

Supplementary information

THE SEARCH FOR TIME-SERIES PREDICTABILITY-BASED ANOMALIES

Javier Humberto OSPINA-HOLGUÍN^{ID*}, Ana Milena PADILLA-OSPINA^{ID}

Appendix A: Minimum transaction costs that return non-positive alphas

The results presented in the main article are dependent on transaction costs. If a given transaction cost is high enough, even if there were patterns in the time series returns of a decile portfolio, the patterns may not be profitably exploitable. In other words, if transaction costs rise sufficiently above zero, alpha will eventually disappear, and then alpha will turn negative. Table A1 reports the minimum threshold transaction cost, in basis points, that makes the algorithm deliver an economically non-positive alpha for the first time, as recorded while successively raising the level of transaction cost from zero. This measure is reported for each decile portfolio in the four settings examined in Table 2 in the main article. Nonetheless, it is important to note that, for some portfolios, the algorithm can occasionally deliver an economically positive alpha again when the level of transaction cost rises above those reported in Table A1. This phenomenon is due to alpha not always being a monotonically decreasing function of transaction cost.

Table A1. Minimum transaction cost that returns a non-positive alpha for each size decile portfolio in the four settings examined in Table 2 in the main article

Decile	Smallest	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	Largest
Minimum transaction cost (in basis points) that returns a non-positive Fama and French (2015) algorithmic alpha for each equal-weighted size decile portfolio										
MTC	43	9	6	8	8	8	11	7	4	2
Minimum transaction cost (in basis points) that returns a non-positive Fama and French (2015) algorithmic alpha for each value-weighted size decile portfolio										
MTC	35	8	7	8	8	9	10	5	3	–
Minimum transaction cost (in basis points) that returns a non-positive Carhart (1997) algorithmic alpha for each equal-weighted size decile portfolio										
MTC	51	9	7	8	8	11	11	7	4	2

End of Table A1

Decile	Smallest	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	Largest
	Minimum transaction cost (in basis points) that returns a non-positive Carhart (1997) algorithmic alpha for each value-weighted size decile portfolio									
MTC	41	11	8	9	9	9	10	6	4	2

Note: Minimum transaction cost (MTC), as measured in basis points, that renders a non-positive alpha for the first time when transaction cost is successively increased, starting from zero. Alpha is measured in a Fama and French (2015) or a Carhart (1997) time series regression on the algorithmic arbitrage market timing investment (long in the algorithm, short in the underlying decile portfolio). The decile portfolios correspond to each equal-weighted or value-weighted size decile portfolio for the testing sample between May 22, 1991, and April 30, 2019. For each decile, the algorithm was trained on data from July 1, 1963, to June 6, 1991. We used the 30-day T-bill as the risk-free asset, and one of the ten NYSE/AMEX/NASDAQ equal-weighted or value-weighted market-cap decile portfolios as the risky asset.

Appendix B: Moving blocks bootstrap p -values

To compute individual p -values that are robust against heteroscedasticity and autocorrelation (HAC) in the errors, as well as non-normality, we used the “naïve” (Davison & Hall, 1993) moving blocks bootstrap (MBB) (Künsch, 1989). The naïve bootstrap computes robust HAC t -statistics with the same formula in the bootstrap world as in the original data (Gonçalves & Politis, 2011), and, despite its “naivety”, it is sophisticated enough to perform exceptionally well in simulations where it has been found that the naïve bootstrap – and the IID bootstrap, in particular – outperforms the standard normal approximation (Kiefer & Vogelsang, 2005). Furthermore, there is an established theory supporting its suitability for linear regressions in the presence of heteroscedastic and autocorrelated errors (even for the IID case) (Gonçalves & Vogelsang, 2011).

To illustrate the approach, it is necessary to describe the setting for the bootstrap. For each investment rule (IR), we tested the two-sided null hypothesis of no abnormal alpha in the model $IR_t = \mathbf{x}'_t \boldsymbol{\beta} + \varepsilon_t$, $t = 1, \dots, T$, where $\mathbf{x}'_t = (x_{t0}, \mathbf{x}'_{t1})$; $x_{t0} \equiv 1$; \mathbf{x}'_{t1} is a p -dimensional vector of regressors at time t (either the five Fama and French (2015) factors or the four Carhart (1997) factors); $\boldsymbol{\beta}$ is a $p + 1$ -dimensional vector of regression coefficients; and ε_t can be non-normal, autocorrelated and heteroscedastic. (For simplicity, we have assumed that the j index which denotes the decile of the portfolio in the main text is fixed and it is, thus, omitted moving forward). Given the alpha ($\alpha \equiv \boldsymbol{\beta}_0$) is, by definition, the ordinary least squares coefficient corresponding to x_{t0} , the null hypothesis is that $\alpha = 0$.

The procedure for computing a p -value for this hypothesis using the MBB is as follows:

1. Run a linear regression of the model $y_t = \mathbf{x}'_t \boldsymbol{\beta} + \varepsilon_t$, $t = 1, \dots, T$ (where $y_t \equiv IR_t$) and obtain the first component of the least-squares estimate of the regression coefficient vector $\hat{\alpha} \equiv \hat{\boldsymbol{\beta}}_0$, as well as the HAC robust standard error for $\hat{\alpha}$, denoted by $\text{s.e.}(\hat{\alpha})$, using the Newey-West (1987) estimator of the $\boldsymbol{\beta}$ covariance matrix.¹

¹ Hanck et al. (2020), for example, provide detailed instruction on computing such a standard error. We used the Bartlett kernel and a truncation parameter of $\lceil 0.75T^{1/3} \rceil$, where $\lceil \cdot \rceil$ is the ceiling function.

2. Compute the t -statistics (in the original data) as:

$$t_{\hat{\alpha}} = \frac{\hat{\alpha} - 0}{\text{s.e.}(\hat{\alpha})}. \tag{1}$$

3. Compute $B = 5,000$ moving block t -statistic bootstraps as follows:

3.1. Set $l, 1 \leq l \ll T$, as the length of the bootstrap blocks, which are defined as:

$$\Psi_{s,l} \equiv \left\{ \left(y_s, \mathbf{x}'_s \right)', \left(y_{s+1}, \mathbf{x}'_{s+1} \right)', \dots, \left(y_{s+l-1}, \mathbf{x}'_{s+l-1} \right)' \right\}. \tag{2}$$

There are $n-l+1$ of such blocks. When $l=1$, the moving blocks bootstraps corresponds to the standard IID (paired) bootstrap, also known as random-x or case resampling (Fox & Weisberg, 2018).

3.2. Generate a set of b blocks $\Psi_{s,l}$ (each of length l) with uniform probability $1/(n-l+1)$, $s=1, \dots, n-l+1$ (Godfrey, 2009, pp. 207–214). That is, the generated blocks are obtained by random sampling, with replacement, from all possible $n-l+1$ overlapping blocks.

3.3. Generate a bootstrap sample, $S^* = \left\{ \left(y_t^*, \mathbf{x}'_t \right)', t=1, \dots, T^* \right\}$, of size $T^* = bl$ by joining together the b generated blocks (Godfrey, 2009, pp. 207–214). (We used $T^* = T$ and truncated the last block when appropriate.)

3.4. Using the bootstrap sample S^* , run a linear regression of the model $y_t^* = \mathbf{x}'_t \boldsymbol{\beta}^* + \varepsilon_t^*$, $t=1, \dots, T^*$ and obtain the first component of the least-squares estimate of the regression coefficient vector $\hat{\alpha}^* \equiv \hat{\boldsymbol{\beta}}_0^*$, as well as the HAC robust standard error of $\hat{\alpha}^*$, denoted by $\text{s.e.}(\hat{\alpha}^*)$, using the Newey-West (1987) estimator of the covariance matrix of $\boldsymbol{\beta}^*$.

3.5. Compute the bootstrap t -statistic as:

$$t_{\hat{\alpha}^*} = \frac{\hat{\alpha}^* - \hat{\alpha}}{\text{s.e.}(\hat{\alpha}^*)}. \tag{3}$$

4. Compute the value of q as the cumulative distribution function for the distribution D^* evaluated at $t_{\hat{\alpha}^*}$, where D^* is the empirical distribution function of the B bootstrapped values $t_{\hat{\alpha}^*}$.

5. Finally, compute the p -value associated with the null hypothesis that $\alpha = 0$ as follows:

$$p\text{-value} = \begin{cases} 2q & , \text{if } q \leq 1/2 \\ 2(1-q) & , \text{if } q > 1/2. \end{cases} \tag{4}$$

A similar procedure was used to compute the p -values of the other regression coefficients. Although we used a block size of $l=5$ in the reported results, we also compared the reported results with the results obtained using $l=1$ and $l=10$, all being similar to that of $l=5$.

References

- Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of Finance*, 52(1), 57–82. <https://doi.org/10.1111/j.1540-6261.1997.tb03808.x>
- Davison, A. C., & Hall, P. (1993). On studentizing and blocking methods for implementing the bootstrap with dependent data. *Australian Journal of Statistics*, 35(2), 215–224. <https://doi.org/10.1111/j.1467-842X.1993.tb01327.x>
- Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, 116(1), 1–22. <https://doi.org/10.1016/j.jfineco.2014.10.010>
- Fox, J., & Weisberg, S. (2018). Bootstrapping regression models in R. An appendix to *An R Companion to Applied Regression*. <https://socialsciences.mcmaster.ca/jfox/Books/Companion/appendices/Appendix-Bootstrapping.pdf>
- Godfrey, L. (2009). *Bootstrap tests for regression models*. Palgrave Macmillan. <https://doi.org/10.1057/9780230233737>
- Gonçalves, S., & Politis, D. (2011). Discussion: Bootstrap methods for dependent data: A review. *Journal of the Korean Statistical Society*, 40(4), 383–386. <https://doi.org/10.1016/j.jkss.2011.07.003>
- Gonçalves, S., & Vogelsang, T. J. (2011). Block bootstrap HAC robust tests: The sophistication of the naive bootstrap. *Econometric Theory*, 27(4), 745–791. <https://doi.org/10.1017/S0266466610000496>
- Hanck, C., Arnold, M., Gerber, A., & Schmelzer, M. (2020). *Introduction to econometrics with R*. University of Duisburg-Essen. <https://www.econometrics-with-r.org/>
- Kiefer, N. M., & Vogelsang, T. J. (2005). A new asymptotic theory for heteroskedasticity-autocorrelation robust tests. *Econometric Theory*, 21(6), 1130–1164. <https://doi.org/10.1017/S0266466605050565>
- Künsch, H. R. (1989). The jackknife and the bootstrap for general stationary observations. *The Annals of Statistics*, 17(3), 1217–1241. <https://doi.org/10.1214/aos/1176347265>
- Newey, W. K., & West, K. D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55(3), 703–708. <https://doi.org/10.2307/1913610>