Volatility-managed portfolios on French's factors: the replication, the real-time critique, and the 136 months after the sample

Carlo Hofer. November 2026. The note of the project Volatility-managed portfolios. Code, notebooks and the tables every number comes from: github.com/ochofer/volatility-managed-portfolios, results/202609.

Moreira and Muir (2017) scale a factor's position each month by the inverse of its realised variance in the month before, and report alphas, the intercepts of the regression of the scaled factor on the original, that are positive and significant for the market, momentum and profitability among the six factors studied here. Cederburg, O'Doherty, Wang and Yan (2020) show that an investor who had to estimate in real time how to combine the scaled and the original factor would have gained little. This note replicates both papers and Barroso and Santa-Clara's (2015) scaled momentum on Kenneth French's six factors, tests the Sharpe ratio differences on Moreira and Muir's sample and on the months after it (a Sharpe ratio is the mean excess return over its standard deviation, annualised) with Ledoit and Wolf's (2008) bootstrap, and asks what the signal did in the 136 months after Moreira and Muir's sample, May 2015 to August 2026. Six expectations were written before any code judged them, and four are met. The replications hold for 6 of 6 factors in Moreira and Muir's Table I and in 4 of 4 comparisons for scaled momentum, and in 22 of 24 comparisons for the real-time critique, an expectation not met because French has revised the profitability factor's history since the paper. The expectation that the Sharpe ratio test and the alpha's t-statistic agree is not met either, 2 of 6 factors where 5 were expected, because the two test different things: the volatility-managed Sharpe ratio is above the original's beyond its interval for momentum alone, +0.51, annualised, with an interval of 0.07 to 0.96. After the sample, no factor's real-time combination has a Sharpe ratio above the real-time original's by more than its interval allows, the median difference +0.007, annualised, against a median half-width of 0.213, so these months cannot tell the differences from zero. The market's alpha net of 10 basis points per unit of turnover is 4.03 per cent a year (t 2.31), and 3.13 (t 1.81) net of 20.

Terms used in this note, each with the one meaning it carries throughout.

factor: the monthly return of a long-short portfolio that captures one source of common movement in stock returns: the market (Mkt-RF), size (SMB), value (HML), momentum (Mom), profitability (RMW) and investment (CMA) of Fama and French (2015) and Carhart (1997), as Kenneth French publishes them. **excess return**: a return minus the one-month Treasury bill rate; every factor is one. **realised variance**: the variance of a factor computed from its daily returns within one month, each paper with its own definition. **the signal**: last month's realised variance of a factor, and nothing else. **volatility-managed factor**: Moreira and Muir's construction, the factor's return times c over last month's realised variance; the word managed describes the construction and says nothing about its performance. **c**: the constant that gives the volatility-managed factor the original's standard deviation over the whole sample, so it is known only at the sample's end. **position**: the multiple of the factor held in a month, c over last month's realised variance for the volatility-managed factor. The papers call a position above 1 leverage. **real-time combination (S1)**: Cederburg et al.'s construction: the mean-variance weights on the volatility-managed and the original factor, with a risk aversion of 5, estimated each month from the months before it, from 120 months on, the total position capped at 5. **real-time original (S2)**: the original factor alone at its real-time mean-variance weight, S1's benchmark. **ex post optimal combination (S3)**: the mix of the two with the weights that were best over the whole sample, which no investor could have known. **benchmark**: the strategy a construction is compared with. **scaled momentum**: Barroso and Santa-Clara's construction, WML times 12 per cent a year over its realised volatility of the previous 126 trading days. **WML**: winners minus losers, the top momentum decile's return minus the bottom decile's. **alpha**: the intercept of the regression of the volatility-managed factor on its original, in per cent a year; it measures what the volatility-managed factor adds to a holding of the original. **beta**: the slope of that regression. **standard error**: the standard deviation of an estimate across the samples it could have come from; White's allows the residual's variance to change, and Newey and West's at 12 lags also allows a month's residual to be correlated with the 12 before it. **Sharpe ratio**: the mean excess return over its standard deviation, annualised by the square root of 12. **appraisal ratio**: the alpha over the residual's standard deviation, annualised. **Jobson-Korkie test**: a test that two return series over the same months have the same Sharpe ratio, assuming normal and independent months, with Memmel's (2003) correction. **interval**: the 95 per cent interval of Ledoit and Wolf's test for the true Sharpe ratio difference; the test rejects when it excludes zero. **half-width**: the distance from a difference to either end of its interval. **delta method**: the standard error of a function of means, from the function's slopes and the means' covariance. **fat tails**: extreme months that occur more often than a normal distribution allows. **degrees of freedom**: the parameter of a Student-t distribution that sets how fat its tails are; the fewer, the fatter. **Student-t**: a bell-shaped distribution with fatter tails than the normal one, here with five degrees of freedom. **kurtosis**: excess kurtosis, a measure of how often extreme values occur, 0 for the normal distribution. **p-value**: the probability of a difference at least as far from zero as the one computed, if the two true Sharpe ratios were equal. **size**: the share of samples in which a test rejects when the null is true. **turnover**: the absolute change in the position from the month before, the first month carrying none. **basis point**: a hundredth of a percentage point. **break-even cost**: the cost per unit of turnover at which the alpha falls to zero. **maximum drawdown**: the largest fall from a peak of the compounded return. **band**: the distance from a published number within which a computed one counts as matching, fixed before the code. **verdict**: whether an expectation is met; a count reported in no verdict is shown without one. **vintage**: the version of French's files, named by the CRSP release they were built from.

1. Introduction

A volatility-managed portfolio holds a factor in inverse proportion to the variance the factor showed in the month before, less after a volatile month and more after a calm one. Moreira and Muir (2017) applied the rule to the market and to the standard equity factors from 1926 to April 2015 and found alphas against the unscaled factors that are positive and significant for the market, momentum and profitability among the six factors studied here, the market's 4.86 per cent a year with a standard error of 1.56. The rule needs one number from the past, last month's realised variance, and one constant, c, which sets the position's average size and is chosen on the whole sample. The question this note asks is how much of that result an investor could have used: whether it reproduces on French's data as they stand today, whether it survives the combination of the managed and the original factor being estimated in real time, and what the signal did in the months after the paper's sample ended.

The existing evidence answers the second part and leaves the third open. Cederburg, O'Doherty, Wang and Yan (2020) argue that an alpha against the original factor is the gain of an investor who already holds the original and adds the managed factor in the amount the regression implies, an amount known only at the sample's end, like c. When they estimate that amount each month from the months before it, the combination's Sharpe ratio on the six factors of this note falls below the real-time original's for four of them and differs significantly only for momentum, where it is higher. Their test of a Sharpe ratio difference is Jobson and Korkie's, which assumes normal and independent months, and Moreira and Muir report alphas, so neither paper gives a test of the Sharpe ratio difference that allows for fat tails and for volatile months that follow each other. Neither can say what happened after its sample: Moreira and Muir's ends in April 2015 and Cederburg et al.'s in December 2016.

This note replicates both papers and Barroso and Santa-Clara's (2015) scaled momentum, the third paper's version of the same scaling, on French's six factors, each paper on its own sample and with its own definition of the realised variance, and judges each replication against expectations written before the code ran. It then tests the Sharpe ratio differences on Moreira and Muir's sample and on the months after it with Ledoit and Wolf's (2008) studentised block bootstrap, checked first on simulated pairs whose Sharpe ratios are equal by construction, and carries both constructions forward unchanged to the 136 months from May 2015 to August 2026, with c fixed on the months to April 2015. A cost of 10 basis points per unit of turnover, with 20 as the one sensitivity, prices the trades the signal needs. The design estimates no premium and recommends no position: every return reported is a simulation on French's history, with its standard error or interval beside it, and every number comes from a table committed with the code.

French's factors are an informative setting for these questions for two reasons. Both papers built their results on them and French keeps them current, so the replication and the months after the sample run on one source, and French's revisions of history become visible: a factor whose statistics differ from a paper's on the same months has changed since the paper's download, and an archived copy of the file can show by how much. At the same time, the months after the sample hold the highest realised variance of the market in French's history, March 2020, so the signal was tested in a month in which it cut the position to 0.013 for April, when the market returned 13.6 per cent, and the 136 months are few enough that every difference they give has to be read against its interval.

Empirically, the replications hold where French's data have not changed, and the alpha and the Sharpe ratio separate. Moreira and Muir's Table I reproduces for 6 of 6 factors and scaled momentum in 4 of 4 comparisons; Cederburg et al.'s results reproduce in 22 of 24 comparisons, the two misses on RMW, whose mean French's current file puts at 3.12 per cent a year on their months against their 2.92. On Moreira and Muir's sample the volatility-managed Sharpe ratio is above the original's beyond its interval for momentum alone, +0.51, annualised, with an interval of 0.07 to 0.96, while the alpha's t-statistic is above two for the market, momentum and RMW, and an identity shows why the two tests differ. After the sample no factor's real-time combination has a Sharpe ratio above the real-time original's by more than its interval allows, the median difference +0.007, annualised, against a median half-width of 0.213. These 136 months therefore cannot tell the differences from zero, and they rule out neither a positive nor a negative difference of the size estimated on Moreira and Muir's sample.

2. Design

Data. Kenneth French's data library, the vintage built from the 202609 CRSP database: the three-factor and five-factor files, the momentum factor and the ten portfolios sorted on prior returns, each monthly and daily, downloaded at run

time and never committed. The market, SMB and HML start in July 1926, momentum and WML in January 1927 (their daily files on 3 November 1926), RMW and CMA in July 1963. The sample ends in August 2026 whatever the vintage. Each paper's results are compared on its own months: Moreira and Muir's to April 2015, the end their published counts of months imply (1,065 for the market, 621 for RMW and CMA); Cederburg et al.'s to December 2016; Barroso and Santa-Clara's from May 1927 to December 2011. The months after the sample are May 2015 to August 2026.

The signal and the constructions. Each construction uses its paper's own realised variance: Moreira and Muir's sum of squared deviations of the daily returns from the month's mean (their equation 2), Cederburg et al.'s 22 over the month's trading days times the sum of squared daily returns (their equation 4), and Barroso and Santa-Clara's 21 times the mean of the squared daily returns over the previous 126 trading days (their equation 5). The volatility-managed factor holds c over last month's realised variance; c makes its standard deviation over the sample equal to the original's. The real-time combination S1 holds, each month, the mean-variance weights of an investor with a risk aversion of 5 on the volatility-managed and the original factor, estimated from all the months before it once 120 are available, its total position capped at 5; S2 holds the original alone at its real-time weight, and S3 the combination with the weights that were best over the whole sample. Scaled momentum holds WML at 12 per cent a year over its realised volatility of the previous 126 trading days.

The regression and its standard errors. Each volatility-managed factor is regressed on its original, monthly, by least squares. The alpha is the intercept in per cent a year. Moreira and Muir's Table I is judged with White's standard errors, the paper's own; Newey and West's at 12 lags are reported beside them and used for the alpha in sections 3.3 and 3.5. Cederburg et al.'s Sharpe ratio differences are reproduced with their Jobson and Korkie test.

The test of a Sharpe ratio difference. Ledoit and Wolf (2008) write the difference between two Sharpe ratios as a function of four means, those of each series and of its square, and take its standard error by the delta method from the long-run covariance of the four. On the sample the covariance comes from the prewhitened Parzen kernel with Andrews' (1991) bandwidth; the paper proposes the QS kernel and reports the Parzen kernel's performance as very similar. The interval comes from a circular block bootstrap: 2,000 samples of blocks of 12 consecutive months, each sample's difference studentised by its own standard error from its blocks, and the 95th percentile of the absolute studentised differences, z*, replacing the normal table's 1.96. The test rejects when the interval excludes zero. The paper calibrates the block length by its Algorithm 3.1; here it is fixed at 12 months before the run. Before the test touches a factor it is run on 1,000 simulated pairs of 1,065 months with the market's monthly mean and standard deviation, a correlation of 0.601, Student-t shocks with five degrees of freedom and equal Sharpe ratios by construction, to measure its size.

Costs and positions. A month's turnover is the absolute change in the position from the month before, the first month carrying none. The cost, 10 basis points per unit of turnover and 20 as the sensitivity, is subtracted from the volatility-managed return, and the alpha is estimated again on the net return. The factor's own rebalancing is the same for the managed and the original factor and is left out. The break-even cost is the gross alpha over 12 times the average turnover. The position's percentiles are compared with Moreira and Muir's Table V and Cederburg et al.'s Table 1, Panel D.

The expectations. Six expectations were written into the code's constants on 8 and 10 November 2026, each with its reason, before any run that judged them: E1 to E3 on the three replications, E4 on the test's size and on its agreement with the alpha, E5 on the months after the sample, E6 on the costs and the positions. Each is a strict pass or fail, and no level moved after a run. Table 9 states each with what the run found.

3. Results

3.1 The replications on each paper's sample

Table 1 reports Moreira and Muir's Table I beside the computed regression, and the replication holds for all six factors (E1 met). Every alpha has the published sign and 5 per cent significance, lies within one published standard error of the published alpha and has an appraisal ratio within 0.05 of the published one, and the computed months equal the published ones. The market's alpha is 4.93 per cent a year with a White standard error of 1.57, against 4.86 and 1.56, and its volatility-managed Sharpe ratio is 0.52, as in Table V. The largest gap is RMW's, 2.80 against 2.44 per cent a year, 0.43 of a published standard error, the first sign of the revision section 3.2 measures. The Newey-West standard

errors are larger than White's for five factors, by up to 43 per cent for SMB, because the residuals of neighbouring months are correlated, which Newey and West's standard error allows and White's does not.

Table 1. Moreira and Muir's Table I on their sample, published beside computed: the alpha of each volatility-managed factor on its original, in per cent a year

		Alpha		*Standard error*			*Appraisal ratio*		
Factor	*Months*	*Published*	*Computed*	*Published*	*White*	*Newey-West*	*Published*	*Computed*	*E1 holds*
Mkt-RF	1065	4.86	4.93	1.56	1.57	1.76	0.33	0.33	yes
SMB	1065	−0.58	−0.59	0.91	0.91	1.31	−0.07	−0.07	yes
HML	1065	1.97	1.73	1.02	1.05	1.27	0.20	0.17	yes
Mom	1060	12.51	12.83	1.71	1.72	2.05	0.86	0.87	yes
RMW	621	2.44	2.80	0.83	0.89	1.00	0.42	0.44	yes
CMA	621	0.38	0.36	0.67	0.67	0.55	0.08	0.07	yes

Each factor from its first month to April 2015: August 1926 for the market, SMB and HML, January 1927 for momentum, August 1963 for RMW and CMA. The published standard errors are White's, as are the computed ones the verdict uses; the Newey-West standard errors, at 12 lags, are those of sections 3.3 and 3.5. The market's volatility-managed Sharpe ratio is 0.52 computed against Table V's 0.52. Source: results/202609/e1_comparison_202609.csv, notebook 03.

Table 2 reports Barroso and Santa-Clara's Table 3 for momentum and scaled momentum, and their two headline results reproduce (E3 met). Scaling WML to 12 per cent a year raises its Sharpe ratio from 0.54 to 1.00, against the paper's 0.53 and 0.97, and cuts its maximum drawdown from 95.6 to 43.8 per cent, against 96.7 and 45.2. The worst month of scaled momentum is a fall of 24.2 per cent against WML's 77.7, and its excess kurtosis 2.0 against 18.0. Run on to August 2026, the 1,192 months from May 1927 give Sharpe ratios of 0.49 and 0.92 and the same drawdowns, in no verdict.

Table 2. Barroso and Santa-Clara's Table 3 on May 1927 to December 2011, published beside computed: momentum (WML) and scaled momentum

	WML		*Scaled WML*	
Statistic	*Published*	*Computed*	*Published*	*Computed*
Mean, % a year	14.46	14.86	16.50	16.69
Standard deviation, % a year	27.53	27.42	16.95	16.77
Sharpe ratio	0.53	0.54	0.97	1.00
Best month, %	26.18	26.24	21.95	21.71
Worst month, %	−78.96	−77.66	−28.40	−24.16
Skewness	−2.47	−2.39	−0.42	−0.36
Excess kurtosis	18.24	18.04	2.68	2.01
Maximum drawdown, %	−96.69	−95.56	−45.20	−43.83

1,016 months. The paper's scaled series starts in March 1927; French's daily file of the ten momentum portfolios starts on 3 November 1926, so the first scaled month here is May 1927. E3 judges the two Sharpe ratios and the two maximum drawdowns. Source: results/202609/bsc_table_3_computed_202609.csv and e3_comparison_202609.csv, notebook 05.

3.2 The real-time critique, and a revised factor

Tables 3 and 4 report Cederburg et al.'s Tables 1 and 5 beside the computed ones, and 22 of the 24 comparisons hold, so E2 is not met. Both misses are RMW's: its volatility-managed Sharpe ratio is 0.566 against their Table 1's 0.51, 0.006 outside the band of 0.05, and S1's 0.510 against their Table 5's 0.44, 0.020 outside. The other comparisons hold: S1 minus S2 has the published sign for six of six factors, momentum's +0.41 is the only difference significant at 5 per cent by the Jobson and Korkie test, as published, and S1 falls short of S3 by 0.08 to 0.12, against the published 0.07 to 0.14.

Table 3. Cederburg, O'Doherty, Wang and Yan's Table 1 on their sample, published beside computed: annualised Sharpe ratios of the original and the volatility-managed factor

	Original		*Volatility-managed*		*Difference*		*p-value*	
Factor	*Published*	*Computed*	*Published*	*Computed*	*Published*	*Computed*	*Published*	*Computed*
Mkt-RF	0.42	0.420	0.51	0.513	+0.09	+0.092	0.30	0.31
SMB	0.23	0.226	0.08	0.076	−0.15	−0.150	0.09	0.09
HML	0.40	0.401	0.38	0.367	−0.02	−0.033	0.86	0.74
Mom	0.48	0.475	0.99	0.962	+0.50	+0.487	0.00	0.00
RMW	0.38	0.406	0.51	0.566	+0.13	+0.161	0.29	0.20
CMA	0.53	0.534	0.40	0.392	−0.13	−0.142	0.23	0.19

Each factor from its first month to December 2016: 1,085 months for the market, SMB and HML, 1,080 for momentum, 641 for RMW and CMA. The p-value is Jobson and Korkie's with Memmel's correction. E2's clause (i) fails for RMW alone. Source: results/202609/cowy_table_1_computed_202609.csv, notebook 04.

Table 4. Cederburg, O'Doherty, Wang and Yan's Table 5, Panel A, on their sample, published beside computed: the real-time combination S1, the real-time original S2 and their differences

	S1		*S2*		*S1 − S2*		*p-value*		*S1 − S3*	
Factor	*Published*	*Computed*	*Published*	*Computed*	*Published*	*Computed*	*Published*	*Computed*	*Published*	*Computed*
Mkt-RF	0.42	0.415	0.46	0.458	−0.04	−0.043	0.64	0.62	−0.11	−0.113
SMB	0.14	0.128	0.19	0.187	−0.06	−0.059	0.37	0.36	−0.12	−0.122
HML	0.38	0.375	0.43	0.442	−0.06	−0.067	0.41	0.33	−0.12	−0.122
Mom	0.92	0.889	0.49	0.477	+0.44	+0.412	0.00	0.00	−0.07	−0.078
RMW	0.44	0.510	0.31	0.328	+0.13	+0.182	0.53	0.38	−0.14	−0.123
CMA	0.52	0.526	0.56	0.561	−0.03	−0.036	0.20	0.14	−0.11	−0.121

Months as in Table 3, each strategy from 120 months after the factor's first month. S3 is the ex post optimal combination. E2's clause (ii) fails for RMW alone, in S1; clauses (iii) and (iv) hold. Source: results/202609/cowy_table_5_computed_202609.csv, notebook 04.

The cause of the miss is French's revision of RMW since the paper's download. On the 641 months of their Table 1, RMW's mean in the current file is 3.12 per cent a year against their 2.92, with the same standard deviation, 7.70 against 7.71, while the other five factors' means lie within 0.12 of theirs. Two diagnostics, computed after the verdict and carrying none, point the same way. Moving RMW's monthly returns by one constant to their mean closes the volatility-managed Sharpe ratio's gap from 0.056 to 0.022 and S1's from 0.070 to 0.027. The Internet Archive's copy of French's five-factor file of 9 July 2017 gives RMW a mean of 2.89 and a standard deviation of 7.74 on the same months, the mean within 0.03 of theirs. The miss therefore measures a change in French's data, and the replication of the construction holds on the other five factors.

3.3 The alpha and the Sharpe ratio difference

The test of a Sharpe ratio difference passes its size check: on the 1,000 simulated pairs it rejects 37 at the 5 per cent level, 3.7 per cent, inside the range of 3.5 to 6.5 fixed before the run as two standard errors of a share of 0.05. Table 5 reports it on Moreira and Muir's sample beside the alpha. The volatility-managed Sharpe ratio is above the original's beyond its interval for momentum alone, by 0.51, annualised, with an interval of 0.07 to 0.96 and a p-value of 0.03. For the market the difference is +0.10 with an interval of −0.11 to 0.31 and a p-value of 0.33, beside an alpha with a Newey-West t-statistic of 2.80; for RMW it is +0.19 with an interval of −0.11 to 0.49, beside a t-statistic of 2.81. E4 expected the two tests to agree in sign and significance for at least five factors, and they agree for two, SMB and momentum, so E4 is not met; by category, in no verdict, they agree for four. The signs of HML's and CMA's alphas, positive, and of their Sharpe ratio differences, negative, were known from the replication before any bootstrap ran, and they already limited the agreement to four of six.

Table 5. The alpha and the Sharpe ratio difference on Moreira and Muir's sample: the volatility-managed factor against its original

Factor	Alpha, % a year	t (Newey-West)	Sharpe ratio, managed	Sharpe ratio, original	Difference	95% interval	p-value	Correlation ρ	α/σ	(1 − ρ) × Sharpe ratio, original
Mkt-RF	4.93	2.80	0.515	0.419	+0.097	−0.112 to 0.305	0.33	0.601	+0.264	0.167
SMB	−0.59	−0.45	0.084	0.227	−0.143	−0.362 to 0.076	0.19	0.605	−0.053	0.090
HML	1.73	1.36	0.354	0.395	−0.041	−0.270 to 0.189	0.72	0.539	+0.141	0.182
Mom	12.83	6.24	0.993	0.480	+0.514	0.072 to 0.955	0.03	0.450	+0.777	0.264
RMW	2.80	2.81	0.598	0.407	+0.192	−0.108 to 0.491	0.21	0.583	+0.361	0.170
CMA	0.36	0.66	0.414	0.538	−0.124	−0.337 to 0.090	0.24	0.672	+0.052	0.176

Months as in Table 1. The difference, its interval and its p-value are Ledoit and Wolf's test, 12-month blocks, 2,000 bootstrap samples, annualised. The difference equals α/σ minus the last column for every factor, to four decimals. Source: results/202609/sharpe_tests_mm_sample_202609.csv and alpha_and_sharpe_difference_202609.csv, notebook 06.

An identity shows why the two tests differ. With the volatility-managed factor scaled to the original's standard deviation σ, the regression's slope equals the correlation ρ between the two, and the volatility-managed factor's mean is α + ρμ, so its Sharpe ratio minus the original's equals α/σ minus (1 − ρ) times the original's Sharpe ratio. Table 5 computes both sides, equal to four decimals for every factor. The alpha therefore measures what the volatility-managed factor adds to a holding of the original, and the difference measures the volatility-managed factor held in the original's place, which gives up the share 1 − ρ of the original's Sharpe ratio. With correlations of 0.45 to 0.67, the alpha per unit of risk has to exceed a third to a half of the original's Sharpe ratio before the difference turns positive: for the market α/σ is 0.264 and (1 − ρ) times the Sharpe ratio 0.167, so the difference is +0.097; for HML they are 0.141 and 0.182, so a positive alpha of 1.73 per cent a year comes with a difference of −0.041. Cederburg et al.'s Table 1 shows the same separation for the market, a difference with a p-value of 0.30 beside Moreira and Muir's t-statistic of 3.1. The sign of a difference on this sample says nothing about its sign on another.

3.4 The 136 months after the sample

Table 6 reports the four Sharpe ratios on May 2015 to August 2026 and the two comparisons, and E5 is met. No interval for S1 minus S2 lies above zero, where at most two were allowed, and the median difference, +0.007, annualised, lies within the median half-width, 0.213. The six differences run from −0.090 for HML, with an interval of −0.234 to 0.053, to +0.198 for momentum, with an interval of −0.342 to 0.738. S1 and S2 are close where S1's position moves little: for SMB, HML and CMA their returns have correlations of 0.98 or more and the half-widths are 0.022 to 0.144, while S1's market position ran from 0.18 to 1.97 against S2's 0.44 to 0.49, a correlation of 0.798 and a half-width of 0.639.

Table 6. The 136 months after Moreira and Muir's sample, May 2015 to August 2026: Sharpe ratios and their differences with Ledoit and Wolf's intervals, annualised

	Sharpe ratio				*S1 − S2*			*Volatility-managed − original*		
Factor	*Original*	*Volatility-managed*	*S1*	*S2*	*Difference*	*95% interval*	*Half-width*	*Difference*	*95% interval*	*Half-width*
Mkt-RF	0.784	0.554	0.693	0.775	−0.083	−0.722 to 0.556	0.639	−0.231	−1.120 to 0.659	0.889
SMB	−0.199	−0.301	−0.180	−0.211	+0.031	−0.092 to 0.154	0.123	−0.103	−0.690 to 0.485	0.587
HML	−0.019	−0.333	−0.141	−0.051	−0.090	−0.234 to 0.053	0.144	−0.315	−1.032 to 0.402	0.717
Mom	0.207	0.383	0.394	0.197	+0.198	−0.342 to 0.738	0.540	+0.177	−0.449 to 0.803	0.626
RMW	0.270	0.410	0.311	0.258	+0.052	−0.229 to 0.334	0.282	+0.140	−0.394 to 0.675	0.535
CMA	−0.061	−0.317	−0.093	−0.075	−0.018	−0.040 to 0.005	0.022	−0.257	−1.046 to 0.532	0.789

c fixed on the months to April 2015; S1 and S2 carried forward from their expanding windows. Each test uses 2,000 bootstrap samples of 11 blocks of 12 months. E5 judges the S1 − S2 columns; the volatility-managed columns carry no verdict. These months cannot tell any of the twelve differences from zero, and the sign of a difference on them says nothing about another sample. Source: results/202609/after_sample_sharpe_202609.csv, e5_real_time_tests_202609.csv and e5_volatility_managed_tests_202609.csv, notebook 07.

The volatility-managed factor with c fixed on the months to April 2015 gives the same counts, in no verdict: no interval above zero, and a median difference of −0.167 within a median half-width of 0.672. Its Sharpe ratio is below the original's for four factors and above it for momentum and RMW. For the market it is 0.554 against the original's 0.784,

a difference of −0.231 with an interval of −1.120 to 0.659. The three months with the largest gaps between the two returns are October 2018, April 2020 and January 2018. In October 2018 the position was 3.56, after a calm September, and the market fell 7.65 per cent while the volatility-managed market fell 27.3 per cent. In April 2020 the position was 0.013, after March 2020, the highest realised variance of the market in French's history, and the market rose 13.6 per cent while the volatility-managed market rose 0.18 per cent. In January 2018 the position was 3.38, and the market rose 5.58 per cent while the volatility-managed market rose 18.9 per cent.

The intervals are wide because the months are 136. A standard error falls with the square root of the number of months, so 136 months in place of 1,065 multiply it by 2.80, and the volatility-managed comparisons' half-widths are 0.53 to 0.89, annualised, here, against 0.21 to 0.44 on Moreira and Muir's sample. These 136 months therefore cannot tell the differences from zero: an interval of −1.12 to 0.66 for the market is consistent with a negative difference and with a positive one of the size estimated on Moreira and Muir's sample, and the sign of each difference says nothing about another sample.

3.5 What trading the signal costs, and how large the position gets

Table 7 reports the turnover and the alpha gross and net of costs, and the market's alpha survives 10 basis points as E6 asks. The volatility-managed market's position changes by 0.752 of the factor a month on average, against Moreira and Muir's 0.73. At 10 basis points per unit of turnover its alpha is 4.03 per cent a year with a Newey-West standard error of 1.74 (t 2.31), against the paper's 3.98, and at 20 basis points it is 3.13 with a standard error of 1.73 (t 1.81); the break-even cost is 55 basis points, against 56. Each basis point of cost takes 12 times the average turnover, 9.0 basis points a year for the market, from the alpha. Momentum's alpha is 11.70 per cent a year net of 10 basis points (t 5.82) and 10.56 net of 20 (t 5.37), with a break-even of 112 basis points; RMW's is 2.06 (t 2.09) and 1.32 (t 1.35); those of SMB, HML and CMA lie within two standard errors of zero gross and net.

Table 7. Turnover and the alpha gross and net of costs on Moreira and Muir's sample: the volatility-managed factor against its original

		Gross		*Net of 10 bp*		*Net of 20 bp*		
Factor	*Turnover, a month*	*Alpha, % a year*	*t*	*Alpha*	*t*	*Alpha*	*t*	*Break-even, bp*
Mkt-RF	0.75	4.93	2.80	4.03	2.31	3.13	1.81	55
SMB	0.55	−0.59	−0.45	−1.25	−0.94	−1.92	−1.42	none
HML	0.69	1.73	1.36	0.90	0.70	0.07	0.06	21
Mom	0.95	12.83	6.24	11.70	5.82	10.56	5.37	112
RMW	0.62	2.80	2.81	2.06	2.09	1.32	1.35	38
CMA	0.57	0.36	0.66	−0.32	−0.60	−1.01	−1.86	5

Months as in Table 1. Turnover is the absolute change in the position from the month before, the first month carrying none; the cost is charged on it, per unit, and the factor's own rebalancing, the same for the managed and the original factor, is left out. t-statistics with Newey-West standard errors at 12 lags. The break-even cost is the gross alpha over 12 times the average turnover; SMB's gross alpha is negative and has none. Moreira and Muir's Table IV gives the market 0.73, 3.98 per cent net of 10 basis points and 56. Source: results/202609/costs_mm_sample_202609.csv, notebook 08.

Table 8 reports the size of the position, and every percentile holds its band, so E6 is met in all 17 of its comparisons, the market's net alpha and 16 percentiles. The market's median position on Moreira and Muir's sample is 0.94 and its 99th percentile 7.27, against 0.93 and 6.39; on Cederburg et al.'s sample every factor's median lies within 0.03 of Panel D and every 99th percentile within 0.48. The largest position, 16.8 for the market in March 1964, followed the market's calmest month of the sample, February 1964. For the market, a position of 7 holds seven times the capital in stocks, six of them borrowed at the bill rate, and a mandate that capped the position at 2 would have bound in 183 of the 1,065 months; the capped strategy is a different one, with its own alpha, which this note does not estimate.

Table 8, Panel A. The size of the volatility-managed position, published beside computed: percentiles of c over last month's realised variance

	Median		*75th / 90th percentile*		*99th percentile*		
Factor and sample	*Published*	*Computed*	*Published*	*Computed*	*Published*	*Computed*	*Largest*
Market, Moreira and Muir's sample	0.93	0.94	1.59 / 2.64	1.62 / 2.71	6.39	7.27	16.8
Mkt-RF, Cederburg et al.'s sample	0.96	0.96			6.47	6.00	13.1
SMB, Cederburg et al.'s sample	0.81	0.79			5.07	5.08	8.7
HML, Cederburg et al.'s sample	1.02	1.00			5.89	5.60	8.9
Mom, Cederburg et al.'s sample	1.01	1.01			8.64	8.32	16.2
RMW, Cederburg et al.'s sample	1.11	1.09			5.02	5.27	8.7
CMA, Cederburg et al.'s sample	0.97	0.94			4.56	4.23	8.7

Moreira and Muir's Table V, Panel A, for the market, with their definition of the realised variance; Cederburg et al.'s Table 1, Panel D, for the six factors, with theirs. E6 asks the median, 75th and 90th percentiles within 0.15 and the 99th within 1.0; all 16 hold. Source: results/202609/position_market_mm_sample_202609.csv and position_cowy_sample_202609.csv, notebook 08.

Table 8, Panel B. The real-time combination's position, its first month to August 2026: percentiles of S1's total position

			Percentile				*Share of months, %*		
Factor	*First month*	*Months*	*5th*	*Median*	*95th*	*Largest*	*Below 0*	*At the cap of 5*	*S2, median*
Mkt-RF	August 1936	1081	0.19	0.56	1.78	5.00	0.8	0.3	0.40
SMB	August 1936	1081	−0.11	0.43	1.45	4.32	6.4	0.0	0.41
HML	August 1936	1081	0.19	0.75	2.08	5.00	2.6	0.1	0.58
Mom	January 1937	1076	0.17	1.07	4.82	5.00	0.5	4.6	0.56
RMW	August 1973	637	−0.57	0.89	3.65	5.00	13.8	2.0	1.05
CMA	August 1973	637	1.01	1.57	1.99	2.47	0.6	0.0	1.57

S1's position is its total holding of the factor, the weight on the volatility-managed factor times its position plus the weight on the original, capped at 5 in either direction. Its first month is 120 months after the factor's first month. No verdict. Source: results/202609/position_real_time_202609.csv, notebook 02.

The real-time combination's position is smaller and moves less, as Panel B of Table 8 reports, in no verdict. From its first month to August 2026, S1's median position runs from 0.43 for SMB to 1.57 for CMA and its 95th percentile from 1.45 to 4.82. It reaches the cap of 5 in at most 4.6 per cent of the months, for momentum, and it is below zero, a short holding of the factor, in 13.8 per cent of RMW's months and 6.4 per cent of SMB's. For the market the real-time original S2 has a median position of 0.40, against S1's 0.56.

Table 9. The six expectations, written before any run that judged them, and what the run found

No.	*Expected (src/sq/constants.py, 8 and 10 November 2026)*	*Found, vintage 202609*	*Verdict*
E1	On Moreira and Muir's sample, for at least five of the six factors: the alpha's sign and 5 per cent significance as in Table I, the alpha within one published standard error, the appraisal ratio within 0.05; for the market, the volatility-managed Sharpe ratio within 0.05 of 0.52	6 of 6 (Table 1)	met
E2	On Cederburg et al.'s sample, for each factor: (i) the original and volatility-managed Sharpe ratios within 0.05 of their Table 1; (ii) S1 and S2 within 0.05 of their Table 5; (iii) S1 − S2 with the published sign on at least five of six, momentum's the only significant difference; (iv) S1 − S3 within 0.05	22 of 24: (i) and (ii) fail for RMW alone, which French has revised since the paper; (iii) and (iv) hold (Tables 3 and 4)	not met
E3	On May 1927 to December 2011, the Sharpe ratios of WML and scaled WML within 0.05 of 0.53 and 0.97, their maximum drawdowns within five points of 96.69 and 45.20 per cent	4 of 4 (Table 2)	met
E4	The size check rejects 3.5 to 6.5 per cent of 1,000 simulated pairs with equal Sharpe ratios; on Moreira and Muir's sample the bootstrap agrees with the alpha's Newey-West t-statistic in sign and significance for at least five of six factors	3.7 per cent; 2 of 6, the alpha and the Sharpe ratio difference being different tests; 4 of 6 by category, in no verdict (Table 5)	not met
E5	On May 2015 to August 2026, S1's Sharpe ratio exceeds S2's by more than the interval allows on at most two factors, and the median difference lies within the median half-width	0 of 6; +0.007 within 0.213; the volatility-managed factor, in no verdict, 0 of 6 and −0.167 within 0.672 (Table 6)	met
E6	The market's alpha net of 10 basis points positive and more than two Newey-West standard errors from zero; the market's position percentiles within 0.15 of Table V (99th within 1.0); each factor's median within 0.15 and 99th percentile within 1.0 of Cederburg et al.'s Panel D	4.03 per cent a year, t 2.31, and 3.13, t 1.81, at 20 basis points, which E6 does not judge; 17 of 17 (Tables 7 and 8)	met

E4 is dated 8 November 2026 and the others 10 November 2026; no level moved after a run. Source: results/202609/e1_verdict_202609.csv to e6_verdict_202609.csv, notebooks 03 to 08.

4. Discussion

The natural comparison is with the two papers' own conclusions. On Moreira and Muir's sample the alphas reproduce for every factor, and the market's survives 10 basis points per unit of turnover at 2.31 standard errors, as the paper reports. Cederburg et al.'s conclusion reproduces as well: an investor who estimated the combination in real time had a Sharpe ratio below the real-time original's for four of six factors, by 0.04 to 0.07, above it for momentum by 0.41, the one significant difference, and above it for RMW by 0.18. Ledoit and Wolf's test adds the reading that separates the two papers: held in the original's place, the volatility-managed factor has a higher Sharpe ratio beyond its interval for momentum alone, so an alpha of more than two standard errors is a statement about adding the managed factor to the original, at weights fixed on the whole sample, and the evidence for holding it alone is confined to momentum on this sample.

What the months after the sample imply is therefore narrow. Neither construction shows a Sharpe ratio difference beyond its interval on any factor, and the intervals are wide enough to hold the paper-sample estimates: the market's volatility-managed difference on Moreira and Muir's sample, +0.097, lies inside its after-sample interval of −1.120 to 0.659, as does momentum's, +0.514, inside −0.449 to 0.803. These months are consistent with the signal doing after 2015 what it did before, and with it doing less; they cannot tell the two apart.

At the same time, the design has limits, each with its cause. First, the costs are charged on the change in the position of a factor, and the factors are long-short portfolios whose own trading in the stocks that make them up is left out because it is the same for the managed and the original factor; a book that holds the stocks pays more, and the break-even costs bound what it could pay. Second, French revises history, so a rerun on a later vintage need not reproduce every digit, as RMW shows; every table names its vintage, and the archived copy shows the size of one revision. Third, the block length of 12 months and the Parzen kernel were fixed before the run, where Ledoit and Wolf calibrate the block length from the data and propose the QS kernel. Fourth, c for the months after the sample is fixed on the months to April 2015, the paper's own constant, so the volatility-managed comparison after the sample asks what the paper's construction did next, and the real-time combination asks what an investor estimating it would have done. Finally, the position is never

capped for the volatility-managed factor, as in the paper, so its largest positions, up to 16.8, are positions no mandate would allow.

What would settle the question for the months after the sample is more of them. The half-width falls with the square root of the months, so halving the market's S1 minus S2 half-width of 0.639 would require four times the 136 months, about 45 years.

Against this evidence the note supports a statement about the signal on French's six factors over these samples. Scaled by last month's realised variance, the market, momentum and RMW earn alphas against their originals of more than two Newey-West standard errors on Moreira and Muir's sample, the market's net of 10 basis points at 2.31 standard errors and net of 20 at 1.81; held in the original's place, the volatility-managed factor has a higher Sharpe ratio beyond its interval for momentum alone; estimated in real time, the combination's Sharpe ratio is below the real-time original's for four of six factors on Cederburg et al.'s sample; and the 136 months since the sample cannot tell any of these differences from zero.

References

Andrews, D. W. K. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica*, 59(3), 817-858.

Barroso, P. and Santa-Clara, P. (2015). Momentum has its moments. *Journal of Financial Economics*, 116(1), 111-120.

Carhart, M. M. (1997). On persistence in mutual fund performance. *Journal of Finance*, 52(1), 57-82.

Cederburg, S., O'Doherty, M. S., Wang, F. and Yan, X. S. (2020). On the performance of volatility-managed portfolios. *Journal of Financial Economics*, 138(1), 95-117.

Fama, E. F. and French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, 116(1), 1-22.

French, K. R. (2026). Data library. mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html, files built from the 202609 CRSP database.

Jobson, J. D. and Korkie, B. M. (1981). Performance hypothesis testing with the Sharpe and Treynor measures. *Journal of Finance*, 36(4), 889-908.

Ledoit, O. and Wolf, M. (2008). Robust performance hypothesis testing with the Sharpe ratio. *Journal of Empirical Finance*, 15(5), 850-859.

Memmel, C. (2003). Performance hypothesis testing with the Sharpe ratio. *Finance Letters*, 1(1), 21-23.

Moreira, A. and Muir, T. (2017). Volatility-managed portfolios. *Journal of Finance*, 72(4), 1611-1644.

Newey, W. K. and West, K. D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55(3), 703-708.

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica*, 48(4), 817-838.