

PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO DE JANEIRO

DEPARTAMENTO DE ECONOMIA

Monografia de Final de Curso

Follow The Crowd

**Estratégias de fund replication para fundos multimercado
brasileiros**



Pedro Henrique de Arruda Torres

No. de matrícula: 1811163

Orientador: Carlos Viana de Carvalho

Coorientador: Márcio Garcia

Novembro 2021

PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO DE JANEIRO
DEPARTAMENTO DE ECONOMIA
Monografia de Final de Curso

Follow The Crowd
**Estratégias de fund replication para fundos multimercado
brasileiros**

Pedro Henrique de Arruda Torres
No. de matrícula: 1811163

Orientador: Carlos Viana de Carvalho
Coorientador: Márcio Garcia

Novembro 2021

**Declaro que o presente trabalho é de minha autoria e que não recorri
para realiza-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

Por sorte ou bênção, em cada etapa da minha vida, não faltaram pessoas ao meu lado para me guiar, proteger e amar. Com a conclusão de mais capítulo da minha vida, gostaria de dedicar esse momento para agradecê-las.

Primeiro, preciso agradecer a minha Mãe, pessoa mais forte, inteligente e carinhosa que já conheci. Não seria nada sem você

Agradeço ao meu Pai, por ser um exemplo de resiliência. Só hoje, percebo sua força

Ao meu irmão, sou grato por te ter como guia, sempre a frente, me mostrando qual caminho traçar

Preciso destacar o papel fundamental dos meus amigos, que nesse momento de transição, no qual a vida adulta vem à tona, me lembram da importância de uma boa risada

Aos meus amigos de infância, agradeço ao Thomaz que sempre caminhou ao meu lado, e o Judice que me surpreende todo dia com sua coragem e ambição

Lembro do Gabriel, Pedro, Leo e Alves que estavam ao meu lado nos meus momentos mais felizes e se mantiveram lá nos momentos ruins

Na PUC, fui muito feliz de me aproximar de um grupo muito especial de pessoas e espero me manter perto de vocês para sempre. Tenho certeza de que a gente continuará se puxando para cima nessa nova fase da vida. São eles: Álvaro, André, Prates, Adinha, Chloé, Bia, Gabi, Luisa, Isadora, Juliana, Gustavo, Lenz, Luccas, Manu, Nic, Pedro, Pimenta e Tom

Obrigado pelos bons momentos

Agradeço ao Fernando Texeira pela ajuda nessa monografia, e por sempre confiar em mim

E por fim, agradeço ao meu Orientador Carlos Viana de Carvalho, pela atenção e

motivação durante todo este projeto

SUMÁRIO

1 - Introdução.....	8
2 - Motivação	9
3 - Revisão de Literatura	10
4 - Dados.....	18
4.1 - Retorno dos Ativos.....	18
4.2 - Retorno dos fundos	19
5 – Metodologia	22
5.1 – Inferência dos Portfólios.....	22
5.2 – Construção das Carteiras Sintéticas.....	25
5.3 – Controles para <i>Look-ahead bias</i>	26
6 – Resultados.....	27
6.1 – Teste da Metodologia	27
6.2 – Retorno por Ativo.....	30
6.3 – Retorno das Carteiras.....	31
6.4 – Análises Adicionais.....	38
7 – Conclusão.....	40
8 – Referências	42

Lista de Figuras

Figura 3-1 Índices Sharpe para os fundos clones e originais do estudo de (Lo [†] , 2006).....	16
Figura 4-2-1 Correlograma dos retornos dos fundos para identificar feeders.....	20
Figura 5-1-1 Coeficientes estimados para gestoras anônimas	23
Figura 5-1-2 R-quadrados obtidos para todas as regressões móveis, em azul está o valor mediano por dia.....	23
Figura 5-1-3 Coeficientes agregados - Média e Mediano	24
Figura 5-1-4 Coeficientes estimados para Média por PL e Top 5	25
Figura 6-1-1 Resultado de teste com carteiras de 2 ativos.....	27
Figura 6-1-2 Resultado do teste com carteiras de 3 ativos.....	28
Figura 6-1-3 Resultado do teste com carteiras dinâmicas de 2 ativos	29
Figura 6-1-4 Correlograma para evidencias diferença de correlação entre ativos.....	29

Lista de tabelas

Tabela 1 – Dados descritivos da performance entre 1982 e 1999 dos mutual funds analisados em (FRANS A. DE ROON, 2000).	11
Tabela 2 Comparação das performances dos mutual funds e das carteiras mímicas em (FRANS A. DE ROON, 2000).....	12
Tabela 3 Informações descritivas dos ativos usados.....	19
Tabela 4 Nome dos Fundos Master, seu CNPJ e o CNPJ do Feeders utilizados em seu lugar	21

“Imitation is the sincerest form of flattery.”

– Charles Caleb Colton

1 -Introdução

Este artigo propõe uma estratégia de investimentos quantitativa como alternativa aos fundos multimercados do Brasil. A metodologia descrita infere as composições dos portfólios das principais gestoras macro, e a partir disso, estima uma carteira sintética que captura mudanças de marginais de alocação. Com as posições reestimadas a partir de janelas móveis, o método propõe manter-se atualizado com as mudanças de cenário dos gestores.

Mais especificamente, trata-se de uma estratégia de *fund replication*, no sentido que o algoritmo proposto não irá compor uma estratégia de investimento de maneira independente, mas sim identificar e replicar a estratégia dos principais gestores do mercado. Seleccionamos uma cesta de ativos que captura grande parte da exposição dos multimercados em relação ao câmbio, juros e ações. Com posições compradas e vendidas, montamos os portfólios sintéticos.

Veremos que há diversas formas de agregar as posições estimadas. Por isso, avaliaremos quais as melhores formas de selecionar e ponderar os fundos usados. Além disso, veremos também qual o tamanho ótimo da janela temporal, a fim de maximizar os retornos. Vale ressaltar também que, decorrente do processo de inferência das posições, geramos muita informação residual que não utilizamos na composição final das carteiras. Abordaremos algumas das análises possíveis a partir desses dados no Capítulo 6.

Apesar da considerável transparência do mercado de fundos no Brasil, especialmente quando comparado a outros países, só temos acesso às verdadeiras posições que compõem os portfólios após o decurso de 3 meses. Por isso, um ceticismo em relação à acurácia do processo de inferência é razoável. Com o intuito de testar a capacidade do modelo de capturar as posições tomadas, montamos carteiras testes com composições conhecidas e testamos o método, cujos resultados serão abordados no Capítulo 6.

Após a construção das diferentes carteiras sintéticas, analisaremos os *backtests* das rentabilidades das estratégias, comparando-as aos rendimentos dos fundos que as originaram. Abordaremos, mais adiante as precauções tomadas para evitar qualquer viés de *look-ahead*, preservando, assim, a confiabilidade de que os resultados obtidos seriam análogos aos resultados da estratégia empregada em um cenário em que não há qualquer informação futura.

2 - Motivação

Em países com sistemas financeiros mais amadurecidos, como os Estados Unidos, percebemos que há uma diminuição da demanda pelos serviços prestados por *hedge funds*. Como consequência, observamos que as rotineiras taxas de 2% de administração e 20% sobre a performance estão caindo em desuso¹. Dito isso, no Brasil, a demanda por fundos multimercados, acelerada pela queda da taxa SELIC dos últimos anos, vem numa forte crescente e as altas taxas cobradas não devem abaixar tão cedo².

Ademais, veremos também, com *papers* como (Young, 2007), que o mercado de fundos expõe os clientes a uma forte seleção adversa, onde gestores “ruins” conseguem aparentar competência e oferecem retornos satisfatórios, sem necessariamente gerar *alpha*. Ademais, as estratégias aplicadas por esses gestores ficam expostas a acentuados riscos de calda, de maneira a justificar o retorno oferecido no curto prazo.

Nesse contexto, temos que o mercado de fundos no Brasil é não só caro, mas pode também expor os investidores à seleção adversa. Assim, a estratégia proposta neste trabalho pode vir a ser uma interessante alternativa para os fundos multimercados, não só na perspectiva de rentabilidade, mas também para evitar que os investidores fiquem expostos aos riscos de calda elevados, devido a gestores incompetentes.

¹ Nishant K., julho, 2020. Hedge Fund Fees in Free Fall Is the New Reality For a Humbled Industry <https://www.bloomberg.com/graphics/2020-hedge-fund-management-performance-fees/>

² Lucas B., agosto, 2021. Impulsionada por demanda pela renda fixa, indústria de fundos capta quase R\$ 300 bi em 2021 até julho <https://www.infomoney.com.br/onde-investir/impulsionada-por-demanda-pela-renda-fixa-industria-de-fundos-capta-quase-r-300-bi-em-2021-ate-julho/>

3 - Revisão de Literatura

A literatura sobre *fund-performance* teve enormes avanços devido às contribuições de William Sharpe, até hoje o trabalho dele é constantemente lembrado com a utilização do Índice Sharpe, métrica que mensura a performance excedente de um portfólio ajustado ao risco ³. Dito isso, seus trabalhos sobre inferência de portfólio e *Factor Analysis* são especificamente úteis para fundamentar a proposta desta monografia.

Ao abordarmos a metodologia por detrás do algoritmo, ficam evidentes as semelhanças com os estudos de *Factor Analysis* introduzidos por Sharpe. Em (Sharpe, 1992), somos apresentados a uma análise de portfólio que consiste em explicar a variância da cota de um fundo utilizando diferentes classes de ativos.

Teríamos:

$$\tilde{R}_i = [b_{i1}\tilde{F}_1 + b_{i2}\tilde{F}_2 + \dots b_{in}\tilde{F}_n] + \tilde{e}_i$$

Onde \tilde{R}_i seria o retorno do portfólio, \tilde{F}_i seria os fatores e b_i seriam os diferentes pesos para cada fator. Dessa forma, uma combinação linear dos fatores descreve o comportamento de R_i . Nos primeiros exemplos, Sharpe usa uma combinação de 12 *Asset Classes*. Dentre elas temos índices de debêntures públicas e privadas de diferentes países e índices de tipos distintos de ação como *Value* e *Growth*. Como veremos a frente, na nossa análise, abordaremos o problema de inferência de maneira mais dinâmica, numa janela móvel e curta de tempo. Por isso, é necessário um número reduzido de ativos. Incluímos índices de câmbio às variáveis explicativas devido à demanda das mesas de multimercado brasileiras por esses ativos.

Esse arcabouço proposto por Sharpe serve como base para a metodologia descrita nessa monografia, além disso, vale ressaltar ainda outros trabalhos que ajudaram a guiar este projeto. Nessa linha temos as contribuições de (FRANS A. DE ROON, 2000) na qual os autores usam *Style Analysis* para construir benchmarks individuais para os diferentes gestores e criam também portfólios que imitam essas carteiras e comparam os retornos com as originais.

No paper, os autores investem considerável energia em comparar o resultado das equações quando há ou não restrições que garantem que os fatores sejam positivos e somam 1.

³Índice Sharpe - disponível em: <https://jpm.pm-research.com/content/21/1/49>

Essa mesma questão foi abordada durante a escolha da metodologia utilizada mais adiante neste paper. Dessa forma, veremos que a maneira como lidamos com as diferentes volatilidades dos ativos pressupõe que sejam possíveis níveis de alavancagem maiores que 1, de modo que não aplicaremos esta restrição.

No artigo citado acima, os autores tentam responder uma pergunta bastante similar à abordada por nós. Nesse sentido, eles buscam averiguar a viabilidade de uma carteira inferida como alternativa aos mutual funds. Com esse intuito, os autores compilam a performance de 18 mutual funds, e calculam as regressões usando somente 2 fundos passivos como contraparte, o Vanguard 500 index (large cap) e o Vanguard International (value stock, ex EUA). Na tabela abaixo temos os fundos comparados.

Mutual Fund	Style	Average Return	Stand. Dev.	Net Assets	Front Load
		(%)	(%)	(mln \$)	(%)
Alliance Global Sm	W	0.94	5.80	74.4	4.25
Alliance Intl	F	1.09	5.04	76.7	4.25
Bailard, Biehl Intl	F	0.99	4.95	113.4	0.00
Evergreen Intl Gr	F	0.94	4.37	66.0	0.00
First Invest Global	W	1.17	4.92	312.4	6.25
Kemper Intl	F	1.10	4.45	398.4	5.75
Nations Intl Gr	F	1.05	4.81	22.4	0.00
New Perspective	W	1.39	3.86	23061.1	5.75
Oppenheimer Global	W	1.33	4.99	3580.5	5.75
Phoenix-Aberdeen	W	1.00	5.67	185.3	4.75
Putnam Global Gr	W	1.35	4.23	3518.3	5.75
Scudder Intl	F	1.24	4.44	3103.7	0.00
T.Rowe Price Intl	F	1.26	4.56	10006.7	0.00
Templeton Global Sm	W	1.14	4.36	1095.8	5.75
Templeton Gr	W	1.28	4.08	12319.5	5.75
Templeton World	W	1.31	4.10	8589.9	5.75
United Intl Gr	F	1.27	4.29	1236.4	5.75
Vanguard Intl Gr	F	1.32	4.77	7601.6	0.00

Tabela 1 – Dados descritivos da performance entre 1982 e 1999 dos mutual funds analisados em (FRANS A. DE ROON, 2000).

Agora, os autores utilizam da regressão abaixo para obter \hat{a}_i que seria a métrica de excesso de retorno dos mutual funds em relação às combinações lineares dos fundos explicativos.

$$r_{i,t} = \hat{a}_i + \hat{b}_1 R^{Vang500} + \hat{b}_2 R^{VangWorld} + e_{i,t}$$

De início, os autores consistentemente encontram valores negativos para \hat{a}_i , contudo eles consideram o fato de que diferenças em volatilidade poderiam estar superestimando a performance da combinação linear, por isso, analisam os Índices Sharpe dos retornos.

Nesse momento, fica evidente que os “*mimicking*” portfólios se mostraram sim como alternativas viváveis para os Mutual Funds. Note que, exceto no caso do fundo New Perspective, um investidor deveria sempre escolher uma combinação linear dos fundos passivos mesmo quando ajustamos por unidade de volatilidade.

Mutual Fund	Jensen Measure	Sharpe Ratio	
	$\eta = 1.43\%$	mutual fund	mim. portfolio
Alliance Global Sm	-0.543*	0.163	0.355
Alliance Intl	-0.256	0.216	0.303
Bailard, Biehl Intl	-0.335*	0.200	0.282
Evergreen Intl Gr	-0.419*	0.215	0.314
First Invest Global	-0.218	0.238	0.338
Kemper Intl	-0.244*	0.248	0.302
Nations Intl Gr	-0.291*	0.218	0.296
New Perspective	-0.036	0.360	0.358
Oppenheimer Global	-0.059	0.267	0.341
Phoenix-Aberdeen	-0.474*	0.176	0.360
Putnam Global Gr	-0.046	0.318	0.341
Scudder Intl	-0.117	0.278	0.307
T.Rowe Price Intl	-0.087	0.276	0.299
Templeton Global Sm	-0.308	0.261	0.362
Templeton Gr	-0.170	0.313	0.362
Templeton World	-0.139	0.320	0.363
United Intl Gr	-0.111	0.295	0.329

Tabela 2 Comparação das performances dos mutual funds e das carteiras mímicas em (FRANS A. DE ROON, 2000).

Vale também ressaltar a diferença entre Factor Analysis e Style Analysis, ambas metodologias buscam explicar o retorno dos fundos por meio de índices que capturam diferentes fontes de risco do mercado. Assim, a natureza destes índices que explicam as diferenças entres *Style* e *Factor*. Dessa maneira, *Factor Analysis* utiliza fatores sistemáticos como Volatilidade, Momentum e tamanho, enquanto isso, *Style Analysis* usa índices que representam segmentos de mercado como *Large / Small Caps* e *Growth / Value stocks*.⁴

Agora, é necessário ressaltar que a análise no paper acima foi acerca de retorno de mutual fundos internacionais, que não são o mesmo produto que os fundos multimercados brasileiros que pretendemos replicar. Dito isso, os resultados obtidos ainda servem como incentivo e validação para a nossa proposta. Nesse prisma, há ainda trabalhos como o de Andrew Lo que oferecem abordagens ainda mais parecidas com a nossa.

Inspirado pela falta de transparência e *accountability* dos hedges funds americanos, os autores se propõem a criar “clones lineares” das carteiras usando combinações de 6 fatores de risco distintos. Dessa maneira, seria criado veículos passivos de baixo custo, note que eles não se comprometem em manter uma estratégia dinâmica de portfólio como vamos propor mais adiante.

Em (Lo[†], 2006), os autores visam exercícios de replicação de portfólios bastante parecidos com o nosso. Com esse objetivo, eles usam 5 ativos explicativos, são eles: USD (índice de dólar, CREDIT (Lehman Corporate AA Intermediate Bond Index), CREDIT (spread entre e índice BAA da Lehman com o índice Treasury da Lehman), S&P500 e CMDTY (Goldman Sachs Commodity Index).

Assim, temos:

$$R_{it} = \beta_1 SP500_t + \beta_2 BOND_t + \beta_3 USD_t + \beta_4 CREDIT_t + \beta_5 CMDTY_t + \varepsilon_{it}$$

$$t = 1, \dots, T$$

$$\text{s.a. } 1 = \beta_1 + \dots + \beta_5 .$$

Considerando a restrição onde a soma de todos os coeficientes de ser 1 e a retirada do intercepto, note que os autores conseguem inferir as posições tomadas a partir de uma simples

⁴ **FTSE Russel Styles vs. Factors** What they are, how they're similar/different and how they fit within portfolios –https://content.ftserussell.com/sites/default/files/research/styles_vs_factors-_insights_final.pdf

regressão. Além disso, eles aceitam coeficientes negativos para capturar posições short, restrição proposta por Sharpe (1992) e abordado com dito anteriormente em (FRANS A. DE ROON, 2000). Em seguida o retorno do clone é calculado e renormalizado a fim de ter a mesma volatilidade que a carteira original, mais a diante será explicado como lidamos com a volatilidade em nosso algoritmo já que é um pouco diferente.

Agora, é necessário levar em consideração a limitação que os autores tiveram de dados, eles usaram os dados mensais de retorno da base TASS⁵, por isso qualquer análise feita seria considerando longos espaços de tempo. Dessa maneira, acredito que temos uma grande vantagem metodológica dado que temos retornos diários com pouquíssimos dias de atraso fornecidos pela CVM.

Assim, eles propõem duas metodologias, na primeira a regressão usa toda a base de dados para estimar as carteiras clones, porém como observado pelos autores, há um gritante *look-ahead bias* nessa comparação, algo que também abordamos em nosso exercício. Em seguida é exposto a alternativa de usar uma *rolling-window* de 24 meses para obter os clones, de novo há de se frisar que isso gera um enorme atraso em relação ao mercado, problema que também contornamos com nosso método.

Por fim, antes de analisarmos o resultado há ainda um grande *look-ahead bias* aparentemente ignorado pelos autores, apesar de usar um *rolling-window*, não é considerado que mesmo dentro desse período há viés. No sentido que a posição tomada para o primeiro mês foi baseada em retornos do 24º mês a frente. Essa pequena observação não pode ser menosprezada especialmente se estamos propondo uma estratégia dinâmica de alocação de recurso. Apesar disso, podemos observar que os resultados obtidos foram bastantes surpreendentes especialmente quando analisamos os clones com pesos fixo.

Abaixo temos o gráfico dos índices Sharpe para os dois tipos de clones e as carteiras originais. Como dito acima, os resultados são excepcionais, podemos ver que os fundos de peso fixos tiveram retornos bem comparáveis e as vezes superiores às carteiras originais e os fundos de *rolling-window* também tiveram bons resultados, porém a performance não foi tão acentuada como nos pesos fixos.

⁵ TASS Hedge Fund Database - (Lož, 2006) <https://library.princeton.edu/resource/14585>

Um possível exemplo de como o *look-ahead bias* do clone de peso fixo pode inflar a performance poderia ser a diminuição do *Equity Risk Premium* observado nas últimas décadas⁶. Considerando que as séries de retorno iniciam em 1986, e terminam somente em 2004, a informação de contida nos portfólios modernos podem aumentar os retornos caso fossem absorvidas desde o início do exercício. Dessa maneira, se hoje ERP é reduzido e os portfólios se expõem mais a esse fator de risco, um portfólio fixo com essa informação aproveitaria todo esse movimento enquanto os fundos contemporâneos no tempo não aproveitaram dessa informação.

⁶ Capital Perspectives: Shrinking equity risk premium -<https://journalrecord.com/2020/02/11/capital-perspectives-shrinking-equity-risk-premium/>

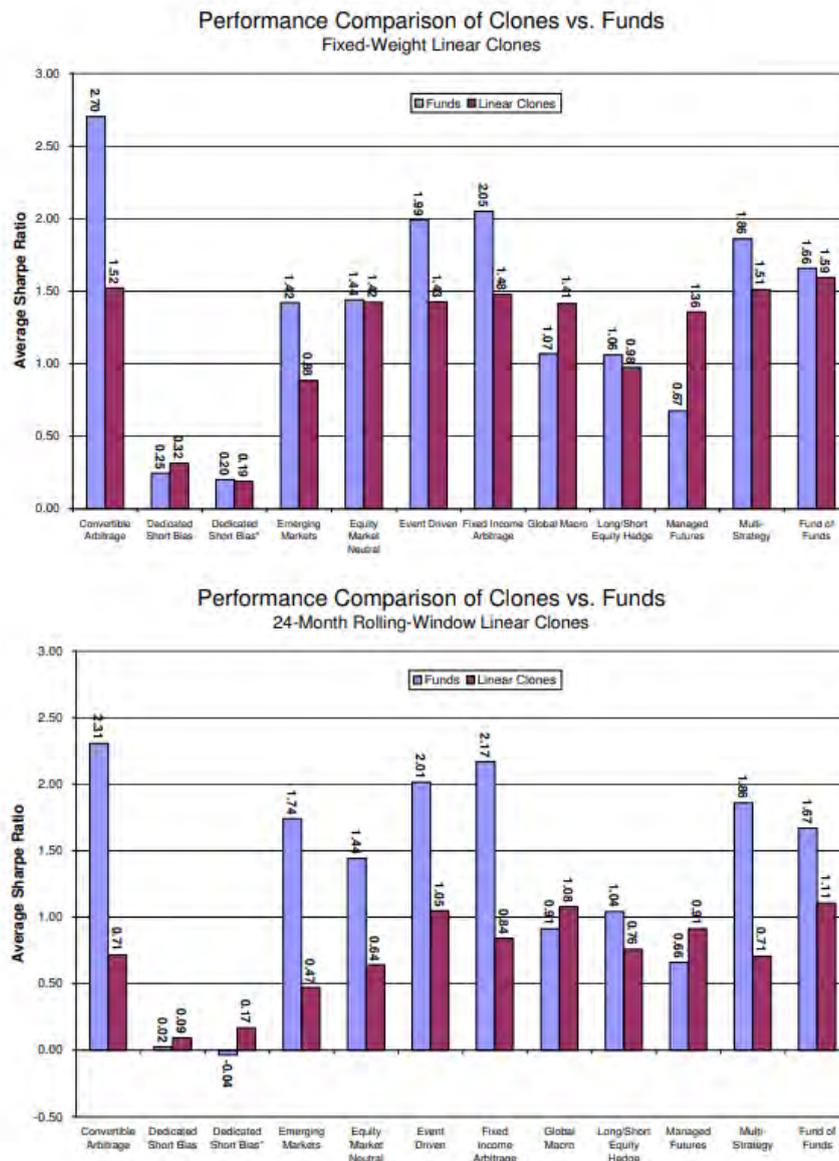


Figura 3-1 Índices Sharpe para os fundos clones e originais do estudo de (Loz, 2006)

Por fim, é útil comentar como há literatura que oferece outras maneiras de replicar retornos de diferentes fundos. Como referenciado anteriormente, o trabalho de (Young, 2007) analisa uma estratégia de “*performance-mimicking*”, onde um gestor conseguiria replicar o retorno de um gestor habilidoso sem precisar criar as próprias estratégias de investimento. Em suma o gestor ruim ficaria exposto a derivativos com grandes riscos de calda de maneira a se aproximar somente à performance do bom gestor. Note que nesse caso, o gestor ruim não tenta replicar a estratégia, mas somente a performance visando tomar proveito da falta de informação dos clientes. Apesar de interessante, esse paper serve mais como uma denúncia de como a atual estrutura de remuneração, por exemplo o sistema (2%, 20%), é facilmente corrompível.

Nesse prisma há outros trabalhos que propõem formas de replicar retornos usando contratos futuros como no caso de (Kat and Palaro, 2005;2006ab) ou como no case de (Bertsimas, 2001) onde são propostas estratégias de trading chamadas de “*epsilon-arbitrage*”. Embora exista essas alternativas, elas consistem em métodos mais teóricos e pouco aplicável para investidores comuns, além disso, elas não se preocupam necessariamente em inferir e replicar os portfólios individuais e por isso não serviram como base para a metodologia apresentada aqui.

4 - Dados

O modelo proposto teria como input dois tipos de dados, o retorno diário de ativos financeiros e o retorno diário dos fundos de investimento. Temos os retornos dos ativos como as variáveis explicativas dos retornos dos fundos. Nesse capítulo, será brevemente exposto como foi o processo de obtenção e formatação desses dados e será analisado também algumas de suas características gerais.

4.1 - Retorno dos Ativos

Com o intuito de explicar os retornos dos fundos, foi necessário buscar um número limitado de ativos que representariam a infinidade de posições possíveis que os fundos multimercado podem tomar. Considerando que as principais vertentes de um fundo multimercado são câmbio, juros e bolsa, foram escolhidos 7 ativos que poderiam se representar essas posições, são eles: USDBRL (Câmbio dólar/real), DXY (Índice de Câmbio do dólar), BZAA2Y (Juro nominal 2 anos), BZAA10Y (Juro nominal 10 anos), USSW10 (Juro americano 10 anos), IBOV (Bolsa Brasileira), SPX (Bolsa Americana). A fonte usada para todos esses ativos foi a Bloomberg, porém há outras formas de obter esse dado caso necessário.

Iniciando o tratamento dos dados, para as séries de Bolsa e câmbio tivemos que pegar a variação diária do preço já que o dado original estava em nível de preço. Para as séries de juros, americano e brasileiro, temos de início as séries das taxas dos vértices móveis, por isso não seria correto calcular a variação percentual dado que estamos lidando com uma taxa e não um nível, não obstante, também não seria ideal usar a taxa em si por falta de estacionariedade. Frente a esse problema, usamos a primeira diferença da taxa o que as deixariam comparáveis às séries de retorno dos fundos. Optamos por essa abordagem já que a maioria dos gestores calculam suas exposições em relação ao PL dos fundos, de forma que uma variação da taxa em pontos percentuais geraria uma variação proporcional no retorno da carteira.

Informações Descritivas dos Ativos

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
BZAA2Y	2,853	-0.0001	0.093	-0.664	-0.040	0.035	1.192
BZAA10Y	2,853	-0.0004	0.074	-0.774	-0.032	0.030	1.040
USSW10	2,853	-0.001	0.050	-0.244	-0.032	0.030	0.260
IBOV	2,853	0.023	1.571	-14.800	-0.800	0.900	13.900
DXY	2,853	0.009	0.443	-2.370	-0.250	0.260	2.050
USDBRL	2,853	0.042	1.017	-5.791	-0.546	0.631	7.371
SPX	2,853	0.056	1.083	-11.980	-0.350	0.540	9.380

Tabela 3 Informações descritivas dos ativos usados

Acima temos algumas informações descritivas das séries de retorno. Note que há uma diferença considerável na volatilidade dos ativos. Mais à frente, no Capítulo 5, veremos que será necessário considerar essas diferenças na hora de ponderar o portfólio para manter a contribuição de cada ativo equivalente.

4.2 - Retorno dos fundos

Enquanto a busca das séries dos ativos foi bem facilitada pela infraestrutura da Bloomberg, organizar os dados dos fundos foi um pouco mais trabalhoso. A fonte principal dos dados foram os dados abertos da CVM⁷ que oferecem as declarações de retorno dos fundos brasileiros desde 2005, porém para esse projeto só foram obtidos os dados desde 2010.

De imediato, a CVM fornece links para o download dos retornos diários até 2017, para obter os próximos 7 anos foi necessário baixar individualmente os arquivos mensais, *des-zipar* e compilar todos numa base agregada. É necessário informa que, mesmo com o auxílio de linguagens de programação como o R, lidar com bases desse tamanho já demanda bastante das máquinas. Ao todo, foram compilados cerca de 150 arquivos em um objeto de quase 40 milhões de linhas com informações de retorno de 35 mil fundos diferentes, por isso fica evidente a urgência de filtrar, logo de início, somente as informações necessárias.

Além dos dados da CVM, utilizamos também os dados descritivos da Quantum⁸ para parear informações relevantes aos dados da CVM. Com esses dados, conseguimos associar os

⁷ **Dados Abertos CVM** - http://dados.cvm.gov.br/dados/FI/DOC/INF_DIARIO/DADOS/

⁸ **Dados descritivos Quantum** - <https://quantumfinance.com.br/bases/>

diferentes CNPJs com os seus respectivos nomes e outras informações da natureza dos fundos, como taxas cobradas e se ele é um *master* ou *feeder*. Após unir as bases podemos iniciar o processo de seleção de fundos.

Já que queremos, ao fim, somente os fundos macro mais relevantes prosseguimos com uma série de etapas com o intuito de filtrar todos os fundos irrelevantes para nossa análise. Devido à alta densidade da base, quase todos os filtros foram feitos em R de maneira automatizada reduzindo o tempo e vieses de seleção. Primeiro, selecionamos os 500 fundos os maiores patrimônio líquidos, em seguida, usamos os dados da Quantum para discriminar os fundos *Masters* dos demais para que possamos analisar todos os retornos livres de taxas e em seguida filtramos para obter somente os fundos macro.

Por fim, percebemos que apesar de selecionar somente os fundos descritos como master, ainda havia alguns *feeders* mal definidos na base. Com o intuito de encontrar e remover esses fundos de maneira automatizada, calculamos uma matriz de correlação dos últimos 300 dias de retornos com todos os fundos restantes. Com a matriz, conseguimos encontrar cotas correlacionadas quase perfeitamente, nesses casos presumimos que os fundos teriam uma relação de *feeder* e *master* e eliminamos os repetidos.

Abaixo temos o resultado de uma das iterações desse processo. Observe que usamos um método de correlograma que agrupou os resultados correlacionados no canto superior esquerdo e rente a diagonal principal. Em seguida, rodamos um filtro que remove todos esses fundos espelhados da base. Note que os nomes dos fundos foram substituídos por nomes genéricos.

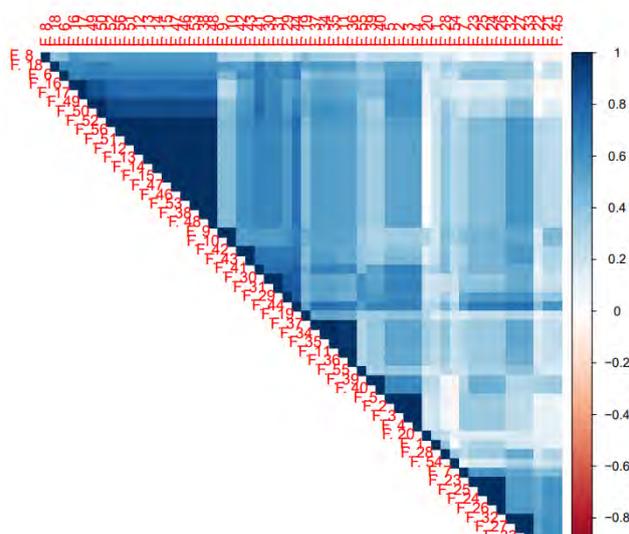


Figura 4-2-1 Correlograma dos retornos dos fundos para identificar feeders

Com a retirada dos últimos *feeders* da base, restam 27 fundos macro dentre eles as casas mais famosas do mercado. Agora com a base devidamente limpa podemos calibrar e testar o método.

No Capítulo 6, veremos que apesar de usarmos os *Masters* para a inferência das posições, será interessante comparar os retornos obtidos com a rentabilidade oferecida pelos respectivos *Feeders*. Com isso em mente, selecionamos a partir dos Masters escolhidos os *Feeders* abaixo

	Fundo	Master CNPJ	Feeder CNPJ
1	A1	36.181.795	35.185.577
2	ABSOLUTE VERTEX	21.470.947	21.470.989
3	ACE CAPITAL	34.687.549	34.774.662
4	BAHIA MARAU RV	17.056.278	17.087.932
5	CAPSTONE MACRO	35.803.288	35.726.908
6	DAHLIA	30.317.436	30.317.454
7	GARDE DARTAGNAN	18.981.231	19.212.817
8	GAUSS	20.216.114	24.592.505
9	GENOA	35.828.652	35.828.684
10	GAVEA MACRO I	08.875.020	8.893.082
11	IBIUNA HEDGE ST	15.487.918	12.154.412
12	JGP STRATEGY	97.548.167	13.812.165
13	KAPITALO II	12.083.748	12.105.940
14	KAPITALO K10 MASTER	29.726.117	29.726.133
15	KINEA CHRONOS	21.624.757	21.624.757
16	LEGACY CSHG FIC	29.236.572	30.329.404
17	OCCAM	17.248.340	35.271.414
18	OPPORTUNITY TOTAL	09.720.828	5.448.587
19	SAFRA GALILEO	27.249.022	10.347.249
20	SANTANDER MACRO FIC	11.714.807	11.714.716
21	SPX NIMITZ	12.798.221	12.831.360
22	VERDE	07.455.507	22.187.946
23	VERDE AM SCENA	26.744.808	8.507.696
24	VINTAGE ALPHA MACRO	19.941.720	37.319.496
25	VISTA MACRO	31.301.673	21.646.715

Tabela 4 Nome dos Fundos Master, seu CNPJ e o CNPJ do Feeders utilizados em seu lugar

5 – Metodologia

5.1 – Inferência dos Portfólios

A fim de montar uma carteira que sintetiza a opinião média dos investidores precisamos antes identificar quais são essas opiniões individualmente. Com esse fim, pegamos os retornos diários dos fundos de investimento e calculamos uma regressão múltipla e sem constante móvel com uma cesta de ativos que representariam as principais exposições dos fundos multimercado. Escolhemos o Índice IBOV, a taxa de câmbio USDBRL, taxa de juro brasileiro nominal móvel de 2 anos BZAA2Y e a taxa de juro brasileiro real móvel de 10 anos BZAA2Y, Índice DXY (dólar) e o índice SPX de bolsa americana.

Devido à enorme diferença de volatilidade entre os ativos, para que o coeficiente estimado seja informativo, foi necessário normalizar todos os retornos. Fazendo isto, podemos interpretar os coeficientes como uma métrica de exposição ao ativo, o que fundamentará a composição da carteira mais adiante.

Calculamos os coeficientes usando uma janela móvel já que miramos inferir as posições dos gestores de maneira dinâmica. De maneira que seja possível capturar mudanças de posicionamento marginais. Perceba que o tamanho da janela deve levar em consideração dois importantes fatores (1) quanto maior a janela, mais observações teremos a disposição e por isso teremos resultados com menor variância e (2) quanto maior a janela maior a probabilidade de o gestor mudar a composição de seu portfólio, o que diminuiria a precisão do coeficiente estimado. Assim, o tamanho da janela precisa ser um equilíbrio entre esses dois fatores.

Em seguida, compilamos os resultados para cada fundo e podemos visualizá-los da seguinte maneira.

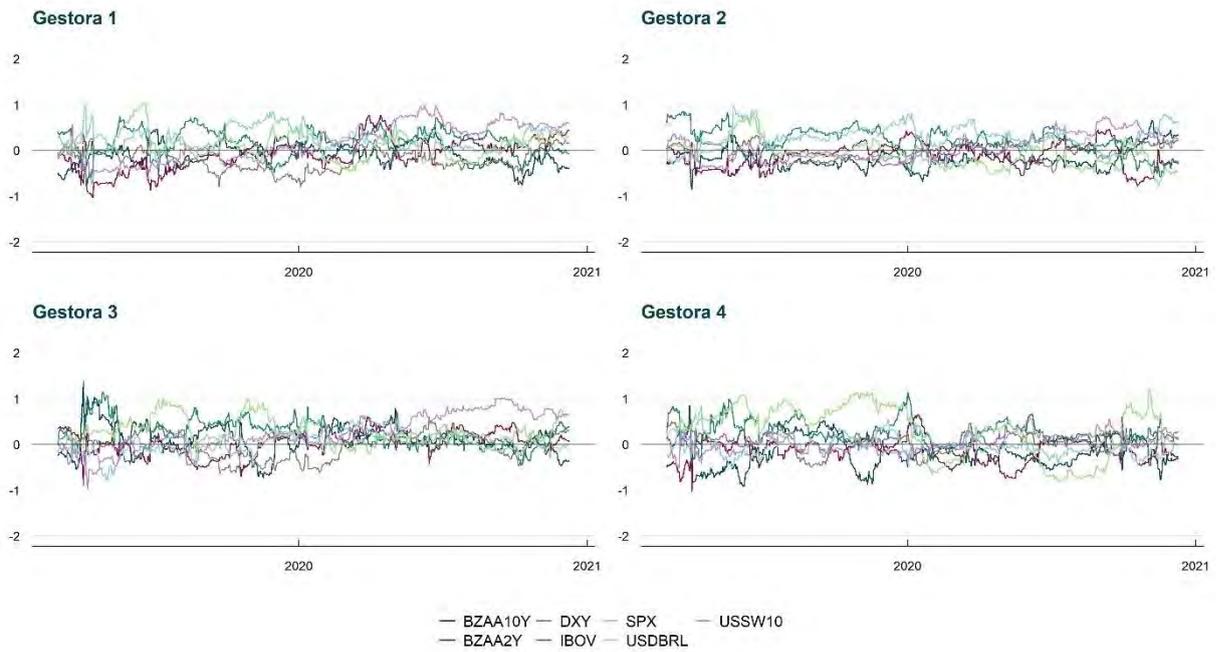


Figura 5-1-1 Coeficientes estimados para gestoras anônimas

Além dos coeficientes estimados, podemos analisar também o R^2 obtido para cada iteração do processo. Dessa maneira, conseguimos compreender quanto da variância total do retorno do fundo pode ser explicado por meio desses ativos. Abaixo, temos a visualização dos R^2 no tempo para cada fundo, além disso, está destacado em azul o R^2 mediano em cada data. Evidencia-se que a variância da qualidade das regressões é bem considerável, contudo, a mediana dos R^2 obtidos é alta, se mantendo entre 0.5 e 0.75 durante a maioria das iterações.

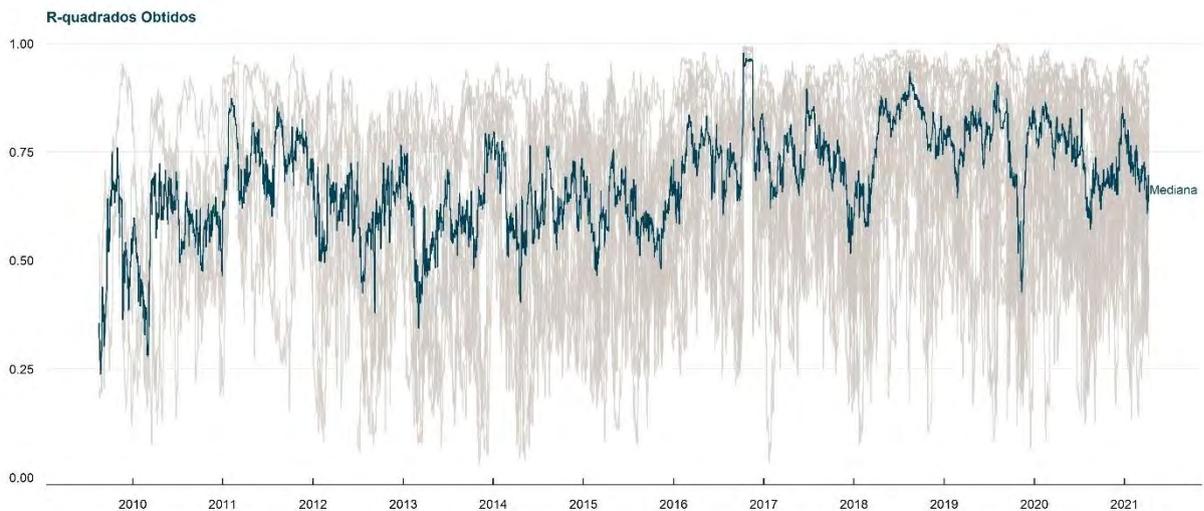


Figura 5-1-2 R-quadrados obtidos para todas as regressões móveis, em azul está o valor mediano por dia

Após calcular os coeficientes móveis de todos os fundos, já podemos começar a agregá-los. Podemos sintetizar as posições de diferentes formas, com o intuito não só de mensurar qual o posicionamento médio do mercado, mas também responder perguntas mais específicas como veremos adiante. Com esse objetivo, podemos calcular, primeiro, a mediana e média das exposições a fim de ter uma estimativa das posições sem nenhuma adição de viés. Abaixo temos os resultados.

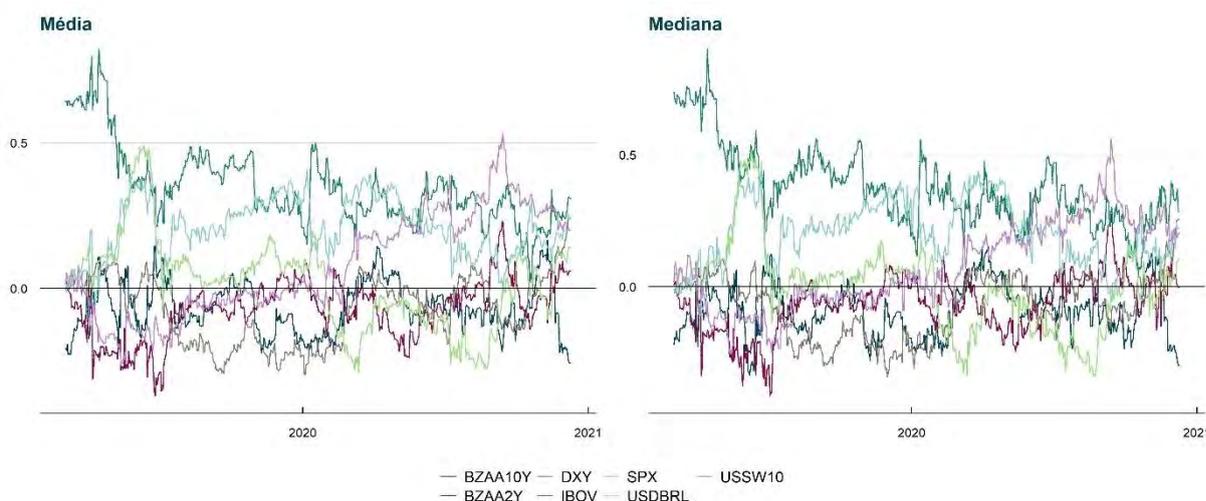


Figura 5-1-3 Coeficientes agregados - Média e Mediano

Podemos também agregar os dados de maneira ponderada pelo tamanho dos patrimônios líquidos dos fundos. Nesse sentido, as gestoras terão um peso proporcional à quantidade de dinheiro sob gestão. Agregar dessa forma pode ser informativo se quisermos inferir onde está alocado o capital do mercado, porém teremos um indicador viesado em relação à opinião dos gestores. Além disso, quando formos construir as carteiras sintéticas, usar as posições das maiores casas pode ser interessante, já que elas apresentam um histórico melhor de geração de *alpha* (Martelanc, 2015).

Uma última agregação interessante seria compilar os fundos que apresentaram uma melhor performance. Sob a hipótese de que no mercado brasileiro há uma certa persistência de resultados, isto é, retornos passados podem ajudar a prever retornos futuros (Segalis, 2018), seria muito interessante selecionar somente as opiniões dos melhores gestores. Assim, calculamos um ranking móvel de 250 dias que determina os 5 melhores fundos da janela. À vista disso, evitamos qualquer *look-ahead bias* durante a seleção dos fundos. Com essa cesta de fundos, calculamos a média das exposições. Observe abaixo a visualização da média ponderada por PL e da média do Top 5.

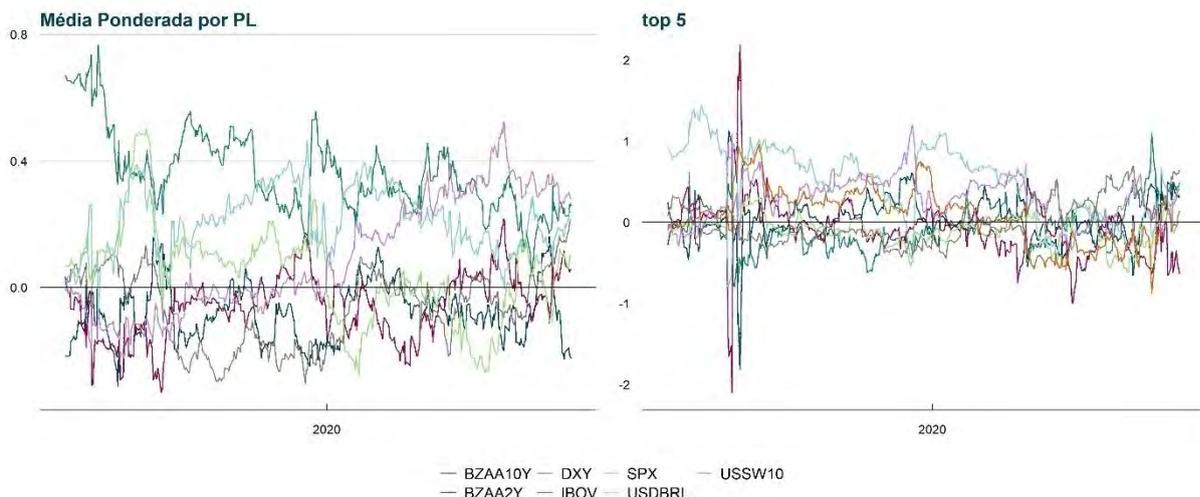


Figura 5-1-4 Coeficientes estimados para Média por PL e Top 5

5.2 – Construção das Carteiras Sintéticas

A partir dessa primeira análise dos coeficientes precisaríamos traduzir esses diferentes valores em uma série única que representaria o retorno da nossa carteira. Normalmente, devido ao fato de que ambas as séries estão em percentual, poderíamos interpretar que os coeficientes estimados nos indicam as magnitudes dos impactos no retorno do fundo, dado uma variação em 1 unidade de cada série explicativa. Contudo, lembramos que as variáveis foram normalizadas antes da estimação dos *betas*, por isso, devemos antes reajustar as variáveis explicativas. Assim fazendo, a volatilidade natural do ativo não influenciará o peso dele no portfólio final.

Considerando a grande diferença de volatilidade dos retornos dos ativos, usamos os 3 primeiros anos de dados (2010-2012) para calcular os diferentes desvios padrões das séries. Com isso, calculamos a volatilidade média dos ativos e a taxa de alavancagem necessária para deixar a contribuição de cada ativo equivalente em relação ao retorno final da carteira sintética. Nesse prisma, percebemos que ativos mais voláteis, como IBOV, teriam uma taxa de alavancagem menor que 1, enquanto os menos voláteis, como o USSW10, teriam uma taxa maior em compensação.

Por fim, temos, para todos os dias, a posição, *long* ou *short*, de cada ativo e suas series ajustadas pela volatilidade. Simulando um cenário onde não seria prático reajustar o portfólio sintético todo dia, fixamos as posições mensalmente e quinzenalmente e calculamos as series de retornos finais.

Podemos expressar o processo de maneira matricial.

Temos que:

$$P_{\lambda} = \begin{pmatrix} r_{11} & r_{12} & \dots & r_{1p} \\ r_{21} & r_{22} & \dots & r_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ r_{n1} & r_{n2} & \dots & r_{np} \end{pmatrix} * \begin{pmatrix} Al_{11} & 0 & \dots & 0 \\ 0 & Al_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & Al_{np} \end{pmatrix} * \begin{pmatrix} C_{11} & C_{12} & \dots & C_{1n} \\ C_{21} & C_{22} & \dots & C_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ C_{p1} & C_{p2} & \dots & C_{pn} \end{pmatrix}$$

$$P_{\lambda} = \begin{pmatrix} \sum_j^p r_{1j} * Al_j * C_{j1} \\ \sum_j^p r_{2j} * Al_j * C_{j2} \\ \vdots \\ \sum_j^p r_{nj} * Al_j * C_{jn} \end{pmatrix}$$

Onde:

P = vetor com a série final de retornos

λ = índice indicando método de coeficiente usado

n = índice indicando data

p = índice indicando ativo

r = retornos diários de cada ativo

Al = desvio padrão médio dos ativos (2010-2012) / desvio padrão do ativo (2010-2012)

C = coeficiente estimado para cada ativo na data especificada

5.3 – Controles para *Look-ahead bias*

Sempre devemos prestar atenção se nosso modelo está utilizando informações futuras na sua estimação. Se esse for o caso, você terá *look-ahead bias* e o resultado não será representativo a uma iteração na margem. Levando isso em consideração, em todas as etapas da metodologia tomamos o cuidado de incluir somente informações passadas nos cálculos. Isso foi considerado nas agregações por patrimônio líquido e do *Top 5* e no cálculo da volatilidade que ajusta os ativos. Por fim, os coeficientes estimados nas janelas móveis são atribuídos à última data da janela, dessa maneira evitamos novamente o uso de informações futuras.

6 – Resultados

6.1 – Teste da Metodologia

Observando os resultados preliminares apresentados no Capítulo 5, cabe a dúvida sobre a capacidade das regressões de capturarem precisamente quais ativos compõem as carteiras. Com isso em mente, construímos carteiras com graus diferentes de complexidade para testar a capacidade de inferência do método.

As carteiras testes são compostas por combinações lineares dos ativos selecionados, em alguns casos, mudamos a composição de uma mesma carteira para analisar a capacidade de capturar essas mudanças na margem. Vale ressaltar que, antes de serem combinados linearmente, os ativos foram normalizados de forma a compensar a diferença de volatilidade entre eles.

Abaixo temos as visualizações dos coeficientes no tempo como no Capítulo 5, veremos primeiro as carteiras mais simples compostas por somente dois ativos. Percebemos que apesar da limitação no número de observações e dificuldades oriundas da multicolinearidade dos ativos, o método teve bastante êxito em inferir as composições simples.

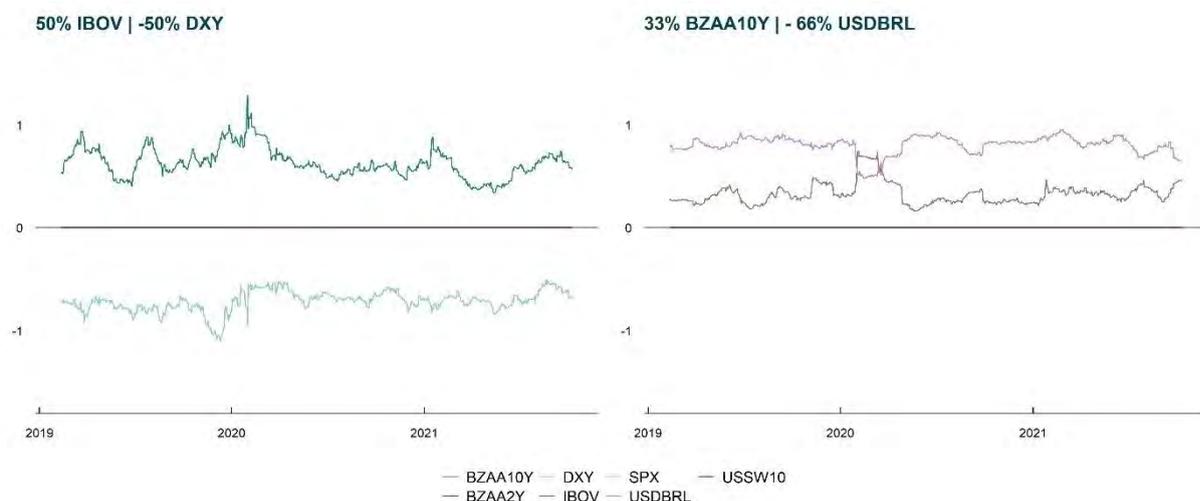


Figura 6-1-1 Resultado de teste com carteiras de 2 ativos

Conferindo as carteiras teste de 3 ativos, percebemos que a capacidade de inferência continua precisa. Dito isso, percebemos que a variância dos coeficientes estimados é ainda maior, isso se novamente pelo fato dos ativos não serem ortogonais entre si e a pequena janela de temporal comparado ao grande número de regressores, 40 observações e 7 ativos.

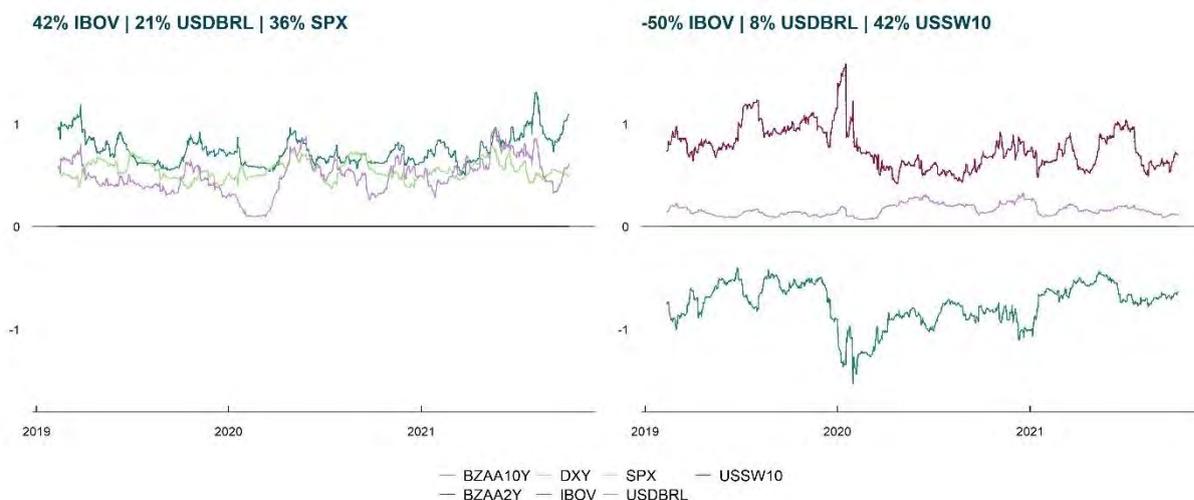


Figura 6-1-2 Resultado do teste com carteiras de 3 ativos

Por fim, analisamos uma carteira teste dinâmica que muda completamente sua composição em junho de 2020. Percebemos que durante o período de transição, quando a janela captura ambas as combinações de ativos, as estimações pioram de qualidade, com até alguns ativos inapropriados aparecendo, como no caso de USSW10 na primeira e USDBRL na segunda.

Outro ponto interessante é a grande variância dos estimadores da segunda metade da primeira carteira. Optamos por testar a capacidade do modelo de perceber estratégias de desinclinação, onde o gesto tomaria posições contrárias ao longo da curva de juros, porém a alta correlação entre as séries de juros proporcionam uma maior dificuldade para o MQO. Com o correlograma, percebemos que a correlação entre as taxas de juros é consideravelmente maior que entre IBOV e USDBRL, 0.73 e -0.46 respectivamente, o que pode explicar a diferença de precisão.

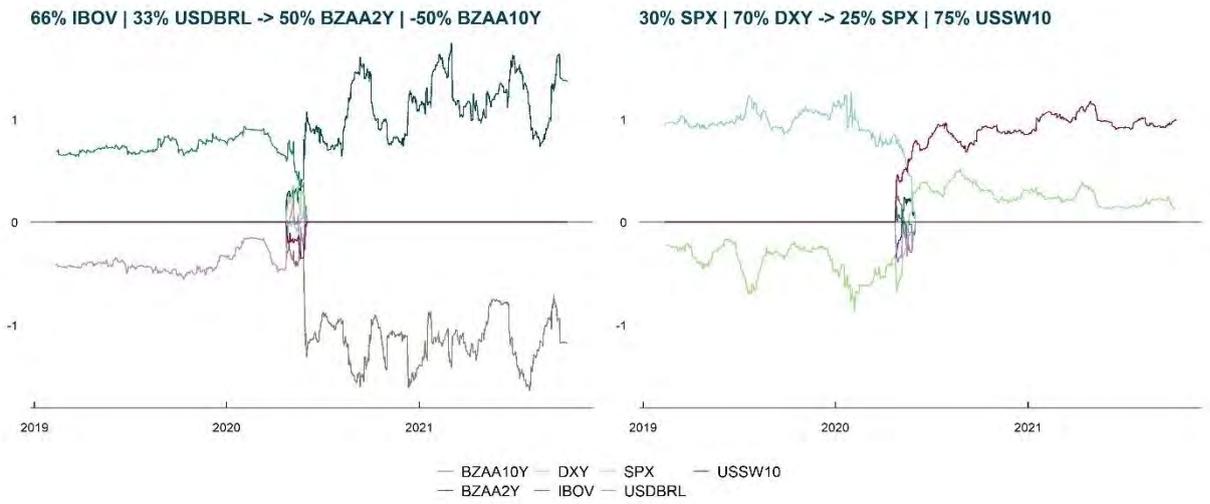


Figura 6-1-3 Resultado do teste com carteiras dinâmicas de 2 ativos



Figura 6-1-4 Correlograma para evidências diferença de correlação entre ativos

6.2 – Retorno por Ativo

Antes de calcularmos o somatório final de P_λ , é possível visualizar o desempenho de cada uma das posições individualmente. Temos abaixo os resultados das carteiras Média Quinzenal e Top5 Mensal.

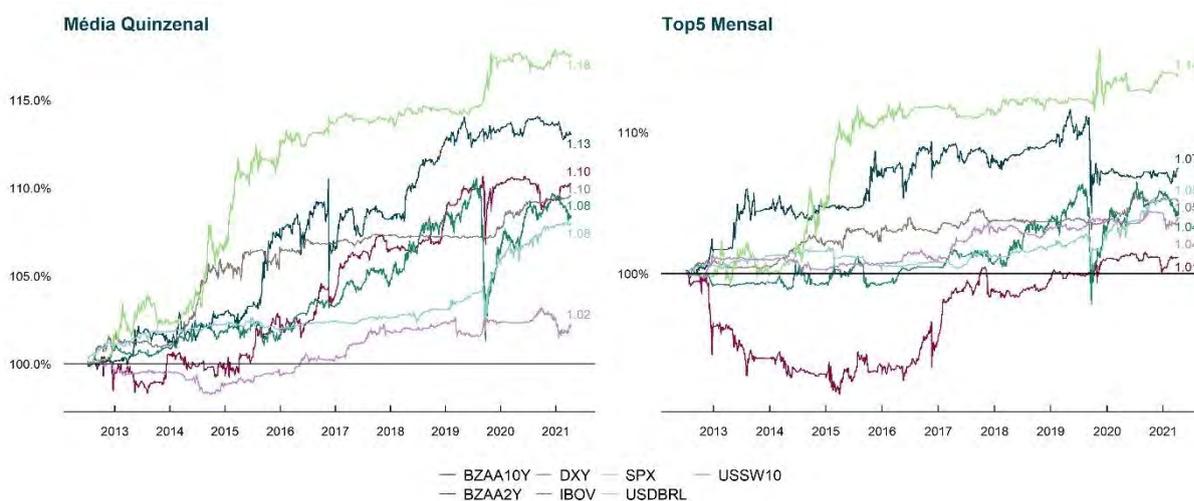


Figura 6-2-1 Retorno acumulado por ativo

Note que, apesar dos cuidados para ajustar a volatilidade dos ativos, fica evidente a dificuldade de prever qual será o comportamento futuro das séries. Assim, dificilmente teremos uma equidade de contribuição para a volatilidade total entre os ativos. Sobre a performance, é interessante notar que, com a exceção dos ativos USDBRL e BZAA10Y, as posições tomadas foram bem diferentes, contribuindo para uma performance final distinta

6.3 – Retorno das Carteiras

Agregando os retornos demonstrados acima, obtemos as séries de retorno P_λ demonstradas anteriormente. Com os retornos diários dos fundos e das carteiras sintéticas, podemos calcular o rendimento acumulado dos ativos do período de janeiro 2013 até outubro de 2021. Calculamos a performance de todas as agregações propostas, com as carteiras sendo recalculadas quinzenalmente e mensalmente. Para fins de comparação calculamos também a média dos retornos diários dos fundos (estratégia *equal weighted* de alocação) e a taxa CDI acumulada no período (preto pontilhado).

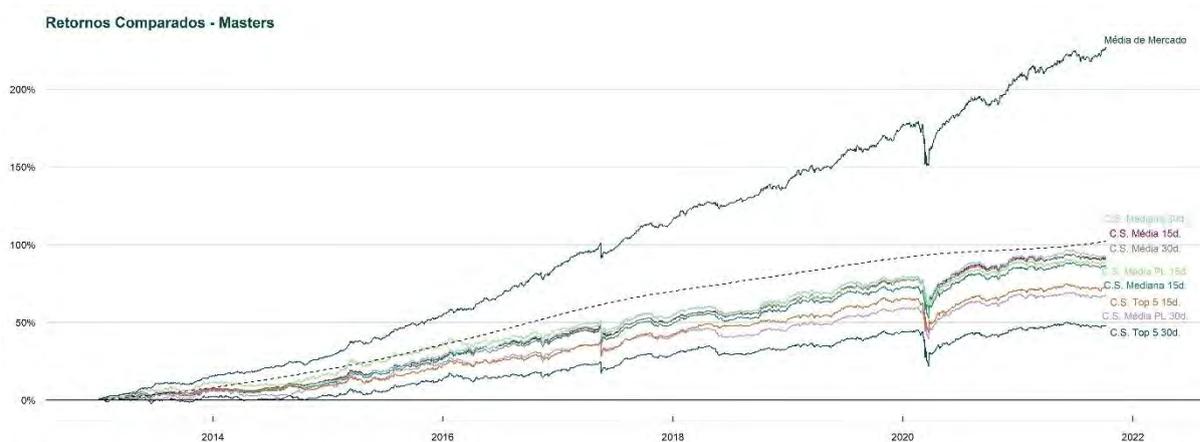


Figura 6-3-1 Retorno Acumulado das carteiras sintéticas com Média dos Masters

Note que ainda estamos comparando os resultados com os retornos dos fundos Master, o que implica que ainda não foram descontadas todas as taxas administrativas e de performance. Considerando a natureza composta dos retornos e a grande *outperformance* em relação ao benchmark mais comum o CDI⁹, podemos supor que a performance do algoritmo será mais competitiva em relação aos *feeders*.

⁹ Definição da taxa CDI - <https://www.infomoney.com.br/guias/cdi/>

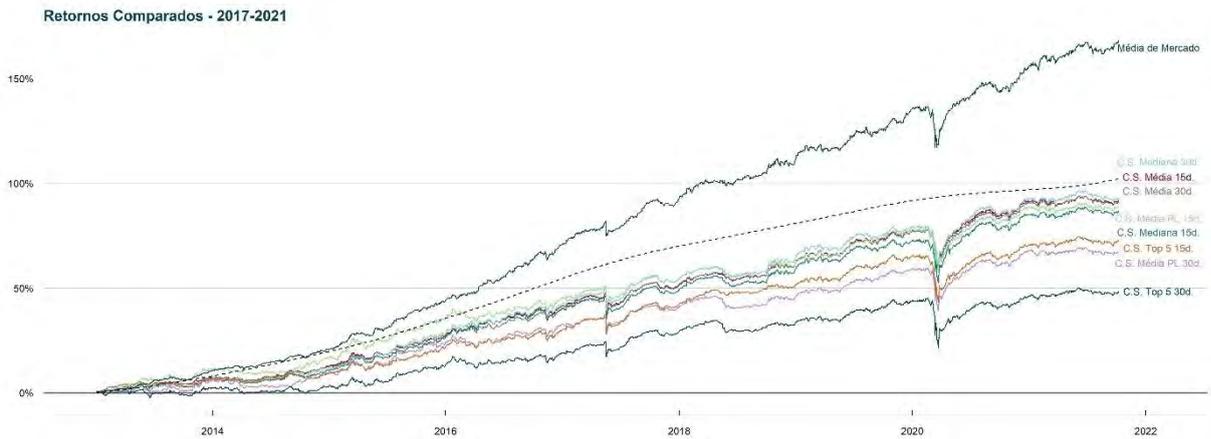


Figura 6-3-2 Retorno Acumulado das carteiras sintéticas com Média dos Masters

Apesar da diferença de performance ser bastante acentuada entre os *Masters* e os *feeders*, nenhuma estratégia de agregação λ conseguiu *outperformar* a performance dos fundos pagos no acumulado do período 2013 - 2021. Dito isso, os resultados se aproximaram consideravelmente. Embora que a rentabilidade acumulada estar num nível inferior, é necessário analisar o retorno ajustado ao risco, caso contrário, conseguiríamos controlar a performance simplesmente aumentando o tamanho de todas as posições. Visando essa métrica ajustada, calcularemos o Índice de Sharpe¹⁰ dos retornos de mercado e das 3 melhores estratégias (Mediana mensal, Média quinzenal e Média Mensal).

Temos o Índice Sharpe definido como:

$$Sharpe = \frac{R_p - R_l}{\sigma}$$

Onde:

R_p = Retorno do portfólio

R_l = Retorno livre de risco (CDI)

σ = Desvio Padrão do retorno

¹⁰ The Sharpe Ratio - <https://jpm.pm-research.com/content/21/1/49>

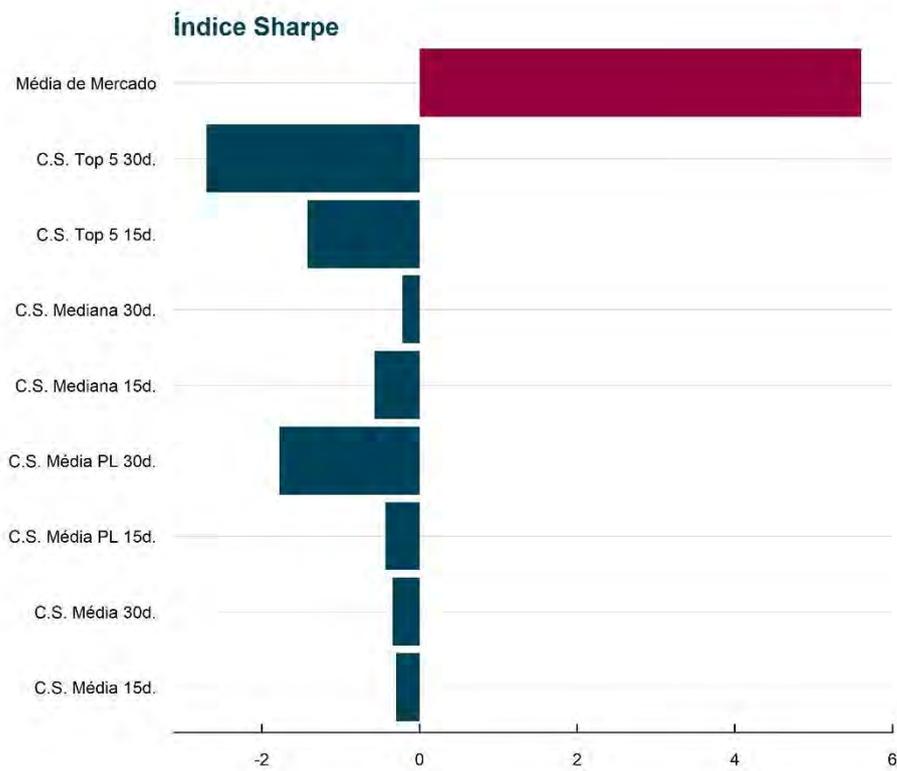


Figura 6-3-3 Bar chart de Índices Sharpe

Analisando os resultados obtidos, percebemos que os fundos multimercado apresentaram Índices Sharpe nos valores de 5.6, média e mediana respectivamente. Enquanto isso, as carteiras sintéticas apresentaram Índices negativos. Isso ocorre devido ao fato de que o retorno apresentado, foi inferior ao retorno acumulado do CDI, dessa maneira, independente da volatilidade realizada, o resultado será dificilmente informativo.

Já que ainda precisamos mensurar a rentabilidade da carteira penalizada pelo risco, propomos uma análise de um indicador mais simples. Temos abaixo o Retorno Ajustado, que seria a rentabilidade total dividida pelo desvio padrão observado.

$$\text{Retorno Ajustado} = \frac{R_p}{\sigma}$$

Onde:

R_p = Retorno do portfólio

σ = Desvio Padrão do portfólio

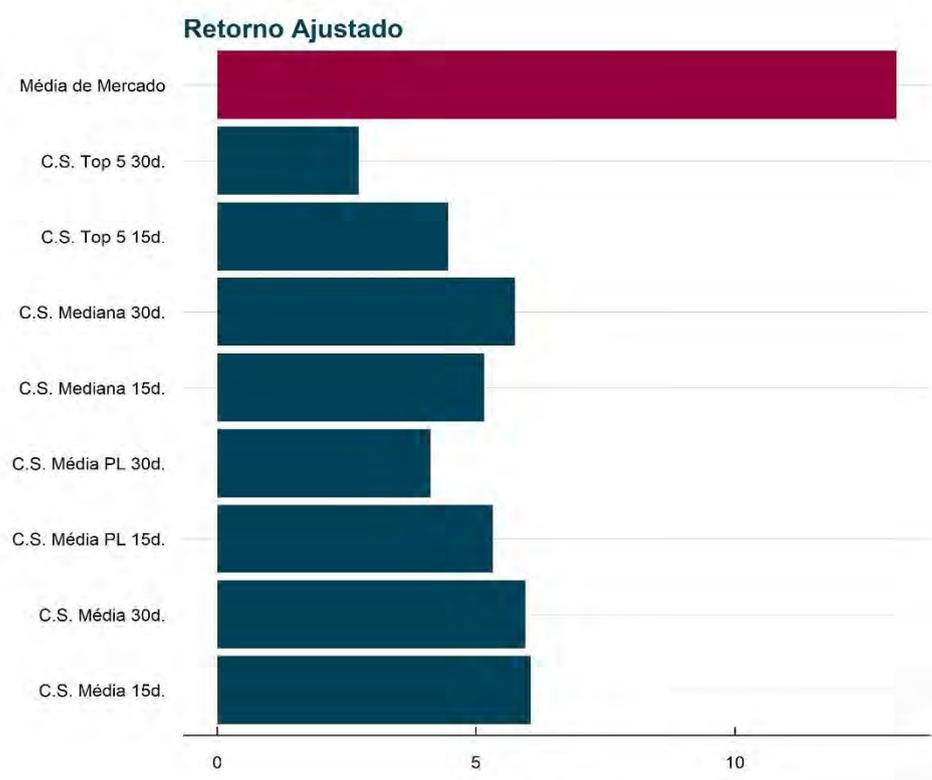


Figura 6-3-4 Bar chart de retornos ajustados

Agora, é possível interpretar os resultados de maneira mais clara. Vemos que o indicador de Retorno Ajustado evidencia valores de 13.1 para os fundos multimercado, média e mediana respectivamente. Enquanto isso, as carteira sintéticas apresentam valores entre 2.7 (Top 5 mensal) e 6.05 (Média quinzenal). As 3 melhores carteiras foram: Média quinzenal: 6.05, Média mensal: 5.94 e Média quinzenal: 5.15.

Após ajustar os retornos pela volatilidade percebemos que um estratégia de alocação equal weighed entre os fundos multimercados, ofereceu retornos maiores durante os período observado.

Dito isso, devido ao grande duração da simulação e da natureza exponencial de retornos no tempo, ganhos pequenos no início da série se acumulariam em diferenças grande ao final do período observado. Esse efeito exageraria a diferença entre os fundos e as carteiras sintéticas, dito isso, podemos analisar a performance anual dos portfólios destacados acima.

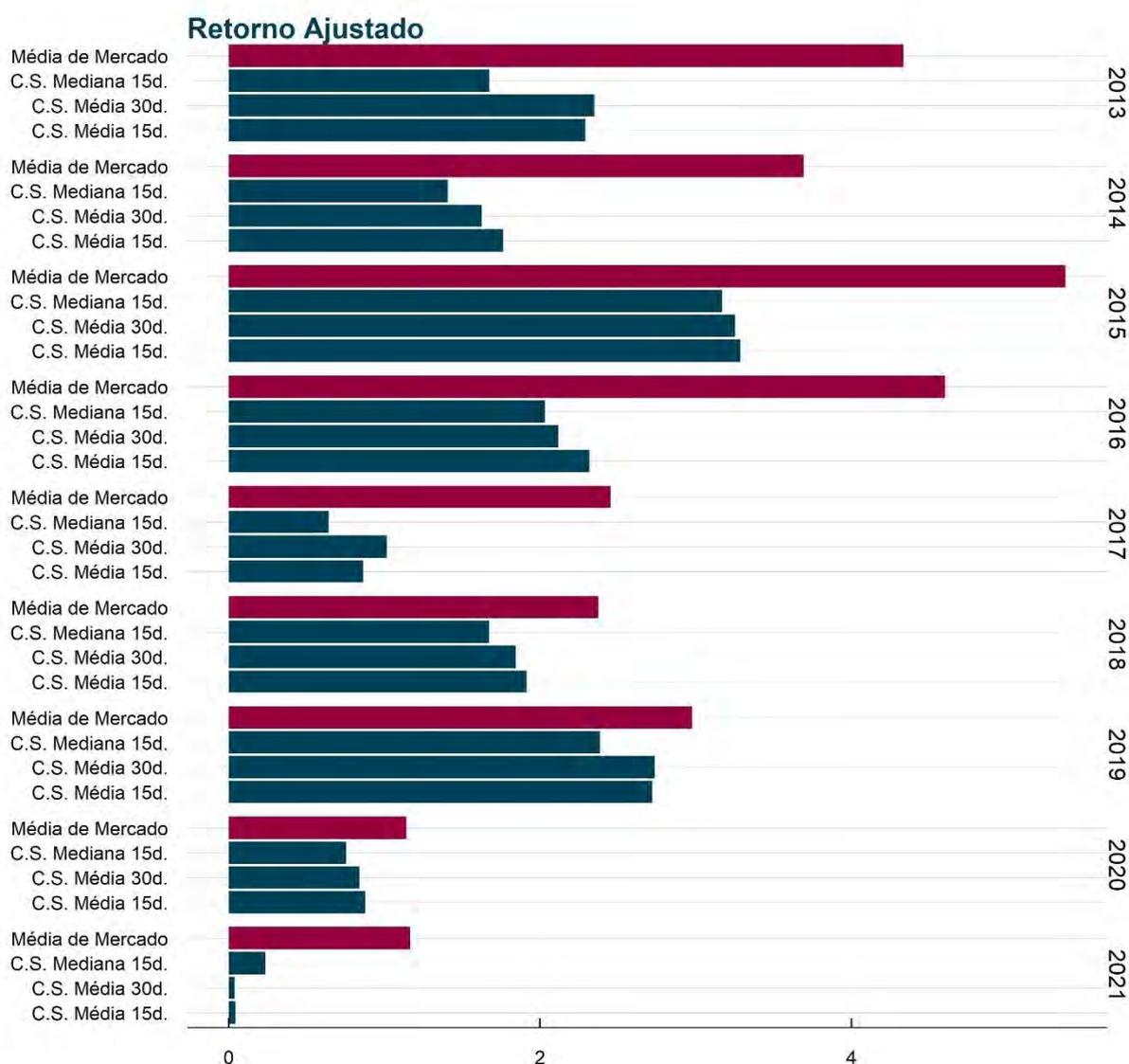


Figura 6-3-5 Bar chart de retornos ajustados por ano

Analisando os retornos ajustados por ano, podemos perceber certos comportamentos interessantes. Destacamos primeiro a boa performance das carteiras sintéticas em 2020, percebe-se que a Média quinzenal apresentou um retorno ajustado superior à Mediana dos Fundos. Esse resultado precisa ser considerado com atenção, pois o algoritmo demonstrou uma surpreendente resiliência durante um momento de extrema volatilidade. Isso é contra intuitivo devido ao desanimador histórico de estratégias quantitativas durante momentos de alta volatilidade (Shah, 2021).

Além disso, podemos observar que conforme o ano é mais recente, menor parece ser a outperformance das fundos. Analisando a base de dados percebemos que pode haver um certo “survivorship bias” para os retornos com maior histórico. Assim, poucos fundos que performaram surpreendentemente bem no passado chegaram até hoje para serem selecionados.

Com isso, os retorno mais antigos podem ser viesados e representar a performance de casas consagradas como Verde, Opportunity e SPX que por sua vez geravam *Alpha* de maneira que nosso algoritmo não conseguiu capturar. Abaixo temos a visualização da quantidade de fundos por ano.

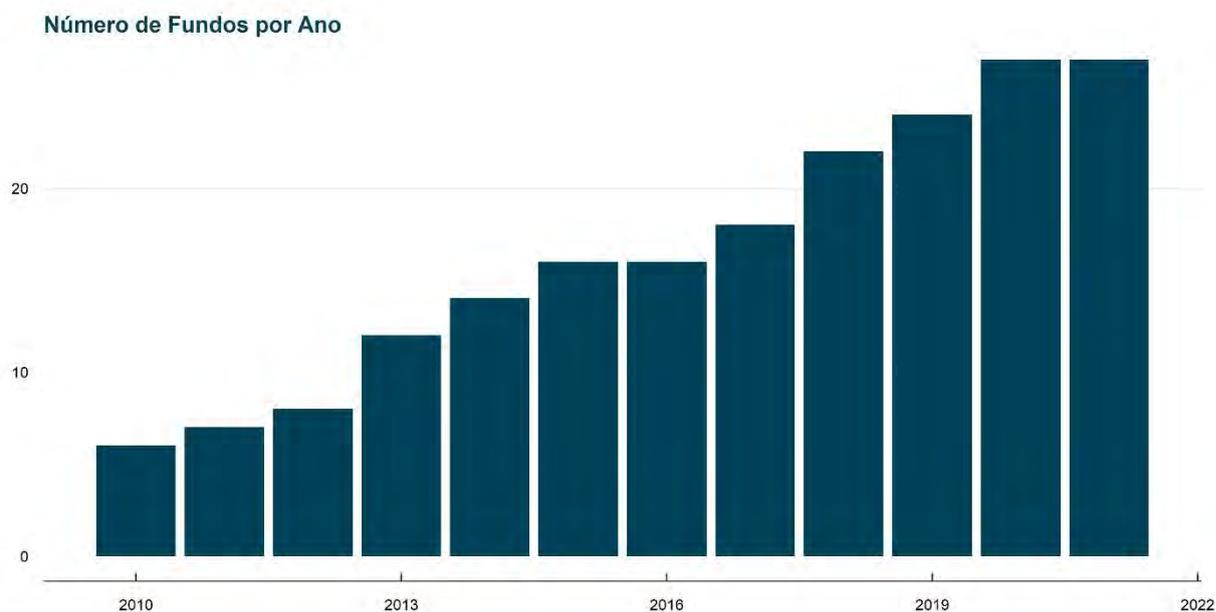


Figura 6-3-6 Bar chart de número de fundos por ano

Outra hipótese válida que explicaria a baixa performance nos primeiros anos é a maior variância da média dos coeficientes devido ao número reduzido de coeficientes estimados. Com uma grande quantidade de fundos, podemos assumir que a perda de precisão devido à variância dos coeficientes será mitigada ao calcular a média. Nesse prisma, podemos concluir que o baixo número de fundos pode piorar a qualidade das posições tomadas nas carteiras sintéticas.

Corroborando com essa tese, temos a decepcionante performance da estratégia Top 5. Percebemos que, apesa delas estarem sempre acompanhando os melhores fundos, o retorno final foi dos piores apresentados. Podemos observar também que os coeficientes apresentados na Figura 5-1-5 têm uma variância muito maior que as outras estratégias apresentadas.

Para que essa hipótese seja válida, teríamos que supor que as principais gestoras se posicionam no mercado de maneira similar. Dessa forma, a média das estimativas convergia a um valor mais próximo da média verdadeira das posições, apesar da variância das estimativas individuais. Embora que o retorno dos diferentes fundos, sejam correlacionados, supor que as gestoras se posicionam de maneira tão similar pode não ser razoável, por isso, seriam

necessários mais testes para averiguar ao certo qual a razão da baixa performance nos primeiros anos.

Outro fator importantíssimo a considerar quando analisamos os resultados é a baixíssima volatilidade da série de retorno média dos fundos. Isso ocorre devido ao fato de que a os resultados são uma média de todos os retornos, dessa maneira a volatilidade individual de cada fundo é suavizada em torno de um valor esperado de retorno. Para facilitar a concepção dessa explicativa, podemos supor que todos os fundos têm valores esperados de retornos parecidos e que a média assintótica desses retornos se aproximaria a esse valor esperado. Assim, a série final deveria apresentar um desvio padrão inferior ao desvio padrão das séries que a compõem. Vejamos abaixo que o desvio padrão da Média de Mercado é inferior à maioria dos fundos.

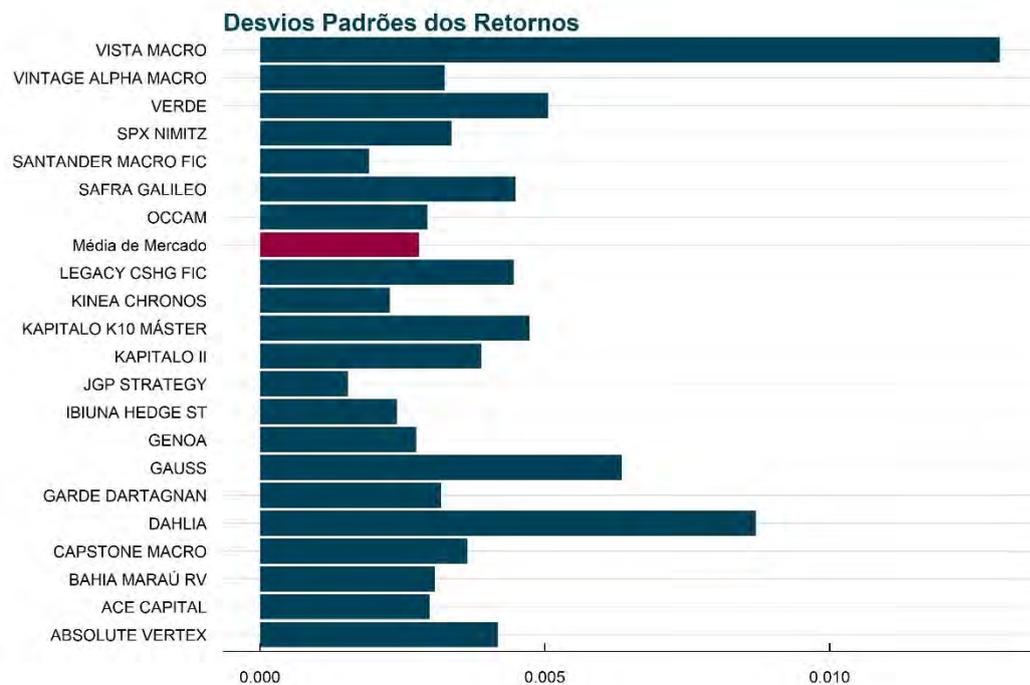


Figura 6-3-7 Bar chart de desvios padrões

6.4 – Análises Adicionais

Ao inferir as posições dos fundos com o objetivo de montar uma carteira sintética, nos deparamos com informações extremamente úteis. Com uma estimativa razoável das exposições dos fundos conseguimos fazer inúmeras análises interessantes de sentimento de mercado. Nessa seção abordaremos algumas das análises possíveis.

Das posições que estimamos, selecionamos o juro nominal de 10 anos BZAA10, o índice IBOV e a taxa de câmbio USDBRL para mensurar como os fundos estão se posicionando em relação aos ativos nacionais. Depois de selecionarmos esses ativos, calculamos um indicador de difusão, que nos mostra a porcentagem de fundos que estimamos que estão com posições *long* para os respectivos ativos. Para construir o indicador de cada ativo, consideramos que qualquer coeficiente positivo significaria uma posição comprada e o contrário seria uma posição *short*. Assim calculamos os indicadores de difusão abaixo. Veja abaixo o resultado desde 2020.



Figura 6-4-1 Gráfico de Difusão para ativos brasileiros

No gráfico acima, vemos que, segundo nossas estimativas, menos fundos estão exposto à bolsa de valores nacional e ao câmbio para os dados mais recentes. Em contrapartida, parece que a quantidade de fundos aplicados em juros longos aumentou na margem.

Outra maneira de estudar esses resultados seria construir um indicador de sentimento dos gestores em relação aos ativos brasileiros. Agora, em vez de inferir as posições de maneira binária, podemos pegar as médias dos coeficientes estimados para cada ativo a fim de criar uma série única que sintetizaria a exposição à ativos nacionais.

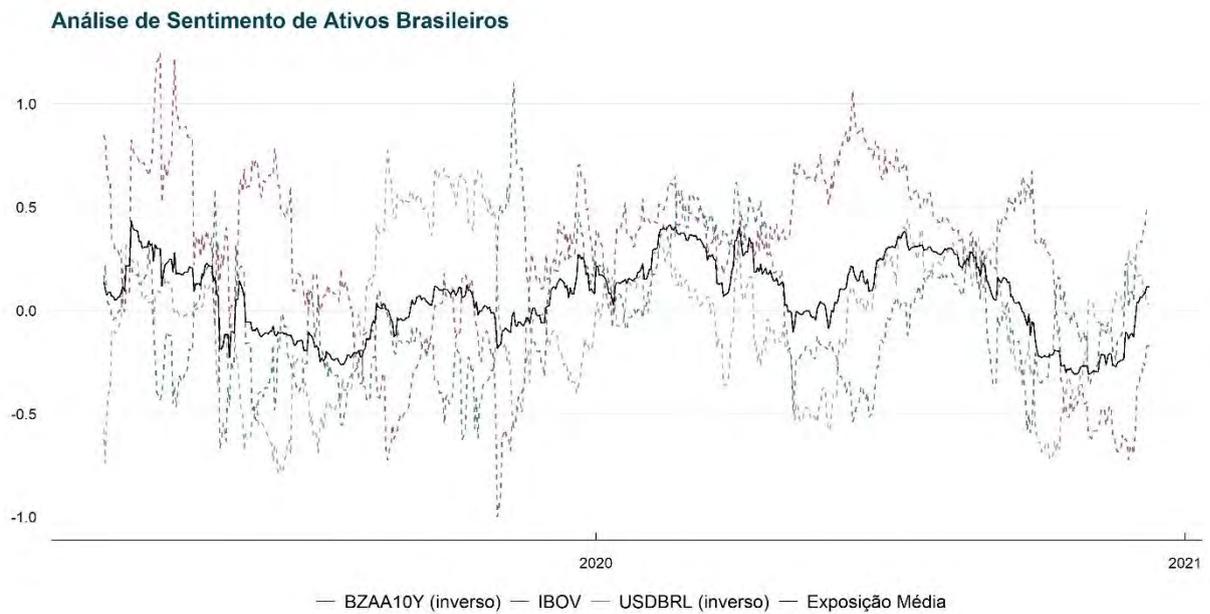


Figura 6-4-2 Gráfico de sentimento para ativos brasileiros

Observe que a partir desse resultado, conseguimos um resultado bem diferente, com o sentimento geral subindo consideravelmente na margem. Isso ocorre claro, devido ao fato de que não simplificamos as estimativas de maneira binária. Assim conseguimos uma medida mais razoável do sentimento médio do mercado.

Embora que essas análises ainda estejam bastante embrionárias, há sem dúvida bastante a se explorar em relação à análise de sentimento usando nossas estimativas. Primeiro, caberia um cuidado especial a fim de selecionar ativos que capturem melhor a exposição nacional com o intuito de calcular índices mais precisos. Em seguida, seria interessante ver se há algum poder preditivo nesses potenciais resultados.

7 – Conclusão

Observamos que diferentemente dos regimes regulamentares de outros países, a transparência imposta no mercado de fundos brasileiro nos permite acesso a inúmeros *insights* sobre as carteiras das principais gestoras.

Com o acesso aos retornos com pouquíssimos dias de atraso, podemos inferir com considerável precisão quais posições cada fundo detém em suas carteiras e, a partir daí, concebemos o posicionamento de diversos subconjuntos de fundos. Assim, estimamos tanto as posições médias e medianas de mercado, quanto as agregações dos melhores *performers* e a média ponderada por patrimônio líquido. Por si só, essa análise de *positioning* de mercado pode ser de extrema utilidade. Na Seção 6.4, ilustramos algumas formas de sintetizar as posições estimadas em séries bastante informativas acerca da exposição do mercado aos ativos nacionais.

Em seguida, usamos essas posições estimadas para construir as carteiras sintéticas. Dessa maneira, podemos verificar se há possibilidade de *piggyback* a performance dos fundos multimercados.

Vimos no Capítulo 6, que apesar de explicarmos grande parte da variância dos fundos com os ativos escolhidos, foi difícil replicar essa performance *out-of-sample*. Apesar do algoritmo apresentar um R-quadrado médio elevado, percebemos dificuldade de acompanhar a performance dos fundos com taxas de performance.

Isso pode se dar por alguns motivos. Primeiro, devemos considerar que os ativos selecionados podem não ter sido ideais. Dessa forma, pode ser que a variância não explicada pela regressão seja grande demais e justifique o menor retorno das carteiras.

Outra hipótese a se considerar seria a grande variância dos coeficientes estimados devido a multicolinearidade dos retornos e a curta janela de tempo. Como abordado no Capítulo 6, pode ser que se tivéssemos uma base de dados mais densa, esse problema poderia ser contornado.

Por fim, devemos também considerar que há limitações naturais para qualquer processo de inferência econométrica. Algumas estratégias de *hedge* ou operações *intra-day* podem gerar retornos não explicáveis sob uma ótica linear. Por isso, pode haver um limite de retorno modelável.

Sobretudo, é necessário frisar que os resultados apresentados pecam de maneira quantitativa e não qualitativa. De forma que, as carteiras conseguiram sim inferir as posições

dos fundos e apresentaram uma serie de retorno *out-of-sample* comparável aos fundos que as originaram, porém de magnitude menor. Com isso, afirmamos que com mais dados e uma melhor calibragem do modelo linear, pode ser que as estratégias apresentem retornos satisfatório quando comparadas aos fundos multimercados oferecidos atualmente.

8 – Referências

- Bertsimas, D. K. (2001). “Hedging Derivative Securities and Incomplete. *Operations Research*.
- de Carvalho, M. R. (2005). Avaliação de desempenho de fundos multimercado: resultados passados podem ser utilizados para definir uma estratégia de investimento? *Revista de Economia e Administração*.
- Elias Cavalcante-Filho, F. M.-L. (2021). Unskilled fund Managers: replicating active fund performance with a few ETFs. *Interantional Review of Financial Analysis*, 50.
- FRANS A. DE ROON, T. E. (2000). EVALUATING STYLE ANALYSIS.
- Loz, J. H. (2006). Can Hedge-Fund Returns Be.
- Martelanc, E. C. (2015). Sorte ou habilidade: uma avaliação dos fundos de investimento no Brasil. *Sci Elo Brazil*.
- Paranhos, Í. A. (2013). O desempenho e a persistência de fundos multimercado multiestratégia brasileiros. *LUME*.
- Segalis, M. G. (2018). Análise de persistência de performance de fundos multimercados macro no Brasil. *Red de Repositorios Latinoamericanos*.
- Shah, L. S. (2021). Fixed income mutual fund performance during and after a crisis: a Canadian case. *Journal of Economics and Finance*.
- Sharpe, W. (1992). ASSET ALLOCATION: MANAGEMENT STYLE AND PERFORMANCE MEASUREMENT. *Journal of Portfolio Management*.
- Young, D. P. (2007). The Hedge Fund Game: . *The Quarterly Journal of Economics*.