

PONTIFÍCIA UNIVERDADE CATÓLICA DO RIO DE JANEIRO

DEPARTAMENTO DE ECONOMIA

MONOGRAFIA DE FINAL DE CURSO

**MODELOS PREDITIVOS DO *EQUITY RISK PREMIUM* NO
MERCADO BRASILEIRO: ROBUSTEZ ESTATÍSTICA E
ANÁLISE DOS PRINCIPAIS FATORES**

Felipe Gama Filho de Carvalho

Nº de matrícula: 1510375

Orientadores: Ruy Ribeiro e Marco Cavalcanti

Dezembro, 2018

PONTIFÍCIA UNIVERDADE CATÓLICA DO RIO DE JANEIRO
DEPARTAMENTO DE ECONOMIA
MONOGRAFIA DE FINAL DE CURSO

**MODELOS PREDITIVOS DO *EQUITY RISK PREMIUM* NO
MERCADO BRASILEIRO: ROBUSTEZ ESTATÍSTICA E
ANÁLISE DOS PRINCIPAIS FATORES**

Felipe Gama Filho de Carvalho

Nº de matrícula: 1510375

Orientadores: Ruy Ribeiro e Marco Cavalcanti

Dezembro, 2018

Declaro que esse trabalho é de minha autoria e que não recorri para realizá-lo, a nenhuma forma de ajuda externa, exceto quando autorizado pelo professor tutor.

As opiniões expressas neste trabalho são de responsabilidade única e exclusiva do autor.

Agradecimentos

À minha família, por me ensinar diariamente o que é amor incondicional.

Aos meus amigos que são como família, por me mostrarem o que é apoiar e estar sempre junto.

Índice

1.	Introdução	6
2.	Revisão da Literatura	8
2.1	Defesa x Inexistência da Previsibilidade de Retornos	8
2.2	Abordagens Diversas para a Previsibilidade de Retornos	11
2.3	Efeitos, Consequências e Possibilidades de Absorção dos Estudos	12
2.4	Welch & Goyal (2008) – “Um olhar compreensivo para a performance empírica de previsão do equity premium”	14
3.	Dados	17
4.	Metodologia	20
5.	Resultados – Cenário Base	23
5.1	Modelos Estatisticamente Insignificantes In-Sample:	32
5.2.	Modelos Estatisticamente Significantes In-Sample:	36
6.	Campbell & Thompson (2005)	40
7.	Modelos com “Memória Estendida”	48
8.	Conclusão	53
	Referências Bibliográficas	54

1. Introdução

Em *A Comprehensive Look at The Empirical Performance of Equity Premium Prediction* (2006), Welch & Goyal buscam fazer uma revisão ampla da principal literatura em torno da previsão do *equity risk premium* – o retorno em excesso do mercado de ações em relação a um referencial tomado como livre de risco – no cenário norte-americano. Os autores partem da premissa de que existe uma clara sensação geral de que “a previsão funciona”, sem que haja clareza em relação a como exatamente isso ocorre.

Olhando para o mercado brasileiro, o entendimento geral da questão pode ser pensado como ainda mais primitivo – não existe literatura comparável que forneça a mesma magnitude de modelos propostos para a previsão, até pelo pouco desenvolvimento do próprio mercado, se comparado à sólida estrutura dos EUA, com dados disponíveis desde o fim do século XIX. Aqui, as instituições, tanto as que compõem o mercado financeiro quanto as empresas cujas ações são (ou já foram) negociadas em bolsa, são menos estáveis, estiveram sujeitas a uma história econômica que passa por diversas moedas oficiais, inflação galopante em períodos curtos de tempo, e dados disponíveis que são muito mais escassos. Ainda assim, já existe uma base suficiente para que ideias similares às aplicadas nos países mais desenvolvidos sejam testadas no Brasil (com as devidas restrições).

Este trabalho se propõe a analisar os mesmos fatores colocados em questionamento por Welch & Goyal (2008) – os principais modelos, a significância estatística e a relevância prática que forneçam mais embasamento em torno da previsibilidade do *equity risk premium*, e se de fato pode se dizer que ela “funciona” ou não. É importante enfatizar as aspas em torno de seu funcionamento, porque mesmo no caso norte-americano, quando os autores encontram um resultado com aparente significância estatística, sugerem ceticismo e uma série de ressalvas antes de tomar isso como dado dali em diante. O mercado acionário por si só, além das demais variáveis que se relacionam com ele, estão em constante mudança, e não há nenhuma garantia de que relações encontradas em dados históricos irão ter aderência no futuro. Especialmente no caso brasileiro, quaisquer resultados (tanto negativos quanto positivos) devem ser encarados como material contínuo de pesquisa e análise – algo a ser revisitado e expandido conforme o mercado evolua ainda mais.

A motivação principal, portanto, gira em torno da fundamentação das bases de uma literatura que já é muito extensa em referência a mercados desenvolvidos como o norte-americano, mas ainda é muito incipiente no caso brasileiro. Através de quaisquer resultados obtidos, o trabalho busca oferecer um olhar mais crítico em torno das análises de previsão. A partir daí, é possível pensar em um legado que inclua a construção de estratégias práticas de atuação (como modelos de *trading* que levem o *equity risk premium* em conta), análise mais rica de conjunturas futuras, ou até um desenvolvimento no próprio entendimento popular do mercado (tanto para os que ainda o enxergam como pura aleatoriedade quanto para os que acreditam em qualquer

capacidade preditiva anunciada como se fosse a invenção da roda, mas muitas vezes falha em sua especificação ou resultados).

2. Revisão da Literatura

Esta seção se divide em quatro itens, cada um com o objetivo de destacar os estudos e artigos de principal influência em torno do tema. São eles (i) o debate acerca da existência concreta ou não da previsibilidade de retornos; (ii) as abordagens diversas para a especificação de modelos preditivos; (iii) os efeitos da divulgação dos trabalhos sobre as próprias relações que buscam analisar, e de que formas são ou podem ser incorporados pelo mercado; e (iv) a influência de Welch & Goyal (2008) como principal base para o projeto.

Dada a extensiva bibliografia, seria possível seguir listando literalmente centenas de referências distintas, mas a ideia aqui é se aprofundar um pouco mais em alguns artigos específicos que trazem consigo intuições, metodologias ou conclusões de influência especial na forma de se pensar o assunto como um todo e fundamentar bases para sua aplicação no contexto brasileiro.

2.1 *Defesa x Inexistência da Previsibilidade de Retornos*

Cochrane (2007) é o ponto de partida – ainda que seja de lançamento relativamente recente, sem dúvidas é um dos mais renomados e discutidos *papers* em torno da previsibilidade de retornos. O autor parte de uma análise de valor presente (assim como outros a serem vistos mais adiante) e acrescenta grande sofisticação argumentativa associada a questões mais técnicas e econométricas. É interessante tê-lo em vista no primeiro momento por se tratar de uma abordagem que, antes da especificação de modelos particulares, tenta focar diretamente na existência da previsibilidade.

Com seu raciocínio engenhoso aliado à econometria da cointegração, ele propõe que a pergunta correta não é se os retornos seriam previsíveis, mas qual variável (ou quanto de cada uma delas), entre retornos e dividendos, teria seu crescimento sujeito a ser predito, e, portanto, passível de explicar a variação observada no *dividend yield* (razão Dividendos/Preços). Dado que essa variável é estacionária, a previsibilidade precisaria ser observada em algum dos termos que a compõe, caso contrário ela não seria “trazida de volta” à sua média e variância após algum choque. Talvez o grande brilhantismo de Cochrane esteja justamente na inversão da lógica comum: não é na observação direta que ele propõe sua defesa da previsibilidade dos retornos, mas sim na forte ausência de previsibilidade dos dividendos.

Voltando quase vinte anos no tempo, Fama e French talvez sejam os responsáveis pela difusão ampla (ou ao menos os principais solidificadores) da noção de previsibilidade dos retornos. Já em 1988 os autores articulavam com clareza que a parcela esperada de retornos de curto prazo (aquela para a qual pode se aplicar a previsibilidade) era muito modesta, mas que para prazos maiores, como a parcela da variação nos retornos de dois a quatro anos, o número já passava a ser mais significativo (inclusive estatisticamente).

Os próprios autores fazem questão de citar (lembrando, ainda em 1988) que seu uso do *dividend yield* como variável preditiva não é novidade e que a ideia já era presente na academia e na prática desde décadas antes, mas oferecem uma abordagem interessante acerca da intuição por trás desse uso e dos diferentes horizontes temporais. Primeiro, explicam que, sob uma ótica de “mercados eficientes”, preços correntes estão baixos em relação aos dividendos quando taxa de desconto e retorno esperado são altos, e vice-versa, de modo que *dividend yields* e retornos esperados tem suas trajetórias relacionadas. Estendendo os horizontes temporais, a alta persistência (alta correlação positiva) dos retornos esperados faz com que sua variância cresça em uma proporção maior do que o crescimento do horizonte.

É importante notar que, nesse contexto, retornos esperados são a parcela indicativa da existência de previsibilidade, e sua variância é medida diretamente pelo R^2 das regressões. O comportamento é o oposto para a variância dos resíduos (a parcela não esperada e, portanto, imprevista) – seu crescimento é atenuado pelo o que chamam de “efeito taxa de desconto”, no qual um choque exógeno nos retornos esperados se relaciona com um choque oposto nos preços correntes. Assim, é como se a parcela esperada dos retornos fosse, ao longo da extensão do horizonte temporal, tomando uma presença maior no todo em relação à sua contraparte inesperada, salientando o aumento no poder de previsão.

Lewellen (2004) reforça as análises anteriores, tanto no âmbito das intuições teóricas e financeiras quanto da estatística. Ainda que seu foco permaneça no *dividend yield*, variável priorizada na literatura, ele generaliza o raciocínio para outras duas métricas semelhantes (valor contábil sobre valor de mercado, B/M, e lucro sobre preço, E/P), que seguem a mesma lógica de comparação entre focos de análise fundamentalista e os respectivos preços (individuais ou agregados). Por estes serem o termo no denominador, as três variáveis deveriam ter uma relação positiva com os retornos esperados.

Além da intuição dos “mercados eficientes” supracitada (que Lewellen chama de “precificação racional”), ele também cita explicação válida para seu oposto, uma ótica de “*mispricing*”, na qual uma métrica como o *dividend yield* estaria baixa quando ações estivessem “*overpriced*” – equivalente a dizer que o termo no denominador estaria superestimado, indicando retornos mais baixos no futuro conforme ele voltasse a estar em linha com a análise fundamentalista (analogamente para o oposto).

O autor parte de um cenário inicial no qual grande parte dos resultados positivos de previsão usando as variáveis acima já havia sido alvo de questionamentos contundentes ou, sob perspectiva mais cética, talvez até desbancados. São incorporados alguns novos pontos focais, como testes condicionais ao grau de autocorrelação do regressor, que reforçam resultados satisfatórios para a previsibilidade. Outro ponto importante surge com a capacidade de alteração dos resultados que é apresentada pela inclusão ou não de diferentes períodos na amostra (com os dados da chamada “bolha da internet” sendo responsáveis por reduções significativas nos valores positivos).

Tomando o caminho oposto – a argumentação que busca enfatizar todas as falhas ou fraquezas envolvidas nas especificações preditivas para o *equity risk premium*, ou para retornos de forma geral – um dos *papers* que melhor suscita e sintetiza essas questões é *The Myth of Long-Horizon Predictability* (Boudoukh, Richardson & Whitelaw 2005). Os autores buscam provocar os fortes indícios para previsibilidade de retornos no longo prazo (ou, num mínimo, mais fortes do que no curto), chamando isso de mito.

Sua motivação principal está na forte autocorrelação entre as regressões, de modo que o aumento do R^2 se torna quase perfeitamente proporcional ao aumento no horizonte temporal – um ponto de partida semelhante a Fama e French (1988), mas com intuítos (ou conclusões) quase que opostos. O que mostram os autores é que alta persistência encontrada nos regressores (por exemplo, uma correlação de quase 99% para os estimadores preditivos de um e dois anos do *dividend yield*) faz com que qualquer erro de amostragem se manifeste da mesma forma ao longo do horizonte temporal.

Algumas das principais linhas de crítica aos estudos de previsibilidade de retornos são colocadas em evidência. Talvez o elemento de mais fácil indicação, o viés de pequenas amostras é abordado e caracterizado com relevância especial quando a relação entre choques contemporâneos aos retornos e a variável usada como preditiva do modelo é de correlação negativa. Seu foco principal, no entanto, se volta especificamente para o debate acerca da previsibilidade nos prazos mais longos. Abordagens mais amplas podem ser vistas, por exemplo, em Stambaugh (1999).

Além de propor um debate acerca do contraste entre a estatística clássica (p-valores, intervalos de confiança) e Bayesiana (dependente de “crenças prévias”), Stambaugh mostra de forma muito simples e intuitiva os vieses que estão envolvidos na abordagem comum de MQO dos modelos preditivos. Segundo ele, a regressão dos preços na variável preditiva em questão, defasada por um período, não apresenta resíduos descorrelacionados com essa variável ao longo do tempo – mais especificamente, um choque exógeno nos preços em dado período t está correlacionado com mudanças na variável preditiva em um intervalo de tempo que inclua esse período t . A abordagem que olha apenas para os preços pode ser facilmente estendida para os retornos, como o próprio autor chega a mencionar, e uma consequência direta é que os estimadores de MQO se tornam enviesados.

Mais além, a relação entre os resultados comumente apresentados para prazos mais curtos (com pouca significância estatística) é posta lado a lado com a inadequação de medidas clássicas para o *goodness-of-fit* dos modelos, devido a um extensivo número de variáveis possíveis a serem escolhidas e testes que não são independentes entre elas. Em *Assessing Goodness-of-Fit of Asset Pricing Models: The Distribution of the Maximal R^2* , referência destacada por Boudoukh et al., Foster, Smith e Whaley (1997) abordam os problemas associados ao que chamam de “*data-snooping*” – os vieses que estão implícitos na própria escolha das “melhores” variáveis dentro dessa gama de

possibilidades muito vasta, quando se analisa sob a ótica das métricas clássicas um modelo que seja fruto dessa seleção.

Uma proposta tão interessante quanto provocativa é apresentada para avaliar o número de variáveis necessárias tal que, ao serem selecionadas aleatoriamente, produzissem determinado R^2 em pelo menos 5% dos casos. Ao confrontar essa avaliação com estudos passados, alguns resultados são positivos (quando o número de variáveis tende a infinito e não é possível determinar com clareza quantas seriam necessárias para compor um modelo com o mesmo *goodness-of-fit*), enquanto outros soam desencorajadores (com a constatação de que, por exemplo, um número de variáveis aleatoriamente escolhidas tão pequeno quanto sete, em pelo menos 5% das vezes, já seria suficiente para replicar medidas de *goodness-of-fit* previamente apresentadas).

O objetivo aqui não é atingir uma conclusão clara acerca do assunto – conforme mostrado acima, argumentações e resultados empíricos demonstram, através dos mais variados caminhos, resultados que sinalizam a favor ou contra a previsibilidade de retornos no geral. Estes devem estar sempre em mente e servirem como pano de fundo para qualquer trabalho que se proponha a num mínimo tangenciar o tema, para que se tenha uma noção clara das possibilidades, do que se pretende atingir, e para que a interpretação de seus resultados seja bem guiada.

2.2 *Abordagens Diversas para a Previsibilidade de Retornos*

Seguindo ainda no último *paper* mencionado, Foster et al. servem como boa indicação da riqueza da literatura: os estudos que se propuseram a revisar em sua análise de *goodness-of-fit* – todos anteriores a 1997 – já mostram uma série de variáveis, especificações de modelos e abordagens diferentes e inventivas em torno da previsibilidade de retornos para casos específicos.

Desde então já era possível identificar algumas variáveis especialmente interessantes: *spreads* entre *treasury bills* de diferentes maturidades (Campbell 1987), a diferença entre o *yield* da *treasury bill* de um mês e o *yield* de bonds com nota de crédito pior que Baa, a razão entre o nível do S&P 500 deflacionado e seu nível no fim de anos anteriores (Keim, Stambaugh 1986), e até uma variável *dummy* para o mês de janeiro (Harvey 1989, Ferson e Harvey 1991). Dentre essas, já é possível identificar ao menos duas das principais áreas de busca por bons previsores de retornos: indicadores do mercado de renda variável (como já seria esperado), além de métricas, ou os próprios retornos, de ativos no mercado de renda fixa.

Os indicadores do mercado de renda variável estão essencialmente ligados ao que já foi explicado acima sob a noção dos “mercados eficientes”, e as principais variáveis documentadas na literatura já foram abordadas. Como extensão, outras como a razão entre lucros e preços ou quaisquer tentativas de normalização utilizando médias temporais ao invés de dados pontuais poderiam ser citadas. Referências importantes na

literatura, além das já mencionadas, se encontram em Rozeff (1984) ou Campbell & Shiller (1988).

A associação com o mercado de renda fixa pode ser explicada a partir de diversas óticas. Na macroeconomia, análises de dívida soberana como as *treasury bills* refletem a produção potencial de um país, os estágios em que se encontram dentro de seus ciclos econômicos ou até questões políticas que podem ter interferência em retornos de mercado. Nos aspectos microeconômicos, levando em conta títulos corporativos, sua precificação, análise de crédito ou panorama consolidado refletem aspectos que dialogam diretamente com o desempenho das empresas – muitas vezes até as mesmas – no mercado acionário. É isso que mostram Fama & Schwert (1977), Keim & Stambaugh (1986), ou Campbell (1987).

Para além das variáveis mais comuns, resultados interessantes foram apresentados com o uso de especificações mais inovadoras. Polk, Thompson & Vuolteenaho (2006), por exemplo, partem da lógica de um modelo CAPM (e, é claro, das hipóteses que ele assume, como investidores que levam em conta apenas média e variância em suas decisões), e tentam utilizar os dados *cross-section* (fundamentalmente os *betas*, como medida de exposição ao risco de mercado) para estimar a série temporal do *equity risk premium*. Baker & Wurgler (2000) utilizam os níveis de emissão de ações, em comparação com a emissão de dívida, e encontram correlação negativa com os retornos, resultado que depõe contra a capacidade corporativa de avaliar o *timing* do mercado de maneira eficiente.

Lettau & Ludvigson (2001), com uma construção ainda mais engenhosa, fazem uso de variáveis agregadas para consumo e riqueza e encontram que as flutuações na razão entre as duas funcionam como boas previsoras das flutuações tanto nos retornos do mercado acionário quanto no *equity risk premium*. Sua abordagem é interessante – e também especialmente complexa – por envolver uma mistura entre questões comportamentais do investidor e questões macroeconômicas, além da série de controles demandada para lidar com variáveis que não necessariamente são observáveis.

O que fica evidenciado, tanto pelos caminhos já muito difundidos quanto pelos mais inovadores, é que não há limite para as possibilidades. Variáveis comuns podem ser remodeladas, repensadas a partir de novas proposições estatísticas, e há margem para que variáveis novas sejam construídas explorando campos ainda não especificados.

2.3 Efeitos, Consequências e Possibilidades de Absorção dos Estudos

Qual é o impacto prático da pesquisa acadêmica acerca da previsibilidade de retornos? É claro que a relevância dos trabalhos até aqui analisados, assim como qualquer prospecto de trabalho futuro, passa obrigatoriamente por algum tipo de resposta positiva a essa pergunta. É preciso que se possa mensurar algum efeito, consequência, ou que se tenha em mente um bom entendimento em relação a como agentes do mercado – para além do mundo acadêmico – atuam de forma a absorver quaisquer inovações.

McLean & Pontiff (2012) se propõe justamente a buscar entender e quantificar o impacto da publicação de trabalhos sobre previsibilidade do retorno de ações em sua própria existência. Sua principal novidade está na comparação não apenas dos dados dentro das amostras de estudos passados com uma extensão de seu horizonte temporal, mas também na inclusão dos efeitos fora da amostra pré e pós-publicação, tentando quantificar de forma direta as possibilidades de absorção das estratégias pelo mercado.

A base teórica é bem fundamentada em torno das intuições já abordadas anteriormente: os vieses estatísticos envolvidos em resultados positivos apresentados no passado deveriam ser expostos por uma performance ruim no período *out-of-sample*, caso fossem o único determinante do sucesso anterior. Além desse possível componente, no entanto, há “empurrões” para a direção oposta: se os resultados positivos da previsibilidade em estudos passados forem fruto da precificação de fatores de risco, eles devem persistir ainda depois de publicados – não haverá mudança no comportamento de precificação dos agentes cujas expectativas são puramente racionais (ou seja, pode se pensar em controlar para os vieses estatísticos da amostra inicial e buscar encontrar resultados que, feito isso, sejam basicamente iguais na amostra inicial e na estendida).

Outra possibilidade, ainda, caso a previsibilidade inicialmente reportada seja fruto de um “*mispicing*”, é que a publicação alerte os agentes do mercado sobre sua existência, levando-os a atuar sobre esse fato até que ele seja mitigado (ou eliminado por completo, a depender do nível das condições que dificultem a atuação sobre oportunidades de arbitragem). Essa acaba por ser a teoria mais consistente com os resultados obtidos pelos autores, fato relevante na condução de estudos semelhantes – por exemplo, em mercados como o brasileiro, talvez se espere encontrar mais situações de “*mispicing*” do que em outros mais desenvolvidos. A relevância se mostra, também, na demonstração de que mercados como o norte-americano apresentam grande integração com o primeiro plano da pesquisa acadêmica que os estuda e afeta diretamente.

Campbell & Thompson (2005) apresentam um panorama prático interessante sob o qual essa integração pode ser materializada. Seu estudo pode ser interpretado como uma tentativa de conciliar toda a literatura de questionamento em torno da previsibilidade, a partir de raciocínios, indicações ou restrições que agreguem valor prático para a questão. O grande mérito está na tentativa de se colocar sob a ótica de um “investidor do mundo real” – quais são as concessões a serem feitas, o que precisa ser adaptado, e o que ainda pode ser aproveitado na literatura de modo a ter utilidade para ele.

Sob a argumentação de que o “investidor real” jamais faria uso de regressões lineares simples para a previsão do *equity risk premium* (no sentido mais direto de simplicidade como ausência de restrições ou controles), ao impor condições como a adequação do sinal de alguns previsores ao que é indicado pela teoria econômica tradicional, os resultados fora da amostra já apresentam melhora relevante. Outro fator

de importância destacável é a tentativa de associação entre as conclusões tiradas *in-sample* e *out-of-sample* – até que ponto estão sujeitas às mesmas interpretações, se são passíveis de apuração sob a mesma ótica, e quando que resultados estatisticamente “pobres” ainda podem ser utilizados, na prática, de forma vantajosa.

Os autores estimam uma medida para o que seria um bom R^2 a se esperar *out-of-sample* – associado à capacidade de uso dos modelos para a otimização da utilidade de determinado investidor – e explicitam a questão do poder estatístico de forma muito mais tangível (especialmente por mostrarem que valores extremamente pequenos na interpretação estatística pura já seriam suficientes para gerar ganhos de utilidade com sua aplicação prática, o que os leva a concluir, de forma bem intuitiva, que pouco poder explanatório nas regressões de previsão é algo que deveria ser esperado). A “tangibilidade” do que é proposto ainda aumenta quando se analisam as restrições usadas nesse trecho do estudo: investidores que levam em conta somente média e variância em sua otimização, não podem vender a descoberto e têm um limite de alavancagem de 50%, por exemplo, se enquadram em um perfil muito mais próximo do investidor “de varejo” do que da sofisticação e do instrumental à disposição de grandes investidores institucionais (evidenciando que os ganhos na aplicação não seriam algo limitado a estes últimos).

Diante do contexto deste trabalho – certamente muito incipiente em comparação com o grau de sofisticação dos estudos analisados nessa revisão da literatura – o esforço de conciliação de Campbell e Thompson funciona como excelente ponto referencial.

2.4 Welch & Goyal (2008) – “Um olhar compreensivo para a performance empírica de previsão do equity premium”

Goyal & Welch (2008) serve como principal inspiração e base metodológica para esse trabalho, não diretamente por suas conclusões, mas sim pelo o que propõe e explicita desde o próprio título: o olhar compreensivo, não apenas no sentido da busca por entendimento, mas também no de inclusão de todos ou quase todos os aspectos do tema em questão.

Não que a conclusão não seja relevante – pelo contrário, talvez possa até ser considerada um ponto de inflexão na literatura de previsibilidade do *equity risk premium*, algo a ser levado em conta por todos os trabalhos posteriores. Suas afirmações são muito contundentes: a falha dos diversos modelos testados em superar a média histórica em capacidade preditiva, e o fato de que um investidor fazendo uso de qualquer um deles teria obtido resultados tão bons (ou até piores) quanto se tivesse assumido que “o *equity premium* é como sempre foi”. Não são propostas sem precedentes, ao menos intrinsecamente, conforme já indicado, mas são inovadoras em abrangência – esse sim o grande ponto de destaque.

Diante de comparações entre o contexto norte-americano e o brasileiro, qualquer que fosse a conclusão atingida por Welch & Goyal, seria difícil extrapolá-la como algo muito além de um mero exemplo. É um direcionamento, representa como algumas das

relações entre as variáveis se comportam e pode pautar a formação de algumas expectativas prévias, mas seria algo muito mais significativo – em termos de resultado objetivo – caso a comparação envolvesse outra região mais desenvolvida, como a União Europeia ou o Japão, para citar apenas dois exemplos. O Brasil é um país de desenvolvimento muito mais embrionário, repleto de idiosincrasias que poderiam permitir narrativas divergentes do que atestam os autores. O próprio estágio ainda muito mais primitivo de seu mercado acionário talvez pudesse ser argumento em prol de processos menos eficientes de alocação e formação de preços, somados a pouco conhecimento e capacidade de explorar esse fato por grande parte dos agentes, dando espaço para sinais mais claros de previsibilidade em alguns momentos (para uma minoria que soubesse explorá-los). Ainda que contextos cuja realidade mais se assemelha à norte-americana fossem abordados, diferentes dados, orientações e histórias econômicas poderiam produzir resultados também diferentes – não existe generalização direta nesse quesito.

É com isso em mente que deve ser destacada a abrangência proposta pelos autores, deixando uma base empírica riquíssima para ser adaptada. Algo do qual definitivamente não podem ser acusados é limitação em suas tentativas de explorar as especificações possíveis. Em relação aos dados, seria impossível contemplá-los de forma mais completa – basicamente um século, ou mais, de informações organizadas em todos os horizontes de tempo relevantes (tanto para composição dos regressores e do *equity risk premium*, de um mês até cinco anos, quanto para as janelas de previsão e teste estatístico, incluindo ou excluindo determinados períodos e testando suas implicações). A metodologia também é calcada em diversas alternativas: modelos univariados, multivariados ou condicionais a critérios que buscam minimizar, de forma contínua, os erros de previsão (a serem abordados com mais profundidade adiante).

As variáveis testadas percorrem tudo que foi enumerado na segunda seção deste capítulo. Dentre as métricas “fundamentalistas” do mercado acionário, estão incluídas a razão entre dividendos e preços (correntes e defasados), dividendos e lucros, lucros e preços (com variações para médias móveis e interanuais), além de valor contábil e valor de mercado (todas as variáveis agregadas para o índice de ações como um todo). Propostas de estudos anteriores são revisitadas pelos autores: o *cross-sectional beta*, conforme Polk, Thompson, & Vuolteenaho (2006); o percentual de emissões de *equity* em relação ao total de emissões na economia, conforme Baker & Wurgler (2000); a expansão líquida em *equity*, medida pela variação na capitalização de mercado descontada pelos retornos (ex-dividendos) no período, em linha com Boudoukh et al. (2007); além da própria variância das ações, agregando o quadrado dos retornos diários como *proxy*.

Dentre os indicadores do mercado de renda fixa, se encontram as taxas de *treasury bills* de curto prazo, de bonds de longo prazo, o *spread* entre as duas, e as taxas de dívida corporativa – títulos com nota de crédito AAA, BAA, e também seu *spread*. Uma terceira categoria de relevância engloba variáveis de definição mais macroeconômica: a construção de Lettau & Ludvigson (2001), a partir de consumo,

riqueza e rendimentos agregados, revisitada; a razão entre investimento em capital fixo (privado e não residencial) e capital agregado da economia, e o índice de preços ao consumidor.

Por fim, nenhuma das análises é feita de forma excessivamente complexa, inventiva ou econometricamente vanguardista – os autores literalmente encerram o *paper* dizendo “esperar que a simplicidade de sua abordagem reforce a credibilidade de suas evidências”. Têm-se, em tudo isso, a fundamentação ideal para se pensar de forma ampla e testar de forma prática a previsibilidade do *equity risk premium* no Brasil.

3. Dados

É válido, inicialmente, pontuar algumas considerações tanto conceituais quanto práticas. Duas complexidades se apresentam na ideia de *equity risk premium*: primeiro, o fato de sua abordagem envolver os conceitos de risco e retorno em relação a todo o mercado acionário. A *proxy* mais confiável que se tem para isso está na utilização de índices representativos, dado que, ainda que não englobem todas as empresas listadas em determinado país (algo cuja construção beira a impossibilidade), não deixam de ser o principal ponto focal do mercado como um todo. A principal ênfase foi dada ao índice MSCI Brasil pela consistência e maior abrangência de seus dados; o índice Ibovespa, ainda que seja o referencial mais popular da bolsa brasileira, possui informações mais dispersas e inconsistentes ao longo do tempo.

O segundo aspecto, menos explícito, mas ainda assim relevante, está no termo “livre de risco” da equação. Não seria absurdo considerar dois pontos de vista distintos que, ainda que sejam mais extremos, não perderiam a sensatez: por um lado, a defesa de que não há nenhuma taxa que seja, para esse caso, de fato 100% livre de risco (o que basicamente inviabilizaria o conceito de *equity risk premium* como um todo); por outro, a de uma maior flexibilização sob a perspectiva de um risco mais residual e menos material (o equivalente a “elevar o sarrafo” para o retorno demandado do mercado acionário). É fato, por exemplo, que atualmente ninguém vê um calote do governo brasileiro em sua dívida como prospecto iminente (apesar de isso não ter sido sempre tão claro ao longo da história). Não seria absurdo se valer dessa hipótese e usar a taxa de algum título público como referencial. O lugar do Brasil na escala “livre de risco”, no entanto, parece distante demais do caso norte-americano para seguir esse caminho. Seguindo com certa prudência e conservadorismo, opta-se aqui por utilizar a taxa CDI mensal.

As bases de dados utilizadas são constituídas a partir das seguintes fontes:

- Sistema Bloomberg, especialmente no que se refere às variáveis diretamente financeiras.
- Instituto Brasileiro de Geografia e Estatística (IBGE) para as séries históricas dos índices de inflação.
- Pacote de dados e manipulação “BETS” (Brazilian Economic Time Series), disponível no software R, organizado por Pedro Costa Ferreira, Thalita F. Speranza e Jonatha A. da Costa com o objetivo de unificar as fontes disponíveis para as mais diversas séries temporais econômicas brasileiras em um repositório único.

Construção e Detalhamento das Variáveis:

A variável dependente dos modelos é sempre o *equity risk premium*: o retorno mensal (total, isto é, incluindo tanto apreciação de preço quanto proventos) de um

índice (MSCI Brasil, na especificação principal, ou Ibovespa) descontado da taxa CDI acumulada no mês em questão. Os retornos são compostos de forma contínua e obtidos através da Bloomberg; a taxa CDI mensal é obtida através das bases de dados do pacote BETS. As variáveis independentes dos diversos modelos são construídas conforme abaixo:

Dividendos: soma móvel de quaisquer distribuições de proventos que tenham tido sua data *ex* nos 12 meses anteriores. A partir daí, calcula-se a razão Dividendos / Preços como a diferença do logaritmo dos dividendos e do logaritmo dos preços; e o *Dividend Yield*, termo utilizada para designar a diferença entre o logaritmo dos dividendos e o logaritmo defasado dos preços.

Lucros: soma móvel dos lucros por ação dos membros que compõe o índice ao longo dos 12 meses anteriores. A partir daí, calcula-se a razão Lucros / Preços como a diferença entre o logaritmo dos lucros e o logaritmo dos preços; e a razão de *payout* de dividendos como a diferença entre o logaritmo dos dividendos e o logaritmo dos lucros.

Volatilidade: a volatilidade acumulada nos últimos trinta dias equivale ao desvio padrão anualizado das oscilações de preço diárias do índice acionário ao longo desse período. Como variações que englobem períodos mais longos, também são usados acúmulos da volatilidade em janelas mais extensas.

Valor contábil: é calculado, de forma agregada para o índice, como a soma do valor contábil por ação de cada empresa constituinte, ponderada por seu peso na composição total do mesmo; para se obter o equivalente à razão entre valor contábil e valor de mercado, utiliza-se o valor contábil por ação (ou seja, seu agregado dividido pela capitalização total do índice) e seu último preço para o período em questão.

Expansão do mercado acionário: como forma de medir a atividade do mercado na emissão de novas ações (seja por ofertas públicas iniciais, expansões de capital, distribuições ou recompras), o valor da expansão líquida (em reais) é calculado conforme proposto em Welch & Goyal (2008):

$$\text{Expansão Líquida (t)} = \text{MarketCap(t)} - \text{MarketCap(t-1)} \times [1 + \text{vwpr(t)}]$$

onde MarketCap representa a capitalização total de mercado do índice em dado período, e vwpr (de *value weighted price return*) representa o retorno (considerando apenas oscilações de preço), ponderado pela capitalização. A variável usada a partir disso é a razão entre a expansão líquida e o valor de mercado no fim do período em questão. Vale notar, conforme destacam os autores, que essa métrica implicitamente assume que o retorno consequente de um fechamento de capital (“delistagem”) seria de -100%. A métrica é, obviamente, uma *proxy* e, para o caso norte-americano, não houve diferença nos resultados utilizando todos os retornos reais em casos desse tipo ou mesmo os ignorando, por isso aqui se segue com a abordagem mais simples.

Inflação: é utilizada a série histórica do Índice de Preços ao Consumidor – Amplo (IPCA). Novamente seguindo Welch & Goyal (2008), como os dados de inflação só são

divulgados com um mês de defasagem, espera-se um mês para que sejam usados nas regressões.

Taxas de juros e mercado de crédito: as séries históricas brasileiras, infelizmente, parecem estar o mais distante da disponibilidade norte-americana quando se trata desse tema. Não há um paralelo aplicável para, por exemplo, a longevidade na qual se obtém as taxas das *treasury bills* de curto prazo, tampouco para os *treasury bonds* de prazo mais longo. Como alternativa, aqui se propõe o uso das seguintes medidas: a variação no tamanho das posições de LTNs em custódia, como indicador do maior ou menor peso alocado em títulos públicos e, por consequência, maior ou menor indisposição dos investidores com o mercado acionário e retorno exigido do mesmo; taxa média mensal para novas operações de crédito *non-earmarked* (sem destinação prévia específica para os recursos), tanto para empresas não-financeiras quanto para domicílios; e os respectivos spreads entre essas duas taxas e o CDI.

4. Metodologia

O ponto central da análise aqui exposta é a robustez dos diversos modelos de previsão do *equity risk premium*, tanto *in-sample* quanto *out-of-sample*. Para isso, se tentam replicar alguns dos métodos propostos por Welch & Goyal (2008) aplicando-os ao contexto brasileiro. Como destaque, o questionamento: algum dos modelos testados representaria um ganho frente à tentativa de previsão usando simplesmente a média histórica?

A especificação dos modelos univariados segue o padrão clássico da literatura, usando como variável independente a primeira defasagem do preditor em questão:

$$\text{Equity Risk Premium (t)} = \beta_0 + \beta_1 * \text{Preditor (t-1)} + U(t)$$

São avaliadas dessa forma todas as variáveis cuja construção foi exposta acima. Além disso, é testada também uma especificação multivariada do tipo “*kitchen sink*” – conforme indicado pelo próprio nome remetente à expressão popular americana, nesse modelo são incluídas todas as variáveis disponíveis, sem maior preocupação formal.

Por fim, especifica-se uma abordagem de seleção de modelos na qual, em cada período t , se testam todas as 2^K possíveis combinações para as K variáveis disponíveis, posteriormente avaliadas a partir da maximização do R^2 ajustado (o melhor modelo é selecionado em cada etapa do processo). Dado que essa especificação é revisada em cada período, seu foco é voltado para a análise *out-of-sample*, deixando de lado a avaliação *in-sample*.

Dois outros pontos se fazem de importante destaque. Primeiro, quanto ao uso de variáveis mensais como a escolha principal. É claro que a principal motivação é maximizar o número de observações e informação disponível, dado que o mercado brasileiro ainda é muito incipiente para oferecer uma análise rica usando janelas mais longas. Corre-se o risco, com isso, de estar sujeito a dados mais ruidosos ou a problemas como a heterocedasticidade dos resíduos, mas para o caso do Brasil parece ser um *trade-off* para o qual é difícil tomar outra escolha.

A outra questão tange os períodos de tempo utilizados ao longo da estimação – decisão importante para a análise dos modelos *out-of-sample* e para a qual, no fim das contas, não existe fórmula exata e precisa, tornando-a sujeita à arbitrariedade. Novamente é um *trade-off* de difícil negociação no contexto dos dados mensais (e nem assim extremamente numerosos). Tenta-se, aqui, não fugir do que são consideradas as melhores práticas da literatura – tanto ao garantir que haja observações suficientes no início da amostra para que possa haver significância estatística, quanto ao assegurar um período de avaliação também grande o suficiente – com a escolha de vinte períodos subsequentes ao início da amostra para o ponto de partida das análises *out-of-sample*.

Avaliação Estatística:

Os coeficientes para cada um dos modelos são estimados através do método de Mínimos Quadrados Ordinários, por regressões lineares tradicionais (salvo exceções mencionadas mais adiante), mas avaliados a partir de estatísticas F e seus níveis críticos empiricamente construídos. Dessa forma, a capacidade de previsão dos casos univariados é posta em comparação com uma especificação que siga apenas a média histórica do *equity risk premium*, em linha com o principal foco de análise, e a dos casos multivariados apresenta um controle para a correlação entre seus termos.

Seja N o subscrito para o vetor de erros de previsão a partir da média histórica (até dado período no qual a previsão é feita) e A o subscrito para o vetor de erros de previsão do modelo condicional (especificado, também, com os dados até o período em questão) – as estatísticas *out-of-sample* são conforme abaixo:

$$R^2 = 1 - (EQM_N / EQM_A)$$

$$\bar{R}^2 (R^2 \text{ Ajustado}) = 1 - \frac{(1 - R^2) * (N - 1)}{T - k - 1},$$

$$\Delta REQM = \sqrt{EQM_N} - \sqrt{EQM_A}$$

$$MSE-F = (T - h + 1) * \left(\frac{EQM_N - EQM_A}{EQM_N} \right)$$

EQM é o erro quadrático médio; T é o número de observações usadas para previsão; k é o número de parâmetros estimados, e h é o grau de sobreposição (equivalente a 1 quando não há sobreposição, ou seja, quando a janela de previsão corresponde a frequência dos dados utilizados na estimação dos modelos).

Tanto a estatística F *in-sample* quanto a MSE-F *out-of-sample* não são diretamente reportadas – como alternativa, valores críticos são construídos empiricamente e tomados como parâmetro na significância estatística, conforme descrito abaixo.

Bootstrap:

A construção empírica dos valores críticos para as estatísticas analisadas segue duas variações da metodologia de *bootstrap*. Ambas tomam como hipótese nula a de não-previsibilidade, ou seja, as seguintes especificações dos processos geradores de dados:

$$Y_{t+1} = \alpha + u_{1,t+1}$$

$$X_{t+1} = \mu + \rho * X_t + u_{2,t+1}$$

Para a análise de estabilidade dos modelos, se supõe que os processos sigam a hipótese alternativa dos testes:

$$Y_{t+1} = \alpha + \beta * X_t + u_{1t+1}$$

$$X_{t+1} = \mu + \rho * X_t + u_{2t+1}$$

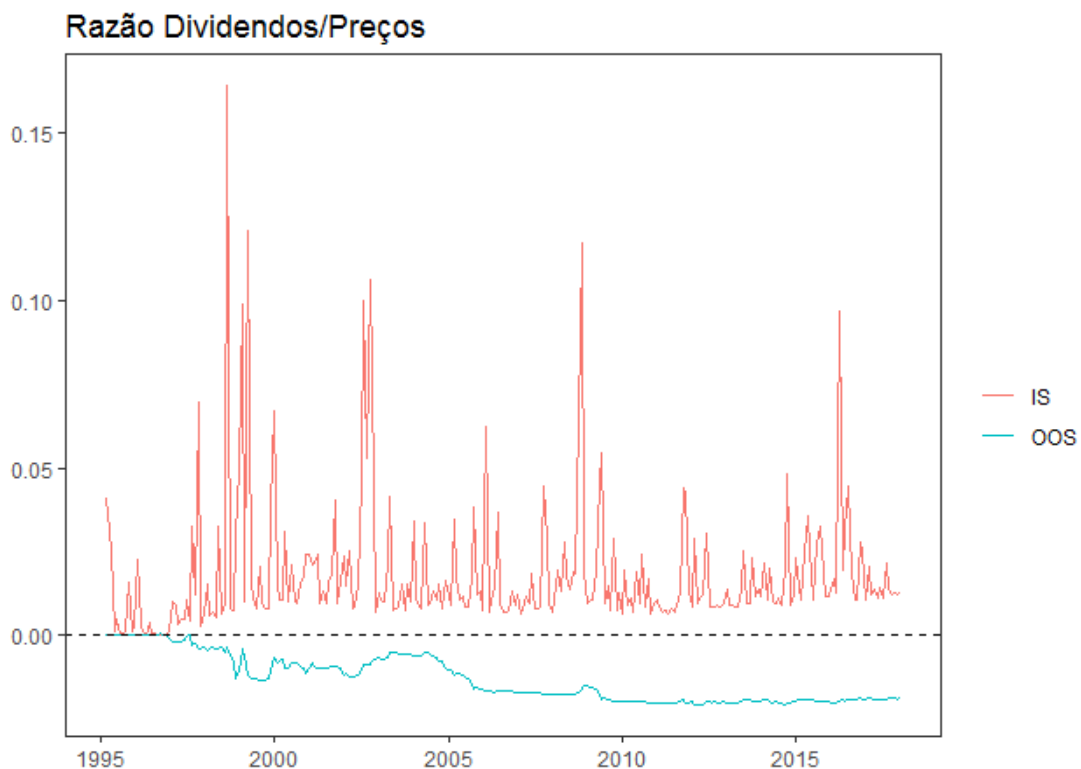
β e ρ são estimados através de Mínimos Quadrados Ordinários usando os dados completos da amostra; os resíduos de suas regressões são armazenados, e é sobre eles que incide o processo de reamostragem. São compostas 10.000 novas séries temporais dos resíduos selecionados (com reposição) a partir da amostra original.

O cenário base é um *bootstrap* em blocos, o que permite controlar para a existência de alguma relação de dependência intertemporal entre os resíduos. Diversos tamanhos foram testados para os blocos e os resultados se mostraram pouco sensíveis a isso. Outra alternativa é a indexação, tal que cada observação das novas amostras esteja alinhada pelo mesmo período. Neste caso, abre-se mão do controle por uma possível dependência intertemporal, de modo que os resíduos são tomados como independentes ao longo do tempo.

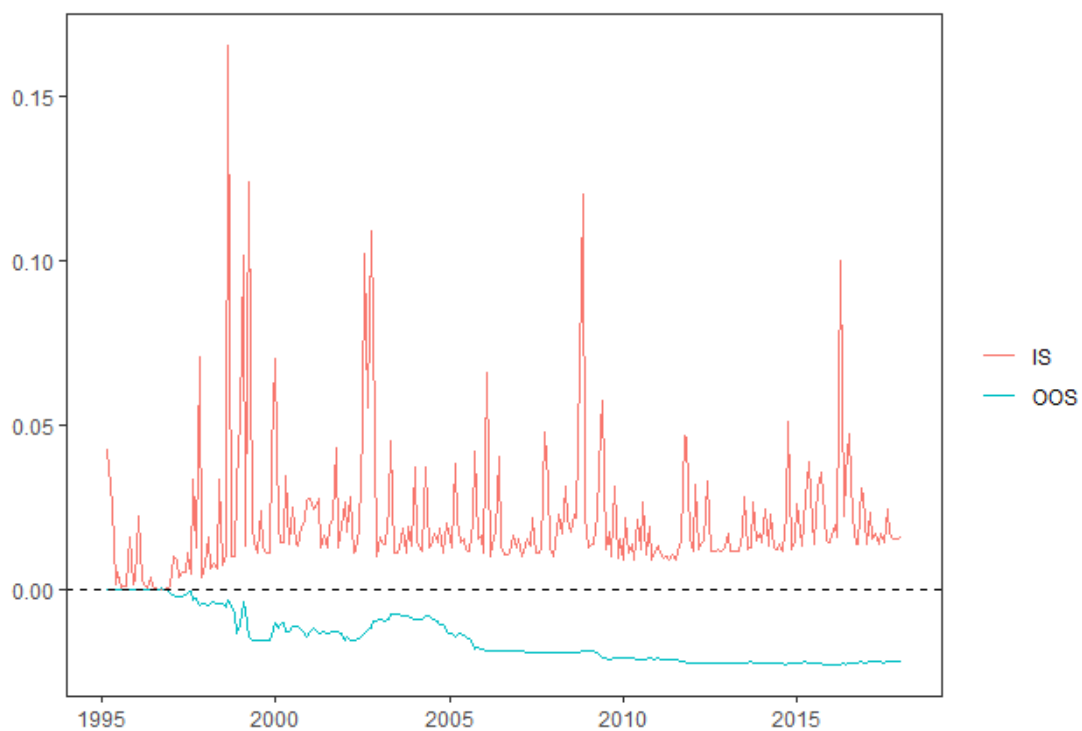
As duas variações na aplicação das reamostragens não apresentam diferenças significantes entre seus resultados. O que as difere é, principalmente, o ganho de eficiência computacional, dado que o primeiro cenário é mais diretamente executável.

5. Resultados – Cenário Base

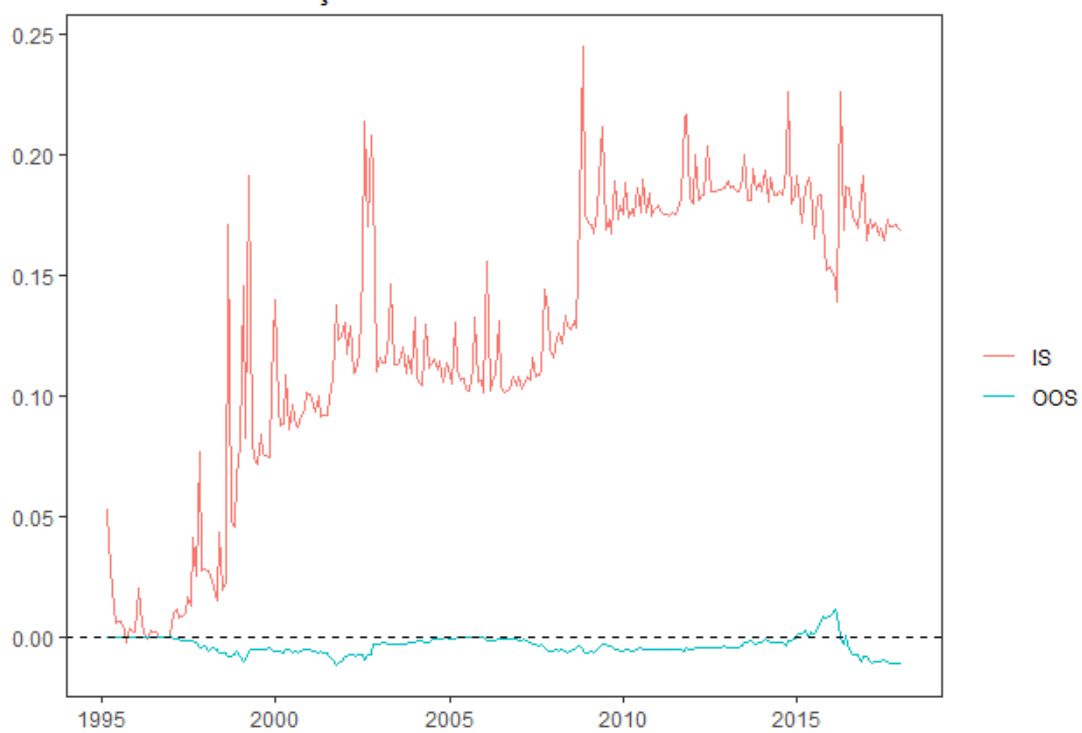
Os gráficos a seguir ilustram a capacidade preditiva dos modelos univariados, dentro e fora da amostra, em comparação com a especificação básica da média histórica. O cenário base de previsão, aqui, segue um intervalo de tempo mensal tanto nas observações das variáveis preditivas quanto no horizonte de previsão. Para a análise *in-sample* (IS), a métrica equivale ao acúmulo do quadrado do desvio do *equity risk premium* em relação à sua média (computada ao longo de todo o período amostral) menos o acúmulo do quadrado dos resíduos das regressões. Para a análise *out-of-sample* (OOS), equivale ao acúmulo do quadrado do erro de previsão da média prevalecte até dado período menos o acúmulo do quadrado do erro de previsão do modelo especificado pela regressão linear. Trabalha-se, aqui, com a mesma medida central na obra de Welch & Goyal (2008), partindo da premissa de uma hipótese nula de melhor capacidade preditiva para a média histórica, e hipótese alternativa de superioridade dos modelos especificados em cada caso. Quando a curva percorre território positivo a hipótese alternativa se mostra melhor em suas previsões, e vice-versa. Os gráficos *in-sample* foram verticalmente ajustados de modo a começarem do zero no mesmo período em que começam as previsões *out-of-sample*, facilitando a comparação.



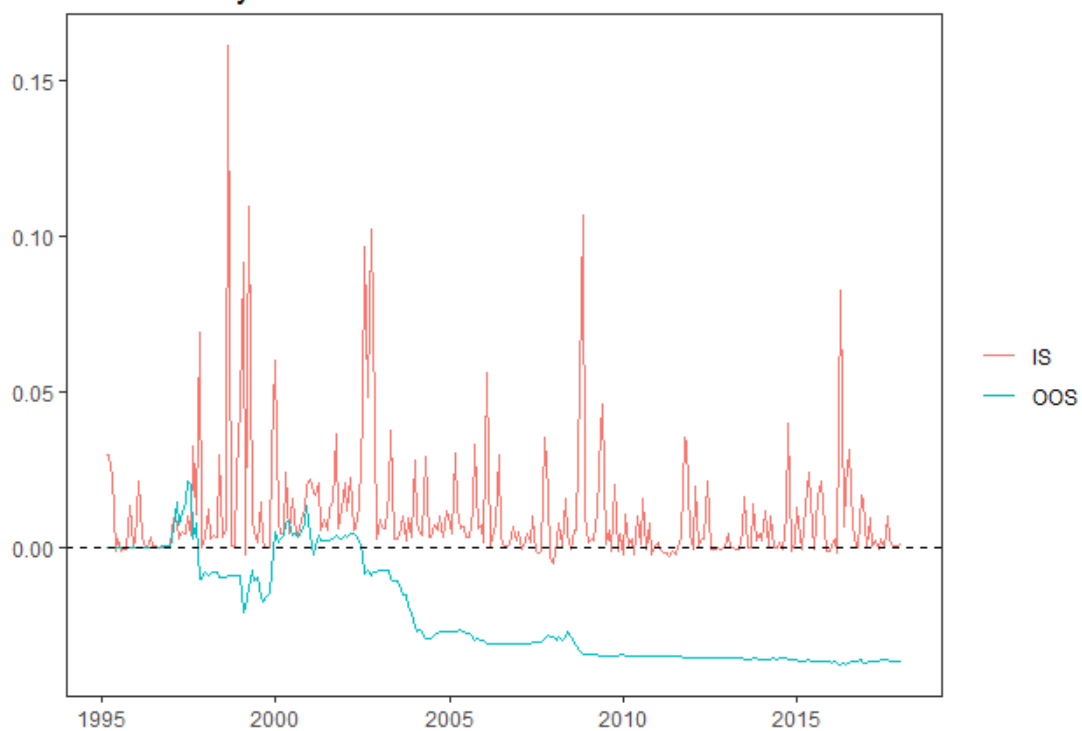
Dividend Yield



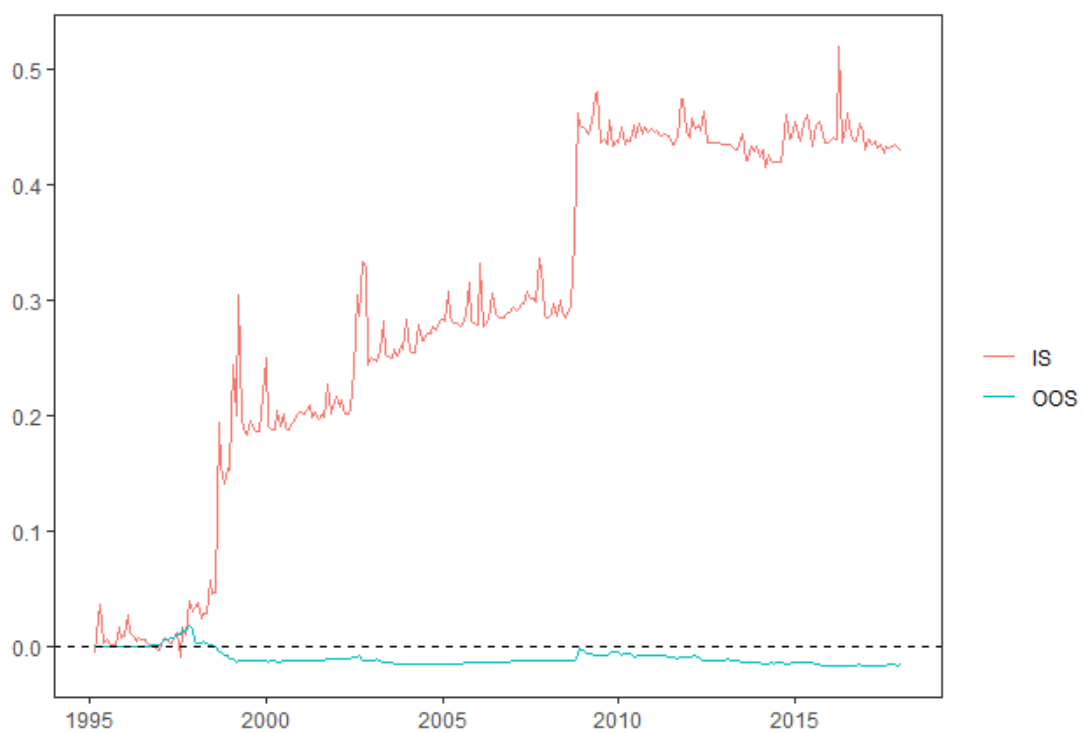
Razão Lucros/Preços



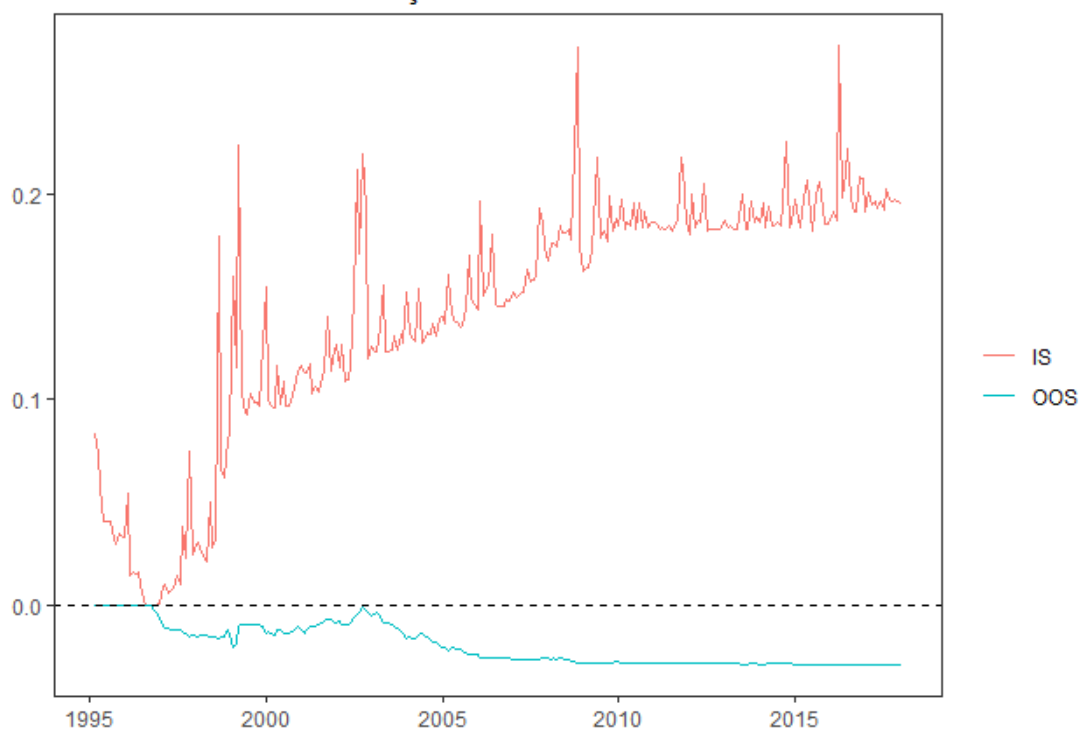
Dividend Payout Ratio



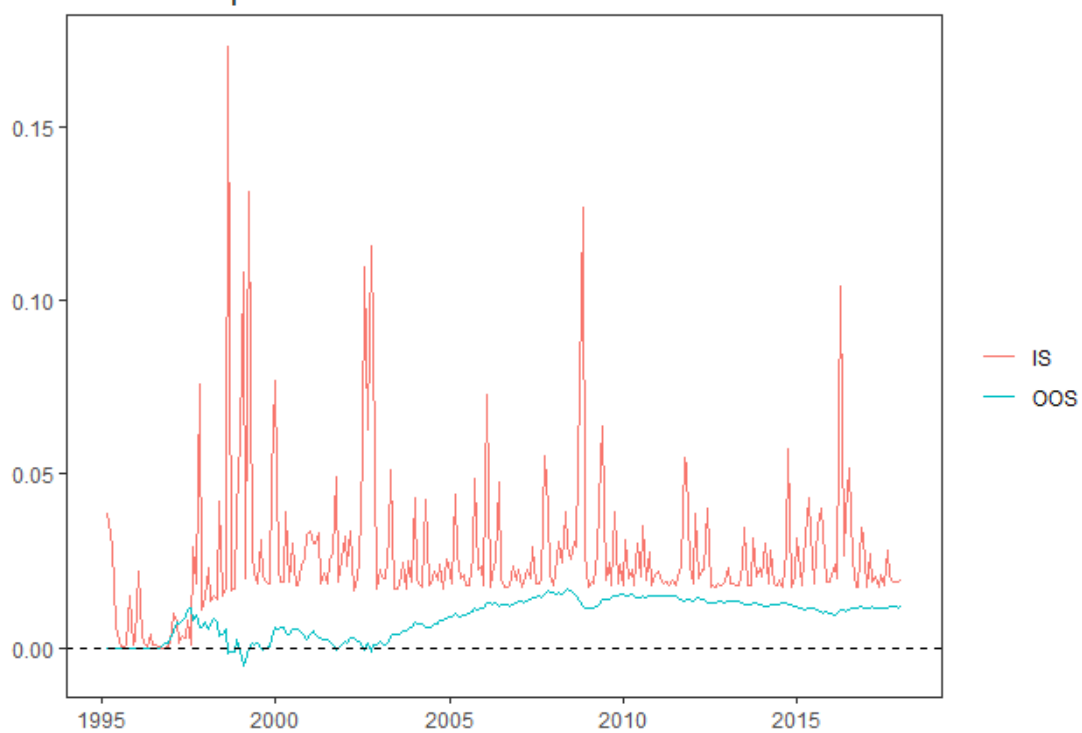
Volatilidade Acumulada Nos Últimos 30 Dias

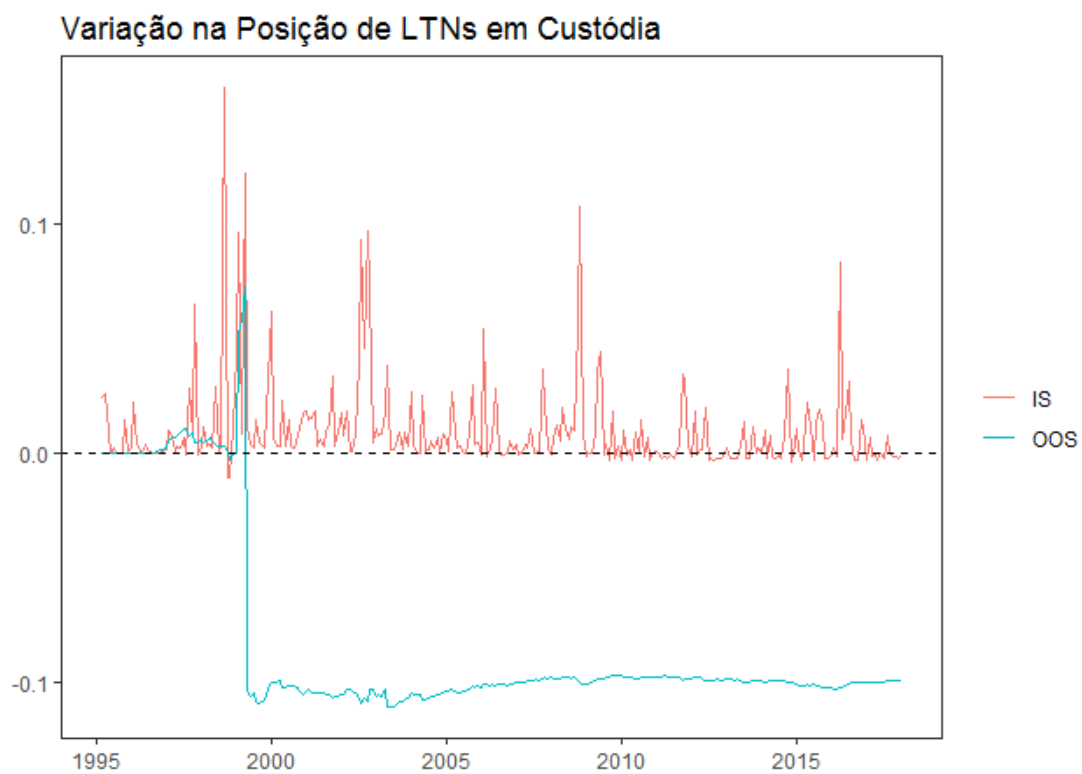
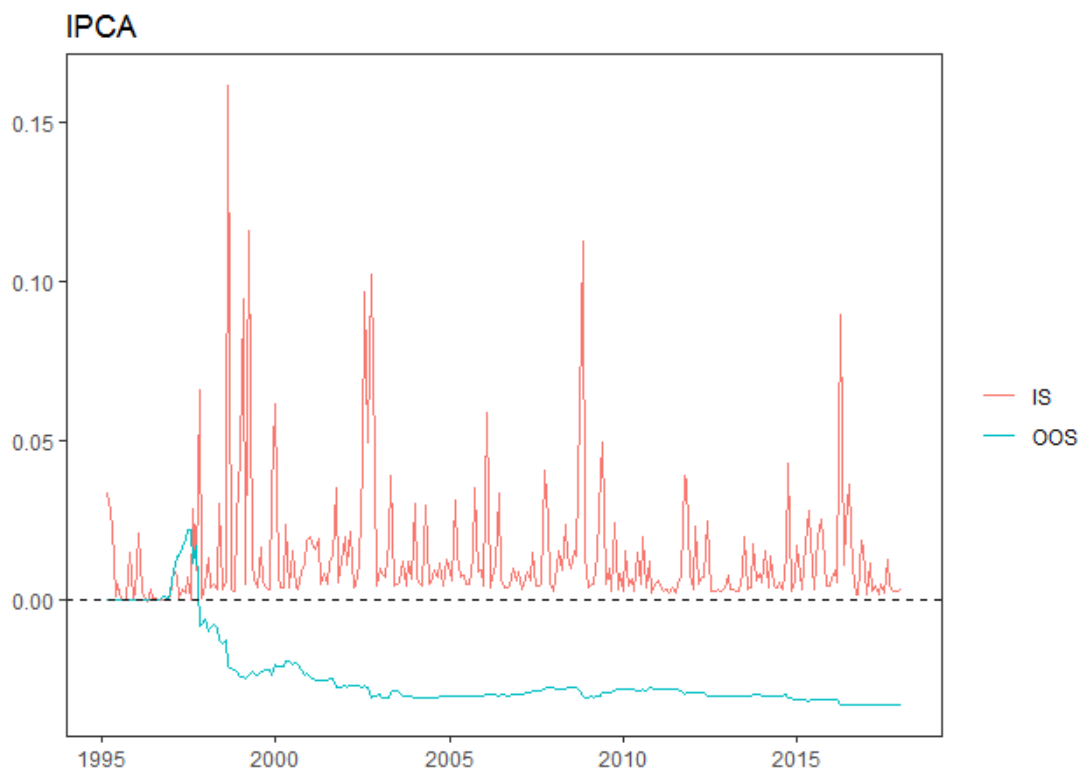


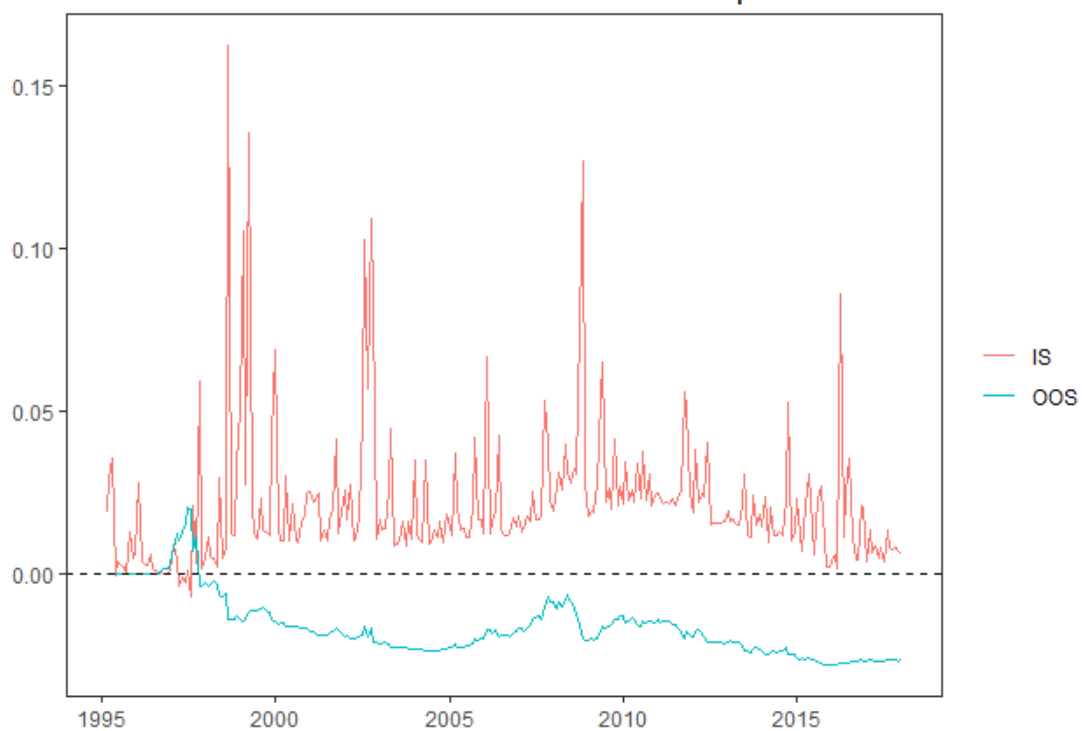
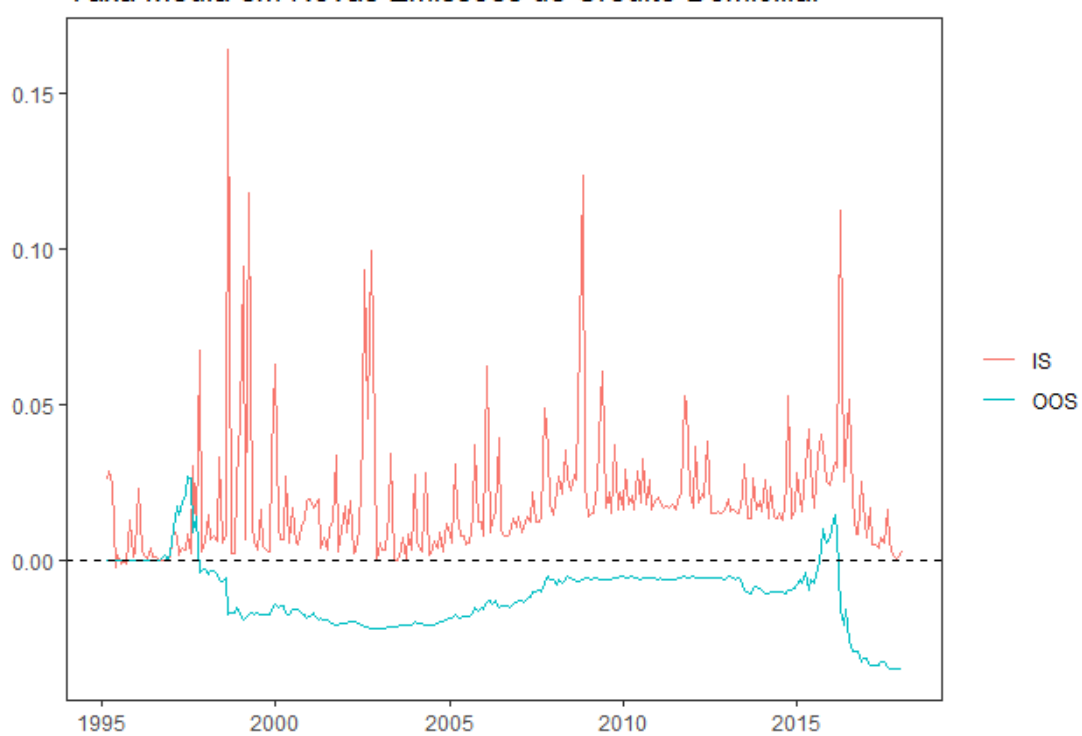
Razão Valor Contábil/Preços



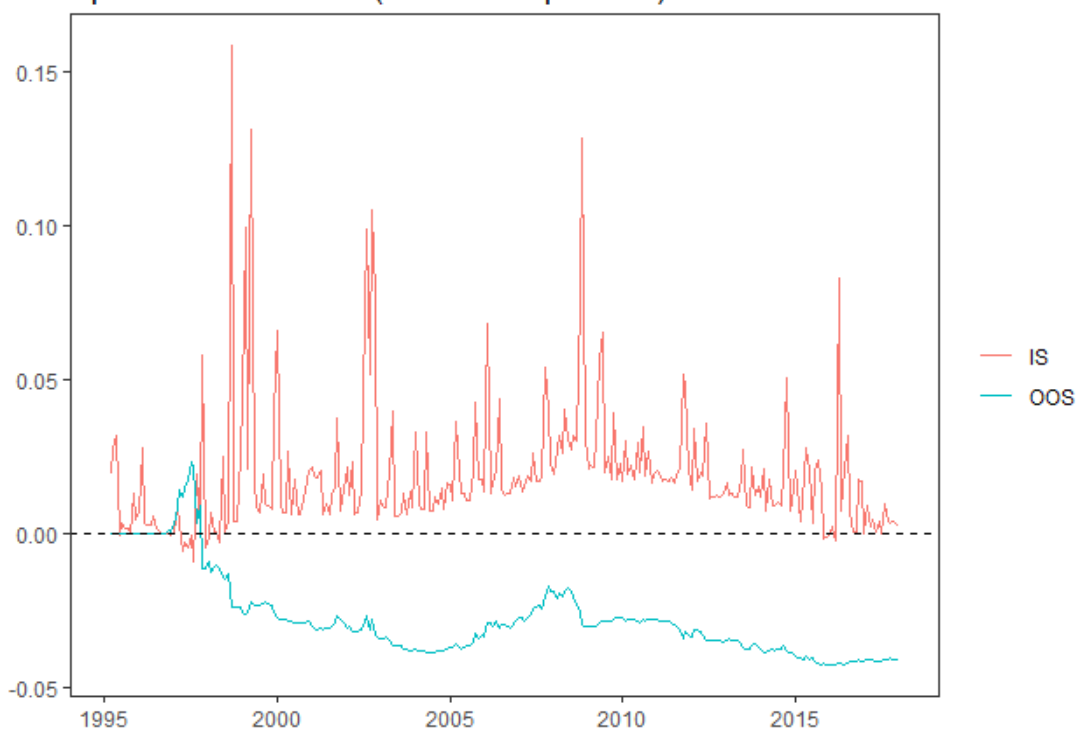
Emissão Líquida



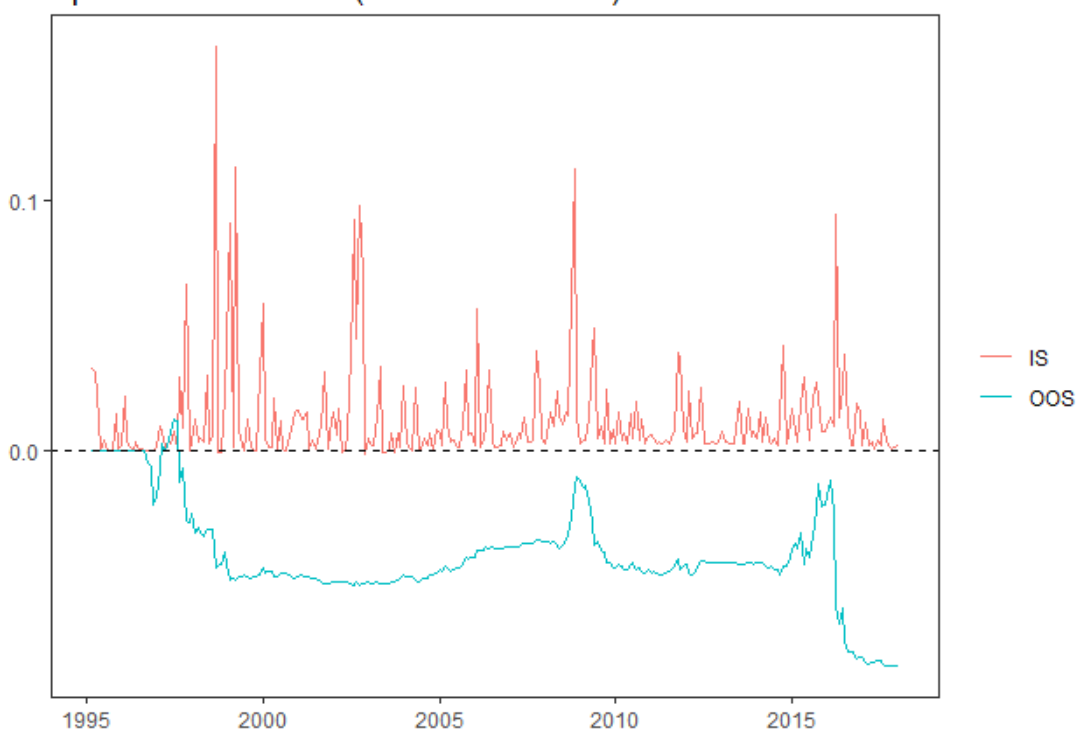


Taxa Média em Novas Emissões de Crédito Corporativo**Taxa Média em Novas Emissões de Crédito Domiciliar**

Spread - Taxa Média (Crédito Corporativo) e CDI



Spread - Taxa Média (Crédito Domiciliar) e CDI



De imediato destacam-se as curvas *in-sample* majoritariamente acima de zero (e acima de suas correspondentes *out-of-sample*), além de, para a grande maioria das variáveis, um comportamento bastante errático e volátil. É algo que deveria ser esperado de antemão, dado o comportamento do próprio *equity risk premium* por si só, também bastante errático e volátil em suas observações mensais. A presença quase que permanente em território positivo não é marcada, na maior parte dos casos, por uma tendência clara de crescimento, mas sim por oscilações um tanto quanto bruscas.

Welch & Goyal fornecem uma espécie de regra de bolso para um modelo passível de inspirar confiança: uma performance significativa *in-sample* e razoável *out-of-sample* ao longo de todo o período da amostra; uma tendência, ainda que irregular, de crescimento; a consistência dessa tendência ao longo de momentos distintos, não representando apenas um crescimento esporádico sujeito a estar condicionado a determinado contexto histórico; e a consistência dessa tendência nos períodos mais recentes, mantendo um viés positivo e evitando a possibilidade do modelo em questão já não mais se aplicar bem ao cenário atual. Nenhum dos gráficos acima aparenta se enquadrar de forma muito completa nessa descrição – as características positivas, quando existentes, aparecem de forma mais isolada e com pouca robustez geral.

A tabela a seguir fundamenta estatisticamente os resultados. A variação na raiz dos erros quadráticos médios (ΔREQM), em especial, fornece uma métrica que segue a mesma unidade de medida das curvas *out-of-sample* e pode ser avaliada com base na mesma intuição – números positivos representam melhor desempenho do modelo condicional à variável preditiva em questão frente ao incondicional (modelo de previsão “ingênuo” que usa somente a média histórica até dado período).

A análise *in-sample*, vale ressaltar, segue as métricas clássicas da literatura (R^2 e R^2 ajustado, nesse caso, não são conforme previamente definidos aqui, mas sim suas especificações tradicionais). A coluna *IS-for-OOS* avalia o R^2 ajustado *in-sample* ao longo do mesmo período em que são realizadas as previsões *out-of-sample* (no caso-base de análise até então, descontando os 20 períodos iniciais da amostra). Não são estimados novos modelos utilizando somente a nova janela – a especificação *in-sample* permanece a mesma, utilizando todos os dados disponíveis; a diferença, portanto, está apenas no uso dos resíduos das regressões para o cálculo do R^2 ($1 - (\text{SQR}/\text{SQT})$), usando somente os resíduos do período de previsão *out-of-sample*).

Variáveis ¹	R ² Ajustado – <i>In-sample</i>	R ² Ajustado – <i>IS-for-OOS</i>	R ² Ajustado – <i>Out-of-sample</i>	ΔREQM
Razão Dividendos/Preços	-0.3248	-0.3594	-1.0277	-0.0341
<i>Dividend Yield</i>	-0.3029	-0.2539	-1.1258	-0.0394
Razão Lucros/Preços	4.1019	4.9682	-0.7719	-0.0203
<i>Dividend payout ratio</i>	-0.2729	-0.7544	-1.6070	-0.0655
Volatilidade	15.1491*	13.8904	-0.9278	-0.0287
Razão Valor Contábil/Preço	4.0604*	5.8875	-1.3416	-0.051
Expansão Líquida do Índice	-0.0018	-0.1288	-0.0029	0.0212
IPCA	-0.3544	-0.6729	-1.4910	-0.059
Variação na Posição de LTNs	-0.1342	-0.8479	-3.7334	-0.1788
Taxa Média em Novas Emissões - Crédito Corporativo	0.4626	-0.5791	-1.2895	-0.0482
Taxa Média em Novas Emissões - Crédito Domiciliar	-0.0371	-0.7092	-1.5722	-0.0634
Spread – Crédito Corporativo e CDI	0.3074	-0.6974	-1.7546	-0.0732
Spread – Crédito Domiciliar e CDI	-0.3408	-0.7221	-3.2751	-0.1544
<i>Kitchen Sink</i>	-2.1886	6.2133	-41.2118	-1.7873
Modelo MS	-	-	-31.6673	-0.8653

¹ Níveis de significância de 90%, 95% e 99% são indicados na tabela através de um, dois ou três asteriscos, respectivamente.

5.1 Modelos Estatisticamente Insignificantes In-Sample:

A grande maioria dos modelos testados não apresenta significância estatística, tanto com base em seus valores críticos empíricos quanto em uma análise focada unicamente na amostra original. É nítido que o número de graus de liberdade tem um peso relevante sobre os resultados expostos: ainda que o número de observações seja maximizado ao focar em uma análise mensal, o poder explanatório testado é, na maior parte dos casos, muito pequeno, gerando diversas ocorrências de R^2 com valores muito baixos e, por consequência, R^2 ajustados negativos.

Welch & Goyal defendem que um modelo com mau desempenho *in-sample* não deveria sequer ser interessante para uma análise *out-of-sample*, mas que, ainda assim, muitos deles são proeminentes ou renomados na literatura. Dessa forma, torna-se válido expor seus resultados, ainda que as estatísticas não aparentem transmitir muita confiança em um primeiro momento.

Razão Dividendos/Preços e *Dividend Yield*:

As principais métricas de dividendos – a razão dividendos/preços e o *dividend yield* – servem como bons exemplos, também por estarem entre as variáveis “canônicas” da literatura: suas trajetórias são muito correlacionadas e são boas representações do comportamento das séries no geral. Destaca-se, primeiro, a volatilidade das curvas *in-sample* nos gráficos. Ainda que se mantenham sempre acima de zero, indicando superioridade ao modelo de média histórica, suas trajetórias são extremamente ruidosas – algo que torna complicada a tarefa de analisar períodos de desempenho superior. A trajetória *out-of-sample*, por outro lado, se mostra estável e constante ao longo de todo o período, mas não da forma que se gostaria – não são atingidos valores positivos.

Há um esboço de melhoria *out-of-sample* entre 2001 e 2003, período essencialmente marcado pela eleição de Lula. O evento e seu impacto nos mercados financeiros de forma geral (sendo o dólar, talvez, o principal exemplo), com o temor de um eventual calote da dívida e de uma gestão de esquerda mais radical, se configura como um importante choque externo nas séries analisadas, de modo que qualquer resposta positiva nos resultados das previsões ao longo desse período seria algo interessante de se observar com mais atenção. Por um lado, é desejável que seu modelo responda de forma robusta à choques exógenos, sem “descarrilar” de bons resultados e perder capacidade preditiva; por outro, torna-se questionável a própria capacidade preditiva a partir do momento que as previsões positivas estão mais relacionadas a fatores externos do que à especificação original. Aqui, no entanto, a discussão é relativamente limitada ao plano teórico – a melhoria nos gráficos *out-of-sample* não passa do esboço mencionado, não sendo muito relevante nem consistente, tal que a trajetória volta a apresentar queda pouco em seguida.

Os resultados no período mais recente apresentam uma leve melhora em relação à amostra completa, mas ainda sem significância suficiente.

Razão Dividendos/Preços	Últimos 5 anos	Amostra Completa
R ² Ajustado IS	0.8281	-0.3248
R ² Ajustado OOS	-0.9984	-1.0277

<i>Dividend Yield</i>	Últimos 5 anos	Amostra Completa
R ² Ajustado IS	0.1898	-0.3029
R ² Ajustado OOS	-1.0966	-1.1258

Dividend payout ratio:

O *dividend payout ratio* se mostra bastante alinhado com as duas variáveis já abordadas, mas com uma trajetória mais volátil, tanto *in-sample* quanto *out-of-sample*. A resposta *out-of-sample* aos choques externos é novamente algo interessante de se observar: além de 2002, esse modelo indica ganhos de desempenho também no período de meados de 2008 até o final de 2010, marcado pela eclosão (e seguidas consequências) da crise financeira mundial. É sabido que o impacto na “economia real” brasileira ao longo do período foi limitado em comparação com o que ocorreu em outros países; a própria queda no índice acionário (acentuada entre maio e novembro de 2008) inverte-se em uma recuperação não desprezível até o final de 2010. Ainda assim, o evento de forma ampla se configura como o principal choque econômico recente, cujos efeitos de transmissão mundial não podem ser relevados.

A maior volatilidade preditiva do *dividend payout ratio* faz com que sua resposta a ambos os eventos seja mais visível do que a apresentada pelas outras duas métricas de dividendos, mas também é responsável pela menor durabilidade desse efeito. Ainda que a trajetória gráfica flerte com a marca de melhor previsão do que o modelo de média histórica *out-of-sample*, isso não chega a ser atingido e ocorre de forma esporádica (além de, novamente, não ser possível deixar de lado a discussão acerca da própria qualidade preditiva quando muito relacionada a eventos externos ao modelo).

Seus resultados no período mais recente são desapontadores e não inspiram confiança em qualquer possibilidade de boa aplicação prática.

<i>Dividend Payout Ratio</i>	Últimos 5 anos	Amostra Completa
R ² Ajustado IS	-1.5981	-0.2729
R ² Ajustado OOS	-1.5829	-1.6070

Taxa Média em Novas Emissões de Crédito Domiciliar:

A taxa média em novas emissões de crédito domiciliar é marginalmente insignificante ao nível estatístico de 90%. Com alguma dose de boa vontade, não seria injusto dar certa validade ao modelo. Alguns de seus trechos apontam para uma direção favorável quando se avalia a aplicação do que ele propõe ao contexto contemporâneo, mas seu desempenho a partir de 2016 coloca em risco essa perspectiva.

Seu comportamento inicia o período amostral de forma errática, mas parece ir se desenvolvendo com a passagem do tempo – isso é especialmente acentuado a partir de meados de 2007, como se o modelo “aprendesse” com o choque exógeno e não estivesse sujeito a apenas um ganho esporádico em seu desempenho. É a janela que vai daí até o final de 2014 (quando, inclusive, são apresentados resultados positivos *out-of-sample*) a principal responsável pela quase significância estatística – fica nítido o descolamento da capacidade preditiva da média histórica. O fim da amostra, no entanto, não é animador - tanto *in-sample* quanto *out-of-sample* a queda é exponencial em um período curto e dá a entender que a “aprendizagem” do modelo em relação ao choque externo (nesse caso, não somente externo à sua especificação, mas também no sentido geopolítico) não se repete para a conjuntura de crise interna do segundo governo Dilma (pelo contrário, a queda é brusca quase como se o modelo não se adaptasse mais ao cenário vigente).

Para esse caso, ao contrário do ditado popular, é difícil escapar da noção de que a última impressão é a que fica.

Taxa Média – Crédito Domiciliar	Últimos 5 anos	Amostra Completa
R ² Ajustado IS	-0.3571	-0.0371
R ² Ajustado OOS	-1.5428	-1.5722

É razoável, em qualquer análise, que haja certa perda de credibilidade – num mínimo um olhar cético – em relação ao modelo que não vai bem ao ser aplicado nos

dados mais recentes. Ainda assim, não é motivo para descarte imediato e só o teste sobre novos dados seria capaz de indicar a capacidade de recuperação do bom desempenho entre 2007 e 2014.

Expansão Líquida do Índice:

O modelo da expansão líquida de ações é o principal a se enquadrar na proeminência já mencionada – seus resultados *in-sample* perdem, mesmo que não por muito, o nível de 90% de significância estatística, mas seus resultados *out-of-sample* são os mais consistentes dentre todos os analisados.

Com um MSE-F de 0.5919, esse é o único caso que chega perto de ultrapassar seu valor crítico empiricamente construído (0.6822) – ainda que não ultrapasse, com exceção de uma curta janela em 1998, conforme ilustrado no gráfico, seu desempenho preditivo supera a média histórica em toda a amostra. É o que indica seu $\Delta REQM$ (0.0212) – o único com resultado consistentemente positivo dentre todas as variáveis testadas. Percebe-se, no entanto, que a magnitude reduzida ainda causa um R^2 ajustado *out-of-sample* negativo (recaindo sobre a questão do número de graus de liberdade e do pouco poder explicativo).

Tendo em mente que a perspectiva aqui é essencialmente ancorada nos resultados *in-sample*, tomando-os como necessários para os modelos relevantes, é preciso também ter clareza sobre a diferença do que eles e sua contrapartida *out-of-sample* representam, além da forma de conciliação entre os dois. A diferença existe, muitas vezes, na própria proposição do estudo: se seu intuito é o de formular evidências documentais de variados contextos históricos, ou de simplesmente encontrar bons previsores dos padrões futuros (que não necessariamente explicam bem os do passado). É a partir disso, também, que alguma leniência a certos casos de fraco desempenho *in-sample* pode vir a ser cabível, e a emissão líquida de ações sem dúvida se mostra como principal candidata.

Seus resultados no período mais recente da amostra inspiram otimismo adicional: o R^2 ajustado *in-sample*, quase nulo na amostra completa, apresenta uma melhora vertiginosa, e o desempenho *out-of-sample* é significativo o suficiente para, pela primeira vez, gerar um R^2 ajustado positivo.

Expansão Líquida do Índice	Últimos 5 anos	Amostra Completa
R ² Ajustado IS	4.0491	-0.0018
R ² Ajustado OOS	0.0259	-0.0029

5.2. Modelos Estatisticamente Significantes In-Sample:

Três modelos possuem significância estatística ao nível de 90% na amostra em questão: a volatilidade acumulada para os 30 dias anteriores, a razão valor contábil/preço e a razão lucros/preços. Dentre os três, apenas os dois primeiros se mantêm significantes quando analisados sob seus níveis críticos empíricos.

As características que os definem se mostram comuns em seus gráficos. Tomando a volatilidade como exemplo, devido ao seu comportamento mais acentuado, percebe-se um padrão que remete a uma escada de quatro degraus – cada um deles relacionado a um dos quatro choques fundamentais para a economia brasileira nos últimos 25 anos. É destacável a capacidade de desenvolvimento dos três modelos diante dos quatro eventos: a crise dos países emergentes de 1997, a eleição de Lula em 2002, a crise financeira mundial de 2008 e a crise política/econômica (com o impeachment e a recessão) de 2016. Cada um deles, no hiato entre choques, percorre trajetória notoriamente positiva *in-sample* e sempre acima do nível no hiato anterior (de forma mais ou menos suave a depender do quão ruidosa é cada série).

Welch & Goyal (2008) utilizam, ainda, uma estatística que chamam de “poder” para, juntamente com os resultados *IS-for-OOS*, avaliar a estabilidade dos modelos que têm resultados significantes. Conforme descrito no método de *bootstrap*, nessa análise as simulações são realizadas sob a hipótese alternativa (supondo que há capacidade preditiva e que o modelo em questão constitui o processo gerador de dados original da série do *equity risk premium*). O poder estatístico representa a quantidade de vezes, dentre as amostras simuladas, que os resultados gerados para o ΔREQM superaram seu valor crítico empírico (previamente construído sob a hipótese nula).

Aqui, além de simplesmente condicionar a estabilidade ao sucesso prévio *in-sample*, se relaxa um pouco esta hipótese ao serem reportados os resultados para a razão lucros/preços e para a expansão líquida do índice (por sua significância estatística, nesses casos, não estar muito distante dos 90%, e por características interessantes *out-of-sample* que motivem sua análise mais profunda).

Variáveis	Poder (%)
Volatilidade	1.15
Razão Valor Contábil/Preço	0.15
Razão Lucros/Preços	0.30
Expansão Líquida do Índice	13.91

Com esse instrumental é possível fundamentar de forma mais completa o que poderia estar gerando aos resultados essencialmente insatisfatórios *out-of-sample*: modelos espúrios (cujo R^2 ajustado, tanto *OOS* quanto *IS-for-OOS*, falha estatisticamente) e/ou uma estimação menos precisa dos parâmetros (que seria acusada por percentuais baixos no poder estatístico).

Razão Lucros/Preços:

Apesar de ser o modelo que, a partir das reamostragens, tem sua significância estatística descartada, esse é o que apresenta um comportamento mais consistente *out-of-sample* entre os três. Sua trajetória é, em diversos trechos da amostra, num mínimo tão boa quanto a especificada pelo modelo de média histórica, o que o torna o principal exemplo prático para a já mencionada importância da distinção entre os dois tipos de análise.

Entre 2002 e 2008, seu desempenho flerta com resultados positivos, quase ultrapassando a média prevalecente em capacidade preditiva. Esse é o intervalo mais consistente ao longo de toda a amostra, em uma trajetória que se destaca por apresentar bastante oscilação, refletida no $\Delta REQM$ de -0.0203 e no R^2 ajustado de -0.7719. A única janela contundentemente positiva começa a se construir em torno de 2013 e se materializa em 2015 (curiosamente no mesmo trecho no qual os resultados *in-sample* demonstram queda relevante). Isso é rapidamente seguido, no entanto, por um declínio a níveis mais baixos do que os vigentes quando o modelo se mostrava mais estável.

Os resultados no período mais recente da amostra não destoam de forma significativa de quando são usadas todas as observações. O mesmo ocorre em relação ao R^2 *IS-for-OOS* (4.9682), ainda que, na margem, haja uma melhora nesse caso. O modelo permanece sendo falho *out-of-sample* ao longo de todos os intervalos relevantes de análise.

Razão Lucros/Preços	Últimos 5 anos	Amostra Completa
R ² Ajustado IS	4.0201	4.1019
R ² Ajustado OOS	-0.7426	-0.7719

Volatilidade acumulada nos últimos 30 dias:

A volatilidade é a variável com resultados empíricos mais consistentes *in-sample*: sua estatística F para a amostra original ultrapassa com relativa folga o valor crítico estabelecido, e seu R^2 ajustado é o maior dentre todos os avaliados.

O R^2 *IS-for-OOS* apresenta ligeira redução frente a amostra cheia, mas ainda está muito acima dos demais modelos. Na especificação que se limita aos últimos cinco anos, ainda que se mantenha positivo, o resultado já apresenta queda mais relevante.

A variável aparenta ser a melhor escolha na busca pela explicação de padrões históricos dentre as outras que são testadas, mas é fundamentalmente falha quando usada como previsor. Mais uma vez, em todas as janelas relevantes, não são encontrados resultados *out-of-sample* positivos.

Volatilidade (30 Dias)	Últimos 5 anos	Amostra Completa
R ² Ajustado IS	2.1862	15.1491
R ² Ajustado OOS	-0.8986	-0.9278

Razão valor contábil/preços:

Apesar de ser, dentre as três, a variável que apresenta trajetória de crescimento *in-sample* mais constante – mais fluida e menos oscilante em torno dos choques – a razão entre valor contábil e preços é a que apresenta os resultados menos consistentes *out-of-sample*, conforme mostrado pelo maior distanciamento de seu gráfico em relação ao campo positivo.

Ela é a que apresenta o ganho de desempenho mais relevante quando se olha para o R^2 *IS-for-OOS* (indo de 4.0604 na amostra completa para 5.8875 nesse caso), algo que também fica nítido em seu gráfico. Os resultados do período inicial são muito insatisfatórios até cerca de 1998, quando há um salto para a trajetória positiva mais estável.

Os últimos cinco anos da amostra são extremamente positivos para os resultados *in-sample*, com o R^2 ajustado se aproximando de uma ordem de grandeza só atingida pela volatilidade até então. *Out-of-sample*, no entanto, o ganho é muito marginal e sem relevância prática (mantendo o padrão encontrado para as demais variáveis com significância estatística – nenhuma delas capaz de superar a expansão líquida do índice como melhor previsor).

Razão Valor Contábil/Preços	Últimos 5 anos	Amostra Completa
R ² Ajustado IS	11.7118	4.0604
R ² Ajustado OOS	-1.3124	-1.3416

De forma geral, os resultados para o poder estatístico parecem exercer um impacto muito contundente na análise. Welch & Goyal (2008), olhando para o mercado norte-americano, reportam números num mínimo em torno dos 40% (já condicionais à significância *in-sample*). Aqui, para os modelos que se enquadram nesse critério, o poder estatístico é basicamente nulo. O único modelo que ultrapassa os 10% é o que tem como previsor a expansão líquida do índice – o que apresenta maior proeminência *out-of-sample*, mas já seria descartado de início se o critério estipulado fosse tratado com maior rigidez, dado que sua significância estatística fica (por pouco) aquém dos 90%.

Aliando isso ao fato de que o R^2 *IS-for-OOS* dá indicações positivas para todos os últimos casos abordados, o problema principal parece ser a incapacidade de estimar os parâmetros da forma ideal, ao invés de alguma relação espúria indicada pelos modelos. É uma conclusão que permite certo otimismo à medida que a precisão estatística só tem a crescer, com o desenvolvimento do mercado e o acúmulo de mais dados, permitindo que as mesmas especificações sejam revisitadas no futuro sob uma ótica mais rica.

6. Campbell & Thompson (2005)

Campbell & Thompson (2005) propõe que, em um cenário de previsão, o que eles chamam de “*reasonable investor*”² não seguiria um modelo que previsse um resultado negativo para o *equity risk premium*. Esse seria o caso no qual não haveria alocação de capital para o mercado acionário, prevalecendo o investimento em renda fixa (no ativo livre de risco). Dessa forma, é proposto que todas as previsões negativas sejam truncadas para zero (tanto no modelo sob teste quanto na hipótese nula, equivalente ao modelo de média histórica).

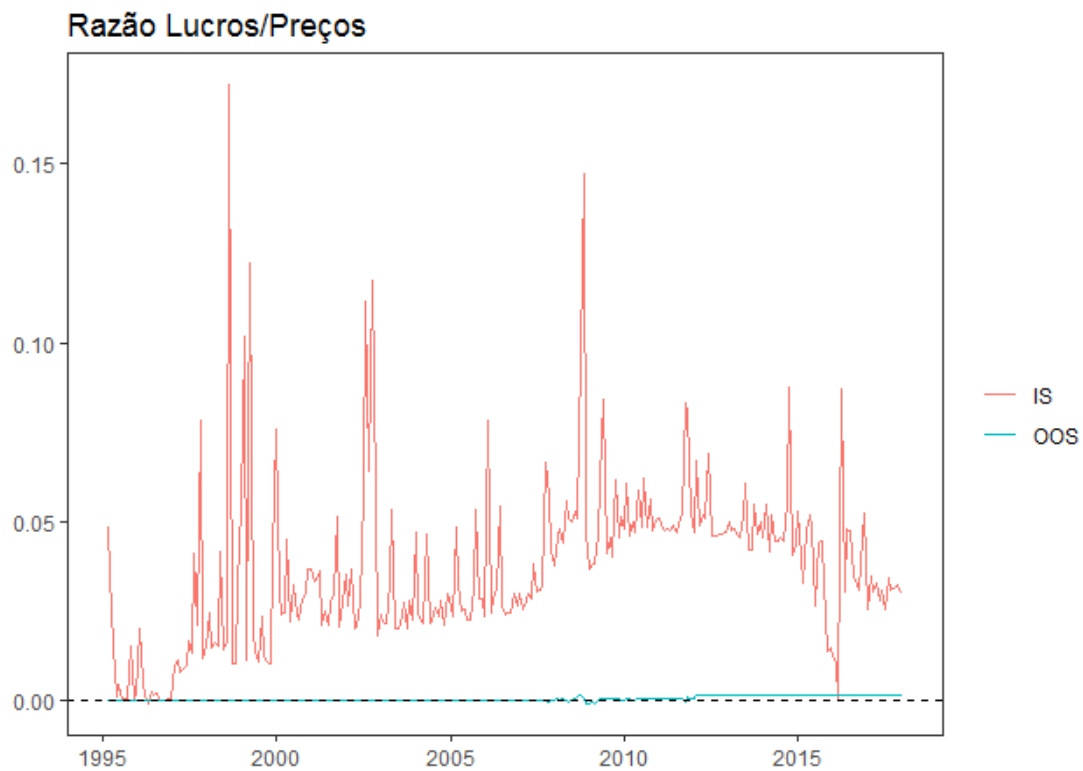
Conforme indicado por Welch & Goyal, essa restrição – que essencialmente transforma a análise básica de regressões lineares em modelos não-lineares – acaba por acrescentar uma dose de cautela nos previsores. Em quaisquer períodos nos quais a realização do *equity risk premium* tenha sido alta (ou, num mínimo, positiva) e nos quais os modelos previam resultados negativos, acabaria por haver uma melhora de performance.

Essa característica, no entanto, acaba por ressaltar algumas especificidades do contexto brasileiro. Em primeiro lugar, o fato de boa parte dos modelos ser essencialmente pessimista em relação ao retorno em excesso da bolsa – alguns deles chegam a indicar um valor negativo ao longo de praticamente todo o período em que se têm observações disponíveis. Dentre os que possuem mais de 80% de suas previsões truncadas para zero estão a razão dividendos/preços (85.09%), o *dividend yield* (81.09%), o IPCA (94.18%) e o diferencial nas posições de títulos públicos (92.72%). São valores surpreendentemente altos e que levam a concluir que, pautado por esses previsores, o “investidor razoável” de Campbell & Thompson jamais teria chegado perto de ações brasileiras (muito por conta disso, esses gráficos não são reportados aqui).

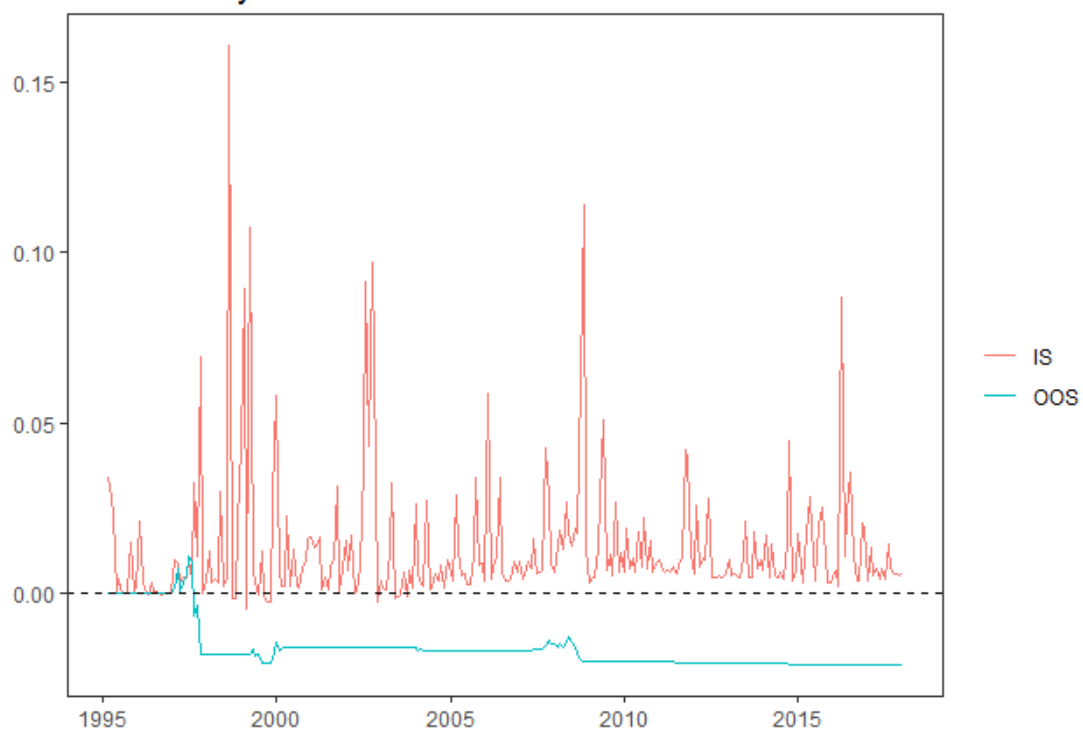
As realizações do *equity risk premium* brasileiro, no entanto, não são totalmente contrárias aos modelos puramente pessimistas – elas são, por si só, negativas em longos períodos da amostra. Dessa forma, o efeito prático do truncamento acaba por ser uma piora nas métricas de erros de previsão *in-sample* – especialmente para as variáveis já mencionadas.

² Algo como um “investidor razoável”, ainda que a tradução livre não seja o ideal.

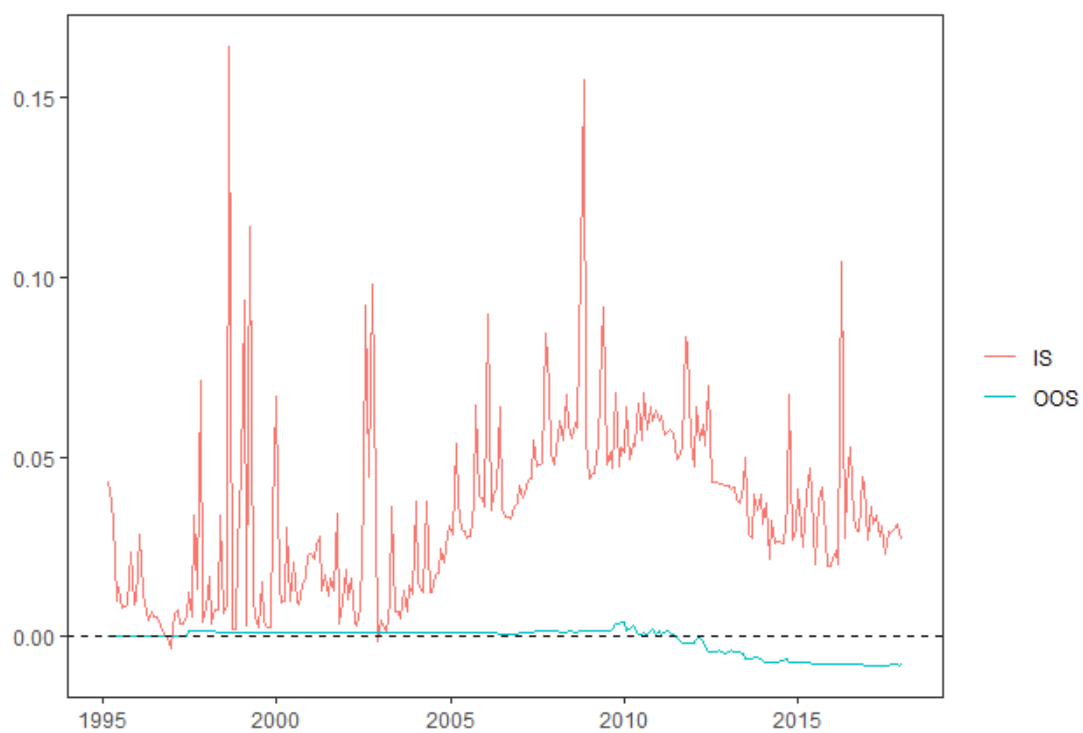
A lógica dos gráficos e da tabela de resultados é exatamente a mesma dos já apresentados anteriormente no cenário principal, mas referente a modelos que se adequam à hipótese de truncamento mencionada acima.



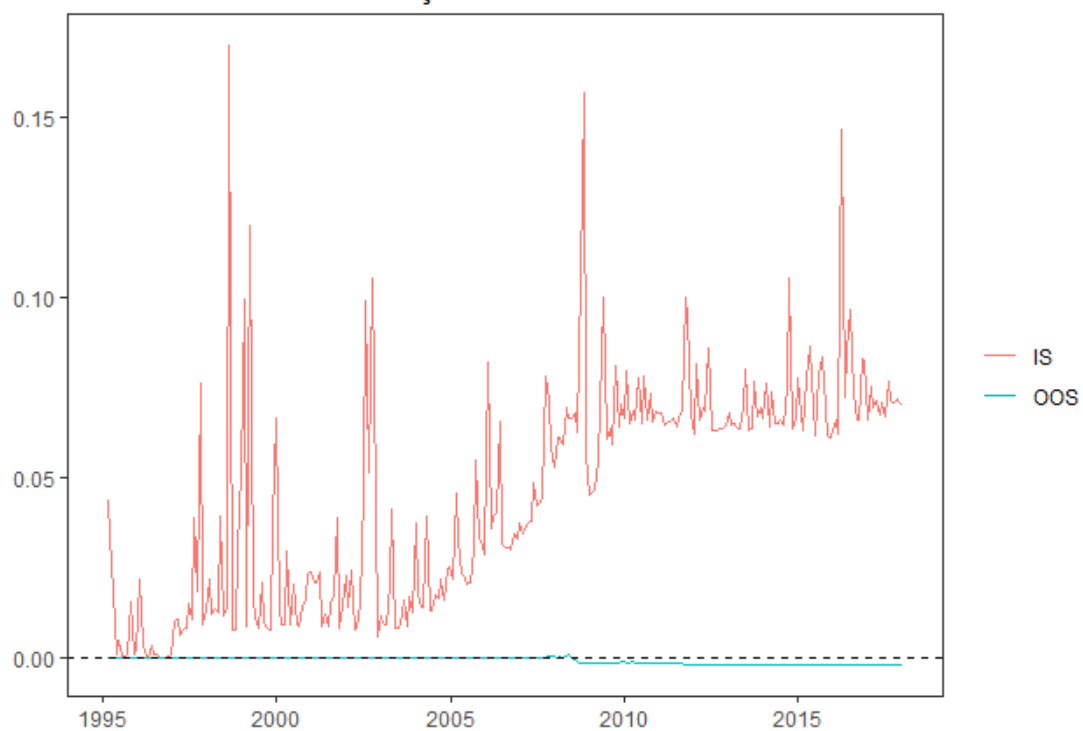
Dividend Payout Ratio



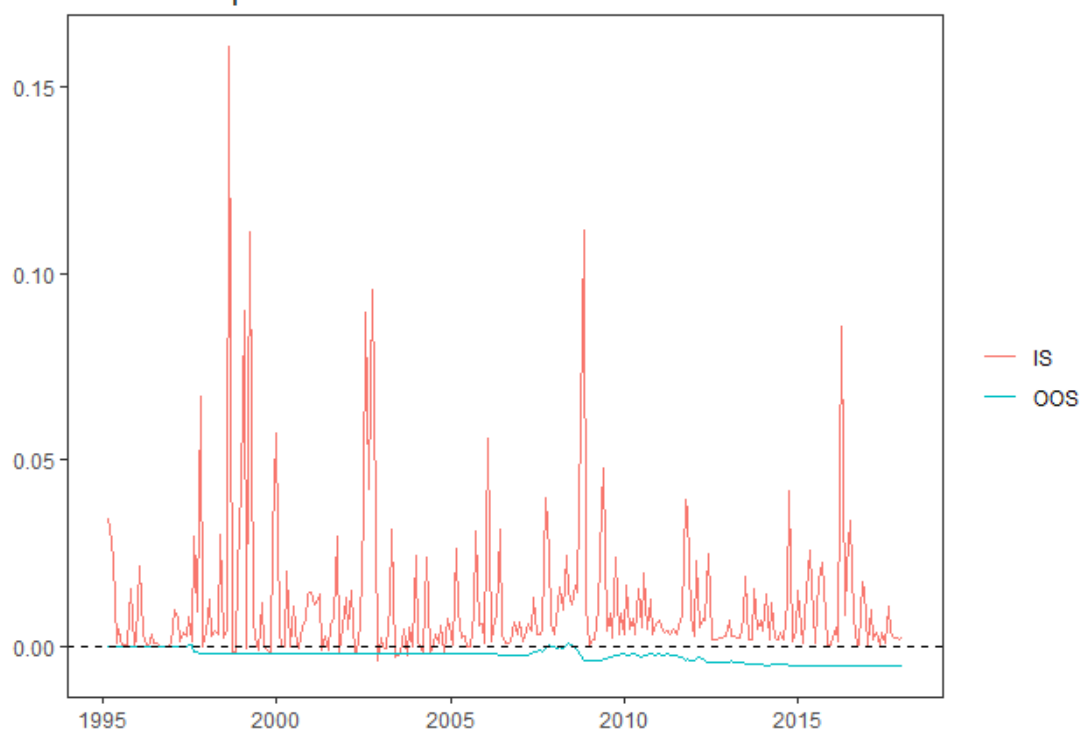
Volatilidade Acumulada Nos Últimos 30 Dias



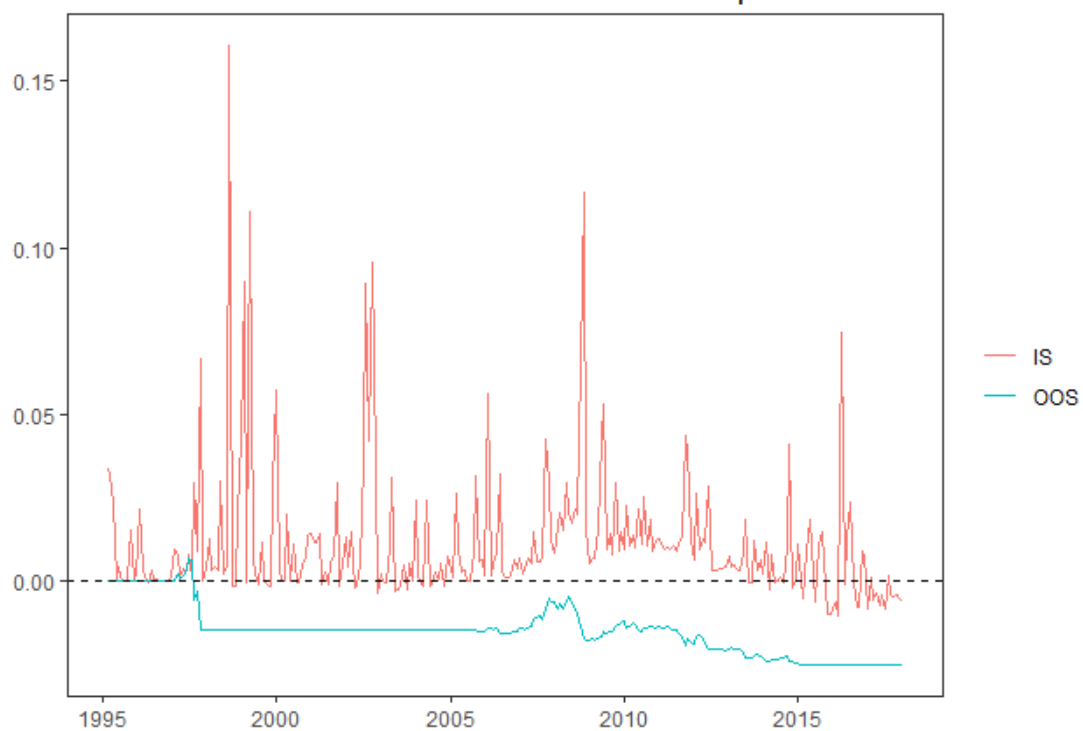
Razão Valor Contábil/Preços



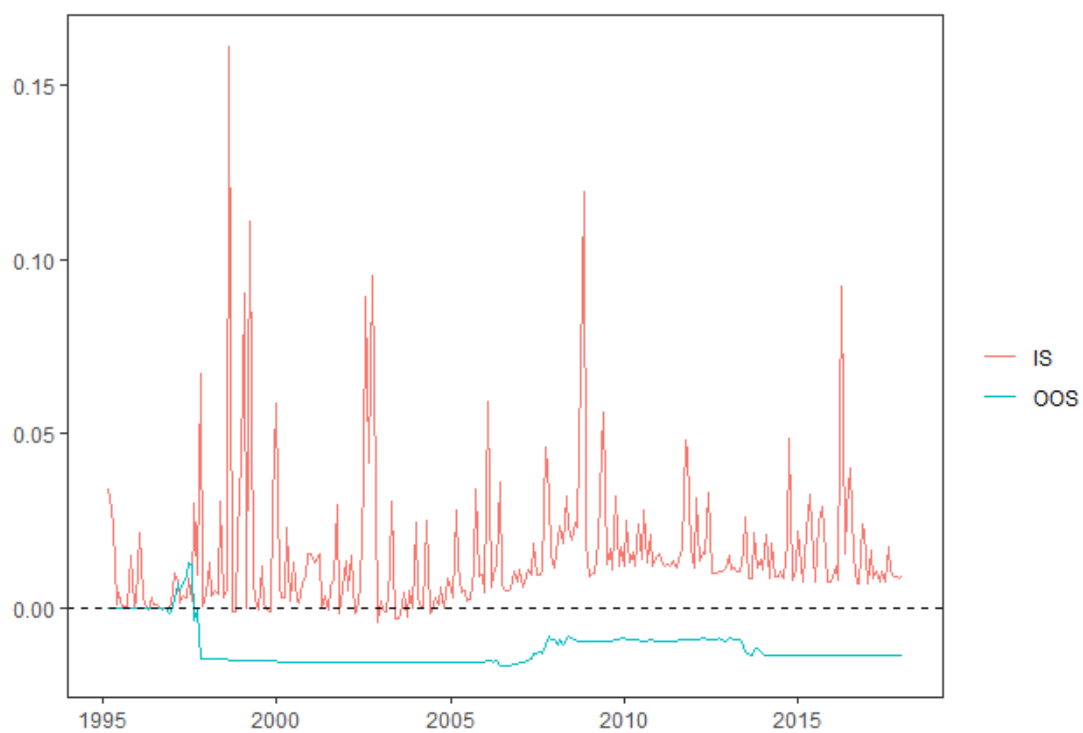
Emissão Líquida



Taxa Média em Novas Emissões de Crédito Corporativo



Taxa Média em Novas Emissões de Crédito Domiciliar



Variáveis	R ² Ajustado – <i>In-sample</i>	R ² Ajustado – <i>IS-for-OOS</i>	R ² Ajustado – <i>Out-of-sample</i>	ΔREQM
Razão Dividendos/Preços	-0.4318	-0.2048	-0.4429	-0.0026
<i>Dividend Yield</i>	-0.4391	-0.1416	-0.4556	-0.0032
Razão Lucros/Preços	-0.1832	0.6206	-0.3498	0.0024
<i>Dividend payout ratio</i>	-0.2288	-0.1968	-1.1138	-0.0385
Volatilidade	0.1525	0.5116	-0.6628	-0.0143
Razão Valor Contábil/Preço	1.3611	1.9954	-0.4682	-0.0039
Expansão Líquida do Índice	-0.3533	-0.3080	-0.5703	-0.0094
IPCA	-0.3516	-0.3186	-1.1981	-0.0430
Variação na Posição de LTNs	0.0178	-0.0558	-0.5509	-0.0083
Taxa Média em Novas Emissões - Crédito Corporativo	-0.5971	-0.5848	-1.2458	-0.0455
Taxa Média em Novas Emissões - Crédito Domiciliar	-0.1201	-0.0822	-0.8532	-0.0245
Spread – Crédito Corporativo e CDI	-0.5185	-0.5020	-1.5367	-0.0610
Spread – Crédito Domiciliar e CDI	-0.3374	-0.3112	-1.8254	-0.0764

Há uma distinção clara no impacto da especificação de Campbell & Thompson nos modelos mais extremos na frequência de suas previsões negativas e nos que demonstram mais equilíbrio. A aplicação não representa um ganho de performance estatística (*in-sample*) em nenhum dos dois casos, mas serve como ênfase para os modelos que, desde o primeiro momento, foram mais significantes: são justamente os que se enquadram no segundo grupo. A volatilidade acumulada nos 30 dias anteriores e a razão valor contábil/preço tiveram 36.72% e 36.36% de suas previsões truncadas para zero, respectivamente. Variáveis que por pouco não foram estatisticamente significantes *in-sample* também apresentam percentuais mais moderados: 46.09% para a razão lucros/preços e 46.18% para a taxa média em novas emissões de crédito domiciliar, por exemplo.

O comportamento *out-of-sample* é um pouco mais ambíguo. O que ele parece indicar é que a introdução de mais conservadorismo nas hipóteses mitiga o impacto negativo das previsões mais extremas (erros de previsão maiores), de modo que o resultado é positivo de forma geral. Isso ocorre com relevância destacável para o diferencial nas posições de títulos públicos (cujo resultado negativo melhora em torno de sete vezes) e em especial para a razão valor contábil/preços (significante *in-sample* e cerca de quatro vezes melhor *out-of-sample* no cenário de Campbell & Thompson do que no cenário base).

As demais variáveis também demonstram ganhos de performance, ainda que em escala menor, em torno de duas a três vezes – as trajetórias no gráfico apresentam uma aproximação relevante da linha tracejada que demarca o zero, o que fica evidenciado também na diminuição dos resultados reportados para o ΔREQM (ainda negativos para todos os modelos menos para a razão lucros/preços, onde se torna minimamente positivo, mas ainda não em magnitude suficiente para que o R^2 ajustado o acompanhe).

A exceção ocorre com a expansão líquida do índice, coincidentemente a variável com melhores resultados *out-of-sample* na análise principal. Tais resultados eram explicados, em princípio, por uma previsão melhor que a média histórica em períodos de realização negativa do *equity risk premium* (o modelo era pessimista de forma mais sensata e precisa). Com o truncamento, os erros de previsão que concediam uma vantagem à variável se perdem ao serem substituídos pelo valor absoluto das observações reais (suas distâncias em relação ao zero em cada um desses períodos).

Campbell & Thompson têm como ponto principal de seu trabalho a aplicação dos modelos testados à *trading strategies* – seu uso prático como motivação para a construção de estratégias de alocação (ou não alocação) de capital no mercado acionário. Eles destacam que “em horizontes mais curtos, até mesmo diferenças muito pequenas para o ΔREQM são suficientes para gerar ganhos ao investidor avesso ao risco” (com o coeficiente de aversão igual a 3 utilizado na hipótese básica que estipulam). Os resultados preditivos são aplicados, de certa forma, com alguma independência da significância estatística que possuem (ou não, na maior parte dos casos), e é muito ampla a gama de hipóteses que pode ser testada – por exemplo,

variando os perfis de aversão ao risco. Esse tipo de análise abre espaço para que, conforme o mercado brasileiro estenda sua história e forneça mais robustez aos seus dados, novos estudos possam surgir tendo como base o mesmo arcabouço aqui descrito e construído.

7. Modelos com “Memória Estendida”

Além dos modelos já abordados, aqui são expostos os resultados para especificações que carregam o que Welch & Goyal chamam de “memória estendida” – significando, basicamente, um horizonte mais amplo no acúmulo da informação usada para a construção de cada variável. Foca-se na volatilidade e na razão lucros/preços, por serem duas das que apresentaram resultados mais relevantes até aqui e que são cabíveis nessa especificação alternativa.

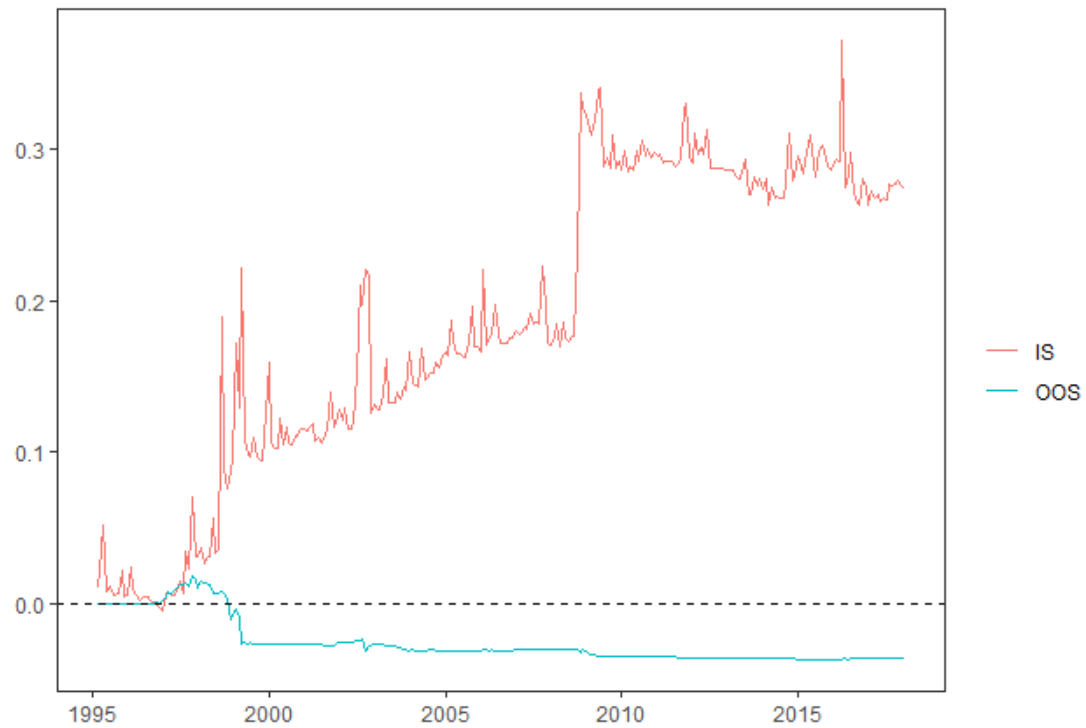
A trajetória da volatilidade é curiosa: quanto mais longo o período de acúmulo na composição da variável, mais ruidosa parece sua representação gráfica, piores são seus resultados *in-sample* e melhores *out-of-sample*. Para as especificações de prazo ainda relativamente curto (60, 90 e 120 dias), o poder explicativo ainda é consideravelmente positivo, só não tanto quanto a proposta original (30 dias).

O uso da média móvel dos 10 anos anteriores dos lucros sobre os preços resulta em valores contundentemente piores em comparação com o uso apenas dos 12 meses anteriores. O R^2 *in-sample* se torna negativo, e o *out-of-sample* (que já era) piora em mais de quatro vezes. Os resultados, de forma geral, deixam claro que o aumento de memória gera valores mais pobres do que seus respectivos pares na especificação original

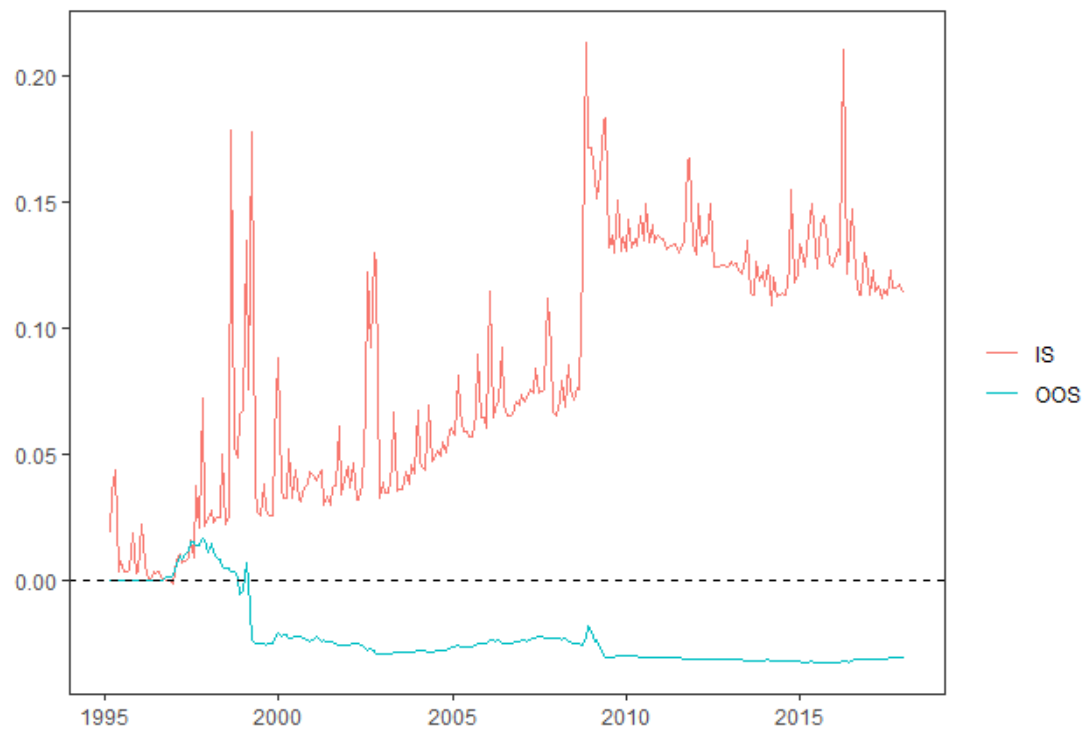
Variáveis	R ² Ajustado – <i>In-sample</i>	R ² Ajustado – <i>IS-for-OOS</i>	R ² Ajustado – <i>Out-of-sample</i>	Δ REQM
Volatilidade Acumulada – 60 dias anteriores	9.4811	8.9026	-1.5638	-0.0629
Volatilidade Acumulada – 90 dias anteriores	3.9382	3.4685	-1.4229	-0.0554
Volatilidade Acumulada – 120 dias anteriores	2.5376	2.1191	-1.2286	-0.0449
Volatilidade Acumulada – 180 dias anteriores	0.3998	0.0947	-1.1919	-0.0429
Volatilidade Acumulada – 200 dias anteriores	0.1667	-0.1475	-1.3332	-0.0505
Volatilidade Acumulada – 260 dias anteriores	0.0263	-0.2870	-0.9819	-0.0316
Volatilidade Acumulada – 360 dias anteriores	-0.2339	-0.4394	-0.7262	-0.0179
Razão Lucros/Preços – Média Móvel de 10 anos	-0.1826	0.7354	-3.7804	-0.1664

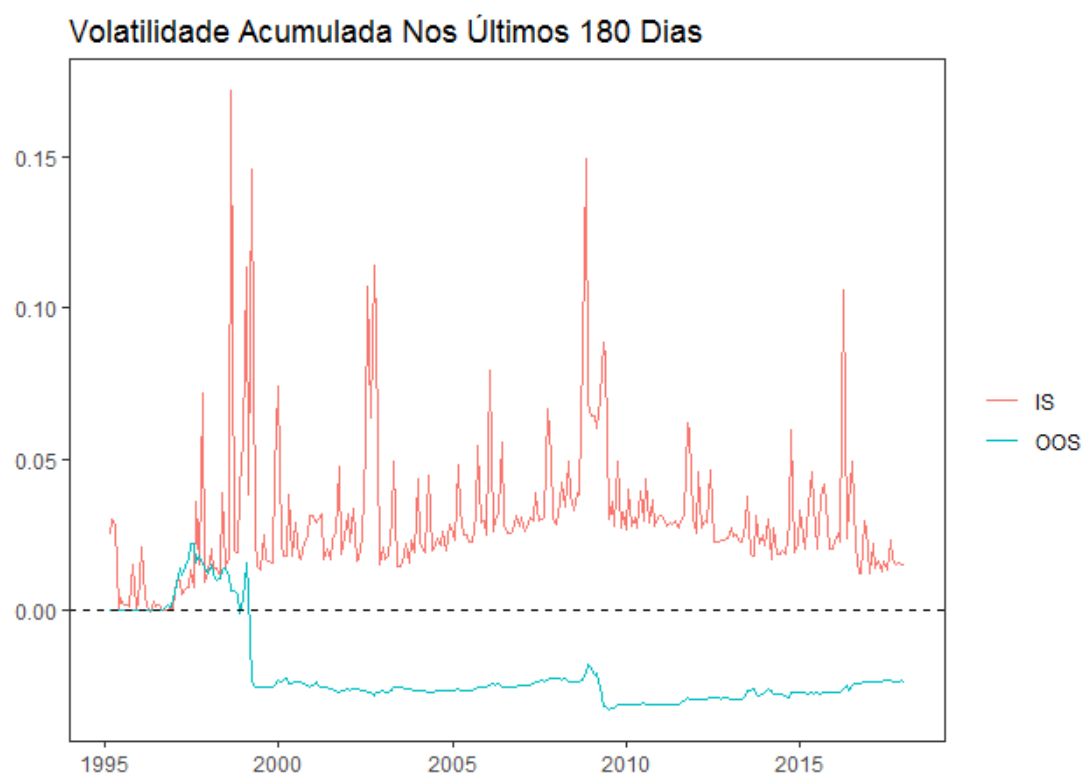
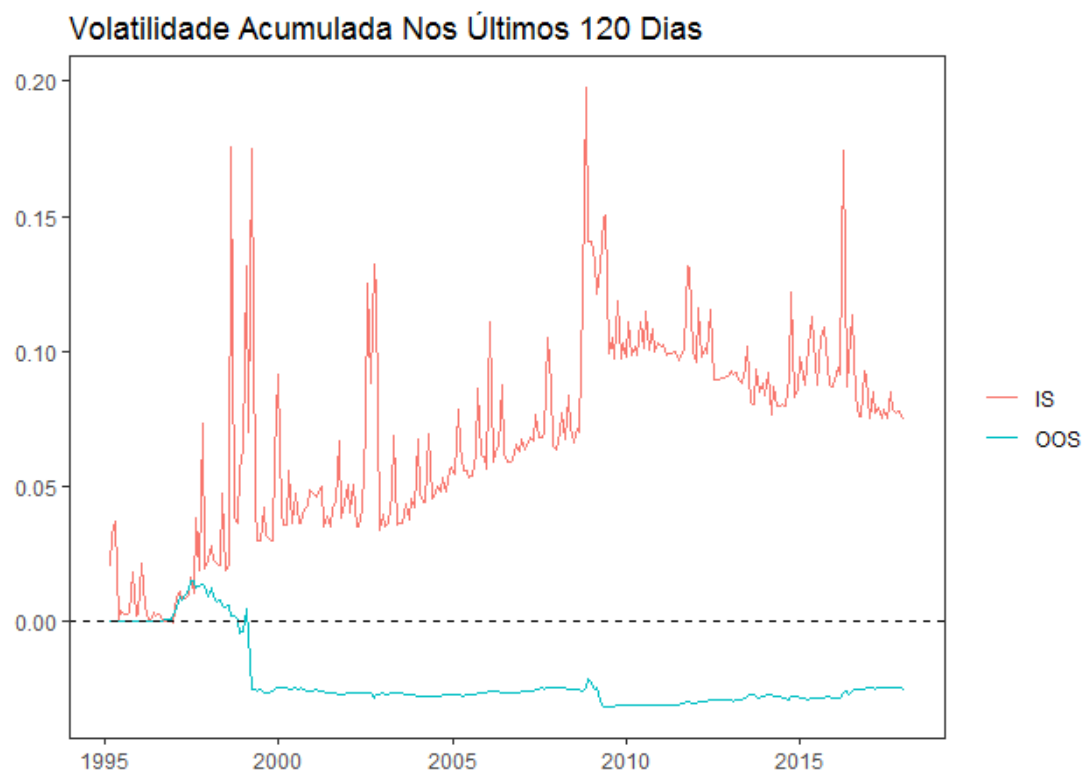
A lógica dos gráficos e da tabela de resultados é exatamente a mesma dos já apresentados anteriormente no cenário principal, mas aplicada aos modelos com “memória estendida” em sua especificação.

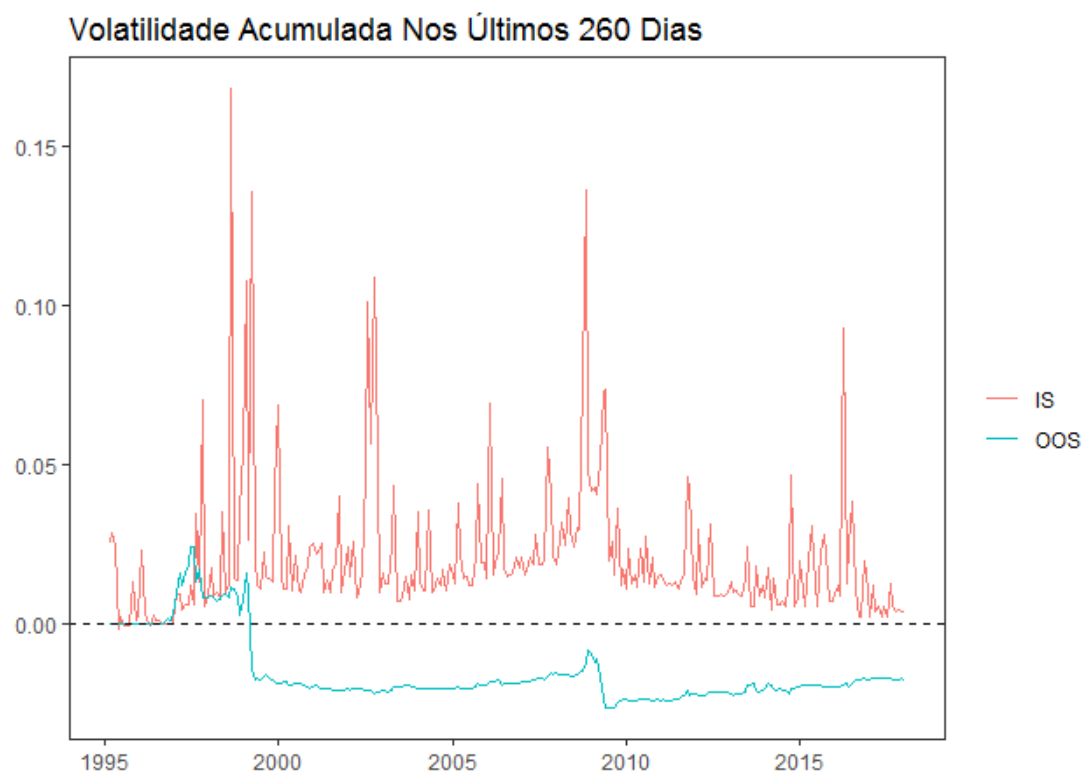
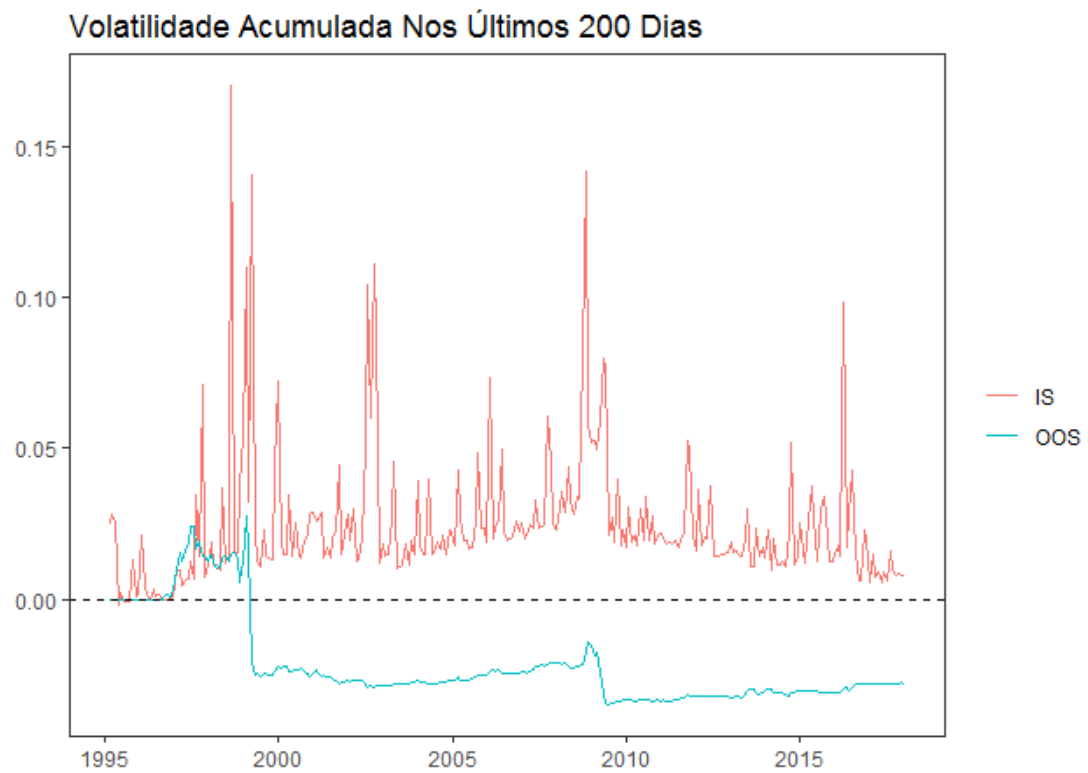
Volatilidade Acumulada Nos Últimos 60 Dias

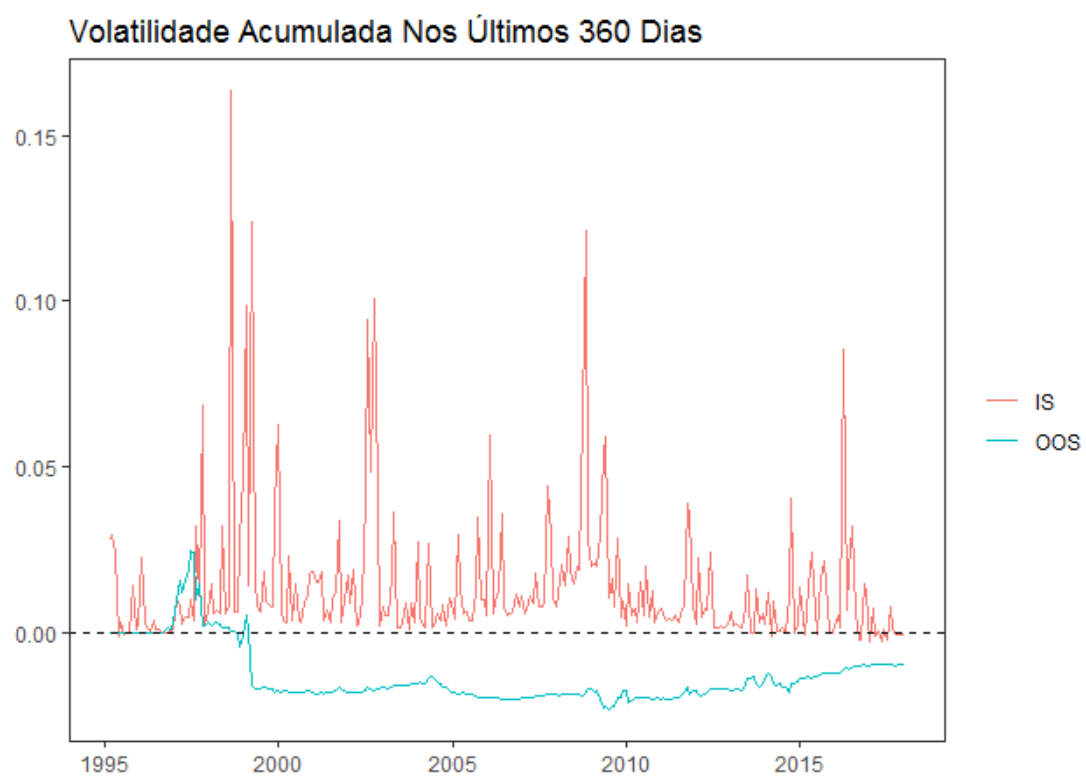


Volatilidade Acumulada Nos Últimos 90 Dias









8. Conclusão

Este trabalho propõe uma extensiva análise do desempenho *in-sample* e *out-of-sample* de regressões lineares (em sua maioria), utilizando o que há de mais famoso na literatura de previsores do *equity risk premium* e testando sua aplicabilidade ao contexto brasileiro. É fruto, principalmente, do arcabouço proposto por Welch & Goyal (2008), se fazendo valer dele na maioria das análises, mas propondo também uma flexibilização de algumas das hipóteses que os autores tomam como necessárias, especialmente aqui na conclusão dos resultados obtidos.

Welch & Goyal postulam que o desempenho *out-of-sample* não é um substituto, mas sim um complemento necessário a bons resultados *in-sample* para que se possa considerar válido um modelo. De acordo com eles, é um “diagnóstico importante nas regressões se, e somente se, o modelo for significante *in-sample*”. De modo a aumentar a abrangência de análise, busca-se aqui uma relação um pouco mais independente entre os dois cenários. Sem que os resultados *in-sample* deixem de ser a base, o modelo que utiliza a expansão líquida do índice, ainda que inicialmente fraco nesse pré-requisito, permite certo otimismo em torno de seu acompanhamento futuro (não apenas pelos resultados *out-of-sample* ao longo de toda a amostra, mas também pela melhora contundente *in-sample* no período final). Os demais, nem com toda a dose de flexibilização que fosse necessária apresentam resultados que atestem para sua validade como bons previsores, estando constantemente atrás da média prevalecente para o *equity risk premium*.

Ainda assim, é possível, também, destacar características que impeçam a invalidação imediata dessas especificações, todas ligadas especialmente ao fato de que qualquer análise sobre o mercado brasileiro olha, ainda, para um cenário muito jovem. Diversos gráficos ilustram que o desempenho *out-of-sample* é especialmente negativo no início das previsões, fato que se liga a uma capacidade mais pobre de seus testes. Aliada a isso está a avaliação de estabilidade dos modelos, com resultados muito fracos e que parecem indicar uma estimação dos parâmetros distante do que seria ideal. São questões que dependem de mais anos de mercado desenvolvido e um volume maior de dados robustos para serem descartadas e permitirem uma avaliação *out-of-sample* mais clara.

Em termos unicamente explicativos e documentais, a volatilidade acumulada nos últimos trinta dias e a razão valor contábil/preços se destacam entre as demais. Não podem ser consideradas boas previsoras futuras, mas apresentam resultados *in-sample* que sem dúvida são relevantes, nem que seja em retrospecto.

De forma geral, os resultados são promissores à medida que cobrem uma área de estudo incipiente no caso brasileiro, fundamentam suas bases práticas e teóricas, e ainda indicam direções que, sob análises futuras, podem vir a se mostrar mais relevantes.

Referências Bibliográficas

- BAKER, M.; WURGLER, J. The equity share in new issues and aggregate stock returns. **Journal of Finance**, vol 55(5), p. 2219–2257. 2000.
- BOUDOUKH, J.; RICHARDSON, M.; WHITELAW, R. F. The Myth of Long-Horizon Predictability. **NBER Working Paper No. 11841**. 2005.
- CAMPBELL, J. Y. Stock returns and the Term Structure. **Journal of Financial Economics**, vol 18(2), p. 373–399. 1987.
- CAMPBELL, J. Y.; SHILLER, R. J. The dividend-price ratio and expectations of future dividends and discount factors. **Review of Financial Studies**, vol 1(3), p. 195–202. 1988.
- CAMPBELL, J. Y.; THOMPSON, S. B. Predicting the Equity Premium Out of Sample: Can Anything Beat the Historical Average? **NBER Working Paper No. 11468**, Julho/2005.
- COCHRANE, J. The Dog That Did Not Bark: A Defense of Return Predictability. 2007. **Review of Financial Studies, Oxford University Press for Society for Financial Studies**, vol. 21(4), p. 1533-1575. Julho/2008.
- FAMA, E. F.; FRENCH, K. R. Dividend yields and expected stock returns. **Journal of Financial Economics**, vol. 22(1), p. 3-25. Outubro/1988.
- FAMA, E. F.; SCHWERT, G. W. Asset returns and inflation. **Journal of Financial Economics**, vol 5(2), p. 115–146. 1977.
- FERSON, W.; HARVEY, C. The Variation of Economic Risk Premiums. **Journal of Political Economy**, vol 99(2), p. 385-415. 1991.
- FOSTER, D.; SMITH, T.; WHALEY, R. E. Assessing Goodness-of-Fit of Asset Pricing Models: The Distribution of the Maximal R^2 . **The Journal of Finance**, vol 52(2), p. 591-607. Junho/1997.
- HARVEY, C. Time-varying conditional covariances in tests of asset pricing models. **Journal of Financial Economics**, vol 24, p. 289—317. 1989.
- KEIM, D. B.; STAMBAUGH, R. F. Predicting returns in the stock and bond markets. **Journal of Financial Economics**, vol 17(2), p. 357–90. 1986.
- LETTAU, M.; LUDVIGSON, S. Consumption, aggregate wealth, and expected stock returns. **Journal of Finance**, vol 56(3), p. 815–849. 2001.
- LEWELLEN, J. Predicting returns with financial ratios. **Journal of Financial Economics**, 74(2), p. 209–235. 2004.

McLEAN, R. D.; PONTIFF, J. Does Academic Work Destroy Stock Return Predictability?. **The Journal of Finance**, vol 71(1), p. 5-32. Outubro/2015.

Pedro Costa Ferreira, Talitha Speranza and Jonatha Costa (2018). BETS: Brazilian Economic Time Series. R package version 0.4.4. <https://CRAN.R-project.org/package=BETS>

POLK, C.; THOMPSON, S. B.; VUOLTEENAHO, T. Cross-Sectional Forecasts of the Equity Premium. **Journal of Financial Economics**, vol 81, p. 101-147. Julho/2006.

R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: <https://www.R-project.org/>.

ROZEFF, M. S. Dividend yields are equity risk premiums. **Journal of Portfolio Management**, vol 11(1), p. 68–75. 1984.

STAMBAUGH, R.F. Predictive Regressions. **Journal of Financial Economics**, vol 54(3), p. 375-421. 1999.

WELCH, I.; GOYAL, A. A Comprehensive Look at The Empirical Performance of Equity Premium Prediction. **Review of Financial Studies, Oxford University Press for Society for Financial Studies**, vol. 21(4), p. 1455-1508. Julho/2008.