

A new of seeing momentum and evaluating metrics used to capture factor exposures

Pedro Teles

Resumo

Utilizando dados de 12 mercados, apresentamos uma nova métrica para captar o fator *momentum* e uma nova metodologia abrangente para avaliar e comparar os diferentes indicadores que são utilizados para captar a sensibilidade dos ativos aos fatores de risco já conhecidos. Mostramos que essa nova métrica possui poder preditivo sobre o retorno do portfólio clássico de *momentum* em 9 dos 12 mercados analisados. Além disso, mostramos que ela não é estatisticamente superior à métrica tradicional (retorno acumulado) no que tange o poder preditivo e que o espaço para os indicadores serem utilizados em conjunto é limitado.

Abstract

Using data from 12 different markets, we present a new metric to capture the momentum factor and a new comprehensive methodology to evaluate and compare indicators used to capture assets' sensibility to already known risk factors. We show that this metric has predictive power over the return of the classic momentum portfolio in 9 of the 12 markets analyzed. After that, we show that this new metric is not statistically superior to the traditional accumulated return regarding predictive power. Finally, there is little space for them to be used together.

1 Introdução

Momentum refere-se a um fenômeno amplamente conhecido e estudado na literatura financeira na qual ativos com um desempenho passado superior - comumente medido pelo retorno acumulado em uma dada janela temporal - tendem a continuar tendo um melhor desempenho que seus pares com desempenho passado inferior. Por mais que esse fator de risco venha recebendo especial atenção a décadas - foi descoberto no início da década de 90 (Jegadeesh and Titman (1993)) -, ainda não existe consenso sobre a origem desse prêmio de risco e o motivo dele ser tão persistente.

Geczy and Samonov (2016) utilizando uma base de dados com ações americanas que começa em 1801 e termina em 2012, mostram, em uma amostra *out of sample*, dado que até então os dados iniciavam apenas em 1926 (*CRSP database* e *SBBI files*), que o prêmio de *momentum* já era positivo e estatisticamente significativo no período pré 1926, além de também ser positivo e significativo no período posterior a 1926. Além disso, Hollstein (2020) expande a análise para outros mercados e mostra que *momentum* e *value* são bastante robustos em diferentes regiões do globo. Por fim, temos fortes evidências da presença de *momentum* em diferentes classes de ativos, como moedas (Menkhoff et al. (2012)), commodities (Blitz and De Groot (2014)) e juros (Van Luu and Yu (2012)).

Por mais que o prêmio de momentum seja estatisticamente e economicamente significativo em diferentes classes de ativos e em diferentes mercados, sua origem ainda não é muito clara. Kelly et al. (2021) afirmam que “apesar de sua influência generalizada na profissão financeira, *momentum* continua sendo um fenômeno misterioso. Uma variedade de teorias positivas, tanto comportamentais como racionais, foram propostas para explicar *momentum*, mas nenhuma é amplamente aceita”. Barberis et al. (2021), por exemplo, desenvolve um modelo de precificação de ativos que incorpora os elementos da teoria do prospecto (Kahneman (1979)) e é capaz de explicar parcialmente diversas anomalias de mercado, incluindo *momentum*.

Assim, o objetivo desse trabalho é apresentar uma explicação alternativa e complementar às teorias atuais por meio de uma nova métrica para captar o fator *momentum*. Além disso, utilizando uma nova metodologia, mostramos que essa métrica possui poder preditivo sobre os retornos do portfólio *Long & Short* de *momentum*. Por fim, comparamos essa nova métrica com o já conhecido retorno acumulado no que tange o poder preditivo de cada estatística e mostramos que elas são similares sob diferentes óticas.

Como sabemos, a teoria financeira pressupõe que os retornos de um ativo se compor-

tam como uma variável aleatória i.i.d. Dessa forma, em uma equação auto regressiva de ordem 1 ($r_t = \phi_0 + \phi_1 r_{t-1} + \epsilon_t$), esperamos que $\phi_0 = \phi_1 = 0$ (ruído branco). Entretanto, esses pressupostos precisam ser testados sob diferentes perspectivas. De fato, podemos considerar $\phi_1 = 0$ - se não pudéssemos, esse seria o parâmetro mais importante para qualquer ativo financeiro. Entretanto, não podemos automaticamente considerar $\phi_0 = 0$.

De fato, quando ordenamos os ativos com base no *drift* (intercepto) de um modelo AR(1), podemos constatar que o *drift* médio dos três primeiros decis é estatisticamente maior que o *drift* médio dos três últimos decis em todos os cenários analisados (*Wilcoxon test*). Com base nessa constatação e na suposição de estacionaridade fraca de ϕ_0 , podemos compreender a origem do prêmio de *momentum* de uma outra maneira: ativos com maior *drift* tendem a ter um termo (ϕ_0) que lhes conferem uma “vantagem” em relação àqueles com um menor *drift*. Desse modo, parte da explicação de *momentum* pode estar relacionada ao Processo Gerador de Dados (PGD) dos retornos dos ativos financeiros em adição às explicações já mais desenvolvidas na literatura relacionadas ao risco dos ativos e ao comportamento dos investidores.

Com efeito, quando fazemos a diferença da média do *drift* - estimado com base no último mês de dados para evitar o “viés da obsolescência” (Kelly et al. (2021)) - dos três primeiros decis e da média do *drift* dos últimos três decis, podemos mostrar, através do teste de Granger (Granger (1969)), que essa estatística Granger causa o retorno do portfólio *Long & Short* de *momentum* clássico com rebalanceamento mensal. A afirmação acima é válida para 9 dos 12 mercados analisados quando o nível de significância é de 5%. Entretanto, quando considerarmos o retorno acumulado (métrica tradicional) ao invés do *drift*, a afirmação passa a ser válida apenas em 7 dos 12 mercados analisados, considerando o mesmo nível de significância de 5%.

O resultado acima, além de ser estatisticamente significativo, também possui uma forte intuição econômica: se a diferença de médias é grande (pequena), isso indica que os ativos presentes nos três primeiros e últimos decis (não) são bastante distintos no que tange a sensibilidade ao fator *momentum*. Por esse motivo, podemos esperar que eles (não) tenham desempenhos futuros bastante desiguais. Desse modo, quando esse *spread* aumenta (diminui), o grau de distinção dos ativos aumenta (diminui), o que gera um(a) aumento (diminuição) na diferença de retorno esperado para os dois tercios de ativos. No presente artigo abordamos essa metodologia apenas para o fator *momentum*, mas nada impede que ela seja estendida para outros fatores de risco como *value*, *size*, lucratividade, etc.

Ademais, utilizando o teste de Vuong (Vuong (1989)), mostramos que o modelo uti-

lizando defasagens da diferença de médias do *drift* mais defasagens do retorno acumulado do portfólio *Long & Short* e o modelo utilizando defasagens da diferença de médias do retorno acumulado - também calculado com base no último mês de dados - mais as defasagens do retorno acumulado do portfólio L&S são indistinguíveis em quase todos os mercados analisados, mesmo o modelo utilizando o *drift* tendo um poder preditivo em mais mercados que aquele utilizando o retorno acumulado.

Por fim, utilizamos o teste ANOVA para verificar se o modelo contendo tanto as defasagens da diferença de médias do *drift* quanto defasagens da diferença de médias do retorno acumulado é superior ao modelo contendo apenas as defasagens de uma das variáveis. Como será apresentado, o teste F indica que o modelo restrito é preferível para grande parte dos mercados presentes na nossa amostra. Ou seja, ambas as variáveis carregam informações de mesma natureza - o que é esperado dado que o objetivo de ambas é captar o mesmo fenômeno - e a adição de uma delas a um modelo contendo a outra tende a não gerar um aumento significativo no poder explicativo do modelo.

O restante do artigo está estruturado da seguinte forma: primeiro, apresentamos os dados - origem, distância temporal e estatísticas descritivas; em seguida, nos aprofundamos na metodologia de cada um dos processos utilizados; depois, apresentamos os resultados; e, por fim, expomos as conclusões e fazemos algumas recomendações de pesquisas futuras possíveis envolvendo o tema.

2 Dados

Os nossos dados foram extraídos da provedora de dados financeiros Bloomberg e envolvem o período de 2002-01-01 até 2021-09-30. Como utilizamos um período de estimação de 12 meses para o portfólio de *momentum*, nossos portfólios começam em 2003-01-01. Para construção dos portfólios consideramos os mercados apresentados na Tabela 1, com seus respectivos índices de mercado - consideramos apenas aqueles mais amplos para evitar algum viés de seleção - e taxas livre de risco - utilizamos o título soberano zero cupom com maturidade de 10 anos. A frequência dos dados é diária, mas a do portfólio é mensal para estar na mesma frequência que o período de rebalanceamento.

Para definição dos ativos elegíveis, aplicamos algumas restrições visando construir portfólios mais verossímeis. As restrições são apresentadas abaixo, mas vale ressaltar que, para evitar o viés do sobrevivente, consideramos aquelas ações que estão ativas

Table 1: Mercados, Índices e Países

	Índice de Mercado	Taxa Livre de Risco
Alemanha	CDAX	GDBR10
Canadá	SPTSX	GCAN10YR
Coreia	KOSPI	GJGB10
Dinamarca	KAX	GDGB10YR
Espanha	MADX	GSPG10YR
Finlândia	HEX	GFIN10YR
França	SBF250	GFRN10
Reino Unido	ASX	GUKG10
Japão	TPX	GVSK10YR
Suécia	SAX	GSGB10YR
Suíça	SPI	GSWISS10
Taiwan	TWSE	GVTW10YR

(são negociadas nos dias de hoje) e também aquelas que deixaram de ser negociadas por qualquer motivo.

- O ativo precisa fazer parte da composição do índice daquele mercado no qual está inserido naquela data;
- O ativo precisa ter um volume médio diário (calculado com base no último mês) de, no mínimo, US\$100.000;
- O ativo precisa existir a, no mínimo, 12 meses;
- O ativo pode ter, no máximo, 10% de dados de preço faltantes.

Após aplicadas as restrições supracitas, apresentamos, na Tabela 2, algumas estatísticas relacionadas a quantidade de ativos elegíveis para cada mercado. Como podemos perceber, Japão, Reino Unido e Taiwan se destacam como mercados com um grande número de ativos respeitando nossas restrições; por outro lado, Dinamarca, Espanha e Finlândia, possuem uma amostra mais restrita. Com uma amostra com propriedades estatísticas bastante diversas e países desenvolvidos e emergentes, podemos estar seguros que temos uma amostra representativa do mercado global de ações.

3 Metodologia

Nessa seção apresentamos os procedimentos estatísticos que serão utilizadas ao longo do artigo, com um enfoque maior na nova metodologia que desenvolvemos para lidar com a relação intertemporal entre o retorno de um portfólio L&S de *momentum* clássico e a evolução de indicadores construídos para captar o fator *momentum* (e.g. retorno acumulado e *drift* do modelo AR(1) dos retornos).

3.1 Portfólio Clássico de *Momentum*

Para definirmos o portfólio de *momentum*, consideramos o padrão da literatura que seria comprar os 30% ativos com maior retorno acumulado entre os meses $t - 12$ e $t - 1$ e vender os 30% ativos com menor retorno acumulado no mesmo período. Além disso, adotamos pesos iguais para todos os ativos (*equal-weighted*) e rebalanceamos o

Table 2: Estatísticas Número de Ativos Elegíveis

	Mínimo	Máximo	Média	Desvio Padrão
Alemanha	144	351	247.9	36.9
Canadá	297	419	352.3	29.4
Coreia	368	825	605.1	101.8
Dinamarca	43	124	70.8	17.1
Espanha	73	116	97.8	10
Finlândia	43	109	72.4	12.1
França	178	350	258.1	32.9
Japão	973	2135	1677.3	244.2
Reino Unido	533	791	664.6	56.7
Suécia	108	326	211	45.6
Suíça	163	326	261.3	27.3
Taiwan	549	870	706.9	61.2

portfólio mensalmente. Optamos por realizar o rebalanceamento em uma frequência maior para garantir que as informações mais recentes acerca da exposição de cada ativo ao fator e risco sejam incorporadas ao portfólio.

3.2 Diferença de Médias

Suponhamos que um determinado mercado tenha n ativos elegíveis em uma dada data t . Para medirmos o poder preditivo de um determinado indicador financeiro (e.g. retorno acumulado) sobre o retorno de um determinado portfólio R_t (e.g. portfólio de *momentum*), primeiro ordenamos os n ativos daquele mercado disponíveis na data t com base nesse indicador, gerando o vetor $x_{i,t}$ $i = 1, 2, \dots, n$. Em seguida, calculamos a seguinte estatística μ_t :

$$\mu_t = \sum_{i=1}^{\text{round}(\frac{3n}{10})} \frac{x_{i,t}}{\text{round}(\frac{3n}{10})} - \sum_{i=\text{round}(\frac{7n}{10})}^n \frac{x_{i,t}}{\text{round}(\frac{3n}{10})} \quad (1)$$

No nosso caso, $t = 1, 2, \dots, 225$, dado que calculamos $x_{i,t}$ mensalmente de Janeiro de 2003 a Setembro de 2021. Além disso, para evitar o “viés da obsolescência” (Kelly et al. (2021)), estimamos tanto o *drift* do modelo AR(1) quanto o retorno acumulado com base no último mês de dados. Para ambas as estatísticas, iremos calcular, todos os mês, a diferença de médias e testaremos o poder preditivo de cada uma sobre o retorno do portfólio de *momentum* tradicional.

3.3 Teste de Granger

O teste de Granger (Granger (1969)) afirma que x_t Granger causa y_t se o modelo contendo tanto as defasagens de x_t quanto as defasagens de y_t possuir uma menor variância que o modelo contendo apenas as defasagens de y_t . Como sabemos, o teste de Granger pode ser aplicado a três situações, mas nesse artigo focaremos apenas na especificação contendo duas variáveis (μ_t e R_t) e suas defasagens. Para definir se a diferença na variância é estatisticamente significativa, iremos comparar ambos os modelos através de um teste F no qual nosso modelo restrito é aquele que inclui apenas as defasagens de R_t e nosso modelo irrestrito inclui as defasagens de R_t e também as defasagens de μ_t . Assim, a estatística do teste é definida da seguinte forma:

$$F = \frac{(SQR_R - SQR_{IR})/g}{(SQR_{IR})/(n - k)} F(g, n - k) \quad (2)$$

onde SQR_R e SQR_{IR} indicam a soma dos quadrados dos resíduos do modelo restrito e irrestrito respectivamente; g é o número de restrições no modelo restrito; n é o número de observações; e, por fim, k é o número total de regressores no modelo irrestrito.

Caso o valor dessa estatística F seja suficientemente grande para um dado nível de significância, teremos evidência suficiente para rejeitar a hipótese nula de validade do modelo restrito e, por consequência, consideraremos o modelo irrestrito como sendo mais adequado para entender a relação entre a variável dependente e os regressores.

No nosso caso, nosso modelo restrito incluirá três defasagens da variável dependente (retorno mensal do portfólio *Long & Short* clássico). O modelo irrestrito, por sua vez, incluirá as três defasagens da variável dependente, mas, também, incluirá três defasagens da diferença de médias, como definido anteriormente. Matematicamente,

$$\text{Modelo Restrito} \rightarrow R_t = \sum_{i=1}^3 R_{t-i} + \epsilon_t \quad (3)$$

$$\text{Modelo Irrestrito} \rightarrow R_t = \sum_{i=1}^3 R_{t-i} + \sum_{i=1}^3 \mu_{t-i} + \varepsilon_t \quad (4)$$

Em que R_t e μ_t já foram definidos e ϵ_t é o resíduo do modelo restrito e ε_t é o resíduo do modelo irrestrito.

3.4 Teste de Vuong

O teste de Vuong (Vuong (1989)) é um teste de razão de verossimilhança para seleção de modelos utilizando o critério de informação de Kullback–Leibler. Esse teste pode ser utilizado tanto para comparar modelos aninhados quanto não aninhados ou, ainda, sobrepostos. O teste irá apontar qual modelo está mais próximo do verdadeiro Processo Gerador de Dados sem, entretanto, definir quão próximo.

No nosso caso, utilizaremos esse teste para comparar uma especificação utilizando o indicador tradicional para captar o fator *momentum* - retorno acumulado - e outra

incorporando o indicador que propomos - *drift* do modelo AR(1) dos retornos. Essa especificação é idêntica à apresentada na Equação 4. Ou seja, iremos comparar dois modelos sobrepostos em que a variável dependente é o retorno mensal do portfólio de *momentum* clássico: o primeiro contendo as três primeiras defasagens da diferença de médias do retorno acumulado e do retorno acumulado do portfólio de *momentum*; o segundo contendo as três primeiras defasagens da diferença de médias do *drift* do modelo AR(1) dos retornos e do retorno acumulado do portfólio de *momentum*.

3.5 Teste F

Após testarmos se a diferença de médias das variáveis retorno acumulado e *drift* do modelo AR(1) possuem poder preditivo sobre o retorno do portfólio de *momentum* clássico (teste de Granger) e se algum modelo contendo defasagens das variáveis supracitadas é superior ao outro (teste de Vuong), analisamos se as defasagens desses indicadores podem ser utilizadas em conjunto para gerar um melhor ajuste ao retorno do portfólio de *momentum* clássico.

Para alcançar esse objetivo iremos utilizar o teste F, já apresentado anteriormente, no qual nosso modelo restrito contém as defasagens da diferença de médias do retorno acumulado mais as defasagens do retorno do portfólio *Long & Short* e o modelo irrestrito contém essas variáveis já citadas mais a diferença de médias do *drift* do modelo AR(1) dos retornos dos ativos. Repetiremos esse teste trocando a “base” do modelo: o modelo restrito terá a diferença de médias calculada com base no *drift* e para o modelo irrestrito adicionaremos a diferença de médias do retorno acumulado.

Caso sejamos capazes de rejeitar a hipótese nula ($p\text{-valor} < 0.05$), teremos evidências para considerar que o modelo contendo ambos os indicadores é superior, no que tange à capacidade de ajuste, a um modelo mais simples com apenas uma dessas variáveis.

4 Resultados

Primeiro, apresentamos os resultados do teste de Granger. Como podemos perceber pelos resultados da tabela 3, a diferença de médias do *drift* do modelo AR(1) dos retornos Granger causa o retorno do portfólio *Long & Short* de *momentum* tradicional em 9 dos 12 mercados analisados quando consideramos um nível de significância de 5%. Esse resultado é bastante interessante, dado a notória dificuldade em encontrar

variáveis que possuem poder preditivo sobre o retorno de variáveis financeiras.

Além disso, pelos resultados da segunda coluna da mesma tabela, podemos inferir que existe causalidade bi-direcional em 4 dos 12 mercados. Ou seja, a diferença de médias Granger causa o retorno do portfólio *Long & Short* ao mesmo tempo em que o retorno do portfólio Granger causa a diferença de médias. Por mais que esse resultado dificulte a interpretação de que o a diferença de médias causa o retorno do portfólio de *momentum*, a diferença de médias continua sendo uma variável valiosa para previsão do retorno do portfólio de *momentum*.

Para não adotarmos uma visão míope, também apresentamos os resultados considerando o retorno acumulado como variável para cálculo da diferença de médias. Nesse caso, a diferença de médias do retorno acumulado Granger causa o retorno do portfólio *Long & Short* de *momentum* em 7 dos 12 mercados, considerando o mesmo nível de significância de 5 %. Esse resultado é também interessante e levanta a pergunta de se algum dos indicadores é superior ao outro e, também, se eles podem ser utilizados em conjunto para melhorar a qualidade do modelo. Essas hipóteses serão testadas em seguida.

Table 3: Estatísticas F do teste de Granger

	AR(1)	AR(1) Reverso	Retorno Acumulado
Alemanha	2.88**	1.32	3.13**
Canadá	10.24***	8.36***	8.39***
Coréia	0.98	0.24	1.72
Dinamarca	0.67	0.59	1.48
Espanha	1.23	3.93***	2.53*
Finlândia	5.77***	1.54	2.06
França	5.78***	2.37*	4.84***
Japão	3.96***	5.83***	3.17**
Reino Unido	5.16***	3.08**	4.24***
Suécia	9.32***	4.04***	5.89***
Suíça	6.63***	1.47	6.96***
Taiwan	3.57**	0.6	2.47*
Graus de Liberdade: 218 (Irrestrito); 215 (Restrito)			

Códigos Signif.: 0.01 *** 0.05 ** 0.1 *

Primeiro, utilizando o teste de Vuong, testamos a hipótese de que o modelo contendo três defasagens da diferença de médias do *drift* do modelo AR(1) mais três defasagens do retorno acumulado do portfólio *Long & Short* é superior àquele contendo as mesmas

três defasagens do portfólio de *momentum* mais a diferença de médias utilizando retorno acumulado. Nesse caso, como podemos observar pela tabela 4, essa afirmação é verdadeira para apenas um mercado considerando um nível de significância de 5%. Ou seja, pelas estatísticas de teste devemos considerar ambos os modelos como sendo indistinguíveis. Também testamos a hipótese contrária de que o modelo contendo as defasagens do retorno acumulado é superior àquele contendo as defasagens do *drift*, mas para nenhum mercado isso foi verdadeiro.

Table 4: Estatística z do teste de Vuong

	LRT
Alemanha	0.15
Canadá	-1.21
Coréia	0.54
Dinamarca	0.64
Espanha	1.19
Finlândia	-2.11**
França	-0.69
Japão	-0.52
Reino Unido	-0.86
Suécia	-1.61*
Suíça	0.18
Taiwan	-0.6
Códigos Signif.: 0.01 *** 0.05 ** 0.1 *	

Por fim, analisamos a possibilidade de ambas as estatísticas (diferença de médias do *drift* e do retorno acumulado) serem utilizadas em conjunto através de um teste F. Os resultados são apresentados na tabela 5 e indicam que o modelo contendo as defasagens de ambas as variáveis é preferível ao contendo apenas as defasagens da diferença de médias do retorno acumulado em 2 e 4 mercados quando consideramos um nível de significância de 5% e 10% respectivamente. Desse modo, para a maioria dos mercados, o modelo mais parcimonioso em parâmetros é preferível, o que corrobora a hipótese levantada anteriormente de que ambas as variáveis carregam informações similares acerca da sensibilidade de cada ativo ao fator *momentum*.

Entretanto, se considerarmos o modelo restrito como sendo aquele em que a diferença de médias é construída com base no *drift* do modelo AR(1) dos retornos e para o modelo irrestrito acrescentamos a diferença de médias do retorno acumulado, a estatística F passa a ser significativa para apenas 1 país ao nível de 5% e 2 ao nível de 10%.

Table 5: Estatística F do teste ANOVA

	Drift AR(1)	Retorno Acumulado
Alemanha	1.39	1.63
Canadá	2.26*	0.61
Coréia	0.8	1.54
Dinamarca	0.35	1.15
Espanha	0.07	1.33
Finlândia	6.66***	2.96**
França	1.69	0.81
Japão	2.61*	1.84
Reino Unido	0.92	0.06
Suécia	5.43***	2.2*
Suíça	1.56	1.87
Taiwan	1.6	0.54
Códigos Signif.: 0.01 *** 0.05 ** 0.1 *		

5 Conclusão

Nesse trabalho, apresentamos uma nova métrica capaz de capturar o fator *momentum*. Além disso, desenvolvemos um arcabouço metodológico bastante robusto e abrangente para avaliar e comparar métricas que visam capturar algum fator de risco.

Utilizando o teste de Granger, mostramos que essa nova métrica (*drift* dos retornos de um modelo AR(1)) possui poder preditivo sobre os retornos do portfólio *Long & Short* de *momentum* tradicional. Também mostramos que a métrica mais tradicional (retorno acumulado) também possui poder preditivo, mas para menos mercados.

Além disso, através do teste de Vuong, mostramos que o modelo utilizando três defasagens da diferença de médias do *drift* dos retornos de um modelo AR(1) mais três defasagens do retorno do portfólio *Long & Short* é indistinguível daquele que contém as três defasagens do retorno do portfólio *Long & Short* mais as três defasagens da diferença de médias do retorno acumulado.

Por fim, utilizando o teste ANOVA, mostramos que para não muitos mercados (até 4 em nossa amostra de 12), haveria um ganho expressivo ao adicionar a diferença de médias do *drift* a um modelo contendo defasagens do retorno do portfólio *Long & Short* mais defasagens da diferença de médias do retorno acumulado. Porém, se considerarmos o modelo restrito como sendo aquele contendo as defasagens da diferença de médias do *drift* e do retorno do portfólio de *momentum*, a adição da diferença de médias do retorno acumulado seria relevante em até 2 mercados apenas.

Desse modo, é simples vislumbrar que um possível próximo passo nessa linha de pesquisa poderia envolver a aplicação dessa metodologia para outros fatores de risco. Poderíamos, por exemplo, ampliar a análise para aquelas métricas utilizadas para definir o fator valor como P/VPA, P/L, Free Cash Flow Yield, etc. Ou seja, testar o poder preditivo desses indicadores sobre o retorno do portfólio de valor e verificar se algum é superior ou, ainda, se eles podem ser utilizados em conjunto.

Além disso, seria bastante interessante testar a precisão do modelo para previsões fora da amostra. Para tal, poderia ser utilizado o teste de Clark e West (Clark and West (2006)), baseado na estatística de Diebold-Mariano-West, que faz uma correção para eliminar um viés existente quando comparamos um modelo preditivo com um modelo do tipo passeio aleatório.

Por fim, outros métodos econométricos poderiam ser utilizados para essa previsão *out-of-sample* como, por exemplo, regressão Ridge ou LASSO. Nesse caso, teríamos um aumento no viés do modelo, mas com a possibilidade de redução expressiva da variância para previsões fora da amostra.

References

- N. Barberis, L. J. Jin, and B. Wang. Prospect theory and stock market anomalies. *The Journal of Finance*, 76(5):2639–2687, 2021.
- D. Blitz and W. De Groot. Strategic allocation to commodity factor premiums. *The Journal of Alternative Investments*, 17(2):103–115, 2014.
- T. E. Clark and K. D. West. Using out-of-sample mean squared prediction errors to test the martingale difference hypothesis. *Journal of econometrics*, 135(1-2): 155–186, 2006.
- C. C. Geczy and M. Samonov. Two centuries of price-return momentum. *Financial Analysts Journal*, 72(5):32–56, 2016.
- C. W. Granger. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: journal of the Econometric Society*, pages 424–438, 1969.
- F. Hollstein. The world of anomalies: Smaller than we think? *Available at SSRN 3670504*, 2020.
- N. Jegadeesh and S. Titman. Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of finance*, 48(1):65–91, 1993.
- D. Kahneman. Prospect theory: An analysis of decisions under risk. *Econometrica*, 47:278, 1979.
- B. T. Kelly, T. J. Moskowitz, and S. Pruitt. Understanding momentum and reversal. *Journal of Financial Economics*, 140(3):726–743, 2021.
- L. Menkhoff, L. Sarno, M. Schmeling, and A. Schrimpf. Currency momentum strategies. *Journal of Financial Economics*, 106(3):660–684, 2012.
- B. Van Luu and P. Yu. Momentum in government-bond markets. *The Journal of Fixed Income*, 22(2):72–79, 2012.
- Q. H. Vuong. Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica: Journal of the Econometric Society*, pages 307–333, 1989.