

5 Previsões

A análise de previsões é concentrada nos quatro últimos anos da amostra, abrangendo 983 dias úteis entre 3 de janeiro de 2000 e 31 de dezembro de 2003. Cada modelo é reestimado diariamente a partir de todos os dados até aquela data e utilizado para gerar previsões pontuais e valores em risco nos horizontes de um, cinco, dez e vinte dias adiante. As especificações dos modelos de regressão em árvore são revisadas mensalmente. O apêndice B descreve o procedimento para geração de previsões a partir das estimações com efeitos de assimetria, baseado em simulações condicionais. Como referência, também são incluídas previsões fornecidas por um modelo GARCH(1,1) e por um modelo de médias móveis exponencialmente amortecidas (EWMA) –diferentemente da abordagem tradicional, entretanto, são calculadas médias móveis da própria volatilidade, as quais geram previsões mais relevantes o modelo para os quadrados dos retornos. O parâmetro de amortecimento é calibrado dentro da amostra para 0,7 em previsões de um e cinco dias e 0,8 (mais suave) para os demais prazos.

5.1 Previsões Pontuais

O cálculo de previsões pontuais gera os resultados da Tabela 5.1. A avaliação das previsões é baseada no critério de erro absoluto médio e na estimação de regressões de Mincer-Zarnowitz:

$$RV_t = \alpha + \beta f_{t,i} + \varepsilon_{t,i}$$

onde RV_t é volatilidade observada no dia t (Hansen e Lunde, 2006a, mostram que embora menos sensível a *outliers*, uma possível comparação através do logaritmo da variável não atende condições necessárias para consistência), e $f_{t,i}$ é a previsão do modelo i para a volatilidade no dia t .¹ Se o modelo i está corretamente especificado, então $\alpha = 0$ e $\beta = 1$. O p-valor do teste F robusto para essa hipótese é computado para previsões um passo adiante (a

¹Verificou-se que a combinação sistemática de previsões pouco contribui para análise, tendo os resultados omitidos.

presença de heterocedasticidade prejudica o cálculo de estatísticas adequadas para horizontes maiores). O R^2 corrigido da regressão é reportado como medida da precisão dos modelos.

Dois testes estatísticos para contraposição entre os modelos são reportados. O teste de Harvey, Leybourne e Newbold (1997) é aplicado para comparação das previsões geradas pelos modelos em contraste com o modelo com integração fracional, onde a qualidade das previsões é medida através dos erros absolutos e dos resíduos da regressão de Mincer-Zarnowitz. Sejam $g(e_{1t})$ e $g(e_{2t})$ as funções de perda para os erros de previsão (e_{1t}, e_{2t}) dos modelos 1 e 2 no dia t . A hipótese nula é dada por $E[g(e_{1t}) - g(e_{2t})] = 0$, onde $g(e_{it}) = |RV_t - f_{t,i}|$ para o critério de erro absoluto e $g(e_{it}) = (RV_t - \hat{\alpha} + \hat{\beta}f_{t,i})^2$ para o critério de R^2 . O teste de habilidade preditiva superior (HPS), derivado por Hansen (2005), examina para cada modelo a hipótese nula de que o mesmo não é inferior a nenhuma outra alternativa em termos da função de perda admitida.

Para previsões de um dia, o modelo com efeitos de assimetria é superlativo tanto em termos de EAM quanto R^2 , significativamente superando o modelo FI pelo teste HLN (através de erros absolutos 5% menores em média) e sendo o único para o qual a hipótese nula do teste HPS não é rejeitada a 5% nos dois critérios, à exceção das médias móveis em R^2 . A superioridade do modelo em árvore com efeitos de assimetria se repete em todas as ações analisadas em termos de erros absolutos (Tabela D.4) e para onze delas em R^2 (Tabela D.6), contra quatro das médias móveis; paralelamente, os modelos FI, EWMA e HAR se alternam como os segundos melhores em R^2 na demais ações, enquanto o modelo HAR leva vantagem em EAMs. Na seqüência das previsões para IBM de um dia, há pouca distinção entre os modelos FI, EWMA, AR e HAR no ajuste medido pelo R^2 (as diferenças não são significantes), embora os dois últimos sejam levemente superiores em EAM (as variações em relação ao FI são significantes a 10%). O modelo com quebras estruturais é significativamente inferior a essas alternativas, padrão repetido em cinco, dez e vinte dias.

A superioridade do modelo com efeitos de assimetria se mantém em cinco dias; o modelo EWMA, entretanto, torna-se adjacente, significativamente superando o modelo FI e não sendo rejeitado pelo teste HPS nos dois critérios. Os modelos FI, AR e HAR permanecem relativamente próximos. Para dez e vinte dias, os resultados são similares entre si: o modelo AE ainda é o melhor em EAM (significativamente superior ao FI no teste HLN), sendo quase idêntico (embora dois décimos inferior) ao EWMA em R^2 . A hipótese nula do teste HPS deixa de ser rejeitada a 5% nos modelos FI e EWMA em EAM e R^2 e no modelo AR em EAM. No geral, o modelo HAR fica atrás por uma diferença diminuta.

De volta aos demais papéis, no horizonte de dez dias o modelo AE é o

melhor em EAM 60% para das ações analisadas (Tabela D.7), sendo transposto quatro vezes pelo EWMA, uma vez pelo HAR e uma vez pelo FI. À exemplo do que ocorre em previsões de um dia, os modelos FI, HAR e EWMA não se destacam consistentemente como segundos melhores, embora o último leve alguma vantagem. Por outro lado, um padrão divergente surge pelo R^2 (Tabela D.9): o modelos EWMA e AE são melhores treze e duas vezes, respectivamente.

Tabela 5.1: Previsões: Um e Cinco dias

| 1 dia | EAM | HLN | HPS | R^2 | HLN | HPS | F |
|-------|-------|-------|-------|-------|-------|-------|-------|
| FI | 0,338 | - | 0,000 | 0,619 | - | 0,025 | 0,158 |
| GARCH | 0,489 | 0,000 | 0,000 | 0,370 | 0,000 | 0,000 | 0,002 |
| AR | 0,333 | 0,091 | 0,001 | 0,619 | 0,498 | 0,005 | 0,001 |
| EWMA | 0,337 | 0,347 | 0,033 | 0,614 | 0,223 | 0,059 | 0,228 |
| AE | 0,321 | 0,000 | 0,903 | 0,643 | 0,004 | 0,855 | 0,009 |
| SB | 0,365 | 0,000 | 0,000 | 0,593 | 0,017 | 0,000 | 0,000 |
| HAR | 0,331 | 0,026 | 0,015 | 0,620 | 0,416 | 0,006 | 0,000 |

| 5 Dias | EAM | HLN | HPS | R^2 | HLN | HPS |
|--------|-------|-------|-------|-------|-------|-------|
| FI | 0,413 | - | 0,026 | 0,479 | - | 0,203 |
| GARCH | 0,527 | 0,000 | 0,000 | 0,289 | 0,000 | 0,000 |
| AR | 0,410 | 0,206 | 0,015 | 0,468 | 0,068 | 0,022 |
| EWMA | 0,406 | 0,097 | 0,397 | 0,493 | 0,032 | 0,647 |
| AE | 0,396 | 0,000 | 0,917 | 0,500 | 0,012 | 0,862 |
| SB | 0,474 | 0,000 | 0,000 | 0,425 | 0,000 | 0,000 |
| HAR | 0,411 | 0,329 | 0,026 | 0,468 | 0,069 | 0,026 |

A tabela reporta previsões fora da amostra para a volatilidade da IBM no período 2000-2003 (983 dias úteis, excluindo datas afetadas por feriados), onde cada modelo é reestimado diariamente e utilizado para gerar previsões para a volatilidade realizada 1, 5, 10 e 20 dias adiante. EAM indica o erro absoluto médio. O R^2 (corrigido) reportado é o da regressão $RV_t = \alpha + \beta f_{t,i} + \varepsilon_{t,i}$, onde $f_{t,i}$ é a previsão do modelo i para a volatilidade realizada do dia t e RV_t é a volatilidade realizada observada naquele dia. A coluna F mostra os p-valores dos testes F robustos para a hipótese conjunta que $\alpha = 0$ e $\beta = 1$. HLN e HPS exibem os p-valores dos testes de Harvey, Leybourne and Newbold (1997) –em comparação com o modelo FI– e Hansen (2005) respectivamente.

Diante das fortes evidências de quebra estrutural em 2003 –ano em que a volatilidade realizada da ação decresceu consistente e fortemente de forma sem precedentes na amostra, induzindo autocorrelações nos resíduos de todos os modelos–, além de variações na persistência da série ao longo do tempo, a Tabela 5.3 abaixo (restrita aos modelos FI e com efeitos de assimetria) examina o impacto dos resultados ano a ano (2000-2002) na análise. No quadro, um, dois ou três asteriscos adjacentes aos números nas colunas EAM e/ou R^2 indicam que o modelo tem EAM/ R^2 inferior em relação à alternativa de forma

Tabela 5.2: Previsões: Dez e Vinte dias

| 10 dias | EAM | HLN | HPS | R^2 | HLN | HPS |
|---------|-------|-------|-------|-------|-------|-------|
| FI | 0,463 | - | 0,211 | 0,371 | - | 0,477 |
| GARCH | 0,555 | 0,000 | 0,000 | 0,230 | 0,000 | 0,001 |
| AR | 0,458 | 0,237 | 0,107 | 0,359 | 0,094 | 0,041 |
| EWMA | 0,463 | 0,461 | 0,291 | 0,390 | 0,029 | 0,845 |
| AE | 0,447 | 0,003 | 0,929 | 0,388 | 0,047 | 0,796 |
| SB | 0,532 | 0,000 | 0,000 | 0,314 | 0,000 | 0,001 |
| HAR | 0,466 | 0,324 | 0,028 | 0,354 | 0,026 | 0,037 |
| 20 Dias | EAM | HLN | HPS | R^2 | HLN | HPS |
| FI | 0,524 | - | 0,435 | 0,238 | - | 0,578 |
| GARCH | 0,591 | 0,000 | 0,000 | 0,149 | 0,000 | 0,005 |
| AR | 0,517 | 0,243 | 0,223 | 0,230 | 0,235 | 0,089 |
| EWMA | 0,535 | 0,091 | 0,188 | 0,253 | 0,107 | 0,836 |
| AE | 0,507 | 0,011 | 0,958 | 0,251 | 0,146 | 0,794 |
| SB | 0,604 | 0,000 | 0,000 | 0,172 | 0,000 | 0,002 |
| HAR | 0,534 | 0,168 | 0,001 | 0,227 | 0,154 | 0,088 |

estatisticamente significante (teste de Harvey, Leybourne e Newbold, 1997) a 10%, 5% e 1% respectivamente.

Em 2000, o modelo com efeitos de assimetria é superior em previsões de um dia pelos dois critérios (significante a 5% no caso do R^2), enquanto as duas estatísticas divergem para cinco, dez e vinte dias: o modelo FI supera o modelo em árvore pelo critério de EAM (significante a 1% em dez e vinte dias) e o contrário ocorre com o R^2 . A contradição sugere um nível de volatilidade não capturado pelas estimações do modelo com efeitos de assimetria, que de outra forma exibiu capacidade superior de responder a variações na volatilidade. Em 2001 e 2002, o modelo com efeitos de assimetria supera consistente e fortemente o modelo FI em todos os horizontes e critérios.

As estatísticas para 2003 são dadas pela Tabela 5.4 e respaldam amplamente a hipótese de quebra estrutural. Todos os modelos são rejeitados nos testes F. Para previsões de um dia, os modelos HAR e AE levam vantagem sobre os demais, embora as diferenças não sejam grandes em relação aos modelos AR e EWMA –o modelo FI fica um pouco atrás, sendo significativamente inferior. Em vinte dias, os modelos EWMA e HAR são significativamente superiores aos outros em erros absolutos, enquanto o modelo FI é significante superior em R^2 .

A tabela revela que os erros absolutos médios em 2003 foram consideravelmente inferiores aos dos anos anteriores, sugerindo uma variância encolhida para a própria volatilidade realizada no período –de fato, as previsões de 20

Tabela 5.3: Previsões por Anos: 2000-2002

| | 2000 | | 2001 | | 2002 | |
|---------|----------|---------|----------|----------|----------|----------|
| | EAM | R^2 | EAM | R^2 | EAM | R^2 |
| 1 dia | | | | | | |
| FI | 0,459 | 0,309 | 0,373 | 0,504 | 0,352 | 0,618 |
| AE | 0,451 | 0,336** | 0,350*** | 0,550*** | 0,328*** | 0,644*** |
| 5 dias | | | | | | |
| FI | 0,536 | 0,129 | 0,465 | 0,390 | 0,454 | 0,357 |
| AE | 0,547 | 0,153* | 0,420*** | 0,405 | 0,428*** | 0,432*** |
| 10 dias | | | | | | |
| FI | 0,567*** | 0,082 | 0,537 | 0,233 | 0,525 | 0,190 |
| AE | 0,608 | 0,095 | 0,485*** | 0,250 | 0,479*** | 0,288*** |
| 20 dias | | | | | | |
| FI | 0,605*** | 0,016 | 0,634 | 0,097 | 0,583 | 0,062 |
| AE | 0,633 | 0,024 | 0,567*** | 0,114 | 0,529*** | 0,148*** |

A tabela reporta os resultados para previsões fora da amostra geradas pelos modelos com efeito de assimetria e FI para a volatilidade da IBM em cada ano entre 2000-2002. EAM indica o erro absoluto médio. O R^2 (corrigido) reportado é o da regressão $RV_t = \alpha + \beta f_{t,i} + \varepsilon_{t,i}$, onde $f_{t,i}$ é a previsão do modelo i para a volatilidade realizada do dia t e RV_t é a volatilidade realizada observada naquele dia. A coluna F mostra os p-valores dos testes F robustos para a hipótese conjunta que $\alpha = 0$ e $\beta = 1$. HLN e HPS exibem os p-valores dos testes de Harvey, Leybourne and Newbold (1997) –em comparação com o modelo FI– e Hansen (2005) respectivamente.

dias para o modelo FI apresentam menores erros absolutos do que as previsões um passo à frente em todos os anos prévios. Finalmente, cabe ressaltar o desempenho inferior do próprio modelo com quebras estruturais no período. A aparente contradição pode ser vista sob o prisma da análise de Granger e Hyung (2004), que mostram que a previsão de modelos com quebras estruturais tendem a ser inferiores mesmo quando o processo verdadeiro é desse tipo; como há dificuldades e atrasos na detecção das quebras, métodos com médias móveis se saem melhor, característica partilhada por estimações FI espúrias.

5.2

Valor em Risco

A avaliação de previsões de valores em risco é centrada nos testes de razão de verossimilhanças para cobertura condicional e independência de Christoffersen (1998). A análise é similar à desenvolvida em Beltratti e Morana (2005), que estudam os benefícios de valor em risco proporcionados por modelos de memória longa. Embora abordagens com funções de perda também tenham

Tabela 5.4: Previsões: 2003

| 1 dia | EAM | HLN | HPS | R^2 | HLN | HPS | F |
|-------|-------|-------|-------|-------|-------|-------|-------|
| FI | 0,170 | | 0,000 | 0,569 | | 0,159 | 0,000 |
| GARCH | 0,322 | 0,000 | 0,000 | 0,413 | 0,001 | 0,000 | 0,000 |
| AR | 0,159 | 0,000 | 0,010 | 0,589 | 0,076 | 0,410 | 0,003 |
| EWMA | 0,158 | 0,001 | 0,668 | 0,583 | 0,119 | 0,521 | 0,006 |
| AE | 0,157 | 0,002 | 0,785 | 0,598 | 0,067 | 0,892 | 0,000 |
| SB | 0,201 | 0,000 | 0,000 | 0,573 | 0,418 | 0,282 | 0,000 |
| HAR | 0,156 | 0,000 | 0,937 | 0,593 | 0,032 | 0,882 | 0,010 |

| 20 Dias | EAM | HLN | HPS | R^2 | HLN | HPS |
|---------|-------|-------|-------|-------|-------|-------|
| FI | 0,274 | | 0,000 | 0,546 | | 0,873 |
| GARCH | 0,527 | 0,000 | 0,000 | 0,276 | 0,000 | 0,005 |
| AR | 0,207 | 0,000 | 0,001 | 0,478 | 0,000 | 0,021 |
| EWMA | 0,200 | 0,000 | 0,524 | 0,479 | 0,000 | 0,002 |
| AE | 0,236 | 0,000 | 0,000 | 0,482 | 0,008 | 0,045 |
| SB | 0,524 | 0,000 | 0,000 | 0,456 | 0,002 | 0,007 |
| HAR | 0,191 | 0,000 | 0,826 | 0,478 | 0,000 | 0,022 |

A tabela reporta os resultados para previsões fora da amostra geradas por seis modelos para a volatilidade da IBM em 2003 (ano afetado por uma possível quebra estrutural). EAM indica o erro absoluto médio. O R^2 (corrigido) reportado é o da regressão $RV_t = \alpha + \beta f_{t,i} + \varepsilon_{t,i}$, onde $f_{t,i}$ é a previsão do modelo i para a volatilidade realizada do dia t e RV_t é a volatilidade realizada observada naquele dia. A coluna F mostra os p-valores dos testes F robustos para a hipótese conjunta que $\alpha = 0$ e $\beta = 1$. HLN exibe os p-valores dos testes de Harvey, Leybourne and Newbold (1997) para igualdade de médias de funções de perda (na tabela, os erros absolutos e os resíduos da regressão acima), onde os modelos são comparados com o modelo fracionalmente integrado.

sendo implementadas na literatura, estas são descartadas por tenderem a ser viesadas em favor de modelos mais conservadores; diante da arbitrariedade na definição da função de perda, o método é considerado mais apropriado para aplicações específicas e checagens internas de modelos apenas.

Considere inicialmente as previsões de um dia. Seja $\hat{q}_{t|t-1}^i(\alpha)$ a previsão de cobertura para o percentil α do modelo i para o dia $t + 1$, condicional na informação do dia t . Na aplicação, são computados valores em risco de 95% e 99%, isto é, $\alpha = 0.05$ e $\alpha = 0.01$. A sequência de falhas de cobertura na cauda inferior α é definida por:

$$F_{t|t-1} = \begin{cases} 1 & \text{if } r_{t+1} < \hat{q}_{t+1|t}^i(\alpha) \\ 0 & \text{if } r_{t+1} > \hat{q}_{t+1|t}^i(\alpha) \end{cases}$$

(5-1)

onde r_t é retorno observado no dia t . O teste de cobertura incondicional (UC) tem como hipótese nula $E(F_{t+1|t}) = \alpha$ contra $E(F_{t+1|t}) \neq \alpha$, enquanto o teste de independência é construído contra uma alternativa Markoviana de primeira ordem.

As previsões de valor em risco para cinco, dez e vinte dias também são computados em bases diárias, mas como os retornos superpostos tornam os eventos correlacionados, são aplicados, à exemplo de Beltratti e Morana (2005), testes baseados em limites de Bonferroni sugeridos por Diebold, Gunther e Tay (1998). Para previsões de k passos, um teste de tamanho limitado por θ pode ser implementado através de testes individuais de tamanho θ/k em cada uma das k subséries

$$\{F_{1+k|1}, F_{1+2k|1+k}, F_{1+3k|1+2k\dots}\}, \{F_{2+k|2}, F_{2+2k|2+k}, F_{2+3k|2+2k\dots}\}, \dots, \\ \{F_{k-1+k|k-1}, F_{k-1+2k|k-1+k}, F_{k-1+3k|k-1+2k\dots}\},$$

onde a hipótese nula é rejeitada se há uma rejeição em qualquer um delas.

A comparação de valores em risco dos modelos com integração fracional, efeitos de assimetria e quebras estruturais está organizada na Tabela 5.5, mostrando que todos os modelos geram previsões adequadas em todos os horizontes.

Tabela 5.5: Análise de Valor em Risco

| | 1% | | | | 5% | | | |
|---------|--------|-------|---|-------|--------|-------|---|-------|
| | Falhas | UC | R | IND | Falhas | UC | R | IND |
| 1 Dia | | | | | | | | |
| FI | 0,006 | 0,186 | 0 | 0,786 | 0,043 | 0,284 | 0 | 0,134 |
| SB | 0,009 | 0,787 | 0 | 0,683 | 0,043 | 0,284 | 0 | 0,388 |
| AE | 0,010 | 0,957 | 0 | 0,650 | 0,054 | 0,578 | 0 | 0,931 |
| 5 Dias | | | | | | | | |
| FI | 0,011 | 1,000 | 0 | 1,000 | 0,058 | 1,000 | 0 | 1,000 |
| SB | 0,010 | 1,000 | 0 | 1,000 | 0,048 | 1,000 | 0 | 0,423 |
| AE | 0,010 | 1,000 | 0 | 1,000 | 0,059 | 1,000 | 0 | 1,000 |
| 10 Dias | | | | | | | | |
| FI | 0,021 | 0,994 | 0 | 0,576 | 0,082 | 0,154 | 2 | 0,630 |
| SB | 0,017 | 1,000 | 0 | 0,576 | 0,070 | 0,370 | 1 | 0,327 |
| AE | 0,015 | 0,994 | 0 | 1,000 | 0,079 | 0,395 | 1 | 0,697 |
| 20 Dias | | | | | | | | |
| FI | 0,032 | 0,289 | 1 | 1,000 | 0,102 | 0,277 | 8 | 1,000 |
| SB | 0,028 | 0,289 | 2 | 1,000 | 0,082 | 0,311 | 2 | 1,000 |
| AE | 0,027 | 1,000 | 0 | 1,000 | 0,111 | 0,134 | 8 | 1,000 |

A tabela reporta os resultados para previsões de valor em risco fora da amostra geradas pelos modelos com efeitos de assimetria, integração fracional e quebras estruturais dentro do período 2000-2003 (983 observações), onde cada modelo é reestimado diariamente e utilizado para o cálculo de valores em risco de 95% e 99% por simulação condicional. As colunas *Falhas* indicam a proporção de dias onde os retornos nos próximos 1, 5, 10 ou 20 dias se situaram abaixo da cauda inferior α da distribuição prevista. Note que as porcentagens de falhas em 5, 10 e 20 são afetadas por sequências superpostas de retornos. UC e IND indicam os p-valores dos testes de razão de verossimilhanças para cobertura incondicional e independência (contra uma série de Markov de primeira ordem) desenvolvida por Christoffersen (1998). Para 5, 10 e 20 dias, são usados testes baseados em limites de Bonferroni sugeridos por Diebold, Gunter and Tay (1998). R é o número de subséries rejeitadas a 5% (dentro 1, 5, 10 ou 20 apropriadamente).