



Victor de Lucca

**Avaliando Previsões de Volatilidade Realizada no
Brasil Usando um Grande Conjunto de Preditores**

Dissertação de Mestrado

Dissertação apresentada ao Programa de Pós-graduação em
Economia como requisito parcial para obtenção do grau de Mestre
em Economia.

Orientador: Prof. Marcelo Cunha Medeiros

Rio de Janeiro
Setembro de 2018



Victor de Lucca

**Avaliando Previsões de Volatilidade Realizada no
Brasil Usando um Grande Conjunto de Preditores**

Dissertação de Mestrado

Dissertação apresentada ao Programa de Pós-graduação em Economia como requisito parcial para obtenção de grau de Mestre pelo Programa de Pós-Graduação em Macroeconomia e Finanças do Departamento de Economia do Centro de Ciências Sociais da PUC-RIO. Aprovado pela Comissão Examinadora abaixo assinada

Prof. Marcelo Cunha Medeiros

Orientador

Departamento de Economia – PUC-Rio

Prof. Cristiano Augusto Coelho Fernandes

Departamento de Engenharia Elétrica – PUC-Rio

Prof. Ricardo Pereira Masini

FGV

Rio de Janeiro, 24 de setembro de 2018

Todos os direitos reservados. É proibida a reprodução total ou parcial do trabalho sem a autorização da universidade, do autor e do orientador

Victor de Lucca

**Graduou-se em Ciências econômicas da PUC-RIO em
2012.**

Ficha Catalográfica

Lucca, Victor de

Avaliando previsões de volatilidade realizada no Brasil usando um grande conjunto de preditores / Victor de Lucca ; orientador: Marcelo Cunha Medeiros. – 2018.

41 f. ; 30 cm

Dissertação (mestrado)–Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Economia, 2018.

Inclui bibliografia

1. Economia – Teses. 2. Volatilidade realizada. 3. HAR. 4. Regularização. 5. Séries temporais. 6. Avaliação de métodos preditivos. I. Medeiros, Marcelo Cunha. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Economia. III. Título.

CDD: 330

Agradecimentos

A realização desta dissertação representa a conclusão de uma fascinante e árdua jornada que moldou significativamente minha trajetória acadêmica e profissional. Cabe aqui agradecer a todos que me ajudaram a completar esta importante fase de meu desenvolvimento. Não por acaso, começo cronologicamente agradecendo aos meus pais, Ellen Jessouroun e Vicente Eugênio de Lucca Neto, por todo suporte e amor incondicional que me foi dado, desde quando me recordo por gente. Meu eterno agradecimento a estas 2 pessoas maravilhosas que acreditaram fielmente, investiram duas vidas em minha formação como pessoa, e me proporcionaram todo o ambiente propício para que eu pudesse chegar aqui hoje. Obviamente não estaria aqui sem eles. Em seguida agradeço ao coordenador do Mestrado Profissional de Finanças e Macroeconomia da PUC-RIO, e meu orientador, Marcelo Medeiros por toda a paciência, empenho e pelo voto de confiança conferido à mim, na minha determinação e força de vontade, desde o início do mestrado até a conclusão deste trabalho. Agradeço enormemente meus grandes colegas de trabalho, Gabriela Mourão, que me deu a oportunidade de me inserir em um ambiente totalmente relacionado à minha área de estudo. Agradeço à Luiza Gazola e ao Francisco Fonseca por serem grandes parceiros neste meu projeto, sendo totalmente compreensivos com meus horários, ocasionalmente heterodoxos, e por proporcionarem todo o suporte para que eu pudesse evoluir de forma determinante na execução desse trabalho.

Resumo

de Lucca, Victor; Medeiros, Marcelo Cunha. **Avaliando Previsões de Volatilidade Realizada no Brasil Usando um Grande Conjunto de Preditores**. Rio de Janeiro, 2018. 41p. Dissertação de mestrado – Departamento de Economia, Pontifícia Universidade Católica do Rio de Janeiro.

Neste estudo avaliamos a literatura referente à modelagem de volatilidade, com foco na metodologia da volatilidade realizada. Implementamos dois métodos de estimação para a previsão de um dia, da volatilidade realizada de uma ação representativa do mercado brasileiro, a ETF BOVA11. Inicialmente usamos o método considerado benchmark para previsão de RV, o modelo auto regressivo heterogêneo, ou HAR. Em seguida, implementamos a modelagem preditiva por métodos de estimação de modelos de regularização e seleção variáveis. O grande benefício desses modelos é a capacidade de lidar com enorme número de variáveis (candidatas) implementando uma função perda para estimação dos coeficientes de variáveis candidatas. Nossos resultados apresentam evidências a favor da superioridade dos modelos obtidos por regularização sobre o modelo HAR. Especialmente quando calibramos os modelos por regularização de modo a não penalizar 2 variáveis (componentes do modelo HAR) dentre as variáveis disponíveis na base.

Palavras-chave

Volatilidade Realizada; HAR; regularização; séries temporais; avaliação de métodos preditivos.

Abstract

de Lucca, Victor; Medeiros, Marcelo Cunha. **Forecast Evaluation of Brazilian Realized Volatility Using Many Predictors**. Rio de Janeiro, 2019. 41p. Dissertação de mestrado – Departamento de Economia, Pontifícia Universidade Católica do Rio de Janeiro.

The motivation of this dissertation is to evaluate the literature regarding volatility modeling, focusing on the methodology of realized volatility and to apply the concepts developed in the literature to the case of a representative asset of the Brazilian stock market. We have implemented two estimation methods for the one-day forecast of the volatility of the BOVA11. As empirical evidence from the literature points out, our preliminary analysis suggests that the asset exhibits a long-range dependence. Thus, we rely on heterogeneous autoregressive processes (HAR) modeling, considered a benchmark for RV prediction, as well as predictive modeling by regression through regularization methods. The great benefit of these models is the ability to deal with an enormous amount of variables (candidates) by implementing a loss function to estimate the final coefficients of relevant variables, and, therefore, selecting the specific dependency structure for the studied process. Our results indicate the superiority of models obtained through regularization, especially when the models are tuned not to penalize 2 variables (components from the HAR) among all the variables from the database.

Keywords

Realized Volatility; HAR; regularization; time series; evaluation of predictive methods.

Sumário

1	Introdução.....	14
2	Métodos e Materiais	18
2.1	Volatilidade Realizada	18
2.2	Metodologias e Modelos.....	20
2.2.1.	HAR.....	20
2.2.2.	Regularização	22
2.3	Base de Dados.....	25
3	Resultados	31
4	Discussão e Conclusão	36
5	Referências	38

Lista de Gráficos

Gráfico 1 – R2 fora da amostra - Todas as variáveis penalizadas	33
Gráfico 2 - R2 fora da amostra - Lag 1 e Lag 5 não penalizados	34
Gráfico 3 - R2 fora da amostra - Lag 5 não penalizado	35

Lista de Tabelas

Tabela 1- Descrição de Variáveis Seleccionadas.....	27
Tabela 2 - Autocorrelação - 5 minutos.....	29
Tabela 3 - MAPEs Modelos.....	31
Tabela 4 - Teste Diebold Mariano Bicaudal	32
Tabela 5 - Teste Diebold Mariano Unicaudal	32
Tabela 6 – Teste SPA – Cenário com todas as variáveis penalizadas	32
Tabela 7 – Teste SPA – Cenário com lag 1 e lag 5 não penalizados	32
Tabela 8 – Teste SPA – Cenário com lag 5 não penalizados.....	32

1 Introdução

Embora o tema de estudo da volatilidade seja prolífico e seu conceito amplamente usado no mercado financeiro, sua definição pode ser considerada ambígua.

Pode-se dizer que a volatilidade é uma forma de quantificar a incerteza, a probabilidade de um estado de natureza mudar rapidamente e de forma imprevisível. Possivelmente por isso, nos munimos dos mais complexos artifícios para endereçar essa questão. Especificamente, a modelagem e previsão de volatilidade é um assunto que recebe muita atenção atualmente, tanto no meio acadêmico quanto no mercado financeiro.

O controle de risco é responsabilidade de importantes departamentos de todas grandes empresas, financeiras ou não, onde a modelagem de volatilidade se mostra como ferramenta absolutamente fundamental.

Riscos associados com ativos negociados em mercados futuros, por exemplo, podem deixar investidores excessivamente expostos a flutuações do mercado e instituições podem facilmente falir, caso a gestão de seu risco seja feita de forma incauta.

Quanto ao ferramental desenvolvido para lidar com esses riscos, exemplo mais expressivo são os contratos futuros e títulos de derivativos, onde as oscilações destes preços se dão exclusivamente em função da volatilidade atribuída à seus respectivos ativos. A partir da possibilidade de se gerir esse componente não observável surge a demanda, quase imperativa, por parte dos agentes econômicos que passam então a fazer escolha sob qual risco incorrer, e em que medida. Modelos de gestão de risco são utilizados por instituições financeiras e, finalmente, exigidas por reguladores financeiros.

Se a volatilidade é uma forma de quantificar a incerteza, pode-se de dizer que existem dois métodos amplamente difundidos de se medir a volatilidade, a saber: Volatilidade implícita e volatilidade histórica ou volatilidade realizada(RV).

A primeira abordagem, volatilidade implícita, modela a volatilidade como variável fim, uma vez que a variância condicional é não diretamente observável. A ideia do modelo Black-Scholes (1973) de precificação de derivativos, como bom exemplo, é partir de variáveis observadas, como valores de preços e juros, e inferir sobre a volatilidade subjacente àquelas condições. Essa é a abordagem de modelos como Generalized Autoregressive Conditional Heteroskedasticity, ou GARCH, desenvolvidos por Engle (1982) e Bollerslev (1986), modelos de volatilidade estocástica (SV) desenvolvido por Taylor (1986) e modelos de média móvel exponencial(EWMA). Ver McAleer (2005) para exposição de ampla variedade de modelos de volatilidade univariados e multivariados, condicionais e estocásticos, modelos de volatilidade. Asai, McAleer and Yu (2006) para uma revisão da crescente literatura relacionada à modelos estocásticos multivariados).

Porém, existem alguns fatores que podem ser considerados negativos à solução de volatilidade observável. Esses modelos são complexos do ponto de vista de implementação se comparados com a metodologia de RV. E, ainda assim, existe muita discussão sobre a eficiência dos mesmos, uma vez que estes não são capazes de descrever ou reproduzir características importantes da distribuição dos retornos quadrados. Principalmente, estes não conseguem descrever a baixa, e lentamente decrescente, auto correlação dos retornos quadrados. Entre outros, Bollerslev (1987), Malmsten e Teräsvirta (2004), Carnero, Peña e Ruiz (2004) discutem extensivamente essa questão.

A Alternativa à essa modelagem é tratar a volatilidade como variável observável. Inicialmente desenvolvida por Merton (1980), a teoria consiste na ideia que a variância ao longo de um intervalo fixo de tempo pode ser estimada, com precisão, como a soma dos retornos quadrados de um dado ativo, desde que os dados disponíveis estejam em frequência de amostragem suficientemente alta.

Tratar a volatilidade como variável observada pode ser algo bastante animador, pois assim pode-se modelar diretamente essa variável, sem precisar depender de modelos complexos de volatilidade implícita. Além disso, os modelos de RV se mostram aptos a capturar as características desejadas da distribuição de probabilidade dos retornos.

Resultados obtidos em Barndorff-Nielsen e Shephard (2002), Andersen, Bollerslev, Diebold e Labys (2003) e Meddahi (2002), abordam as propriedades desta medida de volatilidade, construída a partir de dados em alta frequência.

O ponto principal de atenção, porém, no contexto de volatilidade realizada, é referente ao método de obtenção desse estimador. Efeitos de microestrutura podem afetar esse estimador, resultando em perda de consistência e eficiência dos mesmos. Entre outros, Corsi (2009), Zhang, Mykland e Aït-Sahalia (2005), Bandi e Russell (2006), Hansen e Lunde (2006), Barndorff-Nielsen, Hansen, Lunde, e Shephard (2008) discutem várias soluções para este problema, e em seguida vamos também endereçar esses pontos com sua merecida atenção.

O objetivo do nosso estudo é implementar alguns modelos da literatura mais atual que lida com alta dimensionalidade no contexto da volatilidade realizada no mercado financeiro brasileiro, avaliando possíveis ganhos comparativos desses modelos frente ao modelo HAR, desenvolvido por Corsi (2009) e considerado benchmark para esse tipo de estimação.

Usando uma ação representativa do mercado brasileiro, o ETF BOVA11, no horizonte de previsão de 1 dia, avaliamos a performance e relevância dos estimadores de cada modelo. Inicialmente usamos o modelo auto regressivo heterogêneo, HAR. Como usualmente implementado, incluímos 3 componentes (AR(1)) de defasagem, cada um com uma periodicidade diferente: diária, semanal e mensal. Essa abordagem é motivada pela teoria dos mercados heterogêneos (HMH), apresentada inicialmente por Muller(1997), atribuindo as características distribuições de probabilidade (com caudas pesadas e choques com efeitos persistentes ao longo do tempo) à ampla gama de diferentes agentes econômicos, com objetivos e visões diferentes, que influem na formação de preço.

Em seguida, sob o mesmo objeto de previsão, implementamos os modelos lasso, ridge e elastic net. Estas metodologias muito parecidas entre si, os 3 buscam selecionar variáveis relevantes usando uma penalidade na função de perda, cada um com sua função de perda diferente. Foi observado na literatura que estas metodologias reduzem significativamente os erros e a variância de previsão fora da amostra (“out-of-sample”).

A base de dados disponível foi composta por transações em alta frequência a partir das ações mais líquidas do índice Bovespa, além de indicadores macroeconômicos variados, além de grande número de defasagens das mesmas. O período da base de dados é de março de 2016 a Novembro de 2017, resultando em 558 observações.

Assim, ao incluir variáveis construídas a partir de metodologias alternativas, como por exemplo índices VIX (índices de risco criados pela Chicago Board Exchange), podemos verificar a relevância de cada uma delas para nossa previsão. Além disso, nos permite argumentar sobre a relevância de cada metodologia à luz da teoria e quanto à seleção e estimação de modelos.

Para avaliação e comparação entre diferentes metodologias econométricas é adotada uma janela rolante de forma a gerar sub-amostras a partir da base de dados. Em seguida, a partir da métrica do erro absoluto percentual médio (“MAPE”), são avaliadas as distribuições dos erros dos modelos implementados para nossas amostras e assim comparados os regressores selecionados. Implementamos os testes de Diebold-Mariano (2013) (DM) e o teste de capacidade preditiva de Hansen (2006) (SPA) para avaliação das previsões geradas por cada modelo.

Nossos estudos contribuem para a literatura a medida que nossos resultados apresentam evidências a favor da hipótese de que é possível superar o modelo HAR, usando a metodologia de regularização para a previsão da volatilidade realizada. Adicionalmente, questionamos premissas comuns com relação ao suposto caráter de memória longa do processo de volatilidade realizada e destacamos padrões identificados nas variáveis selecionadas, a luz da teoria econômica.

A dissertação está organizada da seguinte maneira. Na seção 2, descrevemos os Métodos e Materiais, com a definição de volatilidade realizada, descrição das metodologias de modelagem e da base de dados utilizadas. Seção 3, exibimos os resultados do estudo. Por fim, a seção 4 conclui o estudo com discussões e conclusões finais sobre o tema.

2 Métodos e Materiais

2.1 Volatilidade Realizada

Para ilustrar o conceito de volatilidade nocional (ou real), para o dia t $\sigma(t)$, vamos considerar o seguinte processo estocástico:

$$dp(t) = \mu(t)dt + \sigma(t)dW(t)$$

Onde, $p(t)$ é o logaritmo do preço, $\mu(t)$ é um processo contínuo e finito, $W(t)$ é um movimento Browniano padrão, e $\sigma(t)$ é um processo estocástico independente de $dW(t)$. Neste processo de difusão, a volatilidade nocional para o dia t é a integral da volatilidade instantânea durante o intervalo de tempo de um dia $(t-1d;t)$, ou seja:

$$\sigma_t^{(d)} := \left(\int_{t-1d}^t \sigma^2(\omega) d\omega \right)^{1/2}$$

Merton (1980) mostrou que a volatilidade integrada de um movimento browniano pode ser aproximada precisamente usando a soma de retornos (intradiaários) quadrados. Mais recentemente, Andersen et al. (2001), aplicando a teoria da variação quadrática, generalizou esse resultado para classe especial (de média finita) semimartingales. Esta classe engloba processos aplicados na precificação de ativos sem arbitragem como processo de difusão de Itô, e difusões de salto misto.

Em teoria, sob tais condições, a soma dos retornos quadrados intradiaários convergem (à medida que a duração máxima dos retornos chega a zero) para a volatilidade dos preços, em probabilidade, nos permitindo construir uma estimativa livre de erros da volatilidade real sobre um intervalo de tempo fixo. Este estimador não paramétrico é chamado “volatilidade realizada”. E a definição padrão da volatilidade realizada, para uma série de retornos igualmente espaçados sobre um intervalo de tempo de um dia é:

$$RV_t^{(d)} = \sqrt{\sum_{j=0}^{M-1} r_{t-j\Delta}^2}$$

onde M é o período, ou a quantidade de partições dos retornos diários, $\Delta = \frac{1d}{M}$ e $r_{t-j\Delta} = p(t - j\Delta) - p(t - (j + 1)\Delta)$ definem retornos contínuos compostos de frequência Δ , ou, os retornos intradiários de intervalo Δ .

2.2 Metodologias e Modelos

2.2.1. HAR

Á luz da teoria do mercado heterogêneo, apresentado por Muller(1997) e inicialmente implementado por Corsi(2002), o HAR modela a volatilidade através de uma estrutura de agregação de componentes “AR(1)s” heterogêneos. Estes componentes aditivos são calculados como as médias diárias da volatilidade, obtidas à partir de períodos de tamanhos diferentes. Além disso, estes componentes apresentam caráter hierárquico, ou “em cascata” onde, à cada nível, a volatilidade parcial prevista depende da volatilidade passada naquele nível, e também, de componentes em níveis superiores. Ilustrando o exemplo descrito, para o componente semanal (com 5 dias úteis) do HAR temos:

$$RV_t^{(w)} = \frac{1}{5} \left(RV_{t-1d}^{(d)} + RV_{t-2d}^{(d)} + \dots + RV_{t-5d}^{(d)} \right)$$

De forma mais detalhada, o modelo hierárquico considerado possui 3 componentes de volatilidade correspondentes para horizontes temporais, um dia (1d), uma semana (1w) e um mês (1m). A partir de Corsi(2002), caracterizamos estes componentes como os estimadores de volatilidade parcial, e seu objetivo é modelar os reais componentes não observados da volatilidade, denotados respectivamente de $\hat{\sigma}_t^{(d)}$, $\hat{\sigma}_t^{(w)}$ e $\hat{\sigma}_t^{(m)}$.

Supondo que as dinâmicas do mercado são exclusivamente determinadas pelos agentes econômicos, o processo estocástico pode ser determinado pelo componente de maior frequência na cascata, com $\hat{\sigma}_t^{(d)} = \sigma_t^{(d)}$ como volatilidade integrada de um dia. Assim esperamos ter:

$$r_t = \sigma_t^{(d)} \varepsilon_t$$

$$\text{Com } \varepsilon_t \sim NID(0,1)$$

Ou seja, esperamos que eventuais diferenças entre o estimador e a variável não sejam previsíveis. E assim podemos definir:

$$\begin{aligned}\hat{\sigma}_{t+1m}^{(m)} &= c^{(m)} + \phi^{(m)}RV_t^{(m)} + \hat{w}_{t+1m}^{(m)} \\ \hat{\sigma}_{t+1w}^{(w)} &= c^{(w)} + \phi^{(w)}RV_t^{(w)} + \gamma^{(w)}E\left[\hat{\sigma}_{t+1m}^{(m)}\right] + \hat{w}_{t+1w}^{(w)} \\ \hat{\sigma}_{t+1d}^{(d)} &= c^{(d)} + \phi^{(d)}RV_t^{(d)} + \gamma^{(d)}E_t\left[\hat{\sigma}_{t+1w}^{(w)}\right] + \hat{w}_{t+1d}^{(d)}\end{aligned}$$

Onde “ $RV^{(\cdot)}$ ” representam os componentes da volatilidade realizada (observável) e os componentes “ $\hat{w}_{t+1}^{(\cdot)}$ ” componentes de erros. Assim, de forma recursiva podemos chegar à:

$$\sigma_{t+1d}^{(d)} = c^{(d)} + \beta^{(d)}RV_t^{(d)} + \beta^{(w)}RV_t^{(w)} + \beta^{(m)}RV_t^{(m)} + \hat{w}_{t+1d}^{(d)} \quad (1)$$

Essa equação representa a medida de volatilidade diária a partir das modelagens parciais da volatilidade. Assim, o processo da forma funcional do modelo para a volatilidade integrada é obtida a partir da volatilidade realizada como:

$$\sigma_{t+1d}^{(d)} = RV_{t+1d}^{(d)} + w_{t+1d}^{(d)} \quad (2)$$

Onde $w_t^{(d)}$ representa a volatilidade não observada e erros de medida.

Essa equação faz referência a um dos aspectos importantes que devemos ter ao lidar com o modelo HAR. Por mais que exista uma extensa literatura que esse modelo captura os movimentos da volatilidade implícita, não podemos garantir que o fator $w_{t+1d}^{(d)}$ é de fato um ruído branco. Por isso ressalta-se a importância dos cuidados discutidos nas sessões anteriores quanto à amostragem dos dados. A equação acima é a conexão entre a estimativa $RV_{t+1d}^{(d)}$ e a medida da volatilidade não observada $\sigma_{t+1d}^{(d)}$. Substituindo (2) em (1) obtemos o modelo proposto:

$$\mathbf{HAR} : RV_{t+1d}^{(d)} = c + \beta^{(d)}RV_t^{(d)} + \beta^{(w)}RV_t^{(w)} + \beta^{(m)}RV_t^{(m)} + w_{t+1d}^{(d)}$$

$$\text{com } w_{t+1d} = \hat{w}_{t+1d}^{(d)} + w_{t+1d}^{(d)}$$

Algumas características desse modelo são especialmente interessantes, por exemplo, o fato estilizado que os retornos da volatilidade financeira são processos de memória longa onde grandes choques de volatilidade não são rapidamente esquecidos. Ao comparar o modelo à metodologias semelhantes, o mesmo apresenta vantagens importantes, tanto no que tange à implementação e à fundamentação teórica. Por exemplo, produz previsões que com acurácia semelhante, ou melhor, serem

extremamente simples do ponto de vista da implementação e, possibilitando fácil interpretabilidade econômica.

Para citar alguns exemplos, nos modelos ARCH e GARCH a autocorrelação diminui exponencialmente quando deveria ser um decaimento hiperbólico. De fato, nossos dados sugerem um decaimento lento da autocorrelação em retornos diários quadrados, até o atraso de 100, ou aproximadamente 5 meses.

2.2.2. Regularização

Métodos de estimação por regularização fazem parte do campo da estatística chamado de “estatística esparsa” de acordo com Trevor Hastie, Robert Tibshirani e Martin Wainwright no livro “*Statistical Learning with Sparsity – The Lasso and Generalizations*”. De acordo com os autores:

“Um modelo estatístico esparsos consiste em ter apenas um pequeno número de parâmetros ou pesos diferentes de zero dentre os estimadores selecionados. Representa o caso clássico do chamado “menos é mais”. Um modelo esparsos pode ser muito mais fácil de estimar e interpretar do que um modelo denso. Nesta era de big data, o número de características medidas em uma pessoa ou um objeto podem ser grandes e podem ser maiores do que o número de observações. A suposição de esparsidade dos modelos nos permite atacar tais problemas e extrair padrões úteis e reproduzíveis de grandes conjuntos de dados.”

O ponto inicial são os modelos convencionais de Mínimos Quadrados ordinários, baseado na minimização da soma dos erros quadráticos, porém com um fator adicional de restrição, representada pela norma L_q de β :

$$\min_{\beta_0, \beta} \left\{ \frac{1}{2N} \sum_{i=1}^N \left(y_i - \beta_0 - \sum_{j=1}^p x_{ij} \beta_j \right)^2 \right\} \quad (3)$$

sujeito a $\sum_{j=1}^p p_\lambda(|\beta_j|) \leq t$.

A restrição t pode ser considerada como um “orçamento” na distribuição dos pesos distribuídos aos coeficientes β_j do modelo. Assim podemos limitar o ajuste dos dados aos coeficientes. O valor de λ (ou t), por sua vez, é selecionado por BIC.

Essa metodologia é especialmente útil nos modelos esparsos, onde a quantidade de variáveis efetivamente relevantes é significativamente menor do que o número de

variáveis candidatas, e quando temos maior número de variáveis candidatas do que observações.

Nesta metodologia estamos abrindo mão do conceito expresso no Teorema de Gauss-Markov dos estimadores não viesados (BLUE) e passamos a usar estimadores viesados, porém com a interessante característica que estes modelos estimados com métodos de regularização apresentam estimadores com variância significativamente menor do que os modelos obtidos via MQO. Essa abordagem em se buscar o menor grau de variância dos estimadores, em detrimento da preocupação com viés é chamado de “bias-variance tradeoff”.

Isso porque na regressão múltipla por MQO, é mostrado que estimativas de parâmetros baseadas em soma residual dos quadrados tem uma alta probabilidade de ser insatisfatória, se não incorreta, se os vetores de previsão não forem ortogonais.

Com o objetivo de controlar a complexidade dos modelos estimados e prevenir a multicolinearidade nos mesmos a restrição de t (equação (3)) é implementada a fórmula da solução por MQO. Um valor de t alto atua penalizando pouco os parâmetros, permitindo que o modelo se adapte mais aos dados. Por outro lado, um valor t baixo atua de forma mais restritiva especificando modelos mais esparsos. Assim, é de especial interesse obter o valor de t que gera o modelo mais preciso, em termos de previsão a partir de amostras independentes a partir da mesma população. Essa precisão é chamada de habilidade de generalização do modelo. Um valor t muito baixo pode impedir que o lasso capture os sinais relevantes a partir dos dados. Por outro lado, um valor de t muito alto pode levar a um modelo sobrespecificado. Há geralmente um modelo entre esses dois valores que produz a melhor previsão, eventualmente com alguns dos coeficientes da regressão iguais a zero.

Uma relativa simples alteração à função de minimização da soma dos quadrados dos erros, que passa agora a conter um termo adicional, o termo de penalidade. Esse termo impõe limites superiores ao tamanho (da soma) dos betas/estimadores forçando os coeficientes de estimadores não adequados à valores irrelevantes, ou mesmo zerando-os.

Ficamos, portanto, com o seguinte estimador de mínimos quadrados penalizados:

$$\hat{\beta}(\lambda) = \arg \min \sum_{t=1}^T (y_t - \beta' x_t)^2 + \sum_{i=1}^n p\lambda(|\beta_i|),$$

Onde $p\lambda(|\beta_i|)$ é a função de penalidade utilizada, e esta varia de acordo com a metodologia utilizada.

Cronologicamente a metodologia Ridge foi a primeira a ser desenvolvida, e essa usa a norma L2 para penalizar a soma do quadrado dos erros, ou seja:

$$\hat{\beta}_{Ridge}(\lambda) = \arg \min \sum_{t=1}^T (y_t - \beta' x_t)^2 + \lambda \sum_{i=1}^n \beta^2$$

Como os componentes da equação são quadráticos em β , a solução é facilmente encontrada:

$$\hat{\beta}_{Ridge}(\lambda) = (zz' + \lambda I)^{-1} zy.$$

Apesar de reduzir significativamente os coeficientes de regressores irrelevantes, ou altamente correlacionados, a metodologia Ridge apresenta coeficientes viesados e não consistentes para seus regressores.

Em seguida foi desenvolvida a metodologia Lasso Thibishiani (1996), que usa a norma L1 para penalidade:

$$\hat{\beta}_{lasso}(\lambda) = \arg \min \sum_{t=1}^T (y_t - \beta' x_t)^2 + \lambda \sum_{i=1}^n |\beta_i|$$

Assim como o ridge, essa metodologia apresenta estimadores viesados e não consistentes em geral. Porém, ao implementar essa função de penalidade observa-se que a mesma apresenta algumas particularidades importantes quando comparado ao Ridge. Vantagens: (1) O lasso é capaz de efetivamente zerar coeficientes com alta correlação, frequentemente removendo variáveis não cruciais para a regressão e (2) pode se estimar eficientemente o “regularization path”. Desvantagens: (3) lasso irá selecionar, no máximo, a mesma quantidade de variáveis que o número de observações da base, (4) caso um conjunto de variáveis seja altamente correlacionado o lasso tende a selecionar, arbitrariamente, apenas uma das variáveis e (5) para o caso usual com mais observações

do que variáveis e variáveis correlacionadas, foi observado empiricamente que o Ridge domina o lasso em termos de performance de previsão (Tibshirani, (1996)).

A partir dessas duas funções de penalização, ou regularização, foram derivadas também outras funções adaptadas, a partir das descritas acima. Basicamente com o objetivo de corrigir problemas descritos acima.

O elastic-net é uma função híbrida, entre a função ridge e o lasso, onde é determinada uma ponderação entre o termo quadrático e o termo linear, termo α .

Assim como o lasso, o Elastic Net é um método de regularização e seleção de variáveis. Evidências empíricas demonstram que esse método é superior ao lasso em termos de capacidade preditiva (Zou 2004). O elastic net é especialmente útil quando o número de variáveis supera muito o número de observações.

$$\hat{\beta}_{EN}(\lambda) = \operatorname{argmin} \sum_{t=1}^T (y_t - \beta' z_t)^2 + \alpha \sum_{i=1}^n |\beta_i| + (1 - \alpha) \sum_{i=1}^n \beta_i^2,$$

2.3 Base de Dados

A base de dados deste estudo é composta por 1768 variáveis, com 588 observações diárias (sem considerar suas defasagens). A partir de dados de transações de ativos do mercado acionário em alta frequência, de transações ocorridas com 5 min de intervalo, disponíveis no site da Bovespa, foram calculadas 16 séries de volatilidade diária e 15 medidas de covariância (ação i x BOVA11). Adicionalmente, além das variáveis geradas a partir de dados disponibilizados pelo website da BMF&Bovespa (http://www.b3.com.br/pt_br/), incluímos variáveis de indicadores macroeconômicos de juros, câmbio, níveis em volume e níveis do índice IBovespa e cotações de commodities de ouro, obtidos a partir do site do banco central do Brasil. E, finalmente, índices de volatilidade implícita VIX para diferentes ativos, obtidos a partir do site da CBOE. O cálculo deste indicador se baseia em uma média ponderada de preços de contratos de opções de compra e venda para diferentes vencimentos e preços de exercício. Essa metodologia foi inicialmente desenvolvida para medir níveis de risco implícito ao índice de ações do mercado norte americano S&P500. Atualmente essa metodologia é aplicada não somente a este índice dessas ações, mas também a outros ativos e portfólios

relacionados à outros mercados, ou até mesmo relacionados à ações, individuais ou relacionadas à mercados. Neste estudo usamos algumas séries históricas com essa metodologia, a saber: VIX: ETF de mercados de países emergentes (ETF_EM), ETF Brasil, ETF China, além do próprio VIX(S&P500) e o VIX do Google. Adicionalmente, incluímos 40 lags à todas as variáveis descritas acima. E por fim, foi incluída uma variável dummy na base, para o dia de 18/05/2018. Esse dia houve um *Circuit Breaker* na BOVESPA, e como a amplitude do impacto foge totalmente diferente do que se pode considerar “ordinário” decidimos, por controlar essa observação.

Aqui, a escolha de variáveis começa a direcionar algumas hipóteses a serem testadas em nosso estudo. Ao incluir índices VIX em nossa análise por exemplo, buscamos testar a relevância da abordagem de volatilidade implícita sobre a métrica de volatilidade realizada. A relação com a taxa de juros deveria também estar relacionada com a volatilidade realizada, a medida que os juros refletem o custo de oportunidade do dinheiro no tempo, que está diretamente relacionado à incerteza e, portanto, à volatilidade. Para testar a relevância da modelagem de referência, são incluídos também a nossos estudos os componentes do modelo HAR. Esses são variáveis que consistem na média histórica da volatilidade realizada, a partir de diferentes intervalos, 1 dia, uma semana (5 dias) e um mês(22 dias). Finalmente, covariâncias e variâncias condicionais de outras ações (em relação à ação BOVA11) deveriam também se mostrar como variáveis relevantes.

Para calcular a variância e covariância das séries, foi usada base de dados que contém todos os preços de negociação das ações negociadas na BMF&Bovespa. A séries temporais vão de dezembro/2015 à dezembro/2017 e foram selecionadas as 16 ações listadas na Tabela 1.

Tabela 1- Descrição de Variáveis Seleccionadas

Ativo - Código	Nome\Definição	Detalhe/Setor
BOVA 11	ISHARES BOVA	Fundo de Índice - Ibovespa
lag1	Lag 1 do RV BOVA11	Componente modelo HAR
lag5	Média movel 1 semana RV BOVA11	Componente modelo HAR
lag22	Média movel 1 mês RV BOVA11	Componente modelo HAR
ABEV3	Ambev PN	Consumo
BBAS3	Banco do Brasil PN	Financeiro
BBDC4	Bradesco PN	Financeiro
BBSE3	Banco do Brasil Seguridade PN	Financeiro
BRFS3	BR Foods PN	Consumo
CIEL3	Cielo PN	Meios de Pagamento / Financeiro
ITSA4	Itau SA ON	Financeiro
ITUB4	Itau Unibanco Holding ON	Financeiro
KROT3	Kroton PN	Educacional
LREN3	Lojas Renner	Consumo
PETR4	Petrobras PN	Commodities
PETR3	Petrobras ON	Commodities
UGPA3	Supermercado Pão de Açúcar PN	Consumo
VALE3	Vale PN	Commodities
VIVT4	Vivo ON	Consumo
Commodities e Indicadores Macroeconômicos		
BMF_GOLD	Cotação 1 grama de Ouro - A vista BMF	
LND_GOLD	Cotação 1 onça troy de Ouro - Londres	
BV_INDX_VOL	Volume diário Bovespa	
BV_INDX_PNTS	Nível diário Bovespa (fechamento)	
CDI	Certificado de Depósito Interbanário	
SELIC	Taxa Selic Meta diária	
FX	Taxa Ptax \$USD/\$BRL	
Chicago Board of Options Exchange - Índices de risco VIX		
BZ ETF_VIX	Fundo de índice que replica o risco Brasil	
CH ETF_VIX	Fundo de índice que replica o risco China	
EM ETF_VIX	Fundo de índice que replica o risco de Mercados emergentes	
GOO_VIX	VIX do Google	

O critério utilizado foi selecionar 70% da composição do Índice Bovespa, selecionando os ativos com maior participação no Índice, de forma a selecionar os ativos mais líquidos, dentre os disponíveis no mercado brasileiro de ações. A intenção aqui é evitar problemas relacionados à microestrutura de dados, trabalhando na maior frequência possível e minimizando possíveis vieses que poderiam surgir em decorrência de lidar com ações de baixa liquidez.

Adicionalmente, a metodologia de previsão fora a amostra (“out-of-sample”) utiliza validações de robustez com janelas rolantes, e isso exige um grande número de observações para funcionar corretamente. Verificou-se que para o período estudado e variáveis selecionadas, conseguimos atender a essa requisição.

Ao examinar a composição de ativos da amostra selecionada, percebemos uma relativa heterogênea gama de setores, ilustrando a modesta diversificação da indústria brasileira. Destaca-se a concentração na indústria de materiais básicos e outras indústrias, como empresas financeiras e de bens de consumo.

Selecionadas as variáveis a serem incluídas em nosso estudo, devemos nos atentar agora para o melhor método de cálculo das volatilidades realizadas. Na seção anterior, foram discutidos os problemas de se calcular a volatilidade realizada a partir de uma frequência muito alta.

A escolha por usar uma frequência de 5 min é muito usada na literatura, e é também nossa escolha para o presente estudo. Medeiros (2014), Fleming et al.(2003), Andersen et al.(2000), entre outros, e também em estudos aplicados no mercado de ações usaram também essa metodologia.

Liu, Patton, Sheppard (2012) realizam um extenso trabalho focado exclusivamente na avaliação de medidas de volatilidade. Foram avaliadas a performance de 7 metodologias de cálculo de volatilidade realizada, aplicadas à 31 ativos financeiros e 5 frequências diferentes e conclui-se que se o a escolha dos 5 min de intervalo se não é consenso absoluto em termos de performance, certamente é o melhor período de amostragem a ser adotado na maioria dos casos.

Após a limpeza da base de dados, identificamos os preços de transação mais próximos da grade de 5 minutos. Como todas as ações têm alta liquidez, é justo considerar que este preço permanece válido até o final de uma dada grade de 5 min. Ao diferenciar primeiro os preços de log para todas as grades, obtemos os retornos de 5 minutos.

Tabela 2 - Autocorrelação - 5 minutos

AÇÃO	AUTOCORRELAÇÃO SERIAL - Lags:			
	1	2	5	10
BOVA11	0.002893	0.006256	0.00169	0.002455
ABEV3	0.018658	0.010712	0.005224	0.001069
BBAS3	0.003124	0.064769	0.000725	0.001367
BBDC4	0.000817	0.091226	0.005046	0.006109
BBSE3	0.004377	0.002067	0.009926	0.004452
BRFS3	0.025588	0.011565	0.005544	0.004312
CIEL3	0.013011	0.001421	0.001199	0.000279
ITSA4	0.043344	0.001854	0.000956	0.001227
ITUB4	0.001616	0.119707	0.014916	0.003672
KROT3	0.007154	0.008249	0.07955	0.001935
LREN3	0.16063	0.001736	0.001266	0.001131
PETR3	0.008759	0.020938	0.056385	0.010187
PETR4	0.004052	0.076987	0.022392	0.002695
UGPA3	0.026116	0.013294	0.038486	0.002684
VALE3	0.040305	0.017264	0.023817	0.026739
VIVT4	0.013303	0.012225	0.008825	0.004811

Na tabela 2, podemos ver que, auto correlações são insignificantes relativamente próximas de zero na maioria dos casos, e apesar de se manterem diferentes de zero, até o 20 lags atrás.

Com uma dinâmica de média condicional insignificante, como inferido a partir das funções de auto correlação exibidas acima, o modelo para retornos diários pode ser escrito como $r_t^{(d)} = \sigma_t^{(d)} \epsilon_t$ onde $\epsilon_t \sim \text{iid } N(0; 1)$. Portanto, o retorno de 1 dia é condicionalmente gaussiano com variância igual à variância integrada. A normalidade de ϵ_t é justificada recorrendo ao Teorema do Limite Central, argumentando que os retornos agregados sobre um período razoável de tempo deveriam tender a normalidade. Assim a medida RV estima adequadamente a volatilidade integrada σ_t , os retornos padronizados por RV devem ser distribuídos normalmente com variância de 1. A Tabela 3 mostra que é o caso do nosso estimador de volatilidade realizado.

Nos referimos a Liu et al.(2013) para calcular as volatilidades realizadas. Estes autores concluíram que é difícil superar a simples medida de RV de 5 minutos, mesmo sem correção para a microestrutura ou o uso de informações tick-by-tick. A RV é definida para cada ação i e dia τ como:

$$RV_{i,\tau} = \sum_{t=1}^T r_{i,\tau,t}^2 \quad (4)$$

Onde $r_{i,\tau,t}$ é o retorno de 5-min da ação i , BOVA11 para o nosso estudo, e τ a quantidade de partições dentro de um dia. As covariâncias são estimadas como descritas abaixo (com relação ao BOVA11) e definida para cada par de ações (i, j) , no dia τ como:

$$RCOV_{i,j,\tau} = \sum_{t=1}^T r_{i,\tau,t} r_{j,\tau,t} \quad (5)$$

Onde $r_{i,\tau,t}$ é o retorno de 5-min da ação i , e τ a quantidade de partições totalizando um dia.

3 Resultados

Para prever a volatilidade realizada do BOVA11 usamos dois grupos de modelos. O primeiro grupo é composto pelo simples modelo HAR e o segundo grupo é composto por 12 modelos, subdividido, por sua vez, em 3 grupos, Lasso, Ridge e Elastic Net. Estes são modelos por regularização e diferem entre si apenas pela função de penalização utilizada para seleção das variáveis candidatas. O parâmetro α das funções de penalização pondera a função entre as soluções do Ridge (com $\alpha = 0$) e Lasso ($\alpha = 1$). Além dos casos do Lasso e Ridge (casos extremos para os α 's de solução convexa), incluímos mais 10 diferentes valores para α , de 0.1 a 0.9.

Ainda sobre o processo desenvolvido para este estudo, segmentamos a análise em 3 cenários. No primeiro, todas as variáveis foram penalizadas livremente de acordo com o critério de informação BIC. Na segunda implementação, as variáveis “lag1”, a volatilidade defasada do da volatilidade do BOVA11 e o “lag5”, média rolante de 5 dias, também da volatilidade do BOVA11, não foram penalizadas. E na terceira implementação apenas a variável “lag 5” não teve seu coeficiente penalizado. Ao não penalizar uma dada variável esta tem seu coeficiente estimado livremente por mínimos quadrados ordinários.

Dentro de cada cenário, partir das previsões obtidas, calculamos o erro absoluto médio de cada janela de estimação. A partir disso, exibimos abaixo estatísticas descritivas destas medidas de erro, testes de Diebold-Marcus de e o teste SPA de Hansen(2006).

Tabela 3 – MAPEs Modelos

MAPE	HAR	RIDGE	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	LASSO
todos penalizados	0,2042	0,2534	0,2809	0,2508	0,2334	0,223	0,2202	0,2172	0,2145	0,2123	0,2121	0,2124
LAG 5 + 1 FIXO	0,2042	0,1897	0,2028	0,1897	0,1418	0,1076	0,0985	0,0666	0,0486	0,0367	0,0294	0,0247
LAG5 FIXO	0,2042	0,1994	0,2045	0,203	0,176	0,15	0,2007	0,1581	0,0954	0,0678	0,0506	0,0395

Tabela 4 – Teste Diebold Mariano Bicaudal

DM Test Bicaudal	RIDGE	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	LASSO
todos penalizados	0,000	0,000	0,000	0,000	0,014	0,021	0,043	0,097	0,176	0,173	0,146
LAG 5 + 1 FIXO	0,009	0,457	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
LAG5 FIXO	0,468	0,904	0,725	0,000	0,000	0,368	0,000	0,000	0,000	0,000	0,000

No DM Test Bicaudal, H0 é de que ambas as previsões são iguais em capacidade preditiva.

Tabela 5 – Teste Diebold Mariano Unicaudal

DM Test Unicaudal	RIDGE	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	LASSO
todos penalizados	1,00	1,00	1,00	1,00	0,99	0,99	0,98	0,95	0,91	0,91	0,93
LAG 5 + 1 FIXO	0,00	0,23	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
LAG5 FIXO	0,23	0,55	0,36	0,00	0,00	0,18	0,00	0,00	0,00	0,00	0,00

No DM Test Unicaudal, H0 é de que a previsão da coluna não é superior ao modelo HAR

Tabela 6 – Teste SPA – Cenário com todas as variáveis penalizadas

SPA Test - Todos							
Coefficientes penalizados	Rank_M	v_M	MCS_M	Rank_R	v_R	MCS_R	Loss
HAR	1	-1,313856967	1	1	-1,202169475	1	0,204154535
0_7	5	1,90335193	0,0804	5	3,362122675	0,0054	0,214521857
0_8	2	0,792085527	0,567	2	1,202169475	0,6366	0,21233322
0_9	3	0,830163871	0,5424	3	1,228099492	0,6188	0,212137094
LASSO	4	1,242000319	0,3058	4	1,32751582	0,542	0,212449822

Tabela 7 – Teste SPA – Cenário com lag 1 e lag 5 não penalizados

SPA Test - Lag 1 + Lag 5 FIXOS							
	Rank_M	v_M	MCS_M	Rank_R	v_R	MCS_R	Loss
LASSO	1	-20,8568	1	1	-20,8568	1	0,024707

Tabela 8 – Teste SPA – Cenário com lag 5 não penalizado

SPA Test - Lag 5 FIXO							
	Rank_M	v_M	MCS_M	Rank_R	v_R	MCS_R	Loss
LASSO	1	-14,3393	1	1	-14,3393	1	0,039465

Alguns pontos se evidenciam a partir dos resultados. Primeiro, existem implementações pelo método de regularização que superam significativamente a performance preditiva do modelo HAR. Segundo, os modelos por regularização tendem a melhorar seu desempenho, com o aumento do α . Terceiro, é marcante a diferença de performance entre os cenários implementados. Especificamente para os cenários com variáveis não penalizadas, alguns dos modelos obtidos por regularização superam significativamente o modelo HAR.

Os gráficos 1, 2 e 3 exibem os indicadores de ajuste R2 para os modelos em cada cenário.

Gráfico 1 – R2 fora da amostra – Todas as variáveis penalizadas

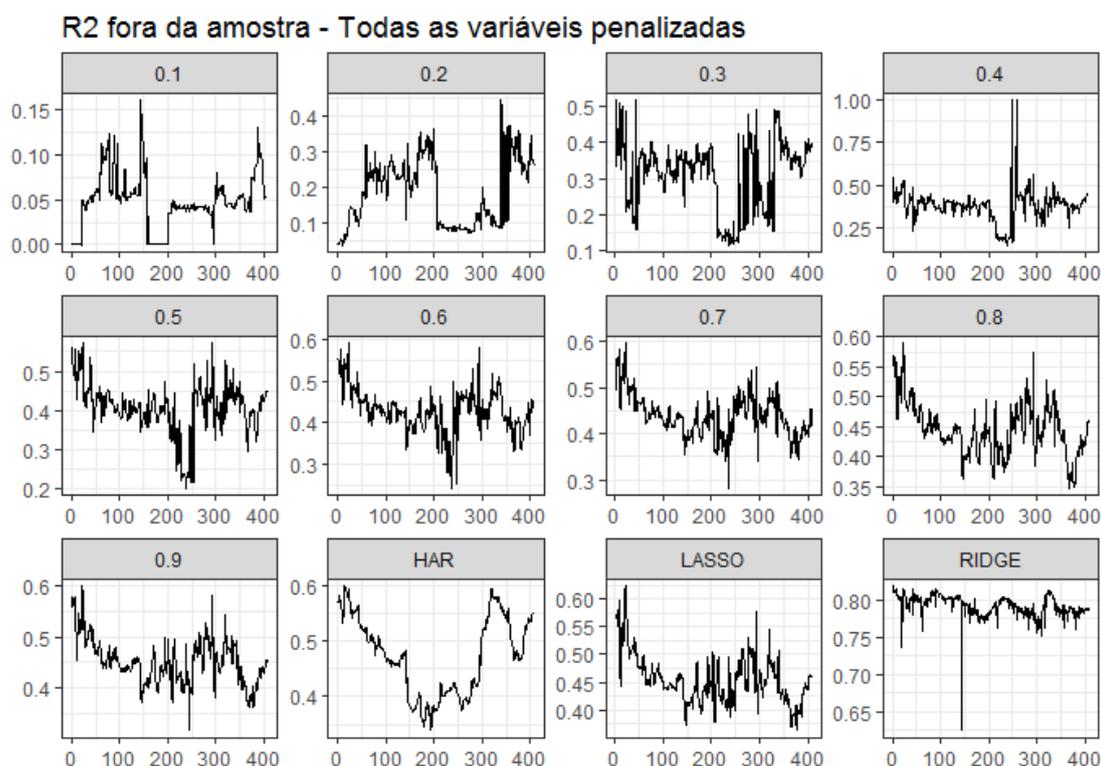


Gráfico 2 – R2 fora da amostra – Lag 1 e lag 5 sem penalização

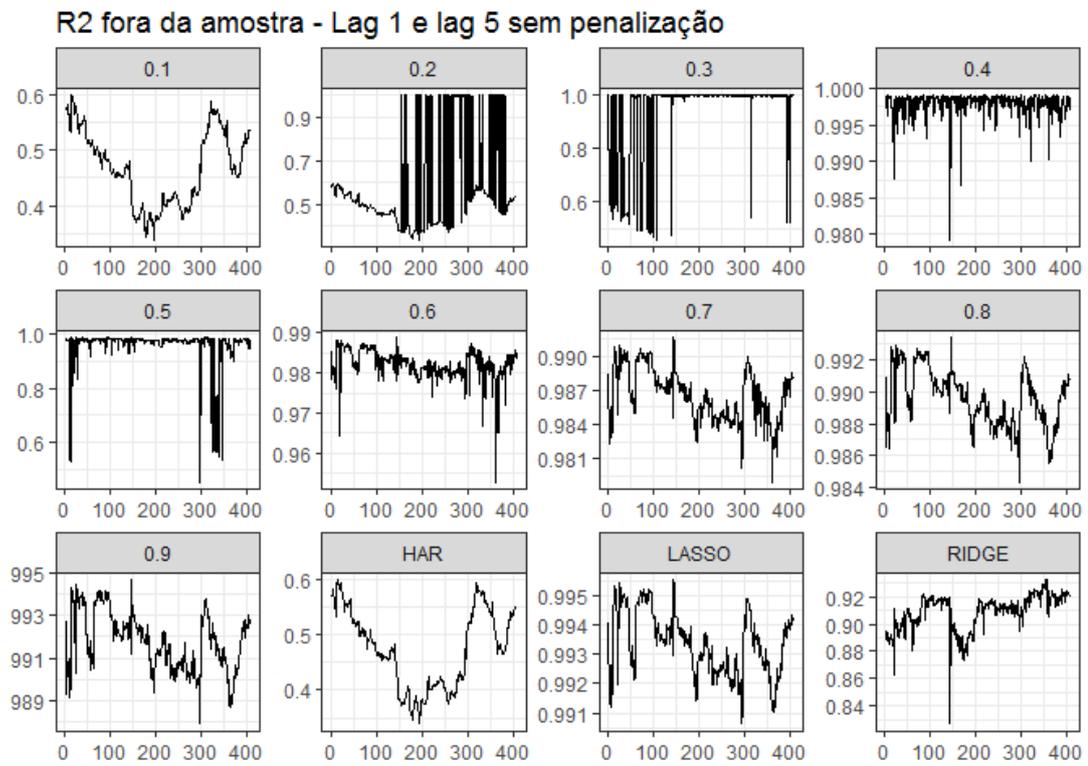
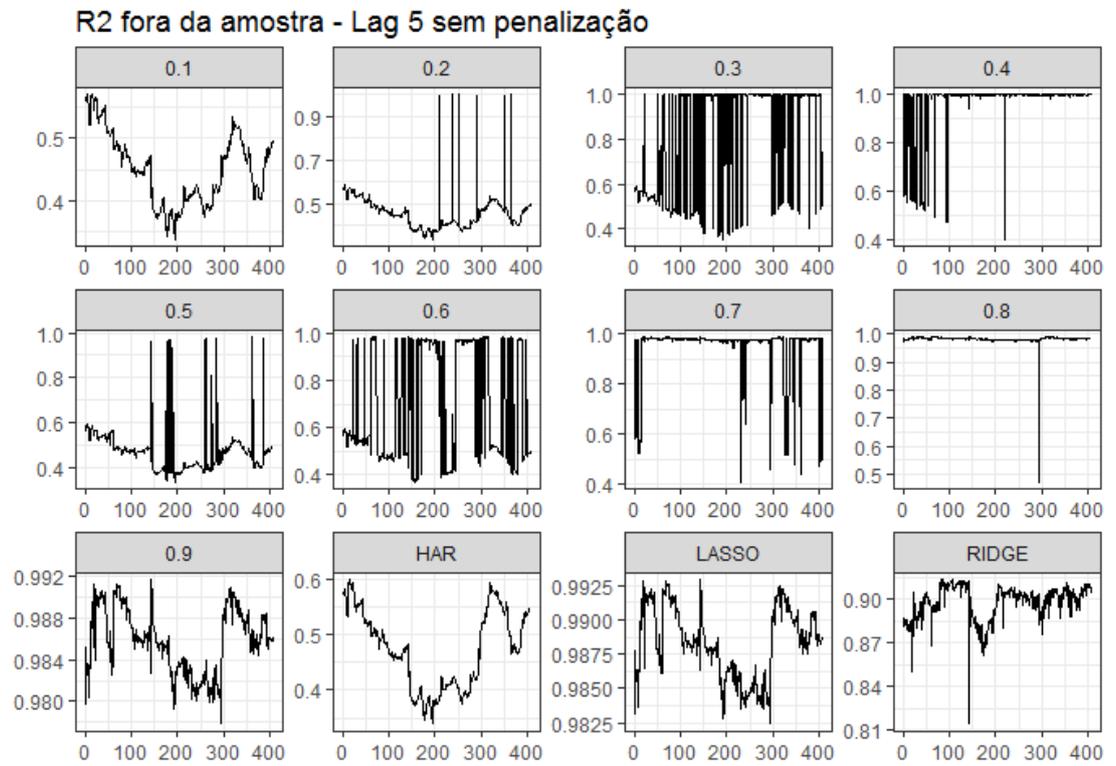


Gráfico 3 – R2 fora da amostra – Lag 5 sem penalização



4 Discussão e Conclusão

Ao examinar os resultados acima, podemos destacar alguns pontos. Primeiro, é marcante a diferença da acurácia dos modelos por regularização, nos diferentes cenários testados. Para todas as métricas observadas, em que componentes do modelo HAR não são penalizadas apresentam resultados significativamente melhores do que o modelo HAR. Especificamente, isso fica evidente com os valores de MAPE notadamente menores do que o modelo HAR, p-valores dos testes Diebold Mariano muito próximos de zero para a maioria dos casos testados e os resultados dos testes SPA de Hansen(2006), que nestes cenários, chegam a selecionar apenas o lasso, como melhor modelo. Adicionalmente, temos coeficientes R2 fora da amostra, com valores muito próximos de 1, também na maioria dos casos testados.

A partir desses resultados, podemos destacar alguns pontos para discussão e conclusões. O primeiro é que os componentes do modelo HAR se mostraram, inegavelmente, importantes para a previsão da volatilidade realizada, pelo menos para a base testada. Segundo, curiosamente, ao não penalizar os coeficientes de variáveis que são incluídas no HAR, a metodologia de regularização conseguiu chegar a modelos com poucas variáveis, mas ainda de caráter heterogêneo e isso contribuiu bastante para a melhora da acurácia de um cenário para o outro. Assim, atentamos para a possibilidade dos modelos obtidos por regularização contribuírem para o tema de previsão de volatilidade por serem capazes de melhorar significativamente a acurácia destas previsões ao selecionar variáveis de caráter heterogêneo e complementares aos componentes do modelo HAR.

Segundo, chama também atenção a diferença de resultados das implementações em cada cenário, especificamente entre os modelos por regularização. Modelos do tipo elastic net tendem a apresentar resultados piores do que seus pares ridge e lasso com a diferença diminuindo a medida que o parâmetro alpha cresce.

Terceiro, os resultados do estudo feito exibem um panorama animador do ponto de vista da evolução na metodologia de previsão da volatilidade realizada. Porém, é importante também destacar a possibilidade dos resultados serem particulares a base testada. Assim sendo, embora possamos concluir que nossos estudos foram capazes de alcançar resultados significativamente melhores do que as previsões obtidas pelo modelo HAR, não podemos generalizar o resultado para todo o contexto de previsão de

volatilidade realizada sem termos um espectro de teste maior. Uma possibilidade seria aumentar a base, atualizando-a com os dados mais recentes, disponíveis no mercado, ou simulando uma base maior, a partir da base disponível.

5 Referências

- Andersen, T. G., Bollerslev, T., Diebold, F. X., & Ebens, H. (2001). The distribution of realized stock return volatility. *Journal of Financial Economics*, 61, 43–76.
- Andersen, T. G., Bollerslev, T., Diebold, F. X., & Labys, P. (Março de 2003). Modeling and Forecasting Realized Volatility. *Econometrica*, p. 579-625.
- Andersen, T., & Bollerslev, T. (1998). Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts. *International Economic Review*, 39, 885-906.
- Andersen, T., Bollerslev, T., & Diebold, F. X. (2003). Modeling and Forecasting Realized Volatility. *Econometrica*, 71, 579-625.
- Andersen, T., Bollerslev, T., Diebold, F. X., & Ebens, H. (2001). The Distribution of Realized Stock Return Volatility. *Journal of Financial Economics*, 61, 43-76.
- Asai, M., McAleer, M., & J., Y. (2006). Multivariate Stochastic Volatility: A Review. *Econometric Reviews*, 25, 145-175.
- Bandi, F. M., & Russell, J. R. (2006). Separating Market Microstructure Noise from Volatility. *Journal of Financial Economics*, 79, 655-692.
- Barndorff-Nielsen, O., & N., S. (2002). Econometric Analysis of Realised Volatility and its Use in Estimating Stochastic Volatility Models. *Journal of the Royal Statistical Society B*, 64, 253-280.
- Barndorff-Nielsen, O., Hansen, P., Lunde, A., & Shephard, N. (2008). Designing Realized Kernels to Measure the Ex-post Variation of Equity Prices in the Presence of Noise. *Econometrica*, 76, 1481-1536.
- Bollerslev, T. (1987). A Conditionally Heteroskedastic Time Series Model for Speculative Prices and Rates of Return. *Review of Economics and Statistics*, 69, 542-547.
- Carnero, M. A., Peña, D., & E., R. (2004). Persistence and Kurtosis in GARCH and Stochastic Volatility Models. *Journal of Financial Econometrics*, 2, 319-342.

- Carvalho, M. C., Freire, M. A., Medeiros, M. C., & Souza, L. R. (2005). Modeling and Forecasting the Volatility of Brazilian Asset Returns: A Realized Variance Approach. *Revista Brasileira de Finanças*, pp. 321-343.
- Corsi, F. (2002). A Simple Approximate Long Memory Model of Realized Volatility. *Journal of Financial Econometrics*, 7, 174-196.
- Corsi, F., Zumbach, G., Müller, U., & Dacorogna, M. (2001). Consistent high-precision volatility from high frequency data. *Economic Notes*, 30, 183–204.
- Diebold, F. X. (2013). *Comparing Predictive Accuracy, Twenty Years Later: A Personal Perspective on the Use and Abuse of Diebold-Mariano Tests*. Pennsylvania: University of Pennsylvania.
- Elliott, G., Gargano, A., & Timmermann, A. (april de 2013). Complete subset regressions. *Journal of Econometrics*, 177, pp. 357–373.
- Engle, R. F. (1982). Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, 50, 987-1007.
- Ferreira, P. G., & de Mattos, D. M. (s.d.). *Instituto Brasileiro de Economia (FGV/IBRE)*. Fonte: Site do Instituto Brasileiro de Economia.
- Fleming, J., Kirby, C., & Ostdiek, B. (2003). The economic value of volatility timing using “realized volatility”. *Journal of Financial Economics*, 67, 473–509.
- Garcia, M. G., Medeiros, M. C., & Santos, F. E. (setembro de 2014). Economic Gains of Realized Volatility in the Brazilian Stock Market. *Revista Brasileira de Finanças (Online)*, 319-349.
- Garcia, M. G., Medeiros, M. C., & Vasconcelos, G. F. (June de 2011). *Real-time inflation forecasting with high-dimensional models: The case of Brazil*. Fonte: International Journal of Forecasting: www.elsevier.com/locate/ijforecast
- Hansen, P. R., & Lunde, A. (2006). Realized Variance and Market Microstructure Noise” (with discussion). *Journal of Business and Economic Statistics*, 24, 127-218.
- Harris, L. (1990). Estimation of stock variance and serial covariance from discrete observations. *Journal of Financial and Quantitative Analysis*, 25, 291–306.

- HILLEBRAND, E., & MEDEIROS, M. C. (2009). *THE BENEFITS OF BAGGING FOR FORECAST MODELS OF REALIZED VOLATILITY*.
- Liuy, L., Pattony, A. J., & Sheppard, K. (5 de December de 2012). *Does Anything Beat 5-Minute RV? A Comparison of Realized Measures Across Multiple Asset Classes*. Fonte: Duke University and University of Oxford.
- Malmsten, H., & T., T. (2004). *Stylized Facts of Financial Time Series and Three Popular Models of Volatility*. Working Paper Series in Economics and Finance 563, Stockholm School of Economics.
- McAleer, M. (2005). Automated Inference and Learning in Modeling Financial Volatility. *Econometric Theory*, 21, 232-261.
- McAleer, M., & Medeiros, M. C. (2008). REALIZED VOLATILITY: A REVIEW. *Econometric Reviews*, 10–45.
- McAleer, M., & Medeiros, M. C. (s.d.). *Forecasting Realized Volatility with Linear and Nonlinear Models*. Fonte: PUC-RIO, DEPARTAMENTO DE ECONOMIA: www.econ.puc-rio.br
- Meddahi, N. (2002). A Theoretical Comparison Between Integrated and Realized Volatility. *Journal of Applied Econometrics*, 17, 479-508.
- Medeiros, M. C., & Mendes, E. F. (s.d.). *Estimating High-Dimensional Time Series Models*. Fonte: PUC-RIO, Departamento de Economia: www.econ.puc-rio.br
- Merton, R. C. (1980). On Estimating the Expected Return on the Market: An Exploratory Investigation. *Journal of Financial Economics*, 8, 323-361.
- Müller, U. A., Dacorogna, M. M., Dave, R. D., Olsen, R. B., Puctet, O. V., & von Weizsacker, J. (1997). Volatilities of Different Time Resolutions – Analyzing the Dynamics of Market Components. *Journal of Empirical Finance*, 4, 213-239.
- Taylor, S. J. (1986). *Modelling Financial Time Series*. Wiley: Chichester.
- Zhang, L., Mykland, P., & Y., A.-S. (2002). A Tale of Two Time Scales: Determining Integrated Volatility with Noisy High-Frequency Data. *Journal of the American Statistical Association*, 100, 1394-1411.

Zhou, B. (1996). High frequency data and volatility in foreign-exchange rates. *Journal of Business and Economic Statistics*, 14, 45–52.