

## 6

### Modelo de Previsão de Preços

#### 6.1

##### Introdução

Com os recentes desenvolvimentos tecnológicos associados ao mercado financeiro, a busca por compreender a dinâmica dos preços dos ativos passou a ser um desafio ainda maior. Atualmente, com a existência de dados de alta frequência, tornou-se possível avaliar a dinâmica das microestruturas e o comportamento dos preços dos ativos. Assim, a principal questão é saber como o ambiente (sistemas de negociação e investidores) pode influenciar o preço dos ativos, tendo em mente que, no mundo real, as imperfeições do mercado financeiro existem, e que, na realidade, são investigadas e consideradas pelos agentes.

Entre os objetivos deste capítulo, pode-se destacar:

- Entender e modelar a dinâmica dos dados de alta frequência, em particular: duração, volume e *spread*;
- Modelar o processo de formação de preços medindo a influência das microestruturas – distribuição das mudanças no preço dos ativos e entender a racionalidade implícita nas ordens de compra e venda;
- Analisar as evidências sobre a capacidade de previsão no que se refere aos ativos do mercado financeiro.

## 6.2

### Modelo proposto

Buscando modelar a dinâmica das microestruturas, propõe-se a utilização da formulação apresentada no capítulo 5 – EMACM. Esta formulação é apropriada em função da capacidade de lidar com casualidades entre as variáveis, além do fato de que a função de previsão pode ser usada diretamente na formulação do modelo *probit* ordenado, tornando possível inferir sobre a dinâmica de alterações de preços.

#### 6.2.1

##### EMACM

Um conjunto de variáveis de alta frequência (duração –  $x_i$ , volume –  $v_i$  e *spread* de compra e venda –  $s_i$ ) seguem um EMACM, se:

$$x_i = \psi_i \cdot \varepsilon_i \rightarrow \varepsilon_i \sim i.i.d.exp(1) \quad (6.2.1)$$

$$v_i = \phi_i \cdot \eta_i \rightarrow \eta_i \sim i.i.d.exp(1) \quad (6.2.2)$$

$$s_i = \varphi_i \cdot \varpi_i \rightarrow \varpi_i \sim i.i.d.exp(1) \quad (6.2.3)$$

onde, a média condicional é:

$$\ln(\mu_i) = \gamma + A_1 \ln(\mu_{i-1}) + \dots + A_q \ln(\mu_{i-q}) + B_0 \ln(\tau_i) + B_1 \ln(\tau_{i-1}) + \dots + B_p \ln(\tau_{i-p}) \quad (6.2.4)$$

$\mu_i' = (\psi_i, \phi_i, \varphi_i)$ ,  $\tau_i' = (x_i, v_i, s_i)$ ,  $\gamma$  é o vetor de coeficientes e  $A_1, \dots, A_q$  e  $B_1, \dots, B_p$  representam as matrizes de coeficientes de cada processo estocástico.

Desta forma, a generalização do modelo pode ser escrita como:

- Equação associada às variáveis de interesse:

$$\begin{bmatrix} x_i \\ v_i \\ s_i \end{bmatrix} = \begin{bmatrix} \psi_i & 0 & 0 \\ 0 & \phi_i & 0 \\ 0 & 0 & \varphi_i \end{bmatrix} \cdot \begin{bmatrix} \varepsilon_i \\ \eta_i \\ \omega_i \end{bmatrix} \quad (6.2.5)$$

Onde,  $\varepsilon_i$ ,  $v_i$  e  $\omega_i \sim \exp(1)$ .

- Equação da média condicional

$$\begin{aligned} \ln \begin{bmatrix} \psi_i \\ \phi_i \\ \varphi_i \end{bmatrix} &= \begin{bmatrix} a_0 \\ b_0 \\ c_0 \end{bmatrix} + \sum_{l=1}^q \begin{bmatrix} a_1^{(l)} & a_2^{(l)} & a_3^{(l)} \\ b_1^{(l)} & b_2^{(l)} & b_3^{(l)} \\ c_1^{(l)} & c_2^{(l)} & c_3^{(l)} \end{bmatrix} \ln \begin{bmatrix} \psi_{i-l} \\ \phi_{i-l} \\ \varphi_{i-l} \end{bmatrix} \\ &+ \begin{bmatrix} 0 & 0 & 0 \\ b_4 & 0 & 0 \\ c_4 & c_5 & 0 \end{bmatrix} \ln \begin{bmatrix} x_i \\ v_i \\ s_i \end{bmatrix} + \sum_{m=1}^p \begin{bmatrix} a_4^{(m)} & a_5^{(m)} & a_6^{(m)} \\ b_5^{(m)} & b_6^{(m)} & b_7^{(m)} \\ c_6^{(m)} & c_7^{(m)} & c_8^{(m)} \end{bmatrix} \ln \begin{bmatrix} x_{i-m} \\ v_{i-m} \\ s_{i-m} \end{bmatrix} \end{aligned} \quad (6.2.6)$$

## 6.2.2

### O Modelo de Apreçamento (Modelo *Probit* Ordenado)

Considerando as equações (3.3.1) e (3.3.10), a formulação proposta é definida como:

$$z_i \sim N(\mu_i, \sigma_i^2) \quad (6.2.7)$$

$$\mu_i = \sum_{k=0}^3 (a_k \cdot x_{i-k} + b_k \cdot v_{i-k} + c_k \cdot s_{i-k}) + \sum_{n=1}^3 (d_n \cdot z_{i-n}) \quad (6.2.8)$$

$$\sigma_i^2 = \gamma_0 \cdot x_i + \gamma_1 \cdot s_{i-1} \quad (6.2.9)$$

Onde,  $z_i$  representa as mudanças nos preços de um ativo específico representadas em unidades do *tick*; ( $a_0, \dots, a_3, b_0, \dots, b_3, c_0, \dots, c_3, d_1, \dots, d_3, \gamma_0$ ,

$\gamma_1$ ) são os parâmetros a serem estimados;  $x_i$  é a duração;  $v_i$  é o volume; e  $s_i$  o *spread* de compra e venda do  $i$ -ésimo evento.

### 6.2.3

#### Testes Empíricos

Os testes empíricos podem ser divididos em diferentes estágios. Dessa forma, tem-se:

1. Cinco dias consecutivos de negociação são selecionados de uma base de dados (01/11/1990 – 08/11/1990) e o modelo EMACM (2,2) é estimado considerando apenas as variáveis existentes no modelo de apreçamento (duração, *spread* de compra e venda e volume). As figuras 6.1 – 6.5 apresentam os resultados da análise dentro da amostra.

- **ACF**
  - **Duration**

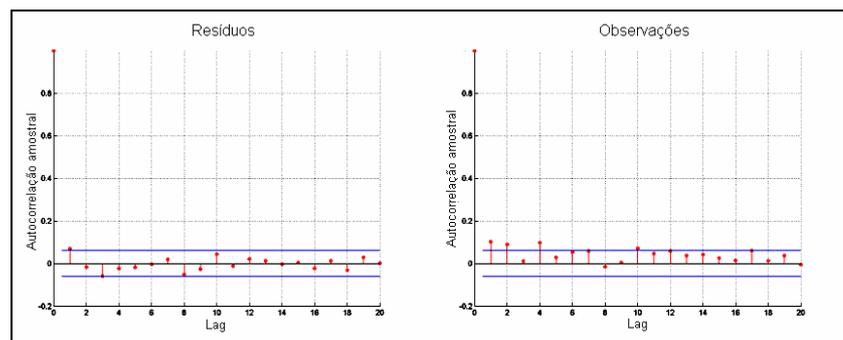


Figura 6.1 – ACF da duração (resíduos x observações)

○ **Volume**

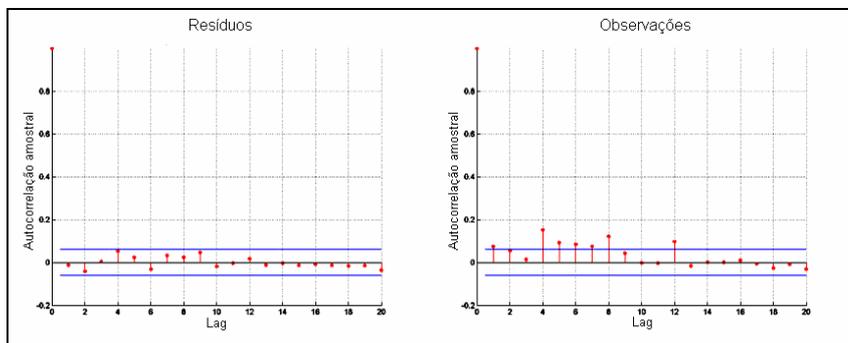


Figura 6.2 – ACF do volume (resíduos x observações)

○ **Spread**

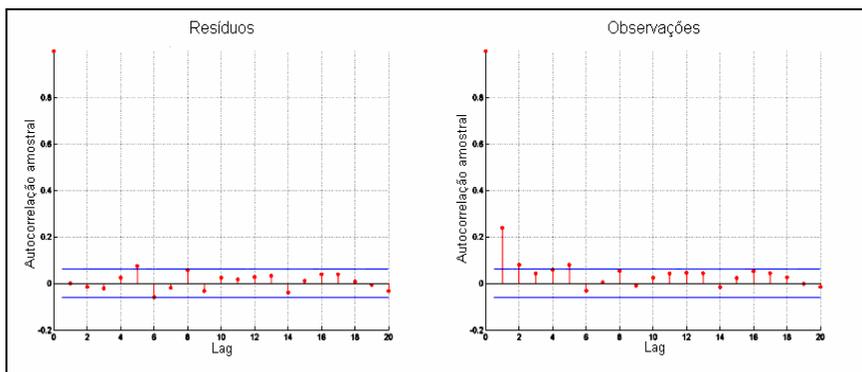


Figura 6.3 – ACF do spread (resíduos x observações)

□ **Ljung-Box:**

|                                  |                | Aceita $H_0$ (95%) | P-Valor | Ljung-Box | Valor Crítico |
|----------------------------------|----------------|--------------------|---------|-----------|---------------|
| <b>Observações</b>               | <b>Duração</b> | Rejeita            | 0,00%   | 56,52     | 25,00         |
|                                  | <b>Volume</b>  | Rejeita            | 0,00%   | 88,94     | 25,00         |
|                                  | <b>Spread</b>  | Rejeita            | 0,00%   | 94,68     | 25,00         |
| <b>Resíduos<br/>(08/11/1990)</b> | <b>Duração</b> | Aceita             | 29,89%  | 17,34     | 25,00         |
|                                  | <b>Volume</b>  | Aceita             | 67,09%  | 12,10     | 25,00         |
|                                  | <b>Spread</b>  | Aceita             | 11,99%  | 21,55     | 25,00         |

Tabela 5 – Análise da dependência linear (Ljung-Box)

2. O modelo *probit* ordenado descrito em (6.2.7), (6.2.8) e (6.2.9) é estimado e os valores resultantes da função de previsão do modelo EMACM são utilizados como substitutos dos valores das variáveis contemporâneas na equação do modelo (função de previsão do modelo de preços). As partições selecionadas durante o processo de estimação são então comparadas com os dados reais e aquele selecionado pela regra NAIVE<sup>3</sup>. Nesta etapa, utiliza-se uma janela móvel de cinco dias para estimação dos coeficientes. O experimento foi realizado, tomando por base o intervalo de tempo entre 08/11/1990 e 16/11/1990. O principal objetivo é apresentar os resultados do modelo de previsão de preços fora da amostra (o Apêndice IV apresenta os principais resultados).
- **Modelo Irrestrito:** os limites das partições e a matriz de coeficientes são estimados conjuntamente. A figura 6.4 apresenta os resultados obtidos. O número de acertos (movimento dos preços um passo à frente) do modelo proposto é comparado com o NAIVE, conforme descrito anteriormente.

---

<sup>3</sup> Regra NAIVE: se uma certa partição X é selecionada em um determinado evento “i”, então, para o evento subsequente, a partição selecionada será igual à anterior.

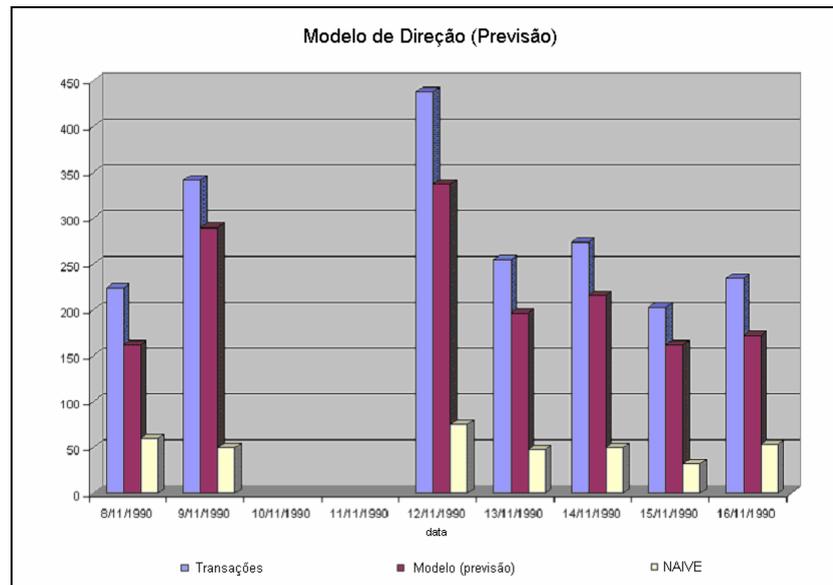


Figura 6.4 – Número de acertos (direção e magnitude)

- **Direção:** o principal objetivo deste procedimento é testar a capacidade de previsão apenas da direção das alterações nos preços, sem levar em consideração suas magnitudes. Nesse experimento, a partição é fixada em zero e apenas os parâmetros da equação precisam ser estimados.

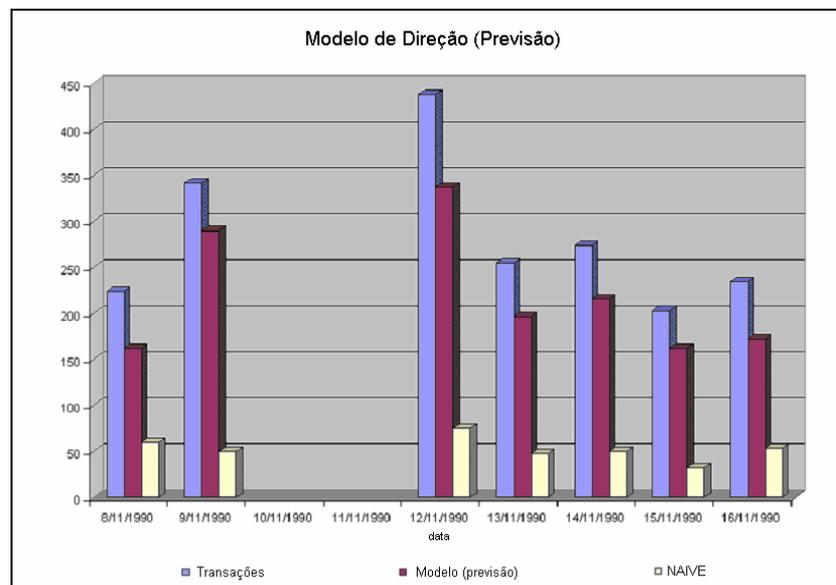


Figura 6.5 – Número de acertos (direção)

Conforme se pode notar, o método proposto captura de forma satisfatória o padrão intradiário e a dinâmica envolvida na mudança dos preços. A figura 6.6 resume os resultados.

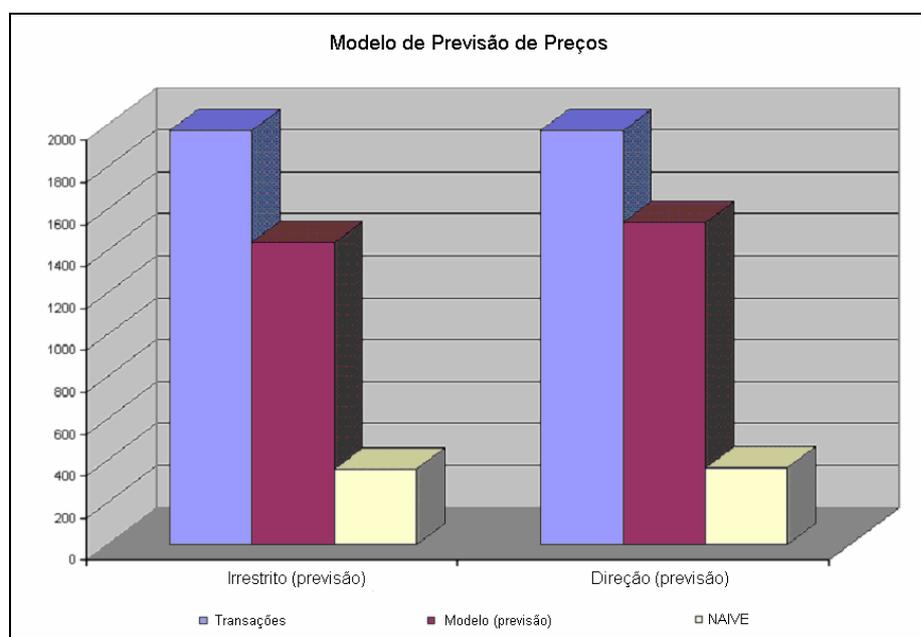


Figura 6.6 – Número total de acertos (direção)

Os resultados encontrados apontam para a existência de um padrão intradiário que é capturado pelo modelo proposto. Apesar do fato das negociações, quando analisadas em conjunto, não apresentarem assimetria significativa (ver os resultados do Apêndice IV), quando a seqüência cronológica de eventos é tomada e modelada apropriadamente, é possível realizar inferências com relação ao comportamento das variações de preços (distribuição condicional).