# ESSAYS ON JUMP RISK FACTORS IN FINANCIAL FORECASTING

by

**BO YU**

**A dissertation submitted to the**

**School of Graduate Studies**

**Rutgers, The State University of New Jersey**

**In partial fulfillment of the requirements**

**For the degree of**

**Doctor of Philosophy**

**Graduate Program in Economics**

**Written under the direction of**

**Norman R. Swanson**

**And approved by**

______________________________

______________________________

______________________________

______________________________

**New Brunswick, New Jersey**

**May, 2020**

## ABSTRACT OF THE DISSERTATION

**Essays on Jump Risk Factors in Financial Forecasting**

**By BO YU**

**Dissertation Director:**

**Norman R. Swanson**

This dissertation consists of two essays that explore issues in empirical asset pricing and portfolio management using high-frequency financial econometrics techniques. The first essay investigates the cross-sectional return predictability of various jump risk factors. The second essay develops sparse portfolio variance forecast models that incorporate informative realized jump risk factors.

In Chapter 2, we study the cross-sectional relationship between (small and large) jump variation measures and future stock returns, based on portfolio sorts and Fama-MacBeth type regressions. We document that a new risk factor, signed small jump variation (i.e., the difference between upside and downside small jump variation measures), strongly predicts the cross-sectional variation in future returns. Constructed based on a data-driven threshold, signed small jump variation has stronger predictive power for future returns than other realized risk measures, in the cross-section. We also conduct various experiments (e.g., event studies, etc.) to further explore the linkages between different jump risk measures and economic factors relating to news in the markets. We show that large jumps are closely associated with "big" news. While such news related information is embedded in large jump variation, the information is generally short-lived, and dissipates too quickly to provide marginal predictive content

for subsequent weekly returns. By contrast, we find that small jumps are more likely to be diversified away than large jumps, thus tend to be more closely associated with idiosyncratic risks, and are therefore more likely to be driven by liquidity conditions and trading activity.

In Chapter 3, we investigate whether the decomposition of realized covariance matrices of portfolios of asset returns into components based on both the signs and magnitudes of the underlying high-frequency returns is useful for forecasting. In particular, our decomposition separates realized covariation into components based on signs (positive and negative) and magnitudes (continuous, small jump, and large jump). Sparse portfolio variance forecast models, which are constructed by utilizing the most informative covariance components, produce significant improvements in predictive accuracy. We show that such predictive gains can be traced to the identification of short-lived pricing signals associated with co-jumps.

# Acknowledgements

I am extremely indebted to my advisor, Professor Norman R. Swanson, for his generous support, invaluable advice, and continuous encouragement throughout my prolonged Ph.D. studies. He guided me into the research fields that I am passionate about, gave me the freedom to think and work independently, encouraged and helped me build up my confidence to develop my work. His enthusiasm for research is contagious and motivational for me, even during the tough times in this journey.

I would like to express my deep gratitude to the rest of my committee members, Professor Bruce Mizrach and Professor Xiye Yang, for their insightful comments and encouraging advice. I feel so fortunate to have had the opportunity to work with them. I am also grateful to Professor Joseph P. Hughes, Professor Roger Klein, Professor John Landon-Lane, Professor Yuan Liao, Professor Daijiro Okada, and Linda Zullinger, for their help and advice over the past few years.

I am greatly thankful to the computing support provided by the Office of Advanced Research Computing (OARC) at Rutgers University. I benefited tremendously from coversations with Doctor Kristina Plazonic and other OARC team members, who helped me overcome many computing challenges that arise when preparing for Chapter 3 of this dissertation.

Last but not least, I would like to thank my parents, for their unconditional love, support, and faith in me.

# Dedication

To my beloved parents.

# Table of Contents

# List of Tables

# List of Figures

# Chapter 1

# Introduction

Volatility or (co)realized variance has been documented as one of the most informative stock-level predictors for future returns or variance (see e.g.,Gu et al. (2019) and Corsi (2009)). With the availability of high-frequency financial data and advances in high-frequency econometrics, different components of the variance, associated with upside/downside and/or continuous/discontinuous price movements, can now be consistently measured. Recent studies have reached a consensus that the signs of the underlying high-frequency returns lead to distinct information content residing in different components of the realized risk measure (see e.g., Bollerslev et al. (2019a) and Bollerslev et al. (2019b)).

In this dissertation, I take a further step by considering both the signs and the magnitudes of the underlying high-frequency stock returns when constructing risk factors and explore the return/variance predictability of these separate risk components. Specifically, in the second chapter, we partition the semi-variances into small and large components. In particular, high frequency intraday data are used to construct various realized jump variation measures, including small upside/downside, large upside/downside, and the difference between upside small (large) and downside small (large) jump variation (i.e., signed small and large jump variation measures). We then investigate the relationship between these various risk measures and future returns, using single sorted and double sorted stock portfolios, and Fama-MacBeth regression analysis. In the third chapter, we decompose the realized covariance into components based on both the signs and magnitudes of the underlying high frequency returns, and construct sparse portfolio variance forecast models by utilizing the most informative components as predictors.

In Chapter 2, entitled *"New Evidence of the Marginal Predictive Content of Small and Large Jumps"* joint with Bruce Mizrach and Norman R. Swanson, we take the additional step of partitioning the semi-variances into small and large components, and explore the possibility that small and large jumps contain different information relative to investing and return predictability. We find that sorting on signed small jump variation (i.e., the difference between the upside and downside small jump variation measures) leads to greater value-weighted return differentials between stocks in our highest and lowest quintile portfolios (i.e., high-low spreads) than when either signed total jump or signed large jump variation is sorted on. Moreover, in a key case, the high-low spread is not significantly different from zero when signed large jump variation is sorted on. Indeed, including large jump variation can actually decrease predictive accuracy, in the sense that average returns and alphas for high-low portfolios are lower when total jump variation is utilized in our prediction experiments rather than small jump variation. These results suggest that there may be a threshold, beyond which "large" jump variation contains no marginal predictive ability, relative to that contained in small jump variation. Analysis of returns and alphas based on industry double-sorts indicates that the benefit of small signed jump variation investing is driven by stock selection within an industry, rather than industry bets. Investors prefer stocks with a high probability of large positive jump variation, but they also tend to overweight safer industries. Additionally, the fact that large and small (signed) jump variation have differing marginal predictive content is explained at least in part by our observation that in double-sorted portfolios, the content of signed large jump variation is negligible when controlling for either signed total jump variation or realized skewness. By contrast, signed small jump variation has unique information for predicting future returns, even when controlling for total jump variation or realized skewness. Further, we find that large jumps are closely associated with "big" news, as might be expected. In particular, large earning announcement surprises increase both the magnitude and occurrence of large jumps. While such news related information is embedded in large jump variation, the information is generally short-lived, and dissipates too quickly to provide marginal predictive content for subsequent weekly returns. Finally, we find that while large jump variation is closely associated with large earnings surprises ("big" news), small jumps tend to be more closely associated with idiosyncratic risks, and can be diversified away.

In Chapter 3, entitled *"Forecasting Portfolio Variance: A New Decomposition Approach,"* we investigate whether the decomposition of realized covariance matrices of portfolio of asset returns into components based on both the signs and magnitudes of the underlying high-frequency returns is useful for forecasting. In particular, our decomposition separates realized covariation into components based on signs (positive and negative) and magnitudes (continuous, small jump, and large jump). The impetus for this decomposition is that certain variation components may be useful for prediction, while others are not; and by including only "information-rich" components in realized (co)-variation forecasting models, predictive accuracy may be improved. Our experiments that are designed to assess the marginal predictive content of different variation components focus on portfolio variance prediction, and utilize various machine learning methods, including: penalized regression type methods such as the least absolute shrinkage operator and the elastic net, as well as dimension reduction methods such as partial least squares and principal components analysis. We find that machine learning methods with key variation components offer limited improvement to out-of-sample fit, relative to benchmark HAR-type forecasting models that do not include our more granular variation measures. However, more sparse models, which are specified using predictors selected using a first "dimension reduction" yield significant improvements in predictive accuracy when our decomposed variation measures are included. These predictive gains can be traced to the identification of short-lived pricing signals associated with co-jumps.

# Chapter 2

# New Evidence of the Marginal Predictive Content of Small and Large Jumps

## 2.1 Introduction

Theoretical models of the risk-return relationship anticipate that volatility should be priced, and that investors should demand higher expected returns for more volatile assets. However, ex-ante risk measures are not directly observable, and must be estimated (see e.g., Rossi and Timmermann (2015)). Given the necessity of estimating volatility, various different risk estimators have been utilized in the empirical literature studying the strength and sign of the risk-return relationship. Unfortunately, the evidence from the literature is mixed, in the sense that researchers have found both negative and positive relationships between return and volatility. One possible reason for these surprisingly contradictory findings is that the risk-return relationship is nonlinear. Examples of papers pursuing this hypothesis include Campbell and Vuolteenaho (2004), who incorporate different factor betas based on good and bad news about cash flows and discount rates; and Woodward and Anderson (2009) who find that bull and bear market betas differ substantially across most industries. This research has helped to spawn the "smart-beta" approach to factor investing.[1] In related research, Feunou et al. (2013) model the effects of volatility in positive and negative return states separately. They define so-called disappointment aversion preferences, and show that investors should demand a higher return for downside variability. These authors find empirical support for their model in the U.S. and several foreign markets using a bi-normal GARCH process to estimate volatility.

In this paper, we focus on the importance of jumps in volatility for understanding the risk-return relationship. We do this by assessing the marginal predictive content

[1] In 2017, Morningstar reported that this approach to investing has attracted over one trillion dollars in assets (see e.g., Jennifer Thompson, Financial Times, December 27, 2017).

of small versus large jump variation, when forecasting one-week ahead cross-sectional equity returns. We also examine earnings announcements as well as carry out various Fama-MacBeth type regressions in order to uncover the linkages between (small and large) jumps and news. Finally, we examine the importance of control variates, including skewness and other firm characteristics, when undertaking to disentangle the relative importance of small, large, positive, and negative jumps for the dynamic evolution of firm specific returns. Much of the empirical research that explores the importance of jumps in this context focuses on estimation of continuous and jump variation components using nonparameteric realized measures constructed with high frequency financial data. A key paper in this area is Bollerslev et al. (2019b), who examine the relationship between signed jumps and future stock returns in the cross-section. They document that signed jump variation, which captures the asymmetric impact of upside and downside jump risks, are good predictors of returns for small and illiquid stocks.[2] In the current paper, we add to this literature by decomposing jump variation into signed small and large components and evaluating the importance of these elements in a cross-section of stock returns. We utilize the cross-section of individual stocks because aggregate index returns may mask small jump effects on return predictability. Indeed, many studies document that aggregation may diversify away idiosyncratic small jumps in the cross-section (see e.g., Aït-Sahalia and Jacod (2012) and Duong and Swanson (2015)).

The motivation for our paper can be traced back to Yan (2011) and Jiang and Yao (2013), who show that large, infrequent jumps are priced in the cross-section of returns. Feunou et al. (2018) take the decomposition used by these authors one step further, and model jumps in the realized semi-variances of market returns. They construct a new measure of the variance risk premium, and find a strong positive premium for downside risk. Fang et al. (2017) find a similar result for Chinese market returns. In a related line of research, various authors study the information content in the upside

[2] In a related paper, Duong and Swanson (2015) construct both small and large jump measures based on some fixed truncation levels. They exploit the risk predictabilities of different jump measures using both index data and Dow 30 stocks and find that small jump variation has more volatility predictability than large jump variation.

and downside jump variation. For example, Guo et al. (2015) document that at the market level, a negative jump component in realized volatility predicts an increase in future equity premia. Bollerslev et al. (2015) identify both left and right jump tail risks under the risk-neutral measure. They find that the left jump tail risk is an appropriate proxy for market fear. Additionally, they find that including a variance risk premium together with jump tail risk measures as predictors significantly improves market return forecasts. Finally, they show that jump risk helps explain the high-low book-to-market and winners versus losers portfolio returns.

Building on the above literature, we decompose jump variation into four distinct components depending on both the direction (semi-variances) and magnitude (small and large) of the jumps.[3] Specifically, we decompose individual stock jump semi-variances into small and large components. High frequency intraday data are used to construct various realized jump variation measures, including large upside/downside, small upside/downside, and the difference between upside large (small) and downside large (small) jump variation. We then investigate the relationship between these various jump measures and future returns, using sorted and double-sorted stock portfolios, and using regression analysis. The reason that we decompose jump semi-variances into small and large components is that this decomposition allows us to explore the possibility that they contain different information relevant to investing and return predictability. As Maheu and McCurdy (2004) note, large jumps may reflect important individual stock and market news announcements. Smaller jumps (or continuous variation) may result from liquidity and strategic trading.

Our key findings can be summarized as follows. First, we find that both small and large upside (downside) jump variation negatively (positively) predict subsequent weekly returns. However, portfolios sorted using signed total jump variation are associated with increased average returns and risk adjusted alphas for high-low portfolios, relative to the cases where upside or downside jump variation is sorted on. This finding

[3]The methods that we implement to separate jump variation utilize recent advances in financial econometrics due to Andersen et al. (2003), Andersen et al. (2007), Jacod (2008), Mancini (2009), Barndorff-Nielsen et al. (2010), Todorov and Tauchen (2010), Aït-Sahalia and Jacod (2012), and Patton and Sheppard (2015).

is in accord with the findings of Bollerslev et al. (2019b).

Our second finding involves the case where jump variation is further decomposed into "small" and "large" components. In this case, sorting on signed small jump variation leads to value weighted high-low portfolios with greater average returns and alphas than when either signed total jump or signed large jump variation is sorted on. Indeed, when the truncation parameter used to differentiate small from large jumps is based on a 5 standard deviation cut-off, we find that average return spreads are 10% higher when signed small jump variation is sorted on rather than signed total jump variation. Moreover, these average return spreads are statistically significant in both cases. However, average return spreads are not significantly different from zero when signed large jump variation is sorted on. Indeed, including large jump variation is actually detrimental to predictive accuracy, as average returns and alphas for high-low portfolios actually decline when total variation is instead utilized in our prediction experiments. These results suggest that there may be a "jump-threshold", beyond which "large" jump variation contains no marginal predictive ability, relative to that contained in small jump variations.[4] In summary, we find that large jump variation has little to no marginal predictive content, beyond a certain threshold. Indeed, when said threshold is judiciously selected, one can actually improve predictive performance in our experiments, leading to increased high-low portfolio average returns and alphas, when sorting portfolios based on small jump variation rather than total jump variation.

Third, industry double-sorts indicate that the benefit of small signed jump variation investing is driven by stock selection within an industry, rather than industry bets. Investors prefer stocks with a high probability of large positive jump variation, but they also tend to overweight safer industries.

---

[4]When equal weighted portfolios are instead examined, sorting on total jump variation yields higher average returns and alphas than when sorting on small or large jump variation. However, deeper inspection of our tabulated results in this case reveals that average returns associated with large jump variation sorts are much smaller (around 1/2 the magnitude) of small and total jump variation sorts, and that the magnitude of average returns associated with small jump variation sorts is much closer (within 10%) to the average returns associated with total jump variation sorts when our truncation parameter uses a 5 standard deviation cut-off instead of a 4 standard-deviation cut-off. This suggests that the "jump-threshold" differs depending upon portfolio type, and indicates that our findings based on equal weighted portfolios are largely in accord with the findings elucidated above.

Fourth, the reason why small and large (signed) jump variation measures have differing marginal predictive content for returns is associated with the importance of realized skewness as a control variable in our experiments. Namely, we find that in double-sorted portfolios, the content of signed large jump variation is negligible when controlling for either signed total jump variation or realized skewness. By contrast, signed small jump variation has unique information for predicting future returns, even when controlling for total jump variation or realized skewness. This finding is consistent with the results from a series of Fama-MacBeth regressions, in which we control for multiple firm characteristics and risk measures.

Finally, small and large jump variation measures are driven by different economic factors and contain different information for predicting future returns. For example, large jumps are closely associated with "big" news. In particular, large earning announcement surprises increase both the magnitude and occurrence of large jumps. While such news related information is embedded in large jump variation, the information is generally short-lived, and dissipates too quickly to provide marginal predictive content for subsequent weekly returns. This is consistent with our finding that filtering out signed small jump variation, which we know to be useful, from signed total jump variation, results in increased predictive ability, relative to the case where only signed total jump variation is utilized in return forecasting, especially for big firms. Additionally, this finding is interesting, given that comparison of aggregated and weighted jump variation measures indicates that small jump variation captures idiosyncratic risks and can be diversified away.[5]

The rest of this paper is organized as follows. In Section 2.2 we discuss the model setup and define the jump risk measures that we utilize. Section 2.3 contains a discussion of the data used in our empirical analysis, and highlights key summary statistics taken from our dataset. Section 2.4 presents our main empirical findings, including discussions of results based on single portfolio sorts, double-sorts, cumulative return and Sharpe ratio analysis, firm-level Fama MacBeth regressions, and finally, jumps and

[5]This result is consistent with the finding of Amaya et al. (2015) that preference for positive asymmetry (skewness) may partially explain the idiosyncratic volatility puzzle, especially for small firms.

news announcements. Section 2.5 concludes.

## 2.2 Model Setup and Estimation Methodology

Following Aït-Sahalia and Jacod (2012), assume that the log price, $X_t$, of a security follows an Itô semimartingale, formally defined as:

$$X_t = X_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s + \int_0^t \int_{\{|x| \leq \epsilon\}} x(\mu - \nu)(ds, dx) + \int_0^t \int_{\{|x| \geq \epsilon\}} x\mu(ds, dx),$$

where $b$ and $\sigma$ denote the drift and diffusive volatility processes, respectively; $W$ is a standard Brownian motion; $\mu$ is a random positive measure with its compensator $\nu$; and $\epsilon$ is the (arbitrary) fixed cutoff level (threshold) used to distinguish between small and large jumps. As pointed out in Aït-Sahalia and Jacod (2012),, the continuous part of this model (i.e., the $\int_0^t \sigma_s dW_s$ term) captures normal hedgeable risk of the asset. The "big jumps" part of the model (i.e., the $\int_0^t \int_{\{|x| \geq \epsilon\}} x\mu(ds, dx)$ term) may capture big news-related events such as default risk, and the "small jumps" part of the model (i.e., the $\int_0^t \int_{\{|x| \leq \epsilon\}} x(\mu - \nu)(ds, dx)$ term) may capture large price movements on the time scale of a few seconds. If jumps are summable (e.g., when jumps have finite activity, so that $\sum_{s \leq t} \Delta X_s < \infty$, for all $t$), then the size of a jump at time $s$ is defined as $\Delta X_s = X_s - X_{s-}$.[6] In this context, the "true" price of risk is often defined by the quadratic variation, $QV_t$, of the process $X_t$. Namely,

$$QV_t = \int_0^t \sigma_s^2 ds + \sum_{s \leq t} \Delta X_s^2,$$

where the variation of the continuous component (i.e., the integrated volatility) is given by $IV_t = \int_0^t \sigma_s^2 ds$, and the variation of the price jump component is given by $QJ_t = \sum_{s \leq t} \Delta X_s^2$.

In the sequel, intraday stock returns are assumed to be observed over equally spaced time intervals in a given day, where the sampling interval is denoted by $\Delta_n$, and the number of intraday observations is $n$. Thus the intraday log-return over the $i$th interval is defined as

$$r_{i,t} = X_{i\Delta_n,t} - X_{(i-1)\Delta_n,t}.$$

[6] A jump process has finite activity when it makes a finite number of jumps, almost surely, in each finite time interval, otherwise it is said to have infinite activity.

It is well known that when the sampling interval goes to zero, the realized volatility, $RV_t$, which is calculated by summing up all successive intraday squared returns, converges to $QV_t$, as $n \to \infty$, where

$$RV_t = \sum_{i=1}^{n} r_{i,n}^2 \to_u QV_t = IV_t + QJ_t,$$

where $\to_u$ denotes convergence in probability, uniformly in time.

To separate jump variation from integrated volatility, Andersen et al. (2007) show that the jump and continuous components of realized variance can be constructed as:

$$RVJ_t = max(RV_t - \widehat{IV}_t, 0)$$

and

$$RVC_t = RV_t - RVJ_t,$$

respectively, where $\widehat{IV}_t$ is an estimator of $\int_0^t \sigma_s^2 ds$. Following Barndorff-Nielsen and Shephard (2004), and Barndorff-Nielsen et al. (2006), we use tripower variation to estimate the integrated volatility. In particular, define

$$\widehat{IV}_t = V_{\frac{2}{3},\frac{2}{3},\frac{2}{3}} \mu_{\frac{2}{3}}^{-3},$$

where $\mu_q = E(|Z|^q)$ is the $q$th absolute moment of the standard normal distribution, and

$$V_{m_1,m_2,...m_k} = \sum_{i=k}^{n} |r_{i,t}|^{m_1} |r_{i-1,t}|^{m_2} ... |r_{i-k+1,t}|^{m_k},$$

where $m_1$, $m_2$ ...$m_k$ are positive, such that $\sum_1^k m_i = q$. Based on the above decomposition approach, Duong and Swanson (2011, 2015) separate jump variation into small and large variation measures, using various truncation levels, $\gamma$. In particular, they define realized small and large jump variation measures as follows:

$$RVLJ_{\gamma,t} = min(RVJ_t, \sum_{i=1}^{n} r_{i,t}^2 I_{\{|r_{i,t}| \geq \gamma\}})$$

and

$$RVSJ_{\gamma,t} = RVJ_t - RVLJ_{\gamma,t},$$

respectively, where $I(\cdot)$ denotes the indicator function, which equals one if the absolute return is larger than the truncation level, and is otherwise equal to zero. We are also interested in upside and downside variation measures associated with positive and negative returns. Thus, following Barndorff-Nielsen et al. (2010) we construct realized semi-variances, defined as: $RS_t^+ = \sum_{i=1}^n r_{i,t}^2 I_{\{r_{i,t}>0\}}$, $RS_t^- = \sum_{i=1}^n r_{i,t}^2 I_{\{r_{i,t}<0\}}$, and $RV_t = RS_t^+ + RS_t^-$. They show that the upside and downside semi-variances ($RS_t^+$ and $RS_t^-$, respectively) each converge to the sum of one-half of the integrated volatility and the corresponding signed jump variation. Namely,

$$RS_t^+ \rightarrow_u \frac{1}{2}\int_0^t \sigma_s^2 ds + \sum_{s\leq t} \Delta X_s^2 I_{\{\Delta X_s>0\}}$$

and

$$RS_t^- \rightarrow_u \frac{1}{2}\int_0^t \sigma_s^2 ds + \sum_{s\leq t} \Delta X_s^2 I_{\{\Delta X_s<0\}}.$$

We construct upside and downside jump variation measures as follows:

$$RVJP_t = max(RS_t^+ - \frac{1}{2}\widehat{IV}_t, 0) \tag{2.1}$$

and

$$RVJN_t = max(RS_t^- - \frac{1}{2}\widehat{IV}_t, 0). \tag{2.2}$$

In addition, signed jump variation can be calculated as the difference between these upside and downside jump measures,

$$SRVJ_t = RVJP_t - RVJN_t. \tag{2.3}$$

This measure captures asymmetry in upside and downside jump variation.

In our analysis, we further decompose upside and downside jump variation measures into small and large components using thresholding method (see Mancini (2009), Duong and Swanson (2015), Li et al. (2017), and the references cited therein for discussion of thresholding methods). In particular, upside large jump variation based on fixed truncation level, $\gamma$, is defined as follows:

$$RVLJP_{\gamma,t} = min(RVJP_t, \sum_{i=1}^n r_{i,t}^2 I_{\{r_{i,t}>\gamma\}}) \tag{2.4}$$

and

$$RVLJN_{\gamma,t} = min(RVJN_t, \sum_{i=1}^n r_{i,t}^2 I_{\{r_{i,t}<-\gamma\}}). \tag{2.5}$$

We use a truncation level, $\gamma$, that is constructed by estimating $\alpha\sqrt{\frac{1}{t}\widehat{IV}_t^{(i)}}\Delta_n^{0.49}$, and is data-driven, accounting for the time-varying diffusive spot volatility of different stocks in the cross-section.[7] In the sequel, we consider three values for $\gamma$, say $\gamma^1$ (with $\alpha = 4$), $\gamma^2$ (with $\alpha = 5$), and $\gamma^3$ (with $\alpha = 6$). Signed large jump variation (i.e., large jump asymmetry) is defined as follows:

$$SRVLJ_t = RVLJP_t - RVLJN_t. \tag{2.6}$$

Our corresponding small jump variation measure is defined as the difference between total jump variation and large jump variation. Namely,

$$RVSJP_t = RVJP_t - RVLJP_t \tag{2.7}$$

and

$$RVSJN_t = RVJN_t - RVLJN_t. \tag{2.8}$$

Signed small jump variation is defined as:

$$SRVSJ_t = RVSJP_t - RVSJN_t. \tag{2.9}$$

In order to analyze the predictability of various jump measures in the cross-section, we normalize each of the jump variation measures discussed above by total realized variation.

Of note, is that a natural alternative to our approach for calculating the upside and downside jump variation measures in (2.1) and (2.2) is to use thresholding. Namely, instead of using tripower variation, one can use truncated realized variation (TRV) as a consistent estimator of integrated volatility, where $TRV_t = \sum_{i=1}^n r_{i,t}^2 I_{\{|r_{i,t}| \le \alpha_n\}} \to_u IV_t = \int_0^t \sigma_s^2 ds$. Upside and downside jump variation measures can then be calculated using:

$$RVJP_t = RS_t^+ - \sum_{i=1}^n r_{i,t}^2 I_{\{0 < r_{i,t} \le \alpha_n\}} \tag{2.10}$$

[7]For each stock, Li et al. (2017) use bipower variation as the fixed value for $\widehat{IV}_t^{(i)}$. We instead use bipower variation as the initial value for the integrated volatility $\widehat{IV}_t^{(0)}$, say, and $\widehat{IV}_t^{(i)}$ is estimated using truncated bipower variation with threshold $\gamma^{(i-1)}$, say, where $\gamma^{(i-1)}$ is fixed only when $|\widehat{IV}_t^{(i)} - \widehat{IV}_t^{(i-1)}|$ is smaller than $5\% \times \widehat{IV}_t^{(i-1)}$.

and

$$RVJN_t = RS_t^- - \sum_{i=1}^{n} r_{i,t}^2 I_{\{-\alpha_n \leq r_{i,t} < 0\}}, \tag{2.11}$$

where $\alpha_n$ is the truncation level.[8] Our empirical findings based on the use of (2.10) and (2.11) to define $RVJP_t$ and $RVJN_t$ are qualitatively the same as those reported in Section 2.4 based on the use of (2.1) and (2.2).

In order to measure skewness and kurtosis, we also construct higher order realized return moments. Following Amaya et al. (2015),

$$RSK_t = \frac{\sqrt{n}\sum_{i=1}^{n} r_{i,t}^3}{RV_t^{\frac{3}{2}}}, \tag{2.12}$$

standardized daily skewness is defined as: and normalized daily realized kurtosis is defined as:

$$RKT_t = \frac{n\sum_{i=1}^{n} r_{i,t}^4}{RV_t^2}. \tag{2.13}$$

Finally, it should be noted that we follow Amaya et al. (2015) and Bollerslev et al. (2019b), and conduct our cross-sectional analysis at the weekly frequency. In particular, on each Tuesday, we compute the following weekly realized measures: $RV_t^W = (\frac{252}{5}\sum_{i=0}^{4} RV_{t-i})^{1/2}$ and $RM_t^W = \frac{1}{5}(\sum_{i=0}^{4} RM_{t-i})$, where $RV_t$ is defined above, and where $RM_t$ denotes any of the realized measures defined above other than $RV_t$ (e.g., $RVJP_t$, $RVJN_t$, $SRVJ_t$, etc.) Hereafter, we shall drop the superscript "W" for the sake of notational brevity. All of the descriptors used to denote the various realized measures constructed in our empirical analysis are summarized in Table 2.1.

As described in detail in Section 2.4, the realized measures outlined above are used in a number of different ways in our empirical analysis. First, we carry out single portfolio sorts, in which we sort stock portfolios on the above realized jump measures, and predict average excess returns, one-week ahead. In these experiments, we also calculate alphas based on regressions that utilize the Fama-French and Carhart factors. In this first part of our analysis, we also examine cumulative returns and Sharpe ratios. In addition to the single portfolio sorts, we carry out double portfolio sorts, in which we sort not only on realized jump risk measures, but also on various control variables, including

[8] Here, the threshold, $\alpha_n = 3\sqrt{\frac{1}{t}\widehat{IV}_t^{(i)}}\Delta_n^{0.49}$, is estimated using the same procedure as in footnote 7.

realized skewness and other firm specific characteristics. Using these double sorts, we also examine the inter-play between individual stock-level jump variation and industry-level jump variation. Needless to say, the purpose of our double-sorts is to examine the robustness of our findings based on single sorts, after controlling for other realized measures. Next, we carry out a series of Fama-MacBeth regressions, in order to check the robustness of our findings to the inclusion of various firm specific characteristics. Finally, we carry out an event study in which the effect of earning surprises on realized jump measures is examined. For complete details, see Section 2.4.

## 2.3 Data

We utilize high frequency trading data obtained from the consolidated Trade and Quote (TAQ) database. In particular, we analyze all stocks in the TAQ database that are listed on the NYSE, Amex, and NASDAQ stock exchanges. There are 15,585 unique stocks during the 1,246 weeks analyzed in this paper.[9] The sample period is from January 4, 1993 to December 31, 2016. Intraday prices are sampled at five minute intervals from 9:30 a.m. to 4:00 p.m. from Monday to Friday. Overnight returns are not considered in this paper, and days with less than 80 transactions at a 5 minute frequency are eliminated. For example, if AAPL has less than 80 trades on a particular day, then AAPL is dropped from our sample, but only for that day. All high frequency data used in this paper are cleaned to remove trades outside of exchange hours, negative or zero prices or volumes, trade corrections and non-standard sale conditions, using the methodology described in Appendix A.1 in Bollerslev et al. (2019b).

We constructed two variants of our dataset. The first is cleaned as discussed above. The second classifies five minute intraday returns greater than 15% as abnormal and replaces them with zeros. In the sequel, results based on analysis of the second dataset are reported. However, results based on utilization of the first dataset are qualitatively the same; and indeed key return results reported in this paper generally change by 1

[9]In some cases, multiple TAQ symbols are matched with a unique Center for Research in Security Prices (CRSP) PERMNO. Over each quarter, the TAQ symbol which has the most observations is kept and the other overlapping observations are dropped.

basis point or less when the former dataset is used in our analysis. Complete results are available upon request from the authors.

Daily and monthly returns, and adjusted numbers of shares for individual securities are collected from the CRSP database. Delisting returns in CRSP are used as returns after the last trading day. Daily Fama-French and Carhart four factor (FFC) portfolio returns are obtained from Kenneth R. French's website.

Following Amaya et al. (2015) and Bollerslev et al. (2019b), we also construct various lower frequency firm level variables that might be related to future returns, such as the market beta (BETA), the firm size, the book-to-market ratio (BEME), momentum (MOM), short-term reversals (REV), idiosyncratic volatility (IVOL), co-skewness (CSK), co-kurtosis (CKT), maximum (MAX) and minimum (MIN) daily return in the previous week, and the Amihud (2002) illiquidity measure (ILLIQ). For a complete list of these firm specific control variables, refer to Table 2.1. For a detailed description of these variables, including the methodology used to construct them, see Appendix A.2 in Bollerslev et al. (2019b).

Note that while the majority of our analysis is based on the examination of individual stocks, in our double sorts, there are some cases (that are reported in Section 2.4.4) where we examine the inter-play between individual stock-level jump variation and industry-level jump variation. In this case, we follow the Fama-French industry classification approach, and group stocks into 49 industries based on their SIC codes, which are obtained from CRSP.

### 2.3.1 Unconditional Distributions of Realized Measures

Figure 2.1 displays kernel density estimates of the unconditional distributions of each of our realized measures, across all firms and weeks. The top two panels in the figure show the distributions of signed jump variation and realized skewness. Both of these distributions are approximately symmetric and peaked around zero. The skewness distribution is more fat-tailed, however.[10] The middle two panels of Figure 2.1

[10] The kurtosis of signed jump variation is 4.36. For realized skewness, the analogous statistic is 12.04.

display the distributions of signed small and large jump variation. Similar to signed jump variation, both signed small and large jump variation mesaures are approximately symmetric around zero, but signed small jump variation is less fat-tailed.[11] Consistent with the results in Amaya et al. (2015) and Bollerslev et al. (2019b), realized volatility and realized kurtosis are both right skewed and very fat-tailed, as shown in the bottom two panels of the figure.[12]

Figure 2.2 shows the time variation in the cross-sectional distribution of each realized measure using 10-week moving averages. In particular, 10th, 50th, and 90th percentiles for each realized measure in the cross-section are plotted. Thus, dispersion at any given time in these plots reflects information about the cross-sectional distribution of the realized measure. Inspection of Panels A and B in the figure reveal that signed jump variation and realized skewness have stable dispersion, for all three cross-sectional percentiles, over time, while the cross-sectional dispersion in realized volatility and kurtosis are rather time-dependent (see Panels C and D). Additionally, similar to the cross-sectional distribution of signed jump variation, the percentiles for signed small and large jump variation measures are quite steady over time, as indicated in Panels E-H.

### 2.3.2 Summary Statistics and Portfolio Characteristics

Table 2.2 contains various summary statistics for all of the realized measures summarized in Table 2.2. In Panel A, the cross-sectional means and standard errors for each of the realized measures is given. This is done for two different truncation levels, denoted as $\gamma^1 = 4\sqrt{\frac{1}{t}\widehat{IV}_t^{(i)}}\Delta_n^{0.49}$ and $\gamma^2 = 5\sqrt{\frac{1}{t}\widehat{IV}_t^{(i)}}\Delta_n^{0.49}$. As might be expected, jump variation is quite sensitive to the choice of $\gamma$. For example, the (normalized) mean of RVSJP (positive (upside) small jump variation) increases from 0.1180 to 0.1715 when the threshold is increased from $\gamma^1$ to $\gamma^2$. Needless to say, various measures remains the same, as they are independent of $\gamma$.

[11] The unconditional kurtosis is 6.43 and 3.51, for signed small and large jump variation, based on truncation level $\gamma^1$; and 8.87 and 3.09 based on truncation level $\gamma^2$, respectively.

[12] The kurtosis is 15.85 and 27.24, for the unconditional distribution of realized volatility and realized kurtosis, respectively.

Panel B of Table 2.2 contains cross-sectional correlations for all of the realized measures. In accord with the findings reported by Amaya et al. (2015) and Bollerslev et al. (2019b), signed jump variation (SRVJ) and realized skewness (RSK) are highly correlated with each other and have significantly positive correlations with the short term reversal variable (REV); as well as with maximum (MAX) and minimum (MIN) daily returns in the previous week.

Interestingly, we also find that signed large jump variation (SRVLJ) is highly correlated with SRVJ and with RSK. However, signed small jump variation (SRVSJ) has lower correlation with SRVJ and much smaller positive correlation with RSK. This finding is consistent with our finding discussed below that realized skewness captures information that is primarily contained in large jumps; and serves as an important distinction between the findings in this paper and those reported in the papers discussed above.

Table 2.3 complements Table 2.2 by sorting stocks into quintile portfolios based on different realized measures. On each Tuesday, stocks are ranked by the realized variation measures, and we calculate the equal-weighted averages of each firm characteristic in the same week. Panels A, B, C and E report summary statistics for portfolios sorted by SRVJ, SRVLJ, SRVSJ, and RSK, respectively. Consistent with the correlations contained in Table 2.2, firms with larger signed small and large jump variation measures tend to have higher signed jump variation, realized skewness, REV, MAX and MIN. Firms with high realized volatility and realized kurtosis (see Panels D and F) tend to be illiquid and small.[13]

## 2.4 Empirical Results

In this section, results based on stocks that are sorted into quintile portfolios based on a single different realized measure are first reported. These single (univariate) portfolio sort results are collected in Tables 2.4 to 2.7. Results based on double sorts are reported

[13]See the Supplementary Appendix for results based on the examination of additional quintile portfolios that are constructed based on ex-ante risk measures and displayed with ex-post risk measures. It is clear that sorting stocks based on jump risk measures results in portfolios with the desired risk exposures.

(in Tables 2.8 to 2.10). We assume a weekly holding period, and return calculations reported in the tables are carried out as follows. At the end of each Tuesday, stocks are sorted into quintile portfolios based on different realized variation measures (see Panel A of Table 2.1). We then calculate equal-weighted and value-weighted portfolio returns over the subsequent week. We report the time series average of these weekly returns for each portfolio (these returns are called "Mean Return" in the tables) In addition, we regress excess return of each portfolio on the Fama-French and Carhart (FFC4) factors to control for systematic risks, using regression of the form

$$r_{i,t}-r_{f,t}=\alpha_i+\beta_i^{MKT}(MKT_t-r_{f,t})+\beta_i^{SMB}SMB_t+\beta_i^{HML}HML_t+\beta_i^{UMD}UMD_t+\epsilon_{i,t} \tag{2.14}$$

where $r_{i,t}$ denotes the weekly return for firm i, $r_{f,t}$ is the risk-free rate; and $MKT_t$, $SMB_t$, $HML_t$, and $UMD_t$ denote FFC4 market, size, value and momentum factors, respectively. The intercepts from these regressions (called "Alpha" in our tabulated results), measure risk-adjusted excess returns, and are also reported in Tables 2.4 to 2.13. Needless to say, our objective in these tables is to assess whether predictability exists, after controlling for various systematic risk factors. Finally, in Tables 2.14, we report the results of cross-sectional (firm level) Fama-MacBeth regressions used to investigate return predictability when simultaneously controlling for multiple realized measures and firm characteristics.

### 2.4.1 Single (Univariate) Portfolio Sorts Based on Realized Measures

In this section, we first discuss the results contained in Table 2.4. Recall that the "Mean-Return" in this table is an average taken over our entire time series of equal-weighted and value-weighted portfolio returns, for single sorted portfolios based on positive jump variation (RVJP), negative jump variation (RVJN) and signed jump variation (SRVJ). Values in parentheses are Newey-West $t$-statistics (see Bollerslev et al. (2015) and Petersen (2009) for further discussion). Panel A provides results for portfolios sorted by RVJP. Inspection of the entries in this panel indicate that mean returns and alphas of high-low portfolios (i.e., the difference in returns (alphas) between the fifth and first quintiles) are all negative, indicating a negative association between RVJP and subsequent stock returns. Interestingly, the alpha of -7.71 basis points (bps) is insignificant for the high-low spread for the equal weighted portfolio, while the mean return of -5.63 bps is only significant at a 10% level for the high-low spread for the value weighted portfolio.

The lack of statistical significance for some of the mean return values reported in Panel A does not characterize our findings when negative and signed jump variation measures are utilized for sorting. Moreover, the magnitudes of the mean returns and alphas are usually three or more times larger when sorting on negative and signed jump variation (to see this, turn to Panels B and C of Table 2.4). In Panel B, the high-low spread of mean returns equals 36.06 bps, with a $t$-statistic of 6.47 for the equal-weighted portfolio, and 15.13 bps with a $t$-statistic of 3.75 for the value-weighted portfolio. Moreover, both equal-weighted and value-weighted portfolios generate significant positive abnormal future returns measured by the alphas. These results clearly point to a statistically significant positive association between negative jump variation and the following week's returns.

Panel C in Table 2.4 contains results for portfolios sorted by signed jump variation. The negative high-low spreads indicate a statistically significant negative association between signed jump variation and future returns. In particular, a strategy buying stocks in the lowest signed jump variation quintile and selling stocks in the highest signed jump variation quintile earns a mean return of 40.82 bps with a $t$-statistic of 9.85 each week for the equal-weighted portfolio and 25.02 bps with a $t$-statistic of 5.78 for the value-weighted portfolio. These results are consistent with the results reported in Bollerslev et al. (2019b). Interestingly, almost all of the mean returns listed in Table 2.4 are "alpha" (see tabulated average alphas in the table), and cannot be explained by standard portfolio risk factors using regressions of the type given above as equation (2.14).

A key question that we provide evidence on in this paper is whether the results summarized in Table 2.4 carry over to the case where small and large jump variation is separately sorted on. First, consider large jumps. Table 2.5 reports the results for portfolios sorted by positive, negative and signed large jump variation, respectively. Similar to positive jump variation, positive large jumps negatively predict subsequent returns, but the predictability is not significant, regardless of the truncation level ($\gamma$) used to separate small and large jumps, and regardless of portfolio weighting used. This is evidenced by the fact that the $t$-statistics for mean returns and alphas of high-low portfolios all indicate insignificance, at a 5% testing level, regardless of truncation level. Thus, there is no ambiguity, as in Panel A of Table 2.4. Positive jump variation is not a significant predictor, under our large jump scenario. On the other hand, we shall see that sorting on small and large negative variation measures yields significant excess returns, as does sorting on small positive jump variation, under both equal and value weighting schemes.

As just noted, equal-weighted high-low portfolios sorted on large negative jump variation generate significant positive returns and alphas (see Panel B of Table 2.5). However, analogous returns and alphas under value weighting are not significant. Signed large jump variation is sorted on in Panel C of Table 2.5. Signed large jump variation is useful for undertaking a long-short trading strategy based on the difference between large upside and downside jump variation measures. Inspection of the results in this panel of the table reveals that the high-low spread for the equal weighted portfolio generates an average risk-adjusted weekly return of -28.36 bps (with a $t$-statistics of -9.39) and -9.25 bps (with a $t$-statistics of -2.87) for the value-weighted portfolio, for truncation level equal to $\gamma^1$. Results based on $\gamma^2$ (i.e., our larger truncation level) are also significant, although magnitudes are lesser and only for our equal weighted portfolio.[14] In particular, observe that when large jump variation is constructed using $\gamma^2$, the high-low spreads for value-weighted portfolios sorted by downside or signed large jump variation measures are insignificant, suggesting that small firms have stronger relationships (than larger firms) between signed (or negative) large jump variation and subsequent returns. This may be due to the fact that smaller firms are in some ways more susceptible to changing market conditions than larger firms.

[14]Empirical findings based on $\gamma^3$ are similar to those discussed above, and hence are not reported. This robustness of our findings to the choice of $\gamma$ also characterizes the other empirical findings discussed in the sequel.

Table 2.6 summarizes results analogous to those reported in Table 2.5, but for positive, negative and signed small jump variation measures. Similar to large jump measures, positive and signed small jump variation measures negatively predict future returns, and negative small jump variation measures positively predict returns in the following week. By contrast, the differences in average (risk-adjusted) returns between equal-weighted and value-weighted long-short portfolios based on RVSJP and RVSJN are smaller than those for portfolios based on large jumps (compare the entries for the high-low quintiles under the two weighting schemes in Panels A and B of Table 2.6 with like entries in Panels A and B of Table 2.5). These results indicate that big firms have a stronger relationship between small jump variation and future returns than that between large jumps and subsequent weekly returns. Since stocks for big firms are more liquid and price discovery more rapid, the predictabilities of large jumps are much weaker or insignificant for big firms. This finding is in line with Bollerslev et al. (2019b), who document that the predictability of signed jump variation is stronger for small and illiquid firms and is driven by investor overreaction. In addition, when using our larger truncation level, $\gamma^2$, value-weighted high-low spreads based on signed small jump variation are larger than those based on signed total jump variation and signed large jump variation. This result implies that a long-short strategy associated with signed small jump variation generates the highest value-weighted risk-adjusted returns, given the use of an appropriate truncation level to separate small and large jumps.

Table 2.7 reports results for portfolios sorted by realized volatility, realized skewness, realized kurtosis, and continuous variance. Consistent with the results in Amaya et al. (2015) and Bollerslev et al. (2019b), there is a significant negative relationship between realized skewness and future returns, while the association is not significant between either realized volatility or realized kurtosis and returns in the following week, regardless of portfolio weighting scheme. In addition, continuous variance significantly and negatively predicts one-week ahead returns for equal-weighted portfolios, but this negative association is not significant for value-weighted portfolios.

### 2.4.2 Cumulative Returns and Sharpe Ratios

Not surprisingly, our findings based on univariate portfolio sorts suggest that strategies that utilize different realized measures deliver different risk-adjusted average returns. In order to investigate this result further, we calculate cumulative returns and Sharpe ratios for short-long portfolios, sorted on various risk measures that are described in Table 2.1, including SRVJ, RSJ, SRVLJ, SRVSJ, and RSK. In addition, for comparison purposes, we also carry out our analysis using the relative signed jump variation measure (called RSJ) that is examined by Bollerslev et al. (2019b). Our experiments are carried out as follows. Beginning in January 1993, various short-long portfolios are constructed, with an initial investment of \$1. These portfolios are re-balanced and accumulated at a weekly frequency, until the end of 2016.[15] Figure 2.3 plots portfolio values over time. Consistent with our results based on single portfolio sorts, inspection of the plots in this figure indicates that for equal-weighted portfolios sorting on signed jump variation (SRVJ) yields the largest portfolio accumulations; and for value-weighted portfolios, sorting on signed small jump variation (SRVSJ) yields the largest portfolio accumulations. [16]

Now, consider the Sharpe ratios reported below, which are reported for various jump measures, and are constructed based on truncation level $\gamma^2 = 5\sqrt{\frac{1}{t}\widehat{IV}_t^{(i)}}\Delta_n^{0.49}$.

Sharpe Ratios

| | SRVJ | RSJ | SRVLJ | SRVSJ | RSK |
|---|---|---|---|---|---|
| Equal-Weighted | 2.1342 | 2.1363 | 1.8556 | 1.8161 | 2.2234 |
| Value-Weighted | 1.1322 | 1.1310 | 0.1611 | 1.2755 | 0.8665 |

The entries in this table are Sharpe ratios for equal and value-weighted short-long portfolios constructed using SRVJ, RSJ, SRVLJ, SRVSJ, and RSK. Recall that RSK is realized skewness (see Table 2.1 for definitions of these measures). The sample of stocks used for Sharpe ratio calculations includes all NYSE, NASDAQ and AMEX listed

[15]Cumulative returns calculations do not include the risk-free rate. For a definition of cumulative returns both with and without the weekly risk-free rate, see Bollerslev et al. (2019b).

[16]Note that RSJ, which measures the same signed jump variation as SRVJ, although using different estimation methodology, generates the highest cumulative return for equal-weighted portfolios, but is dominated by SRVSJ for value-weighted portfolios.

stocks for the period January 1993 to December 2016. At the end of each Tuesday, all of the stocks in the sample are sorted into quintile portfolios based on ascending values of various realized risk measures. A high-low spread portfolio is then formed as the difference between portfolio 1 and portfolio 5, and held for one week, where 1 and 5 refer to quintiles, as in Tables 2.3 to 2.7. The Sharpe ratio is calculated with the one-week ahead returns.

Interestingly, for equal-weighted portfolios, the RSK-based short-long strategy yields the highest Sharpe ratio (i.e., 2.2234), although the ratio of 2.1342 for SRVJ is approximately the same. Still, the success of the RSK measure is likely due to its relatively stable performance, compared with other jump-based strategies. This finding is similar to the findings discussed in Xiong et al. (2016), who show that tail-risks can be substantially reduced by forecasting skewness. Note also that the signed small jump variation (SRVSJ) based portfolio has the highest Sharpe ratio, among all value-weighted portfolios. However, it is clear that all equal-weighted portfolios outperform their corresponding value-weighted counterparts. This result is consistent with the finding discussed above that small and illiquid firms tend to react more strongly to realized risk measures.

### 2.4.3 Double Portfolio Sorts Based on Realized Measures

To further investigate whether small and large jumps are priced differently, we utilize double portfolio sorts. In particular, we carry out double sorts in order to examine the robustness of our findings based on single sorts, after controlling for other realized measures. Table 2.8 reports returns and alphas from various of these sorts in which we alternate the sorting order among SRVJ, SRVLJ and SRVSJ. When we first sort by total jump variation, and then sort stocks based on SRVLJ or SRVSJ, a negative relation only exists between SRVSJ and subsequent weekly returns (see Panels A and B of the table). This result indicates that there is no marginal predictive content associated with large jumps, when conditioning on the predictive content associated with total jump variation, while small jumps have unique information for predicting future returns, even compared to total jumps.

Panel C reports returns and alphas based on sorting on SRVSJ after controlling for SRVLJ. Both the equal- and value-weighted high-low spreads and alphas are statistically significant in this case, while this is not the case if stocks are first sorted by SRVSJ and then by SRVLJ, as shown in Panel D. More specifically, the high-low return is -25.38 bps (with a t-statistic of -6.61), for the value-weighted portfolio in Panel C, and is -3.77 bps with a $t$-statistics of -1.49 in Panel D, for the value-weighted portfolio. This indicates that the predictable content in large jumps becomes negligible after controlling for small jumps.

Bollerslev et al. (2019b) document that the negative association between realized skewness and one-week ahead returns is reversed when controlling for the signed jump variation. To further investigate the relationship between skewness and different jump variation measures, we use double portfolio sorts to control for different effects that are associated with cross-sectional variation in future returns.

Panel A of Table 2.9 reports average returns and corresponding $t$-statistics for 25 portfolios sorted by SRVJ (signed jump variation), controlling for realized skewness (RSK). At the end of each Tuesday, stocks are first sorted into quintiles based on realized skewness, and then within each quintile portfolio, we further sort stocks into quintiles based on signed jump variation. We also report the equal- and value-weighted returns in the following week and Fama-French and Carhart four-factor alphas for the long-short portfolios and the averaged portfolios across quintiles. Inspection of the results in this table indicates that the negative association between SRVJ and future returns still exists, after controlling for RSK, indicating that there is unique predictive information contained in signed jump variation. Panel B in this table reports results for portfolios sorted first by SRVJ and then by RSK. The high-low spreads of the averaged portfolios are positive after controlling for SRVJ, confirming the results reported in Bollerslev et al. (2019b).

Panel A of Table 2.10 contains results for portfolios sorted by SRVLJ (signed large jump variation) after controlling for RSK. As noted above, the negative association between SRVLJ and future returns is reversed after controlling for skewness. By contrast, this issue doesn't exist for portfolios sorted by SRVSJ (signed small jump variation) when controlling for skewness, as shown in panel B of Table 2.10, indicating that signed small jump variation has unique information about future return premia. However, first accounting for skewness negates the usefulness that signed large jump variation has for predicting future returns. This finding serves as an important distinction between the predictive content of small and large jumps. Whereas the former can be forecast by realized skewness, the latter cannot.

Finally, Tables 2.11 contains results for portfolios sorted on RSK, after controlling for SRVLJ and SRVSJ, respectively. Inspection of the entries in this table indicates that the high-low spreads are negative, except in select value weighted portfolio cases, when controlling for SRVSJ. This is not surprising since skewness captures information from both SRVLJ and SRVSJ, while the negative association between realized skewness and subsequent returns remains, when controlling for either SRVLJ or SRVSJ, in most cases. Of note is that this negative association disappears for some value-weighted portfolios, when controlling for SRVSJ, suggesting that signed small jump variation (especially for big firms) is the main driver of the signed total jump variation. These findings are consistent with the findings documented by Bollerslev and Todorov (2011) that S&P 500 market portfolios tend to have symmetric jump tails (large jumps).

### 2.4.4 Using Double Portfolio Sorts to Examine Stock-Level Versus Industry-Level Predictability

In this section, we carry out an additional set of double portfolio sort experiments, in which industry based investing is compared with individual stock based investing. Our earlier findings indicate that low signed jump variation investing (buying stocks with low signed jump variation and shorting stocks with high signed jump variation) can deliver significant risk-adjusted returns (this is similar to low risk investing, and is a result also found by Bollerslev et al. (2019b), for example). In order to examine whether this investment strategy relies on industry betting or stock selection within industries (or both), we form double sorted portfolios based on industry-level and stock-level signed jump risk variation. In particular, each Tuesday we group stocks into 49 industries based on SIC codes. Industry-level signed jump risk is calculated as the value-weighted average of signed (large/small) jump variation measures for stocks within each industry. Thus, stocks in the same industry have the same industry signed (large/small) jump variation during a given week. Stock-level signed jump risk is calculated as outlined in the above. Double sorts are then used to investigate the selection effects at industry- and stock-level. Namely, stocks are sorted into 25 portfolios based on industry- and stock-level signed (large/small) jump variation quintiles. With this particular variety of sorting, results are independent of the order in which stocks are sorted.

Figure 2.4 depicts the percentage of stocks in each portfolio (see Panel A), and the market capitalization in these portfolios (see Panel B). If industry-level selection and stock-level selection lead to different quintile portfolios (i.e. off-diagonal portfolios in the figures have non-zero membership), it is possible to separate these two effects using double sorts. Namely, there are different industry- and stock-level effects. Both panels indicate this to be the case.

Tables 2.12 to 2.13 report our empirical findings based on our double portfolio sort experiments. In particular, Table 2.12 reports results for sorting done on signed jump variation (SRVJ), while Table 2.13 reports results for sorting done on signed large jump variation (SRVLJ) and signed small jump variation (SRVSJ), respectively. Entries in the tables are mean returns and alphas, as in previous tables. However, in these tables we also report industry-level effects and stock-level effects. These are reported in the last two rows of entries in each panel of the tables. The first of these two rows, called “Industry-Level Effect” reports average high-low returns and alphas by averaging across quintiles in the high-low and alpha columns of the table (these are industry-level results). The second of these two rows, called “Stock-Level Effect” reports average high-low returns and alphas by averaging across quintiles in the high-low and alpha rows of the table (these are stock-level results). Summarizing, rows in these tables display portfolios formed by stocks in the same stock-level SRVJ, SRVLJ, or SRVSJ quintiles, while columns report results for portfolios formed by stocks in the same industry-level SRVJ, SRVLJ, or SRVSJ quintiles.

Turning to Table 2.12, notice, for example, that a strategy of buying stocks in the highest industry SRVJ quintile and selling stocks in the lowest industry SRVJ quintile generates an equal-weighted average return of 29.63 bps with a $t$-statistic of 5.66, and the corresponding value-weighed average return is 11.48 bps with a $t$-statistics of 2.23 (see Table 2.12). This finding is interesting, as it suggests that the negative association between SRVJ and future returns is reversed in the industry level. The equal-weighted average of the high-low row (i.e., the average stock-level effect) is -45.28 bps with a $t$-statistic of -11.39 and the alpha is -44.70 bps with a $t$-statistic of -11.50, indicating that the stock-level effect is economically significant. At the stock-level, investors prefer stocks with high SRVJ, requiring lower returns under higher SRVJ, given that there is a large probability of extremely large positive jumps. By contrast, when sorting at the industry-level, investors are more interested in industry exposure with lower SRVJ, or in return distributions concentrated to the right. Lottery-like payoff exposure comes from individual stocks, not from industry bets. These results are mirrored in Table 2.13, where SRVLJ and SRVSJ are the sorting measures. However, average stock- and industry-level returns and alphas are much higher under SRVSJ sorting than under SRVLJ sorting. For example, buying stocks in the highest industry SRVLJ quintile and selling stocks in the lowest industry SRVLJ quintile generates an equal-weighted average return of 14.83 bps with a $t$-statistic of 3.77 under SRVLJ sorting (see panel A of Table 2.13), versus an equal-weighted average return of 26.69 bps with a $t$-statistic of 5.02 under SRVSJ sorting (see panel B of Table 2.13).

### 2.4.5 Firm-Level Fama-MacBeth Regressions

Table 2.14 gathers results based on firm-level Fama-MacBeth regressions, which we run in order to investigate the return predictability associated with variation measures, when controlling for multiple firm specific characteristics. Regressions are carried out as follows. At the end of each Tuesday, we run the cross-sectional regression,

$$r_{i,t+1} = \gamma_{0,t} + \sum_{j=1}^{K_1} \gamma_{j,t} X_{i,j,t} + \sum_{s=1}^{K_2} \phi_{s,t} Z_{i,s,t} + \epsilon_{i,t+1}, \quad t = 1, ..., T, \tag{2.15}$$

where $r_{i,t+1}$ denotes the stock return for firm $i$ in week $t+1$, $K_1$ is the number of potential variation measures, and $X_{i,j,t}$ denotes a relevant realized measure at the end of week $t$. In addition, there are $K_2$ variables measuring firm characteristics, which are denoted by $Z_{i,j,t}$ (see Section 2.3 for details). After estimating the cross-sectional regression coefficients on a weekly basis, we form the time series average of the resulting $T$ weekly $\widehat{\gamma}_{j,t}$ and $\widehat{\phi}_{s,t}$ values, in order to estimate the average risk premium associated with each risk measure. Namely, we construct

$$\widehat{\gamma}_j = \frac{1}{T}\sum_{t=1}^{T}\widehat{\gamma}_{j,t}, \quad \text{and} \quad \widehat{\phi}_s = \frac{1}{T}\sum_{t=1}^{T}\widehat{\phi}_{s,t}, \quad \text{for} \quad j = 1, ..., K_1,\ s = 1, ..., K_2.$$

Panel A of Table 2.14 reports results for regressions on various realized variation measures, without controlling for firm specific characteristics. Consistent with our results based univariate sorting, signed jump variation (SRVJ) significantly negatively predicts cross-sectional variation, in these weekly returns regressions. Additionally, both signed small and large jump variation measures negatively predict future weekly returns. Finally, both small and large upside (downside) jump variation measures negatively (positively) predict subsequent weekly returns. However, when including measures that contain information from both small and large jump variation measures, as well as realized skewness, the negative association between skewness and future returns is reversed (see the results for the regressions labeled IX, XII, XV, XVI). In particular, skewness drives out signed large jump variation in regression XIII by reverting the negative association between the latter and future returns. If only small jumps are considered as control variables, skewness still negatively predicts future returns. This again indicates that signed small jump variation has unique and significant information about future returns.

Panel B of Table 2.14 reports regression results for the same set of regressions in panel A, but controlling for various firm specific characteristics, ranging from BETA to ILLIQ (see Table 2.1 for details). In these regressions, signed (small) jump variation is always significant. Additionally, skewness significantly negatively predicts future returns in regressions that only include small jump variation. This provides yet further evidence that signed small jump variation has unique and significant information about future returns, while large jumps have information in common with realized skewness.

### 2.4.6 Pricing Distinctions Between Small and Large Jumps

The results in previous sections show that small and large jump variation measures contain different information, and thus have different predictive content. To further investigate whether the differences are driven by distinct economic factors, we provide empirical evidence on the inter-relationship between jumps and news.

#### 2.4.6.1 Jumps and News Announcement

We begin by examining the relationship between jumps and firm-level news announcements. In order to do this, we construct event windows using the approach of Bernard and Thomas (1989). We then plot the dynamics of SRVJ, SRVLJ, and SRVSJ around earnings announcements. In particular, following Livnat and Mendenhall (2006), the earning surprise (SUE) for each stock is defined as

$$SUE_{j,t} = \frac{(X_{j,t} - E_{j,t})}{P_{j,t}}, \tag{2.16}$$

where $E_{j,t}$ and $X_{j,t}$ denote the analysts' expectations and reported actual earnings per share, respectively. Here, $P_{j,t}$ is the price per share for stock j at the end of quarter t. In a [-12,12] week event window, where week zero denotes the earning announcement week, stocks are sorted into tertile portfolios by the value of SUE at the end of week zero. We then calculate the equal-weighted and value-weighted average of jump measures for each tertile portfolio at each week. Figure 2.5 displays various jump variation measures of portfolios with the most negative, median, and positive earning surprises. It turns out that large (both positive and negative) jump variation measures are higher during announcement weeks, regardless of news sentiment (i.e., regardless of whether SUE is positive or negative). However, positive large jump variation (RVLJP) is higher on days with the most positive earning surprises, and negative large jump variation (RVLJN) reaches its peak on days with the most negative earning surprises. In contrast, both small positive and negative jump variation measures (RVSJP and RVSJN) have lower magnitudes during announcement weeks. The size of the reduction associated with small positive jump variation (RVSJP) is larger on days with the most negative earning surprises, while small negative jump variation (RVSJN) decreases the most on days with the most positive surprises. For signed jump variation, jump magnitudes increase (relative to non-earnings-surprise weeks) on positive surprise days and decrease on negative surprising days. These results indicate that big news, regardless of sentiment, simultaneously leads to increases in the magnitude of large jump variation, and reductions in the level of small jump variation.

The other direction in which we investigate the linkage between news announcements and jump variation is based on an exploration of whether news announcements affect the frequency of occurrence of either small or large jumps. Table 2.15 reports the average percentage of firms exhibiting particular types of jumps on days with and without earning surprises. Specifically, on each announcement date, all stocks exhibiting earnings are sorted into tertile portfolios based on the absolute value of the earning surprise (SUE). The categories sorted on are denoted as "small", "medium", and "large", with tertiles calculated by appropriate sorting of the firms based on the absolute values of the firms' earnings surprise magnitudes. Then, within each tertile, the percentage of firms exhibiting a particular type of jump (averaged across all earnings surprise days) is calculated and reported. For these calculations, only days in which at least 3 firms report earning surprises and included in our sample.[17] Thus, for example, if 12 firms report earning surprises, then 4 firms will be represented in each of the 3 tertiles. Turning to the results in the table, note, for example, that the entry 0.3042 in the sixth column of Panel A indicates that 30.42% of firms in the "small surprise" tertile portfolio recorded a large jump (measured by SRVLJ) on small surprise days, on average, across the entire daily sample. By contrast, 89.83% of firms exhibit small jumps (measured by SRVSJ) on days with small surprises.

Two clear conclusions emerge upon examination of the results in this table. First, when the magnitude of earning surprises increases, the average percentage of firms with large jumps (SRVLJ) increases from 30.42% to 37.37%. In particular, in Panel A, note that for the "Small" tertile, the percentage of firms exhibiting large jumps (SRVLJ) is 30.42%, while for the "Large" tertile, the percentage is 37.37%. By contrast, the

[17] Results are virtually identical if we only include days in which at least 12 or 24 firms report earnings surprises.

percentage of firms with small jumps decreases as the relative magnitude of earnings surprises increases (i.e., the percentage of firms associated with SRVSJ decreases from 89.83% to 88.29%). This result indicates that "big news" is associated with an increase in the prevalence of large jumps. Second, the prevalence of jumps differs depending upon whether one tabulates results on earnings surprise days (Panel A) or on non-earnings surprise days (Panel B). For example, large news surprises are associated with large jumps for 31.07% of firms on non-announcement days (see Panel B) and 37.38% of firms on announcement days (Panel A). This result is consistent with event study finding that jump magnitudes are larger on announcement days than non-announcement days.

It is also worth noting that Panel C of Table 2.15 reports t-statistics that test whether the differences in percentages of jumps in different portfolios are significant. In this table, "None" refers to the case where percentages are calculated on non-earnings-announcement days. Thus, the fact that the "Large-None" t-statistic associated with SRVLJ is 16.85, indicates that the percentage of large jumps on "large-surprise" earnings announcement days is significantly greater than the percentage of large jumps on non-earnings-announcement days. This in turn implies that large jumps tend to occur on "large-surprise" earnings announcement days. On the other hand, the reverse is true in the case of small jumps. In particular, the "Large-None" t-statistic associated with SRVSJ is -10.85, indicating that small jumps tend to occur on non-earnings-announcement days.

#### 2.4.6.2 Systematic Versus Idiosyncratic Risks

To further explore the unique information embedded in either large or small jump variation measures, and examine their association with systematic and idiosyncratic risks, we identify the effect of diversification on both small and large jumps. In order to do this, we construct two alternative measures of SRVLJ and SRVSL. The ratio of these is plotted in Figure 2.6.

*Method* 1: For jump measures using this method, we simply construct SRVLJ and SRVSJ as done earlier in the paper. Namely, we sort stocks into quintiles based on either weekly SRVLJ or SRVSJ. Then, we construct daily ratios of SRVLJ to SRVSJ for each individual stock in a given quintile. Finally, these ratios are aggregated, forming weekly measures of SRVLJ/SRVSJ. These measures are then used to form equal- or value-weighted ratios of SRVLJ to SRVSJ. These values are depicted in red (solid line) in Figure 2.6.

*Method* 2: For jump measures using this method, we start by constructing the same quintiles (based on weekly SRVLJ and SRVSJ) as done above. Then, we use the 5-minute returns for each stock in a given quintile in order to construct 5-minute aggregate portfolio returns for that quintile. We then construct daily jump measures using these portfolio returns (called SRVLJ and SRVSJ, and SRVLJ/SRVSJ), which are porfolio versions of the similar measures constructed using Method 1. Finally, daily measures are aggregated into weekly measures. These value are depicted in blue (dotted line) in Figure 2.6.

Comparing jump variation ratios constructed in these two different ways allows us to explore the importance of diversification when measuring jump variation. Turning to our findings, Figure 2.6 shows the time series of aggregated (Method 2) and weighted average (Method 1) jump variation measures for the first quintile portfolios. The fact that Method 1 (red line) is much smoother than Method 2 (blue line) means that the small jump component in the ratio of SRVLJ/SRVSJ reamins much larger than in the other case. Thus, the obvious difference between aggregated and weighted averages of SRVLJ/SRVSJ indicates that small jump variation is more likely to be diversified away than large jump variation. This can be immediately seen upon examination of the plots in any of the four panels in the figure. Small jump variation is therefore more closely related to firm specific or idiosyncratic risks, while large jump variation is more likely to be systematic risks. [18]

Another way to explore the relationship between systematic and idiosyncratic risks is to carry out Fama-MacBeth type regressions where the dependent variable is one of our jump variation measures and the independent variables are firm characteristics.[19]

[18]See the Supplementary Appendix for plots of jump variation measures for the other quintile portfolios.

[19]Specifically, our objective in this section is to discuss regressions of the form given in equation (2.15), with the dependent variable replaced by various realized variables.

The results from a number of these sorts of regressions are reported in Table 2.16. Evidently, the firm characteristics always explain more of the dynamics associated with small jumps than with large jumps. This finding is supported by the fact that adjusted $R^2$ are higher when the dependent variable is a small jump variation measure (compare the results of regressions I and II with III and IV). This again suggests that small jump variation is more likely to be associated with idiosyncratic risks.[20]

## 2.5 Concluding Remarks

In this paper, we add to the literature that explores the relationship between equity returns and volatility. In particular, we focus on the strand of this literature that explores the data for evidence of asymmetry (non-linearity) in the return volatility trade-off. Following Bollerslev et al. (2019b), we decompose realized variation into upside and downside semi-variances (good and bad volatilities). We then take the additional step of partitioning the semi-variances into small and large components. Within this context, we examine the marginal predictive content of small and large jump variation measures. We also examine the importance of earnings announcements for examining the linkages between small and large jumps and news.

We find that sorting on signed small jump variation leads to value weighted high-low portfolios with greater average returns and alphas than when either signed total jump or signed large jump variation is sorted on. We also find that there is a threshold, beyond which "large" jump variation contains no marginal predictive ability, relative to that contained in small jump variation. Indeed, including large jump variation can actually be detrimental to predictive accuracy, as average returns and alphas for high-low portfolios actually decline when total variation is instead utilized in some of our prediction experiments. Analysis of returns and alphas based on industry double-sorts

[20]See the Supplementary Appendix for results from double-sorted portfolios that condition on various control variables. In these tables, it is noteworthy that when stocks are first sorted by a control variable (e.g., illiquidity, volatility, firm size and reversal), the SRVJ (SRVLJ and SRVSJ) effect is much higher within quintile portfolios with high illiquidity, high volatility, small firm size, and low reversal. This result suggests that all of these control variables significantly contribute to the predictability of jump variation measures. This result provides additional confirmation to earlier findings reported in Bollerslev et al. (2019b).

indicate that the benefit of small signed jump variation investing is driven by stock selection within an industry, rather than industry bets. Investors prefer stocks with a high probability of large positive jump variation, but they also tend to overweight safer industries. Additionally, we find that the content of signed large jump variation is negligible when controlling for either signed total jump variation or realized skewness. By contrast, signed small jump variation has unique information for predicting future returns, even when controlling for total jump variation or realized skewness. Finally, we find that large jumps are closely associated with “big” news, as might be expected. In particular, large earning announcement surprises increase both the magnitude and occurrence of large jumps. While such news related information is embedded in large jump variation, the information is generally short-lived, and dissipates too quickly to provide marginal predictive content for subsequent weekly returns. Moreover, while large jump variation is closely associated with large earnings surprises (“big” news), small jumps tend to be more closely associated with idiosyncratic risks, and can be diversified away.

Table 2.1: Realized Measures and Firm Characteristics

Panel A: Realized Measures Used in Portfolio Sorts and Fama-MacBeth Regressions

| | |
|---|---|
| RVJP | Positive (upside) jump variation, see (2.1). |
| RVJN | Negative (downside) jump variation, see (2.2). |
| SRVJ | Signed jump variation, $RVJP - RVJN$, see (2.3). |
| RVLJP | Positive (upside) large jump variation, see (2.4). |
| RVLJN | Negative (downside) large jump variation, see (2.5). |
| SRVLJ | Signed large jump variation, $RVLJP - RVLJN$, see (2.6). |
| RVSJP | Positive (upside) small jump variation, see (2.7). |
| RVSJN | Negative (downside) small jump variation, see (2.8). |
| SRVSJ | Signed small jump variation, $RVSJP - RVSJN$, see (2.9). |
| RVOL | Realized volatility |
| RSK | Realized skewness, see (2.12). |
| RKT | Realized kurtosis, see (2.13). |

Panel B: Explanatory Variables and Firm Characteristics Used in Fama-MacBeth Regressions

| | |
|---|---|
| BETA | Market beta |
| log(Size) | Natural logarithm of firm size |
| BEME | Book-to-market ratio |
| MOM | Momentum |
| REV | Short-term reversal |
| IVOL | Idiosyncratic volatility |
| CSK | Coskewness |
| CKT | Cokurtosis |
| MAX | Maximum daily return |
| MIN | Minimum daily return |
| ILLIQ | Illiquidity |

*Notes: The realized measures listed in Panel A of this table are defined and discussed in Section 2.2. For detailed descriptions of the explanatory variables and firm characteristics listed in Panel B of this table, refer to Bollerslev et al. (2019b), and the references cited therein.

# Table 2.2: Summary Statistics for Various Realized Measures and Firm Characteristics Based on Two Jump Truncation Levels

Panel A: Cross-Sectional Summary Statistics

| | SRVJ | RVJP | RVJN | SRVLJ | RVLJP | RVLJN | SRVSJ | RVSJP | RVSJN | RVOL | RSK | RKT | BETA | log(Size) | BEME | MOM | REV | IVOL | CSK | CKT | MAX | MIN | ILLIQ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Part I: Jump Truncation Level= $\gamma^1$ | | | | | | | | | | | | | | | | | | | | | | | |
| Mean | 0.0061 | 0.2698 | 0.2637 | 0.0045 | 0.1518 | 0.1472 | 0.0015 | 0.1180 | 0.1165 | 0.9489 | 0.0288 | 8.2569 | 1.0794 | 6.5280 | 0.5969 | 2023.8456 | 70.6077 | 0.0293 | -0.0263 | 1.1438 | 412.1094 | -346.7608 | -5.2826 |
| Std | 0.1537 | 0.1350 | 0.1347 | 0.1424 | 0.1555 | 0.1523 | 0.0635 | 0.0783 | 0.0783 | 2.1211 | 0.8159 | 4.5706 | 0.5566 | 1.8359 | 0.7224 | 7464.5273 | 927.3551 | 0.0250 | 0.3283 | 0.8474 | 572.1454 | 359.6789 | 2.4047 |
| Part II: Jump Truncation Level= $\gamma^2$ | | | | | | | | | | | | | | | | | | | | | | | |
| Mean | 0.0061 | 0.2698 | 0.2637 | 0.0029 | 0.0983 | 0.0954 | 0.0031 | 0.1715 | 0.1684 | 0.9489 | 0.0288 | 8.2569 | 1.0794 | 6.5280 | 0.5969 | 2023.8456 | 70.6077 | 0.0293 | -0.0263 | 1.1438 | 412.1094 | -346.7608 | -5.2826 |
| Std | 0.1537 | 0.1350 | 0.1347 | 0.1303 | 0.1401 | 0.1368 | 0.0859 | 0.0911 | 0.0909 | 2.1211 | 0.8159 | 4.5706 | 0.5566 | 1.8359 | 0.7224 | 7464.5273 | 927.3551 | 0.0250 | 0.3283 | 0.8474 | 572.1454 | 359.6789 | 2.4047 |

Panel B: Cross-Sectional Correlations

| | SRVJ | RVJP | RVJN | SRVLJ | RVLJP | RVLJN | SRVSJ | RVSJP | RVSJN | RVOL | RSK | RKT | BETA | log(Size) | BEME | MOM | REV | IVOL | CSK | CKT | MAX | MIN | ILLIQ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Part I: Jump Truncation Level= $\gamma^1$ | | | | | | | | | | | | | | | | | | | | | | | |
| SRVJ | 1.00 | 0.57 | -0.57 | 0.91 | 0.43 | -0.40 | 0.37 | 0.13 | -0.18 | -0.02 | 0.94 | 0.03 | -0.03 | 0.01 | 0.01 | 0.01 | 0.30 | -0.03 | 0.09 | 0.00 | 0.17 | 0.22 | 0.00 |
| RVJP | | 1.00 | 0.33 | 0.52 | 0.85 | 0.37 | 0.21 | 0.04 | -0.13 | 0.22 | 0.54 | 0.45 | -0.26 | -0.49 | 0.14 | -0.10 | 0.15 | 0.12 | 0.04 | -0.24 | 0.15 | 0.06 | 0.56 |
| RVJN | | | 1.00 | -0.52 | 0.35 | 0.84 | -0.22 | -0.10 | 0.08 | 0.24 | -0.54 | 0.41 | -0.23 | -0.49 | 0.13 | -0.11 | -0.19 | 0.15 | -0.06 | -0.24 | -0.05 | -0.20 | 0.55 |
| SRVLJ | | | | 1.00 | 0.48 | -0.44 | -0.04 | -0.05 | -0.01 | -0.01 | 0.92 | 0.03 | -0.02 | 0.00 | 0.01 | 0.00 | 0.20 | -0.02 | 0.05 | 0.00 | 0.12 | 0.16 | 0.00 |
| RVLJP | | | | | 1.00 | 0.57 | -0.02 | -0.46 | -0.45 | 0.23 | 0.44 | 0.61 | -0.25 | -0.47 | 0.12 | -0.06 | 0.09 | 0.13 | 0.02 | -0.24 | 0.12 | -0.02 | 0.54 |
| RVLJN | | | | | | 1.00 | 0.01 | -0.44 | -0.45 | 0.24 | -0.41 | 0.59 | -0.23 | -0.48 | 0.11 | -0.06 | -0.10 | 0.15 | -0.03 | -0.24 | 0.01 | -0.17 | 0.54 |
| SRVSJ | | | | | | | 1.00 | 0.42 | -0.42 | -0.02 | 0.19 | 0.00 | -0.03 | 0.01 | 0.01 | 0.01 | 0.26 | -0.03 | 0.08 | 0.00 | 0.14 | 0.18 | 0.00 |
| RVSJP | | | | | | | | 1.00 | 0.64 | -0.04 | 0.06 | -0.40 | 0.06 | 0.03 | 0.01 | -0.05 | 0.10 | -0.03 | 0.04 | 0.04 | 0.02 | 0.11 | -0.06 |
| RVSJN | | | | | | | | | 1.00 | -0.03 | -0.10 | -0.40 | 0.08 | 0.02 | 0.01 | -0.06 | -0.12 | -0.01 | -0.03 | 0.03 | -0.09 | -0.04 | -0.06 |
| RVOL | | | | | | | | | | 1.00 | -0.01 | 0.22 | -0.05 | -0.55 | 0.08 | -0.12 | 0.06 | 0.56 | -0.01 | -0.27 | 0.44 | -0.47 | 0.56 |
| RSK | | | | | | | | | | | 1.00 | 0.04 | -0.02 | 0.00 | 0.01 | 0.00 | 0.22 | -0.02 | 0.06 | 0.00 | 0.13 | 0.17 | 0.00 |
| RKT | | | | | | | | | | | | 1.00 | -0.20 | -0.34 | 0.09 | -0.02 | 0.00 | 0.10 | -0.01 | -0.19 | 0.08 | -0.10 | 0.40 |
| BETA | | | | | | | | | | | | | 1.00 | 0.10 | -0.09 | 0.00 | -0.04 | 0.06 | 0.01 | 0.30 | 0.03 | -0.09 | -0.16 |
| ME | | | | | | | | | | | | | | 1.00 | -0.19 | 0.11 | -0.05 | -0.52 | 0.01 | 0.40 | -0.32 | 0.35 | -0.93 |
| BEME | | | | | | | | | | | | | | | 1.00 | 0.03 | 0.02 | 0.05 | 0.00 | -0.06 | 0.05 | -0.03 | 0.18 |
| MOM | | | | | | | | | | | | | | | | 1.00 | 0.00 | -0.08 | -0.07 | 0.06 | -0.05 | 0.05 | -0.15 |
| REV | | | | | | | | | | | | | | | | | 1.00 | 0.12 | 0.16 | -0.04 | 0.49 | 0.29 | 0.05 |
| IVOL | | | | | | | | | | | | | | | | | | 1.00 | 0.02 | -0.35 | 0.50 | -0.47 | 0.47 |
| CSK | | | | | | | | | | | | | | | | | | | 1.00 | 0.01 | 0.07 | 0.07 | 0.00 |
| CKT | | | | | | | | | | | | | | | | | | | | 1.00 | -0.16 | 0.15 | -0.37 |
| MAX | | | | | | | | | | | | | | | | | | | | | 1.00 | -0.28 | 0.34 |
| MIN | | | | | | | | | | | | | | | | | | | | | | 1.00 | -0.35 |
| ILLIQ | | | | | | | | | | | | | | | | | | | | | | | 1.00 |

| | SRVJ | RVJP | RVJN | SRVLJ | RVLJP | RVLJN | SRVSJ | RVSJP | RVSJN | RVOL | RSK | RKT | BETA | log(Size) | BEME | MOM | REV | IVOL | CSK | CKT | MAX | MIN | ILLIQ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Part II: Jump Truncation Level= $\gamma^2$ | | | | | | | | | | | | | | | | | | | | | | | |
| SRVJ | 1.00 | 0.57 | -0.57 | 0.83 | 0.40 | -0.37 | 0.52 | 0.23 | -0.27 | -0.02 | 0.94 | 0.03 | -0.03 | 0.01 | 0.01 | 0.01 | 0.30 | -0.03 | 0.09 | 0.00 | 0.17 | 0.22 | 0.00 |
| RVJP | | 1.00 | 0.33 | 0.48 | 0.77 | 0.33 | 0.30 | 0.30 | 0.01 | 0.22 | 0.54 | 0.45 | -0.26 | -0.49 | 0.14 | -0.10 | 0.15 | 0.12 | 0.04 | -0.24 | 0.15 | 0.06 | 0.56 |
| RVJN | | | 1.00 | -0.47 | 0.31 | 0.75 | -0.31 | 0.03 | 0.33 | 0.24 | -0.54 | 0.41 | -0.23 | -0.49 | 0.13 | -0.11 | -0.19 | 0.15 | -0.06 | -0.24 | -0.05 | -0.20 | 0.55 |
| SRVLJ | | | | 1.00 | 0.49 | -0.44 | -0.04 | -0.05 | -0.01 | -0.01 | 0.89 | 0.04 | -0.01 | 0.00 | 0.01 | 0.00 | 0.16 | -0.02 | 0.04 | 0.00 | 0.09 | 0.13 | 0.00 |
| RVLJP | | | | | 1.00 | 0.56 | -0.02 | -0.36 | -0.34 | 0.20 | 0.43 | 0.64 | -0.24 | -0.40 | 0.11 | -0.05 | 0.06 | 0.11 | 0.01 | -0.22 | 0.10 | -0.02 | 0.47 |
| RVLJN | | | | | | 1.00 | 0.01 | -0.34 | -0.35 | 0.21 | -0.40 | 0.62 | -0.23 | -0.41 | 0.10 | -0.06 | -0.09 | 0.13 | -0.03 | -0.22 | 0.01 | -0.14 | 0.47 |
| SRVSJ | | | | | | | 1.00 | 0.47 | -0.47 | -0.02 | 0.32 | 0.00 | -0.03 | 0.01 | 0.01 | 0.01 | 0.30 | -0.03 | 0.09 | 0.00 | 0.16 | 0.21 | 0.00 |
| RVSJP | | | | | | | | 1.00 | 0.55 | 0.03 | 0.13 | -0.30 | -0.02 | -0.14 | 0.04 | -0.07 | 0.13 | 0.02 | 0.04 | -0.04 | 0.07 | 0.10 | 0.13 |
| RVSJN | | | | | | | | | 1.00 | 0.05 | -0.17 | -0.30 | 0.01 | -0.15 | 0.04 | -0.07 | -0.15 | 0.05 | -0.04 | -0.04 | -0.08 | -0.09 | 0.13 |
| RVOL | | | | | | | | | | 1.00 | -0.01 | 0.22 | -0.05 | -0.55 | 0.08 | -0.12 | 0.06 | 0.56 | -0.01 | -0.27 | 0.44 | -0.47 | 0.56 |
| RSK | | | | | | | | | | | 1.00 | 0.04 | -0.02 | 0.00 | 0.01 | 0.00 | 0.22 | -0.02 | 0.06 | 0.00 | 0.13 | 0.17 | 0.00 |
| RKT | | | | | | | | | | | | 1.00 | -0.20 | -0.34 | 0.09 | -0.02 | 0.00 | 0.10 | -0.01 | -0.19 | 0.08 | -0.10 | 0.40 |
| BETA | | | | | | | | | | | | | 1.00 | 0.10 | -0.09 | 0.00 | -0.04 | 0.06 | 0.01 | 0.30 | 0.03 | -0.09 | -0.16 |
| ME | | | | | | | | | | | | | | 1.00 | -0.19 | 0.11 | -0.05 | -0.52 | 0.01 | 0.40 | -0.32 | 0.35 | -0.93 |
| BEME | | | | | | | | | | | | | | | 1.00 | 0.03 | 0.02 | 0.05 | 0.00 | -0.06 | 0.05 | -0.03 | 0.18 |
| MOM | | | | | | | | | | | | | | | | 1.00 | 0.00 | -0.08 | -0.07 | 0.06 | -0.05 | 0.05 | -0.15 |
| REV | | | | | | | | | | | | | | | | | 1.00 | 0.12 | 0.16 | -0.04 | 0.49 | 0.29 | 0.05 |
| IVOL | | | | | | | | | | | | | | | | | | 1.00 | 0.02 | -0.35 | 0.50 | -0.47 | 0.47 |
| CSK | | | | | | | | | | | | | | | | | | | 1.00 | 0.01 | 0.07 | 0.07 | 0.00 |
| CKT | | | | | | | | | | | | | | | | | | | | 1.00 | -0.16 | 0.15 | -0.37 |
| MAX | | | | | | | | | | | | | | | | | | | | | 1.00 | -0.28 | 0.34 |
| MIN | | | | | | | | | | | | | | | | | | | | | | 1.00 | -0.35 |
| ILLIQ | | | | | | | | | | | | | | | | | | | | | | | 1.00 |

*Notes: See notes to Table 2.1. This table presents cross-sectional summary statistics and correlations for all realized measures and control variables based on two truncation levels: $\gamma^1 = 4\sqrt{\frac{1}{t}\widehat{IV}_t^{(i)}}\Delta_n^{0.49}$ and $\gamma^2 = 5\sqrt{\frac{1}{t}\widehat{IV}_t^{(i)}}\Delta_n^{0.49}$. The entries in the table for realized measures (see columns 2-13) are constructed using 5-min intraday high frequency data. Entries for firm characteristics (see columns 14-24) are constructed using daily data, with the exception of BEME, which is constructed using monthly data. For complete details, see Sections 2.3 and 2.4

# Table 2.3: Realized Measures and Firm Characteristics of Portfolios Sorted by Various Realized Measures

Panel A: Stocks Sorted by SRVJ

| Quintile | RVJP | RVJN | RVLJP | RVLJN | RVSJP | RVSJN | SRVLJ | SRVSJ | SRVJ | RVOL | RSK | RKT | BETA | log(Size) | BEME | MOM | REV | IVOL | CSK | CKT | MAX | MIN | ILLIQ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0.2021 | 0.3959 | 0.1161 | 0.2723 | 0.0860 | 0.1235 | -0.1563 | -0.0375 | -0.1938 | 0.9394 | -0.9324 | 9.8720 | 1.0369 | 6.1326 | 0.6235 | 0.2006 | -0.0363 | 0.0317 | -0.0708 | 1.0482 | 0.0295 | -0.0494 | -4.6903 |
| 2 | 0.2200 | 0.2777 | 0.1015 | 0.1391 | 0.1185 | 0.1385 | -0.0376 | -0.0200 | -0.0576 | 0.9513 | -0.2504 | 7.1729 | 1.1351 | 6.7041 | 0.5711 | 0.2051 | -0.0111 | 0.0293 | -0.0443 | 1.1987 | 0.0346 | -0.0383 | -5.5835 |
| 3 | 0.2435 | 0.2399 | 0.1127 | 0.1103 | 0.1308 | 0.1296 | 0.0023 | 0.0012 | 0.0036 | 1.0360 | 0.0194 | 6.9162 | 1.1301 | 6.8171 | 0.5695 | 0.2030 | 0.0077 | 0.0283 | -0.0242 | 1.2222 | 0.0397 | -0.0328 | -5.7301 |
| 4 | 0.2801 | 0.2138 | 0.1441 | 0.1005 | 0.1360 | 0.1132 | 0.0436 | 0.0227 | 0.0663 | 0.9162 | 0.2954 | 7.2497 | 1.1096 | 6.7855 | 0.5778 | 0.2097 | 0.0266 | 0.0278 | -0.0066 | 1.2055 | 0.0454 | -0.0282 | -5.6668 |
| 5 | 0.4035 | 0.1914 | 0.2846 | 0.1138 | 0.1189 | 0.0775 | 0.1708 | 0.0413 | 0.2121 | 0.9018 | 1.0138 | 10.0739 | 0.9851 | 6.2007 | 0.6427 | 0.1936 | 0.0485 | 0.0293 | 0.0145 | 1.0443 | 0.0569 | -0.0247 | -4.7419 |

Panel B: Stocks Sorted by SRVLJ

| Quintile | RVJP | RVJN | RVLJP | RVLJN | RVSJP | RVSJN | SRVLJ | SRVSJ | SRVJ | RVOL | RSK | RKT | BETA | log(Size) | BEME | MOM | REV | IVOL | CSK | CKT | MAX | MIN | ILLIQ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Part I: Jump Truncation Level= $\gamma^1$ | | | | | | | | | | | | | | | | | | | | | | | |
| 1 | 0.2255 | 0.3959 | 0.1250 | 0.3016 | 0.1005 | 0.0943 | -0.1766 | 0.0062 | -0.1703 | 0.9429 | -0.9075 | 10.2967 | 1.0163 | 6.0403 | 0.6329 | 0.1966 | -0.0221 | 0.0316 | -0.0548 | 1.0357 | 0.0347 | -0.0461 | -4.5515 |
| 2 | 0.2222 | 0.2603 | 0.0874 | 0.1297 | 0.1347 | 0.1307 | -0.0422 | 0.0041 | -0.0382 | 0.9128 | -0.2003 | 6.9378 | 1.1314 | 6.7639 | 0.5678 | 0.2104 | -0.0011 | 0.0288 | -0.0329 | 1.2050 | 0.0374 | -0.0357 | -5.6652 |
| 3 | 0.2296 | 0.2252 | 0.0968 | 0.0936 | 0.1329 | 0.1316 | 0.0032 | 0.0013 | 0.0044 | 1.1510 | 0.0218 | 6.5571 | 1.1308 | 6.9059 | 0.5648 | 0.1997 | 0.0069 | 0.0282 | -0.0245 | 1.2380 | 0.0390 | -0.0326 | -5.8581 |
| 4 | 0.2726 | 0.2229 | 0.1493 | 0.0981 | 0.1233 | 0.1248 | 0.0512 | -0.0016 | 0.0497 | 0.8314 | 0.2602 | 7.2170 | 1.1237 | 6.7993 | 0.5764 | 0.2118 | 0.0169 | 0.0280 | -0.0166 | 1.2086 | 0.0421 | -0.0304 | -5.6962 |
| 5 | 0.4041 | 0.2153 | 0.3149 | 0.1234 | 0.0891 | 0.0920 | 0.1915 | -0.0028 | 0.1887 | 0.9158 | 0.9900 | 10.5179 | 0.9873 | 6.0946 | 0.6471 | 0.1887 | 0.0354 | 0.0299 | -0.0021 | 1.0376 | 0.0533 | -0.0284 | -4.6003 |
| Part II: Jump Truncation Level= $\gamma^2$ | | | | | | | | | | | | | | | | | | | | | | | |
| 1 | 0.2412 | 0.3897 | 0.0841 | 0.2403 | 0.1571 | 0.1494 | -0.1561 | 0.0076 | -0.1485 | 0.9355 | -0.8465 | 10.5599 | 1.0128 | 6.0603 | 0.6360 | 0.1928 | -0.0152 | 0.0313 | -0.0478 | 1.0403 | 0.0366 | -0.0438 | -4.5663 |
| 2 | 0.2189 | 0.2314 | 0.0341 | 0.0519 | 0.1848 | 0.1794 | -0.0178 | 0.0054 | -0.0124 | 0.9321 | -0.0762 | 6.4172 | 1.1366 | 6.8798 | 0.5622 | 0.2132 | 0.0045 | 0.0283 | -0.0278 | 1.2310 | 0.0388 | -0.0340 | -5.8322 |
| 3 | 0.2438 | 0.2405 | 0.0805 | 0.0770 | 0.1633 | 0.1635 | 0.0035 | -0.0001 | 0.0033 | 1.6847 | 0.0202 | 7.7437 | 1.0936 | 6.5831 | 0.6248 | 0.1664 | 0.0063 | 0.0322 | -0.0021 | 1.2281 | 0.0426 | -0.0360 | -5.3661 |
| 4 | 0.2985 | 0.2611 | 0.1233 | 0.0854 | 0.1752 | 0.1756 | 0.0379 | -0.0004 | 0.0375 | 1.0724 | 0.2140 | 8.0742 | 1.0912 | 6.4686 | 0.5995 | 0.1944 | 0.0125 | 0.0294 | -0.0213 | 1.1478 | 0.0418 | -0.0324 | -5.2346 |
| 5 | 0.3985 | 0.2321 | 0.2526 | 0.0837 | 0.1459 | 0.1484 | 0.1690 | -0.0025 | 0.1664 | 0.9201 | 0.9332 | 10.8242 | 0.9935 | 6.1033 | 0.6466 | 0.1896 | 0.0285 | 0.0300 | -0.0084 | 1.0418 | 0.0506 | -0.0301 | -4.6103 |

Panel C: Stocks Sorted by SRVSJ

| Quintile | RVJP | RVJN | RVLJP | RVLJN | RVSJP | RVSJN | SRVLJ | SRVSJ | SRVJ | RVOL | RSK | RKT | BETA | log(Size) | BEME | MOM | REV | IVOL | CSK | CKT | MAX | MIN | ILLIQ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Part I: Jump Truncation Level= $\gamma^1$ | | | | | | | | | | | | | | | | | | | | | | | |
| 1 | 0.2060 | 0.2784 | 0.1085 | 0.0968 | 0.0975 | 0.1815 | 0.0116 | -0.0840 | -0.0724 | 0.7622 | -0.1832 | 7.1817 | 1.1514 | 6.7378 | 0.5654 | 0.2183 | -0.0270 | 0.0292 | -0.0637 | 1.2017 | 0.0302 | -0.0438 | -5.6348 |
| 2 | 0.2711 | 0.2845 | 0.1758 | 0.1649 | 0.0952 | 0.1197 | 0.0110 | -0.0244 | -0.0135 | 1.0260 | -0.0157 | 8.6959 | 1.0715 | 6.4139 | 0.6062 | 0.1933 | -0.0046 | 0.0302 | -0.0417 | 1.1242 | 0.0380 | -0.0375 | -5.1168 |
| 3 | 0.3077 | 0.3003 | 0.2186 | 0.2126 | 0.0891 | 0.0877 | 0.0060 | 0.0014 | 0.0074 | 1.2309 | 0.0345 | 9.7978 | 1.0053 | 6.1262 | 0.6382 | 0.1754 | 0.0066 | 0.0309 | -0.0275 | 1.0469 | 0.0423 | -0.0356 | -4.6619 |
| 4 | 0.2743 | 0.2492 | 0.1413 | 0.1438 | 0.1330 | 0.1053 | -0.0026 | 0.0277 | 0.0251 | 0.9562 | 0.0709 | 8.0830 | 1.0906 | 6.6614 | 0.5854 | 0.2098 | 0.0190 | 0.0283 | -0.0099 | 1.1752 | 0.0436 | -0.0305 | -5.4837 |
| 5 | 0.2830 | 0.1981 | 0.1004 | 0.1051 | 0.1826 | 0.0930 | -0.0047 | 0.0896 | 0.0849 | 0.7380 | 0.2395 | 7.1974 | 1.0947 | 6.8064 | 0.5783 | 0.2217 | 0.0428 | 0.0272 | 0.0136 | 1.1959 | 0.0517 | -0.0251 | -5.6801 |
| Part II: Jump Truncation Level= $\gamma^2$ | | | | | | | | | | | | | | | | | | | | | | | |
| 1 | 0.2058 | 0.3081 | 0.0749 | 0.0650 | 0.1309 | 0.2431 | 0.0099 | -0.1121 | -0.1023 | 0.8104 | -0.3171 | 7.6436 | 1.1207 | 6.5435 | 0.5792 | 0.2179 | -0.0318 | 0.0301 | -0.0682 | 1.1558 | 0.0296 | -0.0461 | -5.3493 |
| 2 | 0.2561 | 0.2841 | 0.1103 | 0.1027 | 0.1458 | 0.1814 | 0.0077 | -0.0356 | -0.0280 | 1.0119 | -0.0747 | 8.4222 | 1.0918 | 6.5103 | 0.5957 | 0.1964 | -0.0079 | 0.0298 | -0.0427 | 1.1482 | 0.0363 | -0.0379 | -5.2661 |
| 3 | 0.2906 | 0.2847 | 0.1423 | 0.1381 | 0.1484 | 0.1465 | 0.0041 | 0.0019 | 0.0060 | 1.1729 | 0.0308 | 9.3041 | 1.0385 | 6.3412 | 0.6204 | 0.1788 | 0.0065 | 0.0301 | -0.0273 | 1.0955 | 0.0413 | -0.0346 | -4.9812 |
| 4 | 0.2821 | 0.2444 | 0.0950 | 0.0979 | 0.1871 | 0.1465 | -0.0029 | 0.0405 | 0.0377 | 0.9731 | 0.1235 | 8.1884 | 1.0851 | 6.6358 | 0.5919 | 0.2046 | 0.0220 | 0.0284 | -0.0086 | 1.1700 | 0.0447 | -0.0299 | -5.4350 |
| 5 | 0.3137 | 0.1959 | 0.0671 | 0.0713 | 0.2466 | 0.1246 | -0.0043 | 0.1220 | 0.1178 | 0.7734 | 0.3845 | 7.6692 | 1.0627 | 6.6238 | 0.5954 | 0.2153 | 0.0469 | 0.0278 | 0.0158 | 1.1527 | 0.0543 | -0.0245 | -5.4040 |

Panel D: Stocks Sorted by RVOL

| Quintile | RVJP | RVJN | RVLJP | RVLJN | RVSJP | RVSJN | SRVLJ | SRVSJ | SRVJ | RVOL | RSK | RKT | BETA | log(Size) | BEME | MOM | REV | IVOL | CSK | CKT | MAX | MIN | ILLIQ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0.2255 | 0.2140 | 0.1013 | 0.0943 | 0.1241 | 0.1197 | 0.0070 | 0.0044 | 0.0114 | 0.2290 | 0.0485 | 6.8794 | 0.8390 | 8.3393 | 0.5407 | 0.1686 | 0.0040 | 0.0137 | -0.0213 | 1.4325 | 0.0177 | -0.0154 | -7.4961 |
| 2 | 0.2375 | 0.2282 | 0.1136 | 0.1071 | 0.1240 | 0.1211 | 0.0065 | 0.0028 | 0.0094 | 0.3596 | 0.0403 | 7.2835 | 1.0471 | 7.4505 | 0.5506 | 0.1853 | 0.0044 | 0.0187 | -0.0213 | 1.3708 | 0.0257 | -0.0227 | -6.4401 |
| 3 | 0.2567 | 0.2493 | 0.1387 | 0.1327 | 0.1180 | 0.1166 | 0.0060 | 0.0014 | 0.0074 | 0.5331 | 0.0341 | 7.9018 | 1.2246 | 6.6274 | 0.5581 | 0.2598 | 0.0051 | 0.0253 | -0.0249 | 1.2373 | 0.0346 | -0.0304 | -5.5465 |
| 4 | 0.2864 | 0.2823 | 0.1717 | 0.1675 | 0.1147 | 0.1148 | 0.0042 | -0.0001 | 0.0041 | 0.8136 | 0.0216 | 8.7067 | 1.2761 | 5.7649 | 0.5997 | 0.3071 | 0.0058 | 0.0338 | -0.0308 | 1.0491 | 0.0464 | -0.0405 | -4.4310 |
| 5 | 0.3429 | 0.3449 | 0.2336 | 0.2347 | 0.1093 | 0.1102 | -0.0011 | -0.0010 | -0.0021 | 2.8115 | -0.0003 | 10.5156 | 1.0101 | 4.4547 | 0.7443 | 0.0910 | 0.0160 | 0.0548 | -0.0331 | 0.6286 | 0.0817 | -0.0645 | -2.4951 |

Panel E: Stocks Sorted by RSK

| Quintile | RVJP | RVJN | RVLJP | RVLJN | RVSJP | RVSJN | SRVLJ | SRVSJ | SRVJ | RVOL | RSK | RKT | BETA | log(Size) | BEME | MOM | REV | IVOL | CSK | CKT | MAX | MIN | ILLIQ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0.2077 | 0.3914 | 0.1200 | 0.2822 | 0.0877 | 0.1091 | -0.1622 | -0.0215 | -0.1837 | 0.9182 | -0.9829 | 10.3573 | 1.0293 | 6.1615 | 0.6212 | 0.2118 | -0.0274 | 0.0312 | -0.0611 | 1.0555 | 0.0324 | -0.0472 | -4.7275 |
| 2 | 0.2203 | 0.2758 | 0.0967 | 0.1366 | 0.1235 | 0.1393 | -0.0398 | -0.0157 | -0.0556 | 0.9332 | -0.2596 | 6.8740 | 1.1303 | 6.7033 | 0.5719 | 0.2052 | -0.0077 | 0.0291 | -0.0412 | 1.1981 | 0.0353 | -0.0374 | -5.5757 |
| 3 | 0.2430 | 0.2394 | 0.1057 | 0.1035 | 0.1373 | 0.1359 | 0.0023 | 0.0014 | 0.0037 | 1.1222 | 0.0189 | 6.5111 | 1.1288 | 6.8070 | 0.5748 | 0.1932 | 0.0073 | 0.0283 | -0.0238 | 1.2155 | 0.0391 | -0.0325 | -5.7121 |
| 4 | 0.2786 | 0.2142 | 0.1418 | 0.0960 | 0.1368 | 0.1182 | 0.0459 | 0.0186 | 0.0644 | 0.8769 | 0.3040 | 6.9460 | 1.1139 | 6.7671 | 0.5790 | 0.2056 | 0.0231 | 0.0280 | -0.0098 | 1.2017 | 0.0444 | -0.0290 | -5.6422 |
| 5 | 0.3996 | 0.1978 | 0.2947 | 0.1179 | 0.1049 | 0.0800 | 0.1769 | 0.0249 | 0.2018 | 0.8942 | 1.0656 | 10.5964 | 0.9944 | 6.2012 | 0.6374 | 0.1963 | 0.0400 | 0.0297 | 0.0045 | 1.0482 | 0.0549 | -0.0272 | -4.7552 |

Panel F: Stocks Sorted by RKT

| Quintile | RVJP | RVJN | RVLJP | RVLJN | RVSJP | RVSJN | SRVLJ | SRVSJ | SRVJ | RVOL | RSK | RKT | BETA | log(Size) | BEME | MOM | REV | IVOL | CSK | CKT | MAX | MIN | ILLIQ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0.1804 | 0.1785 | 0.0257 | 0.0248 | 0.1548 | 0.1536 | 0.0008 | 0.0011 | 0.0019 | 0.6884 | 0.0110 | 4.4470 | 1.1920 | 7.6130 | 0.5303 | 0.1864 | 0.0054 | 0.0248 | -0.0222 | 1.3529 | 0.0339 | -0.0290 | -6.7592 |
| 2 | 0.2242 | 0.2206 | 0.0757 | 0.0738 | 0.1484 | 0.1468 | 0.0020 | 0.0017 | 0.0036 | 0.7285 | 0.0167 | 5.7679 | 1.1586 | 6.9411 | 0.5522 | 0.2203 | 0.0077 | 0.0276 | -0.0252 | 1.2505 | 0.0387 | -0.0324 | -5.9541 |
| 3 | 0.2630 | 0.2582 | 0.1340 | 0.1310 | 0.1291 | 0.1272 | 0.0030 | 0.0019 | 0.0048 | 0.8130 | 0.0215 | 7.0711 | 1.1103 | 6.5028 | 0.5817 | 0.2175 | 0.0083 | 0.0296 | -0.0265 | 1.1627 | 0.0417 | -0.0347 | -5.3464 |
| 4 | 0.3070 | 0.3004 | 0.2067 | 0.2017 | 0.1004 | 0.0987 | 0.0049 | 0.0017 | 0.0066 | 0.9433 | 0.0292 | 8.9744 | 1.0438 | 6.0841 | 0.6186 | 0.2085 | 0.0076 | 0.0313 | -0.0281 | 1.0644 | 0.0440 | -0.0370 | -4.6922 |
| 5 | 0.3745 | 0.3612 | 0.3171 | 0.3051 | 0.0574 | 0.0561 | 0.0119 | 0.0013 | 0.0133 | 1.5722 | 0.0660 | 15.0322 | 0.8918 | 5.4974 | 0.7022 | 0.1793 | 0.0064 | 0.0332 | -0.0295 | 0.8881 | 0.0478 | -0.0403 | -3.6585 |

*Notes: See notes to Table 2.2. Entries in this table are time series averages of equal-weighted realized measures and firm characteristics of stocks sorted by various realized measures. The sample includes all NYSE, NASDAQ and AMEX listed stocks for the period January 1993 to December 2016. At the end of each Tuesday, all of the stocks in the sample are sorted into quintile portfolios, based on ascending values of various realized measures. The equal-weighted realized measures and firm characteristics of each quintile portfolio are calculated over the same week. Additionally, $\gamma^1 = 4\sqrt{\frac{1}{t}\widehat{IV}_t^{(i)}}\Delta_n^{0.49}$ and $\gamma^2 = 5\sqrt{\frac{1}{t}\widehat{IV}_t^{(i)}}\Delta_n^{0.49}$ are jump truncation levels. See Sections 2.2 and 2.4 for further details.

# Table 2.4: Univariate Portfolio Sorts Based on Positive, Negative, and Signed Total Jump Variation

Panel A: Stocks Sorted by RVJP

| | Equal-Weighted Returns and Alphas | | | | | | Value-Weighted Returns and Alphas | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Quintile | 1(Low) | 2 | 3 | 4 | 5(High) | High-Low | 1(Low) | 2 | 3 | 4 | 5(High) | High-Low |
| Mean Return | 33.65 | 30.50 | 33.04 | 28.08 | 20.64 | -13.01*** | 23.52 | 19.48 | 17.93 | 20.91 | 18.27 | -5.25 |
| | (3.54) | (3.28) | (3.49) | (2.87) | (2.19) | (-2.75) | (3.54) | (3.27) | (2.93) | (3.35) | (2.83) | (-1.35) |
| Alpha | 10.59 | 7.47 | 11.24 | 7.52 | 2.88 | -7.71 | 2.88 | -0.63 | -2.30 | 0.67 | -2.75 | -5.63* |
| | (4.16) | (3.64) | (4.34) | (2.33) | (0.72) | (-1.64) | (2.31) | (-0.44) | (-1.22) | (0.32) | (-1.19) | (-1.87) |

Panel B: Stocks Sorted by RVJN

| | Equal-Weighted Returns and Alphas | | | | | | Value-Weighted Returns and Alphas | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Quintile | 1(Low) | 2 | 3 | 4 | 5(High) | High-Low | 1(Low) | 2 | 3 | 4 | 5(High) | High-Low |
| Mean Return | 13.17 | 23.01 | 26.77 | 33.79 | 49.23 | 36.06*** | 16.23 | 25.44 | 26.41 | 26.62 | 31.36 | 15.13*** |
| | (1.52) | (2.59) | (2.84) | (3.36) | (4.62) | (6.47) | (2.55) | (4.11) | (4.08) | (3.93) | (4.29) | (3.75) |
| Alpha | -9.36 | -0.03 | 4.18 | 13.24 | 31.71 | 41.07*** | -3.55 | 4.94 | 5.27 | 5.49 | 10.05 | 13.60*** |
| | (-4.46) | (-0.02) | (1.86) | (3.93) | (6.34) | (7.51) | (-3.08) | (3.02) | (2.64) | (2.37) | (4.13) | (4.52) |

Panel C: Stocks Sorted by SRVJ

| | Equal-Weighted Returns and Alphas | | | | | | Value-Weighted Returns and Alphas | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Quintile | 1(Low) | 2 | 3 | 4 | 5(High) | High-Low | 1(Low) | 2 | 3 | 4 | 5(High) | High-Low |
| Mean Return | 51.85 | 39.02 | 26.15 | 17.86 | 11.02 | -40.82*** | 34.67 | 27.43 | 19.93 | 13.64 | 9.65 | -25.02*** |
| | (5.14) | (3.85) | (2.70) | (1.98) | (1.33) | (-9.85) | (4.85) | (4.12) | (3.10) | (2.16) | (1.59) | (-5.78) |
| Alpha | 30.54 | 17.81 | 4.56 | -3.58 | -9.64 | -40.18*** | 13.44 | 6.94 | -0.52 | -6.53 | -10.25 | -23.69*** |
| | (8.40) | (5.78) | (1.74) | (-1.56) | (-4.05) | (-10.10) | (5.01) | (3.95) | (-0.40) | (-4.48) | (-4.47) | (-5.56) |

*Notes: Entries in this table are average returns and risk-adjusted alphas for single-sorted portfolios based on RVJP, RVJN and SRVJ, which are described in Table 2.2. The sample includes all NYSE, NASDAQ and AMEX listed stocks for the period January 1993 to December 2016. At the end of each Tuesday, all the stocks in the sample are sorted into quintile portfolios based on ascending values of the various jump variation measures listed in the titel of each panel. Each portfolio is held for one week. The row labeled "Mean Return" reports the time series average values of one-week ahead equal-weighted and value-weighted returns for quintile portfolios. The row labeled "Alpha" reports Fama-French-Carhart four-factor alphas, based on the model (2.14), for each of the quintile portfolios, as well as for the difference between portfolio 5 and portfolio 1. Newey-West $t$-statistics are given in parentheses; and *, **, and *** denote means and alphas that are significant at the 10%, 5%, and 1% levels, respectively.

# Table 2.5: Univariate Portfolio Sorts Based on Positive, Negative, and Signed Large Jump Variation

Panel A: Stocks Sorted by RVLJP

| | Equal-Weighted Returns and Alphas | | | | | | Value-Weighted Returns and Alphas | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Quintile | 1(Low) | 2 | 3 | 4 | 5(High) | High-Low | 1(Low) | 2 | 3 | 4 | 5(High) | High-Low |
| Part I: Jump Truncation Level=$\gamma^1$ | | | | | | | | | | | | |
| Mean Return | 31.40 | 30.00 | 29.17 | 31.04 | 23.78 | -7.61** | 21.78 | 20.61 | 20.72 | 18.56 | 17.75 | -4.03 |
| | (3.31) | (3.19) | (3.10) | (3.28) | (2.48) | (-2.05) | (3.41) | (3.21) | (3.33) | (2.95) | (2.66) | (-1.09) |
| Alpha | 10.07 | 7.64 | 6.65 | 9.04 | 6.08 | -3.98 | 1.97 | 0.68 | -0.42 | -2.32 | -2.83 | -4.80* |
| | (4.20) | (3.44) | (2.81) | (3.35) | (1.53) | (-1.08) | (1.95) | (0.52) | (-0.26) | (-1.27) | (-1.21) | (-1.71) |
| Part II: Jump Truncation Level=$\gamma^2$ | | | | | | | | | | | | |
| Mean Return | 29.42 | 43.86 | 30.00 | 29.12 | 25.77 | -3.65 | 20.42 | 27.59 | 20.12 | 22.50 | 20.39 | -0.03 |
| | (3.15) | (2.16) | (3.04) | (3.07) | (2.74) | (-1.11) | (3.25) | (1.77) | (2.98) | (3.63) | (3.14) | (-0.01) |
| Alpha | 7.74 | 32.93 | 7.28 | 7.02 | 7.70 | -0.04 | 0.45 | 13.80 | -0.27 | 1.60 | -0.52 | -0.97 |
| | (3.70) | (3.23) | (2.92) | (2.62) | (2.02) | (-0.01) | (0.72) | (1.35) | (-0.13) | (0.83) | (-0.22) | (-0.37) |

Panel B: Stocks Sorted by RVLJN

| | Equal-Weighted Returns and Alphas | | | | | | Value-Weighted Returns and Alphas | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Quintile | 1(Low) | 2 | 3 | 4 | 5(High) | High-Low | 1(Low) | 2 | 3 | 4 | 5(High) | High-Low |
| Part I: Jump Truncation Level=$\gamma^1$ | | | | | | | | | | | | |
| Mean Return | 23.80 | 23.68 | 27.64 | 28.85 | 41.56 | 17.76*** | 19.65 | 19.58 | 25.63 | 22.38 | 25.24 | 5.59 |
| | (2.65) | (2.59) | (2.98) | (2.95) | (4.10) | (4.47) | (3.10) | (3.10) | (4.00) | (3.27) | (3.44) | (1.41) |
| Alpha | 2.31 | 1.11 | 5.00 | 7.10 | 23.74 | 21.43*** | -0.20 | -1.32 | 4.52 | 0.89 | 4.34 | 4.54 |
| | (1.11) | (0.55) | (2.32) | (2.47) | (5.39) | (5.53) | (-0.18) | (-0.92) | (2.47) | (0.40) | (1.82) | (1.61) |
| Part II: Jump Truncation Level=$\gamma^2$ | | | | | | | | | | | | |
| Mean Return | 24.76 | 6.46 | 27.89 | 29.12 | 38.62 | 13.86*** | 19.69 | 13.21 | 21.46 | 23.20 | 22.34 | 2.66 |
| | (2.71) | (0.33) | (2.75) | (3.03) | (3.88) | (4.01) | (3.13) | (0.93) | (3.05) | (3.61) | (3.15) | (0.75) |
| Alpha | 3.02 | 6.75 | 6.58 | 7.06 | 20.62 | 17.60*** | -0.30 | 6.52 | 1.26 | 2.28 | 1.58 | 1.88 |
| | (1.63) | (1.08) | (2.50) | (2.67) | (4.81) | (5.06) | (-0.45) | (0.91) | (0.57) | (1.06) | (0.66) | (0.70) |

Panel C: Stocks Sorted by SRVLJ

| | Equal-Weighted Returns and Alphas | | | | | | Value-Weighted Returns and Alphas | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Quintile | 1(Low) | 2 | 3 | 4 | 5(High) | High-Low | 1(Low) | 2 | 3 | 4 | 5(High) | High-Low |
| Part I: Jump Truncation Level=$\gamma^1$ | | | | | | | | | | | | |
| Mean Return | 44.35 | 32.94 | 31.08 | 22.72 | 16.04 | -28.31*** | 26.27 | 22.99 | 22.36 | 17.77 | 16.27 | -10.01*** |
| | (4.52) | (3.36) | (3.13) | (2.44) | (1.88) | (-9.00) | (3.92) | (3.58) | (3.29) | (2.79) | (2.71) | (-3.09) |
| Alpha | 23.47 | 11.38 | 8.91 | 0.96 | -4.90 | -28.36*** | 5.00 | 2.42 | 1.83 | -2.64 | -4.26 | -9.25*** |
| | (7.13) | (4.20) | (3.04) | (0.40) | (-2.17) | (-9.39) | (2.24) | (1.64) | (1.01) | (-1.82) | (-2.24) | (-2.87) |
| Part II: Jump Truncation Level=$\gamma^2$ | | | | | | | | | | | | |
| Mean Return | 40.55 | 28.37 | 33.05 | 24.16 | 19.03 | -21.52*** | 22.59 | 20.48 | 16.45 | 18.86 | 20.14 | -2.45 |
| | (4.19) | (2.91) | (1.48) | (2.55) | (2.19) | (-8.22) | (3.40) | (3.18) | (1.14) | (3.02) | (3.27) | (-0.80) |
| Alpha | 19.59 | 8.18 | 24.26 | 2.23 | -1.97 | -21.55*** | 1.82 | 0.76 | 6.85 | -2.41 | -0.26 | -2.08 |
| | (6.15) | (3.29) | (2.16) | (0.82) | (-0.84) | (-8.33) | (0.86) | (0.68) | (1.04) | (-1.23) | (-0.13) | (-0.69) |

*Notes: See notes to Table 2.4. Entries are average returns and risk-adjusted alphas for single-sorted portfolios based on RVLJP, RVLJN and SRVLJ. Jump truncation levels are $\gamma^1 = 4\sqrt{\frac{1}{t}\widehat{IV}_t^{(i)}}\Delta_n^{0.49}$ and $\gamma^2 = 5\sqrt{\frac{1}{t}\widehat{IV}_t^{(i)}}\Delta_n^{0.49}$ .

# Table 2.6: Univariate Portfolio Sorts Based on Positive, Negative, and Signed Small Jump Variation

Panel A: Stocks Sorted by RVSJP

| | Equal-Weighted Returns and Alphas | | | | | | Value-Weighted Returns and Alphas | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Quintile | 1(Low) | 2 | 3 | 4 | 5(High) | High-Low | 1(Low) | 2 | 3 | 4 | 5(High) | High-Low |
| Part I: Jump Truncation Level=$\gamma^1$ | | | | | | | | | | | | |
| Mean Return | 32.01 | 32.51 | 29.03 | 26.59 | 25.72 | -6.29** | 27.93 | 26.10 | 18.44 | 16.52 | 15.37 | -12.55** |
| | (3.44) | (3.43) | (3.09) | (2.85) | (2.65) | (-2.14) | (3.65) | (3.84) | (2.91) | (2.75) | (2.45) | (-2.54) |
| Alpha | 13.11 | 9.39 | 6.37 | 4.77 | 6.13 | -6.98*** | 7.40 | 5.21 | -1.83 | -2.89 | -4.95 | -12.35*** |
| | (3.74) | (4.21) | (3.02) | (2.08) | (1.82) | (-2.65) | (2.44) | (3.28) | (-1.39) | (-1.55) | (-2.32) | (-2.93) |
| Part II: Jump Truncation Level=$\gamma^2$ | | | | | | | | | | | | |
| Mean Return | 34.25 | 31.55 | 28.43 | 27.41 | 24.23 | -10.02*** | 29.37 | 18.76 | 17.76 | 14.75 | 18.98 | -10.40** |
| | (3.72) | (3.37) | (3.04) | (2.90) | (2.48) | (-3.26) | (4.09) | (2.95) | (2.94) | (2.39) | (2.93) | (-2.25) |
| Alpha | 14.08 | 8.85 | 5.90 | 5.75 | 5.08 | -9.00*** | 8.52 | -2.00 | -2.01 | -5.30 | -1.51 | -10.02** |
| | (4.93) | (4.12) | (2.73) | (2.32) | (1.42) | (-3.13) | (3.87) | (-1.52) | (-1.24) | (-2.67) | (-0.62) | (-2.54) |

Panel B: Stocks Sorted by RVSJN

| | Equal-Weighted Returns and Alphas | | | | | | Value-Weighted Returns and Alphas | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Quintile | 1(Low) | 2 | 3 | 4 | 5(High) | High-Low | 1(Low) | 2 | 3 | 4 | 5(High) | High-Low |
| Part I: Jump Truncation Level=$\gamma^1$ | | | | | | | | | | | | |
| Mean Return | 23.66 | 21.60 | 27.21 | 31.73 | 41.77 | 18.10*** | 6.94 | 15.90 | 21.97 | 27.77 | 31.26 | 24.32*** |
| | (2.64) | (2.38) | (2.97) | (3.35) | (3.93) | (5.07) | (1.00) | (2.42) | (3.46) | (4.34) | (4.68) | (5.00) |
| Alpha | 5.26 | -1.04 | 4.35 | 9.56 | 21.62 | 16.36*** | -13.39 | -4.28 | 1.80 | 7.41 | 10.52 | 23.91*** |
| | (1.59) | (-0.51) | (2.22) | (4.02) | (5.49) | (5.46) | (-4.54) | (-2.80) | (1.33) | (3.72) | (4.07) | (5.21) |
| Part II: Jump Truncation Level=$\gamma^2$ | | | | | | | | | | | | |
| Mean Return | 19.42 | 23.04 | 26.65 | 32.48 | 44.37 | 24.96*** | 14.22 | 18.78 | 25.82 | 29.05 | 32.60 | 18.38*** |
| | (2.23) | (2.60) | (2.93) | (3.35) | (4.09) | (6.07) | (2.13) | (3.00) | (4.13) | (4.42) | (4.43) | (3.80) |
| Alpha | -0.37 | 0.70 | 4.18 | 10.38 | 24.84 | 25.22*** | -5.47 | -1.28 | 5.81 | 7.80 | 10.85 | 16.31*** |
| | (-0.14) | (0.37) | (2.18) | (4.02) | (5.85) | (7.21) | (-2.72) | (-1.07) | (3.39) | (3.67) | (4.02) | (4.15) |

Panel C: Stocks Sorted by SRVSJ

| | Equal-Weighted Returns and Alphas | | | | | | Value-Weighted Returns and Alphas | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Quintile | 1(Low) | 2 | 3 | 4 | 5(High) | High-Low | 1(Low) | 2 | 3 | 4 | 5(High) | High-Low |
| Part I: Jump Truncation Level=$\gamma^1$ | | | | | | | | | | | | |
| Mean Return | 46.41 | 40.51 | 25.12 | 19.00 | 13.74 | -32.67*** | 34.54 | 24.08 | 18.26 | 17.84 | 10.42 | -24.12*** |
| | (4.57) | (4.04) | (2.67) | (2.06) | (1.62) | (-8.60) | (5.00) | (3.58) | (2.81) | (2.85) | (1.72) | (-6.60) |
| Alpha | 23.64 | 19.68 | 5.23 | -2.09 | -8.14 | -31.78*** | 13.77 | 3.25 | -2.37 | -2.39 | -9.27 | -23.04*** |
| | (7.62) | (6.10) | (1.53) | (-0.85) | (-4.17) | (-9.01) | (6.18) | (1.91) | (-1.07) | (-1.52) | (-5.00) | (-6.54) |
| Part II: Jump Truncation Level=$\gamma^2$ | | | | | | | | | | | | |
| Mean Return | 47.90 | 41.62 | 27.23 | 17.90 | 11.20 | -36.70*** | 36.88 | 25.13 | 18.45 | 14.98 | 9.41 | -27.47*** |
| | (4.70) | (4.13) | (2.87) | (1.98) | (1.34) | (-9.06) | (5.31) | (3.79) | (2.86) | (2.39) | (1.52) | (-6.94) |
| Alpha | 25.37 | 20.76 | 6.99 | -3.03 | -10.51 | -35.88*** | 16.07 | 4.52 | -1.87 | -5.39 | -10.34 | -26.41*** |
| | (7.79) | (6.65) | (2.23) | (-1.25) | (-5.21) | (-9.49) | (6.72) | (2.60) | (-1.23) | (-3.30) | (-5.00) | (-6.72) |

*Notes: See notes to Table 2.5.

# Table 2.7: Univariate Portfolio Sorts Based on Realized Volatility, Skewness, Kurtosis and Continuous Variance

Panel A: Stocks Sorted by RVOL

| | Equal-Weighted Returns and Alphas | | | | | | Value-Weighted Returns and Alphas | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Quintile | 1(Low) | 2 | 3 | 4 | 5(High) | High-Low | 1(Low) | 2 | 3 | 4 | 5(High) | High-Low |
| Mean Return | 23.36 | 28.00 | 28.89 | 31.78 | 33.91 | 10.55 | 20.72 | 21.42 | 19.75 | 26.78 | 29.19 | 8.47 |
| | (4.47) | (3.92) | (2.96) | (2.59) | (2.24) | (0.81) | (4.09) | (2.83) | (1.84) | (1.98) | (1.92) | (0.64) |
| Alpha | 4.50 | 5.01 | 5.15 | 8.74 | 16.33 | 11.83 | 1.95 | -1.39 | -3.94 | 2.44 | 5.44 | 3.49 |
| | (2.07) | (2.94) | (2.57) | (2.54) | (2.11) | (1.37) | (1.35) | (-0.67) | (-1.01) | (0.43) | (0.67) | (0.40) |

Panel B: Stocks Sorted by RSK

| | Equal-Weighted Returns and Alphas | | | | | | Value-Weighted Returns and Alphas | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Quintile | 1(Low) | 2 | 3 | 4 | 5(High) | High-Low | 1(Low) | 2 | 3 | 4 | 5(High) | High-Low |
| Mean Return | 47.56 | 38.06 | 27.86 | 19.44 | 12.98 | -34.58*** | 29.45 | 27.52 | 19.27 | 14.68 | 13.21 | -16.24*** |
| | (4.85) | (3.82) | (2.86) | (2.12) | (1.54) | (-9.94) | (4.27) | (4.22) | (2.98) | (2.32) | (2.18) | (-4.29) |
| Alpha | 26.22 | 16.77 | 6.73 | -2.15 | -7.90 | -34.12*** | 7.87 | 7.02 | -0.82 | -5.38 | -6.77 | -14.64*** |
| | (7.93) | (5.66) | (2.41) | (-0.96) | (-3.51) | (-10.08) | (3.30) | (4.44) | (-0.60) | (-3.73) | (-3.23) | (-3.85) |

Panel C: Stocks Sorted by RKT

| | Equal-Weighted Returns and Alphas | | | | | | Value-Weighted Returns and Alphas | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Quintile | 1(Low) | 2 | 3 | 4 | 5(High) | High-Low | 1(Low) | 2 | 3 | 4 | 5(High) | High-Low |
| Mean Return | 28.95 | 28.59 | 29.91 | 29.42 | 29.07 | 0.12 | 19.87 | 21.57 | 21.12 | 22.17 | 19.55 | -0.32 |
| | (3.07) | (3.00) | (3.13) | (3.06) | (3.24) | (0.04) | (3.12) | (3.42) | (3.37) | (3.38) | (2.96) | (-0.10) |
| Alpha | 8.55 | 6.42 | 7.47 | 7.92 | 9.36 | 0.81 | 0.21 | 0.65 | -0.10 | 0.94 | -1.92 | -2.13 |
| | (3.21) | (2.87) | (3.05) | (2.87) | (3.07) | (0.28) | (0.20) | (0.46) | (-0.06) | (0.49) | (-0.91) | (-0.81) |

Panel D: Stocks Sorted by RVC

| | Equal-Weighted Returns and Alphas | | | | | | Value-Weighted Returns and Alphas | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Quintile | 1(Low) | 2 | 3 | 4 | 5(High) | High-Low | 1(Low) | 2 | 3 | 4 | 5(High) | High-Low |
| Mean Return | 36.18 | 31.82 | 28.23 | 27.47 | 22.21 | -13.97** | 24.36 | 24.80 | 22.94 | 23.40 | 20.41 | -3.95 |
| | (3.54) | (3.22) | (3.00) | (3.05) | (2.42) | (-2.58) | (3.53) | (3.65) | (3.59) | (3.87) | (3.20) | (-1.00) |
| Alpha | 19.82 | 10.62 | 5.54 | 4.66 | -0.94 | -20.76*** | 4.27 | 3.29 | 2.42 | 3.27 | 0.06 | -4.22 |
| | (4.04) | (3.36) | (2.48) | (2.47) | (-0.41) | (-3.87) | (1.76) | (1.62) | (1.30) | (2.10) | (0.07) | (-1.46) |

*Notes: See notes to Tables 2.5.

# Table 2.8: Double-Sorted Portfolios: Portfolios Sorted by Various Jump Variation Measures

Panel A: Stocks Sorted by SRVLJ, Controlling for SRVJ Based on $\gamma^2$

| | Equal-Weighted Returns and Alphas | | | | | | Value-Weighted Returns and Alphas | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | SRVJ Quintile | | | | | | SRVJ Quintile | | | | | |
| SRVLJ Quintile | 1(Low) | 2 | 3 | 4 | 5(High) | Average | 1(Low) | 2 | 3 | 4 | 5(High) | Average |
| Part I: Mean Return and Alpha | | | | | | | | | | | | |
| 1(Low) | 47.99 | 28.73 | 21.12 | 12.90 | 7.82 | 23.71 | 41.88 | 18.83 | 19.24 | 13.59 | 5.04 | 16.61 |
| 2 | 50.55 | 37.24 | 22.38 | 26.91 | 12.99 | 33.21 | 29.61 | 25.26 | 18.26 | 17.39 | 14.08 | 22.70 |
| 3 | 53.07 | 38.57 | 23.49 | 21.73 | 12.37 | 29.75 | 36.31 | 26.28 | 8.10 | 15.50 | 17.40 | 24.19 |
| 4 | 56.29 | 21.42 | 35.09 | 14.59 | 17.86 | 27.51 | 36.38 | 7.94 | 18.53 | 13.64 | 16.59 | 20.33 |
| 5(High) | 47.70 | 45.92 | 32.70 | 26.10 | 7.59 | 32.00 | 32.64 | 39.20 | 23.00 | 20.20 | 19.84 | 27.00 |
| High-Low | -0.24 | 17.19 | 11.58 | 13.20 | -0.23 | 8.30 | 6.38 | 20.37 | 3.76 | 6.61 | 14.81 | 10.38 |
| Alpha | -6.48 | 16.64 | 10.80 | 13.88 | 2.60 | 7.49 | 3.69 | 19.41 | 2.46 | 6.06 | 13.84 | 9.09 |
| Part II: t-Statistics | | | | | | | | | | | | |
| 1(Low) | 5.15 | 2.90 | 2.26 | 1.48 | 0.92 | 2.68 | 3.45 | 2.66 | 2.89 | 2.08 | 0.75 | 2.67 |
| 2 | 4.81 | 3.62 | 1.94 | 0.81 | 1.43 | 3.49 | 3.97 | 3.40 | 2.28 | 0.88 | 1.99 | 3.49 |
| 3 | 4.87 | 3.03 | 0.75 | 1.89 | 1.39 | 3.03 | 4.65 | 3.06 | 0.37 | 1.92 | 2.65 | 3.73 |
| 4 | 4.64 | 0.63 | 3.07 | 1.53 | 2.04 | 2.81 | 4.26 | 0.27 | 2.18 | 2.05 | 2.50 | 3.04 |
| 5(High) | 4.39 | 4.29 | 3.23 | 2.71 | 0.89 | 3.38 | 3.77 | 4.89 | 3.16 | 2.78 | 2.99 | 4.01 |
| High-Low | -0.04 | 4.24 | 3.31 | 3.87 | -0.06 | 4.18 | 1.09 | 3.71 | 0.79 | 1.54 | 3.09 | 4.15 |
| Alpha | -1.19 | 4.11 | 3.06 | 4.09 | 0.67 | 3.80 | 0.62 | 3.40 | 0.49 | 1.44 | 2.96 | 3.53 |

Panel B: Stocks Sorted by SRVSJ, Controlling for SRVJ Based on $\gamma^2$

| | Equal-Weighted Returns and Alphas | | | | | | Value-Weighted Returns and Alphas | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | SRVJ Quintile | | | | | | SRVJ Quintile | | | | | |
| SRVSJ Quintile | 1(Low) | 2 | 3 | 4 | 5(High) | Average | 1(Low) | 2 | 3 | 4 | 5(High) | Average |
| Part I: Mean Return and Alpha | | | | | | | | | | | | |
| 1(Low) | 56.90 | 44.50 | 34.78 | 26.13 | 19.50 | 36.36 | 41.88 | 35.91 | 26.98 | 18.28 | 22.43 | 29.73 |
| 2 | 55.66 | 47.02 | 30.91 | 18.59 | 10.34 | 32.50 | 40.76 | 34.11 | 20.84 | 13.23 | 16.52 | 25.09 |
| 3 | 56.24 | 41.20 | 28.46 | 19.74 | 10.02 | 31.13 | 30.72 | 23.04 | 20.39 | 13.82 | 14.29 | 20.45 |
| 4 | 53.58 | 37.21 | 16.75 | 12.62 | 10.19 | 26.07 | 28.23 | 23.82 | 12.02 | 14.94 | 6.93 | 17.19 |
| 5(High) | 34.30 | 25.12 | 19.72 | 12.12 | 5.79 | 19.41 | 17.03 | 16.65 | 18.03 | 12.69 | 3.77 | 13.63 |
| High-Low | -22.60 | -19.38 | -15.06 | -14.01 | -13.71 | -16.95 | -28.00 | -19.26 | -8.95 | -5.59 | -18.66 | -16.09 |
| Alpha | -19.26 | -18.83 | -14.19 | -14.86 | -16.20 | -16.67 | -25.86 | -20.41 | -6.87 | -4.79 | -18.22 | -15.23 |
| Part II: t-Statistics | | | | | | | | | | | | |
| 1(Low) | 5.44 | 4.22 | 3.40 | 2.67 | 2.16 | 3.76 | 5.47 | 4.75 | 3.58 | 2.54 | 3.33 | 4.44 |
| 2 | 5.20 | 4.21 | 3.08 | 1.93 | 1.17 | 3.35 | 5.19 | 4.65 | 3.00 | 1.95 | 2.33 | 3.86 |
| 3 | 5.15 | 3.99 | 2.77 | 2.14 | 1.17 | 3.26 | 4.04 | 3.21 | 2.89 | 2.02 | 2.20 | 3.20 |
| 4 | 5.33 | 3.58 | 1.71 | 1.42 | 1.23 | 2.86 | 3.39 | 3.41 | 1.72 | 2.22 | 1.02 | 2.68 |
| 5(High) | 3.46 | 2.60 | 2.12 | 1.39 | 0.70 | 2.19 | 2.22 | 2.32 | 2.67 | 1.89 | 0.56 | 2.16 |
| High-Low | -4.99 | -5.04 | -4.14 | -3.79 | -3.45 | -7.22 | -5.37 | -3.66 | -1.79 | -1.19 | -3.80 | -5.77 |
| Alpha | -4.15 | -4.90 | -3.91 | -4.16 | -4.12 | -7.42 | -4.93 | -3.65 | -1.30 | -1.02 | -3.78 | -5.32 |

# Table 2.8 (Continued)

Panel C: Stocks Sorted by SRVSJ, Controlling for SRVLJ Based on $\gamma^2$

| | Equal-Weighted Returns and Alphas | | | | | | Value-Weighted Returns and Alphas | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | SRVLJ Quintile | | | | | | SRVLJ Quintile | | | | | |
| SRVSJ Quintile | 1(Low) | 2 | 3 | 4 | 5(High) | Average | 1(Low) | 2 | 3 | 4 | 5(High) | Average |
| Part I: Mean Return and Alpha | | | | | | | | | | | | |
| 1(Low) | 60.34 | 51.38 | 40.69 | 44.93 | 31.91 | 47.28 | 41.88 | 35.33 | 35.90 | 34.10 | 30.94 | 36.43 |
| 2 | 57.30 | 38.41 | 58.52 | 28.26 | 28.35 | 40.27 | 23.32 | 25.07 | 37.49 | 18.00 | 25.93 | 25.12 |
| 3 | 35.25 | 26.42 | 4.21 | 28.76 | 14.62 | 25.17 | 16.19 | 21.74 | 9.63 | 20.36 | 21.71 | 19.86 |
| 4 | 28.42 | 16.02 | 48.13 | 10.74 | 10.78 | 19.49 | 17.78 | 15.39 | 97.02 | 14.92 | 14.12 | 22.20 |
| 5(High) | 18.79 | 9.45 | -1.20 | 8.01 | 8.26 | 10.89 | 14.98 | 8.09 | 6.98 | 7.78 | 13.31 | 11.30 |
| High-Low | -41.55 | -41.93 | -40.72 | -38.27 | -23.65 | -36.71 | -26.63 | -27.24 | -27.87 | -27.67 | -17.63 | -25.38 |
| Alpha | -40.15 | -41.33 | -40.17 | -37.25 | -22.90 | -35.45 | -25.59 | -25.58 | -26.01 | -26.35 | -17.73 | -24.07 |
| Part II: t-Statistics | | | | | | | | | | | | |
| 1(Low) | 5.60 | 4.76 | 1.56 | 4.18 | 3.28 | 4.58 | 5.38 | 4.76 | 1.72 | 4.12 | 4.37 | 5.13 |
| 2 | 5.40 | 3.61 | 2.17 | 2.64 | 3.02 | 3.97 | 3.03 | 3.41 | 1.61 | 2.34 | 3.60 | 3.62 |
| 3 | 3.58 | 2.60 | 0.19 | 2.85 | 1.62 | 2.65 | 2.27 | 3.21 | 0.53 | 2.93 | 3.11 | 3.15 |
| 4 | 2.83 | 1.70 | 1.20 | 1.10 | 1.23 | 2.01 | 2.49 | 2.30 | 1.11 | 2.03 | 2.07 | 2.35 |
| 5(High) | 2.12 | 1.06 | -0.06 | 0.93 | 1.01 | 1.31 | 2.12 | 1.22 | 0.46 | 1.08 | 2.02 | 1.82 |
| High-Low | -8.29 | -8.49 | -2.04 | -6.08 | -5.33 | -8.59 | -5.15 | -5.54 | -1.50 | -4.14 | -3.64 | -6.61 |
| Alpha | -8.24 | -8.75 | -2.03 | -6.22 | -5.50 | -8.80 | -4.90 | -5.18 | -1.39 | -3.87 | -3.66 | -6.13 |

Panel D: Stocks Sorted by SRVLJ, Controlling for SRVSJ Based on $\gamma^2$

| | Equal-Weighted Returns and Alphas | | | | | | Value-Weighted Returns and Alphas | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | SRVSJ Quintile | | | | | | SRVSJ Quintile | | | | | |
| SRVLJ Quintile | 1(Low) | 2 | 3 | 4 | 5(High) | Average | 1(Low) | 2 | 3 | 4 | 5(High) | Average |
| Part I: Mean Return and Alpha | | | | | | | | | | | | |
| 1(Low) | 59.99 | 60.01 | 41.84 | 27.66 | 16.62 | 41.23 | 41.88 | 30.53 | 19.67 | 16.32 | 12.01 | 23.80 |
| 2 | 49.75 | 41.95 | 27.20 | 17.65 | 11.51 | 29.31 | 35.69 | 25.91 | 18.85 | 17.68 | 8.92 | 20.89 |
| 3 | 76.41 | 38.05 | 24.45 | 6.48 | 14.41 | 32.06 | 45.15 | 16.78 | 11.30 | 10.47 | 21.45 | 21.03 |
| 4 | 43.99 | 30.88 | 23.29 | 13.82 | 11.84 | 24.86 | 38.54 | 18.80 | 17.59 | 15.79 | 10.15 | 20.55 |
| 5(High) | 32.38 | 28.64 | 10.59 | 8.97 | 6.78 | 17.47 | 31.35 | 25.93 | 18.89 | 13.97 | 10.03 | 20.03 |
| High-Low | -27.61 | -31.38 | -31.25 | -18.69 | -9.84 | -23.75 | -9.14 | -4.60 | -0.78 | -2.35 | -1.98 | -3.77 |
| Alpha | -26.51 | -31.63 | -31.28 | -19.37 | -9.80 | -23.72 | -8.62 | -3.72 | -0.59 | -2.79 | -1.66 | -3.47 |
| Part II: t-Statistics | | | | | | | | | | | | |
| 1(Low) | 5.51 | 5.59 | 4.22 | 2.85 | 1.85 | 4.25 | 5.37 | 4.03 | 2.69 | 2.23 | 1.71 | 3.59 |
| 2 | 4.51 | 3.85 | 2.62 | 1.87 | 1.29 | 3.05 | 4.71 | 3.54 | 2.63 | 2.71 | 1.32 | 3.28 |
| 3 | 2.31 | 1.95 | 1.61 | 0.41 | 0.80 | 2.49 | 1.53 | 1.16 | 1.08 | 0.78 | 1.49 | 2.25 |
| 4 | 4.05 | 2.90 | 2.28 | 1.44 | 1.31 | 2.63 | 4.83 | 2.51 | 2.47 | 2.31 | 1.39 | 3.23 |
| 5(High) | 3.36 | 2.99 | 1.21 | 1.03 | 0.83 | 2.02 | 4.32 | 3.63 | 2.69 | 2.11 | 1.55 | 3.23 |
| High-Low | -6.48 | -7.66 | -6.92 | -4.65 | -2.73 | -8.80 | -1.83 | -0.98 | -0.16 | -0.46 | -0.44 | -1.49 |
| Alpha | -6.16 | -7.50 | -6.94 | -4.94 | -2.83 | -8.94 | -1.70 | -0.79 | -0.12 | -0.53 | -0.37 | -1.38 |

*Notes: See notes to Table 2.5. This table presents average returns (called "Mean Return") and risk-adjusted alphas (called "Alpha") for portfolios sorted by various jump variation measures. The sample includes NYSE, NASDAQ and AMEX listed stocks for the period January 1993 to December 2016. At the end of each Tuesday, all the stocks in the sample are sorted into quintile portfolios based on ascending values of SRVJ (SRVLJ/SRVSJ). Then, within each quintile portfolio, stocks are further sorted based on the values of SRVLJ/SRVSJ (SRVSJ/SRVLJ), resulting in 25 portfolios. Each portfolio is held for one week. The row labeled "High-Low" reports the average values of one-week ahead returns in Part I (corresponding Newey-West $t$-statistics are given in Part II of the panel). The row labeled "Alpha" reports Fama-French-Carhart four-factor alphas in Part I (corresponding Newey-West $t$-statistics are again given in Part II of the panel) for the double-sorted High-Low portfolios. Note that entries given in the "Average" column of the table, are average returns across the 5 quintiles. Finally, note that SRVLJ and SRVSJ are constructed based on jump truncation level $\gamma^2 = 5\sqrt{\frac{1}{t}\widehat{IV}_t^{(i)}}\Delta_n^{0.49}$.

# Table 2.9: Double-Sorted Portfolios: Portfolios Sorted by SRVJ and RSK

Panel A: Stocks Sorted by SRVJ, Controlling for RSK

| | Equal-Weighted Returns and Alphas | | | | | | Value-Weighted Returns and Alphas | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | RSK Quintile | | | | | | RSK Quintile | | | | | |
| SRVJ Quintile | 1(Low) | 2 | 3 | 4 | 5(High) | Average | 1(Low) | 2 | 3 | 4 | 5(High) | Average |
| Part I: Mean Return and Alpha | | | | | | | | | | | | |
| 1(Low) | 56.92 | 55.49 | 42.81 | 31.69 | 20.91 | 41.57 | 41.88 | 39.12 | 30.49 | 25.47 | 17.56 | 30.90 |
| 2 | 56.77 | 46.17 | 33.30 | 22.89 | 15.27 | 34.88 | 41.63 | 35.10 | 21.38 | 17.20 | 13.71 | 25.81 |
| 3 | 49.97 | 38.92 | 23.47 | 17.47 | 11.61 | 28.29 | 37.18 | 27.01 | 22.24 | 14.45 | 10.52 | 22.28 |
| 4 | 42.67 | 29.57 | 21.68 | 13.20 | 12.18 | 23.86 | 28.53 | 20.24 | 15.10 | 11.99 | 11.47 | 17.47 |
| 5(High) | 31.39 | 20.03 | 17.94 | 11.86 | 4.85 | 17.21 | 18.32 | 21.67 | 12.23 | 7.31 | 12.60 | 14.42 |
| High-Low | -25.54 | -35.46 | -24.87 | -19.83 | -16.06 | -24.35 | -23.56 | -17.46 | -18.27 | -18.16 | -4.95 | -16.48 |
| Alpha | -28.79 | -36.20 | -24.40 | -18.40 | -12.75 | -24.11 | -24.22 | -18.52 | -18.40 | -16.82 | -4.88 | -16.57 |
| Part II: t-Statistics | | | | | | | | | | | | |
| 1(Low) | 6.13 | 5.04 | 3.97 | 3.04 | 2.22 | 4.21 | 5.72 | 4.87 | 3.97 | 3.35 | 2.51 | 4.58 |
| 2 | 5.34 | 4.12 | 3.14 | 2.32 | 1.70 | 3.52 | 5.11 | 4.57 | 2.96 | 2.44 | 2.02 | 3.89 |
| 3 | 4.67 | 3.80 | 2.30 | 1.90 | 1.35 | 2.99 | 4.90 | 3.78 | 3.07 | 2.14 | 1.63 | 3.51 |
| 4 | 4.09 | 2.98 | 2.27 | 1.46 | 1.44 | 2.61 | 3.67 | 2.84 | 2.17 | 1.76 | 1.72 | 2.73 |
| 5(High) | 3.27 | 2.20 | 2.02 | 1.36 | 0.59 | 2.01 | 2.59 | 3.24 | 1.76 | 1.07 | 1.90 | 2.36 |
| High-Low | -5.36 | -7.09 | -5.12 | -4.32 | -3.40 | -7.70 | -4.70 | -3.17 | -3.32 | -3.28 | -0.89 | -5.24 |
| Alpha | -6.18 | -7.57 | -5.25 | -4.32 | -2.80 | -8.07 | -4.81 | -3.50 | -3.36 | -3.04 | -0.90 | -5.45 |

Panel B: Stocks Sorted by RSK, Controlling for SRVJ

| | Equal-Weighted Returns and Alphas | | | | | | Value-Weighted Returns and Alphas | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | SRVJ Quintile | | | | | | SRVJ Quintile | | | | | |
| RSK Quintile | 1(Low) | 2 | 3 | 4 | 5(High) | Average | 1(Low) | 2 | 3 | 4 | 5(High) | Average |
| Part I: Mean Return and Alpha | | | | | | | | | | | | |
| 1(Low) | 51.14 | 29.34 | 21.71 | 17.58 | 12.40 | 26.43 | 41.88 | 18.33 | 20.91 | 12.07 | 4.09 | 18.48 |
| 2 | 49.50 | 39.74 | 23.05 | 14.15 | 11.46 | 27.58 | 37.29 | 33.29 | 14.95 | 10.36 | 10.00 | 21.18 |
| 3 | 49.96 | 37.41 | 27.26 | 18.14 | 12.93 | 29.14 | 28.38 | 23.82 | 22.52 | 14.80 | 15.23 | 20.95 |
| 4 | 53.85 | 43.31 | 28.11 | 18.79 | 12.52 | 31.32 | 36.37 | 28.73 | 19.57 | 17.15 | 10.67 | 22.50 |
| 5(High) | 54.80 | 45.32 | 30.65 | 20.66 | 5.78 | 31.44 | 36.75 | 32.79 | 24.63 | 18.92 | 14.26 | 25.47 |
| High-Low | 3.66 | 15.98 | 8.94 | 3.08 | -6.63 | 5.01 | -0.25 | 14.46 | 3.72 | 6.85 | 10.17 | 6.99 |
| Alpha | 0.54 | 16.64 | 8.57 | 2.34 | -4.54 | 4.71 | -0.71 | 15.66 | 4.39 | 6.30 | 9.15 | 6.96 |
| Part II: t-Statistics | | | | | | | | | | | | |
| 1(Low) | 5.56 | 3.07 | 2.36 | 1.99 | 1.42 | 3.02 | 4.98 | 2.61 | 3.15 | 1.74 | 0.60 | 2.97 |
| 2 | 4.76 | 3.75 | 2.33 | 1.57 | 1.30 | 2.94 | 4.89 | 4.50 | 2.06 | 1.53 | 1.52 | 3.29 |
| 3 | 4.76 | 3.60 | 2.69 | 1.94 | 1.50 | 3.07 | 3.59 | 3.33 | 3.29 | 2.24 | 2.33 | 3.30 |
| 4 | 4.99 | 3.98 | 2.77 | 1.97 | 1.49 | 3.27 | 4.70 | 3.89 | 2.69 | 2.39 | 1.63 | 3.45 |
| 5(High) | 5.00 | 4.19 | 2.99 | 2.17 | 0.68 | 3.27 | 4.50 | 4.20 | 3.29 | 2.83 | 2.09 | 3.87 |
| High-Low | 0.85 | 3.98 | 2.31 | 0.84 | -1.61 | 2.35 | -0.05 | 2.85 | 0.74 | 1.48 | 2.19 | 2.87 |
| Alpha | 0.14 | 4.40 | 2.24 | 0.66 | -1.09 | 2.42 | -0.14 | 3.05 | 0.89 | 1.38 | 2.01 | 2.92 |

*Notes: See notes to Table 2.5. This table presents average returns (called "Mean Return") and risk-adjusted alphas (called "Alpha") for portfolios sorted by SRVJ controlling for RSK, and vice versa. The sample includes NYSE, NASDAQ and AMEX listed stocks for the period January 1993 to December 2016. At the end of each Tuesday, all the stocks in the sample are sorted into quintile portfolios based on ascending values of RSK (SRVJ), and then within each quintile portfolio, stocks are further sorted using values of SRVJ (RSK), resulting in 25 portfolios. Each portfolio is held for one week. The row labeled "High-Low" reports the average values of one-week ahead returns in Part I (corresponding Newey-West $t$-statistics are given in Part II of the panel). The row labeled "Alpha" reports Fama-French-Carhart four-factor alphas in Part I (corresponding Newey-West $t$-statistics are again given in Part II of the panel) for each of the quintile portfolios, as well as for the average across 5 RSK (SRVJ) portfolios.

# Table 2.10: Double-Sorted Portfolios: Portfolios Sorted by SRVLJ/SRVSJ, Controlling for RSK

Panel A: Stocks Sorted by SRVLJ, Controlling for RSK Based on $\gamma^2$

| | Equal-Weighted Returns and Alphas | | | | | | Value-Weighted Returns and Alphas | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | RSK Quintile | | | | | | RSK Quintile | | | | | |
| SRVLJ Quintile | 1(Low) | 2 | 3 | 4 | 5(High) | Average | 1(Low) | 2 | 3 | 4 | 5(High) | Average |
| Part I: Mean Return and Alpha | | | | | | | | | | | | |
| 1(Low) | 47.31 | 31.04 | 20.83 | 13.46 | 9.23 | 24.38 | 41.88 | 20.74 | 17.63 | 10.28 | 10.95 | 17.04 |
| 2 | 48.24 | 35.53 | 18.15 | 41.82 | 10.40 | 30.72 | 27.52 | 25.93 | 13.06 | 22.40 | 13.21 | 21.88 |
| 3 | 46.79 | 36.39 | -2.22 | 17.53 | 16.01 | 27.86 | 27.77 | 24.25 | 23.68 | 13.64 | 18.71 | 21.29 |
| 4 | 47.56 | 9.61 | 27.36 | 19.04 | 20.92 | 27.46 | 27.71 | 8.29 | 20.27 | 17.45 | 16.06 | 20.34 |
| 5(High) | 43.55 | 43.48 | 34.14 | 26.41 | 7.84 | 31.13 | 35.80 | 32.07 | 20.61 | 23.70 | 18.89 | 26.28 |
| High-Low | -3.57 | 12.44 | 13.30 | 12.95 | -1.39 | 6.75 | 10.42 | 11.33 | 2.98 | 13.42 | 7.94 | 9.22 |
| Alpha | -8.76 | 11.52 | 13.22 | 13.09 | 1.59 | 6.13 | 8.26 | 11.20 | 1.91 | 12.69 | 7.30 | 8.27 |
| Part II: t-Statistics | | | | | | | | | | | | |
| 1(Low) | 5.10 | 3.13 | 2.17 | 1.51 | 1.08 | 2.73 | 3.34 | 2.85 | 2.51 | 1.59 | 1.63 | 2.68 |
| 2 | 4.67 | 3.47 | 1.41 | 1.61 | 1.16 | 3.21 | 3.71 | 3.71 | 1.47 | 1.12 | 1.98 | 3.42 |
| 3 | 4.48 | 2.95 | -0.07 | 1.54 | 1.79 | 2.91 | 3.81 | 2.87 | 0.85 | 1.56 | 2.89 | 3.30 |
| 4 | 4.32 | 0.37 | 2.27 | 1.96 | 2.38 | 2.89 | 3.68 | 0.43 | 2.12 | 2.52 | 2.45 | 3.17 |
| 5(High) | 4.22 | 4.03 | 3.42 | 2.75 | 0.93 | 3.34 | 4.08 | 3.80 | 2.81 | 3.23 | 2.79 | 3.90 |
| High-Low | -0.66 | 2.93 | 4.28 | 4.11 | -0.37 | 3.64 | 1.71 | 1.89 | 0.66 | 3.03 | 1.62 | 3.76 |
| Alpha | -1.62 | 2.63 | 4.14 | 4.13 | 0.42 | 3.18 | 1.35 | 1.76 | 0.41 | 2.90 | 1.51 | 3.22 |

Panel B: Stocks Sorted by SRVSJ, Controlling for RSK Based on $\gamma^2$

| | Equal-Weighted Returns and Alphas | | | | | | Value-Weighted Returns and Alphas | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | RSK Quintile | | | | | | RSK Quintile | | | | | |
| SRVSJ Quintile | 1(Low) | 2 | 3 | 4 | 5(High) | Average | 1(Low) | 2 | 3 | 4 | 5(High) | Average |
| Part I: Mean Return and Alpha | | | | | | | | | | | | |
| 1(Low) | 56.15 | 51.19 | 42.56 | 31.93 | 24.16 | 41.20 | 41.88 | 40.69 | 31.16 | 31.49 | 21.39 | 33.55 |
| 2 | 54.34 | 50.15 | 35.79 | 24.76 | 12.89 | 35.59 | 32.45 | 33.20 | 21.93 | 19.73 | 23.95 | 26.25 |
| 3 | 55.24 | 38.53 | 26.29 | 18.62 | 11.09 | 29.95 | 27.91 | 22.64 | 18.61 | 12.36 | 16.37 | 19.58 |
| 4 | 40.20 | 31.45 | 18.77 | 14.26 | 12.14 | 23.35 | 18.13 | 22.01 | 15.30 | 13.17 | 10.40 | 15.79 |
| 5(High) | 30.21 | 18.90 | 15.72 | 7.49 | 4.76 | 15.42 | 18.59 | 19.31 | 12.59 | 8.23 | 6.56 | 13.06 |
| High-Low | -25.94 | -32.29 | -26.84 | -24.43 | -19.40 | -25.78 | -24.44 | -21.37 | -18.56 | -23.26 | -14.84 | -20.49 |
| Alpha | -23.86 | -31.76 | -26.49 | -23.94 | -20.71 | -25.35 | -22.27 | -22.89 | -18.30 | -21.83 | -14.44 | -19.95 |
| Part II: t-Statistics | | | | | | | | | | | | |
| 1(Low) | 5.43 | 4.75 | 3.98 | 3.16 | 2.62 | 4.15 | 5.11 | 5.25 | 4.16 | 4.39 | 3.15 | 4.92 |
| 2 | 5.17 | 4.48 | 3.42 | 2.47 | 1.43 | 3.61 | 4.30 | 4.53 | 2.96 | 2.68 | 3.39 | 3.95 |
| 3 | 5.25 | 3.76 | 2.60 | 1.96 | 1.25 | 3.14 | 3.75 | 3.12 | 2.66 | 1.72 | 2.46 | 3.03 |
| 4 | 4.11 | 3.13 | 1.96 | 1.58 | 1.45 | 2.59 | 2.50 | 3.20 | 2.27 | 1.98 | 1.57 | 2.55 |
| 5(High) | 3.18 | 2.06 | 1.71 | 0.87 | 0.58 | 1.79 | 2.56 | 2.86 | 1.79 | 1.23 | 0.97 | 2.09 |
| High-Low | -5.78 | -7.15 | -6.04 | -5.58 | -4.72 | -8.41 | -4.21 | -4.04 | -3.67 | -4.56 | -2.93 | -6.31 |
| Alpha | -5.36 | -7.20 | -6.12 | -5.75 | -5.27 | -8.89 | -3.83 | -4.29 | -3.51 | -4.09 | -2.83 | -6.01 |

*Notes: See notes to Table 2.8. Portfolios are sorted by SRVLJ/SRVSJ, controlling for RSK, and using truncation level $\gamma^2$, as discussed in the footnote to Table 2.2.

# Table 2.11: Double-Sorted Portfolios: Portfolios Sorted by RSK, Controlling for SRVLJ or SRVSJ

Panel A: Stocks Sorted by RSK, Controlling for SRVLJ Based on $\gamma^2$

| | Equal-Weighted Returns and Alphas | | | | | | Value-Weighted Returns and Alphas | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | SRVLJ Quintile | | | | | | SRVLJ Quintile | | | | | |
| RSK Quintile | 1(Low) | 2 | 3 | 4 | 5(High) | Average | 1(Low) | 2 | 3 | 4 | 5(High) | Average |
| Part I: Mean Return and Alpha | | | | | | | | | | | | |
| 1(Low) | 51.98 | 46.26 | 32.88 | 42.99 | 35.40 | 43.72 | 41.88 | 34.45 | 32.97 | 31.19 | 33.46 | 34.64 |
| 2 | 48.38 | 38.05 | 55.49 | 31.86 | 20.58 | 37.03 | 30.03 | 23.16 | 48.38 | 17.47 | 16.06 | 24.56 |
| 3 | 43.16 | 26.83 | 8.78 | 20.69 | 16.54 | 25.94 | 25.91 | 21.10 | 10.09 | 19.92 | 19.16 | 21.15 |
| 4 | 37.70 | 22.51 | 55.61 | 11.20 | 14.84 | 24.67 | 18.23 | 18.69 | 8.48 | 13.82 | 14.50 | 16.23 |
| 5(High) | 21.43 | 8.10 | -0.87 | 14.29 | 7.74 | 12.81 | 13.59 | 7.52 | -8.20 | 15.74 | 15.50 | 12.31 |
| High-Low | -30.55 | -38.16 | -30.25 | -30.06 | -27.66 | -31.07 | -23.12 | -26.93 | -37.56 | -16.82 | -17.96 | -22.43 |
| Alpha | -32.89 | -37.85 | -28.79 | -30.01 | -24.81 | -30.46 | -23.34 | -24.74 | -36.67 | -14.80 | -18.66 | -21.39 |
| Part II: t-Statistics | | | | | | | | | | | | |
| 1(Low) | 5.63 | 4.46 | 1.33 | 4.16 | 3.61 | 4.48 | 4.94 | 4.65 | 1.67 | 4.00 | 4.80 | 5.06 |
| 2 | 4.68 | 3.68 | 2.16 | 3.04 | 2.22 | 3.72 | 3.99 | 3.28 | 2.04 | 2.34 | 2.31 | 3.63 |
| 3 | 4.14 | 2.66 | 0.40 | 2.04 | 1.84 | 2.69 | 3.32 | 3.22 | 0.59 | 2.66 | 2.87 | 3.29 |
| 4 | 3.62 | 2.29 | 1.37 | 1.11 | 1.73 | 2.50 | 2.54 | 2.82 | 0.48 | 1.80 | 2.22 | 2.54 |
| 5(High) | 2.27 | 0.89 | -0.04 | 1.63 | 0.92 | 1.49 | 1.91 | 1.08 | -0.52 | 2.33 | 2.24 | 1.98 |
| High-Low | -7.91 | -9.22 | -1.71 | -5.63 | -5.85 | -9.46 | -4.56 | -6.02 | -2.29 | -2.75 | -3.47 | -6.54 |
| Alpha | -8.43 | -9.50 | -1.64 | -5.86 | -5.51 | -9.57 | -4.59 | -5.54 | -2.22 | -2.37 | -3.65 | -6.10 |

Panel B: Stocks Sorted by RSK, Controlling for SRVSJ Based on $\gamma^2$

| | Equal-Weighted Returns and Alphas | | | | | | Value-Weighted Returns and Alphas | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | SRVSJ Quintile | | | | | | SRVSJ Quintile | | | | | |
| RSK Quintile | 1(Low) | 2 | 3 | 4 | 5(High) | Average | 1(Low) | 2 | 3 | 4 | 5(High) | Average |
| Part I: Mean Return and Alpha | | | | | | | | | | | | |
| 1(Low) | 57.49 | 56.03 | 40.07 | 26.79 | 17.75 | 39.62 | 41.88 | 24.28 | 18.84 | 16.57 | 12.21 | 23.20 |
| 2 | 51.21 | 42.88 | 31.02 | 15.27 | 12.69 | 30.61 | 35.14 | 26.67 | 14.77 | 21.10 | 9.49 | 21.43 |
| 3 | 54.96 | 41.34 | 25.81 | 20.97 | 10.12 | 30.64 | 37.20 | 24.66 | 20.50 | 10.79 | 7.58 | 20.15 |
| 4 | 42.52 | 38.93 | 29.50 | 16.21 | 9.58 | 27.35 | 37.22 | 23.08 | 17.57 | 15.29 | 11.02 | 20.84 |
| 5(High) | 33.23 | 28.85 | 9.62 | 10.20 | 5.82 | 17.54 | 31.94 | 30.24 | 21.43 | 16.36 | 8.89 | 21.77 |
| High-Low | -24.26 | -27.17 | -30.45 | -16.59 | -11.93 | -22.08 | -12.15 | 5.96 | 2.59 | -0.21 | -3.33 | -1.43 |
| Alpha | -22.78 | -27.83 | -30.17 | -17.16 | -11.85 | -21.96 | -10.86 | 7.74 | 2.94 | -0.96 | -3.03 | -0.83 |
| Part II: t-Statistics | | | | | | | | | | | | |
| 1(Low) | 5.53 | 5.32 | 4.11 | 2.82 | 1.99 | 4.19 | 5.79 | 3.21 | 2.52 | 2.21 | 1.73 | 3.49 |
| 2 | 4.84 | 4.00 | 3.06 | 1.62 | 1.44 | 3.20 | 4.45 | 3.56 | 2.02 | 3.25 | 1.42 | 3.30 |
| 3 | 4.96 | 3.79 | 2.58 | 2.21 | 1.15 | 3.15 | 4.86 | 3.35 | 3.01 | 1.55 | 1.13 | 3.12 |
| 4 | 4.01 | 3.79 | 2.86 | 1.73 | 1.12 | 2.87 | 5.03 | 3.26 | 2.44 | 2.17 | 1.62 | 3.22 |
| 5(High) | 3.41 | 3.05 | 1.09 | 1.17 | 0.72 | 2.03 | 4.19 | 4.35 | 2.91 | 2.49 | 1.38 | 3.48 |
| High-Low | -5.79 | -6.79 | -6.74 | -4.40 | -3.12 | -8.56 | -2.27 | 1.17 | 0.48 | -0.04 | -0.69 | -0.49 |
| Alpha | -5.33 | -6.78 | -6.68 | -4.61 | -3.19 | -8.57 | -2.06 | 1.50 | 0.53 | -0.17 | -0.61 | -0.29 |

*Notes: See notes to Table 2.8. Portfolios are sorted by RSK, controlling for SRVLJ/SRVSJ, and using truncation level $\gamma^2$, as discussed in the footnote to Table 2.2.

# Table 2.12: Double-Sorted Portfolios: Portfolios Independently Sorted by Stock- and Industry-Level SRVJ

| | Equal-Weighted Returns and Alphas | | | | | | | Value-Weighted Returns and Alphas | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Industry-Level Quintile | | | | | | | Industry-Level Quintile | | | | | | |
| Stock-Level Quintile | 1(Low) | 2 | 3 | 4 | 5(High) | High-Low | Alpha | 1(Low) | 2 | 3 | 4 | 5(High) | High-Low | Alpha |
| Part I: Mean Return and Alpha | | | | | | | | | | | | | | |
| 1(Low) | 39.23 | 47.42 | 53.00 | 61.43 | 72.18 | 32.94 | 34.74 | 36.50 | 33.40 | 32.78 | 36.58 | 25.03 | 10.19 | 9.28 |
| 2 | 30.55 | 33.28 | 38.69 | 48.44 | 55.77 | 25.22 | 25.86 | 28.37 | 27.31 | 20.24 | 37.10 | 14.64 | 7.20 | 6.68 |
| 3 | 14.07 | 21.47 | 24.27 | 30.98 | 43.15 | 29.07 | 29.60 | 18.16 | 24.85 | 17.56 | 20.35 | 8.69 | 10.64 | 11.98 |
| 4 | 6.90 | 12.45 | 15.65 | 19.51 | 34.51 | 27.61 | 28.30 | 12.97 | 11.96 | 10.38 | 14.58 | 2.67 | 9.54 | 9.99 |
| 5(High) | -7.35 | 1.99 | 9.90 | 16.40 | 25.94 | 33.29 | 32.92 | -4.30 | 9.86 | 4.86 | 13.25 | -3.95 | 19.82 | 21.72 |
| High-Low | -46.58 | -45.43 | -43.09 | -45.03 | -46.24 | | | -40.80 | -23.54 | -27.92 | -23.33 | -31.16 | | |
| Alpha | -44.64 | -45.23 | -41.18 | -45.96 | -46.46 | | | -41.42 | -22.73 | -26.52 | -23.09 | -28.98 | | |
| Industry-Level Effect (average of High-Low column; Alpha column) | | | | | | 29.63 | 30.29 | | | | | | 11.48 | 11.93 |
| Stock-Level Effect (average of High-Low row; Alpha row) | | | | | | -45.28 | -44.70 | | | | | | -29.35 | -28.55 |
| Part II: t-Statistics | | | | | | | | | | | | | | |
| 1(Low) | 3.82 | 4.37 | 4.98 | 5.87 | 6.83 | 5.20 | 5.35 | 4.61 | 3.98 | 3.93 | 4.55 | 4.71 | 1.40 | 1.27 |
| 2 | 2.90 | 3.13 | 3.49 | 4.48 | 5.22 | 3.50 | 3.35 | 3.82 | 3.56 | 2.35 | 4.58 | 3.16 | 1.14 | 1.00 |
| 3 | 1.39 | 2.09 | 2.28 | 3.11 | 4.43 | 4.60 | 4.53 | 2.34 | 3.23 | 2.25 | 2.76 | 1.87 | 1.54 | 1.73 |
| 4 | 0.70 | 1.28 | 1.60 | 2.12 | 3.93 | 4.40 | 4.47 | 1.69 | 1.62 | 1.36 | 1.87 | 0.68 | 1.50 | 1.54 |
| 5(High) | -0.79 | 0.22 | 1.07 | 1.88 | 3.16 | 5.88 | 5.93 | -0.51 | 1.27 | 0.62 | 1.93 | -1.00 | 2.82 | 3.06 |
| High-Low | -9.33 | -8.12 | -8.27 | -7.80 | -8.01 | | | -7.19 | -3.92 | -5.22 | -4.31 | -5.33 | | |
| Alpha, FFC4 | -9.00 | -8.16 | -7.97 | -7.93 | -8.34 | | | -7.37 | -3.73 | -5.06 | -4.31 | -5.47 | | |
| Industry-Level Effect (average of High-Low column; Alpha column) | | | | | | 5.66 | 5.57 | | | | | | 2.23 | 2.20 |
| Stock-Level Effect (average of High-Low row; Alpha row) | | | | | | -11.39 | -11.50 | | | | | | -8.88 | -8.91 |

*Notes: See notes to Table 2.8. This table presents average returns and risk-adjusted alphas for portfolios sorted by stock-level and industry-level SRVJ. The sample includes all NYSE, NASDAQ and AMEX listed stocks for the period January 1993 to December 2016. A stock's industry signed jump variation (SRVJ) is the capitalization-weighted average of the SRVJ of all stocks within the industry. At the end of each Tuesday, all stocks in the sample are sorted into quintile portfolios based on stock-level and industry-level SRVJ, independently, resulting in 25 portfolios. Each portfolio is held for one week. The row labeled "Industry-Level Effect" reports average values of one-week ahead returns (and Fama-French-Carhart four-factor alphas in the High-Low (Alpha) column) in Part I (corresponding Newey-West $t$-statistics are given in Part II). The row labeled "Stock-Level Effect" reports the average values of one-week ahead returns (and alphas) in Part I (corresponding Newey-West $t$-statistics are again given in Part II).

# Table 2.13: Double-Sorted Portfolios: Portfolios Sorted by Stock- and Industry-Level SRVLJ/SRVSJ Independently

Panel A: Portfolios Sorted Based on SRVLJ

| | Equal-Weighted Returns and Alphas | | | | | | | Value-Weighted Returns and Alphas | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Industry-Level Quintile | | | | | | | Industry-Level Quintile | | | | | | |
| Stock-Level Quintile | 1(Low) | 2 | 3 | 4 | 5(High) | High-Low | Alpha | 1(Low) | 2 | 3 | 4 | 5(High) | High-Low | Alpha |
| Part I: Mean Return and Alpha | | | | | | | | | | | | | | |
| 1(Low) | 34.64 | 35.34 | 35.04 | 51.12 | 51.00 | 16.36 | 17.03 | 22.26 | 24.61 | 20.46 | 26.39 | 25.19 | 2.92 | 0.73 |
| 2 | 25.37 | 25.36 | 22.38 | 34.21 | 36.00 | 10.64 | 9.43 | 18.78 | 22.97 | 19.49 | 23.51 | 27.29 | 8.51 | 7.46 |
| 3 | 17.55 | 0.89 | 26.31 | 34.26 | 31.56 | 14.40 | 11.15 | 18.90 | 15.95 | 11.59 | 4.81 | 17.06 | -6.76 | -10.28 |
| 4 | 13.59 | 20.91 | 21.20 | 27.71 | 34.08 | 19.69 | 19.75 | 11.26 | 21.71 | 10.01 | 12.25 | 25.88 | 13.89 | 15.35 |
| 5(High) | 12.22 | 11.90 | 14.43 | 25.57 | 25.57 | 13.36 | 12.93 | 14.16 | 21.60 | 16.13 | 20.63 | 22.85 | 8.69 | 11.30 |
| High-Low | -22.42 | -23.44 | -20.61 | -25.55 | -25.42 | | | -8.10 | -3.00 | -4.33 | -5.75 | -2.33 | | |
| Alpha | -22.02 | -23.37 | -20.45 | -25.74 | -26.12 | | | -9.58 | -3.52 | -4.34 | -5.40 | 1.00 | | |
| Industry-Level Effect (average of High-Low column; Alpha column) | | | | | | 14.83 | 14.31 | | | | | | 7.24 | 6.90 |
| Stock-Level Effect (average of High-Low row; Alpha row) | | | | | | -23.49 | -23.54 | | | | | | -4.70 | -4.37 |
| Part II: t-Statistics | | | | | | | | | | | | | | |
| 1(Low) | 3.63 | 3.39 | 3.33 | 4.94 | 5.23 | 3.40 | 3.41 | 3.25 | 3.06 | 2.48 | 3.25 | 3.16 | 0.49 | 0.12 |
| 2 | 2.55 | 2.44 | 2.14 | 3.46 | 3.73 | 2.17 | 1.90 | 2.70 | 3.17 | 2.57 | 3.20 | 3.73 | 1.60 | 1.32 |
| 3 | 0.78 | 0.04 | 0.93 | 1.41 | 1.46 | 0.85 | 0.65 | 1.08 | 0.73 | 0.51 | 0.25 | 1.05 | -0.37 | -0.54 |
| 4 | 1.32 | 1.95 | 1.94 | 2.74 | 3.73 | 3.67 | 3.73 | 1.49 | 2.62 | 1.20 | 1.45 | 3.69 | 2.43 | 2.65 |
| 5(High) | 1.31 | 1.24 | 1.50 | 2.84 | 2.99 | 2.97 | 2.91 | 1.83 | 2.83 | 2.12 | 2.83 | 3.56 | 1.50 | 2.09 |
| High-Low | -6.14 | -5.13 | -4.66 | -5.85 | -6.19 | | | -1.57 | -0.59 | -0.90 | -1.19 | -0.45 | | |
| Alpha, FFC4 | -6.09 | -5.03 | -4.58 | -5.91 | -6.38 | | | -1.93 | -0.68 | -0.87 | -1.09 | 0.21 | | |
| Industry-Level Effect (average of High-Low column; Alpha column) | | | | | | 3.77 | 3.55 | | | | | | 1.77 | 1.57 |
| Stock-Level Effect (average of High-Low row; Alpha row) | | | | | | -9.05 | -9.14 | | | | | | -2.01 | -1.91 |

Panel B: Portfolios Sorted Based on SRVSJ

| | Equal-Weighted Returns and Alphas | | | | | | | Value-Weighted Returns and Alphas | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Industry-Level Quintile | | | | | | | Industry-Level Quintile | | | | | | |
| Stock-Level Quintile | 1(Low) | 2 | 3 | 4 | 5(High) | High-Low | Alpha | 1(Low) | 2 | 3 | 4 | 5(High) | High-Low | Alpha |
| Part I: Mean Return and Alpha | | | | | | | | | | | | | | |
| 1(Low) | 40.39 | 42.50 | 43.55 | 55.02 | 68.45 | 28.07 | 29.88 | 44.09 | 37.87 | 32.07 | 36.14 | 45.53 | 1.44 | 0.72 |
| 2 | 32.52 | 40.60 | 41.29 | 45.73 | 59.67 | 27.15 | 28.20 | 24.11 | 24.09 | 23.82 | 27.45 | 42.09 | 17.99 | 18.36 |
| 3 | 15.39 | 22.96 | 24.90 | 30.69 | 45.29 | 29.90 | 30.43 | 20.22 | 15.93 | 21.91 | 17.63 | 25.37 | 5.16 | 4.88 |
| 4 | 9.07 | 9.08 | 17.26 | 17.20 | 32.70 | 23.63 | 23.94 | 11.38 | 12.88 | 22.53 | 11.01 | 25.43 | 14.04 | 15.23 |
| 5(High) | -2.44 | 3.16 | 7.13 | 12.21 | 22.30 | 24.73 | 26.19 | 8.91 | 7.91 | 11.60 | 9.76 | 11.23 | 2.32 | 5.33 |
| High-Low | -42.83 | -39.34 | -36.43 | -42.82 | -46.16 | | | -35.18 | -29.95 | -20.47 | -26.38 | -34.30 | | |
| Alpha | -42.27 | -39.03 | -34.75 | -42.27 | -45.96 | | | -36.51 | -30.26 | -18.68 | -25.89 | -31.90 | | |
| Industry-Level Effect (average of High-Low column; Alpha column) | | | | | | 26.69 | 27.73 | | | | | | 8.19 | 8.90 |
| Stock-Level Effect (average of High-Low row; Alpha row) | | | | | | -41.51 | -40.86 | | | | | | -29.26 | -28.65 |
| Part II: t-Statistics | | | | | | | | | | | | | | |
| 1(Low) | 3.91 | 3.90 | 3.98 | 5.06 | 6.20 | 4.04 | 3.99 | 5.58 | 4.66 | 3.66 | 4.40 | 5.62 | 0.21 | 0.10 |
| 2 | 3.00 | 3.83 | 3.88 | 4.42 | 5.68 | 3.89 | 3.88 | 3.19 | 3.04 | 2.90 | 3.50 | 5.39 | 2.84 | 2.77 |
| 3 | 1.55 | 2.26 | 2.46 | 3.11 | 4.68 | 4.91 | 4.94 | 2.52 | 1.97 | 2.73 | 2.33 | 3.43 | 0.70 | 0.68 |
| 4 | 0.92 | 0.93 | 1.74 | 1.84 | 3.64 | 3.72 | 3.74 | 1.46 | 1.78 | 2.85 | 1.47 | 3.72 | 2.18 | 2.43 |
| 5(High) | -0.25 | 0.33 | 0.77 | 1.41 | 2.70 | 3.94 | 4.15 | 1.06 | 0.96 | 1.49 | 1.36 | 1.74 | 0.34 | 0.76 |
| High-Low | -8.38 | -7.85 | -7.04 | -7.99 | -7.60 | | | -6.21 | -5.45 | -3.70 | -5.15 | -6.15 | | |
| Alpha, FFC4 | -8.38 | -7.95 | -6.85 | -8.19 | -8.02 | | | -6.55 | -5.68 | -3.41 | -5.15 | -5.69 | | |
| Industry-Level Effect (average of High-Low column; Alpha column) | | | | | | 5.02 | 5.01 | | | | | | 1.60 | 1.68 |
| Stock-Level Effect (average of High-Low row; Alpha row) | | | | | | -10.65 | -11.14 | | | | | | -8.78 | -8.75 |

*Notes: See notes to Table 2.12. Jumps are decomposed using truncation level $\gamma^2$, as discussed in the footnote to Table 2.2.

# Table 2.14: Fama-MacBeth Cross-Sectional Regressions

Panel A: Regressions Without Control Variables

| | I | II | III | IV | V | VI | VII | VIII | IX | X | XI | XII | XIII | XIV | XV | XVI |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Intercept | 18.54 | 27.95 | 23.99 | 15.77 | 30.31 | 31.03 | 31.46 | 31.20 | 19.74 | 32.04 | 20.01 | 20.04 | 28.88 | 30.01 | 30.48 | 30.17 |
| | (1.94) | (3.07) | (2.82) | (1.69) | (3.32) | (3.37) | (3.41) | (3.39) | (2.03) | (3.73) | (2.14) | (2.08) | (3.29) | (3.40) | (3.44) | (3.41) |
| RVJP | -63.86 | | | | | | | | -128.25 | | | | | | | |
| | (-6.00) | | | | | | | | (-6.24) | | | | | | | |
| RVJN | 107.11 | | | | | | | | 196.57 | | | | | | | |
| | (8.29) | | | | | | | | (8.98) | | | | | | | |
| RVLJP | | -53.42 | | -44.85 | | | | | | 76.84 | | -79.63 | | | | |
| | | (-6.46) | | (-4.46) | | | | | | (6.58) | | (-3.94) | | | | |
| RVLJN | | 71.27 | | 83.09 | | | | | | -30.83 | | 149.92 | | | | |
| | | (8.12) | | (7.45) | | | | | | (-2.40) | | (7.18) | | | | |
| RVSJP | | | -130.77 | -99.16 | | | | | | | -88.56 | -129.39 | | | | |
| | | | (-8.97) | (-6.24) | | | | | | | (-6.64) | (-6.33) | | | | |
| RVSJN | | | 165.05 | 161.39 | | | | | | | 129.24 | 195.31 | | | | |
| | | | (8.22) | (7.19) | | | | | | | (7.94) | (8.26) | | | | |
| SRVLJ | | | | | -50.07 | | -53.94 | | | | | | 72.19 | | -82.69 | |
| | | | | | (-7.98) | | (-8.37) | | | | | | (6.60) | | (-4.48) | |
| SRVSJ | | | | | | -141.69 | -144.75 | | | | | | | -103.72 | -149.56 | |
| | | | | | | (-9.25) | (-9.32) | | | | | | | (-8.25) | (-7.66) | |
| SRVJ | | | | | | | | -81.15 | | | | | | | | -150.59 |
| | | | | | | | | (-10.15) | | | | | | | | (-7.80) |
| RVOL | | | | | | | | | -8.94 | -7.46 | -6.74 | -9.08 | -5.90 | -6.31 | -6.40 | -6.38 |
| | | | | | | | | | (-1.60) | (-1.32) | (-1.21) | (-1.62) | (-1.05) | (-1.12) | (-1.14) | (-1.13) |
| RSK | | | | | | | | | 16.12 | -22.16 | -9.87 | 9.12 | -24.75 | -10.15 | 4.08 | 14.02 |
| | | | | | | | | | (5.59) | (-9.55) | (-9.49) | (3.07) | (-10.41) | (-9.72) | (1.39) | (4.91) |
| RKT | | | | | | | | | -0.68 | -0.68 | 0.46 | -0.68 | 0.12 | 0.09 | 0.08 | 0.09 |
| | | | | | | | | | (-2.25) | (-2.27) | (1.45) | (-2.24) | (0.42) | (0.30) | (0.28) | (0.32) |
| Adjusted $R^2$ | 0.0063 | 0.0033 | 0.0035 | 0.0082 | 0.0005 | 0.0019 | 0.0024 | 0.0016 | 0.0204 | 0.0175 | 0.0185 | 0.0214 | 0.0160 | 0.0168 | 0.0172 | 0.0168 |

# Table 2.14 (Continued)

Panel B: Regressions with Control Variables

| | I | II | III | IV | V | VI | VII | VIII | IX | X | XI | XII | XIII | XIV | XV | XVI |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Intercept | 100.76 | 100.26 | 92.97 | 98.02 | 97.60 | 97.72 | 98.57 | 98.26 | 89.27 | 92.87 | 93.62 | 89.60 | 94.59 | 94.82 | 95.19 | 94.94 |
| | (4.24) | (5.45) | (5.25) | (4.15) | (5.67) | (5.65) | (5.69) | (5.67) | (3.30) | (4.26) | (3.93) | (3.27) | (4.37) | (4.33) | (4.30) | (4.29) |
| RVJP | -30.35 | | | | | | | | -33.59 | | | | | | | |
| | (-3.04) | | | | | | | | (-1.87) | | | | | | | |
| RVJN | 28.77 | | | | | | | | 50.58 | | | | | | | |
| | (3.26) | | | | | | | | (3.48) | | | | | | | |
| RVLJP | | -27.56 | | -27.07 | | | | | | 11.55 | | -28.36 | | | | |
| | | (-4.46) | | (-2.81) | | | | | | (1.20) | | (-1.53) | | | | |
| RVLJN | | 16.42 | | 23.05 | | | | | | -0.32 | | 48.59 | | | | |
| | | (2.67) | | (2.61) | | | | | | (-0.03) | | (2.89) | | | | |
| RVSJP | | | -26.94 | -34.15 | | | | | | | -24.07 | -38.20 | | | | |
| | | | (-2.78) | (-2.51) | | | | | | | (-2.20) | (-2.04) | | | | |
| RVSJN | | | 44.62 | 45.83 | | | | | | | 26.48 | 52.34 | | | | |
| | | | (4.25) | (3.81) | | | | | | | (2.67) | (3.42) | | | | |
| SRVLJ | | | | | -22.63 | | -25.76 | | | | | | 9.90 | | -31.02 | |
| | | | | | (-5.18) | | (-5.67) | | | | | | (1.10) | | (-1.93) | |
| SRVSJ | | | | | | -33.16 | -38.71 | | | | | | | -23.75 | -41.74 | |
| | | | | | | (-3.92) | (-4.45) | | | | | | | (-2.74) | (-2.83) | |
| SRVJ | | | | | | | | -28.64 | | | | | | | | -39.38 |
| | | | | | | | | (-6.26) | | | | | | | | (-2.69) |
| RVOL | | | | | | | | | 4.79 | 5.07 | 4.59 | 4.68 | 4.94 | 4.87 | 4.86 | 4.84 |
| | | | | | | | | | (0.79) | (0.84) | (0.76) | (0.77) | (0.82) | (0.80) | (0.80) | (0.79) |
| RSK | | | | | | | | | 3.02 | -4.90 | -3.58 | 2.67 | -5.53 | -3.67 | 1.42 | 2.46 |
| | | | | | | | | | (1.23) | (-3.01) | (-4.69) | (0.98) | (-3.45) | (-4.79) | (0.50) | (0.95) |
| RKT | | | | | | | | | -0.53 | -0.61 | -0.38 | -0.56 | -0.49 | -0.46 | -0.44 | -0.44 |
| | | | | | | | | | (-2.00) | (-2.27) | (-1.08) | (-1.87) | (-1.78) | (-1.63) | (-1.49) | (-1.46) |
| Beta | -8.28 | -8.11 | -8.09 | -8.27 | -8.07 | -8.26 | -8.29 | -8.16 | -7.71 | -7.75 | -8.18 | -7.72 | -8.10 | -8.18 | -8.22 | -8.14 |
| | (-1.46) | (-1.40) | (-1.38) | (-1.46) | (-1.37) | (-1.41) | (-1.41) | (-1.39) | (-1.36) | (-1.35) | (-1.42) | (-1.37) | (-1.40) | (-1.42) | (-1.43) | (-1.41) |
| log(Size) | -14.93 | -14.77 | -14.73 | -14.94 | -14.76 | -14.76 | -14.67 | -14.67 | -14.36 | -14.25 | -14.29 | -14.48 | -14.13 | -14.11 | -14.04 | -14.05 |
| | (-5.24) | (-5.25) | (-5.21) | (-5.32) | (-5.15) | (-5.14) | (-5.13) | (-5.13) | (-5.19) | (-5.15) | (-5.16) | (-5.26) | (-5.08) | (-5.07) | (-5.06) | (-5.07) |
| BE/ME | -0.76 | -0.75 | -0.65 | -0.76 | -0.67 | -0.61 | -0.59 | -0.62 | -0.72 | -0.57 | -0.57 | -0.69 | -0.56 | -0.54 | -0.56 | -0.58 |
| | (-0.37) | (-0.36) | (-0.32) | (-0.37) | (-0.33) | (-0.30) | (-0.29) | (-0.30) | (-0.34) | (-0.27) | (-0.27) | (-0.33) | (-0.27) | (-0.26) | (-0.27) | (-0.28) |
| MOM | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| | (0.96) | (0.97) | (0.94) | (0.98) | (0.93) | (0.93) | (0.93) | (0.93) | (1.26) | (1.19) | (1.18) | (1.22) | (1.19) | (1.21) | (1.20) | (1.21) |
| REV | -0.01 | -0.01 | -0.01 | -0.01 | -0.01 | -0.01 | -0.01 | -0.01 | -0.01 | -0.01 | -0.01 | -0.01 | -0.01 | -0.01 | -0.01 | -0.01 |
| | (-5.63) | (-5.83) | (-5.63) | (-5.53) | (-5.84) | (-5.66) | (-5.55) | (-5.64) | (-5.74) | (-5.97) | (-5.82) | (-5.70) | (-5.99) | (-5.83) | (-5.77) | (-5.80) |
| IVOL | -301.46 | -293.89 | -297.02 | -304.39 | -291.92 | -292.47 | -298.32 | -298.14 | -425.97 | -416.77 | -422.83 | -424.54 | -419.05 | -421.52 | -424.24 | -425.20 |
| | (-2.18) | (-2.15) | (-2.17) | (-2.21) | (-2.13) | (-2.12) | (-2.16) | (-2.17) | (-4.69) | (-4.57) | (-4.63) | (-4.67) | (-4.59) | (-4.62) | (-4.65) | (-4.67) |
| CSK | -7.52 | -8.67 | -8.13 | -7.29 | -8.69 | -8.23 | -7.62 | -7.87 | -7.00 | -7.84 | -7.44 | -6.89 | -7.91 | -7.45 | -7.35 | -7.52 |
| | (-1.75) | (-2.01) | (-1.89) | (-1.70) | (-2.01) | (-1.91) | (-1.77) | (-1.82) | (-1.66) | (-1.86) | (-1.77) | (-1.64) | (-1.88) | (-1.77) | (-1.74) | (-1.78) |
| CKT | 2.34 | 2.29 | 2.24 | 2.36 | 2.32 | 2.32 | 2.43 | 2.38 | 1.74 | 1.72 | 1.69 | 1.74 | 1.71 | 1.77 | 1.83 | 1.81 |
| | (1.19) | (1.15) | (1.13) | (1.20) | (1.15) | (1.16) | (1.21) | (1.18) | (0.91) | (0.89) | (0.87) | (0.91) | (0.88) | (0.91) | (0.94) | (0.93) |
| MAX | -0.03 | -0.03 | -0.03 | -0.03 | -0.03 | -0.03 | -0.03 | -0.03 | -0.03 | -0.03 | -0.03 | -0.03 | -0.03 | -0.03 | -0.03 | -0.03 |
| | (-5.33) | (-5.66) | (-5.57) | (-5.31) | (-5.64) | (-5.71) | (-5.55) | (-5.54) | (-7.37) | (-7.55) | (-7.45) | (-7.35) | (-7.59) | (-7.55) | (-7.57) | (-7.55) |
| MIN | -0.02 | -0.02 | -0.02 | -0.02 | -0.02 | -0.02 | -0.02 | -0.02 | -0.01 | -0.02 | -0.01 | -0.01 | -0.01 | -0.01 | -0.01 | -0.01 |
| | (-2.79) | (-3.11) | (-3.14) | (-2.79) | (-3.11) | (-3.12) | (-2.83) | (-2.85) | (-2.81) | (-2.94) | (-2.80) | (-2.81) | (-2.88) | (-2.80) | (-2.73) | (-2.75) |
| ILLIQ | -7.84 | -7.68 | -8.08 | -7.86 | -7.99 | -8.03 | -7.87 | -7.88 | -8.94 | -8.87 | -8.68 | -9.10 | -8.54 | -8.51 | -8.41 | -8.43 |
| | (-5.24) | (-5.12) | (-5.22) | (-5.26) | (-5.15) | (-5.16) | (-5.08) | (-5.10) | (-4.79) | (-5.12) | (-4.96) | (-4.87) | (-4.97) | (-4.95) | (-4.86) | (-4.88) |
| Adjusted $R^2$ | 0.0602 | 0.0597 | 0.0597 | 0.0609 | 0.0590 | 0.0592 | 0.0594 | 0.0592 | 0.0647 | 0.0641 | 0.0642 | 0.0652 | 0.0636 | 0.0637 | 0.0639 | 0.0638 |

*Notes: See notes to Tables 2.1 and 2.5. This table reports results for cross-sectional Fama-MacBeth regressions, based on the regression model depicted as equation (2.15) in Section 2.4.5. In these regression models, future weekly returns are regressed on various realized measures and control variates. The two panels utilize jump truncation level $\gamma^2$, as discussed in the footnote to Table 2.2. The regressions that are reported on are of the form: $r_{i,t+1} = \gamma_{0,t} + \sum_{j=1}^{K_1} \gamma_{j,t} X_{i,j,t} + \sum_{s=1}^{K_2} \phi_{s,t} Z_{i,s,t} + \epsilon_{i,t+1}$, $t = 1, ..., T$, where $r_{i,t+1}$ denotes the stock return for firm $i$ in week $t+1$, $K_1$ is the number of potential variation measures, and $X_{i,j,t}$ denotes a relevant realized measure at the end of week $t$. In addition, there are $K_2$ variables measuring firm characteristics, which are denoted by $Z_{i,j,t}$ (see Section 2.3 for details). In the table, time series averages of the coefficient estimates ($\frac{1}{T}\sum_{t=1}^{T} \widehat{\gamma}_{j,t}$ and $\frac{1}{T}\sum_{t=1}^{T} \widehat{\phi}_{j,t}$) are reported, along with Newey-West $t$-statistics (in parentheses). For complete details, see Section 2.4.

Table 2.15: Jumps Associated with (Absolute) Magnitude of Earning Surprises

Panel A: Daily Average Percentage of Firms Exhibiting Various Types of Jumps, on Days Characterized by Earnings Surprises

| A-SUE | RVJP | RVJN | SRVJ | RVLJP | RVLJN | SRVLJ | RVSJP | RVSJN | SRVSJ |
|---|---|---|---|---|---|---|---|---|---|
| Small | 0.8099 | 0.8180 | 0.9849 | 0.1951 | 0.2004 | 0.3042 | 0.7233 | 0.7258 | 0.8983 |
| Medium | 0.8310 | 0.8232 | 0.9841 | 0.2216 | 0.2173 | 0.3289 | 0.7319 | 0.7232 | 0.8928 |
| Large | 0.8605 | 0.8621 | 0.9909 | 0.2618 | 0.2572 | 0.3737 | 0.7455 | 0.7488 | 0.8829 |

Panel B: Daily Average Percentage of Firms Exhibiting Various Types of Jumps, on Days Characterized by No Earnings Surprises

| | RVJP | RVJN | SRVJ | RVLJP | RVLJN | SRVLJ | RVSJP | RVSJN | SRVSJ |
|---|---|---|---|---|---|---|---|---|---|
| | 0.8836 | 0.8786 | 0.9884 | 0.2252 | 0.2220 | 0.3107 | 0.7941 | 0.7900 | 0.9095 |

Panel C: t-Statistics Associated with the Difference in Jump Size Percentages Between Portfolios

| Difference | SRVJ | SRVLJ | SRVSJ |
|---|---|---|---|
| Medium-Small | -0.55 | 4.98 | -1.68 |
| Large-Medium | 5.76 | 9.02 | -3.06 |
| Large-None | 3.54 | 16.85 | -10.85 |

*Notes: See notes to Tables 2.1. Panels A and B of this table report daily average percentages of firms exhibiting various types of jumps, on days with (Panel A) and without (Panel B) earnings surprises. On earning announcement dates for which at least 3 stocks report earning, the "reporting" stocks are sorted into tertile portfolios (called "Small", "Medium", and "Large"), based on the absolute value of earning surprise (A-SUE), where SUE is defined in equation (2.16). Thus, small, medium and large portfolios are only constructed on days for which at least 3 firms are characterized by an earnings surprise. Then, the percentage of firms exhibiting jumps in each of the three earnings surprise size categories is calculated, for various different jump types (i.e., RVJP, RVJN, etc.) Finally, percentages are averages over all reporting days in the sample. Finally, various Newey-West t-statistics measuring the significance of the differences in jump size percentages for SRVJ, SRVLJ, and SRVSJ type jumps are reported in Panel C of the table.

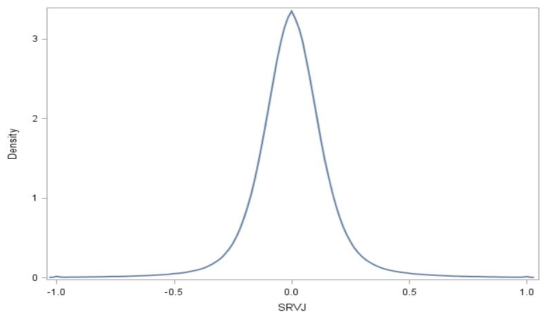
Table 2.16: Fama-MacBeth Type Regressions Using Various Jump Variation Measures as Dependent Variable

| | SRVLJ | SRVLJ | SRVSJ | SRVSJ | SRVJ | SRVJ |
|---|---|---|---|---|---|---|
| | I | II | III | IV | V | VI |
| Intercept | 0.0080 | 0.0176 | 0.0034 | 0.0159 | 0.0115 | 0.0335 |
| | (5.81) | (10.78) | (2.59) | (11.64) | (4.63) | (12.31) |
| RVOL | -0.0050 | 0.0025 | -0.0065 | 0.0006 | -0.0115 | 0.0032 |
| | (-11.02) | (5.36) | (-15.37) | (1.88) | (-14.81) | (4.71) |
| Beta | | 0.0014 | | -0.0011 | | 0.0003 |
| | | (4.40) | | (-3.76) | | (0.79) |
| log(Size) | 0.0013 | 0.0003 | 0.0030 | 0.0017 | 0.0043 | 0.0020 |
| | (5.79) | (1.62) | (15.19) | (9.21) | (12.71) | (6.67) |
| BE/ME | | 0.0007 | | 0.0003 | | 0.0010 |
| | | (4.11) | | (2.72) | | (4.58) |
| MOM | | 0.0007 | | 0.0011 | | 0.0018 |
| | | (3.54) | | (5.94) | | (6.13) |
| REV | 0.25166 | 0.1176 | 0.3172 | 0.1882 | 0.5688 | 0.3058 |
| | (61.05) | (31.43) | (65.55) | (41.52) | (69.66) | (39.62) |
| IVOL | | -0.1424 | | -0.1880 | | -0.3303 |
| | | (-17.05) | | (-23.23) | | (-24.43) |
| CSK | | 0.0142 | | 0.0206 | | 0.0349 |
| | | (19.14) | | (28.33) | | (26.38) |
| CKT | | -0.0008 | | -0.0008 | | -0.0015 |
| | | (-2.28) | | (-1.96) | | (-2.31) |
| MAX | | 0.2456 | | 0.2248 | | 0.4704 |
| | | (43.77) | | (25.28) | | (35.84) |
| MIN | | 0.5115 | | 0.4365 | | 0.9480 |
| | | (54.63) | | (48.47) | | (58.64) |
| ILLIQ | 0.0020 | 0.0016 | 0.0032 | 0.0023 | 0.0052 | 0.0039 |
| | (6.33) | (5.32) | (18.60) | (16.36) | (12.20) | (10.28) |
| Adjusted $R^2$ | 0.0322 | 0.0492 | 0.1070 | 0.1473 | 0.1049 | 0.1517 |

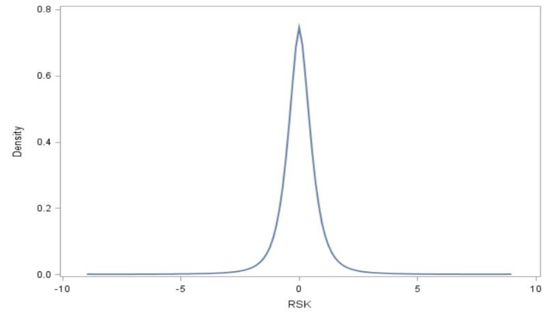
*Notes: See notes to Tables 2.1, 2.5 and 2.14. This table reports results for cross-sectional Fama-MacBeth type regressions using various jump variation measures (listed across the first row of entries in the table) as dependent variables, and for various control variables (listed in the first column of the table). Thus, the regressions in this table mirror those reported in Table 2.14, with one difference. Namely, the dependent variable in the regressions is either SRVLJ, SRVSJ, or SRVJ. Here, SRVLJ and SRVSJ are contructed using jump truncation level $\gamma^2 = 5\sqrt{\frac{1}{t}\widehat{IV}_t^{(i)}}\Delta_n^{0.49}$.

Figure 2.1: Unconditional Distributions of Realized Measures

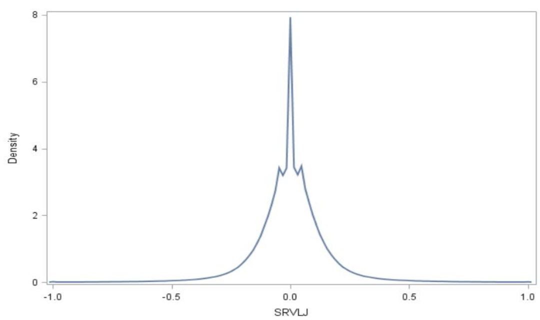
Panel A: SRVJ Kernel Density Estimate



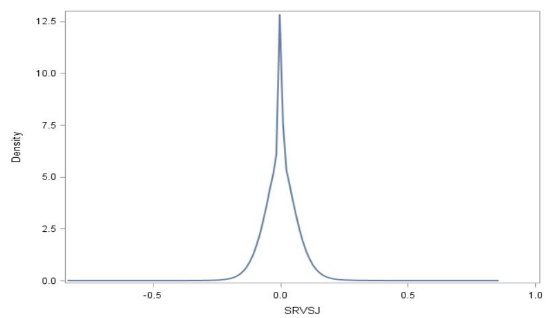

Panel B: RSK Kernel Density Estimate



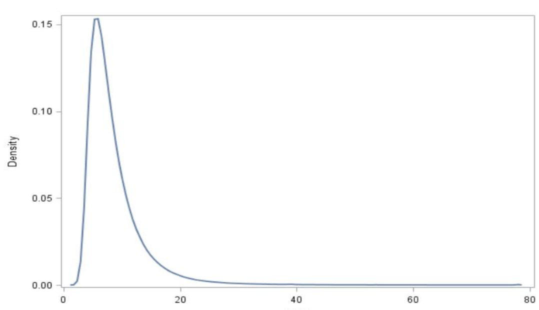

Panel C: SRVLJ Kernel Density Estimate



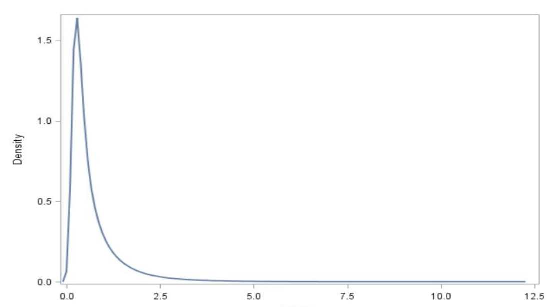

Panel D: SRVSJ Kernel Density Estimate




Panel E: RKT Kernel Density Estimate



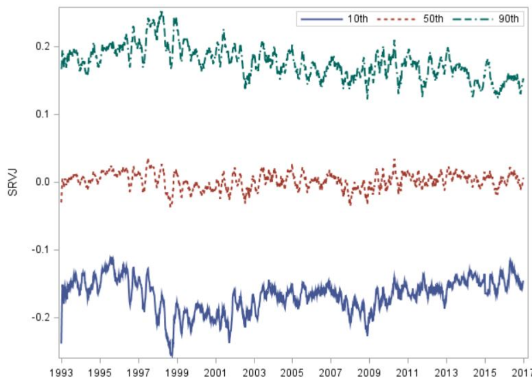

Panel F: RVOL Kernel Density Estimate



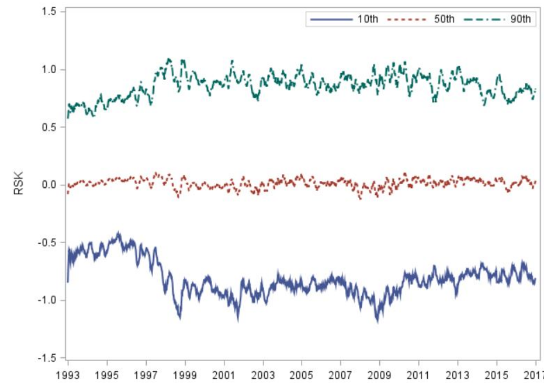

*Notes: See notes to Table 2.1. Panels A-F display unconditional distribution kernel density estimates of various realized measures, for the cross-section of stock returns for the period January 1993 to December 2016. Signed small and large jump variation measures are constructed using truncation levels $\gamma^1 = 4\sqrt{\frac{1}{t}\widehat{IV}_t^{(i)}}\Delta_n^{0.49}$. Distributions are similar when using $\gamma^2 = 5\sqrt{\frac{1}{t}\widehat{IV}_t^{(i)}}\Delta_n^{0.49}$.

Figure 2.2: Percentiles of Realized Measures

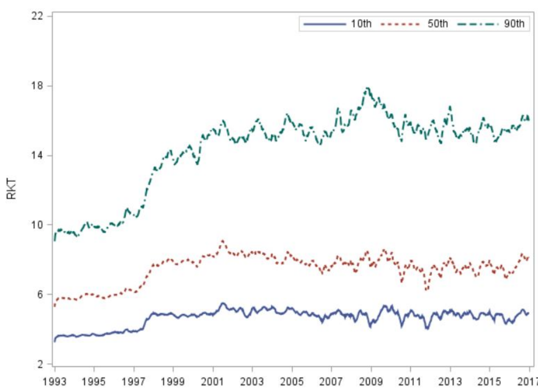
Panel A: Percentiles of SRVJ



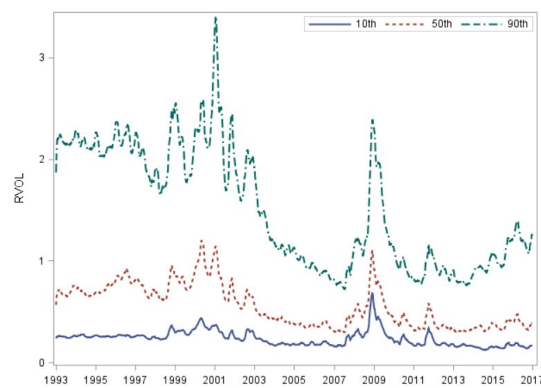

Panel B: Percentiles of RSK



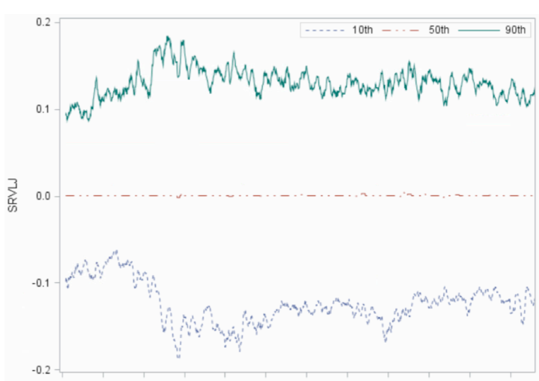

Panel C: Percentiles of RKT



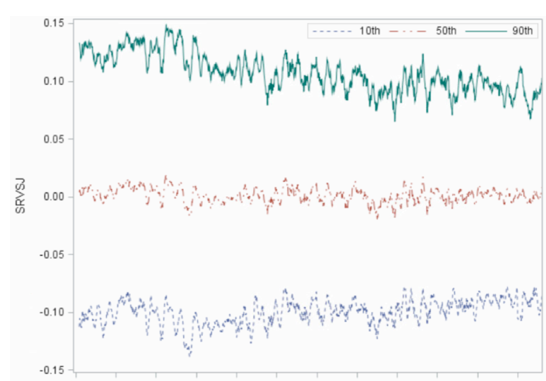

Panel D: Percentiles of RVOL




Panel E: Percentiles of SRVLJ Based on $\gamma^2$



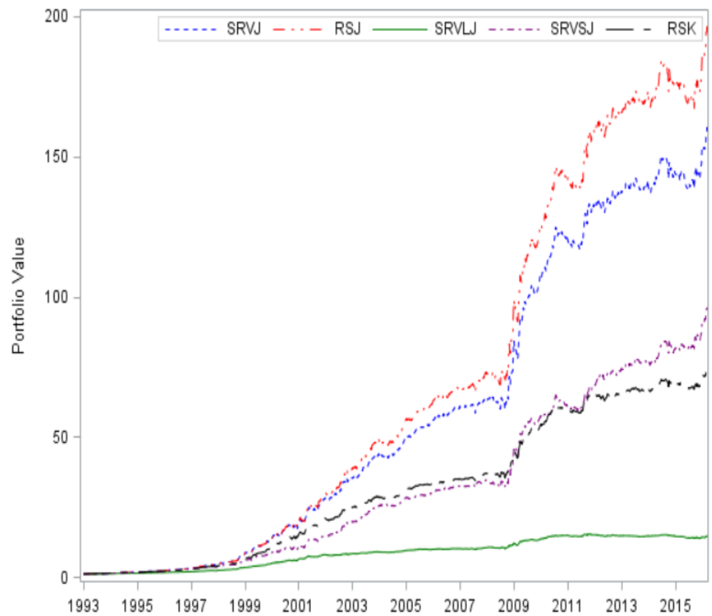

Panel F: Percentiles of SRVSJ Based on $\gamma^2$



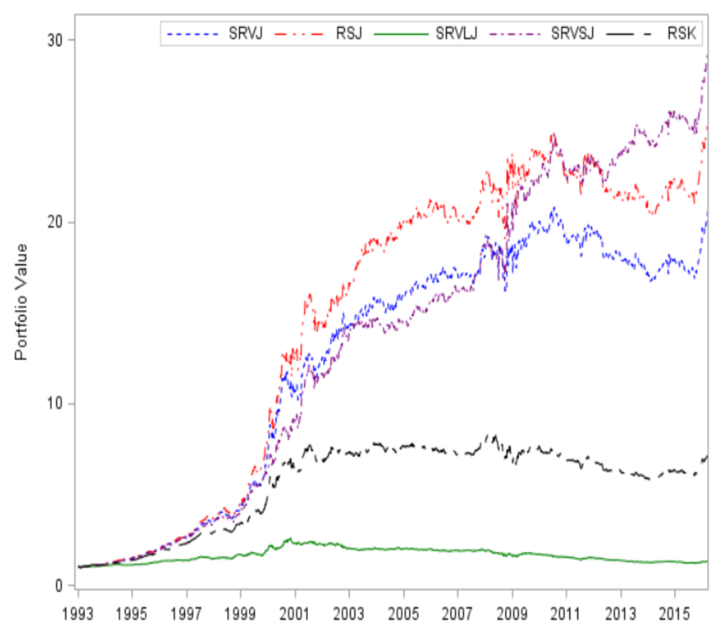

*Notes: See notes to Table 2.1. Panels A-H display 10-week moving averages of percentiles of realized measures, for the cross-section of stocks, for the period January 1993 to December 2016. Signed small and large jump variation measures are contructed based on jump truncation level $\gamma^2 = 5\sqrt{\frac{1}{t}\widehat{IV}_t^{(i)}}\Delta_n^{0.49}$.

Figure 2.3: Cumulative Gains of Short-Long Portfolios

Panel A: Equal-Weighted Mean Return



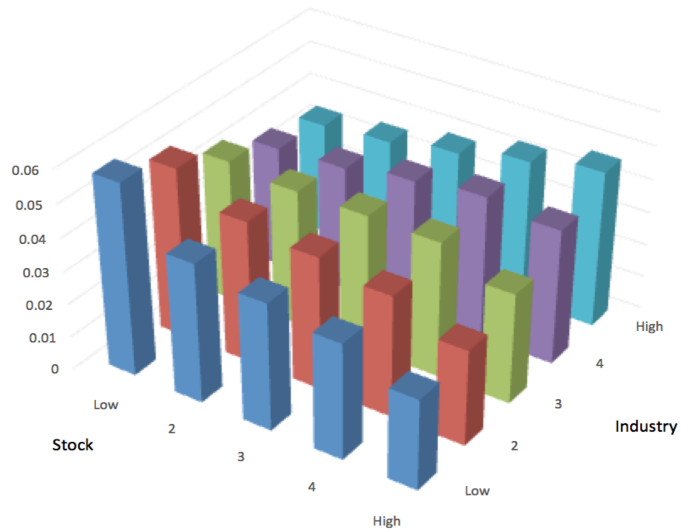

Panel B: Value-Weighted Mean Return



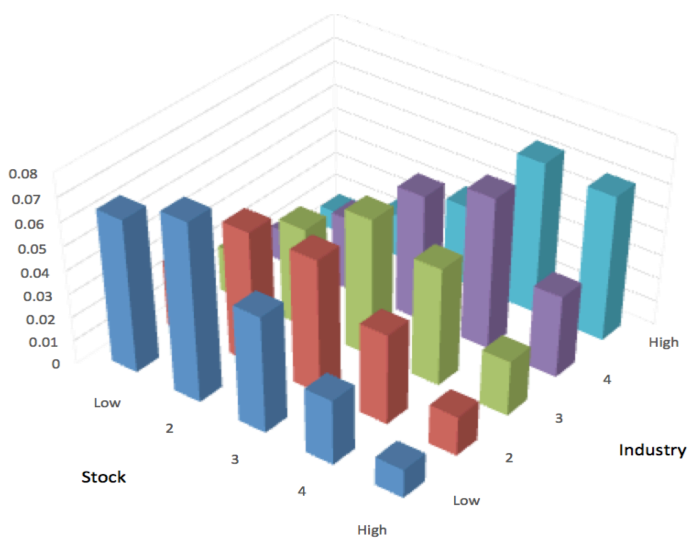

*Notes: Panels A-B display cumulative gains of equal-weighted and value-weighted short-long portfolios constructed using SRVJ, SRVLJ, SRVSJ, and RSK (see Table 2.1 and Section 2.2 for a discussion of these measures). RSJ is the relative signed jump variation measure defined and analyzed in Bollerslev et al. (2019b), who include the risk-free rate in all of their calculations, while we do not (refer to Bollerslev et al. (2019b) for complete details). In all experiments, the initial investment, made on January 1993, is \$1. Each portfolio is re-balanced and accumulated on a weekly basis, through 2016. Signed small and large jump variation measures used in the experiment reported on in this figure are constructed based on truncation level $\gamma^2 = 5\sqrt{\frac{1}{t}\widehat{IV}_t^{(i)}}\Delta_n^{0.49}$. See Section 2.4.2 for further discussion.

Figure 2.4: Distribution of Stocks in Portfolios Formed Based on Stocks' Signed Jump Variation (SRVJ) and Industry Signed Jump Variation

Panel A: Average Distribution of Stocks Across Double-Sorted Portfolios



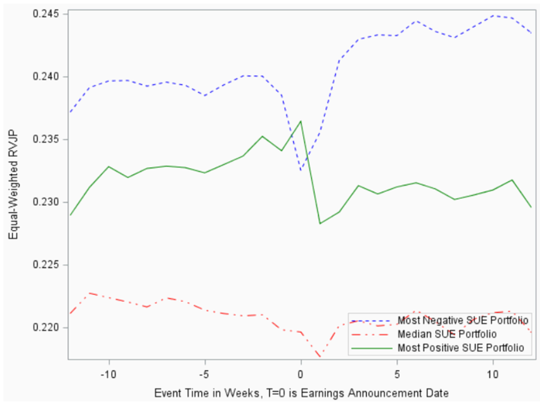

Panel B: Average Distribution of Market Capitalization Across Double-Sorted Portfolios



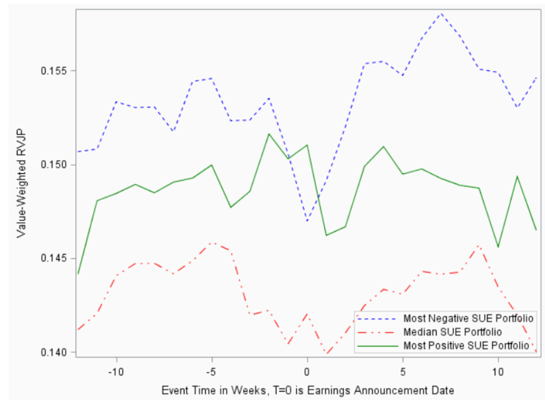

*Notes: See notes to Table 2.13. The vertical axis in Panels A and B measures time series average proportions of stocks and market capitalizations, across double sorted portfolios.

Figure 2.5: Jump Variation Measures Around Earnings Announcement

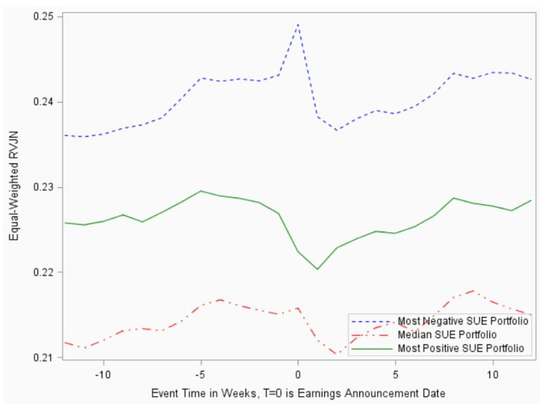
Panel A1: Equal-Weighted RVJP



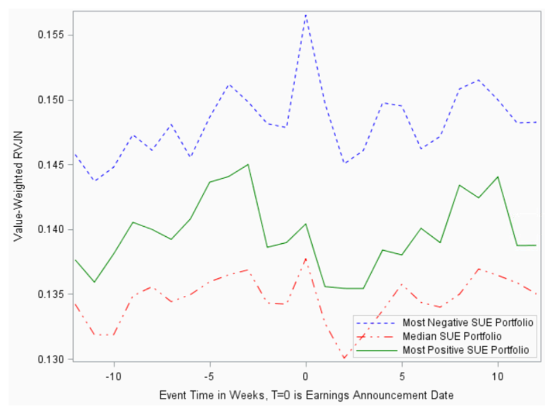

Panel A2: Value-Weighted RVJP



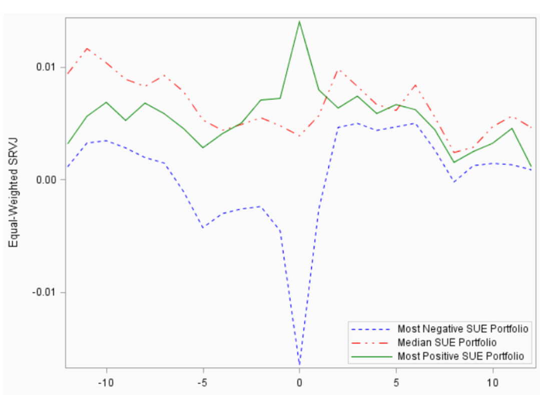

Panel B1: Equal-Weighted RVJN




Panel B2: Value-Weighted RVJN




Panel C1: Equal-Weighted SRVJ




Panel C2: Value-Weighted SRVJ

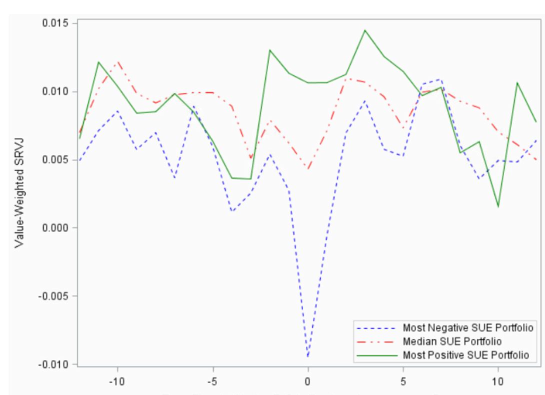
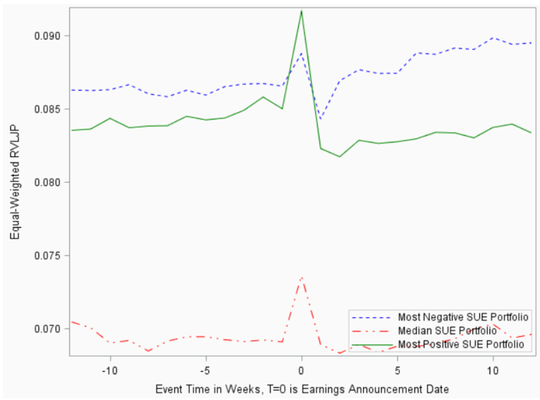
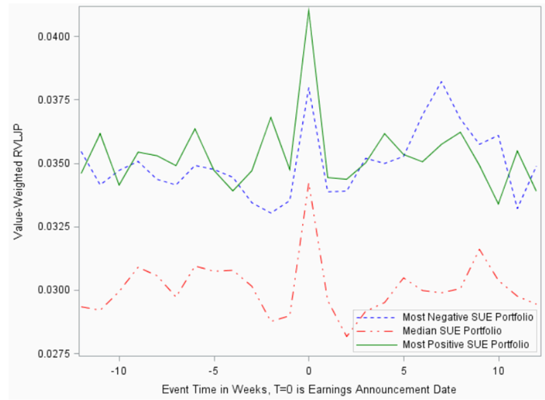


Figure 2.5 (Continued)

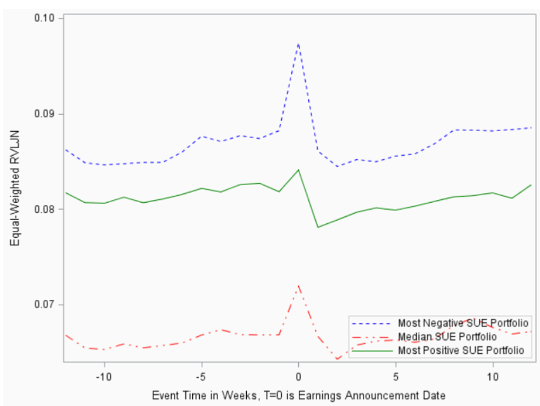
Panel D1: Equal-Weighted RVLJP



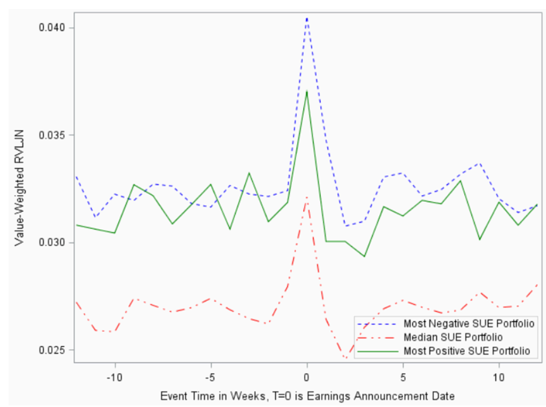

Panel D2: Value-Weighted RVLJP



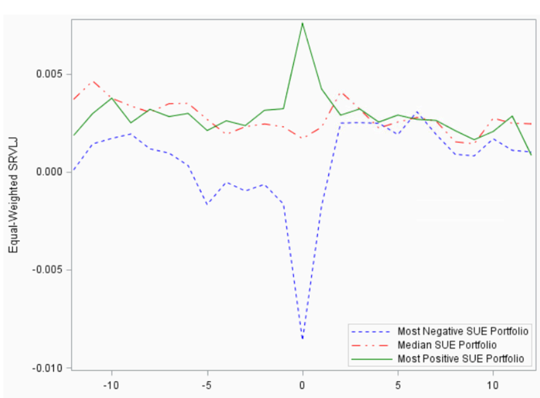

Panel E1: Equal-Weighted RVLJN




Panel E2: Value-Weighted RVLJN




Panel F1: Equal-Weighted SRVLJ




Panel F2: Value-Weighted SRVLJ

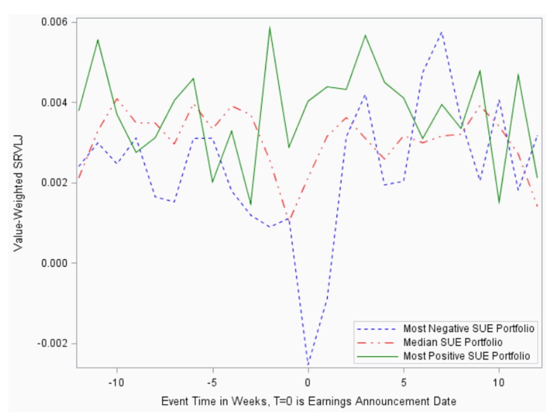
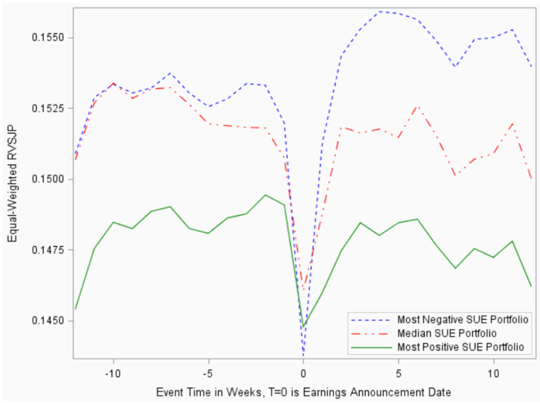
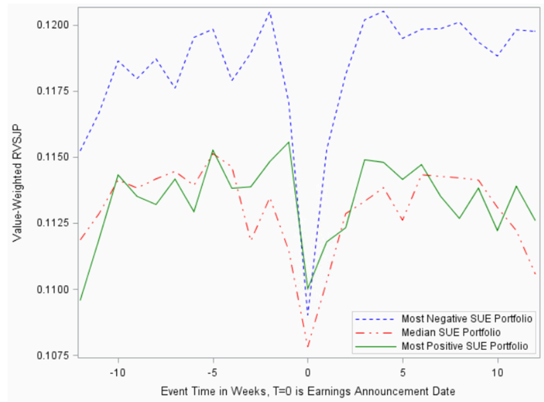


Figure 2.5 (Continued)

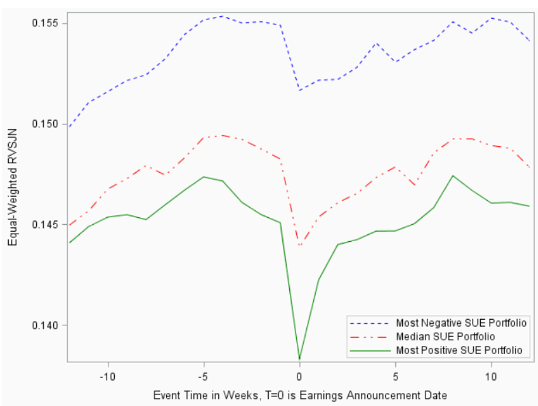
Panel G1: Equal-Weighted RVSJP



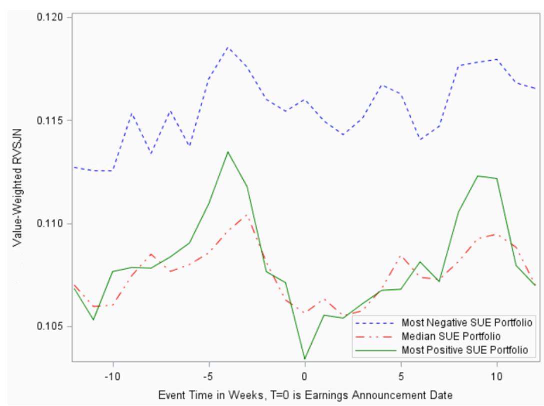

Panel G2: Value-Weighted RVSJP



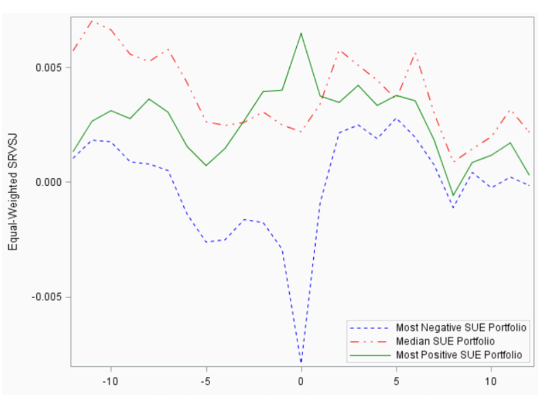

Panel H1: Equal-Weighted RVSJN



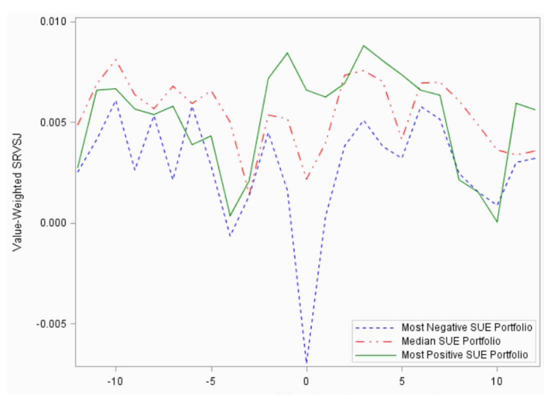

Panel H2: Value-Weighted RVSJN




Panel I1: Equal-Weighted SRVSJ



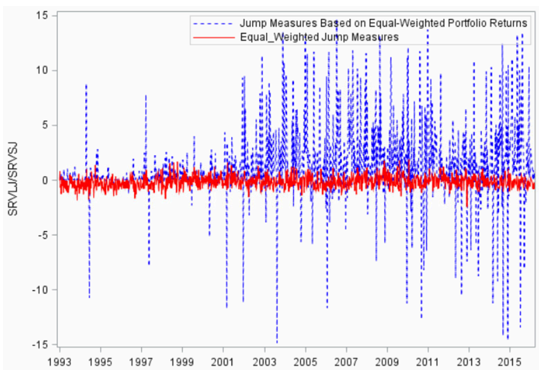

Panel I2: Value-Weighted SRVSJ



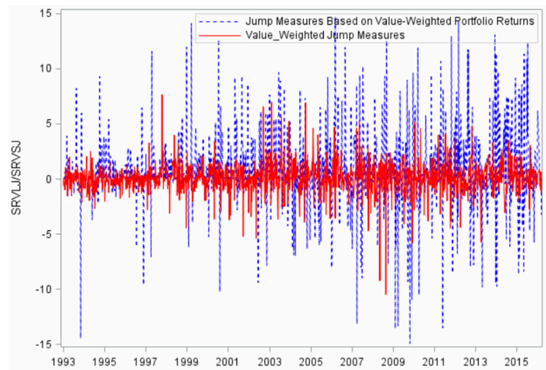

*Notes: See notes to Table 2.1. Panels A-I display equal- or value-weighted averages of various weekly jump variation measures in a [-12, 12] week window around earnings announcement.

Figure 2.6: Aggregated and Weighted Average of Jump Variation Measures

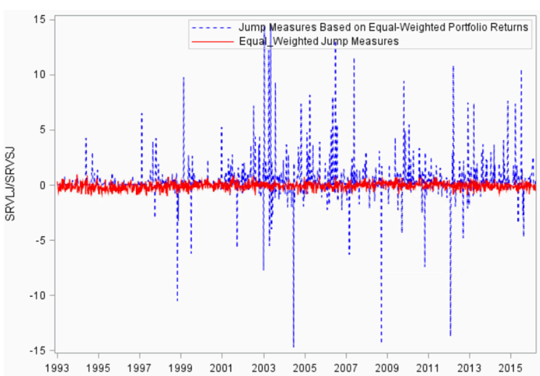
Panel A: Equal-Weighted SRVLJ/SRVSJ (Stocks Sorted by SRVLJ)



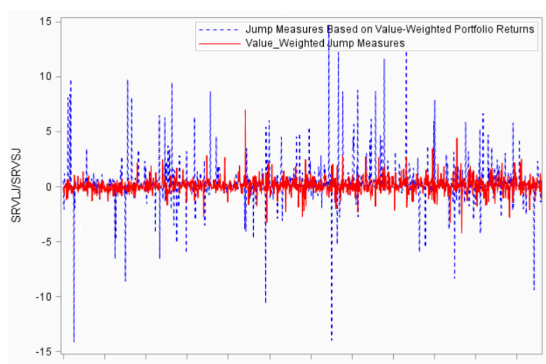

Panel B: Value-Weighted SRVLJ/SRVSJ (Stocks Sorted by SRVLJ)




Panel C: Equal-Weighted SRVLJ/SRVSJ (Stocks Sorted by SRVSJ)




Panel D: Value-Weighted SRVLJ/SRVSJ (Stocks Sorted by SRVSJ)




*Notes: See notes to Table 2.1. Panels A-D display weekly aggregated and weighted averages of the ratio of SRVLJ to SRVSJ for 1st quintile stocks, sorted on SRVLJ to SRVSJ. Aggregated jump measures are depicted in blue (dotted line), and are constructed using 5-minute portfolio returns. Weighted average jump measures are depicted in red (solid line) and are constructed using individual daily jump measures, and then aggregating to weekly. All calculation utilize jump truncation level $\gamma^1 = 4\sqrt{\frac{1}{t}\widehat{IV}_t^{(i)}}\Delta_n^{0.49}$. For complete details, refer to Section 2.4.6.2.

# Chapter 3

# Forecasting Portfolio Variance: A New Decomposition Approach

## 3.1 Introduction

The price movements of an asset depend on the nature of news (sentiment and importance) and on the corresponding information processing mechanism. Different stocks may respond to the same market-level news announcement in different ways (e.g., with upside or downside price drift or with small or large jumps). Thus, when examining the covariance matrix of a portfolio, it is crucial to consider the interactions among different types of price movements, as different components may provide different informational content. In light of this fact, the objective of this paper is to build on the research of Bollerslev et al. (2019a), in which realized covariance matrices are decomposed into constituent variation components. In particular, we analyze both the signs and magnitudes of the underlying high-frequency returns used in the construction of realized covariance matrices. The impetus for our approach is that by including only "information-rich" components in realized (co)-variation forecasting models, predictive accuracy may be improved. Finally, in our prediction experiments, we consider a wide variety of forecasting models, constructed using both standard HAR specifications, as well as various machine learning methods, including: penalized regression type methods such as the least absolute shrinkage operator (LASSO) and the elastic net (EN), and dimension reduction methods such as partial least squares (PLS) and principal components analysis (PCA).

It should be noted that the decomposition of realized components of the covariance matrix used in this paper is closely related to that discussed in recent work on the construction of risk measures based on high-frequency data, including realized skewness and kurtosis (see, e.g. Neuberger (2012) and Amaya et al. (2015)), and jumps (see, e.g. Bollerslev et al. (2019b), Feunou et al. (2018), and Duong and Swanson

(2015)). Moreover, our utilization of various machine learning methods for forecasting financial variables builds on previous research in which stock returns are predicted using shrinkage and variable selection methods (see, e.g., Rapach et al. (2013)), and estimating and testing asset pricing models using dimension reduction methods (see, e.g. Giglio and Xiu (2019) and Kelly et al. (2017)). However, in contrast to these studies, we synthesize machine learning techniques with foreasting portfolio variances using high-frequency based risk measures, including those based on jumps and co-jumps.

Our contributions to the literature are thus twofold. First, investigate the usefulness of a new decomposition that separates the realized covariance matrix of a portfolio into components based on the signs (positive or negative) and magnitudes (continuous, small jump, or large jump) of underlying high-frequency returns. Second, we investigate the importance of sparseness when forecasting covariance matrices using cross section data. This is done by designing forecasting experiments that utilize machine learning, shrinkage, and dimension reduction methods. Our empirical analysis is based on all constituent S&P 500 stocks, for the period January 2005 - December 2013. The candidate predictors used in the specification of our forecasting models include 21 "concordant" and "discordant" variation components measured at daily, weekly, and monthly frequencies, totaling 63 unique variables. We construct portfolios formed using 5 to 200 stocks, based on high-frequency returns at 5-minute, 15-minute, and 30-minute frequencies. Under each of the 117 data frequency, truncation level, and portfolio dimension settings that we analyze, we evaluate the predictive performance of each forecasting model using 200 randomly selected portfolios. Our empirical finding can be summarized as follow.

First, sparsity or parsimony is one of the key factors for improving the portfolio variance forecasting performance. Namely, although each variation component we construct may (in principle) contain marginal predictive content, only a small set of these components actually contain information that is useful for forecasting portfolio variance. This conclusion is predicated on the observation that restricted SCHAR-r models are significantly more accurate than SCHAR models, for example. Additionally, both of these models are highly parsimonious (sparse), and they are our two mean square forecasting error "best" models, when specified using two-step penalized regression in which dimension reduction via the LASSO and EN are carried out in the first step.

Second, co-jumps are the source of all out-of-sample forecasting gains, when sparse models are compared with standard HAR-type models. However, all of our best performing models indicate that negative continuous variation components are the most influential predictors. This does not, of course, mean that jumps do not matter. Indeed, when our variation components are constructed using 5-minute and 15-minute frequency data, the MSFE-best models include both continuous and jump components as predictors. This finding is more pronounced as portfolio dimension increases. Finally, the above findings are less pronounced when 30-minute high-frequency returns are used in covariance matrix estimation. Thus, forecasting gains are driven by the identification of co-jumps through well-diversified portfolios, and the use of higher frequencies of data.

Finally, machine learning methods, including the LASSO and EN provide limited improvement to the out-of-sample fit, unless sparseness is enhanced by using first stage variable selection and dimension reduction when specifying forecasting models. This result further underscores the importance of imposing sparseness, after carrying out decompositions of the variety examined in this paper. Namely, if sparseness is retained by removing predictors associated with extraneous (co)-variation information, in contexts where realized covariance matrices are decomposed into constituent variation components that depend on both the signs and magnitudes of the underlying high-frequency returns, then realized covariance matrices can be more precisely predicted.

The rest of this paper is organized as follows. In Section 3.2 we discuss the model setup and define the jump risk measures that we utilize. Section 3.3 contains a discussion of the data used in our empirical analysis. Section 3.4 presents our main empirical findings, including discussions of results based on panelized regression models, dimension reduction models, and sparse models. Section 3.5 concludes.

## 3.2 Model Setup and Estimation Methodology

### 3.2.1 Components of covariance matrix

Following Aït-Sahalia and Jacod (2012), we assume the d-dimensional log-price process, $X_t = [X_{1,t}, ..., X_{d,t}]^\top$, follows an Itô semimartingaleis, defined as

$$X_t = X_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s + J_t,$$

where $b$ and $\sigma$ denote the drift and diffusive volatility processes, respectively; $W$ is a d-dimensional standard Brownian motion; J denotes the jump part, which can be further decomposed into its small and large components,

$$J_t = \int_0^t \int_{\{|x| \leq \epsilon\}} x(\mu - \nu)(ds, dx) + \int_0^t \int_{\{|x| \geq \epsilon\}} x\mu(ds, dx)$$

where $\mu$ is a random positive measure with its compensator $\nu$; and $\epsilon$ is the (arbitrary) fixed cutoff level (threshold) used to distinguish between large and small jumps. For each trading day t, transaction prices are observed over equally spaced intervals and the corresponding intraday log returns, $r_{t,i} = [r_{1,t,i}, ..., r_{d,t,i}]^\top$, at the $i$th interval $\Delta_n$ are defined as

$$r_{t,i} = X_{i\Delta_n,t} - X_{(i-1)\Delta_n,t},$$

The daily realized covariance matrix is thus defined as

$$RCOV_t = \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} r_{t,i} r_{t,i}^\top.$$

We decompose the covariance matrix into separate components based on the sign and magnitude of the underlying high-frequency returns using thresholding method (see Mancini (2009), Duong and Swanson (2015), Li et al. (2017), and the references cited therein). Let $r_{t,i}^{C^+}$, $r_{t,i}^{C^-}$, $r_{t,i}^{SJ^+}$, $r_{t,i}^{SJ^-}$, $r_{t,i}^{LJ^+}$, and $r_{t,i}^{LJ^-}$ denote the positive continuous, negative continuous, positive small jump, negative small jump, positive large jump, negative large jump return vector, respectively. In particular,

$$\begin{aligned} r_{t,i}^{C^+} &\equiv r_{t,i} \odot I_{t,i}^{C^+} & r_{t,i}^{C^-} &\equiv r_{t,i} \odot I_{t,i}^{C^-} \\ r_{t,i}^{SJ^+} &\equiv r_{t,i} \odot I_{t,i}^{SJ^+} & r_{t,i}^{SJ^-} &\equiv r_{t,i} \odot I_{t,i}^{SJ^-} \\ r_{t,i}^{LJ^+} &\equiv r_{t,i} \odot I_{t,i}^{LJ^+} & r_{t,i}^{LJ^-} &\equiv r_{t,i} \odot I_{t,i}^{LJ^-} \end{aligned} \tag{3.1}$$

where $I_{t,i}^{C+} \equiv [1_{\{0<r_{1,t,i}\leq\alpha\Delta_n^{\varpi}\}}, ..., 1_{\{0<r_{d,t,i}\leq\alpha\Delta_n^{\varpi}\}}]^{\top}$, $I_{t,i}^{LJ+} \equiv [1_{\{r_{1,t,i}>\gamma\}}, ..., 1_{\{r_{d,t,i}>\gamma\}}]^{\top}$, and $I_{t,i}^{SJ+} \equiv [1_{\{\alpha\Delta_n^{\varpi}<r_{1,t,i}\leq\gamma\}}, ..., 1_{\{\alpha\Delta_n^{\varpi}<r_{d,t,i}\leq\gamma\}}]^{\top}$ denote the element-wise indicator functions, with $\alpha\Delta_n^{\varpi}$ and $\gamma$ as the truncation levels to separate jumps from continuous part, and large jumps from small jumps, respectively.[1] $I_{t,i}^{C-}$, $I_{t,i}^{LJ-}$ and $I_{t,i}^{SJ-}$ are defined analogously. The "concordant" semicovariances based on return vectors with same magnitudes are defined as,

$$\begin{aligned}
PC_t &\equiv \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} (r_{t,i}^{C+})(r_{t,i}^{C+})^{\top}, & NC_t &\equiv \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} (r_{t,i}^{C-})(r_{t,i}^{C-})^{\top}, \\
PSJ_t &\equiv \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} (r_{t,i}^{SJ+})(r_{t,i}^{SJ+})^{\top}, & NSJ_t &\equiv \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} (r_{t,i}^{SJ-})(r_{t,i}^{SJ-})^{\top}, \\
PLJ_t &\equiv \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} (r_{t,i}^{LJ+})(r_{t,i}^{LJ+})^{\top}, & NLJ_t &\equiv \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} (r_{t,i}^{LJ-})(r_{t,i}^{LJ-})^{\top},
\end{aligned} \tag{3.2}$$

the "concordant" semicovariances based on return vectors with different magnitudes are defined as,

$$\begin{aligned}
PCSJ_t &\equiv \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} (r_{t,i}^{C+})(r_{t,i}^{SJ+})^{\top}, & NCSJ_t &\equiv \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} (r_{t,i}^{C-})(r_{t,i}^{SJ-})^{\top}, \\
PSJC_t &\equiv \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} (r_{t,i}^{SJ+})(r_{t,i}^{C+})^{\top}, & NSJC_t &\equiv \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} (r_{t,i}^{SJ-})(r_{t,i}^{C-})^{\top}, \\
PCLJ_t &\equiv \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} (r_{t,i}^{C+})(r_{t,i}^{LJ+})^{\top}, & NCLJ_t &\equiv \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} (r_{t,i}^{C-})(r_{t,i}^{LJ-})^{\top}, \\
PLJC_t &\equiv \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} (r_{t,i}^{LJ+})(r_{t,i}^{C+})^{\top}, & NLJC_t &\equiv \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} (r_{t,i}^{LJ-})(r_{t,i}^{C-})^{\top}, \\
PSLJ_t &\equiv \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} (r_{t,i}^{SJ+})(r_{t,i}^{LJ+})^{\top}, & NSLJ_t &\equiv \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} (r_{t,i}^{SJ-})(r_{t,i}^{LJ-})^{\top}, \\
PLSJ_t &\equiv \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} (r_{t,i}^{LJ+})(r_{t,i}^{SJ+})^{\top}, & NLSJ_t &\equiv \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} (r_{t,i}^{LJ-})(r_{t,i}^{SJ-})^{\top},
\end{aligned} \tag{3.3}$$

the "discordant" semicovariances based on return vectors with same magnitudes are

[1] All truncation levels have included time-of-day effects.

defined as,

$$\begin{aligned}
MC_t^+ &\equiv \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} (r_{t,i}^{C^+})(r_{t,i}^{C^-})^\top, & MC_t^- &\equiv \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} (r_{t,i}^{C^-})(r_{t,i}^{C^+})^\top, \\
MSJ_t^+ &\equiv \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} (r_{t,i}^{SJ^+})(r_{t,i}^{SJ^-})^\top, & MSJ_t^- &\equiv \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} (r_{t,i}^{SJ^-})(r_{t,i}^{SJ^+})^\top, \\
MLJ_t^+ &\equiv \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} (r_{t,i}^{LJ^+})(r_{t,i}^{LJ^-})^\top, & MLJ_t^- &\equiv \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} (r_{t,i}^{LJ^-})(r_{t,i}^{LJ^+})^\top,
\end{aligned} \tag{3.4}$$

and the "discordant" semicovariances based on return vectors with different magnitudes are defined as,

$$\begin{aligned}
MCSJ_t^+ &\equiv \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} (r_{t,i}^{C^+})(r_{t,i}^{SJ^-})^\top, & MCSJ_t^- &\equiv \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} (r_{t,i}^{C^-})(r_{t,i}^{SJ^+})^\top, \\
MSJC_t^+ &\equiv \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} (r_{t,i}^{SJ^+})(r_{t,i}^{C^-})^\top, & MSJC_t^- &\equiv \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} (r_{t,i}^{SJ^-})(r_{t,i}^{C^+})^\top, \\
MCLJ_t^+ &\equiv \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} (r_{t,i}^{C^+})(r_{t,i}^{LJ^-})^\top, & MCLJ_t^- &\equiv \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} (r_{t,i}^{C^-})(r_{t,i}^{LJ^+})^\top, \\
MLJC_t^+ &\equiv \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} (r_{t,i}^{LJ^+})(r_{t,i}^{C^-})^\top, & MLJC_t^- &\equiv \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} (r_{t,i}^{LJ^-})(r_{t,i}^{C^+})^\top, \\
MSLJ_t^+ &\equiv \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} (r_{t,i}^{SJ^+})(r_{t,i}^{LJ^-})^\top, & MSLJ_t^- &\equiv \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} (r_{t,i}^{SJ^-})(r_{t,i}^{LJ^+})^\top, \\
MLSJ_t^+ &\equiv \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} (r_{t,i}^{LJ^+})(r_{t,i}^{SJ^-})^\top, & MLSJ_t^- &\equiv \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} (r_{t,i}^{LJ^-})(r_{t,i}^{SJ^+})^\top
\end{aligned} \tag{3.5}$$

Following from the above definitions, the realized covariance matrix equals to the sum of all these "concordant" and "discordant" components. And $PCSJ_t = PSJC_t$, $NCSJ_t = NSJC_t$, $PCLJ_t = PLJC_t$, $NCLJ_t = NLJC_t$, $PSLJ_t = PLSJ_t$, $NSLJ_t = NLSJ_t$ and $MCSJ_t^+ = MSJC_t^-$, $MCSJ_t^- = MSJC_t^+$, $MCLJ_t^+ = MLJC_t^-$, $MCLJ_t^- =$

$MLJC_t^+$, $MSLJ_t^+ = MLSJ_t^-$, $MSLJ_t^- = MLSJ_t^+$, therefore we have several combined components,

$$\begin{aligned}
PCSJ2_t &\equiv PCSJ + PSJC, & NCSJ2_t &\equiv NCSJ_t + NSJC_t,\\
PCLJ2_t &\equiv PCLJ_t + PLJC_t, & NCLJ2_t &\equiv NCLJ_t + NLJC_t,\\
PSLJ2_t &\equiv PSLJ_t + PLSJ_t, & NSLJ2_t &\equiv NSLJ_t + NLSJ_t,\\
MLJ2_t &\equiv MLJ_t^+ + MLJ_t^-, & MSJ2_t &\equiv MSJ_t^+ + MSJ_t^-,\\
MCSJP2_t &\equiv MCSJ_t^+ + MSJC_t^-, & MCSJN2_t &\equiv MCSJ_t^- + MSJC_t^+,\\
MCLJP2_t &\equiv MCLJ_t^+ + MLJC_t^-, & MCLJN2_t &\equiv MCLJ_t^- + MLJC_t^+,\\
MSLJP2_t &\equiv MSLJ_t^+ + MLSJ_t^-, & MSLJN2_t &\equiv MSLJ_t^- + MLSJ_t^+,\\
MC2_t &\equiv MC_t^+ + MC_t^-.
\end{aligned} \tag{3.6}$$

### 3.2.2 Forecasting portfolio variance

Given the high-frequency return vectors of the constituents of a portfolio, the realized covariance matrix of the portfolio can be separated into various "concordant" and "discordant" semivariances based on returns of different signs and magnitudes. For any given portfolio weight vector $w$,

$$\begin{aligned}
RV^p &\equiv w'RCOVw\\
&\equiv w'PV^pw + w'NV^pw + w'MV^pw + w'PVM^pw + w'NVM^pw + w'MVM^pw \\
&\equiv \mathcal{PV}_t^p + \mathcal{NV}_t^p + \mathcal{MV}_t^p + \mathcal{PVM}_t^p + \mathcal{NVM}_t^p + \mathcal{MVM}_t^p
\end{aligned} \tag{3.7}$$

where $PV^p$, $NV^p$, $MV^p$ denote various positive concordant, negative concordant and discordant semivariances constructed by return vectors with same magnitudes, and $PVM^p$, $NVM^p$, $MVM^p$ denote the corresponding semivariances based on return vectors with different magnitudes.

### 3.2.3 Forecasting model comparisons

Our bechmark forecasting model is HAR model of Corsi (2009), in which the one-day-ahead forecast for portfolio variance depends on daily, weekly and monthly lags of portfolio variance,

$$RV^p_{t+1|t} = \theta_0 + \theta_d RV^p_t + \theta_w RV^p_{t-1:t-4} + \theta_m RV^p_{t-5:t-21}. \tag{3.8}$$

We also consider semivariance HAR (SHAR) model of Patton and Sheppard (2015),

$$RV^p_{t+1|t} = \theta_0 + \theta_{d^+}\mathcal{V}^{+p}_t + \theta_{d^-}\mathcal{V}^{-p}_t + \theta_w RV^p_{t-1:t-4} + \theta_m RV^p_{t-5:t-21}. \tag{3.9}$$

where $\mathcal{V}^{+p}_t$ and $\mathcal{V}^{-p}_t$ denote the daily semivariances of the portfolio.

In addition, we also consider the semicovariance HAR (SCHAR) and its restricted version SCHAR-r of Bollerslev et al. (2019a),

$$\begin{aligned} RV^p_{t+1|t} = \theta_0 &+ \theta_{d,\mathcal{P}}\mathcal{P}^p_t + \theta_{w,\mathcal{P}}\mathcal{P}^p_{t-1:t-4} + \theta_{m,\mathcal{P}}\mathcal{P}^p_{t-5:t-21} \\ &+ \theta_{d,\mathcal{N}}\mathcal{N}^p_t + \theta_{w,\mathcal{N}}\mathcal{N}^p_{t-1:t-4} + \theta_{m,\mathcal{N}}\mathcal{N}^p_{t-5:t-21} \\ &+ \theta_{d,\mathcal{M}}\mathcal{M}^p_t + \theta_{w,\mathcal{M}}\mathcal{M}^p_{t-1:t-4} + \theta_{m,\mathcal{M}}\mathcal{M}^p_{t-5:t-21}. \end{aligned} \tag{3.10}$$

The restricted version is constructed as follows,

$$RV^p_{t+1|t} = \theta_0 + \theta_{d,\mathcal{N}}\mathcal{N}^p_t + \theta_{w,\mathcal{N}}\mathcal{N}^p_{t-1:t-4} + \theta_{m,\mathcal{N}}\mathcal{N}^p_{t-5:t-21} + \theta_{m,\mathcal{M}}\mathcal{M}^p_{t-5:t-21}. \tag{3.11}$$

We extend the standard HAR model by incorporating the abovementioned realized components,

$$RV^p_{t+1|t} = \theta_0 + \Theta_d Z^p_t + \Theta_w Z^p_{t-1:t-4} + \Theta_m Z^p_{t-5:t-21}. \tag{3.12}$$

where $Z_t$ denotes a set of daily realized components defined in (3.2)-(3.6).

To estimate these HAR-type models, we use a standard least squares objective function,

$$L(\theta) = \frac{1}{T}\sum_{t=1}^{T}(RV^p_{t+1} - f(\mathcal{Z}^p_t;\theta))^2. \tag{3.13}$$

where $f(\mathcal{Z}^p_t;\theta)$ denotes the predicted realized variance of a portfolio by applying each of the abovementioned models plus machine learning methods with the corresponding predictors up to day t, denoted by $\mathcal{Z}^p_t$. We adopt three criteria to evaluate model performance,

(1) Heteroskedasticity adjusted root mean square error (HARMSE) (see Corsi et al. (2010)),

$$HARMSE = \sqrt{\frac{1}{T}\sum_{t=1}^{T}(\frac{y_t - \hat{y}_t}{y_t})^2},$$

(2) In-sample $R^2$,

(3) Out-of-sample $R^2$ (see Campbell and Thompson (2008)),

$$R^2_{oos} = 1 - \frac{\sum_{t=1}^{T}(y_t - \hat{y}_t)^2}{\sum_{t=1}^{T}(y_t - \bar{y}_t)^2}.$$

For model comparison, we use modified Diebold and Mariano (1995) test following Gu et al. (2019),

$$DM_{12} = \bar{d}_{12}/\hat{\sigma}_{\bar{d}_{12}}$$

$$\bar{d}_{12,t+1} = \frac{1}{n}\sum_{i=1}^{n}((\hat{e}^{(1)}_{i,t+1})^2 - (\hat{e}^{(2)}_{i,t+1})^2)$$

where $\hat{e}^{(1)}_{i,t+1}$ and $\hat{e}^{(2)}_{i,t+1}$ denote the forecasting error for portfolio i at time t based on each model. $\bar{d}_{12}$ and the $\hat{\sigma}_{\bar{d}_{12}}$ is the mean and Newey-West standard error of $d_{12}$ over 200 randomly selected portfolios.

### 3.2.4 Penalized Regression: LASSO and Elastic Net

Arguably, the ultimate goal of regression analysis is to construct a model that can predict well with new data and also to discover variables that contribute to the prediction. When there is a large number of predictors, the least squares method for a linear regression will typically produces non-zero estimates of all parameters, making the interpretation challenging and leading to overfitting as well.

Thus it is crucial to regularize the estimation process by reducing the number of parameters, rendering a parsimonious specification. In contrast to the least-squares estimate, which often has low bias but high variance, penalized regression may improve the prediction accuracy (measured by mean squared error) by introducing some bias but reducing the variance of predicted values (bias-variance trade-off).

Building on simple linear regression models, penalized linear models impose a penalty term on the loss function,

$$\mathcal{L}(\theta; \cdot) = L(\theta) + \Phi(\theta; \cdot), \tag{3.14}$$

where $\Phi(\theta;\cdot)$ is the penalty term, which takes the form,

$$\Phi(\theta;\lambda,\alpha)=\lambda(1-\alpha)\sum_{i=1}^{K}|\theta_i|+\frac{1}{2}\lambda\alpha\sum_{i=1}^{K}\theta_i^2, \tag{3.15}$$

where $\lambda$ and $\alpha$ are two hyperparameters. In this paper, we consider two methods, LASSO and elastic net, corresponding to different values of $\alpha$. When $\alpha = 1$, there is only a $l_2$-penalty on the parameters and this case corresponds to ridge regression, which can shrink all parameters but not set any of them to zero. If $\alpha = 0$, this $l_1$-penalty setting corresponds to the LASSO regression, which can shrink all coefficients and set certain parameters to zero simultaneously. The $\alpha \in (0,1)$ case corresponds to elastic net, which combines characteristics of both the $l_1$ and $l_2$ penalties. Since the main purpose of this paper is to find the most relevant signals for prediction, we focus on LASSO (see Tibshirani (1996)) and elastic net (see Zou and Hastie (2005)) as they can produce simpler models through both shrinkage and variable selection.

Of note is that LASSO tends to randomly select one variable from a group within which variables are correlated.[2] Elastic net is proposed to tackle this problem by assigning similar coefficients to highly correlated variables. Under the assumption that only a small number of predictors are important signals for predicting portfolio realized variance, it's possible that this small set of variables comes from different groups. Thus we adopt both LASSO and elastic net in our variable selection procedure, in hopes that no potential candidate models are missed in this step.

LASSO is applied through a two-stage process, a special case of the relaxed lasso (Meinshausen 2007). As lasso sets a number of the coefficients to be zeros and shrinks the others towards zero relative to the regular least-square estimates, these nonzero estimates by lasso will cause bias towards zero. Relaxed lasso is a method to tackle this issue by separating the two effects of standard lasso (variable selection and shrinkage) into a two-step procedure: a relative large penalty on the full set of variables in the first step for variable selection; a relative small penalty on the selected variables for shrinkage (soft de-biasing). In this paper, we adopt a special case of relaxed lasso, which involves

[2]This is why the variable selection procedure generates a huge quantity of candidate predictor combinations over time and portfolios.

a standard least-squares estimation to the subset of variables selected from the first step, such that the difference in prediction performance between penalized regression and sparse models is mainly due to the selection of predictors. To be consistent, we utilize the same two-stage process for elastic net models.

### 3.2.5 Dimension Reduction: PLS and PCR

When regressors are highly correlated, OLS may result in unstable and unreliable estimates. Penalized linear regressions are one of the possible remedies for multicollinearity by imposing constraints on the magnitudes of parameters. But such models can lead to suboptimal prediction performance especially when input data contains a lot of redundant information.[3] To tackle this issue, we can apply dimension reduction methods that utilize derived mutually orthogonal components as new regressors. Generally, dimension reduction models involve a two-step procedure. First, they produce a number of linear combinations of the original variables. Next, the first few components which can explain most of the variability in independent or dependent variables are used in a regular regression for prediction. Two commonly used methods in this domain are principal component regression (PCR) (see, e.g., Stock and Watson (2002a,b, 2006), and Bai and Ng (2006a,b, 2008)) and partial least squares (PLS) (see, e.g., Kelly and Pruitt (2013, 2015)).

PCR transforms the original T $\times$ K input data matrix $\mathcal{Z}^p$ into a set of derived covariates named principal components based on the singular-value decomposition (SVD) of $\mathcal{Z}^p$. Specifically, $\mathcal{Z}^p = USV^\top$, where $S_{T\times K} = diag[\delta_1, ..., \delta_K]$ with $\delta_K \geq \cdots \geq \delta_1 \geq 0$ denotes the non-negative singular values of $\mathcal{Z}^p$ and $U_{T\times T} = [u_1, ..., u_T]$ and $V_{K\times K} = [v_1, ..., v_K]$ are left and right singular vectors of $\mathcal{Z}^p$ respectively. Columns $u_i (i = 1, ..., T)$ and $v_j (j = 1, ..., K)$ are orthogonal unit vectors with length T and K, respectively. Thus $\mathcal{T}^p_{T\times K} = \mathcal{Z}^p V$ defines the full principal component decomposition of $\mathcal{Z}^p$. The leading L ($L < K$) principal components, corresponding to the first L largest singular values (the

[3] As documented by Tibshirani (1996), ridge regression empirically dominates lasso in terms of forecasting performance when regressors are correlated, indicates that a linear combination of all the original input variables can better represent the dependent variable relative to a subset of the redundant variables. This is consistent with the idea of dimension reduction methods.

squared root of eigenvalues) and eigenvectors, are used as new predictors in a second step of regression, defined as

$$RV^p = \theta_0' + \Theta' \mathcal{T}_L^p = \theta_0' + \Theta'(\mathcal{Z}^p V_L), \tag{3.16}$$

The estimation procedure of PCR is to find a set of K-dimensional vectors $[v_1, ..., v_L]$ such that each derived principal component successively retains the maximum possible variation in $\mathcal{Z}^p$. Thus the $j^{th}$ vector of weights satisfies

$$v_j = \arg\max_v \mathrm{Var}(\mathcal{T}) = \arg\max_v \mathrm{Var}(\mathcal{Z}^p V),\ s.t. v'v = 1,\ v_j' v_l = 0,\ l = 1, ..., j-1. \tag{3.17}$$

The limitation of PCR is that it only considers the variability in the original predictors when constructing orthogonal principal components, thus it may omit information that would be useful in predicting the response variable. In contrast to PCR, PLS takes into account both the independent and the response variables in the dimension reduction procedure. The $j^{th}$ vector of weights ($W = [w_1, ..., w_K]$) used to construct component of PLS satisfies [4]

$$w_j = \arg\max_w \mathrm{Cov}(\mathcal{Z}^p W, RV^p)^2, \quad s.t.\ w'w = 1,\ \ w_j' w_l = 0,\ \ l = 1, ..., j-1. \tag{3.18}$$

PCR and PLS transform the original space of K variables into a new space of K uncorrelated variables and achieve dimension reduction by discarding the last K-L components corresponding to the last few eigenvalues. In contrast to the unsupervised method of PCR, PLS is applied in a supervised way with consideration of the correlation between independent and dependent variables. However, in the case of low signal-to-noise ratio, irrelevant predictors can still get some weights in the first L linear combinations rather than being eliminated completely from all components, making the prediction performance contaminated by noise.

### 3.2.6 Sparse Models

Though we can decompose realized covariance into multiple components in hopes that each of them has unique information, we still assume that only a small number of these

[4]This problem can be solved in an efficient way by using SIMPLS algorithm of De Jong (1993).

separated realized measures plays an important role in predicting portfolio realized variance. Thus one of the main purposes of this paper is to discover the most relevant signals by exploiting sparsity in models for prediction.

The construction of candidate sparse models relies on the variable selection procedure. In the first step, all of the 63 predictors are included in the penalized regression for estimating portfolio variance in the subsequent day, operated on a rolling window scheme. All models are re-estimated daily using the most recent 1000 daily observations, and predictor sets with less than 10 variables selected by either LASSO or elastic net are saved as candidate predictor combinations. In the second step, a regular linear regression with predictors selected from the first step is performed as a candidate sparse model.

The above procedure is performed under different settings related to data frequency, truncation levels, and the number of stocks used to construct a portfolio. Specifically, the construction and estimation steps are based on (1) 3 data frequencies (5-minute, 15-minute, and 30-minute); (2) 3 truncation levels used to separate large jumps from small jumps ($\gamma^1$, $\gamma^2$, and $\gamma^3$);[5] (3) 13 kinds of portfolios which are constructed by different number of stocks (N $\in$ [5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 150, 200]). Thus there are 117 (3*3*13) unique settings and the detailed procedure is as follows:

In the first step of variable selection, penalized regressions are operated under each abovementioned setting for 100 randomly selected portfolios and all resulted variable combinations are collected for later steps. Usually there are around 80,000 different candidate models (variable combinations) for each of the 117 settings. However, a large quantity of these candidate models are not robust over time and in the cross-section of portfolios. An intermediate step, in which most unstable models are filtered out, is necessary to make the estimation process efficient. Thus before the next step of model selection, the forecasting performance of each candidate model is calculated based on 30 randomly selected portfolios. Any model with average statistics worse than benchmark models is discarded. Typically there are hundreds of models left for each of the 117

[5] $\gamma^1 = 4\sqrt{\frac{1}{t}\widehat{IV}_t\Delta_n^{0.49}}$ , $\gamma^2 = 5\sqrt{\frac{1}{t}\widehat{IV}_t\Delta_n^{0.49}}$ , and $\gamma^3 = 6\sqrt{\frac{1}{t}\widehat{IV}_t\Delta_n^{0.49}}$ .

settings, making the following steps less time-consuming. In the next step of model selection, the comparison of prediction performance of all benchmark and candidate models is conducted based on 200 randomly selected portfolios under each setting.

## 3.3 Data

We obtain high frequency trading data for S&P 500 constituents from the consolidated Trade and Quote (TAQ) database. The sample period is from January, 2005 to December 2013, for a total of 2265 trading days. We subsample the data at 5-minute, 15-minute, and 30-minute frequencies using previous tick approach. Intraday prices are sampled from 9:30 a.m. to 4:00 p.m. from Monday to Friday. Overnight returns are not considered in this paper.

## 3.4 Empirical Results

### 3.4.1 Prediction Performance

Table 3.1-3.3 report the prediction performance of models in terms of HARMSE, in-sample and out-of-sample $R^2$ based on 3 data frequencies and 3 truncation levels. There are 23 models presented in each panel, including 4 benchmark models (HAR, SHAR, SCHAR, and SCHAR-r), 3 penalized linear regressions (LASSO, elastic net with $\alpha = 0.2$ and $\alpha = 0.6$, denoted EN1, EN2), PLS with number of components based on 3 criteria (which select the leading components that can cumulatively explain 90%, 80% and 70% of the variance in the dependent variable, respectively, and drop any selected component if the corresponding marginal contribution is less than 5% of the variance; denoted PLS1, PLS2, PLS3), PCR using principal components based on 3 criteria (which choose the leading components that can cumulatively retain 90%, 80% and 70% of the variability in independent variables, respectively, and discard any chosen component with a marginal contribution less than 5% of the variance; denoted by PCR1, PCR2, PCR3),[6]

[6]We also used cross-validation to determine the number of components for both PLS and PCR, the results are qualitatively the same as those in Table 3.1-3.3.

10 best performing sparse models that outperform all benchmark models (M1,..., M10).[7]

Panel A of Table 3.1 presents the average HARMSE, $R^2_{is}$ and $R^2_{oos}$ based on 200 randomly selected portfolios formed by different number of stocks for each model. Among 4 benchmark models, SHAR and SCHAR are dominated by HAR and SCHAR-r in terms of $R^2_{oos}$. Consistent with Bollerslev, Li, Patton, and Quaedvlieg (2019), the restricted SCHAR-r model outperform standard HAR model in terms of predictive accuracy measured by all three evaluation estimators across all portfolio dimensions (from N=5 to N=200), indicating that realized semicovariances provide additional information for improving prediction performance. However, it's the restricted SCHAR-r model rather than the unrestricted SCHAR model produces out-of-sample forecasting improvement, suggesting that many of the realized semicovariances contain irrelevant or redundant information for prediction.

Panelized regression models (LASSO, EN1, and EN2) clearly exhibit overfitting as the out-of-sample $R^2_{oos}$ are much smaller than the in-sample $R^2_{is}$. It is not surprising as these models assign non-zeros values to the coefficients of some irrelevant or redundant variables. Figure 3.1 displays the model complexity of each model on each re-estimation day. For portfolios with a small dimension (formed by 10 stocks), LASSO or elastic net usually select over 10 variables as predictors. This number is between 5 and 20 if portfolios are formed by 200 stocks and before 2013. There is a sharp increase in the number of variables after 2013 for well-diversified portfolios, indicating an increase in the number of reliable features and the benefit of utilizing the identified common factors in the cross-section.

Dimension reduction models (PLS1-3 and PCR1-3) improve the prediction performance in terms of out-of-sample $R^2_{oos}$ relative to penalized regressions. This result further confirms the assumption that many of realized measures inside the covariance matrix are redundant. Linear combinations (components) utilized by PLS or PCR can average out some of the noise. In addition, the number of components used by PLS and PCR is much fewer than the number of features selected by LASSO or elastic net,

[7] When data frequency is 15-minute or 30-minute and the number of stocks used for forming a portfolio is small, there are none or less than 10 models that can beat all the benchmark models.

with this number ranging from 1 to 5. These simple model settings also reduce overfitting for portfolios in different dimensions. However, PLS and PCR only discard the last few components to achieve the goal of dimension reduction, rather than throwing away unnecessary variables, making the leading components contaminated by noisy variables. Thus it's not surprising that dimension reduction models only perform on par with those benchmark models. This result suggests that it is crucial to filter out those relevant signals and build parsimonious models for better prediction.

In Table 3.1-3.3, M1-M10 denote the top 10 sparse models that can outperform all those 4 benchmark models in terms of all evaluation criteria. Each panel reports the average HARMSE, $R^2_{is}$, and $R^2_{oos}$ over 200 randomly selected portfolios formed by different number of stocks. M1-M10 denote the same models in each panel, but the model specifications are varying across different settings/panels. Panel A of Table 3.1 corresponds to a setting of 5-minute data and a small truncation level $\gamma^1$. Under this setting, all top models perform similarly better than benchmark models, with an improvement of 3% relative to standard HAR model (2% relative to SCHAR-r model) in terms of $R^2_{oos}$ for small portfolios constructed by 5 stocks. This out-of-sample prediction improvement increases to 5% relative to the performance of HAR model (2% relative to SCHAR-r model) for portfolios formed by 200 stocks. Table 3.4-3.6 report the predictors utilized by these best-performing sparse models. For small portfolios, the number of predictors is less than 5 regardless of the setting of data frequency and truncation level. When portfolios are large, with number of stocks greater than 30, more predictors (always less than 10) are included in these sparse models. This result indicates that both diversification effect and those additional predictors contribute to the improvement in prediction performance.

Figure 3.2 reports the median value of all evaluation criteria for all benchmark models, elastic net with $\alpha = 0.2$ (EN1), PLS with components that can explain 90% of the variation in dependent variable (PLS1), PCR with principal components that account for 90% of the variability in independent variables (PCR1), the sparse model, together with the 10% and the 90% quantiles for the top sparse models over 200 randomly selected portfolios with dimensions ranging from N=5 to N=200. Sparse models generate a substantial improvement in prediction over all the other models as the portfolio size increases. One exception is that elastic net has the largest in-sample $R^2_{is}$ for portfolios formed by 20 or more stocks, which further confirms the existence of the overfitting problem in penalized regressions.

Of note is that when portfolios are small (N is less than 50), there is a large dispersion of prediction performance among the 10%, median, and 90% quantiles for sparse models. The potential reason is that when portfolios are not well-diversified, detected jumps are most likely to be idiosyncratic jumps, which may affect the prediction to different extent for different portfolios. However, for portfolios with large dimensions, most of the idiosyncratic jumps are diversified out, and the remainings tend to be co-jumps in the cross-section, leading to an improvement in the prediction performance. This result further confirms that co-jumps are critical in forecasting portfolio variance.

To complement Table 3.1-3.3, which only report the quantitative prediction performance of all models, Table 3.7-3.9 show the statistical significance of differences among models. We report the pairwise Diebold-Mariano test statistics. The Diebold-Mariano test compares the forecast accuracy of two forecasting models and the null hypothesis is that two models have the same forecast accuracy. The corresponding test statistic is asymptotically $N$ (0,1) distributed under the null hypothesis. Thus a negative statistic indicates that the column model is dominated by the row model. Regardless of data frequency and portfolio dimension, penalized regressions perform poorly compared to benchmark models. Dimension reduction models significantly outperform the over-parameterized SCHAR model, while there is no significant differences between PLS/PCR and HAR/SCHAR-r models. In contrast, sparse models improve the out-of-sample performance over all benchmark models. Except at 30-minute frequency and for some small portfolios at 15-minute frequency, sparse models produce statistically significant improvement over all benchmark, penalized regressions, and dimension reduction models.

### 3.4.2 Variable Importance

In this section, we investigate the predictors utilized in best-performing sparse models. Table 3.4-3.6 report the selected variables in those top 5 sparse models (M1-M5) under different settings. Regardless of data frequency, truncation level and portfolio dimension, daily, weekly and monthly negative continuous components (dNC, wNC, and mNC) are three predictors appeared in most of the top sparse models.

At the fastest frequency (5-minute), top sparse models also use several components based on the interactions between continuous and jump returns, including positive continuous and small jump part (PCSJ), positive continuous and large jump part (PCLJ), and negative continuous and small jump part (NCSJ). When the portfolio dimension increases to a certain level (N is larger than 50 in Table 3.4), more jump related components are included in those top models, such as positive small and large jumps part (PSLJ), negative small and large jump part (NSLJ), and discordant components based on small and/or large jumps (MSLJN, MSLJP, etc.), indicating that co-jump related components are more likely to be identified through constructing diversified portfolios and such measures play an important role in predicting portfolio variance. Of note is that all these jump related components are weekly or monthly aggregated measures, except daily NCSJ for portfolios formed by 200 stocks. This result suggests that it is hard to identify co-jumps using high-frequency data at fast frequencies due to issues of asynchronicity and microstructure noise. While weekly and monthly jump measures aggregate otherwise weak signals embedded in daily jumps.

Based on 15-minute data and besides those three important continuous components (daily,weekly, and monthly negative continuous parts), top sparse models exploit several jump related measures, including concordant measures (PCSJ, NCSJ, PSLJ, NSLJ, NSJ, and NLJ, etc.) and discordant measures (MLJ, MCLJ, and MCSJ, etc.). For portfolios with small dimensions (N=5 and N=10), jump related measures are almost all weekly and/or monthly aggregations. However, daily jump related measures become influential predictors for portfolios with large size (N is larger than 30 in Table 3.5), indicating that co-jumps can be identified when building well-diversified portfolios using

15-minute data.

Table 3.6 reports the selected variables for top sparse models basde on 30-minute data. At this coarse frequency, there are fewer sparse models that can outperform all benchmark models in terms of all evaluation criteria. When portfolios are in small dimensions (N is less than 50), there are none such sparse models. Though the selected jump related variables are similar to those at higher frequencies, most of them are weekly or monthly aggregations. This is not surprising as it is difficult to detect jumps at such coarse frequency, making aggregated jump measures contain relative more information than rarely detected daily jump measures.

Figure 3.4-3.6 complement Table 3.4-3.6 by showing the variable importance for penalized regression, dimension reduction and sparse models. Specifically, we want to identify covariance components that have an significant influence on predicting portfolio variance while simultaneously controlling other predictors. Following Kelly et al. (2017), we rank each separated covariance measure by the corresponding variable importance, denoted by $VI_i$, which is defined as the reduction in the forecasting $R^2_{is}$ from setting the values of feature i to zeros and keeping the remaining forecasting model fixed.

The most influential predictors are based on negative continuous returns (dNC, wNC, and mNC), an universal agreement among all models. Besides these important negative continuous component, elastic net tends to place similar weights on correlated jump measures, making the number of selected variables much larger than that in sparse models. This is also why penalized regressions are susceptible to overfitting. In contrast to penalized regressions, dimension reduction models place emphasis on a smaller set of variables, including continuous and jump components. However, PLS and PCR set nonzero weights on the other variables, making predictions contaminated by noise. Sparse models exploit much fewer predictors than the other models, extracting predictive information from the most relevant and influential continuous and co-jump components.

### 3.4.3 The Effects of Data Frequency and Truncation Level on Prediction

The construction of realized covariance matrix and the corresponding separated components is based on large dimensional high-frequency datasets. To alleviate the problems of asynchronicity and microstructure noise associated with high-frequency data, we subsample the original dataset at different frequencies using the previous tick approach, including 5-minute, 15-minute, and 30-minute. In this section, we examine the effect of subsampling on prediction performance.

As discussed in Section 3.4.2, the appropriate data frequency to detect co-jumps is at 15-minute. Sparse models can thus draw information directly from daily jump measures (e.g., NSLJ etc.) for prediction. Though there are concerns of asynchronicity and microstructure noise at higher frequency (5-minute), sparse models can alleviate such concerns by using a small set of predictors, among which weekly and monthly jump measures aggregate otherwise incomplete information from daily co-jump measures. Thus top sparse models at 5-minute frequency significantly outperform all the other models in terms of out-of-sample forecasting performance. While at a coarse frequency (30-minute), jumps are less likely to be identified, potentially the reason why sparse models produce indistinguishable improvement over benchmark models. In summary, data frequency affects the detection of jumps, which may contain critical information for prediction, making sparse models not attractive at certain frequencies. This result further confirms that co-jumps are key factors that contribute to the outperformance of sparse models, not only due to the sparseness itself.

For robustness, we use three truncation levels to separate large jumps from small jumps. Table 3.4-3.6 show that predictors of top sparse models, especially jump related measures, are different when applying different truncation levels. In fact, the difference is mainly due to the change in definition of a certain range of returns. A large jump related component when using a small truncation level is actually measuring similar parts of the covariance matrix as does a small jump related component when using a large truncation level. One example is that a discordant component that based on negative small and large jump returns (NSLJ) is an important predictor when using the small truncation level $\gamma^1$ at 15-minute frequency, the corresponding predictor turns into a component based only on negative small jump returns (NSJ) if a larger truncation level $\gamma^3$ is used. Though the selected predictors may be denoted differently, they are providing similar information related to prediction, thus the performance of top sparse models are indistinguishable under different settings regarding truncation level. While it is still necessary to separate a jump into its small and large components as each of them and the interactions between them among different stocks can provide unique information.

## 3.5 Concluding Remarks

We propose a new decomposition approach of the covariance matrix of a portfolio, building on the work of Bollerslev et al. (2019a). The decomposition utilizes information on the sign and magnitude of the underlying high-frequency returns. In this decomposition, interactions among stocks related to negative/positive continuous variation components, as well as small/large jump components, yield a number of new predictors, many of which are found to contain unique information that has marginal predictive content for future portfolio variances. This finding is dependent upon the construction of sparse models that utilize only relevant signals, and drop “noisy” variables. More specifically, our findings are predicated upon the judicious use of machine learning, shrinkage and dimension reduction methods when specifying alternatives to standard HAR-type prediction models. Finally, it is noteworthy that predictive gains are most pronounced when higher frequency data is used in our empirical experiments, and when portfolios with greater numbers of stocks are examined.

# Table 3.1: Prediction Performance (5-minute)

Panel A: Separating Large Jumps from Small Jumps Based on $\gamma^1$

| Model | HAR | SHAR | SCHAR | SCHAR-r | LASSO | EN1 | EN2 | PLS1 | PLS2 | PLS3 | PCR1 | PCR2 | PCR3 | M1 | M2 | M3 | M4 | M5 | M6 | M7 | M8 | M9 | M10 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Number of Stocks: N=5 | | | | | | | | | | | | | | | | | | | | | | | |
| HARMSE | 0.85 | 0.81 | 0.80 | 0.74 | 0.83 | 0.84 | 0.83 | 0.71 | 0.71 | 0.73 | 0.77 | 0.76 | 0.75 | 0.72 | 0.73 | 0.74 | 0.71 | 0.71 | 0.71 | 0.71 | 0.72 | 0.73 | 0.73 |
| $R_{is}$ | 0.55 | 0.58 | 0.60 | 0.58 | 0.65 | 0.65 | 0.65 | 0.55 | 0.56 | 0.58 | 0.56 | 0.55 | 0.55 | 0.59 | 0.59 | 0.59 | 0.59 | 0.59 | 0.60 | 0.60 | 0.60 | 0.60 | 0.60 |
| $R_{oos}$ | 0.61 | 0.60 | 0.53 | 0.62 | 0.50 | 0.49 | 0.49 | 0.62 | 0.59 | 0.58 | 0.63 | 0.62 | 0.62 | 0.64 | 0.64 | 0.64 | 0.64 | 0.64 | 0.64 | 0.64 | 0.64 | 0.64 | 0.64 |
| Number of Stocks: N=10 | | | | | | | | | | | | | | | | | | | | | | | |
| HARMSE | 0.88 | 0.81 | 0.80 | 0.73 | 0.83 | 0.83 | 0.83 | 0.72 | 0.70 | 0.72 | 0.77 | 0.77 | 0.75 | 0.72 | 0.71 | 0.72 | 0.71 | 0.72 | 0.72 | 0.70 | 0.72 | 0.72 | 0.71 |
| $R_{is}$ | 0.57 | 0.61 | 0.63 | 0.60 | 0.67 | 0.67 | 0.67 | 0.57 | 0.58 | 0.60 | 0.58 | 0.57 | 0.57 | 0.62 | 0.62 | 0.62 | 0.62 | 0.62 | 0.62 | 0.62 | 0.62 | 0.62 | 0.62 |
| $R_{oos}$ | 0.65 | 0.63 | 0.57 | 0.66 | 0.52 | 0.52 | 0.52 | 0.64 | 0.63 | 0.63 | 0.65 | 0.64 | 0.64 | 0.68 | 0.68 | 0.68 | 0.68 | 0.68 | 0.68 | 0.68 | 0.68 | 0.68 | 0.68 |
| Number of Stocks: N=30 | | | | | | | | | | | | | | | | | | | | | | | |
| HARMSE | 1.00 | 0.89 | 0.85 | 0.79 | 0.88 | 0.88 | 0.88 | 0.78 | 0.72 | 0.75 | 0.84 | 0.84 | 0.83 | 0.76 | 0.76 | 0.76 | 0.77 | 0.77 | 0.75 | 0.76 | 0.78 | 0.79 | 0.76 |
| $R_{is}$ | 0.58 | 0.63 | 0.65 | 0.61 | 0.68 | 0.68 | 0.68 | 0.57 | 0.59 | 0.60 | 0.58 | 0.57 | 0.57 | 0.63 | 0.63 | 0.63 | 0.63 | 0.63 | 0.63 | 0.63 | 0.63 | 0.63 | 0.63 |
| $R_{oos}$ | 0.65 | 0.63 | 0.57 | 0.67 | 0.54 | 0.53 | 0.52 | 0.64 | 0.64 | 0.65 | 0.65 | 0.63 | 0.63 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 |
| Number of Stocks: N=50 | | | | | | | | | | | | | | | | | | | | | | | |
| HARMSE | 1.04 | 0.92 | 0.88 | 0.81 | 0.87 | 0.88 | 0.87 | 0.80 | 0.73 | 0.77 | 0.87 | 0.87 | 0.86 | 0.78 | 0.79 | 0.78 | 0.79 | 0.77 | 0.78 | 0.79 | 0.77 | 0.80 | 0.79 |
| $R_{is}$ | 0.58 | 0.63 | 0.65 | 0.61 | 0.68 | 0.68 | 0.68 | 0.57 | 0.59 | 0.60 | 0.58 | 0.57 | 0.57 | 0.63 | 0.63 | 0.63 | 0.63 | 0.63 | 0.63 | 0.63 | 0.63 | 0.63 | 0.63 |
| $R_{oos}$ | 0.64 | 0.63 | 0.56 | 0.67 | 0.56 | 0.55 | 0.55 | 0.63 | 0.64 | 0.65 | 0.65 | 0.63 | 0.63 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 |
| Number of Stocks: N=100 | | | | | | | | | | | | | | | | | | | | | | | |
| HARMSE | 1.08 | 0.95 | 0.89 | 0.84 | 0.88 | 0.89 | 0.89 | 0.83 | 0.74 | 0.80 | 0.89 | 0.90 | 0.88 | 0.80 | 0.79 | 0.80 | 0.81 | 0.78 | 0.74 | 0.82 | 0.82 | 0.82 | 0.80 |
| $R_{is}$ | 0.58 | 0.63 | 0.65 | 0.61 | 0.68 | 0.68 | 0.68 | 0.57 | 0.58 | 0.60 | 0.58 | 0.57 | 0.57 | 0.63 | 0.63 | 0.63 | 0.63 | 0.63 | 0.63 | 0.63 | 0.63 | 0.63 | 0.63 |
| $R_{oos}$ | 0.64 | 0.63 | 0.56 | 0.67 | 0.56 | 0.54 | 0.54 | 0.63 | 0.64 | 0.66 | 0.64 | 0.63 | 0.63 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 |
| Number of Stocks: N=200 | | | | | | | | | | | | | | | | | | | | | | | |
| HARMSE | 1.11 | 0.97 | 0.90 | 0.85 | 0.89 | 0.88 | 0.89 | 0.84 | 0.74 | 0.82 | 0.91 | 0.91 | 0.90 | 0.83 | 0.74 | 0.81 | 0.81 | 0.80 | 0.79 | 0.82 | 0.79 | 0.83 | 0.85 |
| $R_{is}$ | 0.58 | 0.63 | 0.65 | 0.61 | 0.68 | 0.68 | 0.68 | 0.57 | 0.59 | 0.61 | 0.58 | 0.57 | 0.57 | 0.63 | 0.63 | 0.63 | 0.63 | 0.63 | 0.63 | 0.63 | 0.63 | 0.63 | 0.63 |
| $R_{oos}$ | 0.64 | 0.63 | 0.55 | 0.67 | 0.58 | 0.51 | 0.55 | 0.63 | 0.64 | 0.66 | 0.64 | 0.62 | 0.62 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 |

Panel B: Separating Large Jumps from Small Jumps Based on $\gamma^2$

| Model | HAR | SHAR | SCHAR | SCHAR-r | LASSO | EN1 | EN2 | PLS1 | PLS2 | PLS3 | PCR1 | PCR2 | PCR3 | M1 | M2 | M3 | M4 | M5 | M6 | M7 | M8 | M9 | M10 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Number of Stocks: N=5 | | | | | | | | | | | | | | | | | | | | | | | |
| HARMSE | 0.85 | 0.81 | 0.80 | 0.74 | 0.85 | 0.85 | 0.85 | 0.70 | 0.70 | 0.72 | 0.75 | 0.75 | 0.74 | 0.71 | 0.72 | 0.73 | 0.72 | 0.70 | 0.72 | 0.73 | 0.72 | 0.70 | 0.72 |
| $R_{is}$ | 0.56 | 0.59 | 0.60 | 0.58 | 0.65 | 0.65 | 0.65 | 0.55 | 0.56 | 0.58 | 0.57 | 0.56 | 0.55 | 0.60 | 0.60 | 0.60 | 0.60 | 0.60 | 0.60 | 0.60 | 0.60 | 0.60 | 0.60 |
| $R_{oos}$ | 0.62 | 0.60 | 0.55 | 0.62 | 0.49 | 0.49 | 0.49 | 0.63 | 0.60 | 0.59 | 0.63 | 0.63 | 0.62 | 0.65 | 0.64 | 0.64 | 0.64 | 0.64 | 0.64 | 0.64 | 0.64 | 0.64 | 0.64 |
| Number of Stocks: N=10 | | | | | | | | | | | | | | | | | | | | | | | |
| HARMSE | 0.88 | 0.81 | 0.81 | 0.73 | 0.83 | 0.83 | 0.83 | 0.71 | 0.69 | 0.71 | 0.76 | 0.77 | 0.74 | 0.71 | 0.72 | 0.72 | 0.71 | 0.72 | 0.72 | 0.70 | 0.72 | 0.72 | 0.72 |
| $R_{is}$ | 0.57 | 0.61 | 0.63 | 0.60 | 0.66 | 0.66 | 0.66 | 0.56 | 0.58 | 0.59 | 0.57 | 0.57 | 0.56 | 0.62 | 0.61 | 0.62 | 0.62 | 0.62 | 0.62 | 0.62 | 0.61 | 0.62 | 0.62 |
| $R_{oos}$ | 0.64 | 0.63 | 0.57 | 0.66 | 0.51 | 0.50 | 0.51 | 0.64 | 0.63 | 0.63 | 0.65 | 0.64 | 0.64 | 0.68 | 0.68 | 0.68 | 0.68 | 0.68 | 0.68 | 0.68 | 0.68 | 0.67 | 0.67 |
| Number of Stocks: N=30 | | | | | | | | | | | | | | | | | | | | | | | |
| HARMSE | 1.00 | 0.89 | 0.86 | 0.79 | 0.85 | 0.85 | 0.85 | 0.78 | 0.71 | 0.76 | 0.84 | 0.84 | 0.82 | 0.78 | 0.76 | 0.76 | 0.77 | 0.77 | 0.77 | 0.77 | 0.76 | 0.78 | 0.76 |
| $R_{is}$ | 0.58 | 0.63 | 0.64 | 0.61 | 0.67 | 0.67 | 0.67 | 0.57 | 0.58 | 0.60 | 0.58 | 0.57 | 0.57 | 0.63 | 0.63 | 0.63 | 0.63 | 0.63 | 0.63 | 0.63 | 0.63 | 0.63 | 0.63 |
| $R_{oos}$ | 0.64 | 0.63 | 0.57 | 0.67 | 0.55 | 0.55 | 0.53 | 0.64 | 0.64 | 0.65 | 0.65 | 0.63 | 0.63 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 |
| Number of Stocks: N=50 | | | | | | | | | | | | | | | | | | | | | | | |
| HARMSE | 1.04 | 0.92 | 0.87 | 0.81 | 0.85 | 0.85 | 0.86 | 0.80 | 0.72 | 0.77 | 0.86 | 0.86 | 0.84 | 0.80 | 0.80 | 0.79 | 0.78 | 0.78 | 0.80 | 0.79 | 0.79 | 0.77 | 0.79 |
| $R_{is}$ | 0.58 | 0.63 | 0.65 | 0.61 | 0.67 | 0.67 | 0.67 | 0.57 | 0.58 | 0.60 | 0.58 | 0.57 | 0.57 | 0.63 | 0.63 | 0.63 | 0.63 | 0.63 | 0.63 | 0.63 | 0.63 | 0.63 | 0.63 |
| $R_{oos}$ | 0.64 | 0.63 | 0.56 | 0.67 | 0.56 | 0.55 | 0.55 | 0.63 | 0.64 | 0.65 | 0.65 | 0.63 | 0.63 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 |
| Number of Stocks: N=100 | | | | | | | | | | | | | | | | | | | | | | | |
| HARMSE | 1.08 | 0.95 | 0.89 | 0.84 | 0.87 | 0.88 | 0.88 | 0.83 | 0.73 | 0.80 | 0.90 | 0.90 | 0.87 | 0.81 | 0.82 | 0.82 | 0.82 | 0.78 | 0.82 | 0.81 | 0.82 | 0.80 | 0.81 |
| $R_{is}$ | 0.58 | 0.63 | 0.65 | 0.61 | 0.67 | 0.67 | 0.67 | 0.57 | 0.58 | 0.60 | 0.58 | 0.57 | 0.57 | 0.63 | 0.63 | 0.63 | 0.63 | 0.64 | 0.63 | 0.63 | 0.63 | 0.63 | 0.63 |
| $R_{oos}$ | 0.64 | 0.63 | 0.56 | 0.67 | 0.58 | 0.57 | 0.57 | 0.63 | 0.64 | 0.65 | 0.64 | 0.63 | 0.63 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 |
| Number of Stocks: N=200 | | | | | | | | | | | | | | | | | | | | | | | |
| HARMSE | 1.11 | 0.97 | 0.90 | 0.85 | 0.85 | 0.86 | 0.86 | 0.84 | 0.73 | 0.84 | 0.91 | 0.91 | 0.88 | 0.79 | 0.82 | 0.83 | 0.81 | 0.80 | 0.84 | 0.83 | 0.83 | 0.83 | 0.83 |
| $R_{is}$ | 0.58 | 0.63 | 0.65 | 0.61 | 0.66 | 0.66 | 0.66 | 0.57 | 0.58 | 0.61 | 0.58 | 0.57 | 0.57 | 0.64 | 0.63 | 0.63 | 0.64 | 0.64 | 0.64 | 0.63 | 0.64 | 0.63 | 0.63 |
| $R_{oos}$ | 0.64 | 0.63 | 0.56 | 0.67 | 0.60 | 0.58 | 0.58 | 0.63 | 0.64 | 0.66 | 0.64 | 0.62 | 0.62 | 0.70 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 |

# Table 3.1 (Continued)

Panel C: Separating Large Jumps from Small Jumps Based on $\gamma^3$

| Model | HAR | SHAR | SCHAR | SCHAR-r | LASSO | EN1 | EN2 | PLS1 | PLS2 | PLS3 | PCR1 | PCR2 | PCR3 | M1 | M2 | M3 | M4 | M5 | M6 | M7 | M8 | M9 | M10 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Number of Stocks: N=5 | | | | | | | | | | | | | | | | | | | | | | | |
| HARMSE | 0.84 | 0.80 | 0.80 | 0.73 | 0.86 | 0.86 | 0.86 | 0.71 | 0.70 | 0.72 | 0.76 | 0.75 | 0.74 | 0.72 | 0.72 | 0.72 | 0.72 | 0.72 | 0.72 | 0.72 | 0.72 | 0.72 | 0.70 |
| $R_{is}$ | 0.56 | 0.59 | 0.61 | 0.58 | 0.65 | 0.65 | 0.66 | 0.55 | 0.56 | 0.58 | 0.57 | 0.56 | 0.55 | 0.60 | 0.60 | 0.60 | 0.60 | 0.60 | 0.60 | 0.60 | 0.60 | 0.60 | 0.60 |
| $R_{oos}$ | 0.62 | 0.60 | 0.55 | 0.63 | 0.44 | 0.45 | 0.43 | 0.63 | 0.60 | 0.59 | 0.63 | 0.63 | 0.63 | 0.65 | 0.65 | 0.65 | 0.65 | 0.65 | 0.65 | 0.65 | 0.65 | 0.65 | 0.64 |
| Number of Stocks: N=10 | | | | | | | | | | | | | | | | | | | | | | | |
| HARMSE | 0.88 | 0.81 | 0.81 | 0.73 | 0.83 | 0.84 | 0.83 | 0.72 | 0.69 | 0.72 | 0.77 | 0.77 | 0.75 | 0.72 | 0.72 | 0.73 | 0.72 | 0.71 | 0.73 | 0.70 | 0.73 | 0.69 | 0.72 |
| $R_{is}$ | 0.57 | 0.61 | 0.63 | 0.60 | 0.67 | 0.67 | 0.67 | 0.57 | 0.58 | 0.59 | 0.58 | 0.57 | 0.56 | 0.62 | 0.62 | 0.62 | 0.62 | 0.62 | 0.62 | 0.62 | 0.62 | 0.62 | 0.62 |
| $R_{oos}$ | 0.65 | 0.63 | 0.57 | 0.66 | 0.52 | 0.48 | 0.52 | 0.65 | 0.63 | 0.63 | 0.65 | 0.65 | 0.64 | 0.68 | 0.68 | 0.68 | 0.68 | 0.68 | 0.68 | 0.68 | 0.68 | 0.68 | 0.68 |
| Number of Stocks: N=30 | | | | | | | | | | | | | | | | | | | | | | | |
| HARMSE | 1.00 | 0.89 | 0.85 | 0.78 | 0.86 | 0.87 | 0.87 | 0.77 | 0.71 | 0.75 | 0.83 | 0.84 | 0.81 | 0.76 | 0.76 | 0.77 | 0.77 | 0.75 | 0.77 | 0.72 | 0.72 | 0.77 | 0.75 |
| $R_{is}$ | 0.58 | 0.63 | 0.64 | 0.61 | 0.67 | 0.67 | 0.67 | 0.57 | 0.58 | 0.60 | 0.58 | 0.57 | 0.57 | 0.63 | 0.63 | 0.63 | 0.63 | 0.63 | 0.63 | 0.63 | 0.63 | 0.63 | 0.63 |
| $R_{oos}$ | 0.64 | 0.63 | 0.56 | 0.67 | 0.53 | 0.52 | 0.52 | 0.63 | 0.64 | 0.65 | 0.64 | 0.63 | 0.63 | 0.69 | 0.69 | 0.69 | 0.69 | 0.68 | 0.68 | 0.68 | 0.68 | 0.68 | 0.68 |
| Number of Stocks: N=50 | | | | | | | | | | | | | | | | | | | | | | | |
| HARMSE | 1.05 | 0.92 | 0.88 | 0.81 | 0.89 | 0.89 | 0.89 | 0.80 | 0.72 | 0.78 | 0.87 | 0.87 | 0.84 | 0.78 | 0.73 | 0.78 | 0.80 | 0.79 | 0.78 | 0.80 | 0.80 | 0.81 | 0.73 |
| $R_{is}$ | 0.58 | 0.63 | 0.65 | 0.61 | 0.67 | 0.67 | 0.67 | 0.57 | 0.58 | 0.60 | 0.58 | 0.57 | 0.57 | 0.63 | 0.63 | 0.63 | 0.63 | 0.63 | 0.63 | 0.63 | 0.63 | 0.63 | 0.63 |
| $R_{oos}$ | 0.64 | 0.63 | 0.56 | 0.67 | 0.54 | 0.53 | 0.53 | 0.63 | 0.64 | 0.65 | 0.64 | 0.63 | 0.63 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 |
| Number of Stocks: N=100 | | | | | | | | | | | | | | | | | | | | | | | |
| HARMSE | 1.08 | 0.95 | 0.89 | 0.84 | 0.88 | 0.89 | 0.88 | 0.83 | 0.74 | 0.81 | 0.90 | 0.90 | 0.86 | 0.75 | 0.80 | 0.76 | 0.82 | 0.82 | 0.78 | 0.80 | 0.81 | 0.82 | 0.83 |
| $R_{is}$ | 0.58 | 0.63 | 0.65 | 0.62 | 0.67 | 0.67 | 0.67 | 0.57 | 0.58 | 0.61 | 0.58 | 0.57 | 0.57 | 0.63 | 0.63 | 0.63 | 0.63 | 0.63 | 0.63 | 0.63 | 0.63 | 0.63 | 0.63 |
| $R_{oos}$ | 0.64 | 0.63 | 0.56 | 0.67 | 0.53 | 0.53 | 0.53 | 0.63 | 0.64 | 0.66 | 0.64 | 0.63 | 0.63 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 |
| Number of Stocks: N=200 | | | | | | | | | | | | | | | | | | | | | | | |
| HARMSE | 1.11 | 0.97 | 0.90 | 0.85 | 0.85 | 0.86 | 0.86 | 0.84 | 0.74 | 0.84 | 0.91 | 0.91 | 0.87 | 0.80 | 0.82 | 0.79 | 0.75 | 0.77 | 0.82 | 0.80 | 0.84 | 0.82 | 0.79 |
| $R_{is}$ | 0.58 | 0.63 | 0.65 | 0.62 | 0.66 | 0.66 | 0.66 | 0.57 | 0.58 | 0.61 | 0.58 | 0.57 | 0.57 | 0.64 | 0.64 | 0.64 | 0.63 | 0.63 | 0.63 | 0.64 | 0.64 | 0.63 | 0.63 |
| $R_{oos}$ | 0.64 | 0.63 | 0.56 | 0.67 | 0.57 | 0.55 | 0.55 | 0.63 | 0.64 | 0.66 | 0.64 | 0.62 | 0.62 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 |

*Notes: This table reports the in-sample and out-of-sample portfolio volatility forecasting performance of benchmark models (HAR, SHAR, SCHAR, and SCHAR-r), LASSO, Elastic Net with $\alpha = 0.2$ and $\alpha = 0.6$ (EN1-EN2), partial least squares using components that can explain 90%, 80%,70% of variance in the response variable (PLS1-PLS3), principal component regression using components that can explain 90%, 80%,70% of variability in explanatory variables (PCR1-PCR3), and the top 10 models in terms of HARMSE, $R^2_{is}$,$R^2_{oos}$ using predictors selected by LASSO or elastic net (M1-M10), at the 5-minute data frequency. The reported numbers are based on 200 randomly selected portfolios constructed by 5, 10, 30, 50, 100, and 200 stocks, respectively, therefore all statistics are the average value over time and all randomly selected portfolios. The truncation level used to separate jumps and continuous variation is $3\sqrt{\frac{1}{t}\widehat{IV}_t}\Delta_n^{0.49}$ . $\gamma^1$, $\gamma^2$, and $\gamma^3$ (in the form of $\alpha\sqrt{\frac{1}{t}\widehat{IV}_t}\Delta_n^{0.49}$ ) are three truncation levels used to split jump variation into large and small components, with $\alpha = 4$, 5, and 6, respectively.

# Table 3.2: Prediction Performance (15-minute)

Panel A: Separating Large Jumps from Small Jumps Based on $\gamma^1$

| Model | HAR | SHAR | SCHAR | SCHAR-r | LASSO | EN1 | EN2 | PLS1 | PLS2 | PLS3 | PCR1 | PCR2 | PCR3 | M1 | M2 | M3 | M4 | M5 | M6 | M7 | M8 | M9 | M10 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Number of Stocks: N=5 | | | | | | | | | | | | | | | | | | | | | | | |
| HARMSE | 1.24 | 1.20 | 1.13 | 1.08 | 1.30 | 1.30 | 1.29 | 1.05 | 1.06 | 1.07 | 1.10 | 1.09 | 1.10 | 1.06 | 1.06 | 1.05 | 1.06 | 1.06 | 1.07 | 1.07 | | | |
| $R_{is}$ | 0.48 | 0.51 | 0.53 | 0.51 | 0.59 | 0.59 | 0.59 | 0.48 | 0.49 | 0.51 | 0.51 | 0.49 | 0.48 | 0.52 | 0.53 | 0.53 | 0.53 | 0.53 | 0.53 | 0.53 | | | |
| $R_{oos}$ | 0.45 | 0.42 | 0.37 | 0.44 | 0.27 | 0.26 | 0.29 | 0.48 | 0.45 | 0.41 | 0.46 | 0.46 | 0.47 | 0.47 | 0.47 | 0.47 | 0.47 | 0.47 | 0.47 | 0.47 | | | |
| Number of Stocks: N=10 | | | | | | | | | | | | | | | | | | | | | | | |
| HARMSE | 1.27 | 1.19 | 1.10 | 1.04 | 1.19 | 1.19 | 1.20 | 0.99 | 1.00 | 1.01 | 1.07 | 1.05 | 1.07 | 1.01 | 0.99 | 1.00 | 1.00 | 0.99 | 1.03 | 1.01 | 0.99 | 1.02 | 1.00 |
| $R_{is}$ | 0.50 | 0.53 | 0.56 | 0.54 | 0.62 | 0.62 | 0.62 | 0.50 | 0.52 | 0.54 | 0.53 | 0.50 | 0.50 | 0.55 | 0.55 | 0.55 | 0.55 | 0.56 | 0.55 | 0.56 | 0.56 | 0.56 | 0.56 |
| $R_{oos}$ | 0.47 | 0.45 | 0.39 | 0.47 | 0.28 | 0.27 | 0.26 | 0.49 | 0.46 | 0.45 | 0.48 | 0.48 | 0.49 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 | 0.50 | 0.50 |
| Number of Stocks: N=30 | | | | | | | | | | | | | | | | | | | | | | | |
| HARMSE | 1.40 | 1.28 | 1.16 | 1.09 | 1.23 | 1.24 | 1.23 | 1.02 | 1.04 | 1.04 | 1.13 | 1.12 | 1.12 | 1.03 | 1.00 | 0.99 | 1.07 | 1.08 | 1.03 | 1.09 | 1.04 | 1.01 | 0.98 |
| $R_{is}$ | 0.50 | 0.54 | 0.57 | 0.54 | 0.62 | 0.62 | 0.62 | 0.50 | 0.52 | 0.56 | 0.53 | 0.50 | 0.49 | 0.56 | 0.56 | 0.57 | 0.57 | 0.57 | 0.57 | 0.57 | 0.56 | 0.56 | 0.57 |
| $R_{oos}$ | 0.45 | 0.42 | 0.36 | 0.45 | 0.26 | 0.23 | 0.24 | 0.47 | 0.43 | 0.44 | 0.44 | 0.45 | 0.46 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 |
| Number of Stocks: N=50 | | | | | | | | | | | | | | | | | | | | | | | |
| HARMSE | 1.45 | 1.32 | 1.18 | 1.11 | 1.25 | 1.28 | 1.25 | 1.04 | 1.07 | 1.07 | 1.15 | 1.16 | 1.15 | 1.04 | 1.10 | 1.10 | 1.04 | 1.01 | 1.01 | 1.00 | 0.99 | 1.09 | 1.03 |
| $R_{is}$ | 0.49 | 0.54 | 0.57 | 0.54 | 0.62 | 0.62 | 0.63 | 0.50 | 0.52 | 0.56 | 0.53 | 0.50 | 0.49 | 0.56 | 0.57 | 0.57 | 0.57 | 0.56 | 0.56 | 0.56 | 0.57 | 0.57 | 0.56 |
| $R_{oos}$ | 0.45 | 0.42 | 0.36 | 0.46 | 0.29 | 0.23 | 0.25 | 0.47 | 0.43 | 0.44 | 0.44 | 0.45 | 0.47 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 |
| Number of Stocks: N=100 | | | | | | | | | | | | | | | | | | | | | | | |
| HARMSE | 1.49 | 1.36 | 1.20 | 1.15 | 1.29 | 1.31 | 1.29 | 1.07 | 1.09 | 1.09 | 1.19 | 1.19 | 1.18 | 1.06 | 1.12 | 1.13 | 1.05 | 1.09 | 1.03 | 1.02 | 1.01 | 1.02 | 1.02 |
| $R_{is}$ | 0.49 | 0.54 | 0.57 | 0.54 | 0.62 | 0.62 | 0.62 | 0.50 | 0.52 | 0.56 | 0.53 | 0.50 | 0.49 | 0.57 | 0.57 | 0.57 | 0.56 | 0.57 | 0.56 | 0.57 | 0.57 | 0.56 | 0.57 |
| $R_{oos}$ | 0.45 | 0.42 | 0.35 | 0.46 | 0.30 | 0.22 | 0.26 | 0.46 | 0.43 | 0.45 | 0.43 | 0.44 | 0.46 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 |
| Number of Stocks: N=200 | | | | | | | | | | | | | | | | | | | | | | | |
| HARMSE | 1.51 | 1.38 | 1.20 | 1.16 | 1.25 | 1.28 | 1.25 | 1.08 | 1.10 | 1.10 | 1.22 | 1.21 | 1.20 | 1.12 | 1.06 | 1.13 | 1.16 | 1.14 | 1.04 | 1.08 | 1.06 | 1.09 | 1.03 |
| $R_{is}$ | 0.49 | 0.54 | 0.57 | 0.54 | 0.62 | 0.62 | 0.62 | 0.50 | 0.52 | 0.56 | 0.54 | 0.50 | 0.49 | 0.57 | 0.57 | 0.57 | 0.57 | 0.57 | 0.58 | 0.57 | 0.56 | 0.57 | 0.57 |
| $R_{oos}$ | 0.44 | 0.41 | 0.35 | 0.45 | 0.29 | 0.20 | 0.27 | 0.46 | 0.43 | 0.44 | 0.42 | 0.44 | 0.45 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 | 0.50 |

Panel B: Separating Large Jumps from Small Jumps Based on $\gamma^2$

| Model | HAR | SHAR | SCHAR | SCHAR-r | LASSO | EN1 | EN2 | PLS1 | PLS2 | PLS3 | PCR1 | PCR2 | PCR3 | M1 | M2 | M3 | M4 | M5 | M6 | M7 | M8 | M9 | M10 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Number of Stocks: N=5 | | | | | | | | | | | | | | | | | | | | | | | |
| HARMSE | 1.23 | 1.18 | 1.14 | 1.07 | 1.27 | 1.26 | 1.26 | 1.03 | 1.04 | 1.06 | 1.09 | 1.08 | 1.09 | 1.04 | 1.04 | 1.06 | 1.07 | 1.07 | 1.06 | 1.05 | 1.05 | 1.05 | 1.04 |
| $R_{is}$ | 0.47 | 0.50 | 0.52 | 0.50 | 0.58 | 0.58 | 0.58 | 0.47 | 0.48 | 0.51 | 0.50 | 0.48 | 0.47 | 0.52 | 0.52 | 0.52 | 0.52 | 0.52 | 0.52 | 0.52 | 0.52 | 0.52 | 0.52 |
| $R_{oos}$ | 0.44 | 0.41 | 0.35 | 0.43 | 0.23 | 0.20 | 0.24 | 0.47 | 0.45 | 0.43 | 0.45 | 0.46 | 0.47 | 0.47 | 0.47 | 0.47 | 0.47 | 0.47 | 0.47 | 0.47 | 0.47 | 0.47 | 0.47 |
| Number of Stocks: N=10 | | | | | | | | | | | | | | | | | | | | | | | |
| HARMSE | 1.27 | 1.19 | 1.12 | 1.05 | 1.22 | 1.22 | 1.22 | 0.99 | 1.00 | 1.01 | 1.07 | 1.05 | 1.07 | 1.00 | 0.99 | 1.00 | 0.98 | 1.01 | 1.00 | 1.00 | 1.00 | 1.01 | 1.04 |
| $R_{is}$ | 0.49 | 0.53 | 0.55 | 0.53 | 0.61 | 0.61 | 0.61 | 0.49 | 0.50 | 0.54 | 0.52 | 0.49 | 0.49 | 0.55 | 0.55 | 0.55 | 0.55 | 0.55 | 0.55 | 0.55 | 0.55 | 0.55 | 0.55 |
| $R_{oos}$ | 0.46 | 0.43 | 0.36 | 0.45 | 0.26 | 0.25 | 0.27 | 0.48 | 0.45 | 0.44 | 0.46 | 0.46 | 0.48 | 0.49 | 0.49 | 0.49 | 0.49 | 0.49 | 0.49 | 0.49 | 0.49 | 0.49 | 0.49 |
| Number of Stocks: N=30 | | | | | | | | | | | | | | | | | | | | | | | |
| HARMSE | 1.40 | 1.29 | 1.16 | 1.09 | 1.27 | 1.32 | 1.29 | 1.03 | 1.04 | 1.05 | 1.13 | 1.12 | 1.13 | 1.05 | 1.04 | 1.02 | 1.01 | 0.99 | 1.03 | 1.01 | 1.08 | 1.00 | 1.01 |
| $R_{is}$ | 0.49 | 0.54 | 0.57 | 0.54 | 0.62 | 0.62 | 0.62 | 0.50 | 0.52 | 0.55 | 0.53 | 0.50 | 0.49 | 0.56 | 0.56 | 0.56 | 0.56 | 0.56 | 0.56 | 0.56 | 0.56 | 0.56 | 0.56 |
| $R_{oos}$ | 0.45 | 0.42 | 0.36 | 0.46 | 0.30 | 0.27 | 0.28 | 0.47 | 0.44 | 0.45 | 0.45 | 0.45 | 0.47 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 |
| Number of Stocks: N=50 | | | | | | | | | | | | | | | | | | | | | | | |
| HARMSE | 1.44 | 1.32 | 1.17 | 1.11 | 1.28 | 1.31 | 1.27 | 1.04 | 1.06 | 1.06 | 1.15 | 1.14 | 1.15 | 1.04 | 1.05 | 1.02 | 1.06 | 1.01 | 1.05 | 1.03 | 0.99 | 1.02 | 1.08 |
| $R_{is}$ | 0.49 | 0.54 | 0.57 | 0.54 | 0.62 | 0.62 | 0.62 | 0.50 | 0.52 | 0.56 | 0.53 | 0.50 | 0.49 | 0.56 | 0.56 | 0.56 | 0.56 | 0.56 | 0.56 | 0.56 | 0.56 | 0.56 | 0.56 |
| $R_{oos}$ | 0.45 | 0.42 | 0.36 | 0.46 | 0.32 | 0.28 | 0.28 | 0.47 | 0.44 | 0.45 | 0.44 | 0.45 | 0.47 | 0.51 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 |
| Number of Stocks: N=100 | | | | | | | | | | | | | | | | | | | | | | | |
| HARMSE | 1.48 | 1.36 | 1.19 | 1.14 | 1.36 | 1.36 | 1.36 | 1.06 | 1.09 | 1.09 | 1.18 | 1.17 | 1.18 | 1.00 | 1.13 | 1.03 | 1.06 | 1.05 | 1.08 | 1.10 | 0.98 | 1.09 | 0.98 |
| $R_{is}$ | 0.49 | 0.54 | 0.57 | 0.54 | 0.62 | 0.62 | 0.62 | 0.50 | 0.52 | 0.56 | 0.53 | 0.50 | 0.49 | 0.57 | 0.57 | 0.56 | 0.57 | 0.57 | 0.58 | 0.57 | 0.57 | 0.56 | 0.57 |
| $R_{oos}$ | 0.45 | 0.42 | 0.35 | 0.46 | 0.32 | 0.28 | 0.31 | 0.46 | 0.43 | 0.45 | 0.44 | 0.44 | 0.46 | 0.51 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 |
| Number of Stocks: N=200 | | | | | | | | | | | | | | | | | | | | | | | |
| HARMSE | 1.51 | 1.38 | 1.20 | 1.16 | 1.29 | 1.36 | 1.33 | 1.08 | 1.11 | 1.11 | 1.21 | 1.20 | 1.20 | 1.09 | 0.99 | 1.02 | 1.10 | 1.09 | 1.06 | 1.16 | 1.16 | 1.10 | 1.11 |
| $R_{is}$ | 0.49 | 0.54 | 0.57 | 0.54 | 0.61 | 0.61 | 0.61 | 0.50 | 0.52 | 0.56 | 0.53 | 0.50 | 0.49 | 0.58 | 0.57 | 0.57 | 0.58 | 0.57 | 0.58 | 0.57 | 0.57 | 0.57 | 0.59 |
| $R_{oos}$ | 0.44 | 0.41 | 0.35 | 0.45 | 0.38 | 0.30 | 0.30 | 0.46 | 0.43 | 0.44 | 0.43 | 0.44 | 0.46 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 |

# Table 3.2 (Continued)

Panel C: Separating Large Jumps from Small Jumps Based on $\gamma^3$

| Model | HAR | SHAR | SCHAR | SCHAR-r | LASSO | EN1 | EN2 | PLS1 | PLS2 | PLS3 | PCR1 | PCR2 | PCR3 | M1 | M2 | M3 | M4 | M5 | M6 | M7 | M8 | M9 | M10 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Number of Stocks: N=5 | | | | | | | | | | | | | | | | | | | | | | | |
| HARMSE | 1.23 | 1.18 | 1.14 | 1.08 | 1.20 | 1.20 | 1.20 | 1.02 | 1.03 | 1.03 | 1.08 | 1.07 | 1.07 | 1.02 | 1.03 | 1.03 | 1.05 | 1.03 | 1.03 | 1.03 | 1.04 | 1.04 | 1.03 |
| $R_{is}$ | 0.48 | 0.51 | 0.53 | 0.51 | 0.59 | 0.59 | 0.59 | 0.48 | 0.49 | 0.51 | 0.51 | 0.49 | 0.48 | 0.53 | 0.53 | 0.53 | 0.53 | 0.53 | 0.53 | 0.53 | 0.53 | 0.53 | 0.53 |
| $R_{oos}$ | 0.46 | 0.44 | 0.38 | 0.45 | 0.32 | 0.32 | 0.32 | 0.49 | 0.48 | 0.46 | 0.47 | 0.47 | 0.48 | 0.49 | 0.49 | 0.49 | 0.49 | 0.49 | 0.49 | 0.48 | 0.48 | 0.48 | 0.48 |
| Number of Stocks: N=10 | | | | | | | | | | | | | | | | | | | | | | | |
| HARMSE | 1.27 | 1.19 | 1.12 | 1.04 | 1.21 | 1.21 | 1.21 | 0.99 | 1.00 | 1.01 | 1.07 | 1.06 | 1.07 | 1.00 | 0.99 | 1.00 | 0.98 | 0.99 | 1.00 | 0.99 | 1.00 | 1.00 | 1.01 |
| $R_{is}$ | 0.49 | 0.53 | 0.56 | 0.54 | 0.61 | 0.61 | 0.61 | 0.50 | 0.51 | 0.54 | 0.53 | 0.50 | 0.49 | 0.55 | 0.55 | 0.55 | 0.55 | 0.55 | 0.55 | 0.55 | 0.55 | 0.55 | 0.55 |
| $R_{oos}$ | 0.46 | 0.44 | 0.37 | 0.46 | 0.31 | 0.29 | 0.30 | 0.48 | 0.46 | 0.45 | 0.46 | 0.46 | 0.48 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 |
| Number of Stocks: N=30 | | | | | | | | | | | | | | | | | | | | | | | |
| HARMSE | 1.39 | 1.28 | 1.15 | 1.08 | 1.27 | 1.29 | 1.29 | 1.02 | 1.04 | 1.04 | 1.13 | 1.13 | 1.12 | 1.04 | 1.07 | 1.01 | 1.07 | 1.05 | 1.06 | 1.05 | 1.01 | 1.00 | 1.07 |
| $R_{is}$ | 0.50 | 0.54 | 0.57 | 0.54 | 0.62 | 0.62 | 0.62 | 0.50 | 0.51 | 0.55 | 0.53 | 0.50 | 0.49 | 0.56 | 0.57 | 0.56 | 0.57 | 0.56 | 0.57 | 0.57 | 0.56 | 0.56 | 0.56 |
| $R_{oos}$ | 0.46 | 0.43 | 0.36 | 0.46 | 0.31 | 0.27 | 0.30 | 0.47 | 0.45 | 0.45 | 0.45 | 0.45 | 0.47 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 |
| Number of Stocks: N=50 | | | | | | | | | | | | | | | | | | | | | | | |
| HARMSE | 1.45 | 1.32 | 1.17 | 1.12 | 1.38 | 1.41 | 1.35 | 1.05 | 1.06 | 1.07 | 1.15 | 1.17 | 1.15 | 1.04 | 1.06 | 1.02 | 1.10 | 1.03 | 1.07 | 1.08 | 1.02 | 1.03 | 1.04 |
| $R_{is}$ | 0.49 | 0.54 | 0.57 | 0.54 | 0.63 | 0.63 | 0.63 | 0.50 | 0.52 | 0.56 | 0.53 | 0.50 | 0.49 | 0.57 | 0.56 | 0.57 | 0.57 | 0.56 | 0.56 | 0.57 | 0.57 | 0.57 | 0.57 |
| $R_{oos}$ | 0.45 | 0.42 | 0.36 | 0.46 | 0.24 | 0.18 | 0.24 | 0.47 | 0.44 | 0.45 | 0.45 | 0.45 | 0.47 | 0.51 | 0.51 | 0.51 | 0.51 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 |
| Number of Stocks: N=100 | | | | | | | | | | | | | | | | | | | | | | | |
| HARMSE | 1.49 | 1.36 | 1.19 | 1.14 | 1.64 | 1.67 | 1.45 | 1.06 | 1.09 | 1.09 | 1.18 | 1.20 | 1.18 | 1.04 | 1.13 | 1.11 | 1.04 | 1.03 | 1.07 | 1.13 | 1.03 | 1.11 | 1.10 |
| $R_{is}$ | 0.49 | 0.54 | 0.57 | 0.54 | 0.63 | 0.63 | 0.63 | 0.50 | 0.52 | 0.56 | 0.53 | 0.50 | 0.49 | 0.57 | 0.57 | 0.57 | 0.57 | 0.57 | 0.57 | 0.57 | 0.57 | 0.57 | 0.57 |
| $R_{oos}$ | 0.44 | 0.42 | 0.35 | 0.45 | 0.22 | 0.14 | 0.22 | 0.46 | 0.43 | 0.45 | 0.43 | 0.44 | 0.46 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 |
| Number of Stocks: N=200 | | | | | | | | | | | | | | | | | | | | | | | |
| HARMSE | 1.52 | 1.38 | 1.20 | 1.16 | 2.86 | 2.71 | 2.16 | 1.08 | 1.10 | 1.10 | 1.20 | 1.23 | 1.19 | 1.12 | 1.14 | 1.11 | 1.05 | 1.14 | 1.05 | 1.03 | 1.09 | 1.09 | 1.12 |
| $R_{is}$ | 0.49 | 0.54 | 0.57 | 0.54 | 0.63 | 0.63 | 0.63 | 0.50 | 0.52 | 0.56 | 0.53 | 0.50 | 0.49 | 0.57 | 0.57 | 0.57 | 0.57 | 0.57 | 0.57 | 0.57 | 0.57 | 0.58 | 0.57 |
| $R_{oos}$ | 0.44 | 0.41 | 0.35 | 0.45 | 0.21 | 0.15 | 0.27 | 0.46 | 0.43 | 0.44 | 0.43 | 0.44 | 0.46 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 |

*Notes: See notes to Table 3.1.

# Table 3.3: Prediction Performance (30-minute)

Panel A: Separating Large Jumps from Small Jumps Based on $\gamma^1$

| Model | HAR | SHAR | SCHAR | SCHAR-r | LASSO | EN1 | EN2 | PLS1 | PLS2 | PLS3 | PCR1 | PCR2 | PCR3 | M1 | M2 | M3 | M4 | M5 | M6 | M7 | M8 | M9 | M10 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Number of Stocks: N=5 | | | | | | | | | | | | | | | | | | | | | | | |
| HARMSE | 1.71 | 1.64 | 1.50 | 1.47 | 1.63 | 1.64 | 1.63 | 1.46 | 1.46 | 1.48 | 1.48 | 1.48 | 1.52 | | | | | | | | | | |
| $R_{is}$ | 0.47 | 0.51 | 0.53 | 0.51 | 0.59 | 0.58 | 0.58 | 0.49 | 0.49 | 0.51 | 0.51 | 0.51 | 0.49 | | | | | | | | | | |
| $R_{oos}$ | 0.48 | 0.44 | 0.35 | 0.44 | 0.30 | 0.36 | 0.36 | 0.49 | 0.49 | 0.44 | 0.49 | 0.49 | 0.49 | | | | | | | | | | |
| Number of Stocks: N=10 | | | | | | | | | | | | | | | | | | | | | | | |
| HARMSE | 1.75 | 1.64 | 1.46 | 1.43 | 1.51 | 1.51 | 1.51 | 1.34 | 1.33 | 1.38 | 1.40 | 1.41 | 1.44 | | | | | | | | | | |
| $R_{is}$ | 0.48 | 0.53 | 0.56 | 0.54 | 0.60 | 0.60 | 0.60 | 0.51 | 0.51 | 0.54 | 0.54 | 0.52 | 0.50 | | | | | | | | | | |
| $R_{oos}$ | 0.50 | 0.49 | 0.39 | 0.49 | 0.37 | 0.38 | 0.37 | 0.51 | 0.49 | 0.46 | 0.51 | 0.51 | 0.51 | | | | | | | | | | |
| Number of Stocks: N=30 | | | | | | | | | | | | | | | | | | | | | | | |
| HARMSE | 1.87 | 1.69 | 1.46 | 1.39 | 1.43 | 1.43 | 1.43 | 1.29 | 1.26 | 1.32 | 1.36 | 1.39 | 1.43 | | | | | | | | | | |
| $R_{is}$ | 0.49 | 0.54 | 0.58 | 0.56 | 0.62 | 0.61 | 0.62 | 0.52 | 0.52 | 0.57 | 0.56 | 0.52 | 0.51 | | | | | | | | | | |
| $R_{oos}$ | 0.51 | 0.51 | 0.39 | 0.51 | 0.44 | 0.44 | 0.44 | 0.51 | 0.48 | 0.47 | 0.53 | 0.51 | 0.50 | | | | | | | | | | |
| Number of Stocks: N=50 | | | | | | | | | | | | | | | | | | | | | | | |
| HARMSE | 1.94 | 1.72 | 1.49 | 1.40 | 1.41 | 1.41 | 1.42 | 1.31 | 1.27 | 1.33 | 1.37 | 1.41 | 1.46 | 1.36 | 1.38 | 1.33 | 1.36 | | | | | | |
| $R_{is}$ | 0.49 | 0.55 | 0.58 | 0.56 | 0.61 | 0.61 | 0.61 | 0.52 | 0.53 | 0.57 | 0.57 | 0.53 | 0.52 | 0.59 | 0.59 | 0.60 | 0.59 | | | | | | |
| $R_{oos}$ | 0.52 | 0.52 | 0.39 | 0.51 | 0.46 | 0.46 | 0.46 | 0.51 | 0.47 | 0.48 | 0.54 | 0.51 | 0.50 | 0.52 | 0.52 | 0.52 | 0.52 | | | | | | |
| Number of Stocks: N=100 | | | | | | | | | | | | | | | | | | | | | | | |
| HARMSE | 2.00 | 1.76 | 1.53 | 1.41 | 1.41 | 1.41 | 1.41 | 1.31 | 1.27 | 1.32 | 1.37 | 1.41 | 1.47 | 1.36 | 1.39 | 1.36 | 1.32 | 1.37 | 1.36 | 1.41 | 1.35 | 1.32 | 1.40 |
| $R_{is}$ | 0.49 | 0.55 | 0.59 | 0.56 | 0.61 | 0.61 | 0.61 | 0.52 | 0.53 | 0.58 | 0.57 | 0.53 | 0.52 | 0.59 | 0.60 | 0.59 | 0.60 | 0.60 | 0.60 | 0.59 | 0.60 | 0.60 | 0.59 |
| $R_{oos}$ | 0.52 | 0.52 | 0.39 | 0.52 | 0.47 | 0.47 | 0.47 | 0.51 | 0.46 | 0.49 | 0.54 | 0.51 | 0.50 | 0.53 | 0.53 | 0.53 | 0.53 | 0.53 | 0.53 | 0.53 | 0.53 | 0.53 | 0.53 |
| Number of Stocks: N=200 | | | | | | | | | | | | | | | | | | | | | | | |
| HARMSE | 2.00 | 1.75 | 1.55 | 1.39 | 1.35 | 1.35 | 1.35 | 1.30 | 1.24 | 1.30 | 1.36 | 1.40 | 1.45 | 1.39 | 1.37 | 1.39 | 1.39 | 1.37 | 1.37 | 1.29 | 1.37 | 1.34 | 1.29 |
| $R_{is}$ | 0.50 | 0.55 | 0.59 | 0.57 | 0.61 | 0.61 | 0.61 | 0.53 | 0.53 | 0.58 | 0.57 | 0.53 | 0.52 | 0.60 | 0.60 | 0.60 | 0.60 | 0.60 | 0.60 | 0.60 | 0.60 | 0.60 | 0.60 |
| $R_{oos}$ | 0.52 | 0.53 | 0.38 | 0.52 | 0.49 | 0.49 | 0.49 | 0.51 | 0.46 | 0.49 | 0.54 | 0.51 | 0.50 | 0.53 | 0.53 | 0.53 | 0.53 | 0.53 | 0.53 | 0.53 | 0.53 | 0.53 | 0.53 |

Panel B: Separating Large Jumps from Small Jumps Based on $\gamma^2$

| Model | HAR | SHAR | SCHAR | SCHAR-r | LASSO | EN1 | EN2 | PLS1 | PLS2 | PLS3 | PCR1 | PCR2 | PCR3 | M1 | M2 | M3 | M4 | M5 | M6 | M7 | M8 | M9 | M10 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Number of Stocks: N=5 | | | | | | | | | | | | | | | | | | | | | | | |
| HARMSE | 1.74 | 1.67 | 1.53 | 1.51 | 1.63 | 1.62 | 1.63 | 1.46 | 1.46 | 1.49 | 1.50 | 1.52 | 1.53 | | | | | | | | | | |
| $R_{is}$ | 0.46 | 0.50 | 0.53 | 0.50 | 0.58 | 0.58 | 0.58 | 0.48 | 0.48 | 0.50 | 0.51 | 0.50 | 0.48 | | | | | | | | | | |
| $R_{oos}$ | 0.47 | 0.45 | 0.36 | 0.44 | 0.32 | 0.32 | 0.34 | 0.49 | 0.48 | 0.44 | 0.48 | 0.49 | 0.48 | | | | | | | | | | |
| Number of Stocks: N=10 | | | | | | | | | | | | | | | | | | | | | | | |
| HARMSE | 1.75 | 1.63 | 1.44 | 1.42 | 1.51 | 1.51 | 1.53 | 1.35 | 1.34 | 1.36 | 1.39 | 1.42 | 1.45 | | | | | | | | | | |
| $R_{is}$ | 0.48 | 0.53 | 0.56 | 0.54 | 0.60 | 0.60 | 0.60 | 0.51 | 0.51 | 0.54 | 0.55 | 0.52 | 0.50 | | | | | | | | | | |
| $R_{oos}$ | 0.51 | 0.50 | 0.40 | 0.49 | 0.38 | 0.40 | 0.40 | 0.51 | 0.50 | 0.47 | 0.52 | 0.52 | 0.51 | | | | | | | | | | |
| Number of Stocks: N=30 | | | | | | | | | | | | | | | | | | | | | | | |
| HARMSE | 1.89 | 1.69 | 1.46 | 1.40 | 1.42 | 1.42 | 1.42 | 1.31 | 1.27 | 1.33 | 1.36 | 1.40 | 1.45 | | | | | | | | | | |
| $R_{is}$ | 0.49 | 0.54 | 0.58 | 0.56 | 0.61 | 0.61 | 0.61 | 0.52 | 0.52 | 0.57 | 0.56 | 0.53 | 0.51 | | | | | | | | | | |
| $R_{oos}$ | 0.52 | 0.52 | 0.40 | 0.52 | 0.46 | 0.46 | 0.46 | 0.51 | 0.49 | 0.49 | 0.54 | 0.51 | 0.51 | | | | | | | | | | |
| Number of Stocks: N=50 | | | | | | | | | | | | | | | | | | | | | | | |
| HARMSE | 1.94 | 1.72 | 1.49 | 1.41 | 1.42 | 1.41 | 1.43 | 1.31 | 1.27 | 1.32 | 1.37 | 1.41 | 1.46 | 1.38 | | | | | | | | | |
| $R_{is}$ | 0.50 | 0.55 | 0.59 | 0.56 | 0.62 | 0.62 | 0.62 | 0.52 | 0.53 | 0.57 | 0.57 | 0.53 | 0.52 | 0.59 | | | | | | | | | |
| $R_{oos}$ | 0.52 | 0.52 | 0.39 | 0.52 | 0.47 | 0.47 | 0.47 | 0.51 | 0.48 | 0.48 | 0.54 | 0.51 | 0.50 | 0.52 | | | | | | | | | |
| Number of Stocks: N=100 | | | | | | | | | | | | | | | | | | | | | | | |
| HARMSE | 1.98 | 1.74 | 1.52 | 1.40 | 1.40 | 1.36 | 1.38 | 1.30 | 1.25 | 1.31 | 1.36 | 1.40 | 1.46 | 1.36 | 1.37 | 1.36 | 1.37 | 1.38 | 1.36 | 1.33 | 1.37 | 1.36 | 1.36 |
| $R_{is}$ | 0.50 | 0.55 | 0.59 | 0.57 | 0.61 | 0.61 | 0.61 | 0.52 | 0.53 | 0.58 | 0.57 | 0.53 | 0.52 | 0.60 | 0.60 | 0.60 | 0.60 | 0.60 | 0.60 | 0.60 | 0.60 | 0.60 | 0.60 |
| $R_{oos}$ | 0.52 | 0.52 | 0.39 | 0.52 | 0.49 | 0.49 | 0.49 | 0.51 | 0.47 | 0.48 | 0.54 | 0.51 | 0.50 | 0.53 | 0.53 | 0.53 | 0.53 | 0.53 | 0.53 | 0.53 | 0.53 | 0.53 | 0.53 |
| Number of Stocks: N=200 | | | | | | | | | | | | | | | | | | | | | | | |
| HARMSE | 2.00 | 1.75 | 1.55 | 1.40 | 1.41 | 1.34 | 1.37 | 1.30 | 1.25 | 1.30 | 1.36 | 1.40 | 1.46 | 1.39 | 1.39 | 1.40 | 1.37 | 1.40 | 1.35 | 1.35 | 1.37 | 1.36 | 1.32 |
| $R_{is}$ | 0.50 | 0.55 | 0.59 | 0.57 | 0.61 | 0.61 | 0.61 | 0.52 | 0.53 | 0.58 | 0.57 | 0.53 | 0.52 | 0.60 | 0.60 | 0.60 | 0.60 | 0.60 | 0.60 | 0.60 | 0.60 | 0.60 | 0.60 |
| $R_{oos}$ | 0.52 | 0.52 | 0.38 | 0.52 | 0.50 | 0.50 | 0.50 | 0.51 | 0.46 | 0.48 | 0.54 | 0.51 | 0.50 | 0.54 | 0.53 | 0.53 | 0.53 | 0.53 | 0.53 | 0.53 | 0.53 | 0.53 | 0.53 |

# Table 3.3 (Continued)

Panel C: Separating Large Jumps from Small Jumps Based on $\gamma^3$

| Model | HAR | SHAR | SCHAR | SCHAR-r | LASSO | EN1 | EN2 | PLS1 | PLS2 | PLS3 | PCR1 | PCR2 | PCR3 | M1 | M2 | M3 | M4 | M5 | M6 | M7 | M8 | M9 | M10 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Number of Stocks: N=5 | | | | | | | | | | | | | | | | | | | | | | | |
| HARMSE | 1.71 | 1.63 | 1.50 | 1.46 | 1.53 | 1.53 | 1.53 | 1.42 | 1.42 | 1.45 | 1.48 | 1.48 | 1.49 | | | | | | | | | | |
| $R_{is}$ | 0.47 | 0.50 | 0.53 | 0.51 | 0.58 | 0.58 | 0.58 | 0.49 | 0.49 | 0.50 | 0.52 | 0.51 | 0.49 | | | | | | | | | | |
| $R_{oos}$ | 0.48 | 0.47 | 0.36 | 0.45 | 0.38 | 0.38 | 0.39 | 0.49 | 0.48 | 0.45 | 0.48 | 0.49 | 0.48 | | | | | | | | | | |
| Number of Stocks: N=10 | | | | | | | | | | | | | | | | | | | | | | | |
| HARMSE | 1.76 | 1.64 | 1.45 | 1.43 | 1.50 | 1.50 | 1.50 | 1.37 | 1.36 | 1.38 | 1.42 | 1.44 | 1.47 | | | | | | | | | | |
| $R_{is}$ | 0.49 | 0.53 | 0.56 | 0.54 | 0.60 | 0.60 | 0.60 | 0.51 | 0.51 | 0.53 | 0.55 | 0.53 | 0.51 | | | | | | | | | | |
| $R_{oos}$ | 0.51 | 0.51 | 0.41 | 0.50 | 0.42 | 0.41 | 0.42 | 0.52 | 0.51 | 0.48 | 0.53 | 0.53 | 0.52 | | | | | | | | | | |
| Number of Stocks: N=30 | | | | | | | | | | | | | | | | | | | | | | | |
| HARMSE | 1.88 | 1.69 | 1.47 | 1.39 | 1.43 | 1.43 | 1.44 | 1.31 | 1.28 | 1.33 | 1.37 | 1.40 | 1.45 | | | | | | | | | | |
| $R_{is}$ | 0.49 | 0.54 | 0.58 | 0.56 | 0.61 | 0.61 | 0.61 | 0.52 | 0.52 | 0.56 | 0.56 | 0.53 | 0.51 | | | | | | | | | | |
| $R_{oos}$ | 0.51 | 0.52 | 0.39 | 0.51 | 0.45 | 0.45 | 0.45 | 0.52 | 0.50 | 0.48 | 0.54 | 0.51 | 0.51 | | | | | | | | | | |
| Number of Stocks: N=50 | | | | | | | | | | | | | | | | | | | | | | | |
| HARMSE | 1.94 | 1.73 | 1.50 | 1.42 | 1.41 | 1.42 | 1.43 | 1.33 | 1.29 | 1.34 | 1.38 | 1.42 | 1.48 | 1.40 | 1.40 | | | | | | | | |
| $R_{is}$ | 0.50 | 0.55 | 0.59 | 0.56 | 0.61 | 0.61 | 0.61 | 0.52 | 0.53 | 0.57 | 0.56 | 0.53 | 0.52 | 0.59 | 0.60 | | | | | | | | |
| $R_{oos}$ | 0.52 | 0.52 | 0.39 | 0.52 | 0.48 | 0.48 | 0.48 | 0.51 | 0.49 | 0.48 | 0.54 | 0.51 | 0.50 | 0.52 | 0.52 | | | | | | | | |
| Number of Stocks: N=100 | | | | | | | | | | | | | | | | | | | | | | | |
| HARMSE | 1.99 | 1.76 | 1.52 | 1.42 | 1.41 | 1.40 | 1.41 | 1.33 | 1.28 | 1.33 | 1.38 | 1.43 | 1.49 | 1.41 | 1.40 | 1.41 | 1.40 | 1.37 | 1.41 | 1.42 | 1.39 | 1.42 | 1.36 |
| $R_{is}$ | 0.49 | 0.55 | 0.59 | 0.56 | 0.61 | 0.61 | 0.61 | 0.52 | 0.53 | 0.57 | 0.57 | 0.53 | 0.52 | 0.60 | 0.60 | 0.60 | 0.60 | 0.60 | 0.60 | 0.60 | 0.59 | 0.60 | 0.60 |
| $R_{oos}$ | 0.52 | 0.53 | 0.39 | 0.52 | 0.49 | 0.49 | 0.49 | 0.51 | 0.48 | 0.48 | 0.54 | 0.51 | 0.50 | 0.53 | 0.53 | 0.53 | 0.53 | 0.53 | 0.52 | 0.52 | 0.52 | 0.52 | 0.52 |
| Number of Stocks: N=200 | | | | | | | | | | | | | | | | | | | | | | | |
| HARMSE | 2.00 | 1.75 | 1.55 | 1.39 | 1.36 | 1.33 | 1.34 | 1.30 | 1.25 | 1.30 | 1.34 | 1.40 | 1.46 | 1.39 | 1.38 | 1.39 | 1.39 | 1.39 | 1.39 | 1.38 | 1.38 | 1.39 | 1.34 |
| $R_{is}$ | 0.50 | 0.55 | 0.59 | 0.57 | 0.61 | 0.61 | 0.61 | 0.52 | 0.53 | 0.58 | 0.57 | 0.53 | 0.52 | 0.60 | 0.60 | 0.60 | 0.61 | 0.61 | 0.61 | 0.60 | 0.60 | 0.60 | 0.60 |
| $R_{oos}$ | 0.52 | 0.53 | 0.38 | 0.52 | 0.51 | 0.51 | 0.51 | 0.51 | 0.47 | 0.48 | 0.54 | 0.51 | 0.50 | 0.54 | 0.53 | 0.53 | 0.53 | 0.53 | 0.53 | 0.53 | 0.53 | 0.53 | 0.53 |

*Notes: See notes to Table 3.1.

## Table 3.4: Predictors in Best-Performing Models (5-minute)

| | $\gamma^1$ | $\gamma^2$ | $\gamma^3$ |
|---|---|---|---|
| | | Number of Stocks: N=5 | |
| M1 | dNC wNC mNC | dNC wNC mNC | dNC wNC mNC |
| M2 | dNC wNC mPC | dNC wNC wMSLJN2 mNC | dNC wNC mPC |
| M3 | dNC wNC mPCSJ2 | dNC wNC mNC mMLJ2 | dNC wNC wMSLJP2 mNC |
| M4 | dNC wNC mNCSJ2 | dNC wPC wNC mNC | dNC wNC mNC mPCSJ2 |
| M5 | dNC wNC mMC2 | dNC wNC mNCSJ2 | dNC wNC mNC mMSLJP2 |
| | | Number of Stocks: N=10 | |
| M1 | dNC wNC mPCSJ2 | dNC wNC mNC | dNC wNC mNC |
| M2 | dNC wNC mNC | dNC wNC mPC | dNC wNC mNC mPCSJ2 |
| M3 | dNC wNC mPC | dNC wNC mNC mPCSJ2 | dNC wNC wNCSJ2 mNC |
| M4 | dNC wPC wNC mNC | dNC wPC wNC mNC | dNC wPC wNC mNC |
| M5 | dNC wNC wPCSJ2 mNC | dNC wNC wNCSJ2 mNC | dNC wNC wMC2 mPCSJ2 |
| | | Number of Stocks: N=30 | |
| M1 | dNC wNC mPCSJ2 | dNC wNC wNCSJ2 mPCSJ2 | dNC wNC mNC |
| M2 | dNC wNC mPCLJ2 | dNC wNC mNC | dNC wNC mNC mPCSJ2 |
| M3 | dNC wNC mNC | dNC wNC mNC mPCSJ2 | dNC wNC wNCSJ2 mNC |
| M4 | dNC wNC wPCLJ2 mPCLJ2 | dNC wNC mPC | dNC wNC mPC |
| M5 | dNC wNC mPC | dNC wNC wNCSJ2 mNC | dNC wPC wNC mNC |
| | | Number of Stocks: N=50 | |
| M1 | dNC wNC mPCLJ2 | dNC wNC wNCSJ2 mPCSJ2 | dNC wNC mNC mPCSJ2 |
| M2 | dNC wNC wPCLJ2 mPCLJ2 | dNC wNC wPSLJ2 wMSLJN2 mNC | dNC wNC mPC mNC mPCSJ2 |
| M3 | dNC wPC wNC mPCLJ2 | dNC wPC wNC wNCSJ2 mPCSJ2 | dNC wNC mNC |
| M4 | dNC wNC mPCSJ2 | dNC wNC mNC mPCSJ2 | dNC wNC wNCSJ2 mNC |
| M5 | dNC wNC mNC mPCLJ2 | dNC wNC mNC | dNC wNC mPC |
| | | Number of Stocks: N=100 | |
| M1 | dNC wNC mPCLJ2 | dNC dNCSJ2 wNC wPSLJ2 wNSLJ2 mNC | dNC wNC mPC mNC mPCSJ2 |
| M2 | dNC wNC mNC mPCLJ2 | dNC wNC wPSLJ2 wMSLJN2 mNC | dNC wNC mNC mPCSJ2 |
| M3 | dNC wPC wNC mPCLJ2 | dNC wNC wNCSJ2 mPCSJ2 | dNC wNC wMLJ2 mPC mNC mPCSJ2 |
| M4 | dNC wNC wPCLJ2 mPCLJ2 | dNC wPC wNC wNCSJ2 mPCSJ2 | dNC wNC wMSLJP2 mNC |
| M5 | dNC wPC wNC mNC mPCLJ2 | dNC wNC mPC mNC mMC2 mMSLJP2 | dNC wNC wNCSJ2 mNC |
| | | Number of Stocks: N=200 | |
| M1 | dNC dNCSJ2 wNC wPLJ wMCSJP2 mNC mPSJ | dNC wNC mPC mNC mMC2 mMSLJP2 | dNC wNC mNC mMCSJN2 mMSLJP2 |
| M2 | dNC dNCSJ2 wNC mPC mNC mPCLJ2 | dNC dNCSJ2 wNC wPSLJ2 wNSLJ2 mNC | dNC wPC wNC mNC mMLJ2 mMCSJP2 mMSLJP2 |
| M3 | dNC wNC mPCLJ2 | dNC wNC wPSLJ2 wMSLJN2 mNC | dNC wPC wNC wMCLJP2 mPSJ mMSLJP2 |
| M4 | dNC wPC wNC mPCLJ2 | dNC wNC mNC mMCSJN2 mMSLJP2 | dNC wNC mPC mNC mPCSJ2 |
| M5 | dNC wNC mNC mPCLJ2 | dNC wPC wNC mNC mMCSJN2 mMSLJP2 | dNC wNC wMLJ2 mPC mNC mPCSJ2 |

*Notes: See notes to Table 3.1.

## Table 3.5: Predictors in Best-Performing Models (15-minute)

| | $\gamma^1$ | $\gamma^2$ | $\gamma^3$ |
|---|---|---|---|
| | | Number of Stocks: N=5 | |
| M1 | dNC wNC mPC | dNC wNC mNC | dNC wNC mNC |
| M2 | dNC dMSLJN2 wNC mNC | dNC wNC wMSLJP2 mNC | dNC wNC mNCSJ2 |
| M3 | dNC wNC mNC mMCLJN2 | dNC wNC mNCSJ2 | dNC wNC mPC |
| M4 | dNC wNC wPCSJ2 mNC | dNC wNC mPC mMSLJP2 | dNC wNC mPCSJ2 |
| M5 | dNC wNC wMCLJP2 mNC | dNC wNC mPCSJ2 | dNC wNC mNC mNLJ |
| | | Number of Stocks: N=10 | |
| M1 | dNC wNC mPCSJ2 | dNC wNC mPCSJ2 | dNC wNC mPCSJ2 |
| M2 | dNC wNC mNC | dNC wNC mNC | dNC wNC mNC |
| M3 | dNC wNC mPC | dNC wNC mPC | dNC wNC mPC |
| M4 | dNC wNC mNCSJ2 | dNC wNC mNCSJ2 | dNC wNC mNCSJ2 |
| M5 | dNC dPSLJ2 wNC mNC | dNC wNC wMSLJP2 mPCSJ2 | dNC wNC mNC mNLJ |
| | | Number of Stocks: N=30 | |
| M1 | dNC dNSLJ2 wNC | dNC wNC mMLJ2 | dNC wNC mNLJ |
| M2 | dNC dNSLJ2 wNC mPC | dNC wNC | dNC dPSLJ2 wNC mNC mNLJ |
| M3 | dNC dNSLJ2 wNC mNC | dNC wNC mNLJ | dNC wNC mNC mNLJ |
| M4 | dNC dNSLJ2 wNC wMC2 mNC | dNC wNC mPCSJ2 | dNC dPSLJ2 wPC wNC mNC mNLJ |
| M5 | dNC dNSLJ2 wNC wMC2 mPC | dNC wNC mNC mNLJ | dNC wPC wNC mNLJ |
| | | Number of Stocks: N=50 | |
| M1 | dNC dNSLJ2 wNC | dNC wNC mMLJ2 | dNC dNSJ wNC wNSJ |
| M2 | dNC dNSLJ2 wNC wMC2 mNC | dNC wNC | dNC wNC mNLJ |
| M3 | dNC dNSLJ2 wNC wMC2 mPC | dNC wNC mPCLJ2 | dNC dNSJ wNC mNC mNLJ |
| M4 | dNC dNSLJ2 wNC wNSLJ2 | dNC wNC wMCSJP2 mPCSJ2 | dNC dNSJ wNC wNSJ wMC2 mPC |
| M5 | dNC dNSLJ2 wNC mPC | dNC wNC mPCSJ2 | dNC wNC mNC mNLJ |
| | | Number of Stocks: N=100 | |
| M1 | dNC dNSLJ2 wNC wNSLJ2 | dNC dPSJ dNSJ wNC wNSJ mMLJ2 | dNC dNSJ wNC wNSJ |
| M2 | dNC dNSLJ2 wNC wMC2 mNC | dNC wPC wNC wMC2 mNCLJ2 | dNC dNSJ wNC wNSJ wMC2 mPC |
| M3 | dNC dNSLJ2 wNC wMC2 mPC | dNC wNC mMLJ2 | dNC dNSJ wNC wNSJ wMCSJP2 mPCSJ2 |
| M4 | dNC dNSLJ2 wNC | dNC dNSJ wNC wNSJ mNC mPCSJ2 | dNC dNSJ wNC mNC mNLJ |
| M5 | dNC dNSLJ2 wPC wNC wMCLJN2 mNC | dNC dNCLJ2 wNC wNSJ | dNC dNSJ wNC mNC mNLJ mMSLJP2 |
| | | Number of Stocks: N=200 | |
| M1 | dNC dNSLJ2 wPC wNC wMCLJN2 mNC | dNC dPSJ dNSJ wNC wNSJ wMC2 mNC mMCLJP2 | dNC dNSJ wNC wNSJ wMCSJP2 mPCSJ2 |
| M2 | dNC dNSLJ2 wNC wNSLJ2 | dNC dPSJ dNSJ wNC wNSJ mMLJ2 | dNC dNSJ wNC wNSJ wMC2 mPC |
| M3 | dNC dNSLJ2 wNC wMC2 mNC | dNC dNSJ wNC wNSJ mPCLJ2 | dNC dNSJ wNC wNCSJ2 wMC2 mNCSJ2 mNCLJ2 |
| M4 | dNC dNSLJ2 wNC wMC2 wMCSJP2 mNC mPCSJ2 | dNC dNLJ dNSLJ2 wNC wNSJ wMCSJN2 mNSJ mMSLJP2 | dNC dNSJ wNC wNSJ |
| M5 | dNC dNSLJ2 wNC wMC2 mPC | dNC dNLJ dNSLJ2 wNC wNSJ mNLJ | dNC dNSJ wNC wNSJ wMCSJP2 |

*Notes: See notes to Table 3.1.

## Table 3.6: Predictors in Best-Performing Models (30-minute)

| | $\gamma^1$ | $\gamma^2$ | $\gamma^3$ |
|---|---|---|---|
| | | Number of Stocks: N=50 | |
| M1 | dNC wNC wPCLJ2 mNC mNSLJ2 | dNC wNC wPCLJ2 mNSJ | dNC wNC wPSLJ2 mNSJ |
| M2 | dNC wNC wPCLJ2 mPC mNSLJ2 | | dNC dNSJ wNC wPCLJ2 mNSJ |
| M3 | dNC wNC wPCLJ2 mPC mNC mNSLJ2 | | |
| M4 | dNC wNC wPCLJ2 mNCSJ2 mNSLJ2 | | |
| M5 | | | |
| | | Number of Stocks: N=100 | |
| M1 | dNC wNC wPCLJ2 mNC mNSLJ2 | dNC wNC wPCLJ2 mNCSJ2 mNSLJ2 | dNC wNC wPCLJ2 mNSJ mNLJ |
| M2 | dNC wNC wPCLJ2 mPC mNSLJ2 | dNC wNC wPCLJ2 mPC mNSLJ2 | dNC wNC wPCLJ2 mNSJ |
| M3 | dNC wNC wPCLJ2 mNCSJ2 mNSLJ2 | dNC wNC wPCLJ2 mNC mNLJ mNSLJ2 | dNC dMC2 wNC wPCLJ2 wMCSJN2 mNC mNSJ |
| M4 | dNC wNC wPCLJ2 mPC mNC mNSLJ2 | dNC wNC wPCLJ2 mNSJ | dNC dMC2 wNC wPSJ wMC2 mNCSJ2 |
| M5 | dNC dNSJ wNC wPCLJ2 mNC mNSLJ2 | dNC wNC wPCLJ2 mNSJ mNLJ | dNC wNC wPCLJ2 mPC mNSJ |
| | | Number of Stocks: N=200 | |
| M1 | dNC wNC wPSJ wPCLJ2 wMCSJN2 mNC mNSLJ2 | dNC dNSJ dMC2 wNC wPCLJ2 wMCSJN2 mPC mNSJ | dNC wNC wPCLJ2 wMCSJN2 mNLJ mNCSJ2 |
| M2 | dNC dNSJ dMC2 wNC wPCLJ2 wMSJ2 mNC mNSLJ2 | dNC dNSJ dMC2 wNC wPCLJ2 wMCSJN2 mPC mNSJ mMCLJP2 | dNC dNSJ dMC2 wNC wPCLJ2 wMCSJN2 mPC mNSJ |
| M3 | dNC dNSJ wNC wPCLJ2 mNSLJ2 | dNC wNC wPCLJ2 wMCSJN2 mNCSJ2 | dNC dMC2 wNC wPCLJ2 wMCSJN2 mNLJ mNCSJ2 mNSLJ2 |
| M4 | dNC wNC wPCLJ2 mNSLJ2 | dNC wNC wPCLJ2 wMCSJN2 mNCSJ2 mMCSJP2 | dNC dMC2 wNC wPCLJ2 wMLJ2 wMCSJN2 mNC mNSJ |
| M5 | dNC dNSJ dMC2 dMCLJN2 wNC wPCLJ2 mNSJ | dNC wNC wPCLJ2 wMCSJN2 mNCSJ2 mMSJ2 mMCSJP2 | dNC dNSJ dMC2 wNC wPCLJ2 wMCSJN2 mPC mNSJ mPLJ |

*Notes: See notes to Table 3.1.

# Table 3.7: Comparison of Out-of-Sample Prediction Performance (5-minute)

Panel A: Portfolios Constructed by 10 Stocks

| | SHAR | SCHAR | SCHAR-r | LASSO | EN1 | EN2 | PLS1 | PLS2 | PLS3 | PCR1 | PCR2 | PCR3 | M1 | M2 | M3 | M4 | M5 | M6 | M7 | M8 | M9 | M10 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| HAR | -0.76 | -2.55 | 1.38 | -2.90 | -2.98 | -3.05 | 0.10 | -1.17 | -1.23 | 0.79 | 0.09 | -0.21 | **2.18** | **2.10** | **2.09** | **1.97** | **2.03** | **2.04** | **2.04** | **2.01** | **1.94** | **1.91** |
| SHAR | | -3.53 | **1.94** | -2.82 | -2.94 | -3.00 | 0.67 | 0.03 | -0.01 | 1.03 | 0.68 | 0.50 | **3.11** | **3.09** | **3.07** | **2.99** | **2.98** | **3.00** | **2.90** | **3.10** | **2.91** | **2.98** |
| SCHAR | | | **3.64** | -1.22 | -1.26 | -1.22 | **2.17** | **2.11** | **2.59** | **2.57** | **2.20** | **1.99** | **4.64** | **4.64** | **4.60** | **4.59** | **4.56** | **4.47** | **4.46** | **4.46** | **4.53** | **4.61** |
| SCHAR-r | | | | -3.39 | -3.51 | -3.64 | -1.06 | -2.19 | -2.54 | -0.72 | -1.13 | -1.24 | **1.90** | **1.77** | **1.76** | 1.64 | **1.66** | **1.67** | **1.71** | 1.45 | 1.52 | 1.46 |
| LASSO | | | | | -0.17 | 0.87 | **3.15** | **2.85** | **2.89** | **3.33** | **3.17** | **3.02** | **3.90** | **3.88** | **3.88** | **3.86** | **3.87** | **3.87** | **3.86** | **3.81** | **3.84** | **3.84** |
| EN1 | | | | | | 1.08 | **3.25** | **2.96** | **3.01** | **3.43** | **3.27** | **3.12** | **4.02** | **4.00** | **4.00** | **3.98** | **3.99** | **3.99** | **3.98** | **3.94** | **3.96** | **3.96** |
| EN2 | | | | | | | **3.32** | **3.04** | **3.12** | **3.54** | **3.35** | **3.17** | **4.21** | **4.19** | **4.19** | **4.18** | **4.18** | **4.18** | **4.18** | **4.12** | **4.16** | **4.15** |
| PLS1 | | | | | | | | -1.42 | -1.10 | **1.80** | -0.09 | -2.69 | **1.68** | 1.61 | 1.61 | 1.54 | 1.57 | 1.61 | 1.61 | 1.53 | 1.50 | 1.45 |
| PLS2 | | | | | | | | | -0.12 | **2.12** | 1.39 | 1.01 | **2.96** | **2.87** | **2.91** | **2.78** | **2.87** | **2.93** | **2.94** | **2.96** | **2.75** | **2.73** |
| PLS3 | | | | | | | | | | **1.73** | 1.11 | 0.80 | **4.06** | **3.96** | **3.98** | **3.80** | **3.95** | **3.93** | **4.00** | **4.17** | **3.81** | **3.81** |
| PCR1 | | | | | | | | | | | -2.27 | -2.11 | 1.60 | 1.52 | 1.52 | 1.43 | 1.47 | 1.50 | 1.50 | 1.40 | 1.38 | 1.32 |
| PCR2 | | | | | | | | | | | | -1.62 | **1.75** | **1.68** | **1.68** | 1.61 | 1.64 | **1.67** | **1.67** | 1.59 | 1.56 | 1.51 |
| PCR3 | | | | | | | | | | | | | **1.76** | **1.70** | **1.70** | 1.63 | **1.66** | **1.70** | **1.70** | 1.62 | 1.59 | 1.55 |
| M1 | | | | | | | | | | | | | | -1.26 | -1.96 | -1.90 | -4.11 | -1.00 | -2.08 | -0.96 | -3.31 | -3.33 |
| M2 | | | | | | | | | | | | | | | -0.77 | -1.44 | -1.84 | -0.60 | -1.23 | -0.71 | -2.72 | -2.75 |
| M3 | | | | | | | | | | | | | | | | -0.57 | -1.13 | -0.46 | -1.04 | -0.58 | -1.46 | -2.24 |
| M4 | | | | | | | | | | | | | | | | | -0.07 | -0.01 | -0.39 | -0.29 | -1.18 | -1.38 |
| M5 | | | | | | | | | | | | | | | | | | 0.03 | -0.55 | -0.33 | -1.17 | -1.82 |
| M6 | | | | | | | | | | | | | | | | | | | -0.34 | -0.38 | -0.44 | -0.83 |
| M7 | | | | | | | | | | | | | | | | | | | | -0.17 | -0.33 | -0.74 |
| M8 | | | | | | | | | | | | | | | | | | | | | 0.00 | -0.20 |
| M9 | | | | | | | | | | | | | | | | | | | | | | -0.55 |

Panel B: Portfolios Constructed by 200 Stocks

| | SHAR | SCHAR | SCHAR-r | LASSO | EN1 | EN2 | PLS1 | PLS2 | PLS3 | PCR1 | PCR2 | PCR3 | M1 | M2 | M3 | M4 | M5 | M6 | M7 | M8 | M9 | M10 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| HAR | -0.40 | -1.94 | 1.60 | -1.48 | -1.79 | -1.61 | -0.60 | -0.14 | 0.96 | -0.01 | -0.79 | -0.78 | **2.58** | **2.24** | **2.32** | **2.56** | **2.54** | **2.44** | **2.35** | **2.42** | **2.29** | **2.32** |
| SHAR | | -2.73 | 1.50 | -0.75 | -1.44 | -1.42 | 0.04 | 0.26 | 1.42 | 0.34 | -0.07 | -0.06 | **2.32** | **2.45** | **2.37** | **2.31** | **2.32** | **2.31** | **2.48** | **2.31** | **2.48** | **2.48** |
| SCHAR | | | **2.86** | 1.26 | 0.52 | 0.51 | 1.36 | **1.69** | **3.14** | **1.68** | 1.24 | 1.25 | **3.48** | **3.69** | **3.61** | **3.35** | **3.36** | **3.39** | **3.55** | **3.38** | **3.56** | **3.60** |
| SCHAR-r | | | | -2.85 | -2.56 | -2.38 | -1.70 | -1.18 | -0.39 | -1.50 | -1.77 | -1.76 | **1.68** | 1.59 | **1.72** | **1.79** | **1.78** | 1.61 | **1.88** | 1.54 | **1.86** | **1.73** |
| LASSO | | | | | -1.33 | -1.07 | 0.83 | 1.20 | **2.45** | 1.38 | 0.63 | 0.64 | **4.12** | **3.49** | **3.58** | **4.15** | **4.17** | **4.26** | **4.03** | **4.27** | **4.00** | **3.74** |
| EN1 | | | | | | -0.05 | 1.42 | **1.68** | **2.53** | **1.81** | 1.25 | 1.25 | **3.59** | **3.21** | **3.18** | **3.59** | **3.59** | **3.54** | **3.47** | **3.53** | **3.46** | **3.34** |
| EN2 | | | | | | | 1.27 | 1.51 | **2.42** | 1.62 | 1.12 | 1.12 | **3.33** | **3.07** | **3.02** | **3.32** | **3.32** | **3.26** | **3.27** | **3.25** | **3.27** | **3.16** |
| PLS1 | | | | | | | | 0.46 | 1.01 | **2.17** | -2.54 | -2.41 | **2.32** | **1.96** | **2.02** | **2.37** | **2.35** | **2.23** | **2.13** | **2.21** | **2.10** | **2.02** |
| PLS2 | | | | | | | | | 0.93 | 0.15 | -0.69 | -0.67 | **2.32** | **1.79** | **1.83** | **2.22** | **2.20** | **2.11** | **1.91** | **2.09** | **1.88** | **1.80** |
| PLS3 | | | | | | | | | | -0.74 | -1.11 | -1.10 | **2.41** | **2.61** | **2.49** | **2.11** | **2.12** | **1.91** | **2.24** | **1.88** | **2.19** | **2.47** |
| PCR1 | | | | | | | | | | | -2.33 | -2.29 | **2.24** | **1.86** | **1.94** | **2.33** | **2.30** | **2.16** | **2.06** | **2.13** | **2.03** | **1.93** |
| PCR2 | | | | | | | | | | | | **2.69** | **2.34** | **2.00** | **2.06** | **2.40** | **2.38** | **2.27** | **2.16** | **2.25** | **2.13** | **2.06** |
| PCR3 | | | | | | | | | | | | | **2.34** | **1.99** | **2.05** | **2.39** | **2.37** | **2.26** | **2.15** | **2.24** | **2.13** | **2.05** |
| M1 | | | | | | | | | | | | | | -0.10 | -0.24 | -0.55 | -0.59 | -0.64 | -0.61 | -0.73 | -0.57 | -0.55 |
| M2 | | | | | | | | | | | | | | | -0.47 | -0.14 | -0.15 | -0.22 | -0.40 | -0.26 | -0.41 | -0.68 |
| M3 | | | | | | | | | | | | | | | | -0.01 | -0.03 | -0.12 | -0.26 | -0.17 | -0.27 | -0.53 |
| M4 | | | | | | | | | | | | | | | | | -0.38 | -0.28 | -0.33 | -0.40 | -0.31 | -0.29 |
| M5 | | | | | | | | | | | | | | | | | | -0.25 | -0.31 | -0.37 | -0.29 | -0.27 |
| M6 | | | | | | | | | | | | | | | | | | | -0.08 | -1.02 | -0.09 | -0.11 |
| M7 | | | | | | | | | | | | | | | | | | | | -0.01 | -0.08 | -0.14 |
| M8 | | | | | | | | | | | | | | | | | | | | | 0.00 | -0.05 |
| M9 | | | | | | | | | | | | | | | | | | | | | | -0.11 |

*Notes: This table reports pairwise Diebold-Mariano test statistics comparing the daily out-of-sample forecasting performance among models. Portfolios are constructed by 10 and 200 stocks using 5-minute high frequency data. Positive number indicates that the column model outperform the row model. Bold font indicates the difference is significant at 10% level or better.

# Table 3.8: Comparison of Out-of-Sample Prediction Performance (15-minute)

Panel A: Portfolios Constructed by 10 Stocks

| | SHAR | SCHAR | SCHAR-r | LASSO | EN1 | EN2 | PLS1 | PLS2 | PLS3 | PCR1 | PCR2 | PCR3 | M1 | M2 | M3 | M4 | M5 | M6 | M7 | M8 | M9 | M10 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| HAR | -1.18 | -2.95 | -0.28 | -2.66 | -2.47 | -2.62 | 1.02 | -0.23 | -0.84 | 0.25 | 0.29 | 0.89 | 1.02 | 0.96 | 0.96 | 0.97 | 0.94 | 0.92 | 0.88 | 0.90 | 0.87 | 0.87 |
| SHAR | | -3.26 | **1.77** | -2.75 | -2.50 | -2.70 | **1.76** | 1.15 | 0.63 | 1.37 | 1.20 | 1.61 | **2.80** | **2.72** | **2.70** | **2.70** | **2.68** | **2.70** | **2.60** | **2.65** | **2.57** | **2.57** |
| SCHAR | | | **4.03** | -1.48 | -1.40 | -1.43 | **3.05** | **2.86** | **2.87** | **2.89** | **2.54** | **2.93** | **4.09** | **4.08** | **4.06** | **4.05** | **4.03** | **4.02** | **4.02** | **4.05** | **4.00** | **4.01** |
| SCHAR-r | | | | -2.91 | -2.66 | -2.87 | 1.31 | 0.02 | -1.17 | 0.49 | 0.53 | 1.11 | **1.94** | **1.84** | **1.82** | **1.89** | **1.80** | **1.78** | **1.69** | **1.74** | **1.67** | 1.62 |
| LASSO | | | | | -0.77 | 1.17 | **3.04** | **2.78** | **2.69** | **2.87** | **2.83** | **2.98** | **3.22** | **3.20** | **3.18** | **3.19** | **3.18** | **3.18** | **3.16** | **3.17** | **3.14** | **3.13** |
| EN1 | | | | | | 1.09 | **2.80** | **2.55** | **2.46** | **2.64** | **2.62** | **2.75** | **2.96** | **2.94** | **2.93** | **2.94** | **2.93** | **2.93** | **2.91** | **2.92** | **2.90** | **2.89** |
| EN2 | | | | | | | **2.99** | **2.73** | **2.64** | **2.82** | **2.79** | **2.93** | **3.18** | **3.15** | **3.14** | **3.15** | **3.14** | **3.14** | **3.12** | **3.13** | **3.10** | **3.09** |
| PLS1 | | | | | | | | -2.13 | -1.95 | -1.17 | -2.47 | -1.46 | 0.46 | 0.39 | 0.38 | 0.39 | 0.36 | 0.33 | 0.28 | 0.31 | 0.27 | 0.28 |
| PLS2 | | | | | | | | | -1.28 | 0.54 | 0.81 | **1.79** | 1.64 | 1.57 | 1.53 | **1.65** | 1.52 | 1.51 | 1.44 | 1.49 | 1.41 | 1.37 |
| PLS3 | | | | | | | | | | 1.37 | 1.17 | **1.73** | **3.54** | **3.50** | **3.41** | **3.59** | **3.39** | **3.43** | **3.33** | **3.39** | **3.24** | **3.07** |
| PCR1 | | | | | | | | | | | 0.14 | 0.91 | 1.15 | 1.08 | 1.07 | 1.09 | 1.05 | 1.01 | 0.97 | 0.99 | 0.96 | 0.94 |
| PCR2 | | | | | | | | | | | | **2.24** | 0.85 | 0.79 | 0.78 | 0.81 | 0.76 | 0.74 | 0.70 | 0.73 | 0.68 | 0.67 |
| PCR3 | | | | | | | | | | | | | 0.53 | 0.46 | 0.45 | 0.47 | 0.43 | 0.41 | 0.36 | 0.39 | 0.35 | 0.36 |
| M1 | | | | | | | | | | | | | | -1.40 | -1.41 | -1.03 | -3.72 | -4.01 | -2.56 | -2.20 | -2.41 | -1.47 |
| M2 | | | | | | | | | | | | | | | -0.04 | -0.24 | -0.70 | -1.14 | -3.71 | -2.56 | -2.02 | -0.81 |
| M3 | | | | | | | | | | | | | | | | -0.19 | -0.73 | -1.07 | -2.44 | -1.36 | -3.59 | -0.89 |
| M4 | | | | | | | | | | | | | | | | | -0.09 | -0.46 | -0.82 | -0.50 | -0.73 | -0.42 |
| M5 | | | | | | | | | | | | | | | | | | -0.77 | -1.51 | -0.73 | -1.49 | -0.61 |
| M6 | | | | | | | | | | | | | | | | | | | -0.59 | -0.14 | -0.58 | -0.24 |
| M7 | | | | | | | | | | | | | | | | | | | | 0.98 | -0.08 | 0.13 |
| M8 | | | | | | | | | | | | | | | | | | | | | -0.64 | -0.18 |
| M9 | | | | | | | | | | | | | | | | | | | | | | 0.18 |

Panel B: Portfolios Constructed by 200 Stocks

| | SHAR | SCHAR | SCHAR-r | LASSO | EN1 | EN2 | PLS1 | PLS2 | PLS3 | PCR1 | PCR2 | PCR3 | M1 | M2 | M3 | M4 | M5 | M6 | M7 | M8 | M9 | M10 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| HAR | -0.74 | -2.09 | 0.45 | -1.92 | -1.70 | -1.50 | 0.60 | -0.41 | 0.00 | -0.67 | -0.17 | 0.50 | **1.77** | **1.74** | **1.74** | 1.54 | 1.59 | 1.51 | **1.71** | 1.45 | **1.85** | 1.46 |
| SHAR | | -2.17 | **1.78** | -0.99 | -1.80 | -1.55 | 1.07 | 0.42 | 1.10 | 0.37 | 0.51 | 0.98 | **2.37** | **2.33** | **2.33** | **2.62** | **2.53** | **2.52** | **2.37** | **2.42** | **2.38** | **2.53** |
| SCHAR | | | **3.05** | 0.58 | -0.68 | -0.54 | **2.02** | 1.63 | **2.24** | 1.56 | 1.51 | **1.94** | **3.25** | **3.21** | **3.20** | **3.34** | **3.40** | **3.39** | **3.19** | **3.33** | **3.30** | **3.29** |
| SCHAR-r | | | | -2.45 | -2.06 | -1.83 | 0.18 | -1.03 | -0.60 | -1.01 | -0.55 | 0.07 | **1.99** | **2.09** | **2.11** | **2.00** | **2.07** | **1.93** | **2.03** | **1.74** | **1.95** | **1.85** |
| LASSO | | | | | -1.14 | -0.95 | **2.42** | **1.67** | **1.66** | 1.49 | **1.76** | **2.33** | **2.78** | **2.82** | **2.84** | **2.59** | **2.62** | **2.51** | **2.77** | **2.38** | **2.78** | **2.49** |
| EN1 | | | | | | 0.40 | **1.94** | 1.61 | **1.78** | 1.60 | **1.70** | **1.90** | **2.16** | **2.17** | **2.17** | **2.30** | **2.23** | **2.21** | **2.16** | **2.16** | **2.15** | **2.26** |
| EN2 | | | | | | | **1.77** | 1.45 | 1.62 | 1.44 | 1.54 | **1.73** | **2.00** | **2.02** | **2.02** | **2.15** | **2.08** | **2.06** | **2.00** | **2.01** | **1.99** | **2.12** |
| PLS1 | | | | | | | | -1.77 | -0.52 | -1.45 | -2.44 | -1.21 | 1.43 | 1.60 | 1.59 | 1.35 | 1.38 | 1.26 | 1.41 | 1.13 | 1.33 | 1.25 |
| PLS2 | | | | | | | | | 0.51 | -0.07 | 0.42 | 1.53 | **2.25** | **2.60** | **2.58** | **2.09** | **2.21** | **2.04** | **2.27** | **1.84** | **2.10** | **1.97** |
| PLS3 | | | | | | | | | | -0.57 | -0.15 | 0.41 | **2.80** | **2.80** | **2.86** | **2.83** | **2.94** | **2.84** | **2.87** | **2.66** | **2.67** | **2.68** |
| PCR1 | | | | | | | | | | | 0.40 | 1.26 | **2.25** | **2.24** | **2.26** | **2.02** | **2.06** | **1.96** | **2.23** | **1.89** | **2.32** | **1.92** |
| PCR2 | | | | | | | | | | | | **2.36** | **1.68** | **1.85** | **1.85** | 1.60 | 1.63 | 1.52 | **1.69** | 1.40 | 1.60 | 1.51 |
| PCR3 | | | | | | | | | | | | | 1.44 | 1.61 | 1.60 | 1.35 | 1.38 | 1.27 | 1.42 | 1.15 | 1.35 | 1.26 |
| M1 | | | | | | | | | | | | | | -0.08 | -0.19 | -0.12 | -0.19 | -0.22 | -0.59 | -0.24 | -0.55 | -0.21 |
| M2 | | | | | | | | | | | | | | | -0.37 | -0.08 | -0.17 | -0.17 | -0.27 | -0.15 | -0.20 | -0.17 |
| M3 | | | | | | | | | | | | | | | | -0.02 | -0.08 | -0.11 | -0.20 | -0.10 | -0.14 | -0.11 |
| M4 | | | | | | | | | | | | | | | | | -0.09 | -0.16 | -0.07 | -0.11 | -0.08 | -0.69 |
| M5 | | | | | | | | | | | | | | | | | | -0.12 | -0.04 | -0.07 | -0.06 | -0.13 |
| M6 | | | | | | | | | | | | | | | | | | | 0.00 | -0.03 | -0.03 | -0.08 |
| M7 | | | | | | | | | | | | | | | | | | | | -0.02 | -0.06 | -0.03 |
| M8 | | | | | | | | | | | | | | | | | | | | | -0.02 | -0.02 |
| M9 | | | | | | | | | | | | | | | | | | | | | | 0.00 |

*Notes: See notes to Table 3.7.

# Table 3.9: Comparison of Out-of-Sample Prediction Performance (30-minute)

Panel A: Portfolios Constructed by 10 Stocks

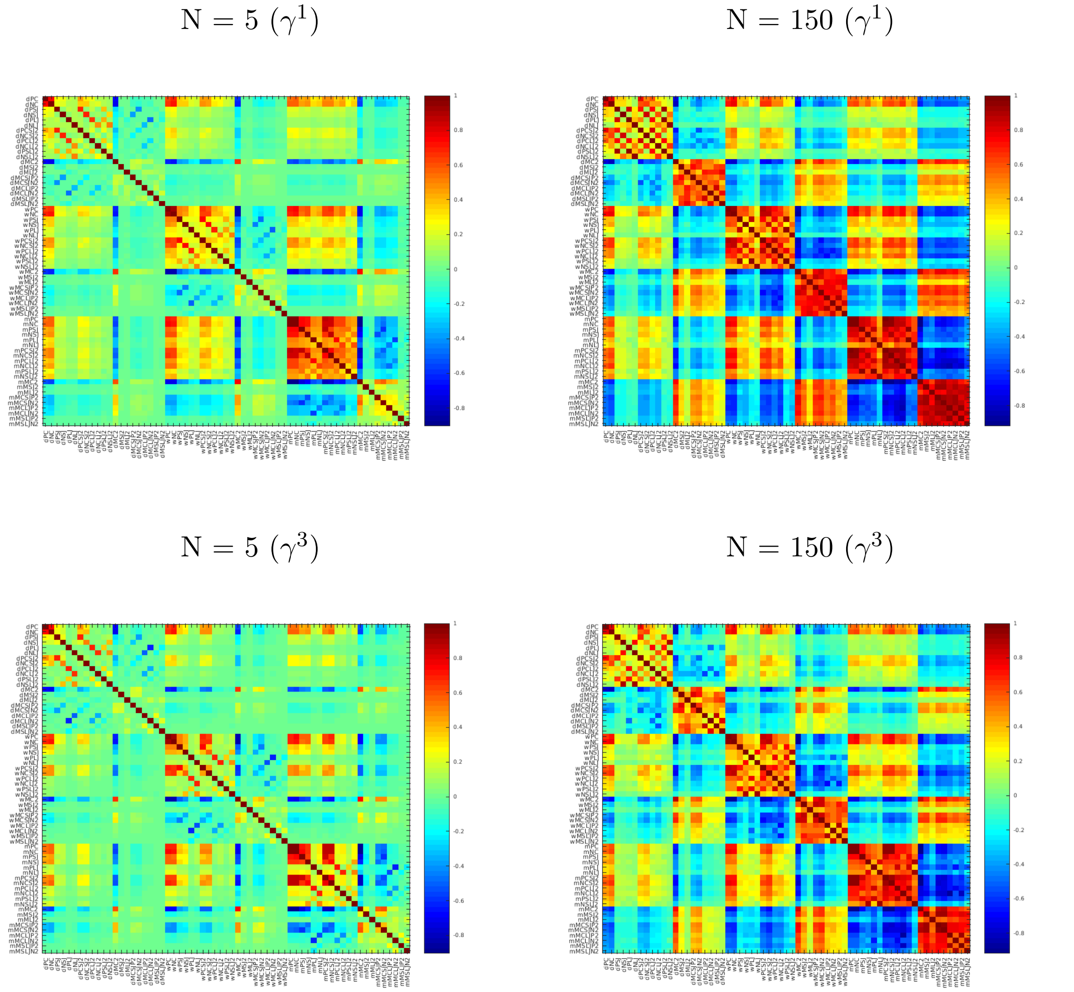
| | SHAR | SCHAR | SCHAR-r | LASSO | EN1 | EN2 | PLS1 | PLS2 | PLS3 | PCR1 | PCR2 | PCR3 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| HAR | -0.58 | -2.61 | -0.62 | -2.69 | -2.65 | -2.64 | 0.82 | -0.11 | -1.24 | 0.29 | 1.01 | 0.39 |
| SHAR | | -3.37 | -0.06 | -4.16 | -4.31 | -4.38 | 0.85 | 0.57 | -1.45 | **1.70** | 1.07 | 0.65 |
| SCHAR | | | **3.45** | -1.04 | -0.44 | -0.38 | **2.84** | **2.87** | **2.45** | **4.03** | **3.16** | **2.61** |
| SCHAR-r | | | | -4.18 | -4.55 | -4.34 | 0.97 | 0.67 | -1.36 | **1.80** | 1.23 | 0.73 |
| LASSO | | | | | **1.91** | 1.23 | **2.90** | **2.95** | **3.54** | **4.36** | **3.17** | **2.69** |
| EN1 | | | | | | 0.07 | **2.89** | **2.93** | **3.24** | **4.55** | **3.21** | **2.65** |
| EN2 | | | | | | | **2.89** | **2.98** | **4.79** | **5.13** | **3.23** | **2.64** |
| PLS1 | | | | | | | | -1.89 | -1.53 | -0.07 | 0.46 | -1.83 |
| PLS2 | | | | | | | | | -1.41 | 0.40 | **1.93** | 0.80 |
| PLS3 | | | | | | | | | | **3.16** | **1.79** | 1.29 |
| PCR1 | | | | | | | | | | | 0.23 | -0.11 |
| PCR2 | | | | | | | | | | | | -0.94 |

Panel B: Portfolios Constructed by 200 Stocks

| | SHAR | SCHAR | SCHAR-r | LASSO | EN1 | EN2 | PLS1 | PLS2 | PLS3 | PCR1 | PCR2 | PCR3 | M1 | M2 | M3 | M4 | M5 | M6 | M7 | M8 | M9 | M10 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| HAR | 0.17 | -2.10 | 0.08 | -0.34 | -0.29 | -0.38 | -0.37 | -1.45 | -0.56 | 0.56 | -0.58 | -0.85 | 0.37 | 0.33 | 0.30 | 0.28 | 0.26 | 0.26 | 0.25 | 0.26 | 0.24 | 0.24 |
| SHAR | | -2.82 | -0.32 | -1.09 | -1.03 | -1.13 | -0.31 | -1.70 | -1.34 | 0.48 | -0.37 | -0.48 | 0.58 | 0.48 | 0.39 | 0.35 | 0.29 | 0.29 | 0.26 | 0.27 | 0.25 | 0.23 |
| SCHAR | | | **2.83** | **2.38** | **2.50** | **2.30** | **1.96** | 1.32 | **1.99** | **2.95** | **1.88** | **1.73** | **3.10** | **3.08** | **3.00** | **3.04** | **2.94** | **2.91** | **2.91** | **2.78** | **2.89** | **2.90** |
| SCHAR-r | | | | -1.02 | -0.95 | -1.08 | -0.24 | -1.69 | -1.26 | 0.66 | -0.31 | -0.42 | 0.71 | 0.63 | 0.57 | 0.53 | 0.47 | 0.53 | 0.50 | 0.54 | 0.48 | 0.45 |
| LASSO | | | | | 0.61 | -0.57 | 0.21 | -1.14 | -1.08 | 1.25 | 0.14 | 0.02 | **2.81** | **2.75** | **2.75** | **2.60** | **2.12** | **3.68** | **3.55** | **2.37** | **3.56** | **3.45** |
| EN1 | | | | | | -0.85 | 0.16 | -1.22 | -1.14 | 1.22 | 0.08 | -0.04 | **2.75** | **2.65** | **2.58** | **2.41** | **2.03** | **4.07** | **3.94** | **2.55** | **3.86** | **3.66** |
| EN2 | | | | | | | 0.25 | -1.09 | -0.98 | 1.27 | 0.17 | 0.06 | **2.63** | **2.58** | **2.60** | **2.47** | **2.10** | **3.26** | **3.16** | **2.30** | **3.14** | **3.12** |
| PLS1 | | | | | | | | -1.31 | -0.44 | 0.71 | -1.03 | -1.72 | 0.48 | 0.45 | 0.42 | 0.40 | 0.39 | 0.39 | 0.38 | 0.41 | 0.38 | 0.37 |
| PLS2 | | | | | | | | | 0.60 | **1.80** | 1.18 | 1.01 | **1.89** | **1.86** | **1.81** | **1.81** | **1.76** | **1.88** | **1.86** | **1.80** | **1.85** | **1.87** |
| PLS3 | | | | | | | | | | 1.43 | 0.37 | 0.26 | **2.55** | **2.51** | **2.62** | **2.56** | **2.29** | **2.42** | **2.38** | **1.97** | **2.40** | **2.37** |
| PCR1 | | | | | | | | | | | -0.76 | -0.86 | -0.10 | -0.17 | -0.22 | -0.25 | -0.31 | -0.29 | -0.31 | -0.33 | -0.32 | -0.33 |
| PCR2 | | | | | | | | | | | | -1.57 | 0.53 | 0.50 | 0.47 | 0.45 | 0.44 | 0.45 | 0.44 | 0.46 | 0.43 | 0.43 |
| PCR3 | | | | | | | | | | | | | 0.62 | 0.58 | 0.55 | 0.54 | 0.53 | 0.54 | 0.53 | 0.56 | 0.52 | 0.52 |
| M1 | | | | | | | | | | | | | | -2.48 | -0.53 | -0.68 | -0.68 | -0.69 | -0.75 | -0.46 | -0.82 | -0.85 |
| M2 | | | | | | | | | | | | | | | -0.25 | -0.40 | -0.45 | -0.47 | -0.53 | -0.32 | -0.59 | -0.63 |
| M3 | | | | | | | | | | | | | | | | -0.29 | -0.32 | -0.28 | -0.35 | -0.22 | -0.40 | -0.39 |
| M4 | | | | | | | | | | | | | | | | | -0.21 | -0.16 | -0.23 | -0.14 | -0.27 | -0.28 |
| M5 | | | | | | | | | | | | | | | | | | -0.01 | -0.07 | -0.05 | -0.09 | -0.11 |
| M6 | | | | | | | | | | | | | | | | | | | -0.66 | -0.06 | -1.19 | -0.46 |
| M7 | | | | | | | | | | | | | | | | | | | | 0.02 | -0.21 | -0.21 |
| M8 | | | | | | | | | | | | | | | | | | | | | -0.05 | -0.07 |
| M9 | | | | | | | | | | | | | | | | | | | | | | -0.10 |

*Notes: See notes to Table 3.7.

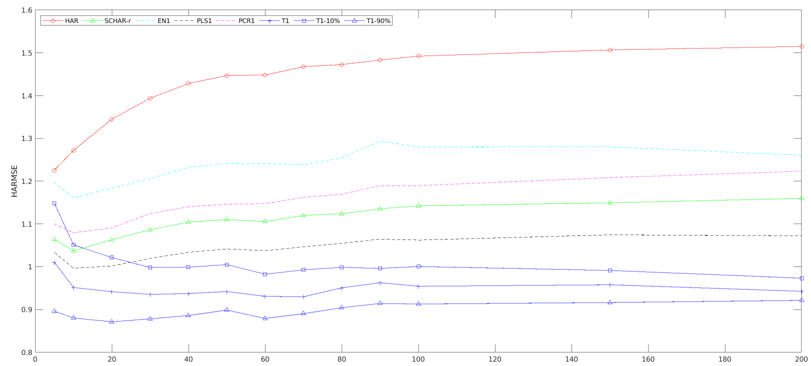
Figure 3.1: Correlations Between Realized Components (15-minute)



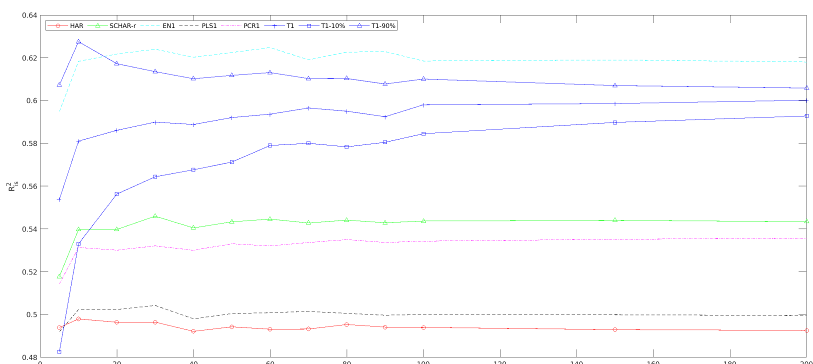

*Notes: See notes to Table 3.1. This figure displays the average of correlations between each two separated realized components based on 200 randomly constructed portfolios.

Figure 3.2: Median Value of Forecasting Performance

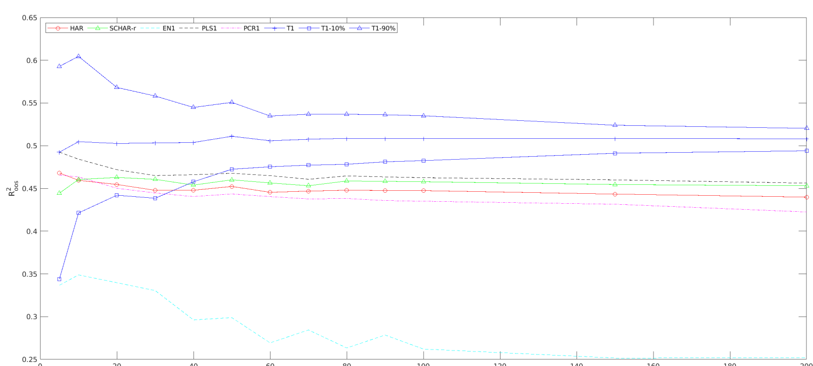
HARMSE




$R^2_{is}$

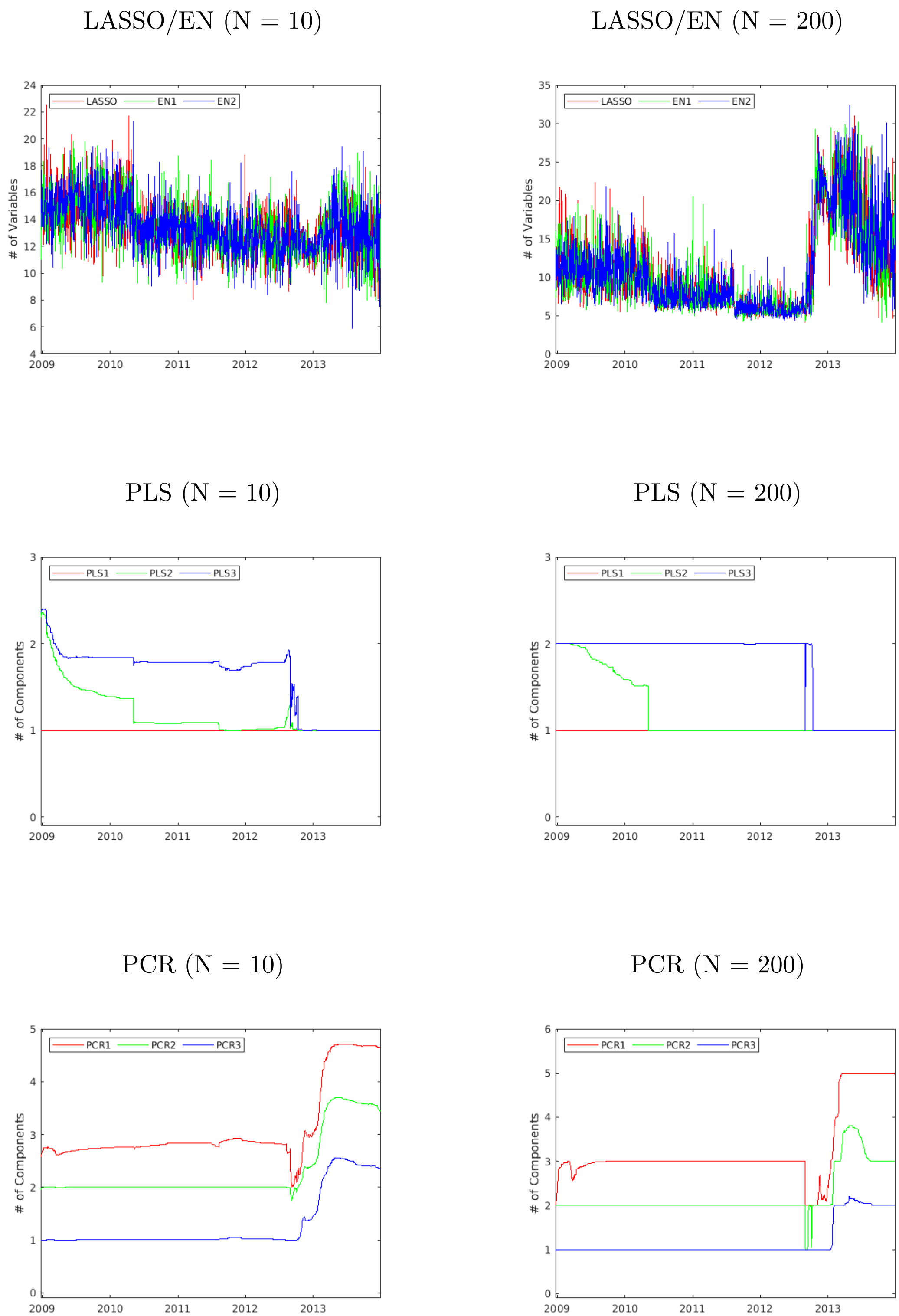



$R^2_{oos}$




*Notes: See notes to Table 3.1. This figure reports the median value of HARMSE, $R^2_{is}$, and $R^2_{oos}$, respectively, for each model based on 200 randomly selected portfolios. The range of portfolio dimension is from 5 to 200. T1 refers to the best-performing sparse model, T1-10% and T1-90% denote the 10% and 90% quantiles of each evaluation criterion for the best-performing sparse model.

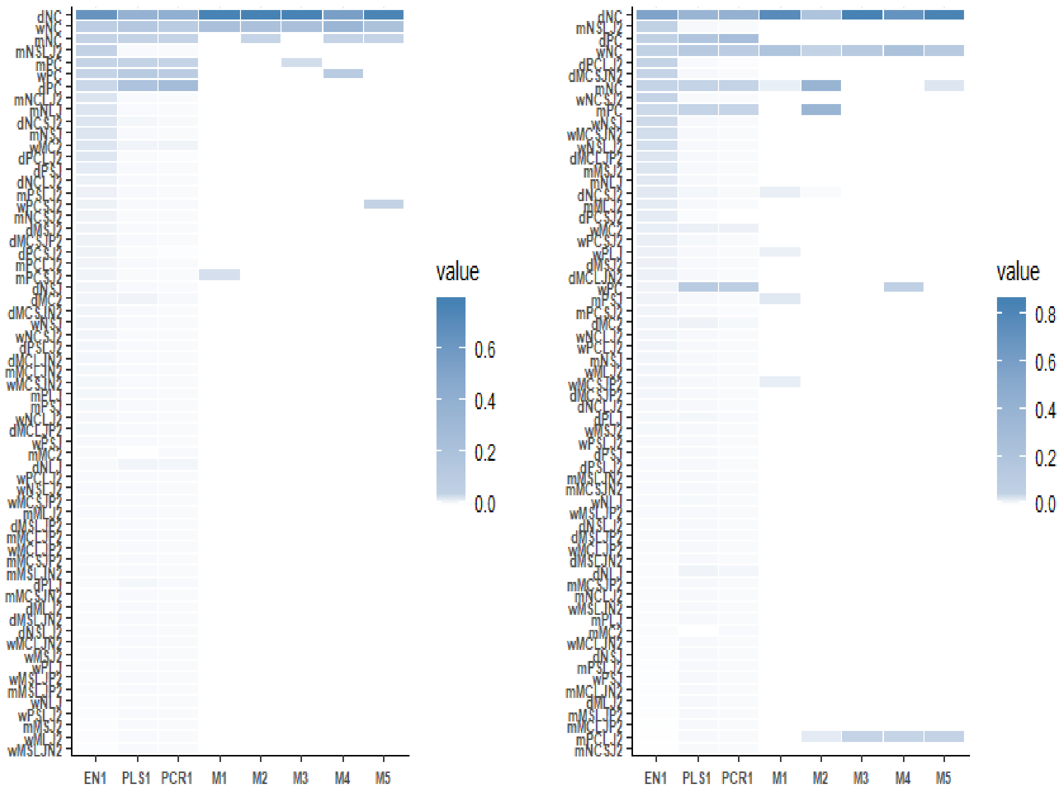
Figure 3.3: Model Complexity (15-minute, $\gamma^1$)




*Notes: See notes to Table 3.1. This figure displays the average of correlations between each two separated realized components based on 200 randomly constructed portfolios.

Figure 3.4: Feature Importance (5-minute, $\gamma^1$)

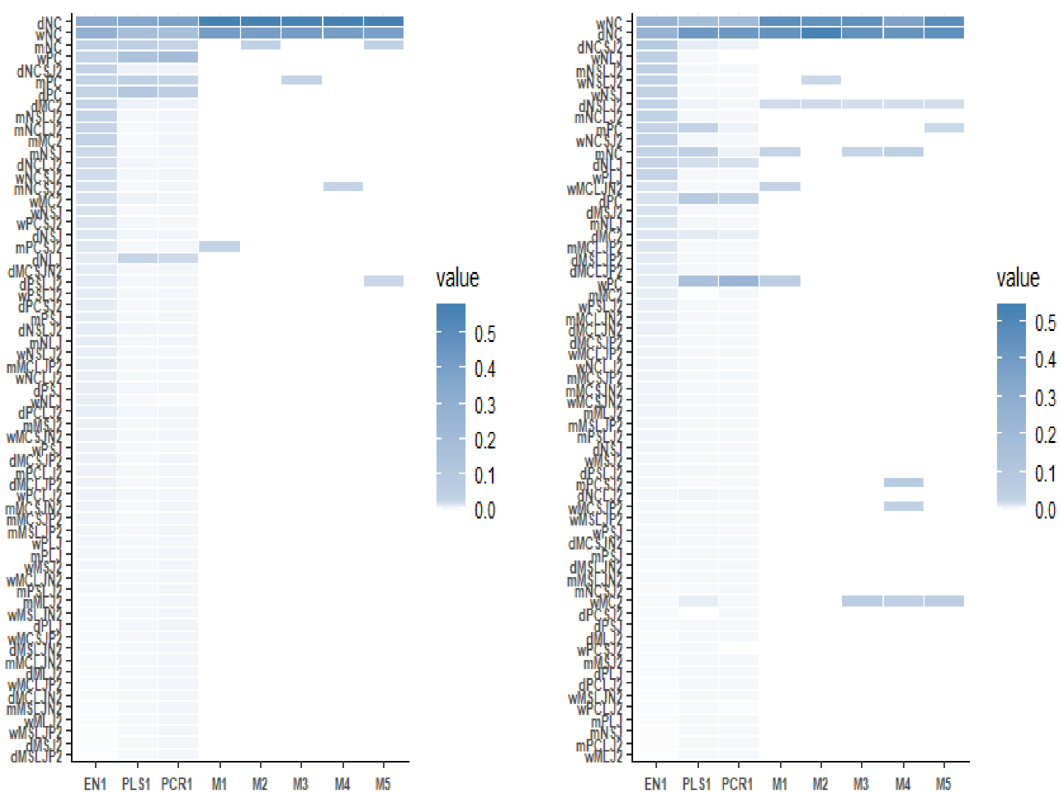
N = 10 N = 200




*Notes: See notes to Table 3.1. This figure displays the average of rankings of each decomposed realized component in terms of in-sample prediction contribution over time and 200 randomly constructed portfolios. Each column represents a model, with predictors ordered by the ranks based on model EN1. The color gradients indicate the most influential (dark blue) to the least influential (white) predictors.

Figure 3.5: Feature Importance (15-minute, $\gamma^1$)

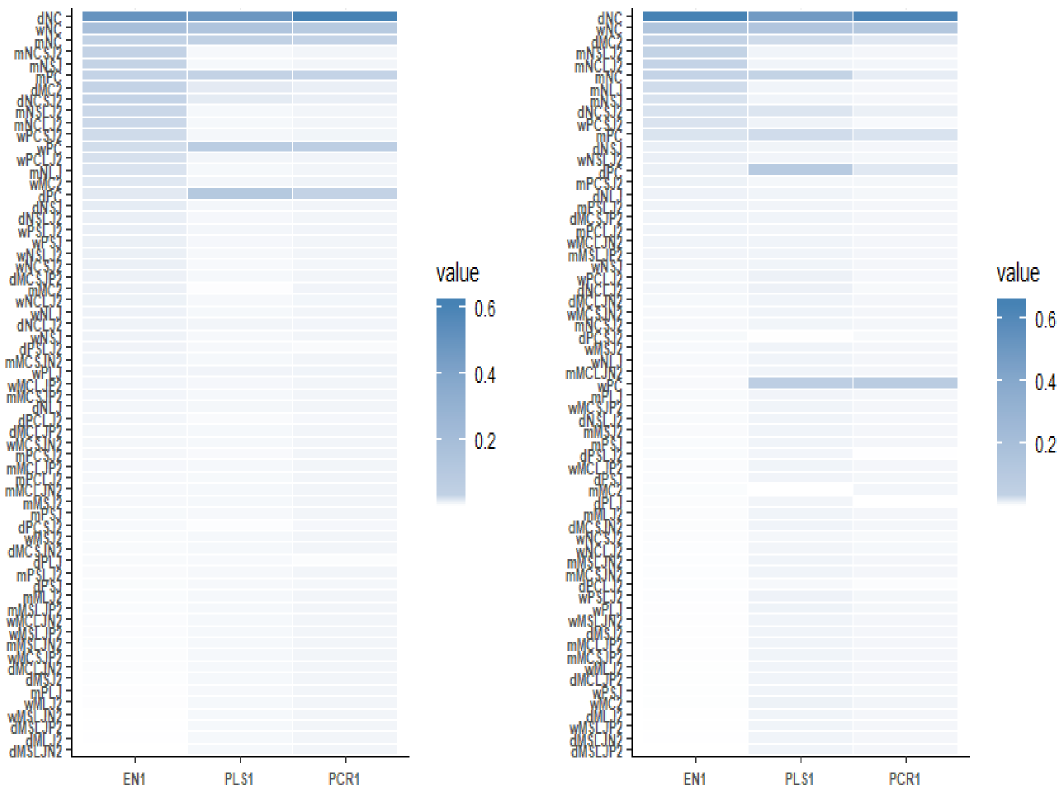
N = 10 N = 200




*Notes: See notes to Figure 3.4.

Figure 3.6: Feature Importance (30-minute, $\gamma^1$)

N = 10 N = 200




*Notes: See notes to Figure 3.4.

# Bibliography

Aït-Sahalia, Y. and Jacod, J. 2012. Analyzing the spectrum of asset returns: Jump and volatility components in high frequency data. *Journal of Economic Literature* 50:1007–1050.

Amaya, D., Christoffersen, P., Jacobs, K., and Vasquez, A. 2015. Does realized skewness predict the cross-section of equity returns? *Journal of Financial Economics* 118:135–167.

Andersen, T. G., Bollerslev, T., and Diebold, F. X. 2007. Roughing it up: Including jump components in the measurement, modeling, and forecasting of return volatility. *The review of economics and statistics* 89:701–720.

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Labys, P. 2003. Modeling and forecasting realized volatility. *Econometrica* 71:579–625.

Bai, J. and Ng, S. 2006a. Confidence intervals for diffusion index forecasts and inference for factor-augmented regressions. *Econometrica* 74:1133–1150.

Bai, J. and Ng, S. 2006b. Evaluating latent and observed factors in macroeconomics and finance. *Journal of Econometrics* 131:507–537.

Bai, J. and Ng, S. 2008. Forecasting economic time series using targeted predictors. *Journal of Econometrics* 146:304–317.

Barndorff-Nielsen, O. E., Graversen, S. E., Jacod, J., and Shephard, N. 2006. Limit theorems for bipower variation in financial econometrics. *Econometric Theory* 22:677–719.

Barndorff-Nielsen, O. E., Kinnebrouk, S., and Shephard, N. 2010. Measuring downside risk: realised semivariance. pp. 117–136.

Barndorff-Nielsen, O. E. and Shephard, N. 2004. Power and bipower variation with stochastic volatility and jumps. *Journal of financial econometrics* 2:1–37.

BERNARD, V. L. AND THOMAS, J. K. 1989. Post-earnings-announcement drift: Delayed price response or risk premium? *Journal of Accounting Research* 27:1–36.

BOLLERSLEV, T., LI, J., PATTON, A. J., AND ROGIER, Q. 2019a. Realized semicovariances. *Working Paper, Duke University, Erasmus University Rotterdam* .

BOLLERSLEV, T., LI, S. Z., AND ZHAO, B. 2019b. Good volatility, bad volatility, and the cross-section of stock returns. *Journal of Financial and Quantitative Analysis,* forthcoming.

BOLLERSLEV, T. AND TODOROV, V. 2011. Estimation of jump tails. *Econometrica* 79:1727–1783.

BOLLERSLEV, T., TODOROV, V., AND XU, L. 2015. Tail risk premia and return predictability. *Journal of Financial Economics* 118:113–134.

CAMPBELL, J. Y. AND THOMPSON, S. B. 2008. Predicting excess stock returns out of sample: Can anything beat the historical average? *Review of Financial Studies* 21:1509–1531.

CAMPBELL, J. Y. AND VUOLTEENAHO, T. 2004. Bad beta, good beta. *American Economic Review* 94:1249–1275.

CORSI, F. 2009. A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics* 7:174–196.

CORSI, F., PIRINO, D., AND RENO, R. 2010. Threshold bipower variation and the impact of jumps on volatility forecasting. *Journal of Econometrics* 159:276–288.

DIEBOLD, F. X. AND MARIANO, R. S. 1995. Comparing predictive accuracy. *Journal of Business & Economic Statistics* 13:134–144.

DUONG, D. AND SWANSON, N. R. 2011. Volatility in discrete and continuous-time models: A survey with new evidence on large and small jumps. 27B:179–233.

Duong, D. and Swanson, N. R. 2015. Empirical evidence on the importance of aggregation, asymmetry, and jumps for volatility prediction. *Journal of Econometrics* 187:606–621.

Fang, N., Jiang, W., and Luo, R. 2017. Realized semivariances and the variation of signed jumps in china's stock market. *Emerging Markets Finance and Trade* 53:563 – 586.

Feunou, B., Jahan-Parvar, M. R., and Tedongap, R. 2013. Modeling market downside volatility. *Review of Finance* 17:443–481.

Feunou, B., Jahan-Pravar, M. R., and Okou, C. 2018. Downside variance risk premium. *Journal of Financial Econometrics* 16:341–383.

Giglio, S. and Xiu, D. 2019. Asset pricing with omitted factors. *Working Paper, Yale University, University of Chicago* .

Gu, S., Kelly, B., and Xiu, D. 2019. Empirical asset pricing via machine learning. *Working Paper, University of Chicago, Yale University, University of Chicago* .

Guo, H., Wang, K., and Zhou, H. 2015. Good jumps, bad jumps, and conditional equity premium. *Working Paper, University of Cincinnati* .

Jacod, J. 2008. Asymptotic properties of realized power variations and related functionals of semimartingales. *Stochastic Processes and their Applications* 118:517–559.

Jiang, G. J. and Yao, T. 2013. Stock price jumps and cross-sectional return predictability. *Journal of Financial and Quantitative Analysis* 48:1519–1544.

Kelly, B. and Pruitt, S. 2013. Market expectations in the cross-section of present values. *The Journal of Finance* 68:17211756.

Kelly, B. and Pruitt, S. 2015. The three-pass regression filter: A new approach to forecasting using many predictors. *Journal of Econometrics* 186:294316.

Kelly, B., Pruitt, S., and Su, Y. 2017. Some characteristics are risk exposures, and the rest are irrelevant. *Working Paper, University of Chicago* .

Li, J., Todorov, V., Tauchen, G., and Chen, R. 2017. Mixed-scale jump regressions with bootstrap inference. *Journal of Econometrics* 201:417–432.

Livnat, J. and Mendenhall, R. R. 2006. Comparing the post-earnings announcement drift for surprises calculated from analyst and time series forecasts. *Journal of Accounting Research* 44:177–205.

Maheu, J. M. and McCurdy, T. H. 2004. News arrival, jump dynamics, and volatility components for individual stock returns. *Journal of Finance* 59:755–793.

Mancini, C. 2009. Non-parametric threshold estimation for models with stochastic diffusion coefficient and jumps. *Scandinavian Journal of Statistics* 36:270–296.

Neuberger, A. 2012. Realized skewness. *Review of Financial Studies* 25:3423–3455.

Patton, A. J. and Sheppard, K. 2015. Good volatility, bad volatility: Signed jumps and the persistence of volatility. *Review of Economics and Statistics* 97:683–697.

Petersen, M. A. 2009. Estimating standard errors in finance panel data sets: Comparing approaches. *Review of Financial Studies* 22:435–480.

Rapach, D. E., Strauss, J. K., and Zhou, G. 2013. International stock return predictability: what is the role of the united states? *The Journal of Finance* 68:16331662.

Rossi, A. G. and Timmermann, A. 2015. Modeling covariance risk in merton's icapm. *Review of Financial Studies* 28:1428–1461.

Stock, J. H. and Watson, M. W. 2002a. Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association* 97:1167–1179.

Stock, J. H. and Watson, M. W. 2002b. Macroeconomic forecasting using diffusion indexes. *Journal of Business & Economic Statistics* 20:147–162.

Stock, J. H. and Watson, M. W. 2006. Forecasting with many predictors. *Handbook of Economic Forecasting* 1:515–554.

Tibshirani, R. 1996. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 58:267–288.

Todorov, V. and Tauchen, G. 2010. Activity signature functions for high-frequency data analysis. *Journal of Econometrics* 154:125–138.

Woodward, G. and Anderson, H. M. 2009. Does beta react to market conditions? estimates of 'bull' and 'bear' betas using a nonlinear market model with an endogenous threshold parameter. *Quantitative Finance* 9:913–924.

Xiong, J. X., Idzorek, T. M., and Ibbotson, R. G. 2016. The economic value of forecasting left-tail risk. *Journal of Portfolio Management* 42:114–123.

Yan, S. 2011. Jump risk, stock returns, and slope of implied volatility smile. *Journal of Financial Economics* 99:216–233.

Zou, H. and Hastie, T. 2005. Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 67:301–320.