Semivolatility-managed portfolios*

Daniel Batista[†1,2] and Marcelo Fernandes[‡3]

[1] Université de Genève

[2] Swiss Finance Institute

[3] São Paulo School of Economics, FGV

January, 2026

Abstract

There is ample evidence that volatility management helps improve the risk-adjusted performance of momentum portfolios. However, it is less clear that it works for other factors and anomaly portfolios. We show that controlling by the upside and downside components of volatility yields more robust risk-adjusted performances across a broad set of factors and anomaly portfolios, as well as exchange-traded funds. In particular, we propose semivolatility-managed portfolios that, apart from deleveraging when downside volatility is high, exploit the higher expected returns in times of good volatility. We also study the asymptotic behavior of a Sortino ratio estimator to develop a test for direct comparison between two given portfolios. We find that our semivolatility-managed portfolios that control for both skewness and downside volatility perform better than the original portfolios and extant (semi)volatility management proposals.

* The authors thank Federico Baldi Lanfranchi, Daniele Bianchi, and Miloš Božović for excellent discussions. They are also indebted to seminar and conference participants at Northwestern University, University of Wisconsin-Madison, University of Liverpool, UCB, Insper, IMECC/Unicamp, UFPB, UFRGS, UFSC, as well as at the High-Voltage Econometrics Workshop (Zurich, 2022), Advances in Financial Econometrics (Copenhagen Business School, 2023), FMA (Aalborg 2023), IFABS (Oxford 2023), FinEML (Rotterdam, 2023), SoFiE (Seoul, 2023), SBE Meetings (Rio de Janeiro, 2023), SFI Research Days (Gerzernsee, 2024), IAAE (Thessaloniki, 2024), and LubraFin (Cabo de Santo Agostinho, 2024) for their helpful comments. Fernandes acknowledges financial support from FAPESP (2023/01728-0) and CNPq (302278/2018-4 and 302153/2022-5). Batista acknowledges financial support from the Swiss National Science Foundation (grant no. 215573). The usual disclaimer applies.

†E-mail: daniel.batistadasilva@unige.ch

‡E-mail: marcelo.fernandes@fgv.br

1 Introduction

There is a recent debate on whether volatility management leads to higher risk-adjusted performance and utility gains for the investor (Barroso and Santa-Clara, 2015; Moreira and Muir, 2017; Liu et al., 2019; Cederburg et al., 2020; Eisdorfer and Misirli, 2020; Barroso and Detzel, 2021). The motivation for such a strategy is simple. Volatility is more persistent than expected returns, and hence timing risk exposure makes more sense than expected returns to boost the return-to-variability ratio. This should work as long as past volatility does not positively correlate with future returns (Ai et al., 2022).

In view that asset returns are not Gaussian, investors should perhaps beware more of downside risk (Roy, 1952; Markowitz, 1959; Sortino and van der Meer, 1991; Ang et al., 2006). Accordingly, we offer a framework that relies on semivolatility management in order to reflect how investors perceive risks. Wang and Yan (2021) take a similar route. They find that scaling by downside volatility yields significantly better performance relative to scaling by volatility because the former is relatively better at predicting future returns. We complement their analysis by showing how to exploit upside risks, in line with Zeisberger's (2023) findings on investors' risk perception. Despite arguing that investors primarily focus on the probability of loss rather than volatility alone, Zeisberger finds that incorporating upside risk in risk management can positively influence investment decisions by providing a more comprehensive view of potential outcomes.

Our approach rests on realized partial second moments, allowing us to control not only for downside volatility, but also for tail risk. In particular, we consider three partial second moments: downside volatility (or negative semivolatility), upside volatility (or positive semivolatility), and partial volatility below a certain quantile in the tail. Downside volatility has a long pedigree in finance (Roy, 1952; Markowitz, 1959). See, among others, Markowitz et al. (1993), Ballestero (2005), and Estrada (2007) for recent applications in portfolio allocation. There is also a more recent literature on decomposing realized volatility into their downside and upside components (Barndorff-Nielsen et al., 2008; Bollerslev et al., 2020b, 2022b). For instance, although Bollerslev et al. (2020a) find that downside and upside variances behave very differently over time, they both help predict future realized volatility. As for finer decompositions based on partial second moments, Bollerslev et al. (2022a) show how to employ machine-learning methods to choose quantiles that maximize the out-of-sample forecast performance of time series models based on realized partial (co)variances.

By combining realized partial variations, we are able to control for higher moments. For instance, the downside-to-upside volatility ratio reflects skewness, whereas the partial volatility below a low quantile (say, first quartile) gauges tail thickness, and hence to some extent downside kurtosis. Accordingly, apart from scaling portfolios by the total variance as in Moreira and Muir (2017), we propose scaling by the ratio of the downside to upside variance, the ratio of the downside variance to upside *volatility*, and by the product between the downside-to-upside variance ratio and the partial variance below the first quartile. Our first proposal exploits upside risk, apart from controlling for downside risk. It does not affect much the investment position in times of conditional symmetry, though. The second scaling strategy handles this issue by using upside volatility rather than variance. Even if the downside and upside volatilities coincide, risk exposure changes because of the extra downside volatility. The same happens in our third scaling scheme, which focuses more on the tail of the conditional distribution. This means that our second and third proposals control to some extent for higher-order moments: skewness and downside variance in the second, whereas skewness and downside kurtosis in the third.

We empirically assess our semivolatility-managed portfolios using a very broad set of test strategies. Apart from factors and anomaly portfolios as in Barroso and Santa-Clara (2015), Moreira and Muir (2017), Liu et al. (2019), Cederburg et al. (2020), Barroso and Detzel (2021), and Wang and Yan (2021), we also entertain exchange-traded funds (ETFs) mainly for two reasons. First, they are directly tradeable, in contrast to factors and anomaly portfolios. Second, we can obtain very precise realized estimates of partial variances using ETF returns at the 1-minute frequency.

To carry out a robust assessment, we account for the three main points that Cederburg et al. (2020) raise against Moreira and Muir's (2017) empirical methodology. First, they argue that estimating the scaling constant using the full sample, as in Moreira and Muir (2017), induces a look-ahead bias. In principle, this is just a minor point because the scaling constant does not affect risk-adjusted performance measures. However, the optimal weight between the managed and unmanaged portfolios that expands the mean-variance frontier depends on the alpha estimate (i.e., intercept of the spanning regression).[1] The latter obviously depends on the full sample, casting doubt on the genuine ability of spanning regressions to distinguish performance. We tackle this issue by employing only real-time information in the estimation of the scaling constant. Perhaps surprisingly, this actually improves the performance gains of managed portfolios. Second, volatility-management strategy can reach impracticable leverage levels. To make sure our managed portfolios are feasible, we truncate scaling to two. For the sake of comparison, we discuss the performance of managed portfolios both with and without leverage restriction in our empirical analysis.

Third, spanning regressions do not offer a direct comparison between risk-adjusted performances. As such, we do not rely exclusively on spanning regression tests to check whether semivolatility management entail performance gains relative to unmanaged and volatility-managed portfolios. Instead, we provide direct comparisons of Sharpe ratios and, due to our interest in downside risk, of Sortino ratios. Unlike Cederburg et al. (2020),

[1] We follow the literature in referring to the original portfolios as unmanaged, even though they are actually characteristic-managed portfolios.

we do not rely on Jobson and Korkie's (1981) JK test given it unrealistically requires iid Gaussian returns. Alternatively, we evaluate whether (semi)volatility management improves Sharpe ratios using Ledoit and Wolf's (2008) bootstrap-based test, which allows for serial correlation, conditional heteroskedasticity, and heavy tails. In addition, we also extend their test to deal with Sortino ratios as well. Apart from assessing each managed portfolio individually, we also carry out a joint assessment by testing for stochastic dominance in the cross-section distribution of Sharpe and Sortino ratios of both managed and unmanaged portfolios.

We find that our semivolatility-managed proposal that controls for both downside variance and skewness offers more robust risk-adjusted performance gains than the extant methods in the literature. In particular, it yields significant gains in both Sharpe and Sortino ratios, especially for the worst-performing factors and anomaly portfolios. Scaling by the downside-to-upside volatility ratio has little effect for the bulk of the factors and anomaly portfolios, though. Controlling for higher-order partial moments pay off across the board, as opposed to volatility management that helps mostly the best-performing factors and anomaly portfolios.

These performance gains stem chiefly from the negative/positive correlation between past downside/upside semivolatility and future returns, in line with Wang and Yan (2021). Predictive regressions show that, in general, decomposing risk into upside and downside volatilities yields more predictive power than controlling exclusively for total volatility. For instance, simple regressions reveal that market portfolio returns do not seem to react much either to total volatility or to downside volatility. However, once we add upside volatility to the equation, we observe that downside pushes future returns down inasmuch as upside volatility push them up, both in a significant manner.

We also discuss how the pricing errors of the Fama-French three- and five-factor models change as we move from unmanaged to managed factors. To do so, we employ Barillas et al.'s (2020) procedure for nonnested factor model comparisons. We find that, using the three-factor model, semivolatility management pays off at the 5% significance level if we time either downside volatility or both volatility and skewness. For the five-factor model, scaling by downside volatility yields Sharpe ratio improvements at the 5% level, whereas managing both volatility and skewness significantly ameliorates the Sharpe ratio at the 1% level. In particular, the maximum attainable Sharpe ratio of the Fama-French five-factor model increases from 1.08 to 1.38 once we replace unmanaged factors by our semivolatility-managed factor that controls for both volatility and skewness.

Finally, it is important to understand through which channels (semi)volatility management generates performance gains. Following Wang and Yan (2021), we decompose spanning-regression alphas into their return-timing (RT) and volatility-timing (VT) components. While volatility timing captures gains arising from reducing exposure during periods of heightened risk, return timing reflects the ability to increase exposure when expected returns are high. This decomposition is particularly informative in our setting, as different timing strategies may achieve similar overall alphas while relying on distinct economic mechanisms. By examining the relative contributions of RT and VT, we shed light on whether performance improvements stem primarily from exploiting upside return opportunities, mitigating downside risk, or a combination of both, thereby providing a clearer interpretation of how semivolatility management operates across factors, anomalies, and exchange-traded funds.

Our study on (semi)volatility management is part of a larger literature on factor timing. Moreira and Muir (2017) claim that investors can improve Sharpe ratios by decreasing exposure to risk factors when their volatility is high. Barroso and Detzel (2021) show however that these gains disappear in times of low sentiment once we account for transaction costs. Bianchi et al. (2022) show how to improve the volatility-managed momentum portfolio by controlling for conditional skewness. DeMiguel et al. (2024) entertain weights on multifactor portfolios that depend on the market volatility, instead of focusing on individual portfolios. They show that volatility-management of the multifactor portfolio weights outperforms, net of transaction costs, regardless of whether sentiment is high. Volatility is not the unique timing variable in the literature, though. For instance, Haddad et al. (2020) time risk factors using the book-to-market spread of the principal components of a large panel of equity factors, whereas Bass et al. (2017), Amenc et al. (2019), Bender et al. (2019), and Gómez-Cram (2022) condition risk premia on different macroeconomic indicators.

The remainder of the paper proceeds as follows. Section 2 describes the data and methodology, whereas Section 3 discusses whether (semi)volatility management indeed yields risk-adjusted performance gains across different factors, anomaly portfolios, and exchange-traded funds. Section 4 concludes.

2 Data and methodology

In this section, we first describe the data and then discuss the extant volatility and semivolatility management strategies in the literature, before proposing novel schemes based on partial variances. Next, we go through our assessment methodology. Apart from spanning regressions as in Moreira and Muir (2017) and the JK test as in Cederburg et al. (2020), we also entertain Ledoit and Wolf's (2008) LW test for Sharpe ratio differentials as well as Barillas et al.'s (2020) cross-section relative test for nonnested factor models. Finally, given our interest in downside risk, we also extend the LW test to compare Sortino ratios, assessing size and power by means of Monte Carlo simulations.

2.1 Data description

We employ a wide array of test assets that fall in three groups. The first set comprises ten equity factors: market (MKT), size (SMB), and value (HML) from Fama and French (1993); momentum (MOM) from Carhart (1997); betting against beta (BAB) from Frazzini and Pedersen (2014); investment (CMA and IA) and profitability (RMW and ROE) from Fama and French (2015) and Hou et al. (2015); and expected growth (EG) from Hou et al. (2021). Daily data on MKT, SMB, HML, MOM, RMW, and CMA returns are from `http://mba.tuck.dartmouth.edu/pages/faculty/ken.french`, whereas BAB returns are available at `http://people.stern.nyu.edu/afrazzin` and IA, ROE and EG data at `http://global-q.org/factors.html`. This is exactly the factors that Moreira and Muir (2017) consider, except for the EG addition. Table 1 reports their descriptive statistics. The EG factor stands out for displaying the highest average return, with the lowest volatility. Indeed, the other factors with highest average returns — namely, MKT, MOM and BAB — also exhibit high volatility and semivolatility. The SMB, HML and MOM factors entail the most asymmetric returns, with extremely heavy tails.

The second refers to a large panel of anomaly portfolios that helps characterize the cross-section of expected returns (Kozak et al., 2020). We extract daily returns on 207 anomaly portfolios from Chen and Zimmermann's (2022) database, available at `https://www.openassetpricing.com`. The sample period varies across anomaly portfolios. The longest time spans are for the *price* and *spinoff* anomaly portfolios (Blume and Husic, 1973; Cusatis et al., 1993), with 25,297 daily observations from January 2, 1926 to December 31, 2021. The shortest samples are for *deferred revenues* (Prakash and Sinha, 2013) , *takeover vulnerability* and *active shareholders* (Cremers and Nair, 2005), with 4,137 daily observations from October 1st, 1990 to February 28, 2007.

Figure 1 exhibits box plots for the main descriptive statistics of the 207 anomaly portfolios. Average returns are typically between 3% and 8% per year, with very few anomaly portfolios displaying negative returns. The exceptions are *governance index* (Gompers et al., 2003); *firm age* (Barry and Brown, 1984); *pension funding status* (Franzoni and Marin, 2006); *sales growth over overhead growth* (Abarbanell and Bushee, 1998); and *cash-flow-to-price variance* (Haugen and Baker, 1996). The volatility of most anomaly portfolios range from 8% to 16% per year, with downside semivolatility between 5% and 11%. Anomaly portfolios typically display a large amount of excess kurtosis, whereas we observe a lot of dispersion in skewness.

The third group of test assets consists of 71 exchange-traded funds (ETFs), which we list in Table 2. We only entertain ETFs that trade for a period longer than 5 years, with an average of at least 300 observations per day (out of a maximum of 390). We collect their 1-minute returns from `https://pitrading.com`. Access to intraday data helps alleviate concerns about the precision of the monthly realized partial moments. By computing monthly realized measures using 1-minute returns, we ensure small-sample bias does not drive the results.

Figure 2 displays box plots for the main descriptive statistics of the 71 exchange-traded funds. There is a lot of dispersion in average returns, with the bulk of the cross-section distribution around 5% per year. There are many ETFs with very negative mean returns, though. Consistently, we also observe many ETFs with very negative skewness. Both volatility and downside semivolatility are typically higher for ETFs than for factors and anomaly portfolios, ranging mostly from 20% to 40% and from 15% to 30%, respectively. ETF returns are also very far from Gaussian, exhibiting both negative skewness and excess kurtosis.

2.2 Timing downside and upside risks

Moreira and Muir (2017) claim that targeting variance yields slightly better performance than timing volatility. As such, they entertain the following strategy:

$$f_{(\sigma),t} = \frac{c_{(\sigma)}}{\hat{\sigma}^2_{t-1}} f_t, \tag{1}$$

where $\hat{\sigma}^2_{t-1}$ is the realized variance at month $t-1$ and f_t is the excess return on the buy-and-hold anomaly portfolio. The scaling constant $c_{(\sigma)}$ standardize the managed portfolio to match the unconditional variance of the unmanaged portfolio.

Wang and Yan (2021) argue that managing downside risk should perform better than total risk because it displays a relatively stronger negative correlation with future returns. In particular, they time downside risk by restricting attention to the negative semivolatility. To keep in line with the seminal papers in the literature, we instead use the negative semivariance:[2]

$$f_{(-),t} = \frac{c_{(-)}}{\hat{\sigma}^2_{(-),t-1}} f_t, \tag{2}$$

where $\hat{\sigma}^2_{(-),t} = \frac{1}{T}\sum_{d=1}^{T} \min(0, r_{d,t})^2$ and $r_{d,t}$ are returns at day d (or minute in the ETF case) within month t. The scaling constant $c_{(-)}$ now standardizes the managed portfolio to match unconditional negative semivariance of the unmanaged portfolio.

Next, we introduce our timing strategies based on partial second moments that account to some extent for higher-order moments. The first attempts to control for both downside and upside risks, deleveraging if the former is high, while scaling up if the latter is high. In particular, we scale by the downside-to-upside variance ratio:

$$f_{(+/-),t} = c_{(+/-)} \frac{\hat{\sigma}^2_{(+),t-1}}{\hat{\sigma}^2_{(-),t-1}} f_t, \tag{3}$$

[2] The empirical findings are quantitatively very similar if we replace the semivariance with semivolatility, as in the original timing strategy of Wang and Yan (2021). The results are available from the authors upon request.

where $\hat{\sigma}^2_{(+),t} = \frac{1}{T}\sum_{d=1}^{T} \max(0, r_{d,t})^2$ and $c_{(+/-)}$ ensures the unconditional downside variance remains the same, as before. A problem with (3) is that, under distributional symmetry, the positive and negative semivariances coincide, keeping the portfolio unchanged irrespective of the volatility levels. As such, it essentially times skewness, scaling up or down according to whether symmetry is positive or negative.

Our second proposal deals with the symmetry issue by replacing the upside variance by the upside volatility in the numerator:

$$f_{(\sqrt{+}/-),t} = c_{(\sqrt{+}/-)} \, \frac{\hat{\sigma}_{(+),t-1}}{\hat{\sigma}^2_{(-),t-1}} \, f_t \tag{4}$$

with $c_{(\sqrt{+}/-)}$ denoting the scalar that keeps constant the unconditional downside semivariance. To understand exactly which risks we control with (4), we decompose the reciprocal of the investment position into the product of the semivolatility ratio $\hat{\sigma}_{(-)}/\hat{\sigma}_{(+)}$ and the downside volatility $\hat{\sigma}_{(-)}$. The downside-to-upside volatility ratio accounts for skewness, exploiting upside risk when realized symmetry is positive and deleveraging when it is negative. At the same time, the second term in the decomposition controls for (downside) volatility, regardless of the realized skewness.

Our last strategy aims to control for tail risk, by contemplating the partial variance below the first quartile: namely, $\hat{\sigma}^2_{(q),t} = \frac{1}{T}\sum_{d=1}^{T} \mathbb{1}\{r_{d,t} < q\} r^2_{d,t}$, where q denotes the first quartile of the empirical distribution of the returns at the highest frequency (Bollerslev et al., 2022a). It is obviously much harder to estimate partial moments at the tails. This is why we restrict attention to the first quartile and exclusively to exchange-traded funds, for which we can estimate monthly partial variances very precisely using 1-minute returns.[3] This results in managed portfolios of the form

$$f_{(q),t} = \frac{c_{(q)}}{\hat{\sigma}^2_{(q),t-1}} \, \frac{\hat{\sigma}^2_{(+),t-1}}{\hat{\sigma}^2_{(-),t-1}} \, f_t \tag{5}$$

with $c_{(q)}$ denoting the scaling constant that ensures that the managed and unmanaged portfolios have the same unconditional negative semivariance. As before, while the downside-to-upside volatility ratio scales leverage up and down according to whether skewness is positive or negative, the first fraction ensures that we deleverage if there is too much mass on the negative tail. This implies that (5) controls to some extent for both skewness and downside kurtosis.

In the empirical application, we consider several configurations for the above timing strategies, in order to take Liu et al.'s (2019) and Cederburg et al.'s (2020) critiques into account. We indeed fix the scaling constant using only real-time information and do not allow leverage to run amok by constraining their magnitude to at most two. The

[3] Our findings are nonetheless similar for values of q between the 5th and 25th percentiles. The results are available from the authors upon request.

aim is to mitigate concerns about unrealistic investment positions in practice. Although the value of the scaling constant should not affect much Sharpe and Sortino ratios given that they do not depend on the unconditional (semi)variance of the portfolio returns, it becomes relevant once we constrain leverage. By using exclusively real-time information, our estimate of the unconditional (semi)variance of the unmanaged portfolio becomes more local, helping alleviate leverage constraints (and hence transaction costs).

2.3 Assessing performance gains

Our interest lies on understanding whether timing risk leads to investors' utility gains. As such, we check whether (semi)volatility-managed portfolios indeed outperform their unmanaged counterparts. To do so, we not only employ the testing approaches previously used in the literature, but also conduct some additional analyses.

The first test rests on spanning regressions, as in Moreira and Muir (2017). We regress returns to the managed portfolio on returns to the unmanaged portfolio:

$$f_{(\cdot),t} = \alpha + \beta\, f_t + u_t, \tag{6}$$

where u_t is a generic white noise. If α differs from zero in (6), the managed strategy helps expand the mean-variance frontier. This means there is a linear combination of the managed and unmanaged portfolios that achieves a higher Sharpe ratio than before. The spanning regression test checks whether there is enough statistical evidence to reject the null hypothesis $\mathbb{H}_0 : \alpha = 0$ in favor of the alternative hypothesis $\mathbb{H}_1 : \alpha \neq 0$. This does not boil down to a direct comparison of Sharpe ratios, though. Cederburg et al. (2020) show that a positive alpha in the spanning regression does not necessarily lead to Sharpe Ratio improvements.

The second set of tests explicitly compares the risk-adjusted performances of managed and unmanaged portfolios. Cederburg et al. (2020) employ Jobson and Korkie's (1981) test to assess whether managed portfolios improve on the Sharpe ratio of unmanaged portfolios. Let the returns on strategies i and j be stationary with mean $\boldsymbol{\mu} = (\mu_i, \mu_j)'$ and covariance matrix

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_i^2 & \sigma_{ij} \\ \sigma_{ij} & \sigma_j^2 \end{pmatrix}.$$

The interest is in the difference between their Sharpe ratios, namely, $\Delta = \mathrm{SR}_i - \mathrm{SR}_j$, with $\mathrm{SR} = \mu/\sigma$. We estimate both means and variances by their realized counterparts.

Let $\boldsymbol{\theta} = (\mu_i, \mu_j, \sigma_i^2, \sigma_j^2)'$. Under the assumption of iid Gaussian returns, Memmel (2003) first establishes that $\sqrt{T}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \xrightarrow{d} N(0, \boldsymbol{\Omega})$, with

$$\boldsymbol{\Omega} = \begin{pmatrix} \sigma_i^2 & \sigma_{ij} & 0 & 0 \\ \sigma_{ij} & \sigma_j^2 & 0 & 0 \\ 0 & 0 & 2\sigma_i^4 & 2\sigma_{ij}^2 \\ 0 & 0 & 2\sigma_{ij}^2 & 2\sigma_j^4 \end{pmatrix},$$

and then obtains the asymptotic variance of $\widehat{\Delta}$ by the delta method. The main problem with the JK test is that Memmel's derivation assumes iid Gaussian returns. However, if (semi)volatility changes over time in a persistent fashion, returns are neither iid, nor unconditionally Gaussian. Ledoit and Wolf (2008) derive a bootstrap version that allows not only for serial correlation and conditional heteroskedasticity in the returns, but also heavy tails. Accordingly, although we also report the JK test, we rely mostly on the Ledoit-Wolf (LW) test in our empirical assessment.

Given our interest in downside risk, we extend both JK and LW tests for Sortino ratio comparisons. The interest now lies on $\Delta_{(-)} = \mathrm{SR}_{(-),i} - \mathrm{SR}_{(-),j}$, with $\mathrm{SR}_{(-)} = \mu/\sigma_{(-)}$. As before, we stack the parameters of interest into a vector $\boldsymbol{\theta}_{(-)} = (\mu_i, \mu_j, \sigma^2_{(-),i}, \sigma^2_{(-),j})$ and then estimate both means and semivariances by their realized counterparts. We then establish the asymptotic behavior of $\widehat{\boldsymbol{\Delta}}_{(-)}$ under the assumptions that asset prices follow a Brownian semimartingale process, as in Barndorff-Nielsen et al. (2008).

For iid Gaussian returns, $\sigma_k = \sigma_{k,s}$ is constant over time and $\mathrm{var}(\hat{\sigma}^2_{(-)}) = (5/4)\sigma^4$; see Proposition 2 in Barndorff-Nielsen et al. (2008). In addition, normality also implies that $\mathrm{cov}(\sigma^2_{(\cdot),i}, \sigma^2_{(\cdot),j}) = \mathrm{cov}(\sigma^2_i, \sigma^2_j)/4 = \sigma^2_{ij}/2$, given that $\sigma^2_{(-),k} = \sigma^2_{(+),k} = \sigma^2_k/2$ for $k = i, j$ and $\mathrm{cov}(\sigma^2_i, \sigma^2_j) = \mathrm{cov}(\sigma^2_{(+),i}, \sigma^2_{(+),j}) + \mathrm{cov}(\sigma^2_{(+),i}, \sigma^2_{(-),j}) + \mathrm{cov}(\sigma^2_{(-),i}, \sigma^2_{(+),j}) + \mathrm{cov}(\sigma^2_{(-),i}, \sigma^2_{(-),j})$. As such, the covariance matrix of $\hat{\boldsymbol{\theta}}_{(-)}$ reads

$$\boldsymbol{\Omega}_{(-)} = \begin{pmatrix} \sigma^2_i & \sigma_{ij} & 0 & 0 \\ \sigma_{ij} & \sigma^2_j & 0 & 0 \\ 0 & 0 & (5/4)\sigma^4_i & \sigma^2_{ij}/2 \\ 0 & 0 & \sigma^2_{ij}/2 & (5/4)\sigma^4_j \end{pmatrix}.$$

Applying the delta method then yields that, under iid Gaussian returns, the asymptotic variance of $\widehat{\Delta}_{(-)}$ is $\frac{1}{T}\left(\frac{\sigma^2_i}{\sigma^2_{(-),i}} - \frac{2\sigma_{ij}}{\sigma_{(-),i}\sigma_{(-),j}} + \frac{\sigma^2_j}{\sigma^2_{(-),j}} + \frac{5\mu^2_i\sigma^4_i}{16\sigma^6_{(-),i}} - \frac{\mu_i\mu_j\sigma^2_{ij}}{4\sigma^3_{(-),i}\sigma^3_{(-),j}} + \frac{5\mu^2_j\sigma^4_j}{16\sigma^6_{(-),j}}\right)$. To relax the iid Gaussian assumption, we entertain a block-bootstrap variant of the testing procedure, as in Ledoit and Wolf (2008).

To assess the finite-sample properties of the proposed tests, we conduct a Monte Carlo study with 1,500 replications under both iid and conditionally heteroskedastic returns. Specifically, we simulate bivariate portfolio returns generated by (i) iid Gaussian innovations, (ii) iid skewed and heavy-tailed innovations obtained from a Gaussian mixture distribution, and (iii) conditionally heteroskedastic innovations following a GARCH(1,1) process. For each data-generating process, we vary the sample size ($T = 120$ or 240), the cross-asset correlation ($\rho = 0$ or 0.3), and the mean difference ($\Delta\mu$), so as to study both size ($\Delta\mu = 0$) and power ($\Delta\mu \neq 0$). For the block-bootstrap procedure, we consider $B = 1,500$ artificial samples and a data-driven block length $\ell \approx T^{1/3}$.

Table 3 reports the empirical rejection frequencies of the JK and LW tests. Under iid Gaussian returns, the JK test exhibits rejection rates close to nominal levels across sample sizes and correlation settings, with only mild distortions at the 10% level when

$\rho = 0.3$. Under the heavy-tailed mixture distribution, the JK test becomes conservative, particularly at the 1% significance level. The LW block-bootstrap test substantially mitigates these size distortions under non-Gaussianity and serial dependence. Under GARCH dynamics, at very small nominal levels, the LW test remains conservative, especially in smaller samples. Overall, size distortions are moderate and decrease with the sample size, supporting the reliability of the bootstrap-based inference in empirically relevant settings.

Figure 3 further assesses power by plotting how the rejection frequency at the 5% level changes with $\Delta\mu$. Three main patterns emerge: (i) power increases monotonically with $|\Delta\mu|$; (ii) larger samples substantially steepen the rejection curves; and (iii) the test remains robust for non-Gaussian and dependent returns. The Gaussian case displays the highest power, while the mixture distribution yields somewhat lower power due to inflated downside dispersion. Under GARCH errors, the block-bootstrap test still delivers competitive power, despite the asymmetry in the rejection curves. In particular, it is easier to detect negative shifts in the mean ($\Delta\mu < 0$) because of the interaction between volatility clustering and downside semivariance in the Sortino ratio.

Shifting expected returns downward implies more frequent and larger negative realizations, which lowers the numerator (average return) while inflating the denominator (downside deviation) of the Sortino ratio. These forces act in the same direction, amplifying the signal of underperformance and producing very high rejection frequencies. By contrast, although shifting expected returns upward increases the numerator, the denominator remains largely governed by occasional volatility bursts that generate negative returns. As such, downside semivariance does not shrink proportionally to the mean shift, reducing to some extent the improvement in the Sortino ratio. That explains why we observe less power for $\Delta\mu > 0$ than for $\Delta\mu < 0$, even though both directions imply the same absolute change in mean returns. At any rate, the block-bootstrap test remains valid under the null, with power growing strongly as $|\Delta\mu|$ increases.

Overall, the simulation evidence supports the practical validity of the JK and LW tests of Sortino difference. The former is simple, fast, and well-behaved when returns are approximately iid. The latter provides robust inference under serial dependence, volatility clustering, and heavy tails. Both approaches maintain satisfactory size and power in finite samples, reinforcing their applicability in empirical portfolio analysis.

In addition to the individual Sharpe and Sortino comparisons, we also run a joint analysis by testing for first-order stochastic dominance in the empirical distributions of Sharpe and Sortino ratios of the anomaly portfolios and ETFs. More specifically, we carry out Barrett and Donald's (2003) test to check whether each managed portfolio displays first-order stochastic dominance with respect to their unmanaged counterparts. The plots of the empirical distributions are also very helpful to understand how each (semi)volatility management strategy behaves as we move from the lowest- to highest-performing factors, anomaly portfolios and exchange-traded funds. See, for instance, de Castro et al. (2022)

for a similar quantile-based analysis.

Finally, our last assessment rests on testing the difference in the squared maximum Sharpe ratio that we can attain from two nonnested factor models, in order to determine which model yields smaller pricing errors. Barillas et al. (2020) extend the asymptotic theory of Barillas and Shanken (2017) to accommodate nonnested models and nontraded factors, under the assumption of jointly stationary and ergodic returns with finite fourth moments. For a given set of factors F, the square of the maximum-attainable Sharpe ratio is $\boldsymbol{\mu}_F' \boldsymbol{\Sigma}_F^{-1} \boldsymbol{\mu}_F$, where $\boldsymbol{\mu}_F$ is the vector of average returns and $\boldsymbol{\Sigma}_F$ is the covariance matrix of the factors. The test then gauges the differences in squared Sharpe ratios. In particular, we assess whether the squared maximum-attainable Sharpe ratios of the Fama-French three- and five-factor models significantly increase once we replace the original factors with their managed counterparts.

3 How does each timing strategy fare empirically?

For the sake of brevity, we omit the results for the less realistic settings without leverage restrictions and using the full sample for the estimation of the scaling factor. A very brief summary is as follows. If we do not restrict attention to real-time information, the risk-adjusted performance worsens very slightly across the board. If we keep leverage unchecked, investment positions become impractically large in magnitude, resulting in extremely unrealistic alphas in the spanning regressions.

To estimate the scaling factor using only real-time information, we burn in the first 24 months for initialization purposes. We then expand the estimation window as we move along the sample so as to include the entire past up to that point. In what follows, we discuss our findings for the feasible portfolios that restrict attention only to real-time information and truncate the magnitude of the investment position. In particular, we build managed portfolios by estimating the (monthly) realized partial variances at month t using daily returns on factors and anomaly portfolios and 1-minute ETF returns.

3.1 Factors

The first panel of Table 4 reports the OLS estimates of the spanning-regression alphas for the sample of factor returns, with p-values based on heteroskedasticity-and-autocorrelation consistent (HAC) standard errors. The second and third panels document respectively the differences in Sharpe and Sortino ratios between managed and unmanaged factors, which we assess using both JK and LW tests. There are two clear patterns. First, a statistically nonzero alpha in the spanning regression does not automatically translate into significant changes in risk-adjusted performance. There are many managed factors with positive alphas that entail lower Sharpe ratios than their unmanaged counterparts. Second, it is much harder to find significant evidence of better performance with direct comparisons than with spanning regressions. Along the same lines, the LW bootstrap-based test sets a much higher bar than the JK test based on the iid Gaussian assumption.

As in Moreira and Muir (2017), we find that volatility management yields significantly positive alphas for MKT, HML, MOM, RMW, BAB, IA, ROE, and EG. However, once we focus on the direct comparisons, volatility-managed factors do not perform so well. Timing volatility improves the Sharpe ratio of the momentum factor at the 5% significance level, and to some extent the Sortino ratios of MOM and ROE with p-values only slightly above 5%. Timing downside variance as in Wang and Yan (2021) fails to produce alpha for only two factors in the spanning regressions (CMA and IA). In contrast, we find little evidence that it ameliorates risk-adjusted performance of the factors, apart from RMW and BAB. Timing only skewness has the worst performance in the spanning regressions, failing to produce significantly positive alphas for MOM, CMA, IA and EG. It nonetheless enhances the risk-adjustment performance of the HML, RMW and BAB factors. Timing both skewness and downside volatility seems to pay off, though. It fails to produce alpha only for MOM and CMA, whereas it yields at the 5% significance level higher Sharpe and/or Sortino ratios for HML, RMW, BAB, and ROE.

Altogether, timing both skewness and downside volatility yields the best overall performance for managed factors. Indeed, timing only (partial) second moments does not work so well for MKT, SMB, CMA, IA and EG, whereas only volatility management improves on MOM. Table 5 offers a potential explanation. We find that, in general, decomposing risk into upside and downside volatilities yields more predictive power than controlling only for total volatility. For instance, MKT returns do not apparently respond to total and downside volatilities. However, once we control for both downside and upside volatilities, the picture changes drastically. Both partial effects are significant and with the expected signs in that downside volatility pushes future returns down and upside volatility push them up.

The predictive regressions actually display opposite signs for the partial effects of downside and upside volatilities for every factor but momentum. For the latter, the impact of total volatility on future returns is much stronger than that of downside volatility. This explains why volatility management works better than semivolatility management for the MOM portfolio. Besides, accounting for both components of the total volatility does not enhance much the predictive power (see Figure 4). Although the partial effect of the upside volatility is of a similar magnitude to that of the downside volatility, it is negative rather than positive. As such, it is not surprising that trying to exploit upside risk fails to improve MOM returns.

The predictive regressions also help us understand better the HML and RMW cases. The simple regressions show that their total volatility affects positively their returns, whereas their downside volatility does not have a significant impact. These results help explain why timing either total volatility or downside volatility does not help improve much the performance of the HML and RMW portfolios. In contrast, the partial effects of both downside and upside volatilities are significantly negative and positive, respectively. And this is exactly the situation we expect our semivolatility-management strategy that accounts for both downside and upside risks to work best.

Finally, we explore Wang and Yan's (2021) spanning-alpha decomposition into return- and volatility-timing components: namely, $\alpha = RT + VT$, with

$$RT = \left(1 + \frac{\mathbb{E}^2(f_t)}{\text{var}(f_t)}\right) \text{cov}(w_t, f_t) \;\; \text{and} \;\; VT = -\frac{\mathbb{E}(f_t)}{\text{var}(f_t)} \text{cov}(w_t, f_t^2)$$

for w_t denoting the amount of leverage implied by the timing strategy. Return timing boosts alpha when leverage positively correlates with contemporaneous returns, whilst volatility timing kicks in at times of negative covariance between leverage and the magnitude of factor returns. Figure 5 shows that return- and volatility-timing contribute in a similar amount to spanning-regression alphas for most managed factors. In contrast, return timing responds for most of the alpha in the HML and SMB managed factors, whereas volatility timing is what drives the alpha of the MOM managed factor.

3.2 Anomaly portfolios and ETFs

We now turn our attention to the broader set of test portfolios. Table 6 documents how the different timing strategies perform relative to just buying and holding anomaly portfolios and exchange-traded funds. Spanning regressions yield significantly positive alphas for 142 anomaly portfolios (out of 207) when timing downside risk. This number slightly decreases to 138 if we time both skewness and downside risk, but drops drastically to 110 and 102 if we time only total volatility or skewness, respectively. The direct comparisons of Sharpe and Sortino ratios reveal similar relative performances, with $f_{(-)}$ and $f_{(\sqrt{+}/-)}$ easily outclassing both $f_{(\sigma)}$ and $f_{(+/-)}$. The JK test indicates a very small advantage to timing only downside variance, which improves the Sharpe ratio of 63 anomaly portfolios and the Sortino ratio of 89 portfolios, relative to timing both skewness and downside variance (63 and 86, respectively). The more realistic assessment based on the LW test flips the advantage to timing both downside variance and skewness: 34 against 30 for the Sharpe ratio and 38 against 35 for the Sortino ratio, respectively.

We next examine the performance of timing (partial) second moments for the ETF returns in excess over the 1-month T-bill rate. The motivation is twofold. First, our realized measures of partial variances are much more precise due to the higher frequency of the ETF returns. Second, investors can easily trade ETFs as opposed to the above risk factors and anomaly portfolios. Table 6 reveals however that (semi)volatility management does not perform as well as one would expect given previous findings. Timing both skewness and downside variance yields the largest number of significantly positive alphas (17) and differences in Sharpe and Sortino ratios (17 and 19, respectively). Timing skewness and kurtosis performs at par with $f_{(\sigma)}$ and $f_{(-)}$, whereas timing only skewness remains performing worst.

We complement the above time-series analyses with some cross-sectional tests. Table 7 documents the t-statistics of the difference in the squared maximum Sharpe ratio we can attain from managed and unmanaged Fama-French three- and five-factor models. We compute the t-statistics for nonnested models with nontraded factors, as in Barillas et al. (2020). Although every timing strategy helps reduce pricing errors, and hence improve the maximum-attainable Sharpe ratio, these enhancements are statistically significant only for $f_{(-)}$ and $f_{(\sqrt{+}/-)}$.

Table 8 reports Barrett and Donald's (2003) pairwise tests of first-order stochastic dominance on the cross-section of Sharpe and Sortino ratios. Given that these tests are noncommutative and do not account for local dominance, we check whether the risk-adjusted performances of managed portfolios stochastically dominate those of unmanaged portfolios, and vice-versa. We find no evidence at the usual significance levels that unmanaged portfolios dominates any of the managed portfolios. In contrast, managed portfolios stochastically dominate their unmanaged counterparts at least once. In particular, we find evidence of stochastic dominance in both Sharpe and Sortino ratios for the volatility-managed anomaly portfolios at the 10% significance level and for the volatility-managed ETFs at the 5% significance level.

Timing downside risk with $f_{(-)}$ yield stochastic dominance in both Sharpe and Sortino ratios for anomaly portfolios at the 1% significance level and in Sortino ratio for the ETFs at the 5% significance level (10% if Sharpe ratio). Targeting only skewness with $f_{(+/-)}$ does not perform so well, stochastically dominating in terms of Sortino ratio only the unmanaged anomaly portfolios. Timing both volatility and skewness with $f_{(\sqrt{+}/-)}$ yields the smallest p-values across the board, with stochastic dominance ensuing for every performance measure and for both anomaly portfolios and ETFs at the 5%, if not 1%, significance level. Finally, we also observe stochastic dominance at the 5% significance level when timing skewness and kurtosis with $f_{(q)}$ for ETFs.

Perhaps more important than the stochastic dominance analysis between managed and unmanaged portfolios is to understand which anomaly portfolios and ETFs each timing strategy improves. Figure 6 reports the cross-sectional distribution of Sharpe and Sortino ratios of the managed and unmanaged anomaly portfolios. They reveal that timing only skewness helps only the worst-performing anomaly portfolios, whereas volatility management works best with the top-performing anomaly portfolios. Timing both skewness and downside volatility also enhances the Sharpe ratios of the worst-performing portfolios, while working as well as timing only downside volatility in the bulk of the distribution. In addition, timing risk seems to have a stronger impact on Sortino ratios than on Sharpe ratios. Figure 7 unveils a slightly different pattern for exchange-traded funds. Timing skewness helps only funds with negative Sharpe ratios, whereas the remaining timing strategies perform best for funds with positive Sharpe ratios. Figure 8 documents the expected leftward shift in the empirical distribution of Sharpe ratios once transaction costs are taken into account. Although transaction costs reduce performance across all timing strategies, the relative ordering remains unchanged, and the improvement among ETFs with positive Sharpe ratios is still clearly visible.

Turning to the spanning-regression alphas, the first panel of Figure 9 shows that $f_{(\surd+/-)}$ delivers alpha estimates of similar magnitude to those obtained under volatility and downside-variance management, with a lower cross-sectional dispersion. A comparable pattern is observed for the appraisal ratios in the second panel. Figure 10 decomposes these alphas into return-timing and volatility-timing components and shows that both components contribute to alpha, with substantial heterogeneity across anomaly portfolios. Table 9 reports the average contributions by firm-characteristic category in the Chen and Zimmermann's (2022) database. For anomalies related to fundamentals and expected returns, the return-timing component tends to be larger on average, whereas for anomalies related to risk and market reactions to information, the volatility-timing component is relatively more prominent.

To sum up, timing higher-order (partial) moments yield quite encouraging results, especially if we manage risk exposures to both skewness and downside volatility. Despite methodological differences, this is in line with Bianchi et al.'s (2022) evidence that adjusting for conditional skewness improves the risk-adjusted performance of the volatility-managed MOM portfolio.

3.3 Value of timing skewness

A natural concern is whether our timing strategy based on $f_{(\surd+/-)}$ truly adds value relative to timing downside variance with $f_{(-)}$ as in Wang and Yan (2021). To address this issue, we run pairs of spanning regressions between these managed portfolios, classifying the results as follows: (i) both span such that neither has significant alpha against the other ($f_{(\surd+/-)} \not\gtrless f_{(-)}$); (ii) ours outspans theirs in that ours have significant incremental alpha, but theirs don't ($f_{(\surd+/-)} \succ f_{(-)}$); (iii) theirs outspans ours ($f_{(\surd+/-)} \prec f_{(-)}$); and (iv) neither spans such that both alphas are significant ($f_{(\surd+/-)} \gtrless f_{(-)}$).

We find that over half of the anomaly portfolios (112 out of 207) fall into the *both-span* category $f_{(\surd+/-)} \not\gtrless f_{(-)}$, whereas only 9 into the *neither-span* category $f_{(\surd+/-)} \gtrless f_{(-)}$. In turn, Wang and Yan's (2021) managed portfolios outspans ours relatively more frequently: 48 cases of $f_{(\surd+/-)} \prec f_{(-)}$ versus 38 of $f_{(\surd+/-)} \succ f_{(-)}$. Still, timing skewness apart from downside volatility offers incremental alpha for a nontrivial fraction of anomaly portfolios (viz., 18%). Adjusting for false discovery rate as in Storey (2002) increases the number of *both-span* instances to 152, while shrinking the proportion of times in which one timing strategy dominates the other: 7% for $f_{(\sqrt{\mp}/-)} \succ f_{(-)}$ and 16% for $f_{(\sqrt{\mp}/-)} \prec f_{(-)}$. We view these findings as strong evidence that our timing strategy is as far from redundant given that, even after accounting for multiple testing, there remains a non-negligible set of anomaly portfolios for which timing both downside volatility and skewness provides incremental spanning.

It remains to investigate in which situations timing both downside volatility and skewness outspans timing only downside volatility. To do so, we run logit regressions that attempt to predict whether $f_{(\sqrt{\mp}/-)} \succ f_{(-)}$ given standardized distributional features of the original portfolio returns: namely, mean, volatility, skewness, kurtosis, downside semi-deviation, and first-order autocorrelation. We find a clear and consistent pattern in Table 10. Timing both downside volatility and skewness is more likely to outspan for anomaly portfolios whose returns are more positively skewed and persistent. In contrast, factors with very high kurtosis are less likely to benefit from our timing strategy. These findings are robust across specifications, regardless of whether we employ the FDR correction.

Table 11 breaks down the instances in which $f_{(\sqrt{\mp}/-)} \succ f_{(-)}$ by characteristic type. It is apparent that accounting for skewness is key for anomalies relating to fundamentals, valuation, and structural attributes (e.g., value, profitability, leverage, size, and long-term reversal). These are exactly the firm characteristics for which upside potential and persistence matter most, with the return-timing channel dominating the volatility-timing channel. In contrast, restricting attention exclusively to downside volatility is more likely to perform best for anomalies that have to do with short-term informational shocks and market frictions (e.g., earnings events, forecast revisions, liquidity, option risk, and short-term reversals) because they react more to downside spikes than to upside asymmetry. Put differently, our timing strategy adds value when good risk drives returns (positive skewness and persistent fundamentals), whereas Wang and Yan's (2021) strategy excels when bad risk prevails (downside shocks and transient volatility). As such, they are complementary in that exploiting conditional skewness captures the expected-return channels that timing only for downside volatility misses.

4 Conclusion

The benefits of volatility timing on active portfolio management remain a topic of ongoing debate. We explore timing strategies that target different combinations of partial second moments to control for higher-order moments. Additionally, apart from proposing cross-sectional tests of stochastic dominance (Barrett and Donald, 2003) and relative pricing (Barillas et al., 2020), we also extend Ledoit and Wolf's (2008) bootstrap-based procedure to carry out direct comparisons of Sortino ratios.

Our empirical findings indicate that volatility management does not enhance the risk-adjusted performance for most factors, anomaly portfolios, and exchange-traded funds. In comparison, our strategy that controls for both downside volatility and skewness shows robust performance across the board, particularly improving the Sharpe and Sortino ratios of the worst-performing portfolios. In addition, we demonstrate the gains of accounting for both positive and negative semivolatilities to predict future returns. Exploiting positive semivolatility enables us to capture moments of positive drift in prices, in contrast to previous strategies that mainly help with negative drifts in prices. This enhances the performance of the anomaly portfolios for which the return timing contributes more than risk timing.

There are two natural directions in which one could extend our analysis. The first is to consider multiple rather than individual tests as in DeMiguel et al. (2024). The second is to propose a more general timing strategy that explicitly accounts for jumps. As it stands, we deal with jumps only implicitly by looking either at the downside-to-upside variance ratio or at the partial variance below a certain quartile.

Data Availability Statement

The data used in this study come from multiple sources. Daily factor returns are publicly available from the Fama–French Data Library (https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html), the AQR website (http://people.stern.nyu.edu/afrazzin), and the Global Q-Factors database (http://global-q.org/factors.html). Daily anomaly portfolio returns are obtained from the Open Asset Pricing database (https://www.openassetpricing.com). Intraday exchange-traded fund (ETF) data are obtained from PiTrading (https://pitrading.com) and are subject to licensing restrictions; they cannot be redistributed by the authors but can be accessed directly from the data provider.

References

Abarbanell, J.S., Bushee, B.J., 1998. Abnormal returns to a fundamental analysis strategy. Accounting Review 73, 19–45.

Ai, H., Han, L.J., Xu, L., 2022. Information-driven volatility. Available at SSRN, doi:http://dx.doi.org/10.2139/ssrn.3961096.

Amenc, N., Esakia, M., Goltz, F., Luyten, B., 2019. Macroeconomic risks in equity factor investing. Journal of Portfolio Management 45, 39–60.

Ang, A., Chen, J., Xing, Y., 2006. Downside risk. Review of Financial Studies 19, 1191–1239.

Ballestero, E., 2005. Mean-semivariance efficient frontier: A downside risk model for portfolio selection. Applied Mathematical Finance 12, 1–15.

Barillas, F., Kan, R., Robotti, C., Shanken, J., 2020. Model comparison with Sharpe ratios. Journal of Financial and Quantitative Analysis 55, 1840–1874.

Barillas, F., Shanken, J., 2017. Which alpha? Review of Financial Studies 30, 1316–1338.

Barndorff-Nielsen, O.E., Kinnebrock, S., Shephard, N., 2008. Measuring downside risk: Realised semivariance, in: Bollerslev, T., Russell, J., Watson, M. (Eds.), Volatility and Time Series Econometrics: Essays in Honor of Robert F. Engle. Oxford University Press, pp. 117–136.

Barrett, G.F., Donald, S.G., 2003. Consistent tests for stochastic dominance. Econometrica 71, 71–104.

Barroso, P., Detzel, A., 2021. Do limits to arbitrage explain the benefits of volatility-managed portfolios? Journal of Financial Economics 140, 744–767.

Barroso, P., Santa-Clara, P., 2015. Momentum has its moments. Journal of Financial Economics 116, 111–120.

Barry, C.B., Brown, S.J., 1984. Differential information and the small firm effect. Journal of Financial Economics 13, 283–294.

Bass, R., Gladstone, S., Ang, A., 2017. Total portfolio factor, not just asset, allocation. Journal of Portfolio Management 43, 38–53.

Bender, J., Le Sun, J., Thomas, R., 2019. Asset allocation vs. factor allocation: Can we build a unified method? Journal of Portfolio Management 45, 9–22.

Bianchi, D., De Polis, A., Petrella, I., 2022. Taming momentum crashes. Available at SSRN, doi:http://dx.doi.org/10.2139/ssrn.4182040.

Blume, M.E., Husic, F., 1973. Price, beta, and exchange listing. Journal of Finance 28, 283–299.

Bollerslev, T., Li, J., Patton, A.J., Quaedvlieg, R., 2020a. Realized semicovariances. Econometrica 88, 1515–1551.

Bollerslev, T., Medeiros, M.C., Patton, A.J., Quaedvlieg, R., 2022a. From zero to hero: Realized partial (co)variances. Journal of Econometrics 231, 348–360.

Bollerslev, T., Patton, A.J., Quaedvlieg, R., 2020b. Multivariate leverage effects and realized semicovariance GARCH models. Journal of Econometrics 217, 411–430.

Bollerslev, T., Patton, A.J., Quaedvlieg, R., 2022b. Realized semibetas: Disentangling "good" and "bad" downside risks. Journal of Financial Economics 144, 227–246.

Carhart, M.M., 1997. On persistence in mutual fund performance. Journal of finance 52, 57–82.

de Castro, L., Galvão, A.F., Montes-Rojas, G., Olmo, J., 2022. Portfolio selection in quantile decision models. Annals of Finance 18, 133–181.

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

Chen, A.Y., Zimmermann, T., 2022. Open source cross-sectional asset pricing. Critical Finance Review 27, 207–264.

Cremers, K.J.M., Nair, V.B., 2005. Governance mechanisms and equity prices. Journal of Finance 60, 2859–2894.

Cusatis, P.J., Miles, J.A., Woolridge, J.R., 1993. Restructuring through spinoffs: The stock market evidence. Journal of Financial Economics 33, 293–311.

DeMiguel, V., Martin-Utrera, A., Uppal, R., 2024. A multifactor perspective on volatility-managed portfolios. Journal of Finance Forthcoming.

Eisdorfer, A., Misirli, E., 2020. Distressed stocks in distressed times. Management Science 66, 2452–2473.

Estrada, J., 2007. Mean-semivariance behavior: Downside risk and capital asset pricing. International Review of Economics and Finance 16, 169–185.

Fama, E.F., French, K.R., 1993. Common risk factors in the returns on stocks and bonds. Journal of Financial Economics 33, 3–56.

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

Franzoni, F., Marin, J.M., 2006. Pension plan funding and stock market efficiency. Journal of Finance 61, 921–956.

Frazzini, A., Pedersen, L.H., 2014. Betting against beta. Journal of Financial Economics 111, 1–25.

Gompers, P., Ishii, J., Metrick, A., 2003. Corporate governance and equity prices. Quarterly Journal of Economics 118, 107–156.

Gómez-Cram, R., 2022. Late to recessions: Stocks and the business cycle. Journal of Finance 77, 923–966.

Haddad, V., Kozak, S., Santosh, S., 2020. Factor timing. Review of Financial Studies 33, 1980–2018.

Haugen, R.A., Baker, N.L., 1996. Commonality in the determinants of expected stock returns. Journal of Financial Economics 41, 401–439.

Hou, K., Mo, H., Xue, C., Zhang, L., 2021. An augmented q-factor model with expected growth. Review of Finance 25, 1–41.

Hou, K., Xue, C., Zhang, L., 2015. Digesting anomalies: An investment approach. Review of Financial Studies 28, 650–705.

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

Kozak, S., Nagel, S., Santosh, S., 2020. Shrinking the cross-section. Journal of Financial Economics 135, 271–292.

Ledoit, O., Wolf, M., 2008. Robust performance hypothesis testing with the Sharpe ratio. Journal of Empirical Finance 15, 850–859.

Liu, F., Tang, X., Zhou, G., 2019. Volatility-managed portfolio: Does it really work? Journal of Portfolio Management 46, 38–51.

Markowitz, H.M., 1959. Portfolio Selection: Efficient Diversification of Investments. Yale University Press.

Markowitz, H.M., Todd, P., Xu, G., Yamane, Y., 1993. Computation of mean-semivariance efficient sets by the critical line algorithm. Annals of Operations Research 45, 307–317.

Memmel, C., 2003. Performance hypothesis testing with the Sharpe ratio. Finance Letters 1, 21–23.

Moreira, A., Muir, T., 2017. Volatility-managed portfolios. Journal of Finance 72, 1611–1644.

Prakash, R., Sinha, N., 2013. Deferred revenues and the matching of revenues and expenses. Contemporary Accounting Research 30, 517–548.

Roy, A.D., 1952. Safety first and the holding of assets. Econometrica 20, 431–449.

Sortino, F.A., van der Meer, R., 1991. Downside risk. Journal of Portfolio Management 17, 27.

Storey, J.D., 2002. A direct approach to false discovery rates. Journal of the Royal Statistical Society, Series B: Statistical Methodology 64, 479–498.

Wang, F., Yan, X.S., 2021. Downside risk and the performance of volatility-managed portfolios. Journal of Banking and Finance 131. Article 106198.

Zeisberger, S., 2023. What is risk? Available at SSRN, doi:http://dx.doi.org/10.2139/ssrn.2811636.

Table 1

Descriptive statistics for the monthly factor returns

We describe the set of factors we consider by reporting the number of monthly time-series observations in 'size', sample period in columns 'start' and 'end', annualized average, minimum, and maximum turns, annualized standard deviation in 'volatility', annualized downside deviation in 'semivolatility', skewness, and excess kurtosis. MKT is the Fama-French market factor, as measured by the excess returns on the market portfolio; SMB the Fama-French size factor, as measured by the small-minus-big mimicking portfolio; HML is the Fama-French value factor, as measured by the high-minus-low mimicking portfolio; MOM is the momentum factor, as measured by the winners-minus-losers rtfolio; RMW is the Fama-French profitability factor, as measured by the robust-minus-weak mimicking portfolio; CMA is the Fama-French investment factor, as measured by the conservative-minus-aggressive mimicking portfolio; BAB is the betting-ainst-beta factor; IA is the investment q-factor; ROE is the profitability q-factor; and EG is the expected growth q-factor.

factor	size	start	end	mean	minimum	maximum	volatility	semivolatility	skewness	excess kurtosis
MKT	1,139	07/1926	05/2021	8.26%	-29.13%	38.85%	18.52%	12.27%	0.17	7.60
SMB	1,139	07/1926	05/2021	2.45%	-16.82%	36.70%	11.02%	6.59%	1.88	18.82
HML	1,139	07/1926	05/2021	4.09%	-13.96%	35.46%	12.17%	6.89%	2.07	18.24
MOM	1,133	01/1927	05/2021	7.66%	-52.27%	18.36%	16.30%	12.48%	-2.96	26.82
RMW	695	07/1963	05/2021	3.10%	-18.48%	13.38%	7.55%	4.91%	-0.30	11.95
CMA	695	07/1963	05/2021	3.17%	-6.86%	9.56%	6.93%	4.19%	0.32	1.53
BAB	1,086	12/1930	05/2021	8.21%	-21.95%	18.65%	11.29%	7.44%	-0.68	7.02
IA	648	01/1967	12/2020	4.01%	-7.16%	9.24%	6.55%	3.93%	0.15	1.22
ROE	648	01/1967	12/2020	6.07%	-14.46%	10.38%	8.87%	5.98%	-0.91	5.72
EG	648	01/1967	12/2020	9.65%	-9.72%	11.51%	6.85%	3.51%	0.11	3.95

Table 2

List of exchange-traded funds in the sample

We list the ticker of every exchange-traded fund (ETF) in the sample, as well as their size and first month. The sample ends on February 2017 for every ETF.

ticker	ETF	size	start	ticker	ETF	size	start
DBC	PowerShares DB Commodity Index Fund	133	02/2006	OIH	Market Vectors Oil Services ETF	193	02/2001
DIA	SPDR Dow Jones Industrial Average ETF	224	07/1998	PFF	iShares S&P U.S. Preferred Stock Index Fund	120	03/2007
DUST	Direxion Daily Gold Miners Bear 3X Shares	75	12/2010	QID	ProShares UltraShort QQQ	128	07/2006
DVY	iShares Dow Jones Select Dividend Index Fund	160	11/2003	QLD	ProShares Ultra QQQ	129	06/2006
DXD	ProShares UltraShort Dow30	128	07/2006	QQQ	PowerShares QQQ	216	03/1999
EDC	Direxion Daily Emerging Markets Bull 3X Shares	99	12/2008	RSX	Market Vectors Russia ETF	119	04/2007
EDZ	Direxion Daily Emerging Markets Bear 3X Shares	99	12/2008	SCO	ProShares UltraShort DJ-UBS Crude Oil	100	11/2008
EEM	iShares MSCI Emerging Markets Index Fund	167	04/2003	SDS	ProShares UltraShort S&P500	128	07/2006
EFA	iShares MSCI EAFE	171	12/2002	SH	ProShares Short S&P500	129	06/2006
ERX	Direxion Daily Energy Bull 3X Shares	100	11/2008	SLV	iShares Silver Trust	131	04/2006
ERY	Direxion Daily Energy Bear 3X Shares	100	11/2008	SMH	Market Vectors Semiconductor ETF	202	05/2000
EWJ	iShares MSCI Japan Index Fund	224	07/1998	SPLV	PowerShares S&P 500 Low Volatility Portfolio	70	05/2011
EWT	iShares MSCI Taiwan Index Fund	171	12/2002	SPXL	Direxion Daily S&P 500 Bull 3X Shares	100	11/2008
EWW	iShares MSCI Mexico Index Fund	171	12/2002	SPXS	Direxion Daily S&P 500 Bear 3X Shares	100	11/2008
EWY	iShares MSCI South Korea Index Fund	171	12/2002	SPY	SPDR S&P 500	224	07/1998
EWZ	iShares MSCI Brazil Index Fund	171	12/2002	SSO	ProShares Ultra S&P500	129	06/2006
FAS	Direxion Daily Financial Bull 3X Shares	100	11/2008	TLT	iShares Lehman 20+ Year Treasury Bond Fund	176	07/2002
FAZ	Direxion Daily Financial Bear 3X Shares	100	11/2008	TNA	Direxion Daily Small Cap Bull 3X Shares	100	11/2008
FXI	iShares FTSE/Xinhua China 25 Index Fund	149	10/2004	TWM	ProShares UltraShort Russell2000	122	01/2007
GDX	Market Vectors Gold Miners ETF	130	05/2006	TZA	Direxion Daily Small Cap Bear 3X Shares	100	11/2008
GDXJ	Market Vectors Junior Gold Miners ETF	88	11/2009	UCO	ProShares Ultra DJ-UBS Crude Oil	100	11/2008
GLD	SPDR Gold Shares	148	11/2004	UNG	United States Natural Gas Fund	119	04/2007
HYG	iShares iBoxx $ High Yield Corporate Bond Fund	119	04/2007	USO	United States Oil Fund	131	04/2006
IBB	iShares Nasdaq Biotechnology Index Fund	171	12/2002	VXX	iPath S&P 500 VIX Short-Term Futures ETN	98	01/2009
IJR	iShares S&P SmallCap 600 Index Fund	171	12/2002	XHB	SPDR S&P Homebuilders ETF	133	02/2006
IVV	iShares S&P 500 Index Fund	171	12/2002	XIV	VelocityShares Daily Inverse VIX Short-Term ETN	76	11/2010
IWD	iShares Russell 1000 Value Index Fund	171	12/2002	XLB	Materials Select Sector SPDR Fund	171	12/2002
IWF	iShares Russell 1000 Growth Index Fund	171	12/2002	XLF	Financial Select Sector SPDR Fund	218	01/1999
IWM	iShares Russell 2000 Index Fund	201	06/2000	XLI	Industrial Select Sector SPDR Fund	171	12/2002
IWN	iShares Russell 2000 Value Index Fund	171	12/2002	XLK	Technology Select Sector SPDR Fund	171	12/2002
IWO	iShares Russell 2000 Growth Index Fund	171	12/2002	XLP	Consumer Staples Select Sector SPDR Fund	171	12/2002
JNK	SPDR Barclays Capital High Yield Bond ETF	111	12/2007	XLY	Consumer Discretionary Select Sector SPDR Fund	171	12/2002
KBE	SPDR KBW Bank ETF	136	11/2005	XME	SPDR S&P Metals and Mining ETF	129	06/2006
KRE	SPDR KBW Regional Banking ETF	129	06/2006	XOP	SPDR S&P Oil & Gas Exploration & Production ETF	129	06/2006
MDY	SPDR S&P MidCap 400 ETF	217	02/1999	XRT	SPDR S&P Retail ETF	129	06/2006
NUGT	Direxion Daily Gold Miners Bull 3X Shares	75	12/2010				

Table 3

Rejection rates of the JK and LW tests of difference in Sortino ratios

We report the empirical rejection frequencies of the JK- and LW-type tests for difference in Sortino ratios across 1,500 Monte Carlo replications. We entertain iid returns with Gaussian and mixture distributions, as well as with GARCH effects. Sample sizes are either of 120 or 240 observations, whereas cross-asset correlation is $\rho \in \{0, 0.3\}$.

				JK			LW		
T	correlation	returns	$\Delta\mu$	10%	5%	1%	10%	5%	1%
120	$\rho = 0$	Gaussian	0.0000	0.105	0.050	0.013	0.090	0.039	0.007
			0.0025	0.135	0.079	0.022	0.137	0.069	0.017
			0.0050	0.265	0.172	0.066	0.269	0.171	0.044
		Mixture	0.0000	0.063	0.023	0.003	0.080	0.036	0.003
			0.0025	0.088	0.042	0.007	0.113	0.043	0.004
			0.0050	0.157	0.097	0.029	0.183	0.091	0.019
		GARCH	0.0000	0.116	0.056	0.010	0.081	0.035	0.003
			0.0025	0.387	0.293	0.138	0.260	0.139	0.033
			0.0050	0.717	0.641	0.473	0.478	0.308	0.067
	$\rho = 0.3$	Gaussian	0.0000	0.104	0.053	0.010	0.118	0.051	0.011
			0.0025	0.165	0.091	0.029	0.168	0.096	0.023
			0.0050	0.335	0.237	0.086	0.334	0.229	0.099
		Mixture	0.0000	0.066	0.028	0.005	0.068	0.025	0.005
			0.0025	0.073	0.039	0.008	0.123	0.056	0.012
			0.0050	0.167	0.083	0.024	0.211	0.123	0.036
		GARCH	0.0000	0.099	0.045	0.009	0.096	0.040	0.009
			0.0025	0.500	0.403	0.242	0.293	0.174	0.046
			0.0050	0.799	0.730	0.569	0.561	0.392	0.152
240	$\rho = 0$	Gaussian	0.0000	0.102	0.053	0.009	0.083	0.037	0.007
			0.0025	0.171	0.094	0.032	0.162	0.089	0.020
			0.0050	0.415	0.307	0.129	0.384	0.263	0.105
		Mixture	0.0000	0.068	0.035	0.002	0.079	0.038	0.007
			0.0025	0.111	0.057	0.017	0.141	0.067	0.013
			0.0050	0.241	0.139	0.038	0.276	0.171	0.057
		GARCH	0.0000	0.102	0.050	0.007	0.090	0.042	0.006
			0.0025	0.619	0.505	0.323	0.364	0.208	0.068
			0.0050	0.885	0.845	0.748	0.627	0.445	0.207
	$\rho = 0.3$	Gaussian	0.0000	0.115	0.061	0.016	0.122	0.055	0.007
			0.0025	0.226	0.141	0.047	0.202	0.119	0.033
			0.0050	0.525	0.395	0.197	0.457	0.335	0.141
		Mixture	0.0000	0.056	0.022	0.005	0.079	0.037	0.004
			0.0025	0.110	0.053	0.010	0.142	0.067	0.015
			0.0050	0.289	0.182	0.052	0.297	0.194	0.066
		GARCH	0.0000	0.100	0.046	0.009	0.103	0.049	0.012
			0.0025	0.710	0.623	0.449	0.408	0.242	0.079
			0.0050	0.928	0.908	0.832	0.700	0.508	0.256

Table 4

Performance of the different timing strategies for each factor

The first panel displays the OLS estimates of the spanning regression alphas (in annualized terms), with their HAC standard errors within parentheses. We respectively denote by $*$, $**$ and $***$ significance at the 10%, 5% and 1% levels. The second and third panels report direct comparisons of Sharpe and Sortino ratios, respectively. For each (semi)volatility management proposal, we document the difference in risk-adjusted performance (namely, Δ and $\Delta_{(-)}$), with the p-values of the JK and LW tests within parentheses and brackets, respectively. We highlight the rejections of the LW test with asterisks as before. The description of the factors is as in Table 1.

	MKT	SMB	HML	MOM	RMW	CMA	BAB	IA	ROE	EG
spanning-regression alpha estimates										
$f_{(\sigma)}$	3.20**	-0.07	1.79**	7.00***	1.38***	0.17	3.40***	1.18**	3.78***	2.90***
	(0.0247)	(0.9224)	(0.0366)	(0.0000)	(0.0097)	(0.6964)	(0.0008)	(0.0191)	(0.0000)	(0.0000)
$f_{(-)}$	2.91**	1.32**	3.37***	4.38**	2.48***	0.02	5.36***	0.51	4.05***	2.86***
	(0.0288)	(0.0447)	(0.0001)	(0.0107)	(0.0001)	(0.9153)	(0.0000)	(0.1061)	(0.0000)	(0.0001)
$f_{(+/-)}$	1.86*	1.26*	3.13***	0.56	2.84***	0.14	4.64***	0.46	2.34***	0.52
	(0.0876)	(0.0776)	(0.0000)	(0.7478)	(0.0005)	(0.4633)	(0.0000)	(0.1358)	(0.0004)	(0.3666)
$f_{(\sqrt{+}/-)}$	2.91**	1.49**	3.87***	2.67	2.68***	0.11	5.79***	0.47*	3.34***	1.86***
	(0.0137)	(0.0258)	(0.0000)	(0.1361)	(0.0001)	(0.4988)	(0.0000)	(0.0984)	(0.0000)	(0.0036)
Sharpe ratio differences Δ										
$f_{(\sigma)}$	0.07	-0.07	0.05	0.39**	0.16	-0.09	0.12	0.07	0.37	0.14
	(0.3847)	(0.3667)	(0.5228)	(0.0000)	(0.1657)	(0.3481)	(0.1178)	(0.4588)	(0.0008)	(0.1849)
	[0.5217]	[0.4177]	[0.7215]	[0.0247]	[0.2918]	[0.3764]	[0.2512]	[0.4963]	[0.1193]	[0.9987]
$f_{(-)}$	0.05	0.08	0.20	0.13	0.31*	-0.30	0.26**	-0.07	0.41	0.01
	(0.5130)	(0.3724)	(0.0082)	(0.1498)	(0.0094)	(0.0511)	(0.0007)	(0.5900)	(0.0009)	(0.9638)
	[0.6895]	[0.3911]	[0.1159]	[0.4777]	[0.0873]	[0.0799]	[0.0280]	[0.6436]	[0.6889]	[0.9880]
$f_{(+/-)}$	0.00	0.06	0.15*	-0.13	0.28**	-0.22	0.23***	-0.06	0.14	-0.26
	(0.9941)	(0.4296)	(0.0155)	(0.1525)	(0.0019)	(0.1397)	(0.0001)	(0.6002)	(0.1908)	(0.0103)
	[1.0000]	[0.5296]	[0.0646]	[0.4977]	[0.0460]	[0.8881]	[0.0100]	[0.6289]	[0.5750]	[0.2159]
$f_{(\sqrt{+}/-)}$	0.06	0.11	0.22**	0.02	0.32**	-0.24	0.31***	-0.05	0.34	-0.09
	(0.4057)	(0.1778)	(0.0004)	(0.8153)	(0.0021)	(0.1166)	(0.0000)	(0.6379)	(0.0031)	(0.4174)
	[0.4963]	[0.2805]	[0.0240]	[0.9154]	[0.0273]	[0.3284]	[0.0020]	[0.7089]	[0.1972]	[0.9960]
Sortino ratio differences $\Delta_{(-)}$										
$f_{(\sigma)}$	0.09	-0.14	0.10	0.74*	0.35	-0.17	0.22	0.15	1.10*	0.98
	(0.4688)	(0.2647)	(0.5096)	(0.0000)	(0.0737)	(0.3094)	(0.0998)	(0.3923)	(0.0000)	(0.0650)
	[0.6835]	[0.3804]	[0.7275]	[0.0553]	[0.2998]	[0.4117]	[0.3438]	[0.4917]	[0.0566]	[0.9567]
$f_{(-)}$	0.07	0.16	0.45	0.24	0.78**	-0.51	0.65**	-0.05	1.51	0.77
	(0.5785)	(0.3125)	(0.0039)	(0.0615)	(0.0011)	(0.0405)	(0.0000)	(0.8419)	(0.0001)	(0.1676)
	[0.7695]	[0.4290]	[0.1093]	[0.4943]	[0.0466]	[0.1552]	[0.0213]	[0.9067]	[0.1692]	[0.5516]
$f_{(+/-)}$	0.16	0.13	0.68**	-0.16	0.87***	-0.34	0.67***	0.04	0.67	-0.17
	(0.2124)	(0.3624)	(0.0013)	(0.1567)	(0.0002)	(0.1872)	(0.0000)	(0.8729)	(0.0073)	(0.6660)
	[0.9554]	[0.6069]	[0.0120]	[0.5716]	[0.0087]	[0.8421]	[0.0067]	[0.9167]	[0.1912]	[0.8688]
$f_{(\sqrt{+}/-)}$	0.21	0.19	0.65**	0.06	0.87***	-0.38	0.84***	-0.01	1.22**	0.41
	(0.0869)	(0.1748)	(0.0001)	(0.6446)	(0.0002)	(0.1537)	(0.0000)	(0.9762)	(0.0001)	(0.3944)
	[0.2818]	[0.3884]	[0.0120]	[0.8581]	[0.0073]	[0.4557]	[0.0027]	[0.9887]	[0.0393]	[0.9900]

Table 5

Predictive regressions for factor returns

We report the results of predictive regressions for monthly factor returns. For each factor, we estimate three specifications. The first employs past realized volatility to predict returns. The second considers only downside risk, using past negative semivolatility as a predictor. The third specification decomposes total volatility into downside and upside risks, letting their partial effects differ. We report the p-values of the t-statistics within parentheses, highlighting with $*$, $**$ and $***$ significance at the 10%, 5% and 1% levels. See Table 1 for the description of the factors.

	MKT			SMB			HML			MOM			RMW		
	total	down	down+up	total	down	down+up	total	down	down+up	total	down	down+up	total	down	down+up
intercept	0.0055**	0.0084***	0.0047*	-0.0003	0.0007	-0.0006	0.0005	0.0042**	0.0018	0.0143***	0.0127***	0.0150***	-0.0010	0.0010	-0.0012
	(0.0477)	(0.0011)	(0.0895)	(0.8609)	(0.6244)	(0.6975)	(0.7660)	(0.0121)	(0.2925)	(0.0000)	(0.0000)	(0.0000)	(0.4960)	(0.4750)	(0.4070)
σ_{t-1}	0.1521			0.5030*			0.6153***			-1.3352***			1.1539**		
	(0.5609)			(0.0751)			(0.0344)			(0.0000)			(0.0039)		
$\sigma_{(-),t-1}$		-0.0533	-0.3014***		0.0861	-0.0982		-0.0533	-0.7890***		-0.3238***	-0.1896*		0.1637	-0.3825***
		(0.4527)	(0.0049)		(0.2455)	(0.3785)		(0.5463)	(0.0000)		(0.0000)	(0.0766)		(0.1760)	(0.0206)
$\sigma_{(+),t-1}$			0.3622***			0.2785**			0.8348***			-0.2508*			0.7198***
			(0.0020)			(0.0272)			(0.0000)			(0.0861)			(0.0000)

	CMA			BAB			IA			ROE			EG		
	total	down	down+up	total	down	down+up	total	down	down+up	total	down	down+up	total	down	down+up
intercept	-0.0016	-0.0002	-0.0019	0.0059***	0.0076***	0.0032*	0.0005	0.0015	0.0000	0.0108***	0.0103***	0.0098***	0.0083***	0.0086***	0.0077***
	(0.3037)	(0.8666)	(0.2045)	(0.0004)	(0.0000)	(0.0576)	(0.7626)	(0.2690)	(0.9865)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
σ_{t-1}	1.3368***			0.2020			0.8977**			-1.6200***			-0.0790		
	(0.0016)			(0.4541)			(0.0304)			(0.0009)			(0.8537)		
$\sigma_{(-),t-1}$		0.2932**	-0.0048		-0.0502	-0.4859***		0.1858*	-0.0321		-0.4904***	-0.5337***		-0.0573	-0.1571
		(0.0167)	(0.9757)		(0.4804)	(0.0000)		(0.1096)	(0.8349)		(0.0001)	(0.0006)		(0.6349)	(0.3120)
$\sigma_{(+),t-1}$			0.4492***			0.6708***			0.3478**			0.0848			0.1616
			(0.0037)			(0.0000)			(0.0321)			(0.6236)			(0.3076)

Table 6

Timing risk of a broader set of test portfolios

We report the number of anomaly portfolios and exchange-traded funds for which we find significantly *positive* alphas and differences in Sharpe and Sortino ratios at the 5% significance level. For the direct mparisons of risk-adjusted performances, we employ both JK and LW tests.

	spanning alpha	Δ		$\Delta_{(-)}$	
		JK	LW	JK	LW
207 anomaly portfolios					
$f_{(\sigma)}$	110	51	25	64	24
$f_{(-)}$	142	63	30	89	35
$f_{(+/-)}$	102	44	19	63	25
$f_{(\sqrt{+}/-)}$	138	63	34	86	38
71 exchange-traded funds					
$f_{(\sigma)}$	14	9	1	13	0
$f_{(-)}$	11	6	1	10	0
$f_{(+/-)}$	3	3	1	3	0
$f_{(\sqrt{+}/-)}$	17	17	3	19	2
$f_{(q)}$	12	9	0	12	0

Table 7

Comparison of the pricing errors with and without timing

We report the t-statistics of the difference in the squared maximum Sharpe ratios that we can attain from managed and unmanaged Fama-French factor models. We compute the t-statistics for the case of nonnested models with nontraded factors (Barillas et al., 2020). Asymptotic-valid critical values rest on the standard Gaussian distribution. We denote by $*$, $**$ and $***$ significance at the 10%, 5% and 1% levels.

	timing strategy			
	$f_{(\sigma)}$	$f_{(-)}$	$f_{(+/-)}$	$f_{(\sqrt{+}/-)}$
Fama-French three-factor model	1.17	1.86**	0.65	1.82**
Fama-French five-factor model	0.23	1.79**	1.05	2.00***

Table 8

Stochastic dominance between managed and unmanaged portfolios

We report Barrett and Donald's (2003) test of first-order stochastic dominance in the cross-section distribution of Sharpe and Sortino ratios of managed and unmanaged portfolios. We denote by $*$, $**$ and $***$ significance at the 10%, 5% and 1% levels, whose asymptotic-valid critical values are respectively 1.073, 1.2239 and 1.5174.

		timing strategy				
		$f_{(\sigma)}$	$f_{(-)}$	$f_{(+/-)}$	$f_{(\sqrt{+}/-)}$	$f_{(q)}$
anomaly portfolios						
Sharpe	dominates	1.13*	1.67***	0.74	1.67***	
	dominated	0.34	0.05	0.49	0.20	
Sortino	dominates	1.13*	2.21***	1.67***	2.31***	
	dominated	0.25	0.05	0.10	0.05	
exchange-traded funds						
Sharpe	dominates	1.35**	1.18*	0.34	1.27**	0.93
	dominated	0.25	0.25	0.42	0.25	0.17
Sortino	dominates	1.35**	1.27**	0.25	1.27**	1.27**
	dominated	0.34	0.34	0.25	0.42	0.17

Table 9

Average return- and volatility-timing contributions by characteristic type

We decompose the spanning alphas into their return- and volatility-timing components. The former boosts alpha when leverage positively correlates with contemporaneous returns, whereas the latter kicks in at times of negative covariance between leverage and the magnitude of factor returns.

characteristic type	RT	VT	characteristic type	RT	VT
accruals	0.3251	0.4247	momentum	1.7121	2.2661
asset composition	1.5166	0.3588	option risk	-0.1878	1.9721
cash-flow risk	6.6056	-0.7960	other	1.0213	0.6976
composite accounting	1.5578	-0.3606	ownership	0.0727	3.1894
default risk	2.1070	2.1251	payout indicator	0.2544	1.3043
earnings event	-1.3183	3.1706	profitability	2.8750	0.8868
earnings forecast	1.1043	0.7821	profitability alt	1.3286	0.4723
earnings growth	1.5631	0.5207	R&D	1.6221	0.5541
external financing	1.3507	0.8345	recommendation	0.4899	1.2260
info proxy	1.6546	-0.2742	risk	0.6878	1.1672
investment	1.2284	0.3087	sales growth	0.5695	0.2157
investment alt	0.5704	0.7135	short-sale constraints	1.1543	0.8443
investment growth	0.0902	0.5493	short-term reversal	-1.9760	10.6572
lead lag	0.2520	1.4385	size	2.5680	0.6823
leverage	2.3030	0.4136	valuation	2.9929	1.0066
liquidity	0.6951	1.0928	volatility	2.1361	1.4699
long-term reversal	2.2511	0.2784	volume	0.3383	1.4305

Table 10

Logit regressions for the probability that $f_{(\sqrt{+}/-)} \succ f_{(-)}$

We report the estimation results of the logit regression for the probability that timing both downside volatility and skewness outspans timing only downside volatility, but not vice-versa. We compute the $f_{(\sqrt{+}/-)} \succ f_{(-)}$ set both with and without multiple-testing FDR correction (Storey, 2002). As regressors, we employ standardized distributional features of the portfolio returns. We document standard errors within parentheses, highlighting significance at the 10%, 5% and 1% levels by *, **, and ***, respectively.

control	logit estimates	(with FDR correction)
intercept	-1.364***	-4.343*
	(0.506)	(2.509)
mean	-1.441**	-7.280**
	(0.592)	(3.560)
standard deviation	-2.008	43.633
	(2.556)	(28.522)
skewness	4.213***	3.192
	(1.426)	(4.512)
kurtosis	-2.646***	-5.901
	(0.861)	(4.193)
downside semi-deviation	1.973	-38.572
	(2.518)	(26.031)
first-order autocorrelation	2.050***	2.920**
	(0.607)	(1.461)

Table 11

Breakdown of the $f_{(\sqrt{+}/-)} \succ f_{(-)}$ set by characteristic type

We report how many anomaly portfolios belong to the $f_{(\sqrt{+}/-)} \succ f_{(-)}$ set for each characteristic type, as well as the average return- and volatility-timing components of their spanning alphas.

	no multiple-testing adjustment				with FDR correction			
characteristic type	count	RT	VT	RT−VT	count	RT	VT	RT−VT
valuation	6	2.99	1.01	1.99	3	2.99	1.01	1.99
other	5	1.02	0.70	0.32	4	1.02	0.70	0.32
leverage	3	2.30	0.41	1.89	2	2.30	0.41	1.89
long-term reversal	3	2.25	0.28	1.97	2	2.25	0.28	1.97
size	1	2.57	0.68	1.89	1	2.57	0.68	1.89
investment	4	1.23	0.31	0.92	1	1.23	0.31	0.92
liquidity	3	0.70	1.09	-0.40	1	0.70	1.09	-0.40
external financing	2	1.35	0.83	0.52	1	1.35	0.83	0.52
profitability	4	2.88	0.89	1.99				
R&D	2	1.62	0.55	1.07				
profitability alt	3	1.33	0.47	0.86				
short-sale constraints	2	1.15	0.84	0.31				
earnings growth	1	1.56	0.52	1.04				
info proxy	1	1.65	-0.27	1.93				
momentum	1	1.71	2.27	-0.55				
option risk	1	-0.19	1.97	-2.16				
payout indicator	1	0.25	1.30	-1.05				
investment growth	1	0.09	0.55	-0.46				

Figure 1: Descriptive statistics of the monthly returns on anomaly portfolios
We report their box plots for the average returns (mean), volatility (vol), and downside semivolatility (semivol) in percentage per annum, as well skewness and excess kurtosis.

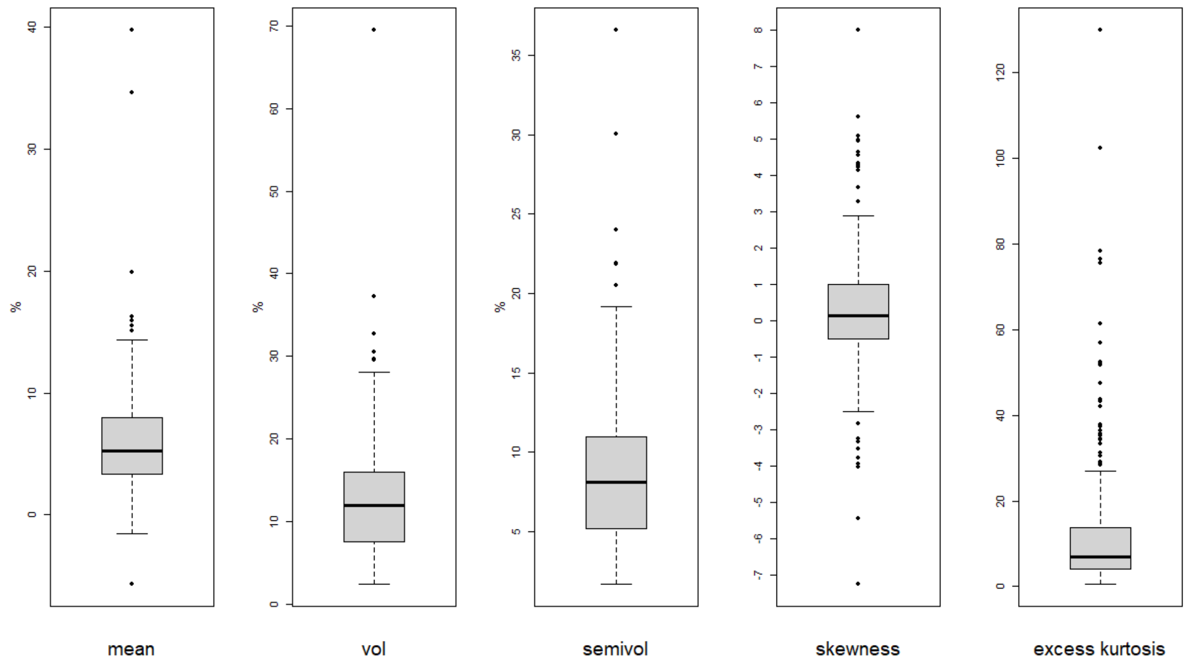


Figure 2: Descriptive statistics of the monthly returns on ETFs
We report their box plots for the average returns (mean), volatility (vol), and downside semivolatility (semivol) in percentage per annum, as well skewness and excess kurtosis.

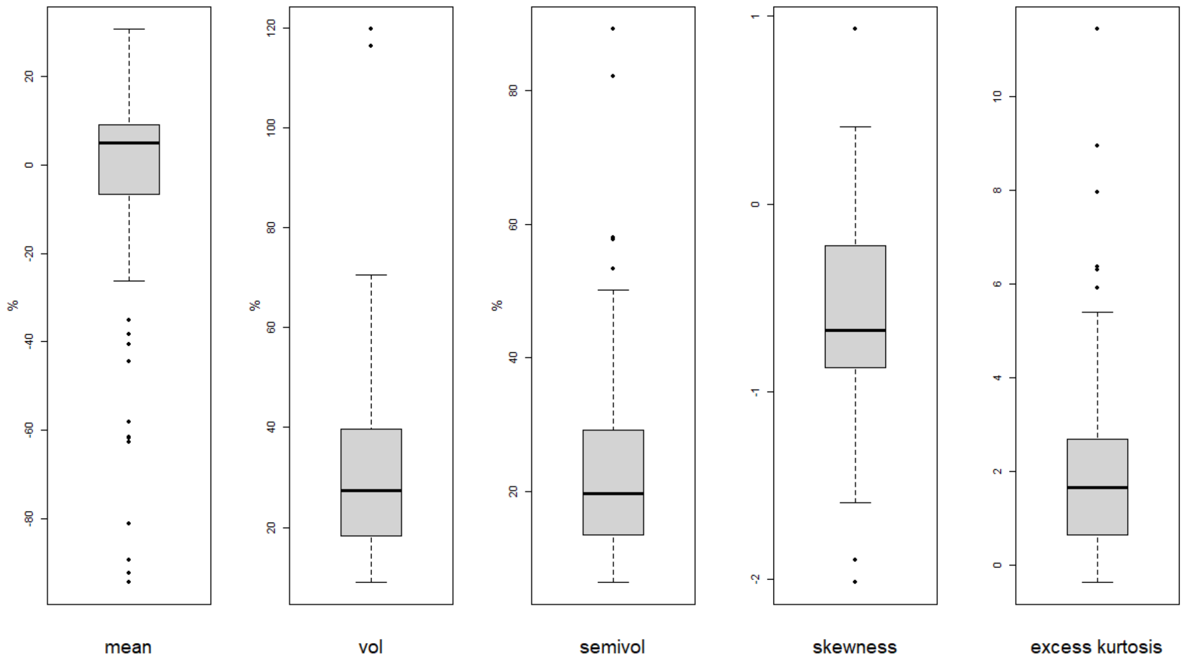


Figure 3: Power curves for the JK and LW tests of difference in Sortino Ratio estimates
The first and second plots display the rejection rates of the JK and LW tests at the 5% level, respectively. We simulate $T = 240$ time-series observations of two portfolio returns either with iid errors from Gaussian and mixture distributions, or with GARCH errors. Results rest on 1,500 Monte Carlo replications, with the LW test using $B = 1,500$ bootstrap samples.

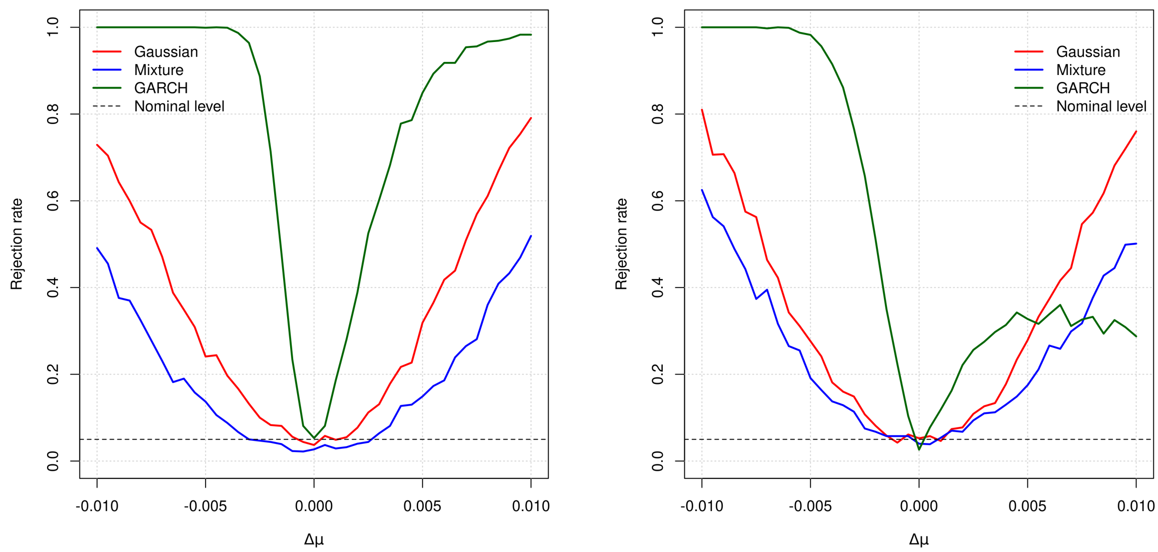


Figure 4: Adjusted R^2's from regressions of factor returns on their past (semi)volatilities
We display the adjusted R-squared from the regressions of factor returns at time t onto their measures of risk at time $t-1$. Beige bars corresponds to values if controlling only for total volatility, red bars if accounting only for downside volatility, and blue bars if regressing on both downside and upside volatilities.

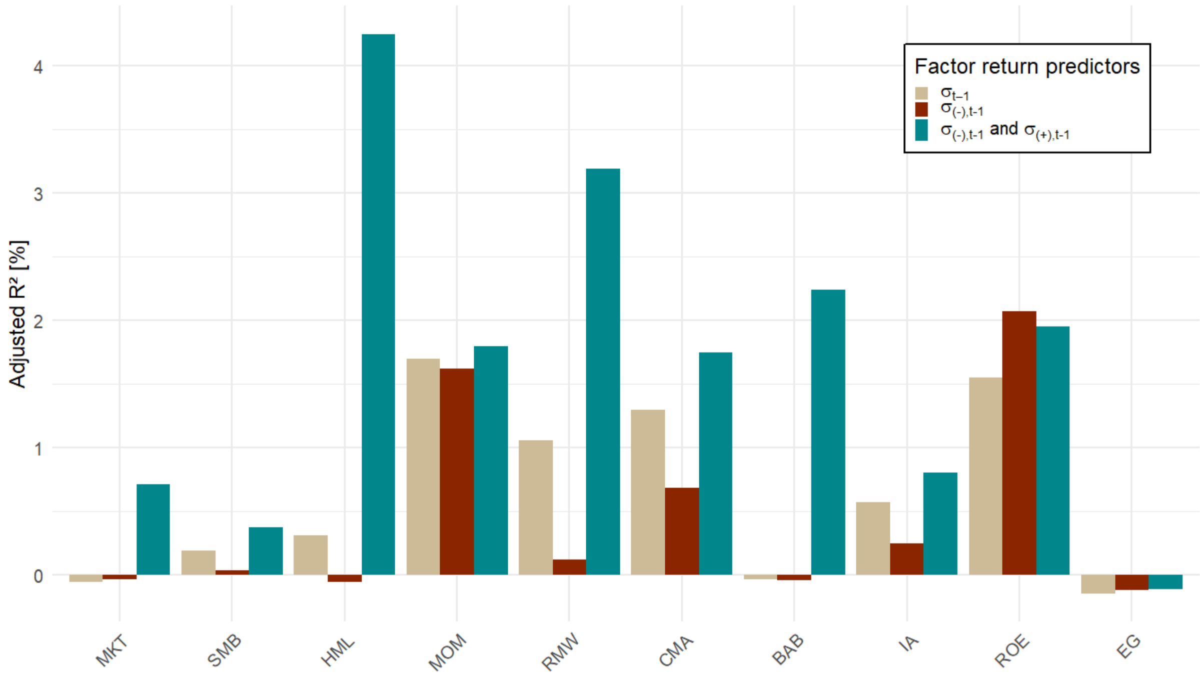


Figure 5: Decomposition of the spanning-regression alphas of the semivolatility-managed strategy
We display the return- and volatility-timing components of the spanning-regression alphas for each factor (in beige and red, respectively(, as well as their sums (which correspond to the alphas, in green).

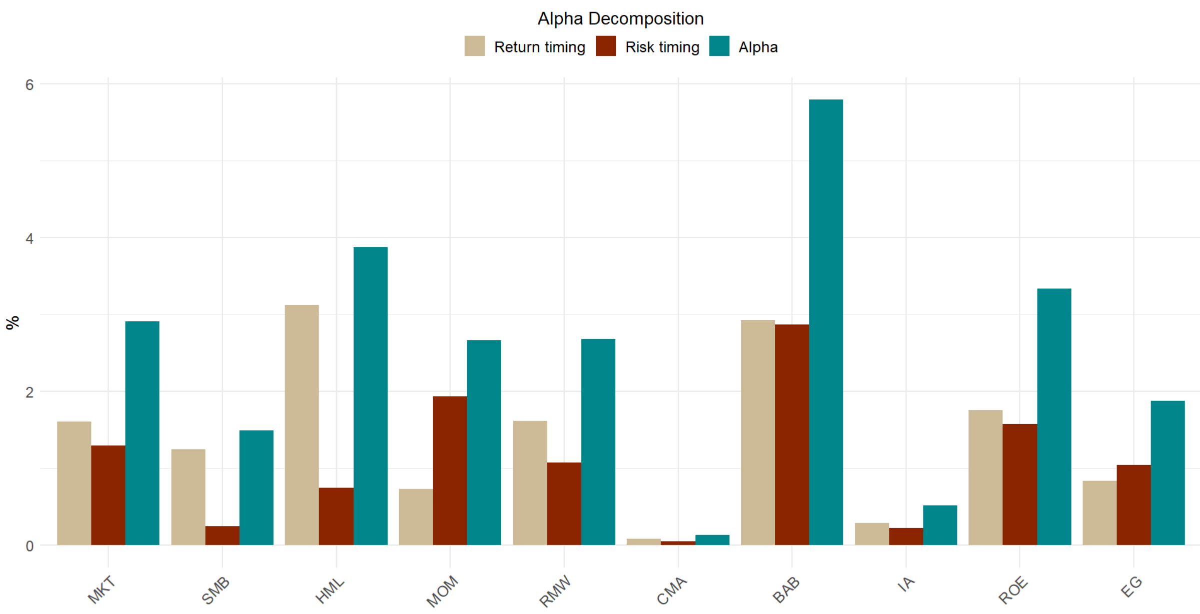


Figure 6: Distribution of the Sharpe and Sortino ratios of managed and unmanaged anomaly portfolios
We report the empirical distribution of the annual Sharpe and Sortino ratios of the 207 anomaly portfolios for the different timing strategies.

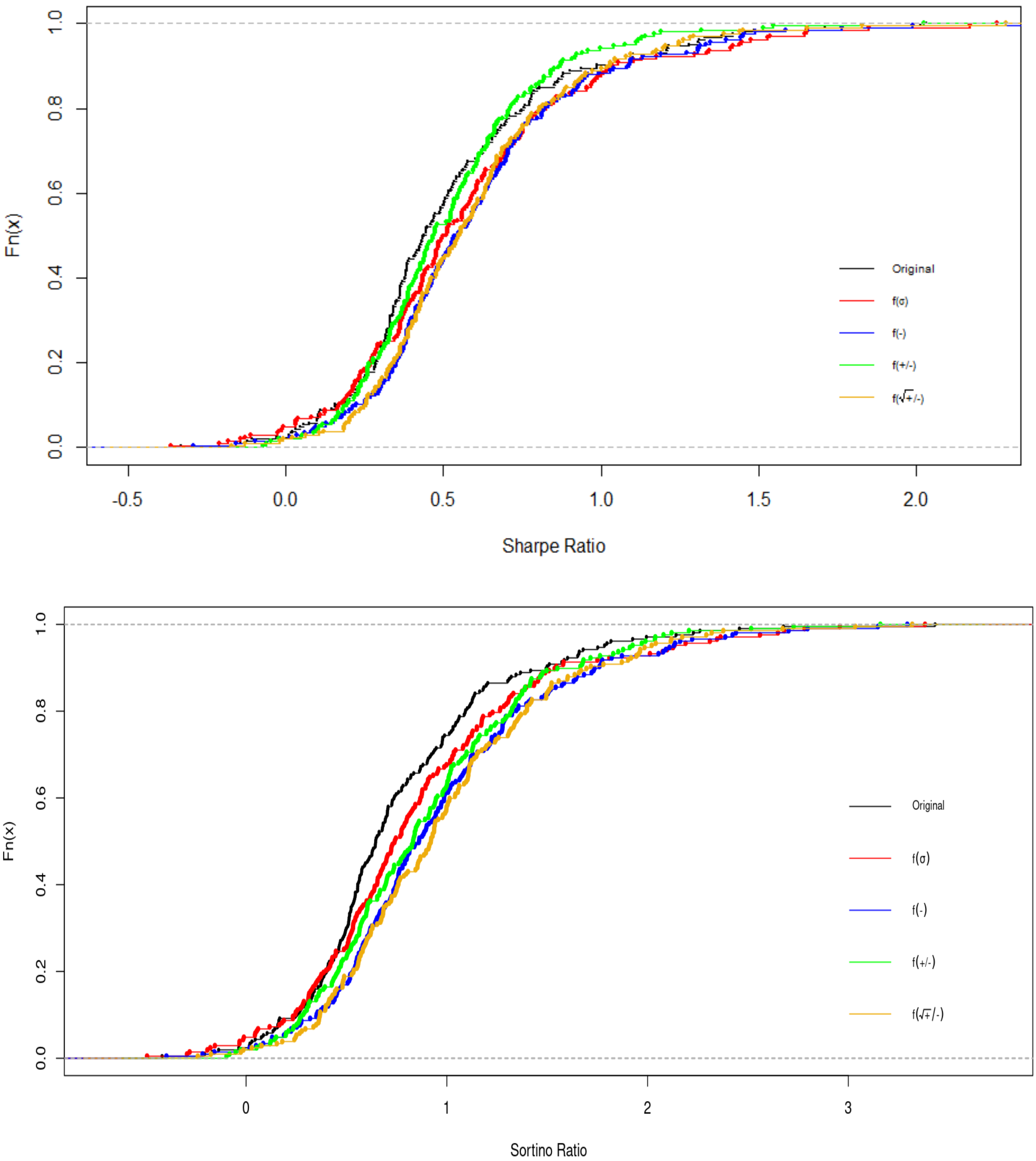
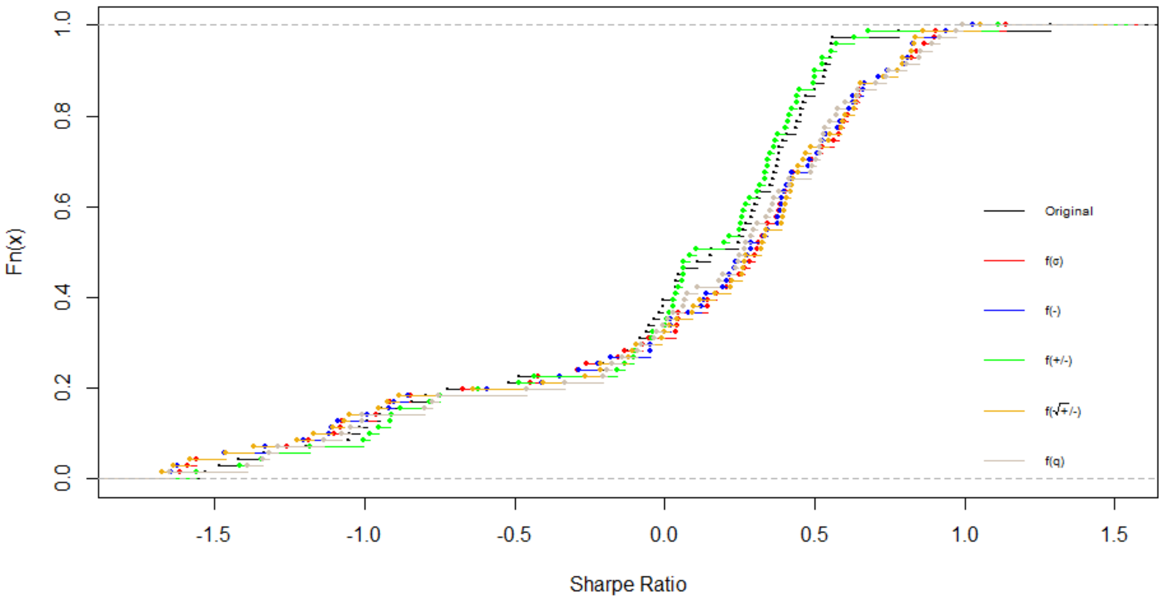


Figure 7: Distribution of the Sharpe and Sortino ratios of managed and unmanaged ETFs
We report the empirical distribution of the annual Sharpe and Sortino ratios of the 71 exchange-traded portfolios for the different timing strategies.



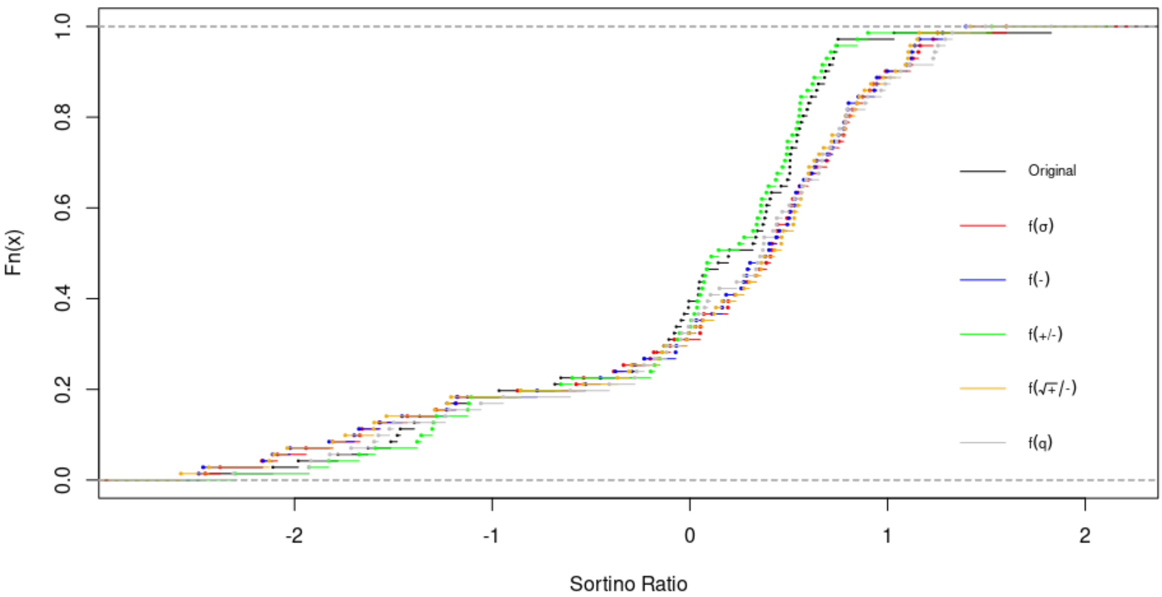


Figure 8: Distribution of the Sharpe ratios of managed and unmanaged ETFs, after transaction costs
We report the empirical distribution of the annual Sharpe ratios of the 71 exchange-traded portfolios for the different timing strategies after considering transaction costs.

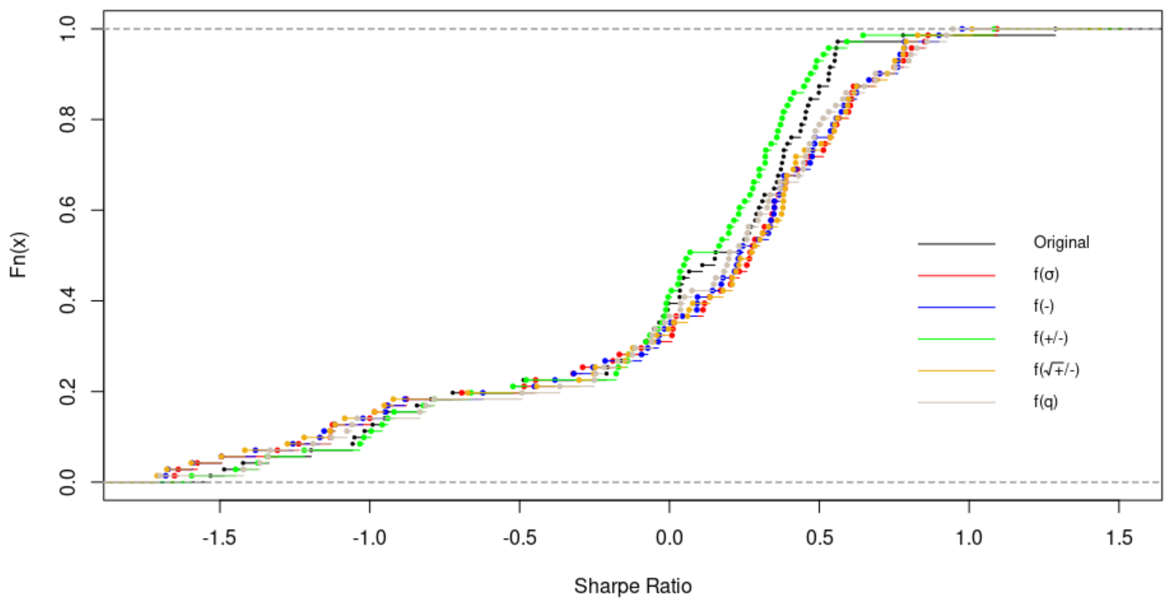


Figure 9: Distribution of the spanning-regression alphas and appraisal ratios
We report the empirical distributions of spanning-regression alphas and appraisal ratios of the 207 anomaly portfolios for the different timing strategies.

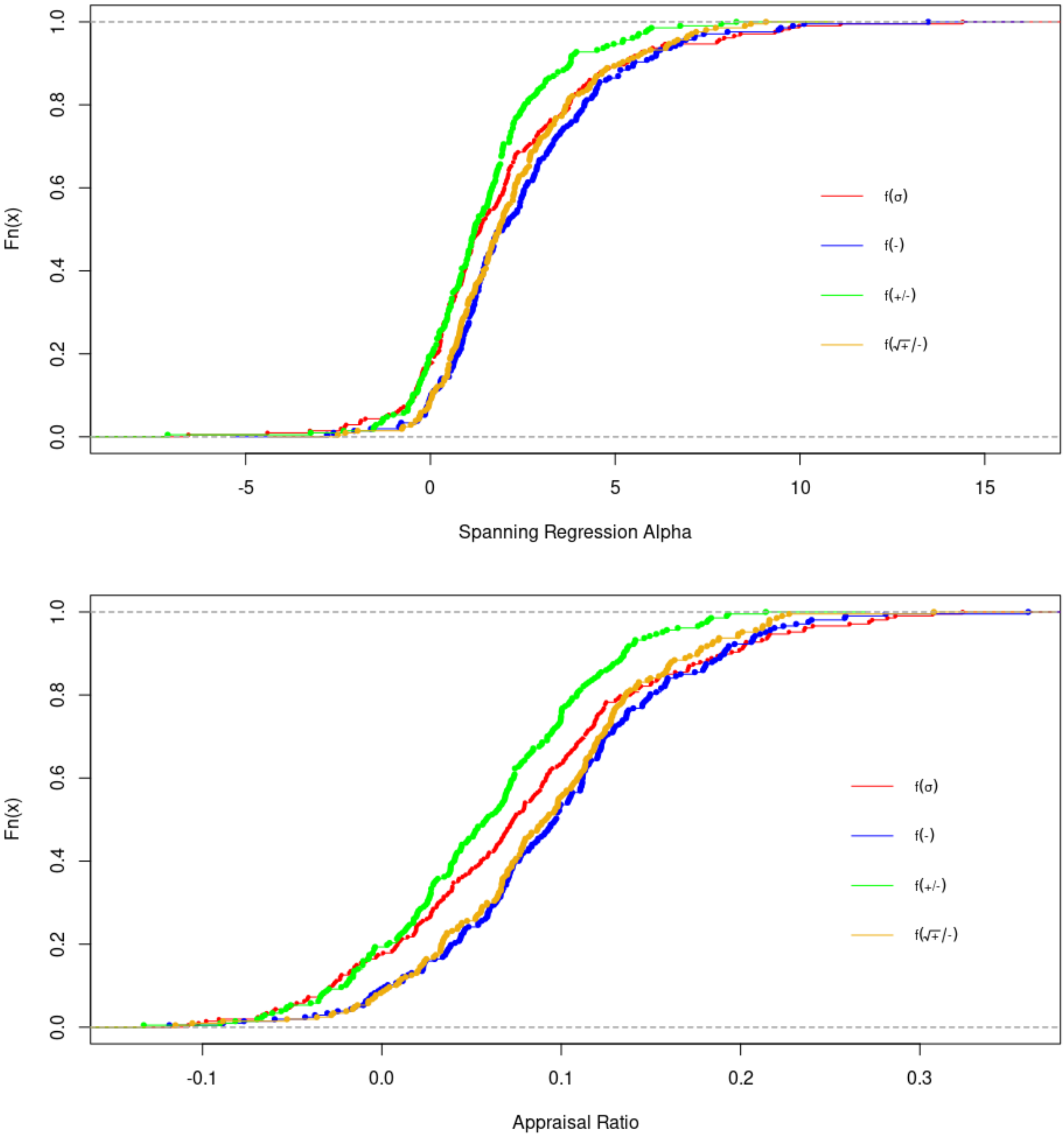


Figure 10: Decomposition of the spanning-regression alphas of the semivolatility-managed portfolios
We display the distributions of the return- and volatility-timing components of the spanning-regression alphas for the 207 anomaly portfolios.