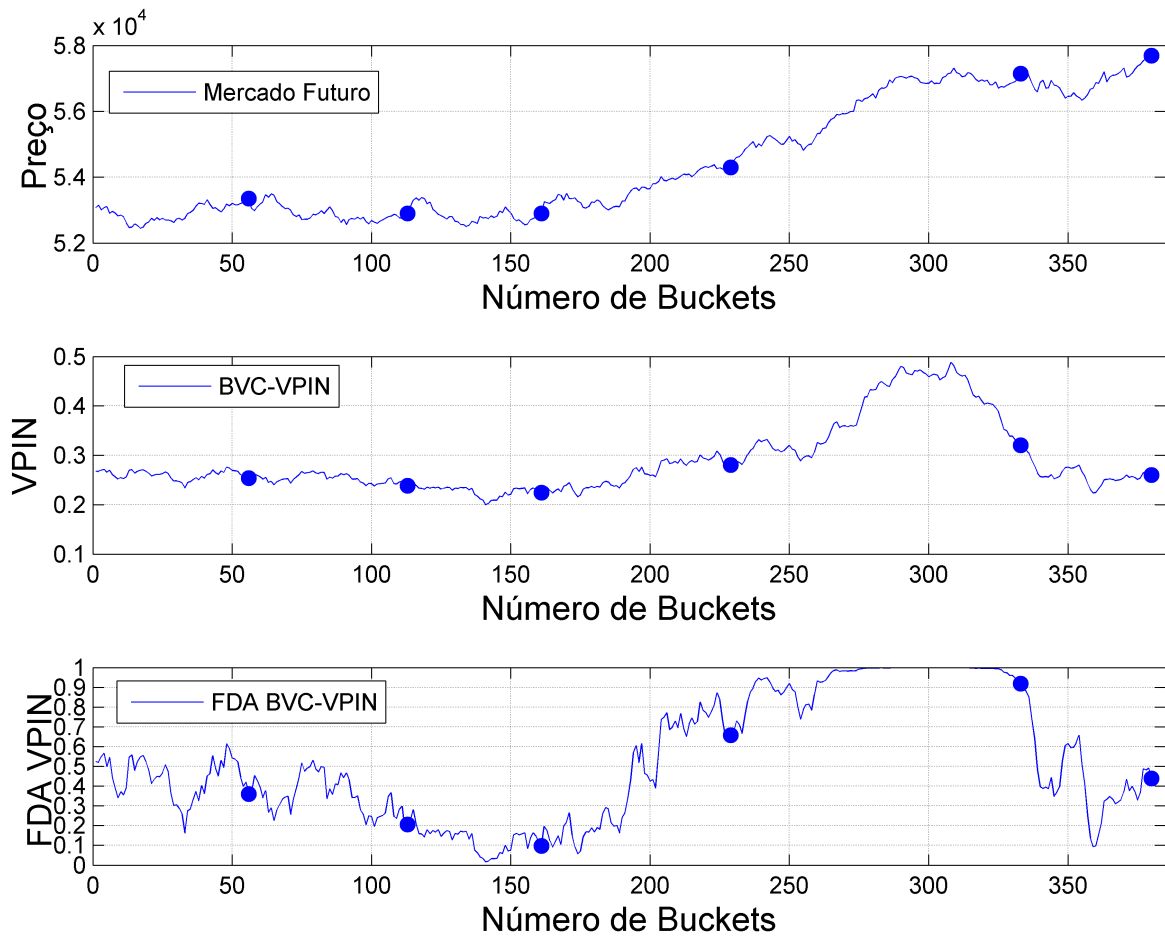


Figura 6.22: Comportamento do BVC-VPIN - dia 23/07/2012 até o dia 30/07/2012 no índice bovespa futuro.

O coeficiente de correlação de Pearson do VPIN e do logaritmo do VPIN com o módulo do retorno do *bucket* subsequente foram calculados:

$$\rho(VPIN_t, |\frac{P_{t+1}}{P_t} - 1|) = -0.1114$$

$$\rho(\ln(VPIN_t), |\frac{P_{t+1}}{P_t} - 1|) = -0.1147$$

6.3 Taxa Juros DI

Foi usado o contrato futuro de DI de um dia negociado na BM&F Bovespa. Conforme foi mostrado no Capítulo 5, o mercado futuro de DI possui um leilão, com um volume considerado relevante, sendo realizado num único preço. Esse leilão causa uma distorção na medida do TR-VPIN e do BVC-VPIN. Para o caso do TR-VPIN, será classificado esse volume inteiro do leilão como agressão de compra ou de venda, gerando um salto na medida do VPIN. Para o caso do BVC-VPIN, ocorrerão vários buckets sem alteração de preço, isso gera uma diminuição instantânea na medida do VPIN. Tendo em vista isso, retiramos o leilão no nosso cálculo do VPIN. Consideraremos todos os negócios do dia anteriores ao leilão.

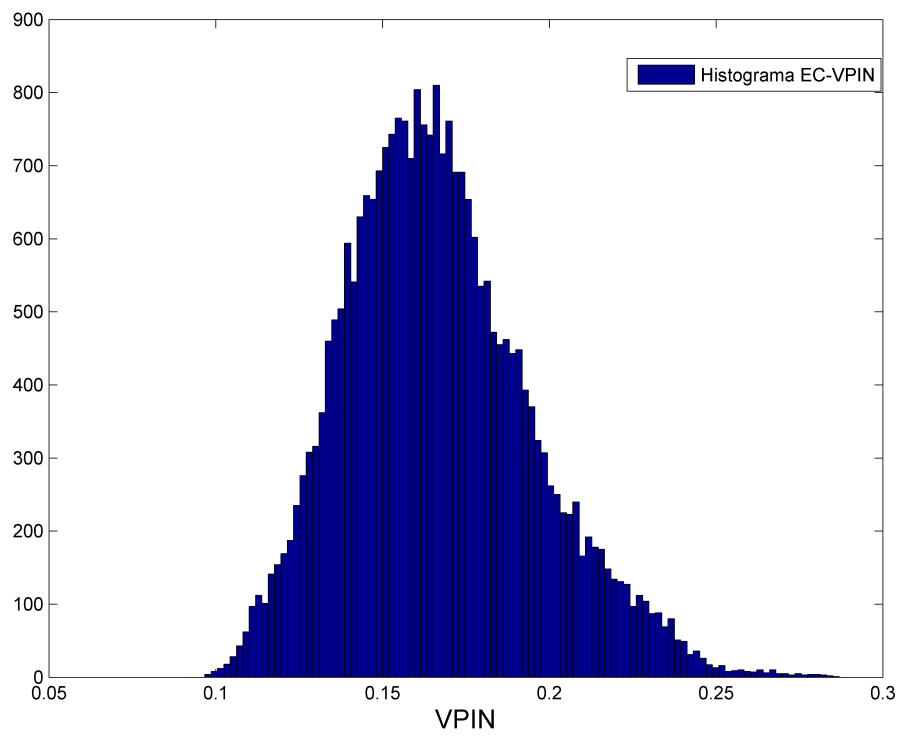


Figura 6.23: Histograma do EC-VPIN no índice bovespa futuro.

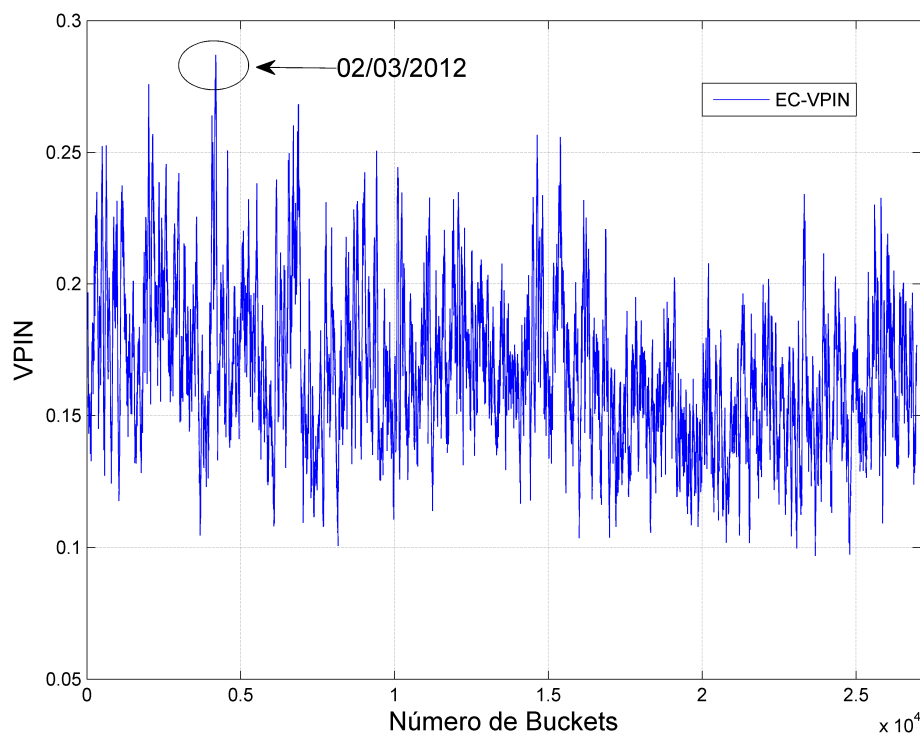


Figura 6.24: Máximo atingido pelo EC-VPIN no índice bovespa futuro.

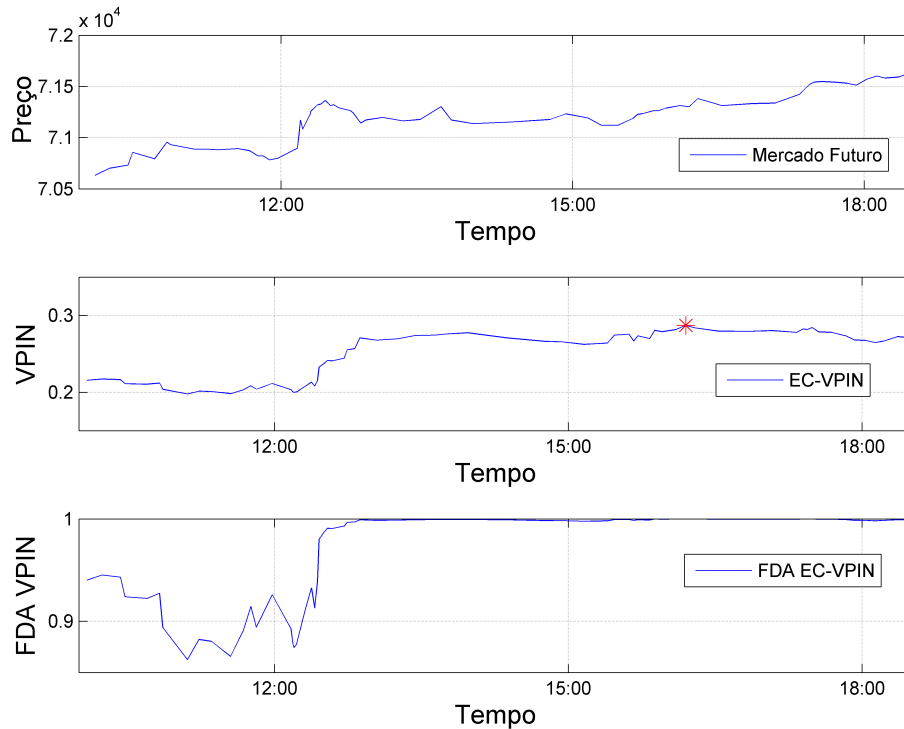


Figura 6.25: Comportamento intradiário do EC-VPIN do dia 02/03/2012 no índice bovespa futuro.

Mercado	DI futuro
Início da Amostra	01/11/2011
Fim da Amostra	30/12/2013
Número de dias de negociação	535
Número Total de <i>Buckets</i>	27766

A Figura 6.31 mostra o histograma da quantidade de 1 *bucket* e a Figura 6.32 mostra o valor da quantidade de 1 *bucket* no decorrer dos 535 dias amostrados. Em períodos de muita liquidez, o estimador de volume diário (a média de volume dos últimos 10 dias) acaba aumentando e gera um aumento nessa quantidade. Todavia, em momentos de pouca liquidez o estimador de volume diário acaba diminuindo e, conseqüentemente, faz a quantidade de 1 *bucket* diminuir.

BVC-VPIN

Teste de Jarque-Bera ($\alpha = 5\%$)	Rejeitado
Média	0.3483
Mediana	0.3453
Desvio Padrão	0.0836
Simetria	0.1432
Curtose	4.0403
Máximo	0.7329

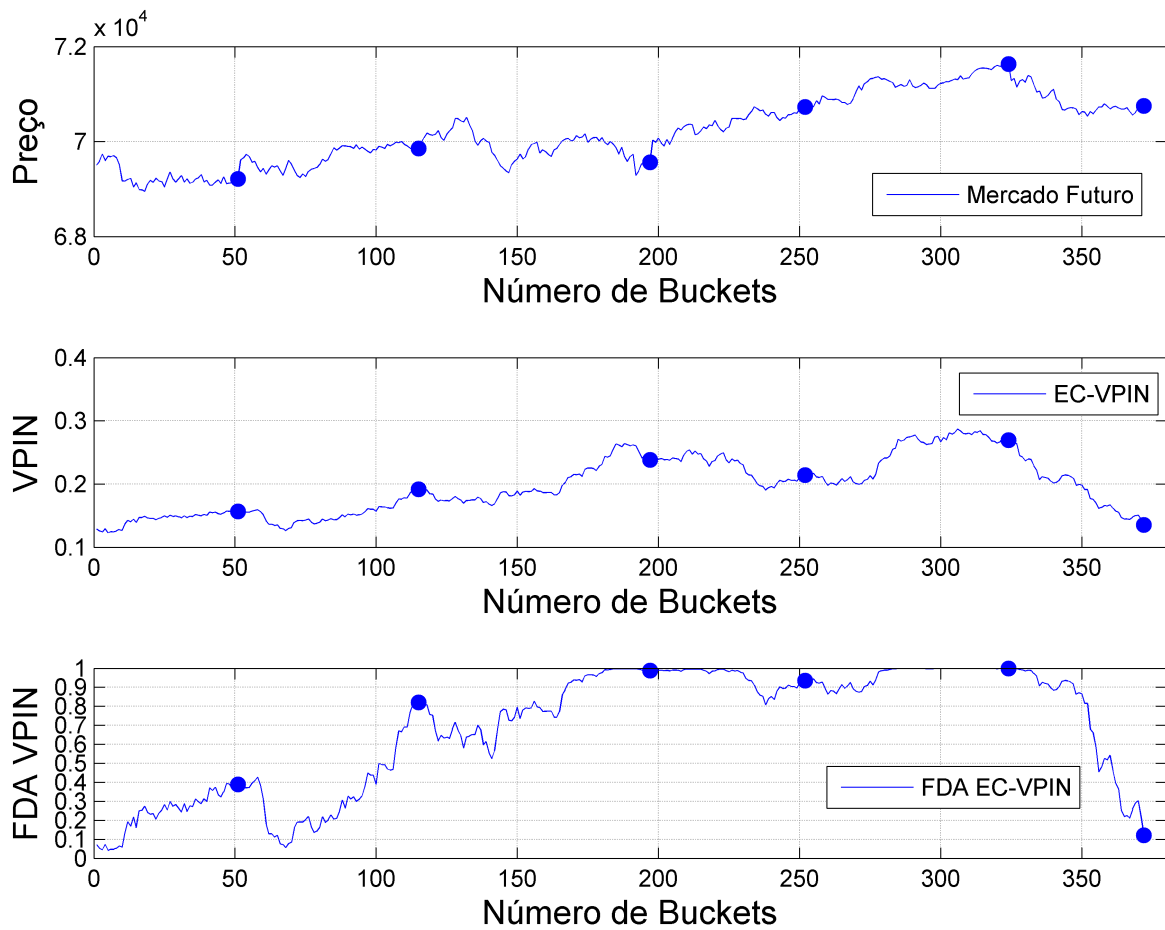


Figura 6.26: Comportamento do EC-VPIN - 6 dias úteis - dia 27/02/2012 até o dia 05/03/2012 no índice bovespa futuro.

Vemos na Figura 6.34 um indicação mostrando que o BVC-VPIN atinge o valor máximo do dia 06/01/2012. Mais precisamente no *bucket* que inicia: 06/01/2012 15:53:00.000 e termina : 06/01/2012 15:54:00.000;

O coeficiente de correlação de Pearson do VPIN e do logaritmo do VPIN com o módulo do retorno do *bucket* subsequente foram calculados:

$$\rho(VPIN_t, |\frac{P_{t+1}}{P_t} - 1|) = -0.0248$$

$$\rho(\ln(VPIN_t), |\frac{P_{t+1}}{P_t} - 1|) = -0.0147$$

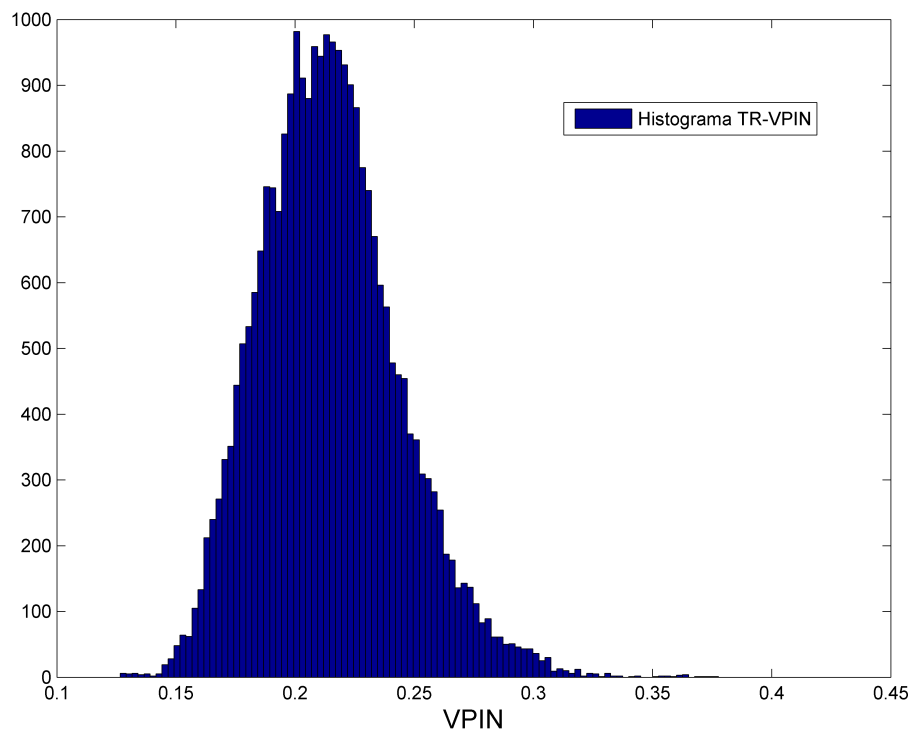


Figura 6.27: Histograma do TR-VPIN no índice bovespa futuro.

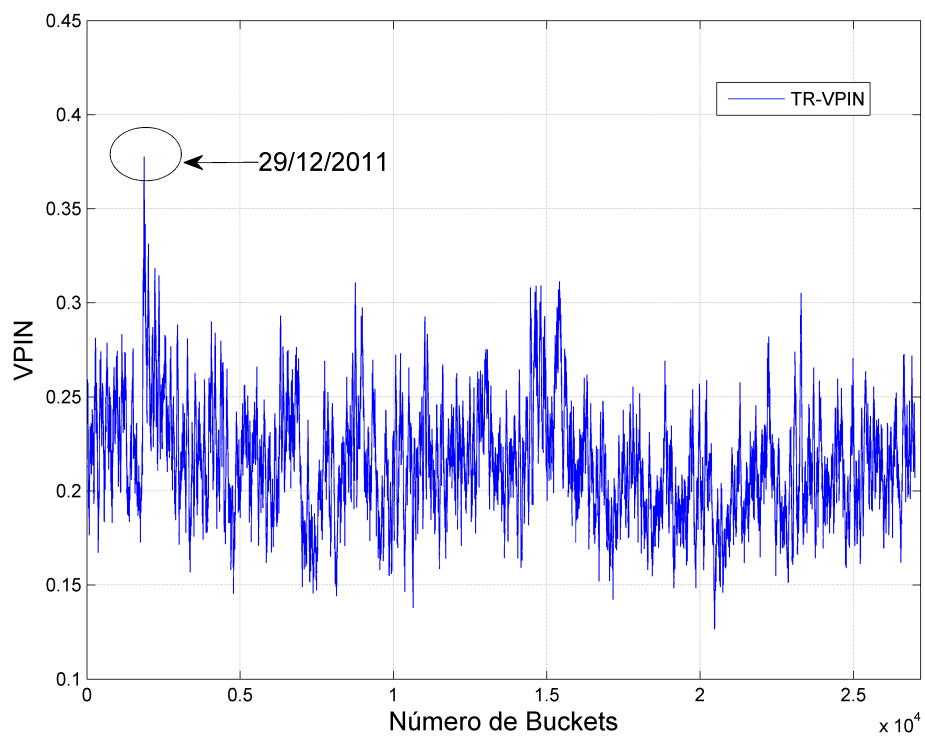


Figura 6.28: Máximo atingido pelo TR-VPIN no índice bovespa futuro.

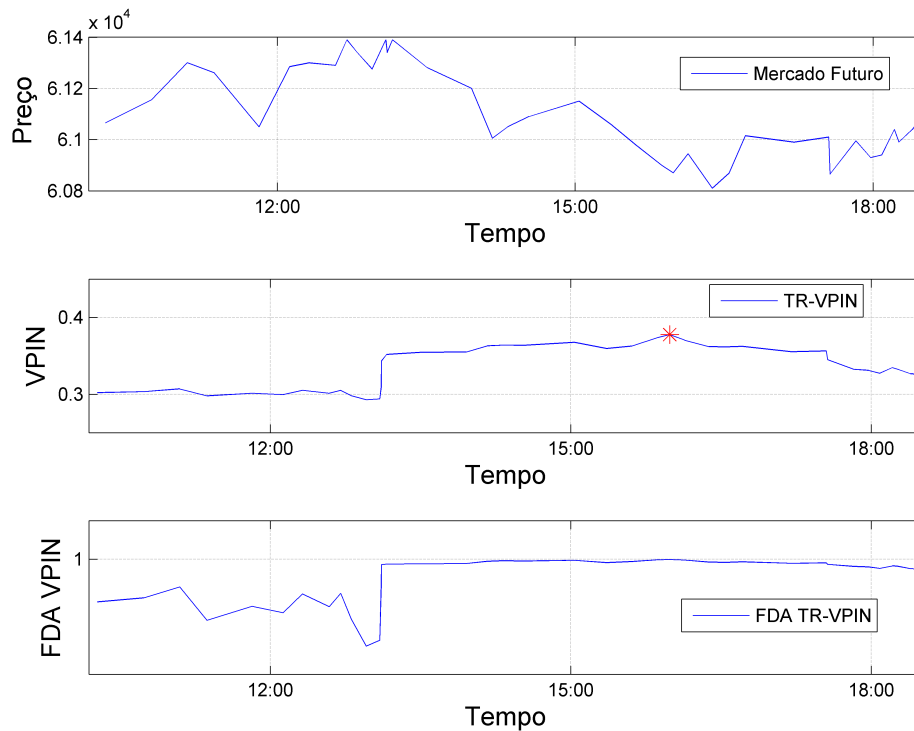


Figura 6.29: Comportamento intradiário do TR-VPIN do dia 29/12/2011 no índice bovespa futuro.

EC-VPIN

Teste de Jarque-Bera ($\alpha = 5\%$)	Rejeitado
Média	0.6076
Mediana	0.6111
Desvio Padrão	0.1098
Simetria	-0.0406
Curtose	2.5956
Máximo	0.9704

Vemos na Figura 6.38 um indicação mostrando que o EC-VPIN atinge o valor máximo do dia 06/01/2012. Mais precisamente no *bucket* que inicia: 06/01/2012 15:43:25.997 e termina : 06/01/2012 15:43:27.034;

O coeficiente de correlação de Pearson do VPIN e do logaritmo do VPIN com o módulo do retorno do *bucket* subsequente foram calculados:

$$\rho(VPIN_t, |\frac{P_{t+1}}{P_t} - 1|) = -0.2543$$

$$\rho(\ln(VPIN_t), |\frac{P_{t+1}}{P_t} - 1|) = -0.2661$$

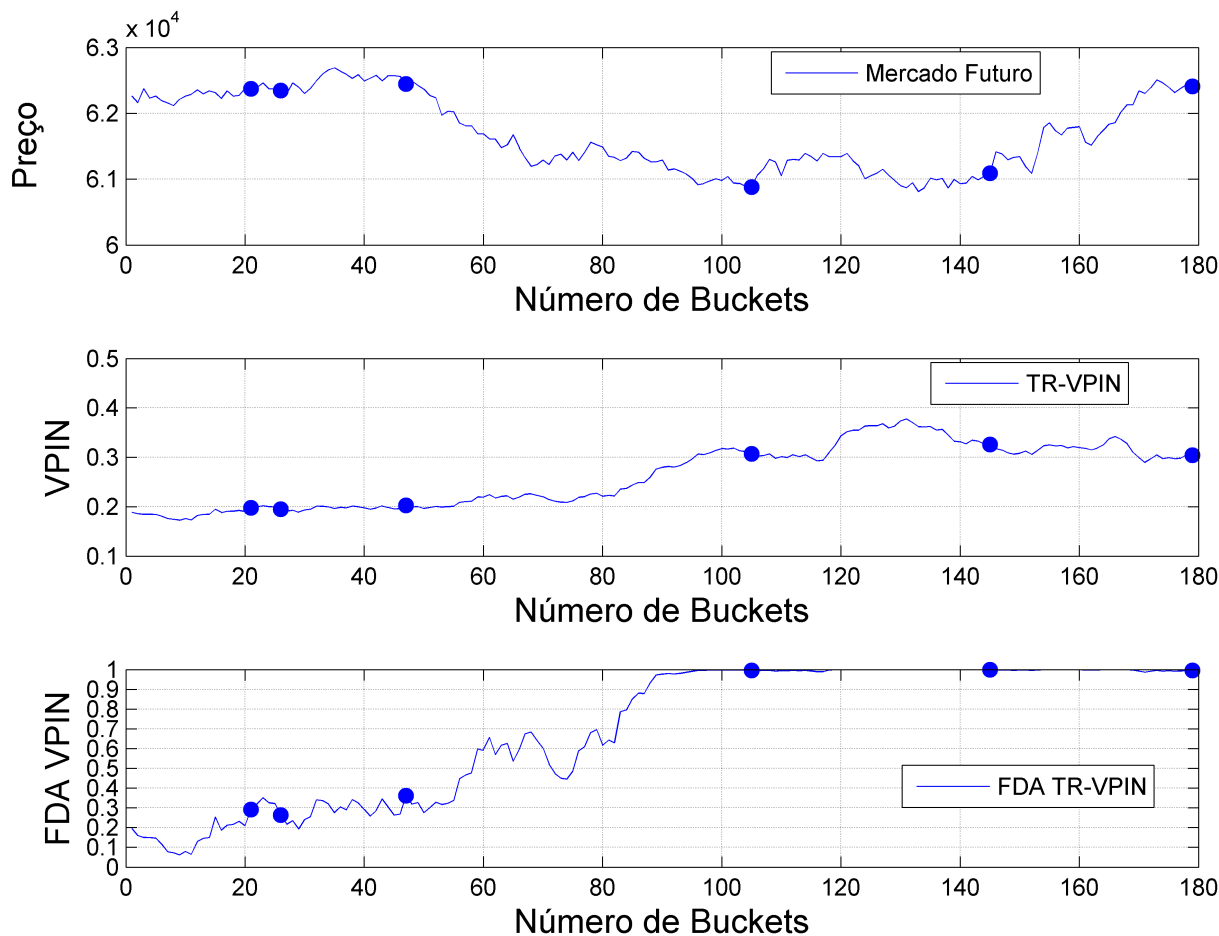


Figura 6.30: Comportamento do TR-VPIN - 6 dias úteis - dia 23/12/2011 até o dia 02/01/2012 no índice bovespa futuro.

TR-VPIN

Teste de Jarque-Bera ($\alpha = 5\%$)	Rejeitado
Média	0.7411
Mediana	0.7473
Desvio Padrão	0.0885
Simetria	-0.3696
Curtose	3.2004
Máximo	1

Vemos na Figura 6.42 um indicação mostrando que o TR-VPIN atinge o valor máximo do dia 31/01/2012. Mais precisamente no *bucket* que inicia: 31/01/2012 15:08:23.971 e termina : 31/01/2012 15:08:23.971;

O coeficiente de correlação de Pearson do VPIN e do logaritmo do VPIN com o módulo do retorno do *bucket* subsequente foram calculados:

$$\rho(VPIN_t, |\frac{P_{t+1}}{P_t} - 1|) = -0.2358$$

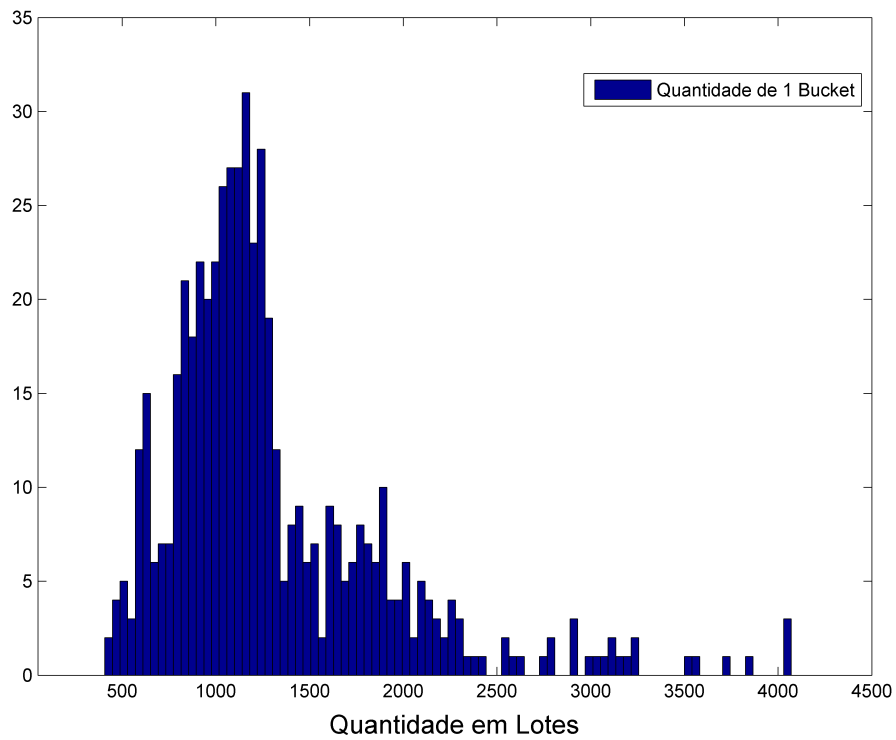


Figura 6.31: Histograma da quantidade de 1 *bucket* no DI futuro.

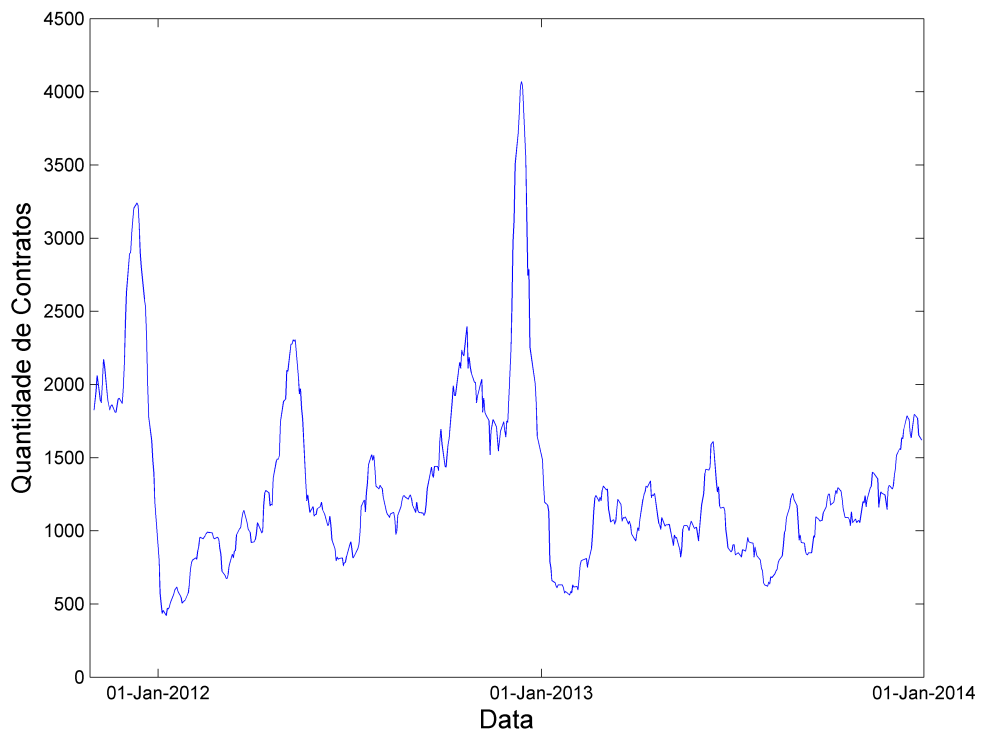


Figura 6.32: Gráfico da quantidade de 1 *bucket* no DI futuro.

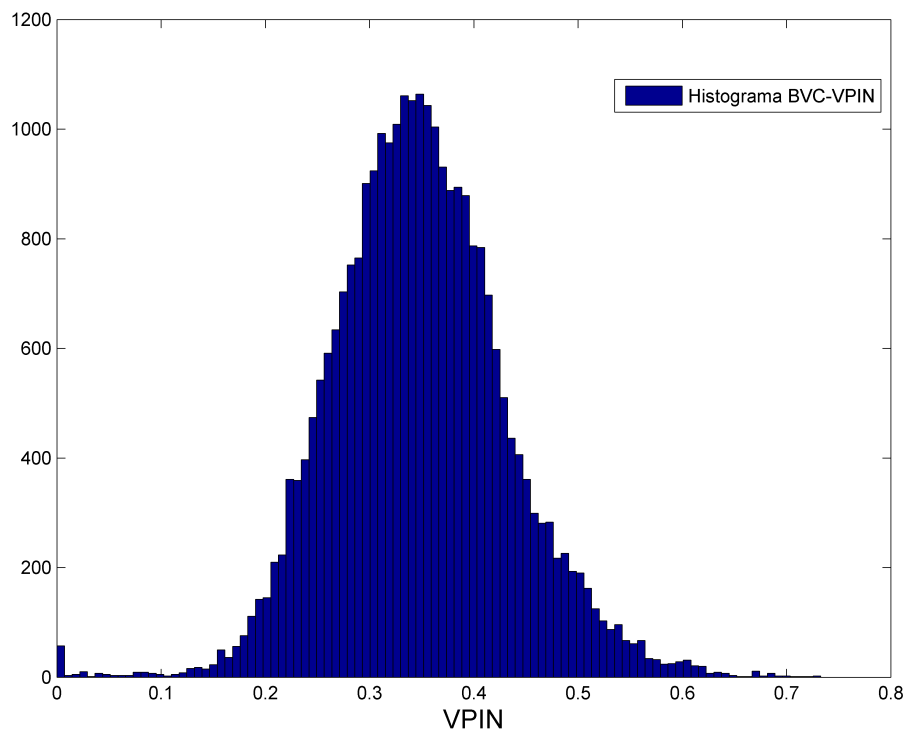


Figura 6.33: Histograma do BVC-VPIN no DI futuro.

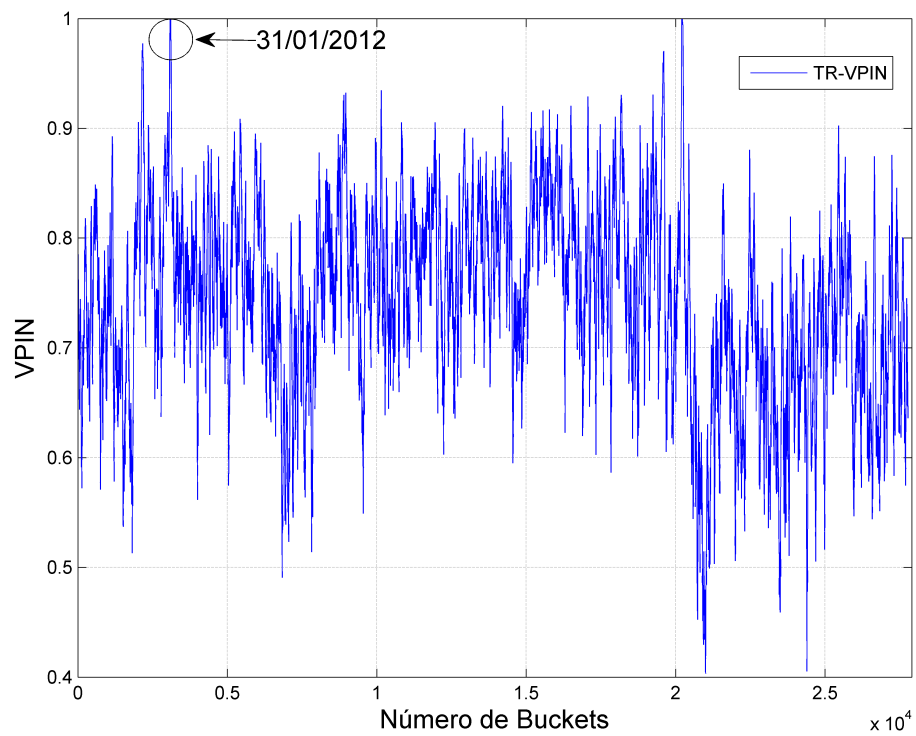


Figura 6.34: Máximo atingido pelo BVC-VPIN no DI futuro.

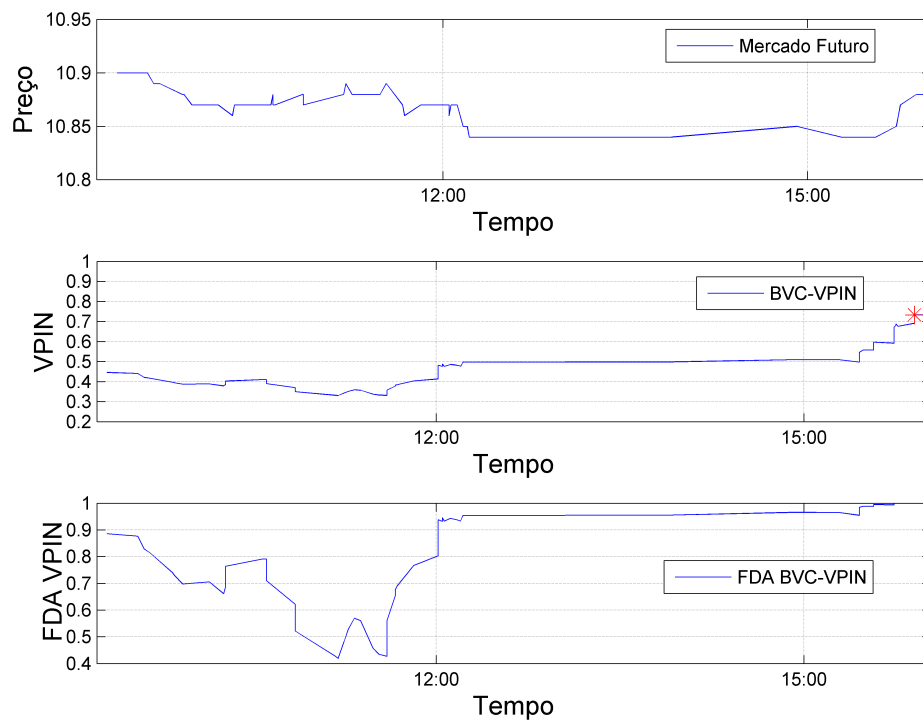


Figura 6.35: Comportamento intradiário do BVC-VPIN do dia 06/01/2012 no DI futuro.

$$\rho(\ln(VPIN_t), |\frac{P_{t+1}}{P_t} - 1|) = -0.2283$$

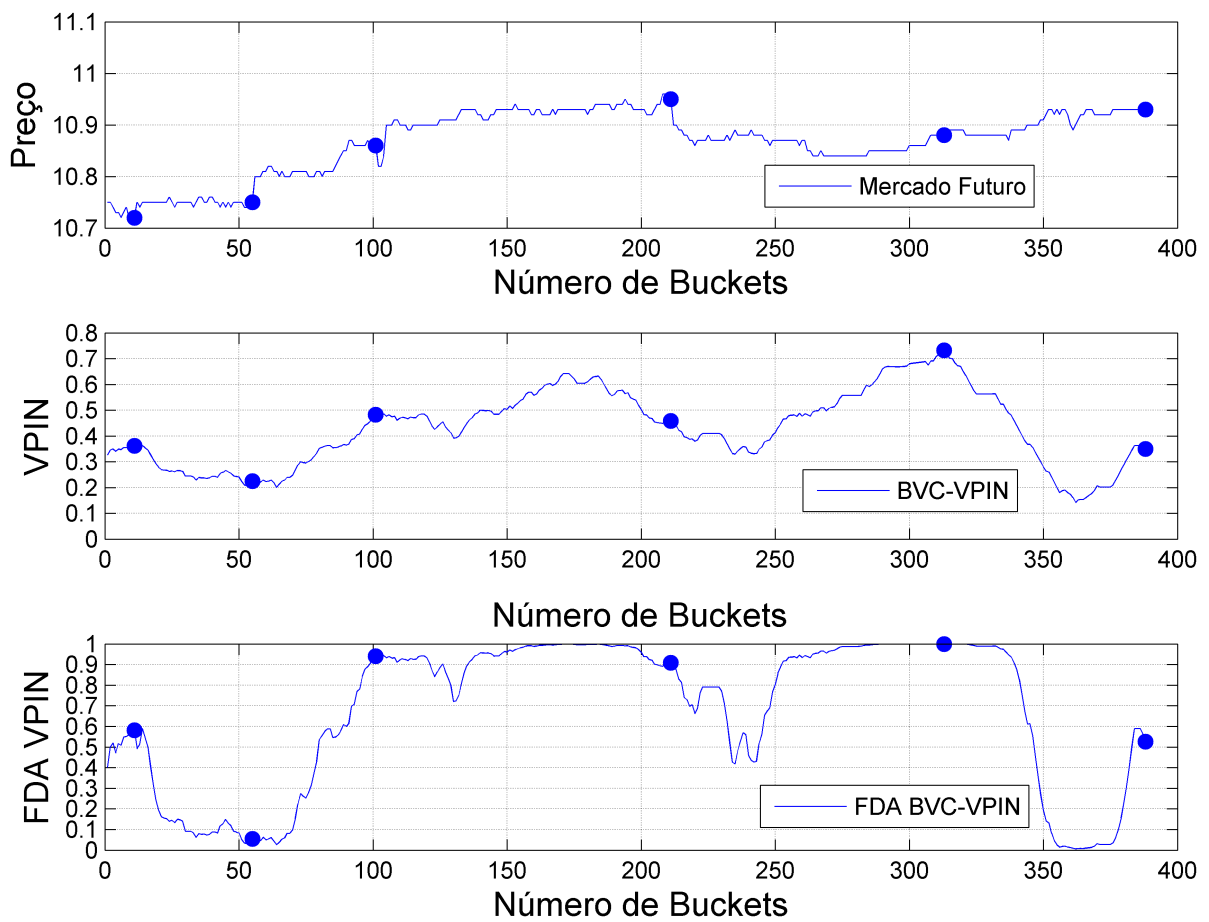


Figura 6.36: Comportamento do BVC-VPIN - 6 dias úteis - dia 02/01/2011 até o dia 09/01/2012 no DI futuro.

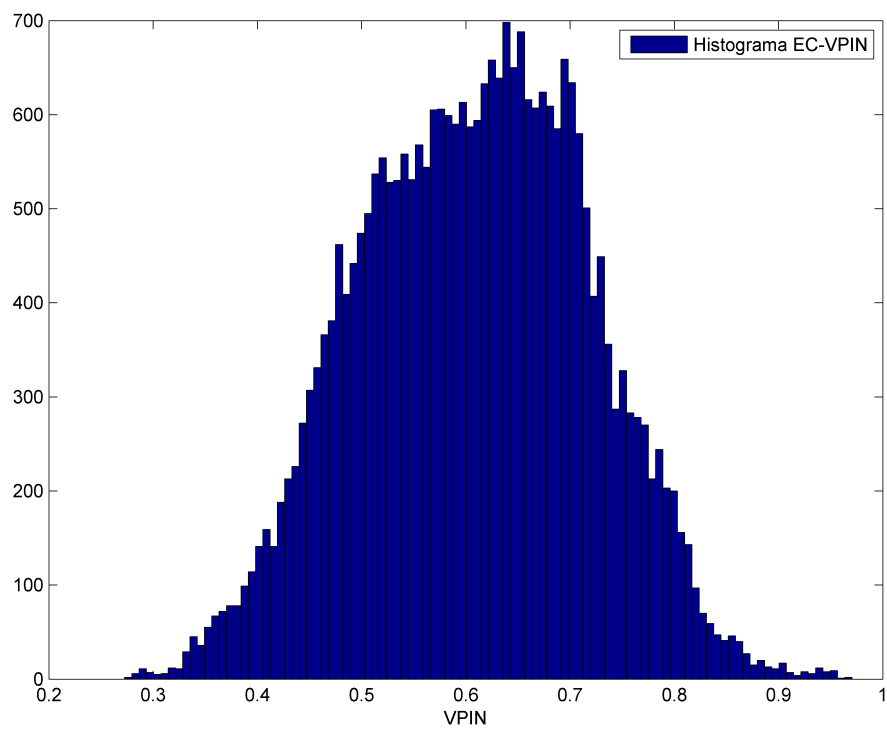


Figura 6.37: Histograma do EC-VPIN no DI futuro.

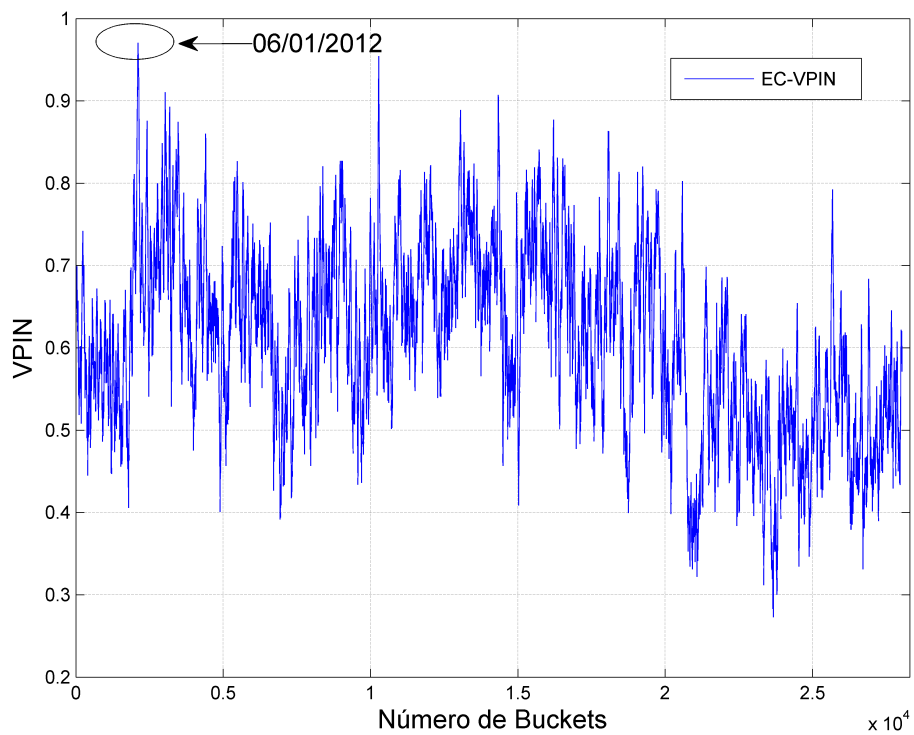


Figura 6.38: Máximo atingido pelo EC-VPIN no DI futuro.

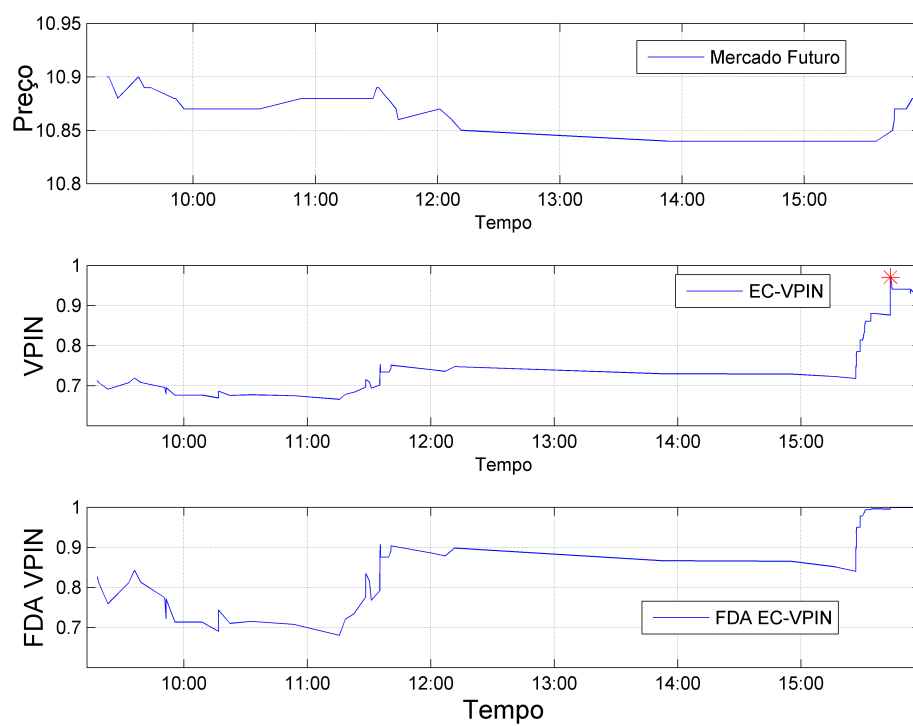


Figura 6.39: Comportamento intradiário do EC-VPIN do dia 06/01/2012 no DI futuro.

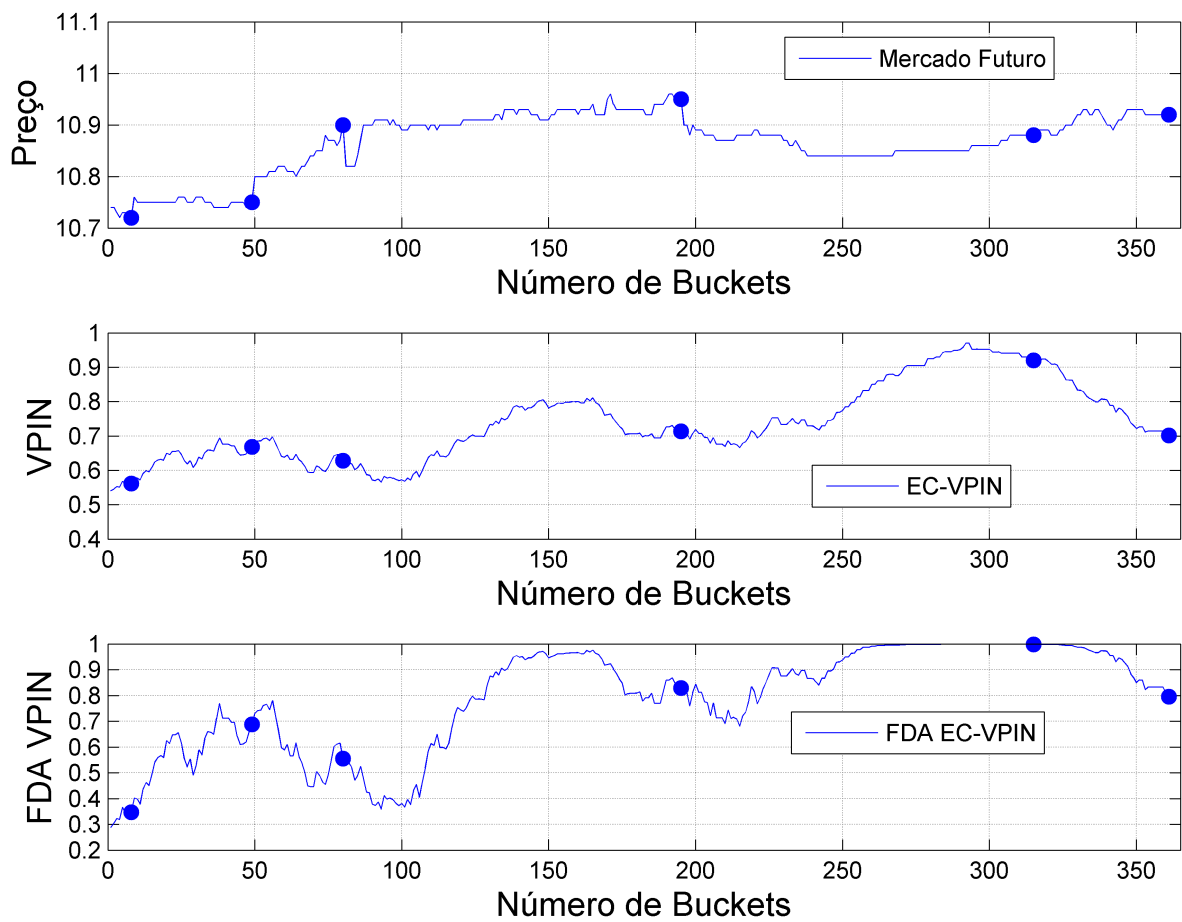


Figura 6.40: Comportamento do EC-VPIN - 6 dias úteis - dia 02/01/2011 até o dia 09/01/2012 no DI futuro.

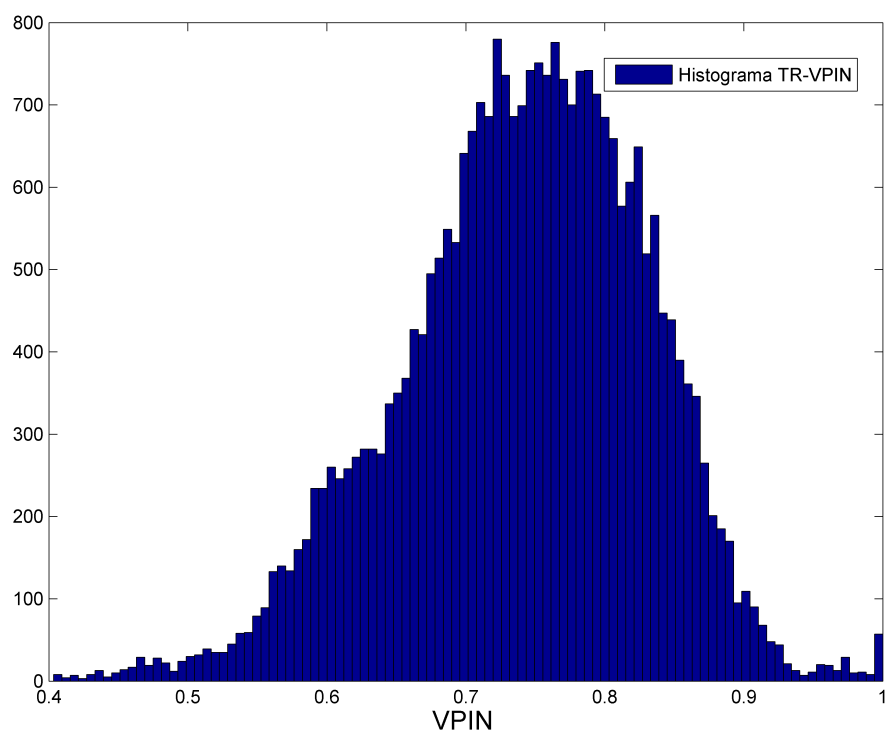


Figura 6.41: Histograma do TR-VPIN no DI futuro.

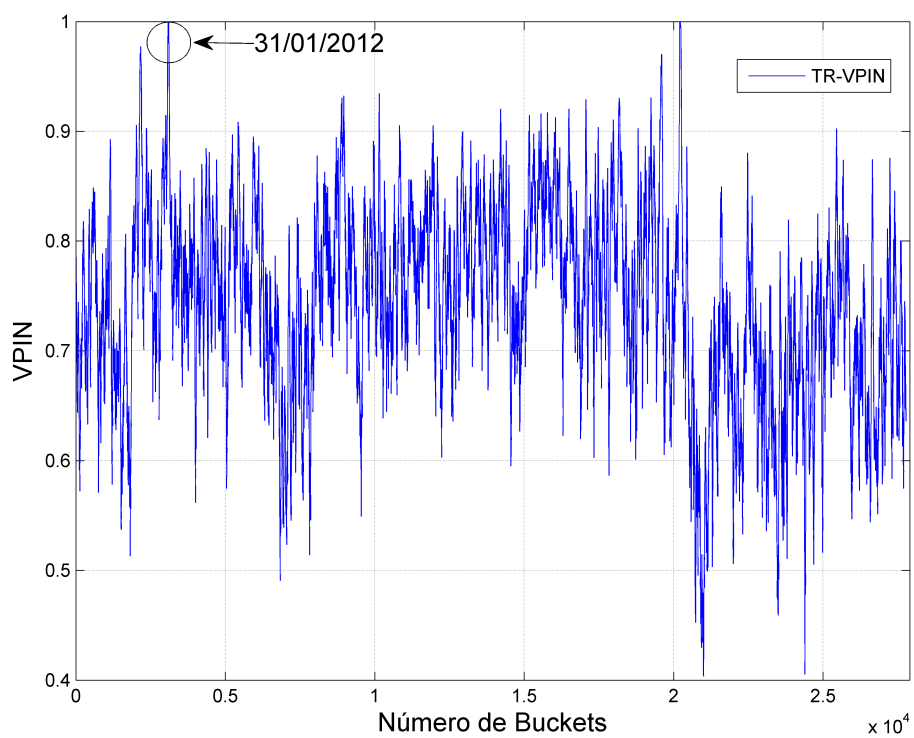


Figura 6.42: Máximo atingido pelo TR-VPIN no DI futuro.

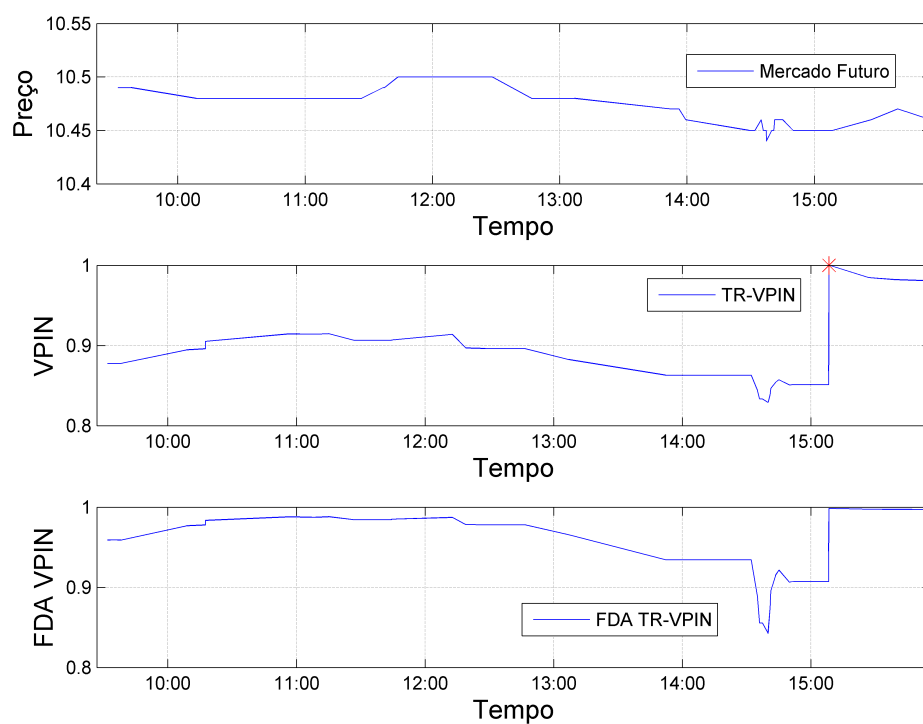


Figura 6.43: Comportamento intradiário do TR-VPIN do dia 31/01/2012 no DI futuro.

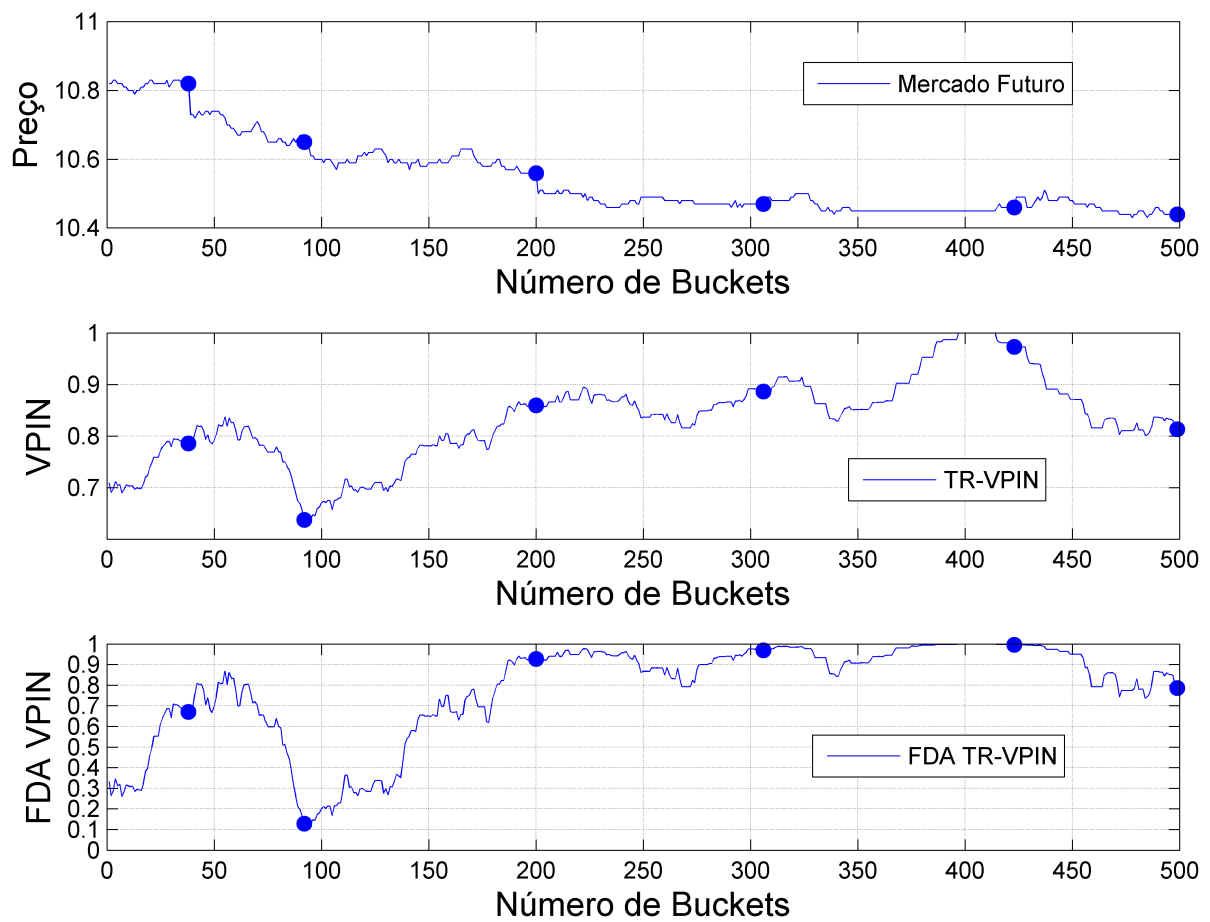


Figura 6.44: Comportamento do TR-VPIN - 6 dias úteis - dia 24/01/2011 até o dia 01/02/2012 no DI futuro.

Capítulo 7

Toxicidade em Eventos

7.1 Programa de Leilões de Swap Cambial - Banco Central do Brasil

No fim do primeiro semestre de 2013, os Estados Unidos anunciaram que retirariam parte dos estímulos financeiros colocados na época da crise de imobiliária de 2008. Por conta disso, ocorreu uma grande desvalorização do real frente ao dólar devido à fuga de capital estrangeiro do país (movimento análogo em quase todos os países emergentes). Tendo em vista esse cenário, a partir do mês de junho de 2013, o Banco Central do Brasil começou a intervir no mercado de câmbio com mais frequência. De junho de 2013 até dezembro de 2013 são 146 intervenções com leilões de swap. Enquanto de novembro de 2011 até maio de 2013, o Banco Central interveio 34 vezes no mercado de cambio com leilões de swap. O Banco central do Brasil anunciou um programa de leilões de swap cambial após o fechamento do mercado do dia 22/08/2013. Esse programa pré-estabelecia leilões de swap cambial no valor de 500 milhões de dólares de segunda-feira a quinta-feira e leilões de linha no valor de 1 bilhão de dolares toda sexta-feira até o fim de dezembro 2013. Esses leilões anunciados funcionam como uma venda no mercado do contrato futuro, pressionando o mercado para uma valorização do real frente ao dólar. Para a consulta de datas com swap cambial ver a referência [10].

Para fazermos o estudo de cada período, as tabelas abaixo mostram estatísticas do mercado futuro de dólar antes e durante as intervenções gorvenamentais.

Síntese das principais estatísticas do período anterior e posterior ao anúncio do Banco Central.

Amostra	Antes	Depois
Número de intervenções	34	146
Tamanho (N_1 e N_2)	19885	7297
Média (μ_1 e μ_2)	0.2419	0.2540
Mediana	0.2354	0.2528
Desvio Padrão (σ_1 e σ_2)	0.0455	0.0389
Simetria	0.9420	0.1775
Curtose	4.7172	2.8537
Máximo	0.4849	0.3954

Note que a mediana e a média aumentaram a partir do momento em que o governo começou a fazer intervenção. Fazendo o seguinte teste de Hipótese para as médias:

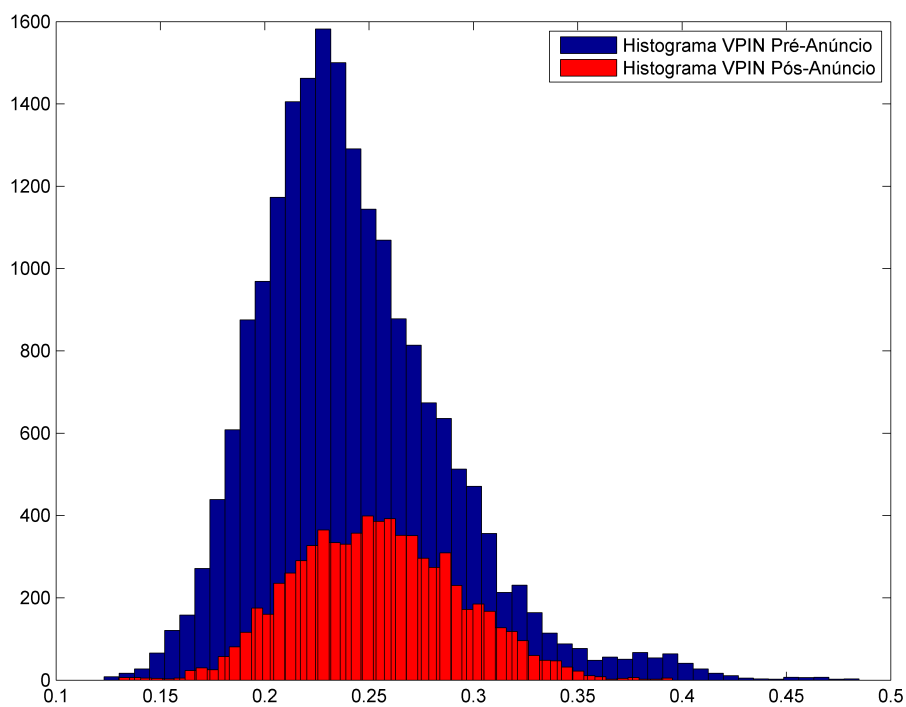
Teste de Hipótese: Bicaudal, $\alpha = 1\%$, $\nu = 27180$.

$$H_0 : \mu_1 = \mu_2$$

$$H_1 : \mu_1 \neq \mu_2$$

$$T = \frac{\mu_2 - \mu_1}{\sqrt{\left(\frac{\sigma_1^2}{N_1} + \frac{\sigma_2^2}{N_2}\right)}} = 21.68 > T_{\frac{\alpha}{2}} = 2.57$$

Podemos dizer que houve uma mudança na média da amostra, pois teste acima rejeita H_0 ($T > T_{\frac{\alpha}{2}}$). Além disso, a Figura 7.1 mostra visivelmente a diferença numérica de curtose entre as duas amostras:



Capítulo 8

Considerações Finais

8.1 Síntese Final

Podemos resumir os quatros principais resultados do trabalho da seguinte forma:

- A classificação de negócio pela regra do tick nos mercados estudados é estritamente mais precisa que a classificação maciça de volume;
- O mercado futuro de taxa de juros DI de um dia é o mais tóxico dos mercados estudados;
- O VPIN nos mercado estudados não é um bom preditor de retorno absoluto;
- O mercado do futuro de dólar ficou mais tóxico durante o programa de leilão de *swap* cambial do Banco Central do Brasil.

8.2 Conclusões

Conforme definições do Capítulo 1, a teoria de microestrutura de mercado divide os negociadores, de acordo com o seu nível de informação, em dois grupos: negociadores informados e negociadores não-informados. Um mercado é chamado de tóxico, quando existe uma alta probabilidade dos provedores de liquidez estarem transacionando com negociadores informados. Um negócio é dito informado, quando um dos negociadores dele é informado.

Mostramos três tipos de classificação de negócios no Capítulo 2. Foi feito um estudo de precisão em cada mercado estudado mostrando que a classificação TR classifica com mais precisão que a classificação BVC. Isso também foi achado no mercado de ações da NASDAQ no Trabalho [14] e no futuro de *S&P* 500 no Trabalho [12]. Segue a Tabela achada:

Precisão Classificação de Negócio:

Mercado	DOL	IND	DI
η^{TR}	79.57%	75.84%	80.50%
η^{BVC}	62.34%	71.68%	55.86%

Mostramos no Capítulo 6 que o mercado de taxa de juros DI de um dia se comportou como o mais tóxico independentemente do tipo de classificação escolhida. Conforme pode ser visto abaixo, a média e a mediana do VPIN da taxa de juros DI é maior que a dos outros mercados independentemente da classificação de negócio.

Média de VPIN:

Mercado	DOL	IND	DI
$VPIN_{BVC}$	0.2451	0.2686	0.3483
$VPIN_{EC}$	0.2459	0.1668	0.6076
$VPIN_{TR}$	0.3197	0.2144	0.7411

Mediana de VPIN:

Mercado	DOL	IND	DI
$VPIN_{BVC}$	0.2398	0.2656	0.3453
$VPIN_{EC}$	0.2350	0.1642	0.6111
$VPIN_{TR}$	0.3095	0.2128	0.7473

Desvio Padrão do VPIN:

Mercado	DOL	IND	DI
$VPIN_{BVC}$	0.0441	0.0373	0.0836
$VPIN_{EC}$	0.0611	0.0281	0.1098
$VPIN_{TR}$	0.0661	0.0292	0.0885

Pode ser constatado o pequeno número de formadores de mercado presentes no mercado de taxa de juros, em relação ao dólar e ao índice bovespa, observando as quantidades, as corretoras e as mudanças no livro de ofertas eletrônico das ordens limitadas. Isso está de acordo com o argumento mostrado de que o mercado de taxa de juros é o mais tóxico, pois a existência de uma maior probabilidade de o negociador ser informado, gera um valor esperado de ganho menor para os fornecedores de liquidez.

Foi visto no Capítulo 6 que a correlação do $VPIN_{BVC}$ (e do logaritmo neperiano) com o retorno absoluto do bucket subsequente pode ser considerada nula. Isso mostra que no mercado brasileiro o VPIN não é um bom preditor do retorno absoluto subsequente. Não ocorre o que foi achado no mercado futuro *S&P* 500 em [22]. Também foi achado uma significativa correlação negativa entre o $VPIN_{TR}$ e do $VPIN_{EC}$ (e do logaritmo neperiano) com o retorno absoluto do bucket subsequente. Isso está de acordo com o comportamento achado na Referência [12] para o mercado futuro *S&P* 500. Levando em conta essas correlações achadas, podemos questionar a medida VPIN quanto a robustez em relação a mudança de classificação de negócio, pois achamos para o $VPIN_{TR}$ e $VPIN_{EC}$ um comportamento diferente ao achado para o $VPIN_{BVC}$.

Por fim, no Capítulo 7 foi mostrado evidências estatísticas que o mercado de taxa de câmbio ficou mais tóxico no período em que o Banco Central do Brasil atuou de forma mais frequente. Rejeitou-se o teste de hipótese criado e mostrou-se mudanças na curtose. A Figura 7.1 é um histograma das duas amostras (antes do anúncio do programa de leilões de depois do anúncio) e

permite uma rápida visualização das mudanças ocorridas. Esse aumento de toxicidade continuando, poderemos ver crises de liquidez momentâneas com mais facilidade (devido a diminuição natural do número de formadores de mercado) e posteriormente uma geração de volatilidade induzida por esse evento de falta de liquidez.

Apêndice A

Resultados Técnicos e Notações

A.1 Códigos Classificação de Negócios e Cálculo do VPIN

Todos os códigos serão mostrados em matlab.

Entrada: A série de dados tick-by-tick foi colocada na forma de uma matriz M cada coluna representa uma informação:

- Coluna 1: Tempo da transação no formato datenum do matlab
- Coluna 2: Preço negociado
- Coluna 3: Quantidade negociada
- Coluna 4: xxxxxxxxxxxxxxxxxxxxxxxx
- Coluna 5: xxxxxxxxxxxxxxxxxxxxxxxx
- Coluna 6: Lado agressor (quando informado pela bolsa)

VPIN: Classificação da Bolsa

```
1 function [M] = Gerador_stats(M, IF_M)
2
3     %COLOCANDO A CLASSIFICACAO DO NEGOCIO DADA PELA BOLSA NA COLUNA 7
4     for i = 1:1:length(IF_M)
5         ini = IF_M(i,1);
6         fim = IF_M(i,2);
7
8         if M(ini,6) ≠ 0
9             for j = ini:1:fim
10                if M(j,6) == 1
11                    M(j,7) = -1;
12                elseif M(j,6) == 2
13                    M(j,7) = 1;
14                elseif M(j,6) == 3
15                    M(j,7) = 0;
16                end
17            end
18        end
19    end
```

```

20
21
22     for i = 1:length(IF_M)
23         %DEFININDO INICIO E FIM DO DIA
24         ini = IF_M(i,1);
25         fim = IF_M(i,2);
26
27         %CASO PARTICULAR DO PRIMEIRO INDICE
28         M(ini,8) = M(ini,3);
29         M(ini,9) = M(ini,3);
30
31         for j = ini+1:1:fim
32             %SOMA ACUMULADA
33             M(j,8) = M(j,3) + M(j-1,8);
34             %COPIA DA QUANTIDADE
35             M(j,9) = M(j,3);
36         end
37     end
38
39
40 end

```

```

1 function [buckets, IF_buckets, qtyVec] = ...
   Gerador_Buckets(M, IF_M, dataIni, dataFim)
2
3     %-----
4     %PROCURANDO O INDICE DO DIA "dataIni"
5     [Y1, M1, D1, -, -, -] = datevec(dataIni);
6     for i = 1:1:length(IF_M)
7         [Y2, M2, D2, -, -, -] = datevec(M(IF_M(i,1),1));
8         %VERIFICANDO SE CADA DIA E IGUAL AO "dataIni"
9         if Y1 == Y2 && M1 == M2 && D1 == D2
10            %ENCONTROU O DIA "dataIni"
11            inicio = i;
12            break;
13        end
14    end
15
16    %PROCURANDO O INDICE DO DIA "dataFim"
17    [Y1, M1, D1, -, -, -] = datevec(dataFim);
18    for i = length(IF_M):-1:0
19        [Y2, M2, D2, -, -, -] = datevec(M(IF_M(i,1),1));
20        %VERIFICANDO SE CADA DIA E IGUAL AO "dataFim"
21        if Y1 == Y2 && M1 == M2 && D1 == D2
22            %ENCONTROU O DIA "dataFim"
23            fim = i;
24            break;
25        end
26    end
27
28    %-----
29    %SER USADO A MEDIA DE VOLUME DOS ULTIMOS "leg" DIAS
30    leg = 10;
31    %CADA BUCKET TERA A FRACAO "fator" DO VOLUME DO DIA
32    fator = 1/50;
33    %AS QUANTIDADES SAO SEMPRE MULTIPLAS DE "qtdMin"
34    qtdMin = 5;
35    %CONTADOR INCREMENTAL DO NUMERO DE BUCKETS
36    contador = 0;

```

```

35  %-----
36
37  %-----
38  %ALOCANDO
39  IF_buckets = zeros(length(IF_M),2);
40  qtyVec = zeros(length(IF_M),1);
41
42  %ITERANDO PELOS DIAS
43  for i = inicio:1:fim
44      %SOMANDO AS QUANTIDADES NEGOCIADAS NOS DIAS
45      qtd = sum(M(IF_M(i-leg,1):IF_M(i-1,2),3));
46      %CALCULANDO A MEDIA
47      qtd = int32(fator*qtd/leg/qtdMin)*qtdMin;
48      %SALVANDO NO VETOR DE QUANTIDADES
49      qtyVec(i,1) = qtd;
50
51      ini = IF_M(i,1);
52      fim = IF_M(i,2);
53
54      %CALCULANDO O NUMERO DE BUCKETS DO DIA "i"
55      nBucketsDia = ...
56          cast(cast(M(fim,8),'double')/cast(qtd,'double'),'int32');
57      if M(fim,8) - nBucketsDia*qtd > 0.01
58          nBucketsDia = nBucketsDia + 1;
59      end
60
61      ultimo_ind = ini;
62      %ITERANDO PELOS BUCKETS DO DIA "i"
63      for j = 1:1:nBucketsDia
64          %-----
65          %CRIANDO A ESTRUTURA BUCKET ZERADA
66          bucket = struct ...
67              ( ...
68              'ind_IF',0, ...
69              'ind_ini',0, ...
70              'ind_fim',0, ...
71              'Open',0, ...
72              'High',0, ...
73              'Low',1000000, ...
74              'Close',0, ...
75              'DeltaPx',0, ...
76              'Qty',0.0, ...
77              'Qty_Buy',0.0, ...
78              'Qty_Sell',0.0, ...
79              'Ratio',0, ...
80              'Ratio_VPIN',0, ...
81              'Type',0, ...
82              'LastBucket',0, ...
83              'VPIN',-1, ...
84              'Ret',0 ...
85          );
86          %-----
87          %QUANTIDADE QUE FALTA PARA COMPLETAR O BUCKET "j"
88          V = qtd;
89
90          %LIMITES EM QUANTIDADES
91          s_ini = (j-1)*qtd+1;
          s_fim = j*qtd;

```

```

92
93     %PREENCHENDO BUCKET DE INICIO E FIM DE DIA
94     if j == 1
95         %BUCKET DE INICIO DE DIA
96         bucket.LastBucket = -1;
97         IF_buckets(i,1) = contador+1;
98     elseif j == nBucketsDia
99         %BUCKET DE FIM DE DIA
100        bucket.LastBucket = 1;
101        IF_buckets(i,2) = contador+1;
102    end
103
104    barra_inicial = 0;
105
106    %ITERANDO NO NEGOCIO QUE PAROU ATE O ULTIMO DO DIA "i"
107    for z = ultimo_ind:1:fim
108        %TESTANDO SE E A PRIMEIRA BARRA DAQUELE BUCKET
109        if M(z,8) ≥ s_ini && barra_inicial == 0
110            barra_inicial = 1;
111            bucket.Open = M(z,2);
112            bucket.ind_ini = z;
113        end
114
115        if M(z,8) ≥ s_ini
116            %ATUALIZANDO "V"
117            if M(z,9) ≤ V
118                if M(z,7) == -1
119                    bucket.Qty_Buy = bucket.Qty_Buy + M(z,9) * 1;
120                    bucket.Qty_Sell = bucket.Qty_Sell + M(z,9) * 0;
121                elseif M(z,7) == 1
122                    bucket.Qty_Buy = bucket.Qty_Buy + M(z,9) * 0;
123                    bucket.Qty_Sell = bucket.Qty_Sell + M(z,9) * 1;
124                else
125                    bucket.Qty_Buy = bucket.Qty_Buy + M(z,9) * 0;
126                    bucket.Qty_Sell = bucket.Qty_Sell + M(z,9) * 0;
127                end
128
129                V = V-M(z,9);
130                M(z,9) = 0;
131            else
132                if M(z,7) == -1
133                    bucket.Qty_Buy = bucket.Qty_Buy + ...
134                        cast(V,'double') * 1;
135                    bucket.Qty_Sell = bucket.Qty_Sell + ...
136                        cast(V,'double') * 0;
137                elseif M(z,7) == 1
138                    bucket.Qty_Buy = bucket.Qty_Buy + ...
139                        cast(V,'double') * 0;
140                    bucket.Qty_Sell = bucket.Qty_Sell + ...
141                        cast(V,'double') * 1;
142                else
143                    bucket.Qty_Buy = bucket.Qty_Buy + ...
144                        cast(V,'double') * 0;
145                    bucket.Qty_Sell = bucket.Qty_Sell + ...
146                        cast(V,'double') * 0;
147                end
148
149                M(z,9) = M(z,9) - V;

```

```

144         V = 0;
145     end
146
147     %TESTA SE O BUCKET "j" ESTA PREENCHIDO
148     if V ≤ 0
149         bucket.Close = M(z,2);
150         ultimo_ind = z;
151         bucket.ind_fim = z;
152         bucket.ind_IF = i;
153         bucket.Ratio_VPIN = cast((bucket.Qty_Buy - ...
154             bucket.Qty_Sell), 'double')/cast(qtd, 'double');
155         bucket.Qty = bucket.Qty_Buy + bucket.Qty_Sell;
156         buckets_aux(j,1) = bucket;
157         contador = contador+1;
158         buckets(contador,1) = bucket;
159         break;
160     end
161
162     if j == nBucketsDia && z == fim && V > 0
163         bucket.Close = M(z,2);
164         ultimo_ind = z;
165         bucket.ind_fim = z;
166         bucket.ind_IF = i;
167         bucket.Ratio_VPIN = cast((bucket.Qty_Buy - ...
168             bucket.Qty_Sell), 'double')/cast(qtd, 'double');
169         bucket.Qty = bucket.Qty_Buy + bucket.Qty_Sell;
170         buckets_aux(j,1) = bucket;
171         contador = contador+1;
172         buckets(contador,1) = bucket;
173     end
174 end
175 end
176 end

```

VPIN: A Regra do Tick

```

1  %-----
2  % MATRIZ "M" - tick-by-tick
3  % 9 Colunas
4  %-----
5  %COLUNA 1  =  TEMPO DO NEGOCIO
6  %COLUNA 2  =  PRECO
7  %COLUNA 3  =  QUANTIDADE
8  %COLUNA 4  =  _____
9  %COLUNA 5  =  _____
10 %COLUNA 6  =  LADO AGRESSOR (QUANDO A BOLSA FORNECER)
11 %COLUNA 7  =  LADO AGRESSOR CLASSIFICADO PELO TICK RULE
12 %COLUNA 8  =  QUANTIDADE ACUMULADA DO DIA
13 %COLUNA 9  =  COPIA DA QUANTIDADE
14 %-----
15
16
17 function [M] = Gerador_stats(M, IF_M)

```

```

18
19     EPS = 0.001;
20
21     %COLOCANDO A CLASSIFICACAO DO NEGOCIO NA COLUNA 7
22     for i = 1:length(IF_M)
23         ini = IF_M(i,1);
24         fim = IF_M(i,2);
25
26         for j = ini:1:fim
27             if j == ini
28                 M(j,7) = 0;
29             else
30                 ΔPx = M(j,2)-M(j-1,2);
31                 if ΔPx > EPS
32                     M(j,7) = 1;
33                 elseif ΔPx < -EPS
34                     M(j,7) = -1;
35                 elseif abs(ΔPx) < EPS
36                     M(j,7) = 0;
37                     for k = j-1:-1:ini
38                         if M(k,7) == 1
39                             M(j,7) = 1;
40                             break;
41                         elseif M(k,7) == -1
42                             M(j,7) = -1;
43                             break;
44                         end
45                     end
46                 end
47             end
48         end
49     end
50
51     Ind2 = find(M(:,7) == 0);
52     M(Ind2,3) = 0;
53
54     for i = 1:length(M)
55         [Y1, M1, D1, H1, MN1, S1] = datevec(M(i,1));
56         if H1 > 16 || (H1 == 16 && MN1 > 1)
57             M(i,3) = 0;
58         end
59     end
60
61     for i = 1:length(IF_M)
62         %DEFININDO INICIO E FIM DO DIA
63         ini = IF_M(i,1);
64         fim = IF_M(i,2);
65
66         %CASO PARTICULAR DO PRIMEIRO INDICE
67         M(ini,8) = M(ini,3);
68         M(ini,9) = M(ini,3);
69
70         for j = ini+1:1:fim
71             %SOMA ACUMULADA
72             M(j,8) = M(j,3) + M(j-1,8);
73             %COPIA DA QUANTIDADE
74             M(j,9) = M(j,3);
75         end

```

```

76     end
77
78
79 end

```

```

1 function [buckets, IF_buckets, qtyVec] = ...
   Gerador_Buckets (M, IF_M, dataIni, dataFim)
2
3   %-----
4   %PROCURANDO O INDICE DO DIA "dataIni"
5   [Y1, M1, D1, -, -, -] = datevec (dataIni);
6   for i = 1:1:length(IF_M)
7       [Y2, M2, D2, -, -, -] = datevec (M(IF_M(i,1),1));
8       %VERIFICANDO SE CADA DIA E IGUAL AO "dataIni"
9       if Y1 == Y2 && M1 == M2 && D1 == D2
10          %ENCONTROU O DIA "dataIni"
11          inicio = i;
12          break;
13       end
14   end
15
16   %PROCURANDO O INDICE DO DIA "dataFim"
17   [Y1, M1, D1, -, -, -] = datevec (dataFim);
18   for i = length(IF_M):-1:0
19       [Y2, M2, D2, -, -, -] = datevec (M(IF_M(i,1),1));
20       %VERIFICANDO SE CADA DIA E IGUAL AO "dataFim"
21       if Y1 == Y2 && M1 == M2 && D1 == D2
22          %ENCONTROU O DIA "dataFim"
23          fim = i;
24          break;
25       end
26   end
27
28   %-----
29   %SERA USADO A MDIA DE VOLUME DOS ULTIMOS "leg" DIAS
30   leg = 10;
31   %CADA BUCKET TERA A FRACAO "fator" DO VOLUME DO DIA
32   fator = 1/50;
33   %AS QUANTIDADES SAO SEMPRE MULTIPLAS DE "qtdMin"
34   qtdMin = 5;
35   %CONTADOR INCREMENTAL DO NUMERO DE BUCKETS
36   contador = 0;
37
38   %-----
39   %ALOCANDO
40   IF_buckets = zeros (length(IF_M), 2);
41   qtyVec = zeros (length(IF_M), 1);
42
43   %ITERANDO PELOS DIAS
44   for i = inicio:1:fim
45       %SOMANDO AS QUANTIDADES NEGOCIADAS NOS DIAS
46       qtd = sum (M(IF_M(i-leg,1):IF_M(i-1,2), 3));
47       %CALCULANDO A MEDIA
48       qtd = int32 (fator*qtd/leg/qtdMin)*qtdMin;
49       %SALVANDO NO VETOR DE QUANTIDADES
50       qtyVec(i,1) = qtd;
51
52       ini = IF_M(i,1);

```



```

52     fim = IF_M(i,2);
53
54     %CALCULANDO O NUMERO DE BUCKETS DO DIA "i"
55     nBucketsDia = ...
56         cast(cast(M(fim,8), 'double')/cast(qtd, 'double'), 'int32');
57     if M(fim,8) - nBucketsDia*qtd > 0.01
58         nBucketsDia = nBucketsDia + 1;
59     end
60
61     ultimo_ind = ini;
62     %ITERANDO PELOS BUCKETS DO DIA "i"
63     for j = 1:nBucketsDia
64         %-----
65         %CRIANDO A ESTRUTURA BUCKET ZERADA
66         bucket = struct ...
67             ( ...
68             'ind_IF', 0, ...
69             'ind_ini', 0, ...
70             'ind_fim', 0, ...
71             'Open', 0, ...
72             'High', 0, ...
73             'Low', 1000000, ...
74             'Close', 0, ...
75             'DeltaPx', 0, ...
76             'Qty', 0.0, ...
77             'Qty_Buy', 0.0, ...
78             'Qty_Sell', 0.0, ...
79             'Ratio', 0, ...
80             'Ratio_VPIN', 0, ...
81             'Type', 0, ...
82             'LastBucket', 0, ...
83             'VPIN', -1, ...
84             'Ret', 0 ...
85         );
86         %-----
87         %QUANTIDADE QUE FALTA PARA COMPLETAR O BUCKET "j"
88         V = qtd;
89
90         %LIMITES EM QUANTIDADES
91         s_ini = (j-1)*qtd+1;
92         s_fim = j*qtd;
93
94         %PREENCHENDO BUCKET DE INICIO E FIM DE DIA
95         if j == 1
96             %BUCKET DE INICIO DE DIA
97             bucket.LastBucket = -1;
98             IF_buckets(i,1) = contador+1;
99         elseif j == nBucketsDia
100             %BUCKET DE FIM DE DIA
101             bucket.LastBucket = 1;
102             IF_buckets(i,2) = contador+1;
103         end
104
105         barra_inicial = 0;
106
107         %ITERANDO NO NEGOCIO QUE PAROU ATE O ULTIMO DO DIA "i"
108         for z = ultimo_ind:fim
109             %TESTANDO SE E A PRIMEIRA BARRA DAQUELE BUCKET

```

```

109         if M(z,8) ≥ s_ini && barra_inicial == 0
110             barra_inicial = 1;
111             bucket.Open = M(z,2);
112             bucket.ind_ini = z;
113         end
114
115         if M(z,8) ≥ s_ini
116             %ATUALIZANDO "v"
117             if M(z,9) ≤ V
118                 if M(z,7) == -1
119                     bucket.Qty_Buy = bucket.Qty_Buy + M(z,9) * 1;
120                     bucket.Qty_Sell = bucket.Qty_Sell + M(z,9) * 0;
121                 elseif M(z,7) == 1
122                     bucket.Qty_Buy = bucket.Qty_Buy + M(z,9) * 0;
123                     bucket.Qty_Sell = bucket.Qty_Sell + M(z,9) * 1;
124                 else
125                     bucket.Qty_Buy = bucket.Qty_Buy + M(z,9) * 0;
126                     bucket.Qty_Sell = bucket.Qty_Sell + M(z,9) * 0;
127                 end
128
129                 V = V-M(z,9);
130                 M(z,9) = 0;
131             else
132                 if M(z,7) == -1
133                     bucket.Qty_Buy = bucket.Qty_Buy + ...
134                         cast(V,'double') * 1;
135                     bucket.Qty_Sell = bucket.Qty_Sell + ...
136                         cast(V,'double') * 0;
137                 elseif M(z,7) == 1
138                     bucket.Qty_Buy = bucket.Qty_Buy + ...
139                         cast(V,'double') * 0;
140                     bucket.Qty_Sell = bucket.Qty_Sell + ...
141                         cast(V,'double') * 1;
142                 else
143                     bucket.Qty_Buy = bucket.Qty_Buy + ...
144                         cast(V,'double') * 0;
145                     bucket.Qty_Sell = bucket.Qty_Sell + ...
146                         cast(V,'double') * 0;
147                 end
148
149                 M(z,9) = M(z,9) - V;
150                 V = 0;
151             end
152
153             %TESTA SE O BUCKET "j" ESTA PREENCHIDO
154             if V ≤ 0
155                 bucket.Close = M(z,2);
156                 ultimo_ind = z;
157                 bucket.ind_fim = z;
158                 bucket.ind_IF = i;
159                 bucket.Ratio_VPIN = cast((bucket.Qty_Buy - ...
160                     bucket.Qty_Sell), 'double')/cast(qtd, 'double');
161                 bucket.Qty = bucket.Qty_Buy + bucket.Qty_Sell;
162                 buckets_aux(j,1) = bucket;
163                 contador = contador+1;
164                 buckets(contador,1) = bucket;
165                 break;
166             end
167         end

```

```

160         end
161
162         if j == nBucketsDia && z == fim && V > 0
163             bucket.Close = M(z,2);
164             ultimo_ind = z;
165             bucket.ind_fim = z;
166             bucket.ind_IF = i;
167             bucket.Ratio_VPIN = cast((bucket.Qty_Buy - ...
168                 bucket.Qty_Sell), 'double')/cast(qtd, 'double');
169             bucket.Qty = bucket.Qty_Buy + bucket.Qty_Sell;
170             buckets_aux(j,1) = bucket;
171             contador = contador+1;
172             buckets(contador,1) = bucket;
173         end
174     end
175 end
176 end

```

VPIN: Classificação Maciça por Volume

```

1  %GERADOR DE BARRAS DE TEMPO
2
3  %-----
4  % MATRIZ "M" -- tick-by-tick
5  % 6 Colunas
6  %-----
7  %COLUNA 1 = TEMPO DO NEGOCIO
8  %COLUNA 2 = PRECO
9  %COLUNA 3 = QUANTIDADE
10 %COLUNA 4 = -----
11 %COLUNA 5 = -----
12 %COLUNA 6 = LADO AGRESSOR (QUANDO A BOLSA FORNECER)
13 %-----
14
15 %-----
16 % MATRIZ "B"
17 % 13 Colunas
18 %-----
19 %COLUNA 1 = TEMPO DA BARRA
20 %COLUNA 2 = ABERTURA
21 %COLUNA 3 = MAXIMO
22 %COLUNA 4 = MINIMO
23 %COLUNA 5 = FECHAMENTO
24 %COLUNA 6 = QUANTIDADE
25 %COLUNA 7 = INICIO/FIM DIA (INICIO = 1/FIM = -1)
26 %COLUNA 8 = QUANTIDADE INTRADIARIA ACUMULADA
27 %COLUNA 9 = REPLICA DA QUANTIDADE
28 %COLUNA 10 = VARIACAO DE PRECO (FECHAMENTO x FECHAMENTO)
29 %COLUNA 11 = RETORNO (FECHAMENTO x FECHAMENTO)
30 %COLUNA 12 = FRACAO AGRESSAO DE COMPRA
31 %COLUNA 13 = FRACAO AGRESSAO DE VENDA
32 %-----
33
34 function [B IF_B] = Gerador_Barras(M, IF_M, freq, mkt)

```

```

35 %-----
36 %*A FREQUENCIA ESTA EM SEGUNDOS
37 %-----
38 %NMEROS DE DIAS DE NEGOCIACAO
39 nbDias = length(IF_M);
40
41 % BARRA PARA 9 HORAS DE MERCADO
42 if strcmp(mkt,'DI1')== 1
43     B = zeros(nbDias*(7*(3600/freq)),13);
44 elseif strcmp(mkt,'DOL')== 1
45     B = zeros(nbDias*(9*(3600/freq)),13);
46 elseif strcmp(mkt,'IND')== 1
47     B = zeros(nbDias*(11*(3600/freq)),13);
48 end
49
50 % DELIMITADOR DE CADA DIA NAS BARRAS
51 IF_B = zeros(nbDias, 2);
52 %-----
53 % 2 Colunas
54 %-----
55 %COLUNA 1 = INDICE DO INICIO DO DIA
56 %COLUNA 2 = INDICE DO FIM DO DIA
57 %-----
58
59 %ITERANDO PELOS DIAS
60 for d = 1:nbDias
61     % NDICE NA MATRIZ "M" DO INICIO DO DIA "d"
62     iniM = IF_M(d, 1);
63     % NDICE NA MATRIZ "M" DO INICIO DO DIA "d"
64     fimM = IF_M(d, 2);
65
66     %DATA INICIO DO DIA DO DIA "d"
67     [A1,M1,D1,H1,MN1,S1] = datevec(M(iniM,1));
68     %DATA FIM DO FIA DO DIA "d"
69     [A2,M2,D2,H2,MN2,S2] = datevec(M(fimM,1));
70
71     if strcmp(mkt,'DI1')== 1
72         %COMECANDO 9:00:00
73         tBeg = datenum([A1,M1,D1,9,0,0]);
74         %TERMINANDO 16:00:00
75         tEnd = datenum([A2,M2,D2,16,0,0]);
76     elseif strcmp(mkt,'DOL')== 1
77         %COMECANDO 9:00:00
78         tBeg = datenum([A1,M1,D1,9,0,0]);
79         %TERMINANDO 18:00:00
80         tEnd = datenum([A2,M2,D2,18,0,0]);
81     elseif strcmp(mkt,'IND')== 1
82         %COMECANDO 9:00:00
83         tBeg = datenum([A1,M1,D1,9,0,0]);
84         %TERMINANDO 19:00:00
85         tEnd = datenum([A2,M2,D2,20,0,0]);
86     end
87
88     %GERANDO O TEMPOS DE CADA BARRA DO DIA "d"
89     dt ...
90         =( (tBeg+(freq/(24*3600))):(freq/(24*3600)):(tEnd+(freq/(24*3600))))';
91     nbBars = length(dt);

```

```

92     %PREENCHENDO O IF_B DO DIA "d"
93     if d == 1
94         IF_B(d,1) = 1;
95         IF_B(d,2) = nbBars;
96     else
97         IF_B(d,1) = IF_B(d-1,2)+1;
98         IF_B(d,2) = IF_B(d,1) + nbBars-1;
99     end
100
101     iniB = IF_B(d,1);
102     fimB = IF_B(d,2);
103
104     %GERANDO A COLUNA 1 DOS TEMPOS DA BARRAS
105     B(iniB:fimB,1) = dt;
106
107     %ITERANDO DENTRO DE UM DIA
108     for i = iniB:fimB
109         %-----
110         %PROCURANDO NEGOCIOS DENTRO DO INTERVALO DE TEMPO
111         if i == iniB
112             ind = (iniM-1) + find(M(iniM:fimM,1) < B(i,1));
113         else
114             ind = (iniM-1) + find(and(M(iniM:fimM,1) ≥ B(i-1,1), ...
115                                     M(iniM:fimM,1) < B(i,1)));
116         %-----
117         %OCORRENCIA DE NEGOCIOS NA BARRA
118         if isempty(ind) ≠ 1
119             %CALCULANDO OHLC E VOLUME
120             open = M(ind(1),2);
121             high = max(M(ind,2));
122             low = min(M(ind,2));
123             close = M(ind(end),2);
124             volume = sum(M(ind,3));
125
126             %SETANDO COLUNAS 2,3,4,5 e 6
127             B(i, 2) = open;
128             B(i, 3) = high;
129             B(i, 4) = low;
130             B(i, 5) = close;
131             B(i, 6) = volume;
132         %-----
133         %NAO OCORRENCIA DE NEGOCIOS NA BARRA
134         else
135             if i == iniB
136                 B(i,2:7) = [0 0 0 0 0 0];
137             else
138                 B(i,2:5) = [B(i-1,5) B(i-1,5) B(i-1,5) B(i-1,5)];
139                 B(i, 6:7) = [0 0];
140             end
141         end
142         %-----
143         %SETANDO COLUNA 7
144         if i == iniB
145             %INICIO DO DIA
146             B(i, 7) = -1;
147         elseif i == fimB
148             %FIM DO DIA

```

```

149         B(i, 7) = 1;
150     end
151 end
152 end
153
154 %GERANDO AS COLUNAS 8 e 9
155 B = qtd_acum(B, IF_B);
156 %GERANDO AS COLUNAS 10 e 11
157 B = Δ_Preco(B);
158 %GERANDO AS COLUNAS 12 e 13
159 B = fracao_volume(B);
160
161 end
162
163 function B = qtd_acum(B, IF_B)
164     %-----
165     %GERANDO AS COLUNAS 8 e 9:
166     %-----
167     %COLUNA 8 = QUANTIDADE INTRADIARIA ACUMULADA
168     %COLUNA 9 = REPLICA DA QUANTIDADE
169     %-----
170
171     %ITERANDO PELOS DIAS
172     for i = 1:length(IF_B)
173         %BARRA DO INICIO DA DIA
174         ini = IF_B(i,1);
175         %BARRA DO FIM DO DIA
176         fim = IF_B(i,2);
177
178         B(ini,8) = B(ini,6);
179         B(ini,9) = B(ini,6);
180         %ITERANDO PELAS BARRAS DAQUELE DIA
181         for j = ini+1:1:fim
182             B(j,8) = B(j,6) + B(j-1,8);
183             B(j,9) = B(j,6);
184         end
185     end
186 end
187
188 function B = Δ_Preco(B)
189     %-----
190     %GERANDO AS COLUNAS 10 e 11:
191     %-----
192     %COLUNA 10 = VARIACAO DE PRECO (FECHAMENTO x FECHAMENTO)
193     %COLUNA 11 = RETORNO (FECHAMENTO x FECHAMENTO)
194     %-----
195
196     %ITERANDO PELAS BARRAS
197     for i = 1:1:length(B(:,1))
198         %-----
199         %VERIFICANDO SE E A PRIMEIRA BARRA DO DIA (FECHAMENTO X ABERTURA)
200         if B(i,7) == -1
201             if B(i,2) > 0
202                 B(i,10) = B(i,5)-B(i,2);
203                 B(i,11) = B(i,5)/B(i,2) - 1;
204             else
205                 B(i,10) = 0;
206                 B(i,11) = 0;

```

```

207         end
208     %-----
209     %VERIFICANDO SE NAO E A PRIMEIRA BARRA DO DIA (FECHAMENTO X ...
        FECHAMENTO)
210     else
211         if B(i-1,5) > 0 && B(i,5) > 0
212             B(i,10) = B(i,5)-B(i-1,5);
213             B(i,11) = (B(i,5)-B(i-1,5))/B(i-1,5);
214         else
215             B(i,10) = 0;
216             B(i,11) = 0;
217         end
218     end
219 end
220 end
221
222 function B = fracao_volume(B)
223     %-----
224     %GERANDO AS COLUNAS 12 e 13:
225     %-----
226     %COLUNA 12 = FRACAO AGRESSAO DE COMPRA
227     %COLUNA 13 = FRACAO AGRESSAO DE VENDA
228     %-----
229
230     %CALCULANDO O DESVIO PADRAO BASEADO EM 60 AMOSTRAS
231     ini = 60;
232     for i = ini+1:1:length(B(:,1))
233         %CALCULANDO O DESVIO PADRAO
234         sigma = std(B(i-ini:i-1,10));
235         if sigma ≠ 0
236             B(i,12) = normcdf(B(i,10)/sigma);
237         elseif B(i,10) == 0
238             B(i,12) = 0.5;
239         elseif B(i,10) > 0
240             B(i,12) = 1;
241         elseif B(i,10) < 0
242             B(i,12) = 0;
243         end
244         B(i,13) = 1-B(i,12);
245     end
246 end

```

```

1 function [buckets, IF_buckets, qtyVec] = ...
    Gerador_Buckets(M, IF_M, B, IF_B, dataIni, dataFim)
2
3     %-----
4     %PROCURANDO O INDICE DO DIA "dataIni"
5     [Y1, M1, D1, -, -, -] = datevec(dataIni);
6     for i = 1:1:length(IF_M)
7         [Y2, M2, D2, -, -, -] = datevec(M(IF_M(i,1),1));
8         %VERIFICANDO SE CADA DIA E IGUAL AO "dataIni"
9         if Y1 == Y2 && M1 == M2 && D1 == D2
10             %ENCONTROU O DIA "dataIni"
11             inicio = i;
12             break;
13         end
14     end

```

```

15     %PROCURANDO O INDICE DO DIA "dataFim"
16     [Y1, M1, D1, -, -, -] = datevec(dataFim);
17     for i = length(IF_M):-1:0
18         [Y2, M2, D2, -, -, -] = datevec(M(IF_M(i,1),1));
19         %VERIFICANDO SE CADA DIA E IGUAL AO "dataFim"
20         if Y1 == Y2 && M1 == M2 && D1 == D2
21             %ENCONTROU O DIA "dataFim"
22             fim = i;
23             break;
24         end
25     end
26     %-----
27     %SER USADO A MEDIA DE VOLUME DOS ULTIMOS "leg" DIAS
28     leg = 10;
29     %CADA BUCKET TERA A FRACAO "fator" DO VOLUME DO DIA
30     fator = 1/50;
31     %AS QUANTIDADES SAO SEMPRE MULTIPLAS DE "qtdMin"
32     qtdMin = 5;
33     %CONTADOR INCREMENTAL DO NUMERO DE BUCKETS
34     contador = 0;
35     %-----
36
37     %-----
38     %ALOCANDO
39     IF_buckets = zeros(length(IF_M),2);
40     qtyVec = zeros(length(IF_M),1);
41
42     %ITERANDO PELOS DIAS
43     for i = inicio:1:fim
44         %SOMANDO AS QUANTIDADES NEGOCIADAS NOS DIAS
45         qtd = sum(B(IF_B(i-leg,1):IF_B(i-1,2),6));
46         %CALCULANDO A MEDIA
47         qtd = int32(fator*qtd/leg/qtdMin)*qtdMin;
48         %SALVANDO NO VETOR DE QUANTIDADES
49         qtyVec(i,1) = qtd;
50
51         ini = IF_B(i,1);
52         fim = IF_B(i,2);
53
54         %CALCULANDO O NUMERO DE BUCKETS DO DIA "i"
55         nBucketsDia = ...
56             cast(cast(B(fim,8),'double')/cast(qtd,'double'),'int32');
57         if B(fim,8) - nBucketsDia*qtd > 0.01
58             nBucketsDia = nBucketsDia + 1;
59         end
60
61         ultimo_ind = ini;
62         %ITERANDO PELOS BUCKETS DO DIA "i"
63         for j = 1:1:nBucketsDia
64             %-----
65             %CRIANDO A ESTRUTURA BUCKET ZERADA
66             bucket = struct ...
67                 ( ...
68                 'ind_IF',0, ...
69                 'ind_ini',0, ...
70                 'ind_fim',0, ...
71                 'Open',0, ...
72                 'High',0, ...

```



```

72     'Low',1000000, ...
73     'Close',0, ...
74     'DeltaPx',0, ...
75     'Qty',0.0, ...
76     'Qty_Buy',0.0, ...
77     'Qty_Sell',0.0, ...
78     'Ratio',0, ...
79     'Ratio_VPIN',0, ...
80     'Type',0, ...
81     'LastBucket',0, ...
82     'VPIN',-1, ...
83     'Ret',0 ...
84 );
85 %-----
86 %QUANTIDADE QUE FALTA PARA COMPLETAR O BUCKET "j"
87 V = qtd;
88
89 %LIMITES EM QUANTIDADES
90 s_ini = (j-1)*qtd+1;
91 s_fim = j*qtd;
92
93 %PREENCHENDO BUCKET DE INICIO E FIM DE DIA
94 if j == 1
95     %BUCKET DE INICIO DE DIA
96     bucket.LastBucket = -1;
97     IF_buckets(i,1) = contador+1;
98 elseif j == nBucketsDia
99     %BUCKET DE FIM DE DIA
100    bucket.LastBucket = 1;
101    IF_buckets(i,2) = contador+1;
102 end
103
104 barra_inicial = 0;
105
106 %ITERANDO DA BARRA QUE PAROU ATE O ULTIMO DO DIA "i"
107 for z = ultimo_ind:1:fim
108     %TESTANDO SE E A PRIMEIRA BARRA DAQUELE BUCKET
109     if B(z,8) ≥ s_ini && barra_inicial == 0
110         barra_inicial = 1;
111         bucket.ind_ini = z;
112     end
113
114     if B(z,8) ≥ s_ini
115         %ATUALIZANDO "V"
116         if B(z,9) ≤ V
117             bucket.Qty_Buy = bucket.Qty_Buy + B(z,9) * B(z,12);
118             bucket.Qty_Sell = bucket.Qty_Sell + B(z,9) * B(z,13);
119
120             V = V-B(z,9);
121             B(z,9) = 0;
122         else
123             bucket.Qty_Buy = bucket.Qty_Buy + ...
124                 cast(V,'double') * B(z,12);
125             bucket.Qty_Sell = bucket.Qty_Sell + ...
126                 cast(V,'double') * B(z,13);
127
128             B(z,9) = B(z,9) - V;
129             V = 0;

```

```
128         end
129
130         %TESTA SE O BUCKET "j" EST PREENCHIDO
131         if V ≤ 0
132             ultimo_ind = z;
133             bucket.ind_fim = z;
134             bucket.ind_IF = i;
135             bucket.Ratio_VPIN = cast((bucket.Qty_Buy - ...
136                 bucket.Qty_Sell), 'double')/cast(qtd, 'double');
137             bucket.Qty = bucket.Qty_Buy + bucket.Qty_Sell;
138             buckets_aux(j,1) = bucket;
139             contador = contador+1;
140             buckets(contador,1) = bucket;
141             break;
142         end
143
144         if j == nBucketsDia && z == fim && V > 0
145             ultimo_ind = z;
146             bucket.ind_fim = z;
147             bucket.ind_IF = i;
148             bucket.Ratio_VPIN = cast((bucket.Qty_Buy - ...
149                 bucket.Qty_Sell), 'double')/cast(qtd, 'double');
150             bucket.Qty = bucket.Qty_Buy + bucket.Qty_Sell;
151             buckets_aux(j,1) = bucket;
152             contador = contador+1;
153             buckets(contador,1) = bucket;
154         end
155     end
156 end
157 end
```


Referências Bibliográficas

- [1] *BM&F Bovespa: boletim diário informativo 09/08/2011*, <http://www.acervobmfbovespa.com.br/Archive/ReadPDF?archiveID=29046>.
- [2] *BM&F Bovespa: boletim diário informativo 27/07/2012*, <http://www.acervobmfbovespa.com.br/Archive/ReadPDF?archiveID=48883>.
- [3] *BM&F Bovespa: novo recorde de negócios 27/07/2012*, <http://www.acervobmfbovespa.com.br/Archive/ReadPDF?archiveID=48159>.
- [4] *Cálculo de ajuste dos contratos da BM&F Bovespa*, http://www.bmf.com.br/bmfbovespa/pages/boletim1/arquivos/Metodologia_janeiro-2011.pdf.
- [5] *Especificações contrato futuro de taxa de câmbio de reais por dolar da BM&F Bovespa*, http://www.bmf.com.br/bmfbovespa/pages/contratos1/Financeiros/PDF/Cambio_Futuro.pdf.
- [6] *Especificações contrato futuro de taxa di de um fia da BM&F Bovespa*, <http://www.bmf.com.br/bmfbovespa/pages/contratos1/Financeiros/PDF/Difuturo.pdf>.
- [7] *Especificações contrato futuro de Índice bovespa da BM&F Bovespa*, http://www.bmf.com.br/bmfbovespa/pages/contratos1/Financeiros/IND_futuro.pdf.
- [8] *ftp BM&F Bovespa*, <ftp://ftp.bmf.com.br/marketdata/>.
- [9] *Metodologia PTAX*, <https://www3.bcb.gov.br/normativo/detalharNormativo.do?N=110082375&method=detalharNormativo>.
- [10] *Sisbacen*, <http://www.bcb.gov.br/?SISBACEN>.
- [11] *Volume mensal BM&F Bovespa*, <http://www.bmfbovespa.com.br/shared/iframeBoletim.aspx?altura=3200&idioma=pt-br&url=www.bmf.com.br/bmfbovespa/pages/boletim1/VolumeGeral/VolumeGeral.asp>.
- [12] Torben Andersen and Oleg Bondarenko, *Reflecting on the VPIN dispute*, *Jornal of Financial Markets* **17** (2014), 53–64.
- [13] Andriy Shkilko Bidisha Chakrabarty, Roberto Pascual, *Order imbalance and individual stock returns: Theory and evidence*, *Jornal of Financial Economics* **72** (2004), 485–518.
- [14] Bidisha Chakrabarty, Roberto Pascual, and Andriy Shkilko, *Trade classification algorithms: A horse race between the bulk-based and tick-based-rules*, *Johnson School Research Paper Series* (2012), <http://ssrn.com/abstract=2182819>.

- [15] Maureen O'Hara David Easley, Soeren Hvidkjaer, *Is information risk a determinant of asset returns ?*, The Journal of Finance **57** (2002), no. 5.
- [16] David Easley, Marcos López de Prado, and Maureen O'Hara, *Flow toxicity and liquidity in high-frequency trading world*, Review of Finance Studies **25** (2012), 1457–1493.
- [17] ———, *The volume clock: Insights into the high frequency paradigm*, Journal of Portfolio Management **9** (2012).
- [18] David Easley, Robert Engle, Maureen O'Hara, and Liuren Wu, *Time-varying arrival rates of informed and uninformed traders*, Journal of Financial Econometrics **6** (2012), no. 2, 171–207.
- [19] Quan Gan, David Johnstone, and Wang Chun Wei, *Cluster pin: A new estimation method for the probability of informed trading*, (2013), <http://ssrn.com/abstract=2363814>.
- [20] Lawrence Glosten and Paul Milgrom, *Bid, ask and transaction prices in a specialist market with heterogeneously informed traders*, Journal of Financial Economics **13** (1985), 71–100.
- [21] Maureen O'Hara, *Market Microstructure theory*, Blackwell, 1997.
- [22] Marcos M. L. Prado, *Advances in High Frequency Strategies*, Complutense University, 2011.
- [23] Álvaro Cartea, Ryan Donnelly, and Sebastian Jaimungal, *Robust market making*, (2013).