# Conditional Benchmarks and the Identification of Skill in Active Management

Scott Cederburg* Michael S. O'Doherty† N. E. Savin‡ Ashish Tiwari§

February 16, 2016¶

**Abstract**

Recent studies find that it is possible to identify superior performing equity mutual funds ex ante based on fund characteristics, and this evidence often takes the form of a large spread in unconditional alphas for characteristic-sorted portfolios. Unconditional benchmarks can, however, produce misleading evidence on underlying manager skill for strategies that require frequent rebalancing and exhibit unstable exposures to the benchmark factors. We propose an approach to performance attribution in which lagged factor loadings for constituent funds are used as instruments to capture predictable changes in the exposures of mutual fund portfolios. In comparison to existing methods, our benchmarks yield superior tracking performance and a more powerful statistical assessment of manager skill. We apply our method to reevaluate the predictive content of multifactor model $R^2$ and return volatility for fund performance and find no evidence of superior selection ability in the related strategies.

**JEL classifications:** G11, G23

*Eller College of Management, University of Arizona and Freeman School of Business, Tulane University. Email: cederburg@email.arizona.edu.

†Trulaske College of Business, University of Missouri. Email: odohertym@missouri.edu.

‡Tippie College of Business, University of Iowa. Email: gene-savin@uiowa.edu.

§Tippie College of Business, University of Iowa. Email: ashish-tiwari@uiowa.edu.

¶We thank Brad Goldie, Haim Kassa, Xiaolu Wang, and seminar participants at Iowa State University and Miami University for helpful comments and suggestions.

# 1 Introduction

Prior literature on the performance of equity mutual funds has generated a long list of apparently successful strategies for identifying skill in active management. Recently suggested predictors of mutual fund performance include its manager's resemblance to successful managers (Cohen, Coval, and Pástor (2005)), industry concentration (Kacperczyk, Sialm, and Zheng (2005)), unobserved actions of funds (Kacperczyk, Sialm, and Zheng (2008)), active share (Cremers and Petajisto (2009)), contractual incentives for managers (Massa and Patgiri (2009)), risk shifting (Huang, Sialm, and Zhang (2011)), mutual fund $R^2$ (Amihud and Goyenko (2013)), skill to pick stocks in booms and time the market in recessions (Kacperczyk, Van Nieuwerburgh, and Veldkamp (2014)), managerial activeness (Doshi, Elkamhi, and Simutin (2015)), fund return volatility (Jordan and Riley (2015)), and contrarian strategies (Wei, Wermers, and Yao (2015)).

A common empirical approach in these papers is to sort mutual funds into portfolios based on the proposed predictor variable and evaluate the performance of the resulting trading strategies. This analysis typically involves computing style-adjusted returns via time-series regressions of the portfolio returns on market, size, value, and momentum factors (i.e., Carhart (1997) four-factor model regressions). Realized portfolio returns are effectively decomposed into a component that reflects the return that could be earned from exposures to the four passive benchmark factors and a residual component (i.e., alpha). The key evidence for distinguishing a useful predictor is that the extreme portfolios exhibit an economically large and statistically significant difference in alphas. The results from recent literature suggest that it is possible to identify a subset of managers with security selection or market timing skill.[1]

A potential concern with these findings relates to the benchmarking procedure used to style-adjust the portfolio returns. Specifically, for each of the proposed predictors, the identity of the "skilled" funds changes through time. The associated strategies are designed to maintain exposure to the skilled managers by trading in and out of individual mutual funds over the sample period. As the portfolios are rebalanced each month or year, however, these strategies can exhibit pronounced shifts in both the identity of the constituent funds and their underlying factor exposures. Importantly, these changes in portfolio loadings can occur solely as a result of portfolio turnover, even if the underlying mutual funds are not adjusting their factor exposures over time. These style dynamics are not reflected in the standard Carhart (1997) four-factor benchmarking approach,

[1] Other characteristics with predictive content for mutual fund performance include reliance on public information (Kacperczyk and Seru (2007)) and ability to pick outperforming industries (Busse and Tong (2012)). The findings in these studies, however, are not based on the unconditional Carhart (1997) regression approach.

which assigns constant exposures to a given strategy over the full sample period. Consequently, the reported alphas and apparent managerial skill may simply reflect portfolio style drift and a benchmark model that performs poorly in tracking the strategy returns.

To address this concern, we extend the conventional conditional performance evaluation framework to account for these predictable changes in style exposures. As we discuss in more detail below, the use of conditional benchmarks that closely track shifts in portfolio factor loadings leads to dramatically different conclusions regarding the ex ante identification of superior managerial skill. In particular, focusing on two recently documented predictors of fund performance, we show that neither fund $R^2$ nor fund return volatility is related to future performance after accounting for shifts in style exposures of the relevant fund portfolio strategies.

In the typical conditional performance evaluation implementation (e.g., Ferson and Schadt (1996), Ferson (2010), and Ferson (2013)), one or more of a portfolio's factor loadings is modeled as a linear function of state variables, such as the dividend yield, default spread, and term spread. These variables have a long history of use in forecasting asset returns (e.g., Fama and French (1989)), and the motivation for including them in a conditional benchmark is to control for mechanical factor timing strategies based on publicly available information. As Ferson and Schadt (1996) explain, any portfolio strategy that can be replicated using such information should not be deemed as having superior performance.

Our innovation is to replace or complement the traditional conditioning variables with another set of instruments based on lagged factor loading estimates for a strategy's constituent mutual funds. We specifically propose using the portfolio-weighted average lagged factor loadings across funds held in a particular portfolio as instruments for that portfolio's exposures in a Carhart (1997) model regression. The use of lagged loadings as instruments for the conditional factor exposures in a performance evaluation context was first proposed by Boguth, Carlson, Fisher, and Simutin (2011), who employ this method to reexamine the performance of stock momentum strategies.

Our proposed approach has several attractive features. First, it is easy to implement, as it only requires returns data and is based on standard time-series regression methods used in the mutual fund literature. Second, the lagged factor loadings allow the researcher to incorporate a powerful source of publicly available information in predicting future style exposures and benchmarking performance. In particular, these instruments provide a simple way to pick up high-frequency shifts in style (e.g., at portfolio rebalancing dates) that would be missed by traditional conditioning variables, such as the dividend yield, which tend to be much more persistent in nature. In our empirical applications, we show that models with lagged loadings as instruments exhibit pronounced improvements in tracking strategy returns over unconditional Carhart (1997) model regressions and models with traditional conditioning variables.

Third, our method leads to an intuitive decomposition of a given strategy's unconditional alpha, which tends to be the focus of prior literature as noted above, into performance in security selection (i.e., the conditional alpha), factor timing, and volatility timing. Factor timing and volatility timing have each been considered in isolation in the mutual fund literature.[2] As noted by Ferson and Mo (2015), measures of selectivity that ignore managerial ability in timing either factor returns or factor volatility suffer from an omitted variable bias. Boguth, Carlson, Fisher, and Simutin (2011) make a similar point regarding the importance of market timing and volatility timing in returns-based instrumental variables tests, similar to those in our paper. Finally, because our conditional models with lagged factor loading instruments lead to improved tracking performance (i.e., high time-series regression $R^2$s), this approach also produces extremely precise estimates of conditional alpha. As such, our tests have increased statistical power to identify skill in security selection among the strategies of interest.

To demonstrate the usefulness of our approach, we reevaluate two sets of results from prior literature on the performance of mutual fund strategies—Amihud and Goyenko's (2013) $R^2$ effect and Jordan and Riley's (2015) volatility effect. Amihud and Goyenko (2013) find that mutual funds with low $R^2$ values from Carhart (1997) four-factor regressions subsequently outperform funds with high $R^2$s. That is, funds with greater selectivity in active management over recent periods tend to deliver superior future performance. We focus on this study largely because its results are representative of other findings on the performance of managed portfolios. For example, Cremers and Petajisto (2009) and Petajisto (2013) find that a mutual fund's "active share," or tendency to take diversified positions away from its benchmark, shows a positive association with future abnormal returns. Similarly, Titman and Tiu (2011) and Sun, Wang, and Zheng (2012) show that hedge funds following unique investment strategies have superior investment ability. We also revisit the results in Jordan and Riley (2015), who find that lagged mutual fund volatility is a strong predictor of unconditional alpha over a sample that starts in January 2000. The recent sample period in this paper is ideal for highlighting our conditional performance evaluation approach, as some of the lagged factor loadings we use as instruments require data on daily fund returns, which

[2] A partial list of studies on market timing by mutual funds includes Treynor and Mazuy (1966), Grant (1977), Henriksson and Merton (1981), Henriksson (1984), Ferson and Schadt (1996), Becker, Ferson, Myers, and Schill (1999), and Jiang, Yao, and Yu (2007). Avramov and Chordia (2006) account for predictability in the security selection and factor timing skills of managers. Busse (1999) considers the ability of mutual fund managers to time market volatility.

are not readily available prior to this period.

As a starting point for our empirical analysis, we reproduce the strong inverse relation between portfolio rank and performance for both lagged $R^2$ and lagged volatility. A hypothetical portfolio that takes a long position in the quintile of funds with the lowest $R^2$ values and a short position in the quintile with the highest $R^2$ values earns an average return of 5.02% per year and an unconditional alpha of 3.72%.[3] Similarly, the bottom decile of mutual funds sorted on prior volatility outperforms the top decile by 5.64% after controlling for exposure to the four Carhart (1997) factors in an unconditional framework. We further demonstrate, however, that the extreme $R^2$ and volatility portfolios exhibit considerable portfolio turnover, which generates predictable shifts in style exposures over the sample period. For example, the low-$R^2$ and high-$R^2$ strategies require annualized portfolio turnover of 280% and 315%, respectively. We also find that the superior performance for each of these strategies is concentrated over short sample periods and corresponds with discrete changes in investment styles due to portfolio rebalancing. These results raise the possibility than the prior evidence on superior performance for low-$R^2$ and low-volatility funds is attributable to poorly specified benchmark models, rather than skill in security selection.

Consistent with this possibility, we find that both $R^2$ and volatility are unrelated to future performance relative to our conditional benchmarks that account for predictable shifts in style. The conditional alpha for the low-minus-high $R^2$ (volatility) strategy is reduced by 66% (75%) in comparison to the corresponding unconditional estimate. Neither of the long-short conditional alphas in our most comprehensive conditional models is statistically significant at the 5% level. We also show that our conditional benchmarks for the mutual fund strategies deliver superior tracking performance relative to unconditional benchmark models and traditional conditional methods used in prior literature.

Finally, we present decompositions of the unconditional alphas for the two low-minus-high strategies into factor timing, volatility timing, and security selection effects. In both cases, we find that the large unconditional alphas for these portfolios are mostly attributable to factor timing. That is, whereas the $R^2$ and volatility portfolios tend to have relatively high conditional exposures to the benchmark factors in periods when these factors earn high returns, there is limited evidence that their superior performance is linked to ability in security selection. An important question is whether this apparent success in factor timing is attributable to managers of the underlying funds skillfully shifting their exposures over time or simply to the strategy's rebalancing procedure. To

[3]Following Amihud and Goyenko (2013), we focus on the $R^2$-performance relation among mutual funds with high prior alphas. See Section 3 for details on portfolio formation.

address this issue, we examine the impact of extending the holding period for the $R^2$ and volatility portfolios from one to 24 months. For the $R^2$ strategy, the unconditional alpha declines sharply as we extend the holding period and becomes statistically insignificant after six months. These results suggest that any observed outperformance for low-$R^2$ mutual funds is unlikely to be related to managerial skill in factor timing. In contrast, the unconditional volatility effect remains robust to extending the holding period. Further analysis reveals, however, that the volatility strategy's apparent success in factor timing is entirely attributable to poor style bets in the high-volatility portfolio in 2000 and 2001. Thus, neither strategy appears to robustly identify managers with factor timing skill.

The paper contributes to an extensive literature on the performance evaluation of managed portfolios. We specifically build on the conditional performance evaluation approach to constructing linear returns-based benchmarks introduced by Ferson and Schadt (1996). Our contribution is to incorporate information from a strategy's lagged factor loadings to better characterize its conditional style exposures, and we show that our instruments considerably outperform traditional instruments in tracking portfolio returns. A byproduct of this improved tracking ability is an increase in test power. This method is thus an important tool for investors and researchers who wish to identify significant predictors of managerial skill.

Our approach also lends itself to a useful performance decomposition in terms of selection ability, factor timing, and volatility timing. A paper closely related to ours is Ferson and Mo (2015), who argue for such a decomposition in evaluating mutual fund performance and propose a method based on mutual fund holdings. Our simple returns-based approach should prove useful in disentangling the sources of abnormal unconditional performance documented in past studies and in future research.

The remainder of the paper is organized as follows. Section 2 introduces our estimation approach for performance evaluation and contrasts it with existing methods. Section 3 details our sample selection procedures and the construction of the mutual fund strategies based on $R^2$ and volatility. Section 4 presents our main results on the performance these mutual fund portfolios, and Section 5 concludes.

# 2 Performance evaluation methods and conditioning variables

In this section, we introduce our method for conditional performance evaluation and contrast it with the unconditional and conditional factor regression approaches traditionally applied in the literature. The returns for mutual fund strategies, such as the $R^2$ and volatility portfolios that we consider in this paper, are often benchmark-adjusted in the literature using an unconditional Carhart (1997) four-factor model regression,

$$R_{i,t} = \alpha_i^U + \beta_i^U R_{MKT,t} + s_i^U R_{SMB,t} + h_i^U R_{HML,t} + u_i^U R_{UMD,t} + \varepsilon_{i,t}, \tag{1}$$

where $R_{i,t}$ is the excess return for portfolio $i$ in month $t$, $R_{MKT,t}$ is the return on the market factor, $R_{SMB,t}$ and $R_{HML,t}$ are the size and value factors of the Fama and French (1993) three-factor model, and $R_{UMD,t}$ is a momentum factor. In the spirit of Sharpe (1992), this approach serves to decompose the realized return for a given strategy into two components—a component that reflects the return that could be earned from suitable exposures to the four passive benchmark factors and a residual component (i.e., the unconditional alpha, $\alpha_i^U$) that is often interpreted as managerial skill. Such an approach, however, inherently assumes that the appropriate benchmark portfolio retains constant exposures to the underlying factors. Whereas this assumption is arguably reasonable for benchmarking an individual mutual fund, the results in the literature for predictors of mutual fund performance are typically based on portfolios of funds. These portfolios require trading in and out of individual funds, which often leads to considerable variation in factor exposures.

Ideally, the benchmarks applied to these strategies in assessing managerial skill would account for any predictable changes in factor exposures over time. Conditional benchmarks are likely to track a given strategy's returns better than unconditional benchmarks if factor exposures are time varying.[4] Further, applying the framework of Hansen and Richard (1987) to mutual fund performance evaluation, it is possible for fund managers to appear to have skill relative to unconditional benchmarks but to show no skill after conditioning on the investor information set. More specifically, we know from the asset pricing literature that an unconditional portfolio alpha may be a biased estimate of the conditional alpha if factor loadings vary systematically with the expected returns (i.e., "factor timing") or volatilities (i.e., "volatility timing") of the factors (see, e.g., Grant (1977), Jagannathan and Wang (1996), Lewellen and Nagel (2006), and Boguth, Carlson, Fisher, and Simutin (2011)). The conditional alpha in this case is a direct indicator of skill in security selection. The unconditional alpha, on the other hand, will also reflect factor timing and volatility timing effects, which may or may not indicate managerial timing skill.[5]

[4]For a simple example, consider a strategy that invests entirely in mutual funds implementing value strategies over the first half of the sample period and invests entirely in growth-oriented funds over the second half of the sample. The resulting estimate of the value factor loading, $\hat{h}_i^U$, from the unconditional benchmark in equation (1) might be close to zero, but such a benchmark would likely perform poorly in tracking the strategy's returns.

[5]For a mutual fund that strategically shifts its exposures to the benchmark factors over time, the intercept from

A standard conditional approach to assessing performance is to estimate a version of the Carhart (1997) model that allows factor loadings to vary over time. We assume that the conditional alpha is constant and that the conditional portfolio factor loadings are linear in a set of conditioning variables (e.g., $\beta_{i,t}^{C} \equiv \lambda_{i,0} + \lambda_{i,1}' Z_{i,t-1}^{MKT}$). Specifically, we measure the conditional alpha of each portfolio using the regression,

$$R_{i,t} = \alpha_i^C + (\lambda_{i,0} + \lambda_{i,1}' Z_{i,t-1}^{MKT}) R_{MKT,t} + (\gamma_{i,0} + \gamma_{i,1}' Z_{i,t-1}^{SMB}) R_{SMB,t} + (\eta_{i,0} + \eta_{i,1}' Z_{i,t-1}^{HML}) R_{HML,t} + (\nu_{i,0} + \nu_{i,1}' Z_{i,t-1}^{UMD}) R_{UMD,t} + \varepsilon_{i,t}, \quad (2)$$

where $Z_{i,t-1}^{k}$ is an $n_{i,k} \times 1$ vector of instruments that can vary across portfolios and factors. Importantly, $Z_{i,t-1}^{k}$ is in the investor information set at the beginning of period $t$.

This method of measuring mutual fund performance using conditional factor models was first developed by Ferson and Schadt (1996). The traditional approach in the literature is to use macroeconomic state variables, such as the dividend yield and interest-rate related variables, as instruments for portfolio factor loadings. Our innovation is to borrow from recent advances in the asset pricing literature and add lagged averages of estimated factor loadings for constituent funds as conditioning variables.

For some background on this approach, Lewellen and Nagel (2006) advocate measuring conditional portfolio risk with contemporaneous short-window regression betas to avoid problems associated with conditioning on a subset of the information available to investors. Boguth, Carlson, Fisher, and Simutin (2011) show that this method may cause an "overconditioning" bias because the short-window betas are not known to investors at the beginning of the period. Lagged portfolio loadings are, however, in the investor information set, and Boguth, Carlson, Fisher, and Simutin (2011) demonstrate that these variables serve as good instruments for factor exposures of stock portfolios in tests similar to equation (2).[6]

We use both short-term and longer-term factor loading estimates as conditioning variables in equation (2) following Boguth, Carlson, Fisher, and Simutin (2011). To estimate the lagged short-term factor loading instruments for each strategy portfolio, we first estimate the loadings for each of the mutual funds in the portfolio using an unconditional Carhart (1997) model regression with

equation (1) will reflect ability in both timing and security selection. For portfolios of mutual funds, however, an unconditional alpha that results from predictable changes in factor loadings attributable to portfolio rebalancing is not direct evidence of managerial skill in timing the factors.

[6] An alternative approach that employs lagged portfolio betas as a direct proxy, rather than as instruments for conditional portfolio risk exposures, is also problematic if portfolio risk changes predictably between the prior period and the current holding period (see, e.g., Chan (1988), Grundy and Martin (2001), and Boguth, Carlson, Fisher, and Simutin (2011)).

daily data over the most recent three-month period. The average factor loadings across funds in a strategy portfolio are estimates of the lagged three-month loadings $\beta_{t-1}^{L3}$, $s_{t-1}^{L3}$, $h_{t-1}^{L3}$, and $u_{t-1}^{L3}$ for the portfolio.[7] We also construct longer-term factor loading instruments for each portfolio as the averages of estimated loadings for each fund from a monthly return regression over the most recent 24 months. These instruments are denoted $\beta_{t-1}^{L24}$, $s_{t-1}^{L24}$, $h_{t-1}^{L24}$, and $u_{t-1}^{L24}$. In our empirical tests, we include the three-month and 24-month lagged factor loadings as conditioning variables for the matching factor loading in equation (2).

Estimating conditional benchmarks using lagged loadings as instruments has not yet been applied in the mutual fund literature, but the method is particularly well suited for this setting. Directly using recent factor loadings to predict exposures will perform better when the factor loading estimates have low levels of measurement error. Because mutual funds are diversified portfolios, the factor loadings of funds and strategies that form portfolios of funds can be estimated relatively precisely over short periods. Further, the mutual fund strategies evaluated in the literature often require frequent rebalancing and changes in the identity of constituent funds. The lagged loading instruments for each portfolio are based on the exposures of mutual funds that are currently in the portfolio, so these instruments are designed to rapidly adjust to the inclusion of new funds. Traditional instruments based on macroeconomic variables, on the other hand, tend to move at business cycle or lower frequencies, such that they may provide a poor fit to the short-term movements in factor loadings that result from rebalancing. Finally, choosing instruments for conditional models is subjective and can lead to data-mining concerns (e.g., Ferson, Sarkissian, and Simin (2008) and Cooper and Gubellini (2011)), whereas using lagged loading estimates as conditioning information for factor loadings removes much of the subjectivity from the method.

In our empirical analysis, we measure the unconditional and conditional performance of strategies based on mutual fund $R^2$ and volatility. We estimate the models in equations (1) and (2) using the generalized method of moments (GMM). Each of these regression models is exactly identified, and the GMM parameter estimates correspond to ordinary least squares estimates. We estimate standard errors using the approach in Newey and West (1987) with five lags to account for potential heteroskedasticity and autocorrelation. Our main tests assess whether the conditional alpha of the low-minus-high portfolio, $\alpha_{LH}^{C}$, for each strategy is equal to zero. We also test whether these conditional alphas are significantly smaller than their unconditional counterparts by testing the null hypothesis $\alpha_{LH}^{C} \geq \alpha_{LH}^{U}$ against the alternative $\alpha_{LH}^{C} < \alpha_{LH}^{U}$.[8] Finally, we compare inferences

[7]This approach to calculating portfolio betas is referred to as the "lagged component" approach by Boguth, Carlson, Fisher, and Simutin (2011).

[8]This test involves estimating the models in equations (1) and (2) in a single GMM procedure. See Appendix A.5

from conditional models that use lagged factor loading instruments with models using traditional instruments from the prior literature in equation (2).

## 3 Data and summary statistics

Section 3.1 provides details on our mutual fund sample, and Section 3.2 describes the construction of the portfolios of mutual funds formed on lagged $R^2$ and lagged volatility that are the basis for our empirical tests. Section 3.3 presents summary statistics for the strategies of interest and motivates the use of conditional models to evaluate portfolio performance. Section 3.4 discusses the traditional instruments that are used in some of our conditional models.

### 3.1 Sample construction

Our objective in constructing the mutual fund sample is to reproduce the empirical results from prior literature on the predictive content of $R^2$ and volatility for future fund performance. As such, we largely follow the sample selection methods presented in Amihud and Goyenko (2013) and note that the sample screens in Jordan and Riley (2015) are relatively similar.

We obtain data on monthly mutual fund returns from the Center for Research in Security Prices (CRSP) Survivor-Bias-Free US Mutual Fund Database for the period January 1988 to December 2014. These returns are net of fees, expenses, and brokerage commissions but before any front-end or back-end loads. We convert all net returns to excess returns by subtracting the corresponding risk-free rate.[9] We also collect data on fund characteristics, including total net assets, expense ratio, turnover, and percentage of stocks in the portfolio. We use the MFLINKS database to identify funds with multiple share classes and combine these share classes into portfolios. A fund's total net assets for a given period is the sum of total net assets across share classes, and the fund's returns and other characteristics are asset-weighted averages.

To limit the sample to actively managed equity funds, we follow the approach in Amihud and Goyenko (2013) and screen on the investment objective codes from Wiesenberger, Lipper, Strategic Insight, and Spectrum. We manually check the dataset for index funds and eliminate these observations from the sample. We also eliminate balanced funds, international funds, sector funds, funds with missing names, and funds that have less than 70% of their holdings on average

in Boguth, Carlson, Fisher, and Simutin (2011) for estimation details.

[9] We obtain data on the daily and monthly risk-free rate from Kenneth French's website. See http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/. We thank Kenneth French for making these data available.

in common stocks. We include two additional screens to address potential concerns related to incubation bias (e.g., Evans (2010)). First, we include funds in the sample only after their total net assets reaches $15 million for the first time. Second, we delete any fund-month observation that precedes the fund's first offer date from CRSP.

A number of our tests require data on daily fund returns from the CRSP daily mutual fund return file, which starts in September 1998. We also convert these daily mutual fund returns to returns in excess of the risk-free rate. Finally, we obtain daily and monthly time-series data on the market, size, value, and momentum factors for the Carhart (1997) four-factor model from Kenneth French's website.

## 3.2 Mutual fund portfolios

Section 3.2.1 introduces the mutual fund strategies based on $R^2$, and Section 3.2.2 details the strategies related to mutual fund return volatility.

### 3.2.1 $R^2$ portfolios

We closely follow the portfolio formation procedures in Amihud and Goyenko (2013) to construct our trading strategies related to mutual fund $R^2$. At the beginning of each month, we estimate fund-by-fund regressions of mutual fund excess returns on the four Carhart (1997) factors using the prior 24 months of data.[10] We then rank funds based on $R^2$ from these regressions and eliminate funds that rank below the 0.5th percentile or above the 99.5th percentile.[11] The remaining funds are sorted into five groups based on lagged $R^2$. These portfolios are equal weighted and rebalanced monthly, and our empirical tests examine the performance of these strategies over the period January 1990 to December 2014. Amihud and Goyenko (2013) find that low-$R^2$ funds tend to outperform high-$R^2$ funds, and much of the discussion in their paper focuses on the performance of a hypothetical "low-minus-high" strategy that takes a long position in the low-$R^2$ quintile and a short position in the high-$R^2$ quintile.

The $R^2$ strategy described above does not, however, correspond to the strongest result in Amihud and Goyenko (2013). They specifically show that the predictive content of $R^2$ for future performance is considerably more pronounced among mutual funds with strong prior performance. To replicate these results, we conduct a conditional sort. As above, we first sort mutual funds each

[10]We require a fund to have a valid return for each of the prior 24 months to be included in these portfolios.

[11]As Amihud and Goyenko (2013) note, actively managed funds with extremely high $R^2$ values are essentially closet indexers. Conversely, for funds with extremely low $R^2$ values, the performance attribution model is likely to be grossly misspecified.

month into quintiles based on prior $R^2$ (after censoring the extreme tails of the distribution at the 0.5th and 99.5th percentiles). Within each of these groups, we subsequently sort funds into five groups based on lagged four-factor alpha, obtained from the same Carhart (1997) regression used to estimate $R^2$. This sequential sort generates 25 portfolios, but we only report results for the five high-alpha groups (i.e., the highest alpha quintile within each $R^2$ quintile). Consistent with the findings in Amihud and Goyenko (2013), we also find that the $R^2$ effect is much stronger for high-alpha funds, and most of our empirical work focuses on these conditional strategies.

#### 3.2.2 Volatility portfolios

The volatility strategies are motivated by the evidence in Jordan and Riley (2015) that funds with low past return variance tend to outperform those with high past return variance in subsequent periods. Following Jordan and Riley (2015), we sort sample funds into 10 groups based on lagged standard deviation of daily net returns from the prior 12 months. A mutual fund must have 100 or more valid return observations over the formation period to be included in this trading strategy. These decile portfolios are equal weighted and rebalanced monthly, and we also examine the performance of a hypothetical trading strategy that takes a long position in the low-volatility group and a short position in the high-volatility group. Given that the CRSP daily mutual fund file starts in September 1998, we use 1999 as the initial formation period for the volatility strategies, and our empirical results focus on portfolio performance over the period January 2000 to December 2014. The start of this sample period is identical to the one used in Jordan and Riley (2015).

### 3.3 Portfolio summary statistics

Panel A of Table I reproduces the main results from Amihud and Goyenko (2013) on the association between multifactor model $R^2$ and future mutual fund performance. The table specifically reports results from unconditional Carhart (1997) model regressions for the $R^2$-sorted portfolios. Panel A.1 presents unconditional alpha estimates in percentage per year across quintiles formed from a one-way sort on $R^2$. The low-$R^2$ portfolio earns a modest alpha of 0.24% per year, whereas the high-$R^2$ strategy earns an alpha of -1.41%. The difference in performance of 1.65% is marginally significant at the 10% level using a two-tailed test (Newey–West (1987) corrected $t$-statistic of 1.67).

Consistent with Amihud and Goyenko (2013), however, we find substantially more predictive content for $R^2$ among funds with high prior alphas. Panel A.2 shows the relation between $R^2$ and performance for these conditional sorts. The low-$R^2$ strategy generates an alpha of 2.66% per year compared with -1.06% for high-$R^2$ funds. The hypothetical low-minus-high portfolio earns a statistically significant abnormal return of 3.72% ($t$-statistic of 2.38). The reported unconditional factor loadings suggest that the low-minus-high strategy is tilted modestly toward low-$\beta$ stocks, value stocks, and past winners, and somewhat more strongly toward small-cap stocks. Given the nature of the results in Panel A of Table I, in the remainder of the paper we focus on explaining the stronger $R^2$-performance relation among high-alpha funds.

In Panel B of Table I, we find, consistent with the evidence in Jordan and Riley (2015), that lagged volatility exhibits a pronounced inverse relation with mutual fund performance. The lowest-volatility decile portfolio generates an alpha of 1.53% relative to the Carhart (1997) model, whereas the high-volatility group exhibits benchmark-adjusted returns of -4.11%. The difference in unconditional performance for these two portfolios is 5.64% per year, which is statistically significant at conventional levels ($t$-statistic of 3.19). From the reported factor loadings, we see that the low-minus-high volatility strategy is weighted toward low-$\beta$ stocks, large-cap stocks, and value stocks.

To provide a deeper understanding of the strategies of interest, Table II presents portfolio summary statistics. Panel A shows average net return, average gross return, and standard deviation of net return for the extreme $R^2$ and volatility groups as well as the corresponding difference portfolios. Consistent with the results for unconditional alphas from Panel A.2 of Table I, the low-minus-high $R^2$ portfolio provides an average net return of 5.02% per month and an average gross return of 5.38%. The low- and high-$R^2$ quintiles show similar levels of realized return volatility. For the volatility decile portfolios, the difference in average net (gross) returns is 3.95% (3.74%), and, not surprisingly, the high-volatility strategy exhibits a considerably higher standard deviation relative to the low-volatility portfolio. Panel B reports average characteristics of the mutual funds held in each portfolio. Low-$R^2$ funds tend to be smaller than their high-$R^2$ counterparts and also have higher fund-level turnover and expense ratios.

More importantly, Panel C of Table II shows that the investments based on multifactor model $R^2$ require considerable strategy-level turnover. For example, an investor pursuing Amihud and Goyenko's (2013) proposed strategy of investing in low-$R^2$ mutual funds with high past alphas would see turnover of 280% per year. The high-$R^2$ portfolio exhibits even higher annualized turnover at 315%. These results highlight the dynamic nature of these strategies and have potentially critical implications for the evaluation of their performance. In particular, the high turnover among the extreme $R^2$ groups suggests that the identity and characteristics of the constituent funds are likely to change quite significantly over time. As such, the unconditional risk exposures presented in Table I used to benchmark strategy performance may mask considerable time variation in style exposures for the $R^2$ portfolios over the full sample period.

We see direct evidence of this effect in Figure 1 and Panel D of Table II. Figure 1 shows the three-month and 24-month lagged loading estimates for each of the four factors in the Carhart (1997) model for the $R^2$ portfolios. The lagged three-month (24-month) factor loadings for a given month are estimated fund-by-fund using the prior three months of daily returns (24 months of monthly returns) and then averaged across the constituent funds. As described in Section 2, these variables serve as instruments for portfolio factor exposures in our empirical tests. The three-month lagged betas in Figure 1 begin in January 1999 based on the availability of daily mutual fund return data. Panel D of Table II shows time-series properties of the 24-month lagged factor loadings for each portfolio.

Several of the portfolio factor loadings in Figure 1 show pronounced shifts and trends across time, which provides direct motivation for using a conditional version of the Carhart (1997) model for performance evaluation. Starting with the results for the low-$R^2$ portfolio, we see that this strategy produces large swings in style exposures. For example, Panel D of Table II reports that the lagged 24-month loading on the value factor for the low-$R^2$ portfolio ranges from -0.90, indicating an extreme growth tilt, to 0.43, suggesting a pronounced exposure to value stocks. The loadings for the market, size, and momentum factors show shifts that are similar in magnitude, and the high-$R^2$ portfolio also produces qualitatively similar changes in style exposures. Note that these large swings in loadings are unlikely to be attributable to estimation error. Factor loadings for mutual funds tend to be estimated quite precisely compared to, say, individual stocks, and the estimates presented in Figure 1 and Panel D of Table II are also averages across mutual funds in a given portfolio.

Panels C and D of Table II also show that these issues are relevant for the portfolios formed on lagged volatility. Although the low- and high-volatility strategies require less trading in comparison to the $R^2$ portfolios, the annualized turnovers for the extreme volatility deciles of 87% and 96%, respectively, are still substantial. For the low-minus-high volatility portfolio, each of the four lagged factor loadings takes on an even wider range of values in comparison to the corresponding low-minus-high $R^2$ portfolio loading. Figure 2 shows the lagged three-month and 24-month factor loading estimates for the volatility portfolios. This figure clearly demonstrates the effects of portfolio turnover and shifts in fund style, as we observe several large discrete changes in factor loadings during the sample period. The lagged three-month loading on the momentum factor for the high-volatility portfolio (Panel B of Figure 2), for example, drops from 0.24 in April 2001 to -0.58 in May 2001. As we demonstrate below, accounting for these predictable shifts in factor exposures is critical to assessing portfolio performance.

To gain additional perspective on the predictive ability of $R^2$ and volatility, we plot the time series of returns for the corresponding low-minus-high portfolios in Figures 3 and 4, respectively. Figure 3 shows that the performance for the $R^2$ strategy is highly concentrated around the middle of the 1990 to 2015 sample period. In particular, the most extreme positive returns occur over a four-year period from 1999 to 2002, and three of the four highest monthly returns are realized in a four-month span from November 1999 to February 2000. Figure 4 provides a similar characterization of the volatility effect in mutual fund returns. The evidence on low-volatility funds outperforming high-volatility funds appears concentrated in the initial months of the sample period, as the seven highest monthly returns are realized between March 2000 and March 2001. We also see instances of extreme negative performance around this period, with the low-minus-high volatility portfolio earning returns of -24.9% in February 2000 and -17.7% in June 2000.[12] Notably, the period over which the $R^2$ and volatility strategies realized this extreme performance is also marked by volatile factor returns, such that properly measuring conditional factor exposures is important for assessing managerial skill.

### 3.4 Traditional state variable construction

As noted above, we consider conditional models that are based on either lagged factor loading instruments or traditional instruments used in prior literature. These traditional state variables include the dividend yield, default spread, and term spread. The dividend yield is the sum of dividends accruing to the CRSP value-weighted market portfolio over the prior 12 months divided by the current index level. The default spread is the difference in yields between Moody's Baa- and Aaa-rated bonds, and the term spread is the difference between the 10-year Treasury constant maturity rate and the one-year Treasury constant maturity rate. All bond yields are obtained from the Federal Reserve Bank of St. Louis website.[13]

[12] In untabulated results, we conduct subperiod analyses of the performance of the $R^2$ and volatility strategies. For each strategy, we divide the sample into three equal subsamples. The $R^2$ effect is highly concentrated in the middle subperiod covering May 1998 to August 2006, over which the low-minus-high $R^2$ portfolio earns an unconditional four-factor alpha of 9.14% per year ($t$-statistic of 2.58). In contrast, the unconditional alpha estimates for the low-minus-high $R^2$ strategy in the early and late subperiods are 0.10% and -0.39%, respectively. The low-minus-high volatility strategy also has concentrated performance, earning an annualized alpha of 6.14% in the 2000 to 2004 subperiod ($t$-statistic of 2.35) versus insignificant unconditional alphas of 0.07% and 0.96% in the subsequent two subperiods.

[13] See http://research.stlouisfed.org/fred2/.

# 4 Results

In this section, we apply the empirical methods introduced in Section 2 to reexamine the performance of the $R^2$ and volatility strategies from prior literature. Sections 4.1 to 4.3 present the conditional performance evaluation results and compare our approach based on lagged factor loadings to those adopted in prior studies using traditional instruments. Section 4.4 introduces a decomposition to understand the sources of unconditional alphas for the mutual fund portfolios.

## 4.1 Conditional performance evaluation

Table III reports results from measuring strategy performance using the conditional performance evaluation approach. Panel A shows parameter estimates for the $R^2$ strategy over the sample period January 1999 to December 2014, and Panel B of Table III reports the corresponding figures for the volatility strategy for the period from January 2000 to December 2014. Case 1 in each panel represents the unconditional estimation in which factor loadings are constant. Thus, the results for the volatility portfolios in Panel B correspond to the unconditional Carhart (1997) model results in Panel B of Table I. The shorter sample period for the $R^2$ portfolios relative to the January 1990 to December 2014 period used in the prior tables is based on data availability for the daily returns used to estimate the three-month lagged factor loading instruments. Consistent with the evidence in Figure 3, the Case 1 results in Panel A of Table III demonstrate that the relation between $R^2$ and unconditional alpha is stronger over the 1999 to 2014 period. The remaining three cases in each panel introduce lagged factor loading instruments to model time variation in factor exposures. For each model, we report alphas, factor loadings, and $R^2$s for the extreme portfolios. We also show the low-minus-high portfolio alpha and the $p$-value from a test of whether the conditional alpha is smaller than the corresponding unconditional alpha.

We begin with the $R^2$ strategy results in Panel A of Table III. Case 1 shows that the annualized unconditional alpha of this strategy over the 1999 to 2014 period is 5.65% with a $t$-statistic of 2.6. Case 2 introduces the 24-month lagged factor loading instruments for the low-$R^2$ and high-$R^2$ portfolio loadings. Six of the eight lagged beta instruments are significant predictors of the portfolio factor exposures. Several of the coefficients on the lagged factor loading instruments are close to one, which indicates that these instruments are unbiased predictors of the portfolio loadings. Further, the regression $R^2$ for the low-$R^2$ (high-$R^2$) portfolio increases from 91.9% (95.7%) in Case 1 to 95.0% (98.0%) in Case 2. Modeling time variation in factor exposures thus explains a substantial portion of the remaining variation in portfolio returns. The annualized conditional alpha for the low-minus-high portfolio in Case 2 is 3.16% with a $t$-statistic of 2.5. An additional effect of explaining much of the remaining variation in portfolio returns is that portfolio alpha estimates are more precise relative to the unconditional case. Given this increase in test power, the statistical significance of the alpha does not substantially decline despite a 44% reduction in the magnitude of the alpha estimate. Finally, the conditional alpha is lower than the unconditional alpha, and the difference is statistically significant with a $p$-value of 3.6%. This result suggests that modeling time variation in factor exposures has a significant effect on the portfolio alpha.

The three-month lagged factor exposures are introduced as instruments in Case 3. Each of the eight instruments is a significant predictor of its corresponding factor loading. The short-term instruments appear to provide a noticeably better fit for the portfolios' exposures to the momentum factor, which is perhaps unsurprising given the relatively short-term nature of momentum effects. The instruments are important for explaining variation in portfolio returns, and regression $R^2$s are 96.4% and 98.8% for the low-$R^2$ and high-$R^2$ portfolios. The conditional alpha of the $R^2$ strategy is 1.96%, which is statistically significant at the 10% level with a $t$-statistic of 1.8, and this conditional alpha is significantly lower than the unconditional alpha with a $p$-value of 1.0%.

Finally, Case 4 includes both sets of lagged factor loading instruments. The model generally places more weight on the three-month instruments compared to the 24-month instruments, with the exception of the high-$R^2$ portfolio's loading on the size factor. The regression $R^2$s increase to 96.6% for the low-$R^2$ portfolio and 98.8% for the high-$R^2$ portfolio, compared to 91.9% and 95.7% in Case 1, such that modeling time variation in portfolio betas provides a substantially better fit to portfolio returns. The annualized conditional alpha in this case is an insignificant 1.58% with a $t$-statistic of 1.5. The 72% reduction in the magnitude of alpha from the unconditional to the conditional case is highly statistically significant with a $p$-value of 0.7%. Overall, the conditional model produces substantial evidence that portfolio factor loadings are time varying, and modeling this time variation affects inferences about the $R^2$ strategy.

Panel B of Table III shows corresponding results for the volatility strategy. Case 1 confirms the annualized unconditional alpha of 5.64% from Table I. The 24-month instruments in Case 2 are all statistically significant predictors of the portfolio loadings. The regression $R^2$ increases from 94.2% to 97.3% for the low-volatility portfolio and from 94.7% to 96.0% for the high-volatility portfolio. The conditional alpha is 2.70% ($t$-statistic of 2.6), which is statistically smaller than the unconditional alpha with a $p$-value of 1.6%. Case 3 produces similar results with the three-month instruments. The conditional alpha of 1.53% ($t$-statistic of 1.7) represents a 73% reduction in the magnitude of alpha from the unconditional case.

Case 4 in Panel B of Table III reports results with both sets of lagged factor loading instruments. The regression $R^2$ for the low-volatility (high-volatility) portfolio increases from 94.2% (94.7%) to 97.8% (97.7%), such that more than half of the remaining variation in portfolio returns is explained by time-varying factor exposures. The annualized conditional alpha is 1.39%, which is statistically insignificant at the 5% level with a $t$-statistic of 1.7. Finally, this alpha is significantly lower than the unconditional strategy alpha with a $p$-value of 0.3%.

Before proceeding, we note that the regression $R^2$s in Table III demonstrate that using lagged factor loading estimates as instruments for factor exposures substantially improves tracking performance. A byproduct of this improvement is an increase in the precision of alpha estimates as unexplained return variance declines. In our tests, the standard error of the $R^2$ strategy's conditional alpha in Case 4 of Panel A is 1.07% compared to 2.17% for the unconditional alpha. Similarly, the standard error of the conditional alpha for the volatility strategy in Panel B is 0.81% versus 1.77% in the unconditional model. An increase in the precision of an alpha estimate leads to higher power to reject the null hypothesis of no managerial skill. Our method should thus be useful for researchers and investors seeking to identify fund characteristics that can forecast mutual fund performance.

Taken together, the results in Table III suggest that using conditional benchmarks is important for evaluating strategies that predict mutual fund performance and can have an economically meaningful impact on inferences. In particular, the conditional Carhart (1997) model results for the $R^2$ and volatility strategies suggest that the primary driver of performance is not skill in security selection by mutual fund managers. In Section 4.4, we revisit these results to further examine the potential sources of the unconditional alphas earned by the mutual fund strategies.

### 4.2 Comparison of instruments for conditional factor loadings

We now compare the performance of our conditional performance evaluation approach based on lagged factor loadings to the traditional methods in the literature that rely on macroeconomic predictors. We argue in Section 2 that the lagged loading instruments have desirable features for modeling portfolio betas. In Table IV, we investigate whether or not these instruments outperform the dividend yield, default spread, and term spread in tracking strategy returns. Panel A (Panel B) shows results for the $R^2$ (volatility) strategy. For reference, both panels reproduce the unconditional Carhart (1997) model results (i.e., Case 1 in Table III) and the conditional Carhart (1997) model results with short-term and long-term lagged loadings as instruments (i.e., Case 4 in Table III).

In Panel A of Table IV, the first set of results based on traditional instruments corresponds to models that allow only the market factor loading to vary with conditioning information. That is, the market loading is specified as a linear function of the dividend yield, default spread, term spread, or all three of these traditional state variables. This approach to instrumenting only for the market factor in a Carhart (1997) regression is common in the literature (e.g., Kacperczyk, Sialm, and Zheng (2005), Kosowski, Timmermann, Wermers, and White (2006), Huang, Sialm, and Zhang (2011), and Doshi, Elkamhi, and Simutin (2015)) and is motivated by the extensive evidence on the predictability of market returns using these variables. Using the dividend yield as the sole instrument for market beta, the conditional alpha for the low-minus-high $R^2$ strategy decreases to 5.46% from the unconditional estimate of 5.65%. The difference between these alphas is statistically insignificant with a $p$-value of 35.0%. The regression $R^2$s for the low-$R^2$ portfolio and high-$R^2$ portfolio each exhibit a modest increase of just 0.1%. Similarly, the conditional models incorporating the default spread, term spread, or all three of the traditional instruments for market beta leave inferences unchanged from the unconditional case. Notably, none of these conditional alpha estimates is significantly smaller than the unconditional alpha, and the adjusted $R^2$s are close to their corresponding unconditional model values.

We next consider a set of conditional models that allow each of the four factor loadings to vary with the traditional instruments. This approach is adopted, for example, in Kacperczyk, Sialm, and Zheng (2008) and allows for considerably more flexibility than instrumenting for market beta alone. The conditional alphas across these four models range from 3.88% to 5.70%. Although two of these estimates are marginally insignificant at the 5% level, they remain economically large in comparison to the 1.58% conditional alpha from our conditional approach based on lagged factor loadings. Moreover, the largest regression $R^2$ value for the low-$R^2$ (high-$R^2$) portfolio for any of the models in Panel A of Table IV based on traditional instruments is 93.4% (96.3%). The conditional models with lagged factor loading instruments, for comparison, generate $R^2$ values of 96.6% and 98.8% for the strategies of interest. As such, these conditional models demonstrate a superior ability to track portfolio performance and assess underlying managerial skill.

Similar patterns emerge for the volatility portfolios in Panel B of Table IV. The unconditional alpha for the low-minus-high strategy in this case is 5.64% per year. The models that allow only market beta to vary with the conditioning variables generate alpha estimates between 3.69% and 5.58%. All of these estimates are statistically significant and economically large in comparison to the 1.39% alpha from the conditional model with lagged factor loading instruments. The conditional models that instrument for all four factor loadings with traditional state variables produce a wide range of alphas between 1.80% and 5.28%. One of these four estimates is statistically insignificant and two are significantly smaller than the corresponding unconditional alpha. These results highlight a well-known concern with traditional conditioning methods that inferences can be sensitive to the choice of the information set. Nonetheless, none of the conditional models with traditional state variables produces a long-short alpha as low as the 1.39% estimate from the conditional model that incorporates lagged loadings. As in Panel A, our conditional approach also leads to the highest adjusted $R^2$ value for each portfolio.

Overall, the results in Tables III and IV support the use of lagged portfolio factor loadings as instruments for factor exposures. The lagged beta instruments are highly significant predictors of portfolio factor loadings. Using these instruments also provides a much better fit to portfolio returns, as indicated by regression $R^2$ values, compared to the conditional benchmarks with traditional instruments. This feature is critical, as test power is substantially improved as a result. Further, alternative traditional instruments produce different inferences about conditional alphas, such that the potential for data mining may arise through the subjective choice of conditioning variables.

To provide additional insight on the benefits of the conditional approach with lagged loading instruments, we plot the conditional momentum loadings for the $R^2$ and volatility strategies of interest in Figure 5. In each plot, the solid line shows the momentum loading from the model with 24-month and three-month lagged factor loadings as instruments, and the dashed line shows the momentum loadings as function of the dividend yield, default spread, and term spread. The dashed lines, therefore, correspond to the conditional models in the last line of each panel in Table IV. The key takeaway from the figure is that the benchmarks with lagged loading instruments are able to capture more of the time-series variation in strategy exposures. For example, the conditional momentum loading for the high-volatility portfolio in Panel B of Figure 5 exhibits substantial volatility early in the sample period, moving from 0.62 in November 2000 to -0.57 in June 2001 to 0.57 in January 2004. In comparison, the conditional momentum exposure from the traditional approach is considerably more stable over this period, ranging from -0.11 to 0.22. The traditional instruments tend to be relatively persistent and, therefore, appear less reliable in capturing dramatic shifts in portfolio exposure. Moreover, because the predictive content of volatility for future mutual fund returns is highly concentrated during this early part of the sample (e.g., Figure 4), failing to properly account for the observed shifts in style during this period can lead to misleading inferences about underlying managerial ability.

## 4.3 Full-sample results for the $R^2$ strategy

Our examination of the $R^2$ strategy in Sections 4.1 and 4.2 uses a sample period of January 1999 to December 2014, whereas Amihud and Goyenko (2013) test the strategy over the period January 1990 to December 2010. The sample period of the tests above is limited by the use of daily mutual fund returns to estimate the lagged three-month factor loading instruments. In this section, we expand the sample period for this strategy and develop an approach to using the lagged beta instruments over periods that pre-date the availability of daily mutual fund data.

The three-month lagged factor loading instruments are not available before January 1999, whereas the 24-month lagged beta instruments are available throughout the full sample period. We thus take an approach of using both the three-month and 24-month instruments when they are available and relying only on the 24-month instruments during the early portion of the sample period. We also allow the constant term for each loading and the coefficients on the 24-month instruments to differ in the early and late subperiods, as including the three-month instruments in the late subperiod is likely to affect these other coefficients. To produce this set of instruments, we first construct indicator variables, $1_{I,t}$ and $1_{II,t}$, that take the values of one for January 1990 to December 1998 and January 1999 to December 2014, respectively. We also interact these indicator variables with the three-month and 24-month lagged beta instruments. As an example, the conditional loading for portfolio $i$ on the market factor is

$$\beta_{i,t}^{C} \equiv \lambda_{i,0} + \lambda_{i,1}\beta_{I,t-1}^{L24} + \lambda_{i,2}1_{II,t} + \lambda_{i,3}\beta_{II,t-1}^{L24} + \lambda_{i,4}\beta_{II,t-1}^{L3}, \tag{3}$$

where $\beta_{I,t-1}^{L24} = 1_{I,t}\beta_{t-1}^{L24}$, $\beta_{II,t-1}^{L24} = 1_{II,t}\beta_{t-1}^{L24}$, and $\beta_{II,t-1}^{L3} = 1_{II,t}\beta_{t-1}^{L3}$. The coefficient on the $1_{II,t}$ instrument thus captures the difference in the constant term in the early and late subperiods.

Table V reports unconditional and conditional Carhart (1997) model results over the period January 1990 to December 2014. Panel A shows alpha estimates, regression $R^2$s, and results of a test of whether the conditional alpha is significantly smaller than the unconditional alpha. The annualized unconditional Carhart (1997) model alpha for the low-minus-high portfolio is 3.72% with a $t$-statistic of 2.38. Moving to the conditional model, the estimated alpha is 1.27%, which is insignificant with a $t$-statistic of 1.55. The conditional alpha is significantly lower than the unconditional alpha with a $p$-value of 2.2%. Further, the regression $R^2$ increases substantially from 92.9% (95.4%) for the low-$R^2$ (high-$R^2$) portfolio in the unconditional model to 96.7% (98.6%) in the conditional model. Thus, inferences are similar to those in Table III for the $R^2$ strategy.

Panel B of Table V shows coefficient estimates for the instruments used to model the portfolio factor loadings. The 24-month lagged beta instruments are significant predictors for six of the eight factor exposures in the early part of the sample period. The three-month lagged factor loading instruments appear more influential for forecasting betas in the second part of the sample, as seven of the eight instruments have significantly positive coefficients. Overall, the results in Table V further establish the robustness of our inferences on the $R^2$ effect.

### 4.4 Decomposition and evaluation of strategy performance

Our main results in Section 4.1 show that, whereas the spreads in unconditional alphas for the $R^2$ and volatility strategies are large and significantly positive, the corresponding conditional alphas are substantially smaller in magnitude. Thus, the bulk of the unconditional performance of these strategies is not likely to be attributable to skill in security selection by the mutual fund managers. In this section, we decompose the unconditional alphas of the mutual fund strategies. As shown by Lewellen and Nagel (2006), Boguth, Carlson, Fisher, and Simutin (2011), and others, the difference between the unconditional and conditional alpha of a portfolio is a function of factor timing and volatility timing. In particular, systematic relations between portfolio factor loadings and either the expected returns or volatilities of the factors can produce unconditional alphas that differ from conditional alphas.

For the Carhart (1997) model, the unconditional alpha estimate for a given portfolio can be decomposed as

$$\begin{aligned}
\hat{\alpha}_i^U = \overbrace{\hat{\alpha}_i^C}^{\text{Security selection}} & + \overbrace{\text{cov}(\hat{\beta}_{i,t}^C, R_{MKT,t})}^{\text{Factor timing}} + \overbrace{(\bar{\beta}_{i,t}^C - \hat{\beta}_i^U)\bar{R}_{MKT,t}}^{\text{Volatility timing}} \\
& + \text{cov}(\hat{s}_{i,t}^C, R_{SMB,t}) + (\bar{s}_{i,t}^C - \hat{s}_i^U)\bar{R}_{SMB,t} \\
& + \text{cov}(\hat{h}_{i,t}^C, R_{HML,t}) + (\bar{h}_{i,t}^C - \hat{h}_i^U)\bar{R}_{HML,t} \\
& + \text{cov}(\hat{u}_{i,t}^C, R_{UMD,t}) + (\bar{u}_{i,t}^C - \hat{u}_i^U)\bar{R}_{UMD,t}, \qquad (4)
\end{aligned}$$

where $\hat{\beta}_{i,t}^C$ is the conditional loading on the market factor, $\bar{\beta}_{i,t}^C$ is the average conditional loading, $\hat{\beta}_i^U$ is the unconditional market loading, and the terms for the remaining three factors are defined analogously. For each of the four factors, a direct factor timing term and a factor bias effect term can contribute to differences between the unconditional and conditional portfolio alpha. The direct factor timing terms measure the covariances between factor loadings and factor returns. A positive covariance between a portfolio's exposure to a factor and the factor's realized return will have a positive effect on the measured unconditional alpha. The factor bias terms reflect the differences between the average conditional factor exposures and the unconditional loadings. These terms are related to volatility timing. For example, if the conditional loading on the market factor for a portfolio tends to be high when the market factor is highly volatile, then the portfolio's unconditional market factor loading will overstate its average conditional exposure to the market factor.

Our proposed method for estimating conditional alpha combined with this decomposition of unconditional alpha provides an approach for attribution analysis of a predictor of mutual fund performance. The conditional alpha of a strategy may capture the security selection skill of the managers of mutual funds held in the strategy portfolios.[14] The remainder of the unconditional performance of a strategy is attributable to factor timing or volatility timing, which could potentially, but not necessarily, be indicative of managerial skill as discussed further below. Ferson and Mo (2015) also develop a method for decomposing performance into similar security selection, factor timing, and volatility timing components, and our returns-based approach is complementary to their holdings-based procedure.

Table VI shows results from empirical decompositions of the unconditional alphas for the $R^2$ and volatility portfolios. The conditional models in this table correspond to the Case 4 results from Table III. Beginning with the low-minus-high $R^2$ strategy, the unconditional alpha is 5.65% per year. We decompose this estimate into contributions from security selection, the four direct factor timing terms, and the four factor bias (i.e., volatility timing) effect terms. The $R^2$ strategy portfolio shows large factor timing effects, as these four terms account for 5.46% per year in overall unconditional performance. Timing for the size, value, and momentum factors is particularly strong with contributions of 1.28%, 1.98%, and 2.01% per year, respectively. The factor bias terms, on the other hand, tend to produce a lower unconditional alpha estimate. The estimated unconditional alpha is 1.39% lower because of the factor bias effects, and the unconditional loadings on the size, value, and momentum factors each overstate the average conditional exposures to these factors. The large unconditional alpha of the strategy of 5.65% is thus attributable to a relatively small conditional alpha of 1.58%, which could be related to security selection, along with a large effect of 5.46% from factor timing. Moreover, this factor timing effect reflects an economically large positive timing ability for the low-$R^2$ strategy of 3.45% per year and a large negative timing ability for high-$R^2$ funds at -2.01%.

The results for the volatility strategy are similar in nature. The low-minus-high volatility

[14] As noted by Ferson and Mo (2015) and others, the conditional alpha could also reflect managerial ability to execute low cost trades or manage an efficient securities lending operation.

portfolio's unconditional performance of 5.64% is produced by a relatively small security selection component of 1.39%, a factor timing effect of 4.75%, and a volatility timing effect of -0.50%. Factor timing for the market, size, and value factors are particularly beneficial for unconditional performance with contributions of 1.97%, 0.93%, and 1.97%, respectively. In sum, the results in Table VI suggest that factor timing, rather than security selection, is primarily responsible for the positive unconditional performance estimates of the $R^2$ and volatility strategies.

Given the importance of factor timing for the unconditional performance of these strategies, we consider two additional tests to better characterize the nature of these results. In our first test, we examine whether the positive factor timing performance for each strategy is primarily attributable to the underlying mutual funds or to the portfolio rebalancing procedure. In our second test, we analyze the concentration of each strategy's successful style bets within the sample period to gain perspective on the robustness of the factor timing performance.

We begin by considering the relation between portfolio rebalancing and factor timing for each strategy. Positive factor timing effects for a given portfolio of mutual funds have three potential sources. First, the strategy may tend to invest in managers with the skill to time factors based on information or trading rules that are not publicly known. Second, the strategy may identify mutual fund managers who mechanically follow known timing strategies based on publicly available information. Third, the strategy may shift its investments across mutual funds with different factor exposures when the portfolios are rebalanced such that the strategy's factor exposures change. In our view, only the first of these scenarios lends itself to an interpretation of the portfolio sorting characteristic (e.g., $R^2$ or volatility) as an indicator of managerial skill.

Assuming that the effects of factor timing are relatively stable for mutual fund managers over time (i.e., managers tend to continue to employ a mechanical trading strategy or maintain their timing skill), characteristics of the rebalancing process and performance over alterative holding periods may help to distinguish between fund-level timing and strategy-level timing. We therefore examine the impact of the portfolio holding period on the unconditional performances of the $R^2$ and volatility strategies. Our base results use monthly rebalancing and a one-month holding period. We also consider $n$-month holding periods for $n = 1, \ldots, 24$ by rebalancing $1/n$ of the portfolio weight in each month based on mutual fund $R^2$ or volatility following the portfolio formation rules in Section 3.2. As such, each portfolio in a given month is composed of mutual funds that qualified for inclusion in the portfolio within the past $n$ months.

Figure 6 plots unconditional alphas for the low-minus-high strategies for each monthly holding period length from one to 24 months along with the 95% confidence interval for the alphas. As in Table III, the sample periods are January 1999 to December 2014 for the $R^2$ strategy and January 2000 to December 2014 for the volatility strategy.[15] For the $R^2$ strategy in Figure 6, the abnormal performance declines drastically with the holding period. The unconditional alpha for this low-minus-high portfolio with a one-month holding period is 5.65% per year with a $t$-statistic of 2.61. Moving to a longer period, however, substantially weakens the performance of the $R^2$ strategy. The unconditional alpha is significant only for holding periods up to six months. For example, the unconditional alpha with a 12-month (24-month) holding period is an insignificant 2.14% (1.11%) per year. Considering the dependence on short holding periods and the high portfolio turnover for this strategy discussed in Section 3.3, the results suggest that the positive unconditional performance of the $R^2$ strategy is dependent on timing the factors by relatively rapid rebalancing across mutual funds with differing risk exposures.

In contrast, the performance of the volatility strategy is stable across alternative holding periods. The 12-month and 24-month unconditional alphas of 5.51% ($t$-statistic of 3.10) and 5.37% ($t$-statistic of 3.47), respectively, are similar in magnitude to the one-month unconditional alpha of 5.64% ($t$-statistic of 3.19). As such, our holding-period analysis is unable to rule out the existence of factor timing ability among managers of low-volatility funds. We remain skeptical, however, on the reliability of lagged return volatility as a robust predictor of fund timing ability. In particular, as noted in our discussion of Figure 4, the large unconditional alpha earned by the low-minus-high volatility portfolio is primarily a reflection of the strategy's strong returns during the first two years of the sample period.

Motivated by these concerns, our second set of tests formally investigate whether the factor timing effects for the $R^2$ and volatility strategies are also concentrated over short periods. The factor timing terms in equation (4) are based on sample covariances between conditional factor exposures and factor returns. Taking the market factor as an example, the sample covariance calculation,

$$\text{cov}(\hat{\beta}^C_{i,t}, R_{MKT,t}) = \frac{1}{T}\sum_{t=1}^{T}\hat{\beta}^C_{i,t}R_{MKT,t} - \left(\frac{1}{T}\sum_{t=1}^{T}\hat{\beta}^C_{i,t}\right)\left(\frac{1}{T}\sum_{t=1}^{T}R_{MKT,t}\right), \qquad (5)$$

is dependent on the sum of monthly products of the strategy's exposure to the market factor and the market factor return (i.e., $\hat{\beta}^C_{i,t}R_{MKT,t}$). For a given sample month $\tau$, we can also assess the contribution of a strategy's market timing in months 1 through $\tau$ to its overall market timing ability

[15] For the volatility-sorted portfolios, the formation-period volatility estimates are based on 12 months of prior monthly (daily) excess returns in the pre-January 2000 (post-January 2000) period.

by calculating

$$\text{Cumulative market timing} = \frac{1}{T}\sum_{t=1}^{\tau}\hat{\beta}_{i,t}^{C}R_{MKT,t} - \frac{\tau}{T}\left(\frac{1}{T}\sum_{t=1}^{T}\hat{\beta}_{i,t}^{C}\right)\left(\frac{1}{T}\sum_{t=1}^{T}R_{MKT,t}\right) \qquad (6)$$

and comparing it to the total in equation (5). Summing the right hand side of equation 6 and the analogous relations for the other three factors can provide an indication of which sample months are most informative about factor timing ability. Figure 7 plots this sum for each sample month for the $R^2$ and volatility strategies. To interpret the plots, any upward (downward) trend over a particular period indicates positive (negative) factor timing performance. A strategy with consistent factor timing ability would produce a plot with a positive slope throughout the sample. A strategy with concentrated timing ability, in contrast, may exhibit only a few pronounced upward spikes.

Figure 7 shows that the early portion of the sample period is crucial for producing evidence of factor timing for both long-short portfolios. Focusing on the volatility strategy, nearly all of the full-sample factor timing effect of 4.75% shown in Table VI is attributable to the first two years of the period. The remaining 13 years of the sample produce little evidence of systematic timing ability. The strategy's unconditional alpha, thus, seems to be an artifact of a short run of successful style bets, as opposed to repeated skill in factor timing across the full sample period. Additional analysis shows that the positive factor timing for the low-minus-high volatility strategy in the early portion of the sample is entirely attributable to negative factor timing in the high-volatility portfolio. Thus, the evidence does not support factor timing skill of low-volatility fund managers as an explanation for the positive unconditional performance of the volatility strategy.

Overall, decompositions of the unconditional performance of the $R^2$ and volatility strategies attribute their success over the sample period primarily to factor timing. Neither strategy produces substantial evidence of managerial skill in security selection. Further, the anatomy of the $R^2$ strategy suggests that the portfolio rebalancing procedure, rather than the skill of underlying mutual fund managers, is the source of factor timing. We also show that the positive unconditional performance of each strategy is largely attributable to a short subperiod of successful factor timing. Taken as a whole, our analysis suggests that $R^2$ and volatility do not reliably identify skilled mutual fund managers.

## 5 Conclusion

In this paper, we show that the conventional approach to evaluating portfolios of mutual funds based on unconditional factor-model regressions is problematic. Specifically, these portfolios often require high turnover in their mutual fund holdings and can exhibit both pronounced jumps in style exposures around rebalancing dates as well as predictable trends in these loadings over time. Any evidence of skill in such instances has the potential to be contaminated by a poorly specified benchmark model that fails to account for changes in the portfolios' style exposures.

We introduce a simple returns-based method to evaluating portfolio performance that builds on standard conditional models applied in the literature. This approach successfully incorporates information from lagged factor exposures in assessing managerial skill in security selection, factor timing, and volatility timing. Our method outperforms alternative approaches used in the literature, and the improvements in portfolio return tracking translate to increased power to identify skilled managers. We also demonstrate that the previously documented predictive content of $R^2$ and volatility for future fund alphas is not driven by differences in security selection ability across funds sorted on these characteristics. Our proposed modeling approach should be a useful tool for evaluating similar evidence in past and future research on managed portfolios.

## References

Amihud, Yakov, and Ruslan Goyenko, 2013, Mutual fund's $R^2$ as predictor of performance, *Review of Financial Studies* 26, 667–694.

Avramov, Doron, and Tarun Chordia, 2006, Predicting stock returns, *Journal of Financial Economics* 82, 387–415.

Becker, Connie, Wayne Ferson, David H. Myers, and Michael J. Schill, 1999, Conditional market timing with benchmark investors, *Journal of Financial Economics* 52, 119–148.

Boguth, Oliver, Murray Carlson, Adlai Fisher, and Mikhail Simutin, 2011, Conditional risk and performance evaluation: Volatility timing, overconditioning, and new estimates of momentum alphas, *Journal of Financial Economics* 102, 363–389.

Busse, Jeffrey A., 1999, Volatility timing in mutual funds: Evidence from daily returns, *Review of Financial Studies* 12, 1009–1041.

Busse, Jeffrey A., and Qing Tong, 2012, Mutual fund industry selection and persistence, *Review of Asset Pricing Studies* 2, 245–274.

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

Chan, K. C., 1988, On the contrarian investment strategy, *Journal of Business* 61, 147–163.

Cohen, Randolph B., Joshua D. Coval, and L'uboš Pástor, 2005, Judging fund managers by the company they keep, *Journal of Finance* 60, 1057–1096.

Cooper, Michael J., and Stefano Gubellini, 2011, The critical role of conditioning information in determining if value is riskier than growth, *Journal of Empirical Finance* 18, 289–305.

Cremers, K. J. Martijn, and Antti Petajisto, 2009, How active is your fund manager? A new measure that predicts performance, *Review of Financial Studies* 22, 3329–3365.

Doshi, Hitesh, Redouane Elkamhi, and Mikhail Simutin, 2015, Managerial activeness and mutual fund performance, *Review of Asset Pricing Studies* 5, 156–184.

Evans, Richard B., 2010, Mutual fund incubation, *Journal of Finance* 65, 1581–1611.

Fama, Eugene F., and Kenneth R. French, 1989, Business conditions and expected returns on stocks and bonds, *Journal of Financial Economics* 25, 23–49.

——— , 1993, Common risk factors in the returns on bonds and stocks, *Journal of Financial Economics* 33, 3–53.

Ferson, Wayne E., 2010, Investment Performance Evaluation, *Annual Reviews of Financial Economics* 2, 207–234.

——— , 2013, Conditional performance evaluation, in Cheng-Few Lee, and Alice C. Lee, eds.: *Encyclopedia of Finance* (Springer), United States, chap. 11, pp. 279–285.

Ferson, Wayne E., and Haitao Mo, 2015, Performance measurement with selectivity, market and volatility timing, Forthcoming in *Journal of Financial Economics.*

Ferson, Wayne E., Sergei Sarkissian, and Timothy Simin, 2008, Asset pricing models with conditional betas and alphas: The effects of data snooping and spurious regression, *Journal of Financial and Quantitative Analysis* 43, 331–354.

Ferson, Wayne E., and Rudi W. Schadt, 1996, Measuring fund strategy and performance in changing economic conditions, *Journal of Finance* 51, 425–461.

Grant, Dwight, 1977, Portfolio performance and the "cost" of timing decisions, *Journal of Finance* 32, 837–846.

Grundy, Bruce D., and J. Spencer Martin, 2001, Understanding the nature and the risks of the source of the rewards to momentum investing, *Review of Financial Studies* 14, 29–78.

Hansen, Lars Peter, and Scott F. Richard, 1987, The role of conditioning information in deducing testable restrictions implied by dynamic asset pricing models, *Econometrica* 55, 587–613.

Henriksson, Roy D., 1984, Market timing and mutual fund performance: An empirical investigation, *Journal of Business* 57, 73–96.

Henriksson, Roy D., and Robert C. Merton, 1981, On market timing and investment performance. II. Statistical procedures for evaluating forecasting skills, *Journal of Business* 54, 513–533.

Huang, Jennifer, Clemens Sialm, and Hanjiang Zhang, 2011, Risk shifting and mutual fund performance, *Review of Financial Studies* 24, 2575–2616.

Jagannathan, Ravi, and Zhenyu Wang, 1996, The conditional CAPM and the cross-section of expected returns, *Journal of Finance* 51, 3–53.

Jiang, George J., Tong Yao, and Tong Yu, 2007, Do mutual funds time the market? Evidence from portfolio holdings, *Journal of Financial Economics* 86, 724–758.

Jordan, Bradford D., and Timothy B. Riley, 2015, Volatility and mutual fund manager skill, *Journal of Financial Economics* 118, 289–298.

Kacperczyk, Marcin, and Amit Seru, 2007, Fund manager use of public information: New evidence on managerial skills, *Journal of Finance* 62, 485–528.

Kacperczyk, Marcin, Clemens Sialm, and Lu Zheng, 2005, On the industry concentration of actively managed equity mutual funds, *Journal of Finance* 60, 1983–2011.

——— , 2008, Unobserved actions of mutual funds, *Review of Financial Studies* 21, 2379–2416.

Kacperczyk, Marcin, Stijn Van Nieuwerburgh, and Laura Veldkamp, 2014, Time-varying fund manager skill, *Journal of Finance* 69, 1455–1484.

Kosowski, Robert, Allan Timmermann, Russ Wermers, and Hal White, 2006, Can mutual fund "stars" really pick stocks? New evidence from a bootstrap analysis, *Journal of Finance* 61, 2551–2595.

Lewellen, Jonathan, and Stefan Nagel, 2006, The conditional CAPM does not explain asset-pricing anomalies, *Journal of Financial Economics* 82, 289–314.

Massa, Massimo, and Rajdeep Patgiri, 2009, Incentives and mutual fund performance: Higher performance or just higher risk taking?, *Review of Financial Studies* 22, 1777–1815.

Newey, Whitney K., and Kenneth D. West, 1987, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703–708.

Petajisto, Antti, 2013, Active share and mutual fund performance, *Financial Analysts Journal* 69, 79–93.

Sharpe, William F., 1992, Asset allocation: Management style and performance measurement, *Journal of Portfolio Management* 18, 7–19.

Sun, Zheng, Ashley Wang, and Lu Zheng, 2012, The road less traveled: Strategy distinctiveness and hedge fund performance, *Review of Financial Studies* 25, 96–143.

Titman, Sheridan, and Cristian Tiu, 2011, Do the best hedge funds hedge?, *Review of Financial Studies* 24, 123–168.

Treynor, Jack L., and Kay K. Mazuy, 1966, Can mutual funds outguess the market?, *Harvard Business Review* 44, 131–136.

Wei, Kelsey D., Russ Wermers, and Tong Yao, 2015, Uncommon value: The characteristics and investment performance of contrarian funds, *Management Science* 61, 2394–2414.

**Table I: Unconditional four-factor regressions.**
The table reports unconditional Carhart (1997) four-factor regression results for quintile portfolios sorted on lagged $R^2$ (Panel A) and decile portfolios sorted on lagged volatility (Panel B). Panel A.1 presents results for one-dimensional sorts on lagged four-factor model $R^2$, estimated from the prior 24 months of data. Panel A.2 presents results for conditional sorts. Funds are first sorted into quintiles based on lagged $R^2$. Within each of these groups, funds are then sorted into quintiles based on lagged four-factor alpha. The results in Panel A.2 correspond to the portfolios with the highest prior alphas within each $R^2$ group. "L" refers to the low-$R^2$ portfolio, "H" refers to the high-$R^2$ portfolio, and "L−H" refers to their difference. Panel B presents results for one-dimensional sorts on lagged mutual fund volatility, estimated from the prior 12 months of daily returns. "L" refers to the low-volatility portfolio, "H" refers to the high-volatility portfolio, and "L−H" refers to their difference. The portfolios are equal weighted and rebalanced monthly. The unconditional alpha estimates ($\alpha_i^U$) are reported in percentage per year, and the numbers in parentheses are Newey–West (1987) corrected $t$-statistics with a lag length equal to five. The sample period for Panel A (Panel B) is January 1990 to December 2014 (January 2000 to December 2014).

| Panel A: Quintile portfolios sorted on $R^2$ | | | | | | |
|---|---|---|---|---|---|---|
| | $R^2_{t-1}$ | | | | | |
| | L | 2 | 3 | 4 | H | L−H |
| Panel A.1: Results for all funds | | | | | | |
| Alpha, $\alpha_i^U$ (%) | 0.24 | -0.60 | -0.81 | -1.41 | -1.41 | 1.65 |
| | (0.24) | (-0.82) | (-1.46) | (-2.67) | (-3.40) | (1.67) |
| Panel A.2: Results conditional on high prior alpha | | | | | | |
| Alpha, $\alpha_i^U$ (%) | 2.66 | 0.64 | -0.69 | -1.20 | -1.06 | 3.72 |
| | (1.98) | (0.68) | (-0.99) | (-1.39) | (-1.54) | (2.38) |
| Portfolio factor loadings | | | | | | |
| $R_{MKT}$ loading | 0.99 | 1.03 | 1.03 | 1.05 | 1.04 | -0.04 |
| $R_{SMB}$ loading | 0.41 | 0.42 | 0.30 | 0.25 | 0.07 | 0.34 |
| $R_{HML}$ loading | -0.05 | -0.12 | -0.09 | -0.15 | -0.11 | 0.06 |
| $R_{UMD}$ loading | 0.10 | 0.07 | 0.03 | 0.01 | -0.01 | 0.11 |

| Panel B: Decile portfolios sorted on volatility | | | | | | |
|---|---|---|---|---|---|---|
| | $\sigma^2_{t-1}$ | | | | | |
| | L | 3 | 5 | 7 | H | L−H |
| Alpha, $\alpha_i^U$ (%) | 1.53 | 0.96 | -0.30 | -1.05 | -4.11 | 5.64 |
| | (1.70) | (1.10) | (-0.44) | (-1.26) | (-2.98) | (3.19) |
| Portfolio factor loadings | | | | | | |
| $R_{MKT}$ loading | 0.75 | 0.90 | 0.98 | 1.05 | 1.28 | -0.53 |
| $R_{SMB}$ loading | 0.03 | -0.02 | 0.03 | 0.22 | 0.62 | -0.58 |
| $R_{HML}$ loading | 0.24 | 0.22 | 0.10 | 0.00 | -0.31 | 0.55 |
| $R_{UMD}$ loading | 0.01 | -0.02 | -0.01 | 0.02 | 0.05 | -0.03 |

**Table II: Summary statistics.**

The table reports summary statistics for quintile portfolios sorted on lagged $R^2$ and decile portfolios sorted on lagged volatility. The $R^2$ strategies are based on a double sort on lagged Carhart (1997) four-factor regression $R^2$ and alpha as described in the text. The low-$R^2$ and high-$R^2$ portfolios in the table are those conditioned on having a high prior alpha. For each portfolio, Panel A presents the average net excess return, average gross excess return, and standard deviation of net return in percentage per year. Panel B presents properties of the mutual funds contained in each portfolio. "TNA" is total net assets, "Turnover" is the minimum of aggregated sales or purchases of securities divided by the average 12-month TNA of the fund, "Expense ratio" is the ratio of total investment that shareholders pay for the fund's operating expenses to assets under management, $R^2$ is the $R^2$ value from a Carhart (1997) four-factor regression over the prior 24 months, and "Standard deviation" is the standard deviation of monthly net excess returns over the prior 12 months. The figures in Panel B are time-series averages of the monthly cross-sectional average characteristics for each portfolio. Panel C reports annualized portfolio turnover. Monthly turnover is computed as 0.5 times the sum of the absolute values of the change in portfolio weights in each underlying mutual fund. The annual turnover figures are computed by multiplying the monthly turnover values by 12. Panel D presents time-series properties of the equal-weighted, formation-period factor loadings from Carhart (1997) model regressions using 24 months of prior monthly data.

| | $R^2$ quintile portfolios | | | Volatility decile portfolios | | |
|---|---|---|---|---|---|---|
| | 1990-2014 | | | 2000-2014 | | |
| | L | H | L−H | L | H | L−H |
| **Panel A: Properties of portfolio excess returns (annualized)** | | | | | | |
| Average return (Net, %) | 11.49 | 6.47 | 5.02 | 5.99 | 2.03 | 3.95 |
| Average return (Gross, %) | 12.84 | 7.47 | 5.38 | 7.13 | 3.39 | 3.74 |
| Standard deviation (Net, %) | 17.35 | 16.64 | 7.13 | 12.14 | 25.01 | 17.01 |
| **Panel B: Properties of underlying mutual funds** | | | | | | |
| TNA ($MM) | 1,066.21 | 2,318.84 | | 2,058.41 | 668.74 | |
| Turnover (%) | 90.81 | 68.07 | | 65.02 | 123.71 | |
| Expense ratio (%) | 1.36 | 1.02 | | 1.19 | 1.43 | |
| $R^2$ (%) | 81.09 | 96.87 | | 88.18 | 91.33 | |
| Standard deviation (Monthly, %) | 4.92 | 4.77 | | 3.43 | 7.27 | |
| **Panel C: Portfolio turnover (annualized)** | | | | | | |
| Turnover (%) | 279.75 | 315.22 | | 86.66 | 96.48 | |
| **Panel D: Properties of formation-period factor loadings** | | | | | | |
| $R_{MKT}$ loading | | | | | | |
| Average | 0.93 | 0.99 | -0.07 | 0.83 | 1.16 | -0.33 |
| Minimum | 0.74 | 0.84 | -0.31 | 0.70 | 0.89 | -1.03 |
| Maximum | 1.27 | 1.26 | 0.26 | 1.01 | 1.75 | -0.06 |
| $R_{SMB}$ loading | | | | | | |
| Average | 0.35 | 0.18 | 0.17 | 0.03 | 0.63 | -0.60 |
| Minimum | 0.00 | -0.22 | -0.42 | -0.20 | 0.12 | -1.16 |
| Maximum | 0.80 | 0.72 | 0.63 | 0.52 | 1.05 | 0.21 |
| $R_{HML}$ loading | | | | | | |
| Average | -0.14 | -0.09 | -0.05 | 0.16 | -0.15 | 0.31 |
| Minimum | -0.90 | -0.53 | -0.55 | -0.20 | -0.55 | -0.25 |
| Maximum | 0.43 | 0.25 | 0.80 | 0.78 | 0.11 | 1.15 |
| $R_{UMD}$ loading | | | | | | |
| Average | -0.01 | 0.03 | -0.04 | -0.01 | 0.08 | -0.10 |
| Minimum | -0.47 | -0.24 | -0.37 | -0.17 | -0.28 | -0.60 |
| Maximum | 0.42 | 0.32 | 0.29 | 0.09 | 0.44 | 0.36 |

## Table III: Conditional benchmark models.

The table reports conditional regression results for quintile portfolios sorted on lagged $R^2$ (Panel A) and decile portfolios sorted on lagged volatility (Panel B) using the Carhart (1997) four-factor model. The return regression is given by $R_{i,t} = \alpha_i^C + (\lambda_{i,0} + \lambda'_{i,1} Z_{i,t-1}^{MKT}) R_{MKT,t} + (\gamma_{i,0} + \gamma'_{i,1} Z_{i,t-1}^{SMB}) R_{SMB,t} + (\eta_{i,0} + \eta'_{i,1} Z_{i,t-1}^{HML}) R_{HML,t} + (\nu_{i,0} + \nu'_{i,1} Z_{i,t-1}^{UMD}) R_{UMD,t} + \varepsilon_{i,t}$. The conditioning variables, $Z_{i,t-1}^k$, for a given portfolio include the 24-month and three-month lagged factor loadings. The estimates of $\alpha_i^C$ are reported in percentage per year, and the numbers in parentheses are Newey–West (1987) corrected $t$-statistics with a lag length equal to five. For each regression, $R^2$ is the adjusted $R^2$ value. For each conditional model (i.e., Cases 2 to 4), the table also reports a $p$-value ($p(\alpha_{LH}^C \geq \alpha_{LH}^U)$) for the one-sided test that the conditional low-minus-high alpha is greater than or equal to the corresponding unconditional alpha from Case 1. The sample period for Panel A (Panel B) is January 1999 to December 2014 (January 2000 to December 2014).

| Panel A: Quintile portfolios sorted on $R^2$ | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | $R_{MKT,t}\times$ | | | $R_{SMB,t}\times$ | | | $R_{HML,t}\times$ | | | $R_{UMD,t}\times$ | | | |
| Case | $R^2_{t-1}$ | $\alpha_i^C$ | $p(\alpha_{LH}^C \geq \alpha_{LH}^U)$ | 1 | $\beta^{L24}$ | $\beta^{L3}$ | 1 | $s^{L24}$ | $s^{L3}$ | 1 | $h^{L24}$ | $h^{L3}$ | 1 | $u^{L24}$ | $u^{L3}$ | $R^2$ |
| 1 | L | 3.85 | | 0.98 | | | 0.40 | | | -0.04 | | | 0.10 | | | 91.9 |
| | | (2.0) | | (27) | | | (7.4) | | | (-0.5) | | | (2.1) | | | |
| | H | -1.80 | | 1.06 | | | 0.04 | | | -0.10 | | | -0.02 | | | 95.7 |
| | | (-2.4) | | (33) | | | (1.0) | | | (-2.0) | | | (-0.7) | | | |
| | L−H | 5.65 | n/a | | | | | | | | | | | | | |
| | | (2.6) | | | | | | | | | | | | | | |
| 2 | L | 2.55 | | 0.22 | 0.81 | | 0.33 | 0.15 | | 0.01 | 1.07 | | 0.02 | 0.40 | | 95.0 |
| | | (2.0) | | (1.5) | (5.0) | | (3.6) | (0.6) | | (0.4) | (9.1) | | (0.6) | (1.4) | | |
| | H | -0.61 | | 0.16 | 0.86 | | -0.02 | 0.94 | | 0.01 | 1.02 | | 0.02 | 0.44 | | 98.0 |
| | | (-1.3) | | (0.4) | (2.1) | | (-0.8) | (9.2) | | (0.4) | (4.7) | | (1.1) | (2.0) | | |
| | L−H | 3.16 | 0.036 | | | | | | | | | | | | | |
| | | (2.5) | | | | | | | | | | | | | | |
| 3 | L | 1.14 | | -0.17 | | 1.24 | 0.12 | | 0.56 | -0.04 | | 0.87 | 0.03 | | 1.08 | 96.4 |
| | | (0.9) | | (-1.0) | | (6.6) | (2.3) | | (5.9) | (-1.2) | | (6.6) | (2.5) | | (8.2) | |
| | H | -0.82 | | -0.45 | | 1.47 | 0.03 | | 0.73 | -0.02 | | 1.27 | 0.01 | | 1.07 | 98.8 |
| | | (-2.0) | | (-1.4) | | (4.7) | (1.5) | | (8.9) | (-0.6) | | (4.9) | (1.3) | | (7.3) | |
| | L−H | 1.96 | 0.010 | | | | | | | | | | | | | |
| | | (1.8) | | | | | | | | | | | | | | |
| 4 | L | 0.89 | | 0.06 | -0.06 | 1.06 | 0.25 | -0.43 | 0.60 | 0.01 | 0.45 | 0.45 | 0.02 | -0.04 | 0.96 | 96.6 |
| | | (0.7) | | (0.2) | (-0.4) | (4.9) | (3.1) | (-1.3) | (3.4) | (0.2) | (4.2) | (2.8) | (1.3) | (-0.3) | (11) | |
| | H | -0.69 | | -0.27 | -0.28 | 1.57 | 0.02 | 0.62 | 0.22 | -0.01 | 0.04 | 1.26 | 0.01 | -0.19 | 1.16 | 98.8 |
| | | (-1.6) | | (-0.8) | (-1.3) | (3.3) | (0.8) | (2.8) | (1.1) | (-0.6) | (0.2) | (5.7) | (1.3) | (-1.0) | (7.3) | |
| | L−H | 1.58 | 0.007 | | | | | | | | | | | | | |
| | | (1.5) | | | | | | | | | | | | | | |

(*continued*)

**Table III**—*Continued*

Panel B: Decile portfolios sorted on volatility

| | | | | $R_{MKT,t}\times$ | | | $R_{SMB,t}\times$ | | | $R_{HML,t}\times$ | | | $R_{UMD,t}\times$ | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Case | $\sigma^2_{t-1}$ | $\alpha^C_i$ | $p(\alpha^C_{LH} \geq \alpha^U_{LH})$ | 1 | $\beta^{L24}$ | $\beta^{L3}$ | 1 | $s^{L24}$ | $s^{L3}$ | 1 | $h^{L24}$ | $h^{L3}$ | 1 | $u^{L24}$ | $u^{L3}$ | $R^2$ |
| 1 | L | 1.53 | | 0.75 | | | 0.03 | | | 0.24 | | | 0.01 | | | 94.2 |
| | | (1.7) | | (28) | | | (1.1) | | | (4.8) | | | (1.0) | | | |
| | H | -4.11 | | 1.28 | | | 0.62 | | | -0.31 | | | 0.05 | | | 94.7 |
| | | (-3.0) | | (21) | | | (8.8) | | | (-5.5) | | | (0.9) | | | |
| | L−H | 5.64 | n/a | | | | | | | | | | | | | |
| | | (3.2) | | | | | | | | | | | | | | |
| 2 | L | -0.16 | | 0.33 | 0.58 | | 0.04 | 0.46 | | 0.02 | 0.83 | | -0.01 | 0.24 | | 97.3 |
| | | (-0.3) | | (1.5) | (2.1) | | (1.3) | (3.9) | | (0.7) | (6.4) | | (-1.5) | (3.1) | | |
| | H | -2.86 | | 0.65 | 0.47 | | 0.25 | 0.50 | | -0.04 | 1.12 | | 0.00 | 0.62 | | 96.0 |
| | | (-2.5) | | (2.2) | (1.8) | | (2.0) | (2.8) | | (-0.6) | (3.3) | | (0.1) | (4.2) | | |
| | L−H | 2.70 | 0.016 | | | | | | | | | | | | | |
| | | (2.6) | | | | | | | | | | | | | | |
| 3 | L | -0.45 | | 0.49 | | 0.45 | -0.01 | | 0.81 | 0.02 | | 1.03 | 0.01 | | 0.72 | 97.8 |
| | | (-0.7) | | (2.5) | | (1.9) | (-0.5) | | (7.8) | (0.9) | | (14) | (0.7) | | (5.8) | |
| | H | -1.98 | | -0.32 | | 1.35 | 0.04 | | 0.76 | -0.02 | | 0.85 | 0.03 | | 0.83 | 97.7 |
| | | (-2.0) | | (-0.8) | | (3.9) | (0.4) | | (4.3) | (-0.6) | | (4.1) | (1.3) | | (11) | |
| | L−H | 1.53 | 0.003 | | | | | | | | | | | | | |
| | | (1.7) | | | | | | | | | | | | | | |
| 4 | L | -0.48 | | 0.30 | 0.51 | 0.16 | -0.02 | -0.14 | 0.99 | 0.00 | 0.17 | 0.91 | 0.00 | 0.13 | 0.61 | 97.8 |
| | | (-0.9) | | (1.6) | (1.8) | (0.6) | (-0.8) | (-0.5) | (2.8) | (0.1) | (0.7) | (3.5) | (0.6) | (1.9) | (5.2) | |
| | H | -1.87 | | -0.25 | -0.12 | 1.40 | 0.01 | 0.26 | 0.60 | -0.04 | -0.14 | 0.98 | 0.03 | -0.09 | 0.91 | 97.7 |
| | | (-2.0) | | (-0.7) | (-0.9) | (4.0) | (0.1) | (1.3) | (2.4) | (-0.8) | (-0.5) | (4.9) | (1.5) | (-0.5) | (7.5) | |
| | L−H | 1.39 | 0.003 | | | | | | | | | | | | | |
| | | (1.7) | | | | | | | | | | | | | | |

**Table IV: Comparison of lagged factor loadings and traditional instruments in conditional benchmark models.**

The table reports conditional regression results for quintile portfolios sorted on lagged $R^2$ (Panel A) and decile portfolios sorted on lagged volatility (Panel B) using the Carhart (1997) four-factor model. The return regression is given by $R_{i,t} = \alpha_i^C + (\lambda_{i,0} + \lambda'_{i,1} Z_{i,t-1}^{MKT}) R_{MKT,t} + (\gamma_{i,0} + \gamma'_{i,1} Z_{i,t-1}^{SMB}) R_{SMB,t} + (\eta_{i,0} + \eta'_{i,1} Z_{i,t-1}^{HML}) R_{HML,t} + (\nu_{i,0} + \nu'_{i,1} Z_{i,t-1}^{UMD}) R_{UMD,t} + \varepsilon_{i,t}$. The conditioning variables, $Z_{i,t-1}^k$, for a given portfolio include "traditional instruments" (i.e., the dividend yield, default spread, and term spread) and the 24-month and three-month lagged factor loadings. In each panel, we present results for unconditional models with no instruments for the factor loadings, conditional models with lagged beta instruments for each factor, conditional models with traditional instruments for the market factor, and conditional models with traditional instruments for each factor. The specifications labeled "All" incorporate all three traditional instruments for the indicated factor(s). The estimates of $\alpha_i^C$ are reported in percentage per year, and $t(\alpha_{LH}^C)$ is the Newey–West (1987) corrected $t$-statistic for the low-minus-high alpha. $R_L^2$ and $R_H^2$ are the adjusted $R^2$ values for the low and high portfolios, respectively. For each conditional model, the table reports a $p$-value ($p(\alpha_{LH}^C \geq \alpha_{LH}^U)$) for the one-sided test that the conditional low-minus-high alpha is greater than or equal to the corresponding unconditional alpha. The sample period for Panel A (Panel B) is January 1999 to December 2014 (January 2000 to December 2014).

| | Alpha estimates | | | | | Model fit | |
|---|---|---|---|---|---|---|---|
| Traditional instrument | $\alpha_L^C$ | $\alpha_H^C$ | $\alpha_{LH}^C$ | $t(\alpha_{LH}^C)$ | $p(\alpha_{LH}^C \geq \alpha_{LH}^U)$ | $R_L^2$ | $R_H^2$ |
| Panel A: Quintile portfolios sorted on $R^2$ | | | | | | | |
| Unconditional model | | | | | | | |
| None | 3.85 | -1.80 | 5.65 | 2.61 | n/a | 91.9 | 95.7 |
| Lagged beta instruments | | | | | | | |
| None | 0.89 | -0.69 | 1.58 | 1.48 | 0.007 | 96.6 | 98.8 |
| Traditional instruments for market factor loading | | | | | | | |
| Dividend yield | 3.76 | -1.71 | 5.46 | 2.64 | 0.350 | 92.0 | 95.8 |
| Default spread | 4.17 | -2.05 | 6.22 | 2.98 | 0.852 | 92.1 | 95.8 |
| Term spread | 3.86 | -1.73 | 5.58 | 2.59 | 0.377 | 91.9 | 95.9 |
| All | 4.15 | -1.87 | 6.01 | 2.71 | 0.735 | 92.1 | 96.0 |
| Traditional instruments for all factor loadings | | | | | | | |
| Dividend yield | 2.65 | -1.23 | 3.88 | 1.84 | 0.022 | 93.0 | 96.1 |
| Default spread | 3.18 | -1.87 | 5.05 | 2.41 | 0.268 | 93.0 | 95.9 |
| Term spread | 3.81 | -1.89 | 5.70 | 2.52 | 0.534 | 92.2 | 96.0 |
| All | 2.84 | -1.31 | 4.15 | 1.93 | 0.085 | 93.4 | 96.3 |
| Panel B: Decile portfolios sorted on volatility | | | | | | | |
| Unconditional model | | | | | | | |
| None | 1.53 | -4.11 | 5.64 | 3.19 | n/a | 94.2 | 94.7 |
| Lagged beta instruments | | | | | | | |
| None | -0.48 | -1.87 | 1.39 | 1.71 | 0.003 | 97.8 | 97.7 |
| Traditional instruments for market factor loading | | | | | | | |
| Dividend yield | 1.53 | -4.05 | 5.58 | 3.09 | 0.397 | 94.2 | 94.7 |
| Default spread | 1.40 | -3.97 | 5.37 | 3.07 | 0.215 | 94.2 | 94.7 |
| Term spread | 1.49 | -4.00 | 5.49 | 3.08 | 0.357 | 94.3 | 94.8 |
| All | 1.03 | -2.66 | 3.69 | 2.01 | 0.007 | 94.4 | 95.2 |
| Traditional instruments for all factor loadings | | | | | | | |
| Dividend yield | -0.09 | -3.18 | 3.10 | 2.39 | 0.005 | 96.7 | 95.8 |
| Default spread | 0.93 | -4.35 | 5.28 | 3.22 | 0.282 | 95.3 | 95.4 |
| Term spread | 0.95 | -3.69 | 4.65 | 2.90 | 0.147 | 95.4 | 95.6 |
| All | -0.37 | -2.17 | 1.80 | 1.50 | 0.003 | 97.2 | 96.6 |

**Table V: Conditional benchmark models: Full sample results for portfolios sorted on lagged $R^2$.**

The table reports conditional regression results for quintile portfolios sorted on lagged $R^2$ using the Carhart (1997) four-factor model. The return regression is given by $R_{i,t} = \alpha_i^C + (\lambda_{i,0} + \lambda'_{i,1} Z_{i,t-1}^{MKT}) R_{MKT,t} + (\gamma_{i,0} + \gamma'_{i,1} Z_{i,t-1}^{SMB}) R_{SMB,t} + (\eta_{i,0} + \eta'_{i,1} Z_{i,t-1}^{HML}) R_{HML,t} + (\nu_{i,0} + \nu'_{i,1} Z_{i,t-1}^{UMD}) R_{UMD,t} + \varepsilon_{i,t}$. In estimating these regressions, we define an indicator variable ($1_I$) equal to one for the period 1990 to 1998 and zero otherwise, and an indicator variable ($1_{II}$) equal to one for the period 1999 to 2014 and zero otherwise. The conditioning variables, $Z_{i,t-1}^k$, for a given portfolio include the interactions between $1_I$ and 24-month lagged factor loadings (e.g., $\beta_I^{L24}$), $1_{II}$, the interactions between $1_{II}$ and 24-month lagged factor loadings (e.g., $\beta_{II}^{L24}$), and the interactions between $1_{II}$ and three-month lagged factor loadings (e.g., $\beta_{II}^{L3}$). Panel A reports estimates of alpha and adjusted $R^2$ for an unconditional model with no instruments for the factor loadings and for the conditional model. The estimates of $\alpha_i^U$ and $\alpha_i^C$ are reported in percentage per year, and the numbers in parentheses are Newey–West (1987) corrected $t$-statistics with a lag length equal to five. For the conditional model, the table also reports a $p$-value ($p(\alpha_{LH}^C \geq \alpha_{LH}^U)$) for the one-sided test that the conditional low-minus-high alpha is greater than or equal to the corresponding unconditional alpha. Panel B presents the remaining parameter estimates for the conditional model. The sample period is January 1990 to December 2014.

Panel A: Unconditional and conditional alphas for quintile portfolios sorted on $R^2$

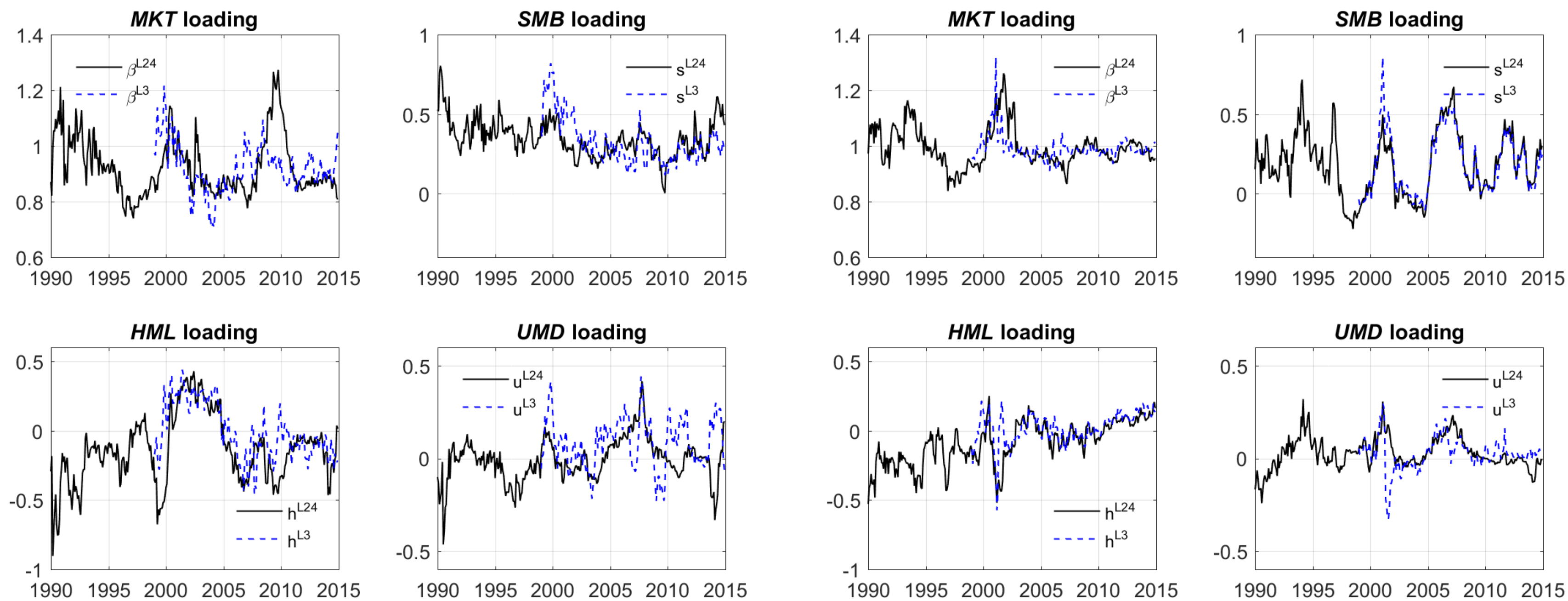
| | Unconditional model | | | | Conditional model | | |
|---|---|---|---|---|---|---|---|
| | L | H | L−H | | L | H | L−H |
| $\alpha_i^U$ | 2.66 | -1.06 | 3.72 | $\alpha_i^C$ | 0.58 | -0.69 | 1.27 |
| | (1.98) | (-1.54) | (2.38) | | (0.67) | (-1.83) | (1.55) |
| $R^2$ | 92.9 | 95.4 | | $R^2$ | 96.7 | 98.6 | |
| | | | | | $p(\alpha_{LH}^C \geq \alpha_{LH}^U) = 0.022$ | | |

Panel B: Factor loadings for conditional model

| | L | H | | L | H | | L | H | | L | H |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $R_{MKT,t}\times$ | | | $R_{SMB,t}\times$ | | | $R_{HML,t}\times$ | | | $R_{UMD,t}\times$ | | |
| 1 | 0.24 | 0.16 | 1 | 0.36 | 0.05 | 1 | 0.03 | -0.03 | 1 | 0.12 | 0.09 |
| | (1.3) | (0.9) | | (2.3) | (2.5) | | (0.6) | (-0.5) | | (3.1) | (3.9) |
| $\beta_I^{L24}$ | 0.83 | 0.84 | $s_I^{L24}$ | 0.16 | 0.78 | $h_I^{L24}$ | 0.64 | 0.65 | $u_I^{L24}$ | 0.22 | 0.51 |
| | (4.0) | (4.9) | | (0.4) | (7.2) | | (3.7) | (3.0) | | (0.8) | (1.7) |
| $1_{II}$ | -0.18 | -0.43 | $1_{II}$ | -0.11 | -0.03 | $1_{II}$ | -0.02 | 0.02 | $1_{II}$ | -0.11 | -0.08 |
| | (-0.6) | (-1.1) | | (-0.6) | (-1.3) | | (-0.4) | (0.3) | | (-2.6) | (-3.2) |
| $\beta_{II}^{L24}$ | -0.07 | -0.28 | $s_{II}^{L24}$ | -0.45 | 0.62 | $h_{II}^{L24}$ | 0.45 | 0.04 | $u_{II}^{L24}$ | -0.04 | -0.19 |
| | (-0.4) | (-1.3) | | (-1.3) | (2.8) | | (4.1) | (0.2) | | (-0.3) | (-1.0) |
| $\beta_{II}^{L3}$ | 1.06 | 1.57 | $s_{II}^{L3}$ | 0.61 | 0.22 | $h_{II}^{L3}$ | 0.46 | 1.26 | $u_{II}^{L3}$ | 0.97 | 1.16 |
| | (5.0) | (3.3) | | (3.4) | (1.1) | | (2.8) | (5.7) | | (10.9) | (7.3) |

**Table VI: Unconditional alpha decompositions.**
The table provides decompositions of unconditional Carhart (1997) four-factor alpha estimates into security selection effects, direct factor timing effects, and volatility timing effects for quintile portfolios sorted on lagged $R^2$ and decile portfolios sorted on lagged volatility. The conditional alphas and factor loadings for the $R^2$ portfolios (volatility portfolios) correspond to Case 4 in Panel A (Panel B) of Table III. We present results for the low ("L"), high ("H"), and low-minus-high ("L−H") portfolios for each sorting variable. Each unconditional alpha ($\alpha_i^U$) is decomposed into a security selection component ($\alpha_i^C$), four factor timing components, and four volatility timing components. For the market factor ($R_{MKT,t}$), the factor timing component is estimated as $\text{cov}(\hat{\beta}_{i,t}^C, R_{MKT,t})$, where $\hat{\beta}_{i,t}^C$ is the conditional market loading for a given portfolio, $\hat{\lambda}_{i,0} + \hat{\lambda}'_{i,1} Z_{i,t-1}^{MKT}$. The volatility timing effect for the market factor is $(\bar{\beta}_{i,t}^C - \hat{\beta}_i^U)\bar{R}_{MKT,t}$, where $\bar{\beta}_{i,t}^C$ is the average conditional loading for the market factor, $\hat{\beta}_i^U$ is the unconditional loading on the market factor (i.e., Case 1 in Table III), and $\bar{R}_{MKT,t}$ is the average return on the market factor. The factor timing and volatility timing effects for the size factor ($R_{SMB,t}$), the value factor ($R_{HML,t}$), and the momentum factor ($R_{UMD,t}$) are estimated analogously. All figures are reported in percentage per year. The sample period for the $R^2$ portfolios (volatility portfolios) is January 1999 to December 2014 (January 2000 to December 2014).

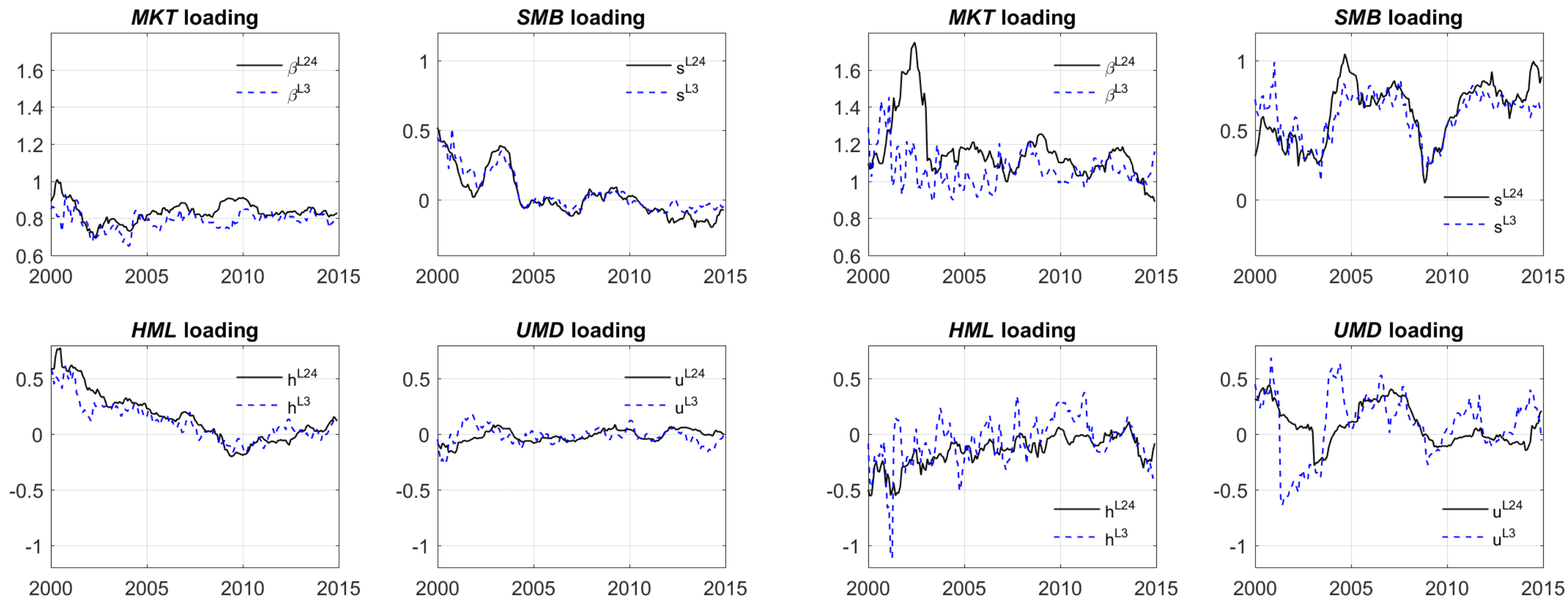
| | $R^2$ portfolios | | | Volatility portfolios | | |
|---|---|---|---|---|---|---|
| | 1999-2014 | | | 2000-2014 | | |
| | L | H | L−H | L | H | L−H |
| Security selection ability | | | | | | |
| $(a)$ Conditional alpha, $\alpha_i^C$ | 0.89 | -0.69 | 1.58 | -0.48 | -1.87 | 1.39 |
| Factor timing ability | | | | | | |
| $\text{cov}(\beta_{i,t}^C, R_{MKT,t})$ | -0.32 | -0.52 | 0.20 | 0.04 | -1.93 | 1.97 |
| $\text{cov}(s_{i,t}^C, R_{SMB,t})$ | 0.71 | -0.57 | 1.28 | 0.43 | -0.50 | 0.93 |
| $\text{cov}(h_{i,t}^C, R_{HML,t})$ | 1.24 | -0.74 | 1.98 | 1.61 | -0.36 | 1.97 |
| $\text{cov}(u_{i,t}^C, R_{UMD,t})$ | 1.83 | -0.18 | 2.01 | 0.14 | 0.27 | -0.13 |
| $(b)$ Total factor timing | 3.45 | -2.01 | 5.46 | 2.23 | -2.53 | 4.75 |
| Volatility timing ability | | | | | | |
| $(\bar{\beta}_{i,t}^C - \beta_i^U)\bar{R}_{MKT,t}$ | -0.03 | -0.27 | 0.24 | 0.40 | -0.73 | 1.13 |
| $(\bar{s}_{i,t}^C - s_i^U)\bar{R}_{SMB,t}$ | -0.47 | 0.60 | -1.07 | 0.01 | -0.35 | 0.36 |
| $(\bar{h}_{i,t}^C - h_i^U)\bar{R}_{HML,t}$ | 0.00 | 0.38 | -0.38 | -0.57 | 1.26 | -1.83 |
| $(\bar{u}_{i,t}^C - u_i^U)\bar{R}_{UMD,t}$ | 0.01 | 0.19 | -0.19 | -0.05 | 0.11 | -0.16 |
| $(c)$ Total volatility timing | -0.49 | 0.90 | -1.39 | -0.21 | 0.29 | -0.50 |
| Unconditional alpha | | | | | | |
| Unconditional alpha, $\alpha_i^U = (a) + (b) + (c)$ | 3.85 | -1.80 | 5.65 | 1.53 | -4.11 | 5.64 |




Panel A: Low-$R^2$ portfolio

Panel B: High-$R^2$ portfolio

**Figure 1: Lagged factor loadings for low-$R^2$ and high-$R^2$ mutual funds.**
The figure shows lagged factor loadings from Carhart (1997) model regressions for the low-quintile (Panel A) and high-quintile (Panel B) portfolios sorted on lagged $R^2$. The strategies are based on a double sort on lagged Carhart (1997) four-factor regression $R^2$ and alpha as described in the text. The low-$R^2$ and high-$R^2$ portfolios are those conditioned on having a high prior alpha. The solid (dashed) lines are lagged loadings estimated from 24 months (3 months) of monthly (daily) data. The lagged loadings for a given portfolio are the equal-weight lagged loadings across constituent mutual funds. The sample period is January 1990 to December 2014.




Panel A: Low-volatility portfolio

Panel B: High-volatility portfolio

**Figure 2: Lagged factor loadings for low-volatility and high-volatility mutual funds.**
The figure shows lagged factor loadings from Carhart (1997) model regressions for the low-decile (Panel A) and high-decile (Panel B) portfolios sorted on lagged volatility. The solid (dashed) lines are lagged loadings estimated from 24 months (3 months) of monthly (daily) data. The lagged loadings for a given portfolio are the equal-weight lagged loadings across constituent mutual funds. The sample period is January 2000 to December 2014.

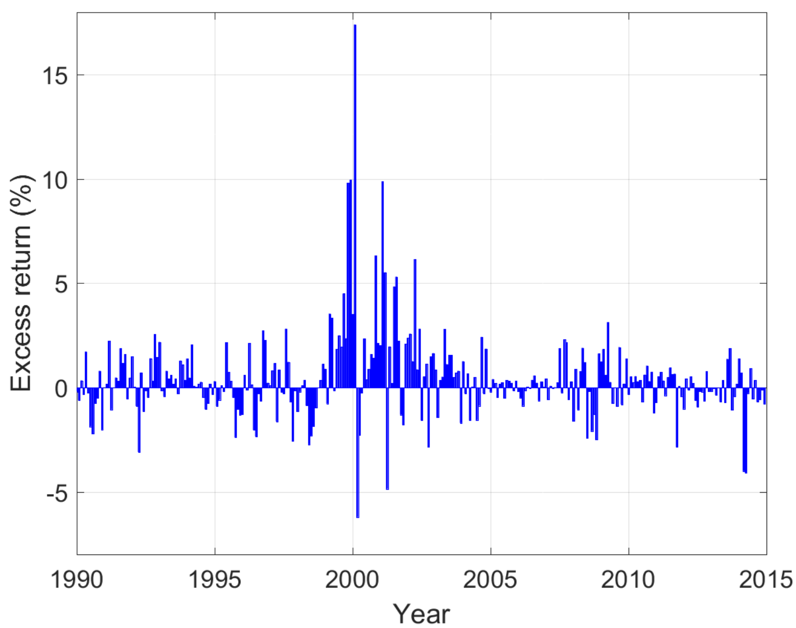



**Figure 3: Difference in returns for low-$R^2$ and high-$R^2$ mutual funds.**
The figure shows differences in net returns in percentage per month for the low-quintile and high-quintile portfolios sorted on lagged $R^2$. The strategies are based on a double sort on lagged Carhart (1997) four-factor regression $R^2$ and alpha as described in the text. The low-$R^2$ and high-$R^2$ portfolios used to estimate the plotted return series are those conditioned on having a high prior alpha. The sample period is January 1990 to December 2014.

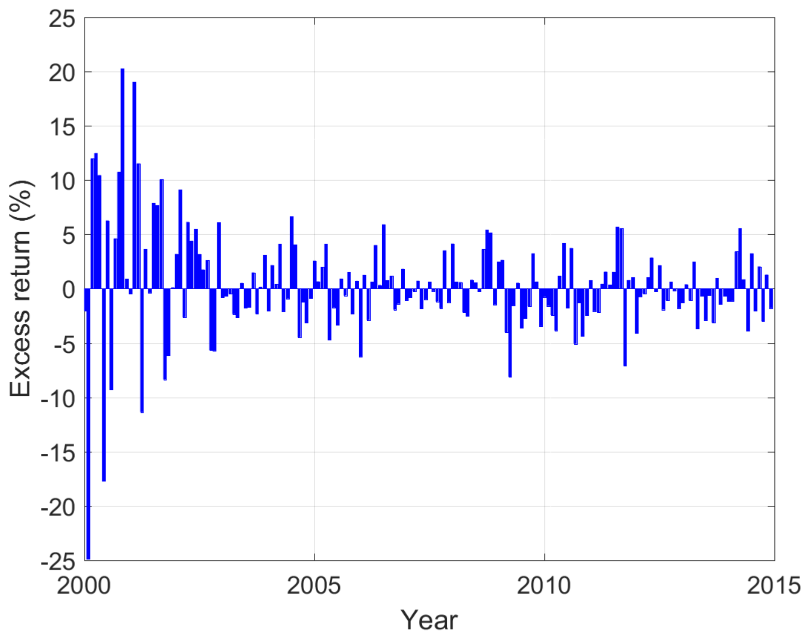
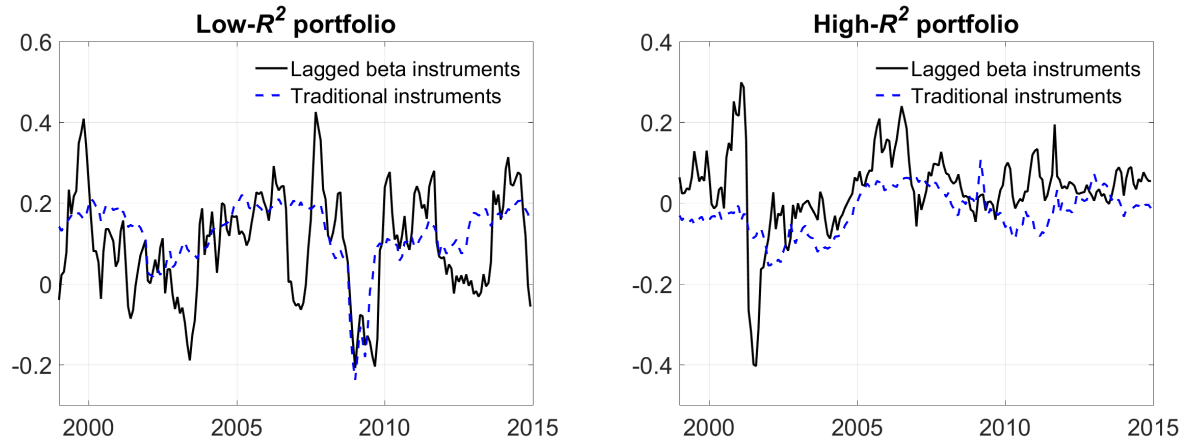
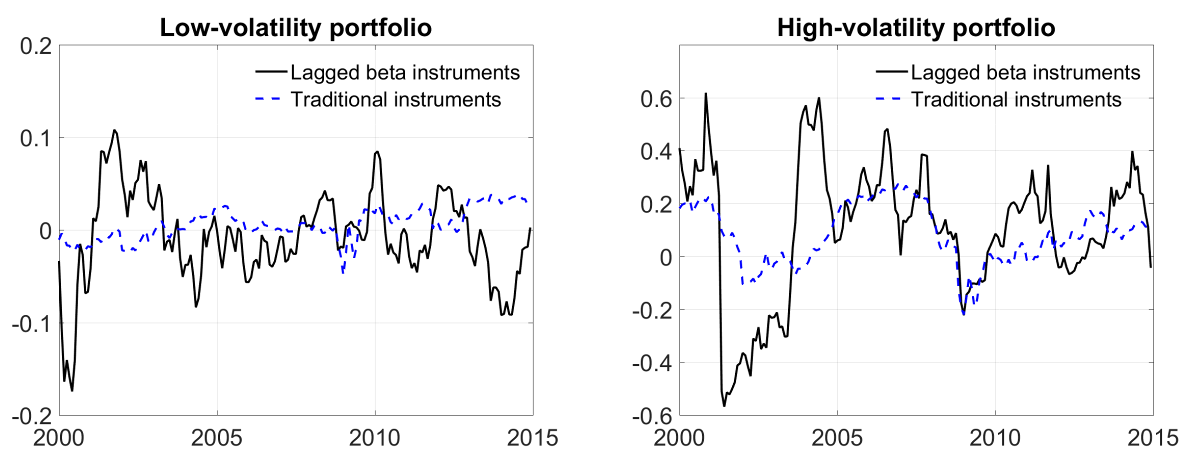



**Figure 4: Difference in returns for low-volatility and high-volatility mutual funds.** The figure shows differences in net returns in percentage per month for the low-decile and high-decile portfolios sorted on lagged volatility. The formation-period volatility estimates are based on 12 months of prior daily excess returns. The sample period is January 2000 to December 2014.




Panel A: $R^2$ portfolios




Panel B: Volatility portfolios

**Figure 5: Comparison of conditional momentum loadings.**
The figure shows conditional momentum loadings, $\hat{u}^{C}_{i,t} \equiv \hat{\nu}_{i,0} + \hat{\nu}'_{i,1} Z^{UMD}_{i,t-1}$, for the low-quintile and high-quintile portfolios sorted on lagged $R^2$ (Panel A) and the low-decile and high-decile portfolios sorted on lagged volatility (Panel B). Each plot presents results for two conditional Carhart (1997) regression models. The solid line corresponds to a conditional model in which factor loadings are modeled as a linear function of the corresponding 24-month and three-month lagged factor loadings. The dashed line corresponds to a conditional model in which factor loadings are modeled as a linear function of the dividend yield, default spread, and term spread. The sample period for Panel A (Panel B) is January 1999 to December 2014 (January 2000 to December 2014).

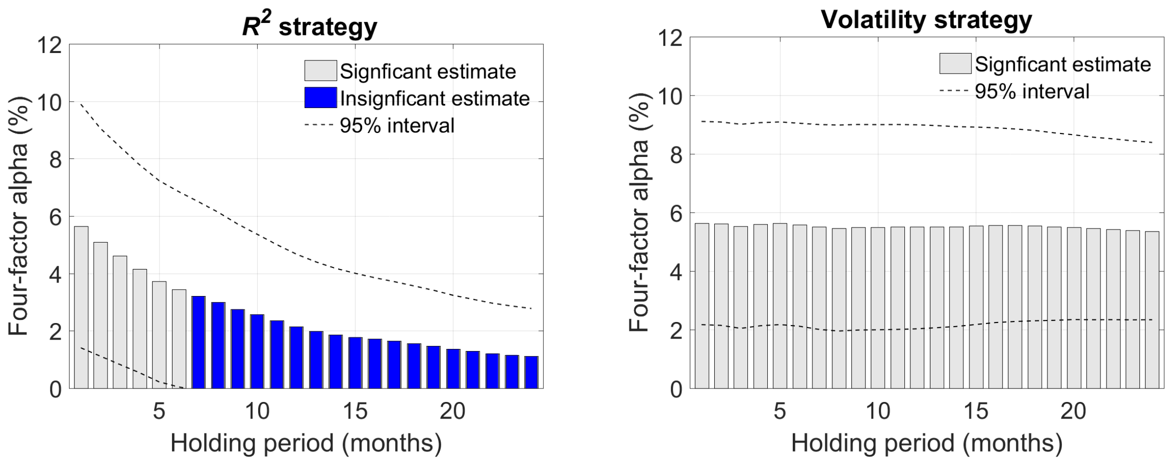



**Figure 6: The impact of holding period on unconditional alpha.**
The figure shows unconditional Carhart (1997) four-factor alphas for the low-minus-high portfolio sorted on lagged $R^2$ (left) and the low-minus-high portfolio sorted on lagged volatility (right) as a function of holding period. Results are provided for portfolio holding periods ranging from one to 24 months. For a holding period of $n$ months, $1/n$ of the portfolio is replaced at the beginning of each month. The $R^2$ strategies are based on a double sort on lagged Carhart (1997) four-factor regression $R^2$ and alpha as described in the text. The low-$R^2$ and high-$R^2$ portfolios used to estimate the low-minus-high alphas are those conditioned on having a high prior alpha. For the volatility-sorted portfolios, the formation-period volatility estimates are based on 12 months of prior monthly (daily) excess returns in the pre-January 2000 (post-January 2000) period. The alpha estimates are reported in percentage per year, and the dashed lines indicate a 95% confidence interval based on Newey–West (1987) corrected $t$-statistics with a lag length equal to five. The sample period for the $R^2$ strategy (volatility strategy) is January 1999 to December 2014 (January 2000 to December 2014).

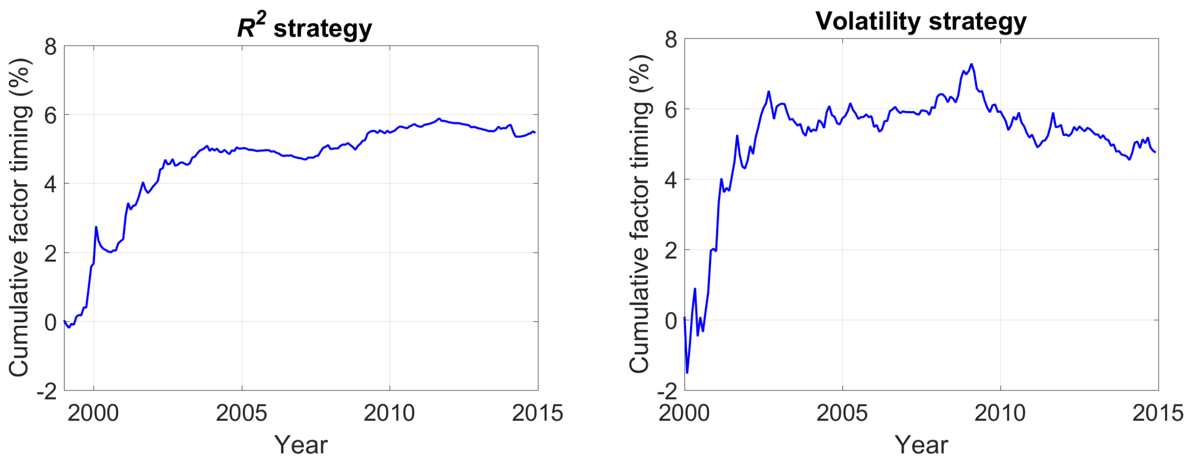



**Figure 7: Cumulative factor timing ability.**
The figure shows the cumulative factor timing ability in percentage per year for the low-minus-high portfolio sorted on lagged $R^2$ (left) and the low-minus-high portfolio sorted on lagged volatility (right). For a given sample month $\tau$, cumulative timing ability for the market factor is given by $\frac{1}{T}\sum_{t=1}^{\tau}\hat{\beta}^{C}_{LH,t}R_{MKT,t} - \frac{\tau}{T}\left(\frac{1}{T}\sum_{t=1}^{T}\hat{\beta}^{C}_{LH,t}\right)\left(\frac{1}{T}\sum_{t=1}^{T}R_{MKT,t}\right)$, where $\hat{\beta}^{C}_{LH,t}$ is the conditional beta and $R_{MKT,t}$ is the return on the market factor in month $t$. Timing abilities for the size, value, and momentum factors are defined analogously, and cumulative factor timing ability is the sum of cumulative timing ability across the four factors. The conditional factor loadings for the $R^2$ (volatility) portfolios correspond to Case 4 in Panel A (Panel B) of Table III. The sample period for the $R^2$ strategy (volatility strategy) is January 1999 to December 2014 (January 2000 to December 2014).