Asymmetries at the core of short-term return predictability*

Johannes Breckenfelder
European Central Bank

Bastien Buchwalter
SKEMA Business School
and Université Côte d'Azur

Roméo Tédongap
ESSEC Business School

June 23, 2025

Abstract

We propose a novel measure that improves short-term equity return predictability by analyzing truncated high-frequency return distributions. We segment returns into core and tail components, focusing on core and tail asymmetries — differences between 'typical' or 'extreme' upside and downside variances. Our empirical findings show that these predictors achieve an in-sample adjusted R^2 of about 7% and an out-of-sample R^2 exceeding 4% for one-month-ahead market return forecasts, outperforming traditional predictors such as valuation ratios and the variance risk premium. The core asymmetry, rather than the tail asymmetry, drives the predictive power.

JEL Classification: C14, C58, G12, G17

Keywords: short-term predictor, equity returns, core asymmetries, tail asymmetries

*We thank John Campbell, Pierre Collin-Dufresne, Simone Manganelli, Albert Menkveld, Marie Hoerova. The views expressed in this paper are those of the authors and do not necessarily reflect those of the European Central Bank or the Eurosystem.
Emails: Johannes Breckenfelder johannes.breckenfelder@ecb.europa.eu; Bastien Buchwalter bastien.buchwalter@skema.edu; Roméo Tédongap tedongap@essec.edu

1 Introduction

The predictability of equity returns has long been a central topic in financial research. Predictors such as valuation ratios and macroeconomic indicators have shown varying degrees of success for different forecasting horizons. Campbell and Thompson (2008) highlighted the long-term return predictability by valuation ratios such as dividend-price and earnings-price ratios. Although robust over longer periods, these metrics struggle to provide predictions at monthly horizons and have limited out-of-sample robustness. Similarly, Bollerslev, Tauchen, and Zhou (2009) investigated the variance risk premium (VRP), showing its usefulness for medium-term forecasts and noting its reduced effectiveness in shorter intervals. We aim to improve the predictability of short-term equity returns by introducing a novel high-frequency-based predictor. We segment the monthly distribution of high-frequency returns into core and tail components while distinguishing between negative and positive returns, i.e. downside and upside of the return distribution.

Previous studies have predominantly emphasized tail risks, focusing on extreme market movements as key drivers of return dynamics. For example, Kelly and Jiang (2014) show the importance of tail risk measures in asset pricing. While this approach captures the kurtosis of the return distribution, it often overlooks the informational value embedded within the return distribution's more typical, less extreme core. Barndorff-Nielsen, Kinnebrock, and Shephard (2008) introduce measures for downside and upside risk of the return distribution containing, respectively, all returns below and above a certain threshold. The difference of upside and downside, henceforth the asymmetry, provides valuable information on the skewness of the return distribution. Our methodology builds on both approaches by distinguishes not only between tails and cores, but also between the downside and upside part of the return distribution, hence, combining the information stemming from skewness and kurtosis in a single setting.

We calculate realized variances across four segments of the return distribution: downside tail, downside core, upside core, and upside tail. The core and tail components allow us to calculate risk measures that capture, respectively, typical and extreme return behaviors. Further, core asymmetry is defined as the difference between typical upside and downside variances.

Similarity, tail asymmetry captures the differences between extreme positive and negative variances. By systematically varying the threshold, α, which defines core and tail components, we find the threshold, α^*, which optimizes the predictive power, measured through adjusted R-square values.

Using a monthly vector-retrogression (VAR) framework, our predictors achieves an adjusted R-squared of 7.25%.The optimal threshold α^* to distinguish between tails and cores is approximately 28.77%, where the adjusted R-squared reaches its maximum. This is significantly higher than what the literature uses as a threshold for defining the tail, e.g. 1%, 5% or 10% thresholds. When we include only tail components, the predictive power declines, with adjusted R-squared values peaking below 4%. In contrast, the predictive power when only the core components are included is over 6%. In the empirical findings we highlight the role of core asymmetries in predicting one-month-ahead market returns.

The model achieves an out-of-sample R^2 of 4.1% at the optimal threshold for one-month-ahead market returns, outperforming traditional predictors like the variance risk premium (VRP). Rolling and expanding window analyses further validate the stability of the model, with OOS R^2 consistently exceeding 3%.

To ensure the robustness of our findings, we performed several additional analyses. These include testing alternative thresholds to distinguish between downside and upside parts of the return distribution, such as median-based and zero-based splits, which yield qualitatively similar results. We also compare traditional and proposed predictors, finding that incorporating conventional variables alongside core and tail asymmetry offers marginal improvements in predictive power. Our measures outperform traditional metrics in both in-sample and out-of-sample analysis, such as VRP and valuation ratios. While our predictor performs well in the short term, traditional predictors perform better in the medium term of three months and onward. This suggests a complementary relationship between the predictors.

Our paper contributes to two intertwined strands of literature: the growing body of work emphasizing the role of asymmetric risks in asset pricing and the vast literature on equity premium predictability. We propose a novel, model-free framework for forecasting short-term equity returns by analyzing truncated high-frequency return distributions. Specifically, we decompose the returns into core and tail components while distinguishing between negative and

positive returns. The latter distinction allows us to construct asymmetry measures for core and tail components. Empirically, we show that core asymmetry- capturing differences in typical upside and downside variation- drives the predictive power of our approach, significantly improving both in-sample and out-of-sample forecasts of monthly market returns.

In the context of asymmetric risk, our work is based on recent contributions highlighting the importance of downside versus upside variation in shaping return dynamics. Kilic and Shaliastovich (2019) and Feunou, Jahan-Parvar, and Okou (2018) decompose the variance risk premium into components "good" (upside) and "bad" (downside), showing their differential predictive power. Bekaert, Engstrom, and Ermolov (2015); Bekaert and Engstrom (2017) develop regime-switching, non-Gaussian models in which bad environments disproportionately affect consumption and asset prices. Feunou, Fontaine, and Tédongap (2013) introduce the BiN-GARCH model to disentangle the downside risk from the upside uncertainty, and Feunou, Jahan-Parvar, and Tédongap (2016) propose and evaluate parametric models where conditional skewness is defined as the standardized difference between the upside and downside variances. In the option-implied domain, Huang and Li (2019) find that variance asymmetry is a priced factor in the cross-section of stock returns. At shorter horizons, Breckenfelder and Tédongap (2012) show that asymmetric realized volatility extracted from high-frequency data improves return forecasts. Our contribution complements these studies by introducing a distributional truncation method to isolate and quantify asymmetry within specific regions of the return distribution, providing a flexible and intuitive forecasting tool.

Our paper also relates to the extensive literature on equity premium prediction, which has evolved significantly in response to critiques such as Goyal and Welch (2008), who questioned the robustness of many popular predictors. In response, Campbell and Thompson (2008) demonstrated that predictors may still be valuable, especially for longer horizons. Subsequent work expanded the set of predictors, drawing from macroeconomic indicators (Cooper and Priestley, 2009; Møller and Rangvid, 2015; Belo and Yu, 2013; Piazzesi, Schneider, and Tuzel, 2007), variance-based measures (Bollerslev, Tauchen, and Zhou, 2009; Bakshi, Panayotov, and Skoulakis, 2011; Bekaert and Hoerova, 2014), sentiment indicators (Hirshleifer, Hou, and Teoh, 2009; Rapach, Ringgenberg, and Zhou, 2016; Huang et al., 2015), and cross-sectional metrics (Kelly and Jiang, 2014; Pollet and Wilson, 2010; Maio, 2016; Colacito et al., 2016). Other

studies have explored technical and commodity-based predictors (Neely et al., 2014; Driesprong, Jacobsen, and Maat, 2008). More recently, Goyal, Welch, and Zafirov (2024) revisited many of these predictors using extended data and found that several lose statistical significance in particular out-of-sample. In contrast, our high-frequency, asymmetry-based predictor shows statistical robustness and economic relevance at short horizons, offering a new direction in searching for reliable return forecasts.

The remainder of the paper is organized as follows: Section 2 outlines the framework and variance measures. Section 3 describes the data and methodology. Section 4 presents the results, including in-sample and out-of-sample predictability and benchmarking against traditional predictors. Section 5 examines robustness, and Section 6 concludes with key findings and implications.

2 Framework

As Campbell and Thompson (2008) pointed out, the abundance of historical data allows us to directly extract risk measures from high-frequency data without relying on additional assumptions. This approach contrasts with the variance risk premium (VRP), which relies on implied volatility and involves numerous assumptions and interpolations.

Over the last decade, forecasting literature has increasingly focused on the behavior of the tails of return distributions, providing both parametric and non-parametric approaches for quantifying these extreme values. Building on these insights, we aim to capture the dynamics of asset returns by considering the entire distribution of high-frequency returns rather than limiting our analysis to the tails.

In line with the risk-return trade-off suggested by Merton (1973), much of the financial literature has focused on total return variance as the primary predictor of equity risk. However, we argue that this approach can be misleading. Behavioral biases cause equity investors to perceive and treat different segments of the return distribution unequally, leading to heterogeneous risk assessments. We extend their approach by decomposing these risks into core and tail components. Specifically, we divide the return distribution into four segments: downside tail, downside core, upside core, and upside tail. While the tail has been analyzed in prior

studies, the core — which represents more typical return behavior— has often been overlooked since Campbell and Thompson (2008) showed that valuation ratios tend to outperform traditional mean-based approaches. Using this insight, we propose an approach based on the skewness and kurtosis of the high-frequency return distribution. Skewness may account for asymmetry between upside and downside variations within a specific period, such as a month. Kurtosis captures the tail thickness of the high-frequency return distribution and is crucial for determining the optimal cut-offs between the tail and core parts of the distribution.

Using this insight, we propose an approach based on the skewness and kurtosis of the high-frequency return distribution. We consider skewness because upside and downside variations may not be symmetric within each time period, such as a month. We consider kurtosis because the thickness of the tails influences the variation within the tails. By segmenting the return distribution using quantiles rather than a fixed value—such as zero—higher kurtosis implies lower variation in the tails, meaning returns are closer to each other. In line with Campbell and Thompson (2008), we use the mean to separate upside and downside variations before determining the optimal threshold, denoted as alpha, within each segment.[1]

We start by considering five-minute log-returns in month t. The i^{th} log-return in month t is denoted $r_{i,t}$ and there a n_t five-minute log-returns in month t. Following Barndorff-Nielsen, Kinnebrock, and Shephard (2008), as a first step, we distinguish between downside and upside risks; that is, we argue that investors perceive downside risk and upside potential differently. There are three thresholds to distinguish between downside and upside parts of the return distribution, i.e. median-based, mean-based and zero-based splits. For the ease of notation we rely on the monthly median of the five-minute log returns, denoted $q_t(1/2)$. We note that our results are qualitatively similar for all three thresholds. Independent of the threshold, downside (upside) parts of the return distribution will henceforth be denoted with a - (+) sign. Distinguishing between downside and upside also considers the skewness of the return distribution.

Building on the literature that emphasizes the role of distributional tails, we further refine the decomposition of the return distribution by partitioning both the downside and upside

[1]Using the mean provides the highest predictive power, though the choice is not critical. Splitting by the median or zero yields qualitatively similar results. Choosing zero, however, places more weight on outliers and diverges from the concept of a truncated return distribution. See the Appendix for more details.

regions into *core* (C) and *tail* (T) components. Specifically, we introduce a threshold parameter $0 < \alpha < 1/2$ and define the corresponding quantiles $q_t(\alpha)$ and $q_t(1-\alpha)$ as the α-th and $(1-\alpha)$-th quantile returns at time t, respectively. This partition divides the monthly return distribution into four regions: downside tail, downside core, upside core, and upside tail. The number of observations that fall into each region reflects the frequency and distribution of extreme and typical returns on both sides of the distribution. Formally, we have:

$$\begin{aligned} n_t^{T-}(\alpha) &= \sum_{i=1}^{n_t} I\left(r_{i,t} < q_t(\alpha)\right) \\ n_t^{C-}(\alpha) &= \sum_{i=1}^{n_t} I\left(q_t(\alpha) \le r_{i,t} < q_t(1/2)\right) \\ n_t^{C+}(\alpha) &= \sum_{i=1}^{n_t} I\left(q_t(1-\alpha) > r_{i,t} \ge q_t(1/2)\right) \\ n_t^{T+}(\alpha) &= \sum_{i=1}^{n_t} I\left(r_{i,t} \ge q_t(1-\alpha)\right), \end{aligned} \tag{1}$$

where $I(\cdot)$ is the indicator function. For record $n_t^{T-} + n_t^{C-} + n_t^{C+} + n_t^{T+} = n_t$. It follows that the average returns with these four components are defined as

$$\begin{aligned} \mu_t^{T-}(\alpha) &= \frac{1}{n_t^{T-}(\alpha)} \sum_{i=1}^{n_t} r_{i,t} I\left(r_{i,t} < q_t(\alpha)\right) \\ \mu_t^{C-}(\alpha) &= \frac{1}{n_t^{C-}(\alpha)} \sum_{i=1}^{n_t} r_{i,t} I\left(q_t(\alpha) \le r_{i,t} < q_t(1/2)\right) \\ \mu_t^{C+}(\alpha) &= \frac{1}{n_t^{C+}(\alpha)} \sum_{i=1}^{n_t} r_{i,t} I\left(q_t(1-\alpha) > r_{i,t} \ge q_t(1/2)\right) \\ \mu_t^{T+}(\alpha) &= \frac{1}{n_t^{T+}(\alpha)} \sum_{i=1}^{n_t} r_{i,t} I\left(r_{i,t} \ge q_t(1-\alpha)\right). \end{aligned} \tag{2}$$

Finally, we define the monthly downside and upside tail realized variances as

$$\begin{aligned} RV_{t-1:t}^{T-}(\alpha) &= \frac{1}{n_t^{T-}(\alpha)} \sum_{i=1}^{n_t} \left(r_{i,t} - \mu_t^{T-}(\alpha)\right)^2 I\left(r_{i,t} < q_t(\alpha)\right) \\ RV_{t-1:t}^{T+}(\alpha) &= \frac{1}{n_t^{T+}(\alpha)} \sum_{i=1}^{n_t} \left(r_{i,t} - \mu_t^{T+}(\alpha)\right)^2 I\left(r_{i,t} \ge q_t(1-\alpha)\right). \end{aligned} \tag{3}$$

Similarly, the monthly downside and upside core realized variances are respectively

$$
\begin{aligned}
RV^{C-}_{t-1:t}(\alpha) &= \frac{1}{n^{C-}_t(\alpha)} \sum_{i=1}^{n_t} \left(r_{i,t} - \mu^{C-}_t(\alpha) \right)^2 I\left(q_t(\alpha) \leq r_{i,t} < q_t(1/2) \right) \\
RV^{C+}_{t-1:t}(\alpha) &= \frac{1}{n^{C+}_t(\alpha)} \sum_{i=1}^{n_t} \left(r_{i,t} - \mu^{C+}_t(\alpha) \right)^2 I\left(q_t(1-\alpha) > r_{i,t} \geq q_t(1/2) \right)
\end{aligned}
\tag{4}
$$

We argue that extracting these risk measures from high-frequency data provides a predictor of equity risk without relying on assumptions, in contrast to traditional predictors like the variance risk premium. By decomposing the return distribution into four parts (downside tail, downside core, upside core, upside tail), we capture a more nuanced view of the return dynamics. Our approach, which uses high-frequency log returns and computes variance across different thresholds, contributes to understanding both extreme and typical return behavior.

Each of the four variances defined in equations (4) is valued differently by equity investors, and this disaggregation of equity risk can be important for enhancing our understanding of the equity risk-return tradeoff. To analyze the risk-return tradeoff, we follow the vector autoregression (VAR) methodology and define the state vector

$$
z_t^\top = \left[\; r_{t-1:t} \quad RV^{T-}_{t-1:t}(\alpha) \quad RV^{C-}_{t-1:t}(\alpha) \quad RV^{C+}_{t-1:t}(\alpha) \quad RV^{T+}_{t-1:t}(\alpha) \; \right] \tag{5}
$$

where the monthly excess log return is given by

$$
r = log(1 + R^{SP500}) - log[(1 + r^f)^{(1/12)}],
$$

and the realized variance measures are likewise expressed in monthly terms. The dependent vector z_t in equation (5) may be augmented to include other potential return predictors such as the VRP, PD ratio, etc. We assume that the vector z_t has the following VAR dynamics

$$
z_t = \theta_0 + \Theta z_{t-1} + u_t \;\text{ where }\; \mathbb{E}[u_t] = 0 \;\text{ and }\; \mathbb{E}\left[u_t u_t^\top \right] = \Sigma \tag{6}
$$

where the first equation addresses returns predictability by the remaining state variables.

We compute the four risk measures for various threshold levels. By systematically varying α, we can determine the optimal α which maximizes the adjusted R-squared of the VAR depicted

in equation (6).

We also use combined risk measures to capture the asymmetry between the upside and downside. Our core and tail asymmetry measures are defined as follows:

$$\begin{aligned} RV^{\Delta C}_{t-1:t}(\alpha) =& RV^{C+}_{t-1:t}(\alpha) - RV^{C-}_{t-1:t}(\alpha) \\ RV^{\Delta T}_{t-1:t}(\alpha) =& RV^{T+}_{t-1:t}(\alpha) - RV^{T-}_{t-1:t}(\alpha). \end{aligned}$$

where $RV^{\Delta C}_{t-1:t}$ and $RV^{\Delta T}_{t-1:t}$ are the core and tail asymmetry respectively. This allows us to rewrite the state vector depicted in equation (5) as

$$z_t^\top = \left[\; r_{t-1:t} \quad 2RV^{T-}_{t-1:t}(\alpha) \quad 2RV^{T+}_{t-1:t}(\alpha) \quad RV^{\Delta C}_{t-1:t}(\alpha) \quad RV^{\Delta T}_{t-1:t}(\alpha) \; \right],$$

a reformulation that preserves the information content in the VAR system.

Finally, consistent with the insights of Bansal and Kiku (2011), multihorizon return predictability fundamentally depends on the distributional properties of multi-period returns, particularly their first and second moments. Subsequently, the necessary term structures of expected returns and (co)variances are readily derived by leveraging the recursive nature of the VAR framework.

3 Data

To compute monthly predictors of tail and core variances of the return distribution, we use high-frequency equity logarithmic (log) returns. Olsen Financial Services provides the raw data. The sample period ranges from January 1, 1990, to December 31, 2019. The raw data set contains 2,954,439 one-minute observations. We delete 28 outliers that are abnormally small or large. Further, we eliminate all trading days with less than 200 observations, yielding 2,952,909 one-minute observations.

Next we establish a five-minute grid according to the following grid xxh04, xxh09, xxh14, etc., henceforth 04-09 grid. In case of a missing value we take the last available value. The first return of the day capture the returns between the opening price and the first observation of the 04-09 grid. Further, no trading prices after closing are considered.

Finally, we remove overnight returns as the paper focuses on high frequency returns. This yields 576,726 five-minute returns over a sample period of 360 months, from January 1990 to December 2019. Unlike other empirical papers such as Ang et al. (2006) and Drechsler and Yaron (2010) which winsorize their variables, we use the returns as they are. We use these returns to compute our monthly risk measures.

Table 1 summarizes the descriptive statistics and correlation matrix for the key variables used in our analysis. These include monthly excess log returns (r), realized variances for downside tail (RV^{T-}), downside core (RV^{C-}), upside core (RV^{C+}), upside tail (RV^{T+}), tail asymmetry ($RV^{\Delta T}$), and core asymmetry ($RV^{\Delta C}$), as well as traditional predictors such as the variance risk premium (VRP) and the price-dividend ratio ($\log(P/D)$).

The upper panel of Table 1 reports descriptive statistics, including the mean, variance, skewness, kurtosis, and first-order autocorrelation for each variable. Monthly excess log returns (r) display a mean (0.01) and variance (0.04), with a negatively skewed distribution (-0.79) and significant excess kurtosis (4.74), reflecting occasional extreme events. Realized variances differ between core and tail components. Tail variances (RV^{T-} and RV^{T+}) exhibit higher means (8.19 and 8.36, respectively) and variances (12.15 and 14.24), capturing the magnitude and frequency of extreme return fluctuations. In contrast, core variances (RV^{C-} and RV^{C+}) have lower means (both 0.25) and variances (0.52 and 0.48). All realized variances are positively skewed and have high kurtosis, with tail variances showing the most extreme deviations, consistent with their focus on rare, large market movements.

The log price-dividend ratio ($\log(P/D)$) has a mean of 0.68 and high autocorrelation (0.98), typical for a slowly varying valuation predictor. The VRP, with a mean of 19.82 and substantial variance (7.85), captures market-implied risk expectations but shows lower skewness and kurtosis than realized variances. These differences suggest a complementary nature of the proposed risk measures and traditional predictors.

The lower panel of Table 1 depicts the correlation matrix. Realized variances across the four components are highly correlated (correlations ranging from 0.91 to 0.97), indicating a shared underlying risk structure despite their segmentation. Asymmetries are much less correlated with other predictors with about 0.5 when related to the four individual components. Excess returns (r) show weak negative correlations with all realized variances, with slightly stronger

relationships for tail variances (r with $RV^{T-} = -0.13$ and $RV^{T+} = -0.14$) compared to core variances (r with $RV^{C-} = -0.09$ and $RV^{C+} = -0.12$). Traditional predictors, including the PD ratio and VRP, are weakly correlated with both excess returns and realized variances.

Figure 1 shows the dynamics of asymmetries by separately analyzing core and tail components of the return distribution. Panel A focuses on core asymmetry, while Panel B highlights tail asymmetry.

Panel A depicts the core asymmetry, which measures the difference between typical upside and downside variances within the core of the distribution. The core excludes extreme values, focusing instead on the more frequent and moderate returns. Core asymmetry fluctuates significantly over time, aligning closely with major market events. For example, we observe sharp spikes in core asymmetry during the late 1990s and early 2000s, coinciding with the bursting of the tech bubble, as well as during the Great Financial Crisis (2007–2009). These periods of heightened asymmetry reflect increased imbalances between typical positive and negative returns, indicative changing investor sentiment. During calmer periods, such as the mid-2000s and the post-crisis recovery, core asymmetry stabilizes, reflecting more balanced and less volatile market conditions.

Panel B highlights tail asymmetry, which measures the difference between extreme positive and negative returns in the tails of the distribution. This predictor captures rare, high-impact events that dominate during periods of significant market stress or exuberance. Tail asymmetry spikes similarly as the core asymmetry.

We now turn to the empirical evaluation of our proposed predictors. By examining the segmented components of high-frequency return distributions, we analyze how they capture both typical and extreme market dynamics. The results presented below aim to show how core and tail asymmetries contribute to the short-term predictability of equity returns.

4 Results

This section outlines the empirical results, divided into in-sample and out-of-sample predictability analyses. The in-sample analysis highlights the role of core and tail asymmetries in explaining one-month-ahead market returns, while the out-of-sample analysis evaluates the model's

robustness and forecasting performance across different thresholds and timeframes.

4.1 In-sample predictability

This section examines the in-sample predictability of our predictor for one-month-ahead market returns, focusing on the period from January 1990 to December 2019. Our analysis highlights the usefulness of tail and core components, both individually and in combination, in forecasting excess market returns. We use our vector autoregression (VAR) framework, systematically exploring how adjusted R-squared varies across different thresholds of our risk measure, denoted by α, within the range $[0.02, 0.48]$.

Figure 2 illustrates the adjusted R-squared values for three distinct regression setups: (i) using only the tail components, (ii) using only the core components, and (iii) using the combined tail and core components. The y-axis represents the adjusted R-squared, while the x-axis corresponds to the threshold α. This framework aims to identify the optimal α that maximizes the adjusted R-squared, thereby enhancing the predictability.

The combined analysis of tail and core components shows that the optimal α is approximately 28%, yielding an adjusted R-squared of 7%. This result is depicted as the black line in Figure 2. When focusing solely on the core components, the adjusted R-squared peaks at 6% with an optimal α of around 30%, represented by the red line. Interestingly, the tail components show a different pattern. For the tail alone, the adjusted R-squared peaks at a much lower α, around 1.5%, and subsequently declines as α increases. This behavior, shown as the blue line, suggests that the tail's predictive power diminishes beyond its narrow range of optimal thresholds. Notably, the tail outperforms the combined model at its peak, but this advantage dissipates as α increases.

These findings provide insights into the interplay between tail and core components in predicting market returns. The distinct behavior of the tail—with its optimal threshold usually restricted to extreme percentiles of the distribution—has likely contributed to an underappreciation of the core's role in return predictability. When additional core components are introduced, their marginal contribution to predictability becomes significant only for α values exceeding 6%, which lie well beyond the conventional thresholds of 1% or 5% typically associated with tail analysis.

Our results challenge traditional notions that focus on extreme events represented by the tails. Instead, we argue for a broader definition of the tail, extending it to encompass thresholds as high as 25% to 30%, thereby incorporating the core. This perspective shows that the core, rather than the tail, plays a more crucial role in driving short-term return predictability. Consequently, truncating the high-frequency return distribution at these expanded thresholds could improve forecasts. By reevaluating the balance between tail and core components, we suggest a framework for understanding the dynamics of short-term market return predictability.

Table 2, Panel A, presents the results of the vector autoregression analysis for monthly excess market returns, focusing on the explanatory power of four key components of the return distribution: downside tail variance, downside core variance, upside core variance, and upside tail variance. The table shows coefficient estimates, t-statistics, and the overall adjusted R-squared. Downside core variance and upside core variance are particularly significant, with a positive and statistically robust coefficient of the downside core variance, while the upside core variance has a negative and similarly significant coefficient. These findings suggest that typical downward movements positively predicts future returns, indicating compensation for perceived downside risks. Conversely, typical upward movements show a negative relationship, implying a premium paid by investors for upside potential. The tail components, show weaker statistical significance, suggesting that extreme values contribute less to short-term return predictability. We will return to this below. When all four components are combined, the model achieves an adjusted R-squared of 7.25

In Table 2, Panel B, we focus on the asymmetry within the core and tail components of variance. To capture this, we reformulate our equation to measure the difference between upside and downside movements within core variance and tail variance. Although the individual variance components exhibit high correlation (as shown in Table 1), their asymmetries are less correlated, enabling us to isolate their predictive significance.

Our analysis shows that core asymmetry is particularly significant, with a negative coefficient, while tail asymmetry, though also negative, is less significant. These results indicate that larger asymmetries, especially within the core, are predictive of future returns. This suggests that investors are compensated for perceived downside risks associated with asymmetric variance.

Table 3 gives the optimal thresholds (α) that define the boundaries between core and tail components, along with their associated predictive power (R^2). The tail components alone, exhibiting by smaller thresholds ($\alpha_T \approx 1.45\%$ for the downside variance and $\alpha_T \approx 2.87\%$ for the upside variance), have limited predictive power, with R^2 values below 4%. This indicates that the predictive value of tails is restricted to extreme market movements, consistent with prior literature on downside risk.

In contrast, core components, with an optimal thresholds of $\alpha_C \approx 30\%$, shows stronger predictive power with both components. The combined core asymmetries achieve an R^2 of 6.44%. Notably, the model incorporating both core and tail components yields the highest R^2 of 7.25%, with an optimal threshold of $\alpha_C \approx 28.77\%$. However, the model performs well with only the tail asymmetry and the core asymmetry with an R^2 of 6.29%. This shows that it is the asymmetries that carry most predictive power.

Figure 3 visualizes the within-sample predictive power of tail and core asymmetries across varying thresholds and models. The red line, representing core asymmetry, shows a steady increase in predictive power as α rises, peaking around 30%. In contrast, the blue line, corresponding to tail asymmetry, shows a peak at a lower α (approximately 20%).

The black line in Figure 3, shows the full model that includes all components (RVT^-, RVC^-, RVC^+, RVT^+), yields the highest predictive power at an optimal threshold of $\alpha^* = 28.77\%$. The gray line, which reflects the combined effect of tail and core asymmetries, closely follows the full model at slightly lower levels, highlighting the marginal value added by considering all components.

Figure 3 also shows that while tail asymmetries contribute to the predictive power of the model, it is the core asymmetries that provide the most. This challenges conventional approaches that focus predominantly on tail risk by showing that incorporating core asymmetry significantly enhances predictability. The results suggest that investors should account for more common market variations in addition to extreme events when assessing short-term return dynamics, in particular the asymmetries within the core variance.

4.2 Out-of-sample predictability

In line with Campbell and Thompson (2008), we calculate the out-of-sample (*OOS*) R^2 to evaluate the model's predictive out-of-sample performance. The *OOS* R^2 is defined as:

$$R^2_{oos} = 1 - \frac{\sum_{t=1}^{T'}(y_t - \hat{y}_t)^2}{\sum_{t=1}^{T'}(y_t - \bar{y})^2},$$

where y_t represents the actual values of the dependent variable in the test set, $\hat{y}_t$ denotes the predicted values, and $\bar{y}$ is the mean of the actual values. The number of observations in the test is T'. Both rolling and expanding window analyses are employed with an initial window length of 15 years ($T' = 180$ months), ensuring the inclusion of the Global Financial Crisis (2007–2008) in the *OOS* period.

Figure 4 depicts the *OOS* R^2 for rolling and expanding windows across various thresholds (α) defining the separation between core and tail components. Both methods show a peak around the in-sample optimal $\alpha^* \approx 28\%$, with the *OOS* R^2 exceeding 4% for monthly vector autoregressions (VAR). The expanding window consistently outperforms the rolling window across most thresholds, achieving higher R^2 values. This result suggests that the expanding window effectively captures persistent risk-return relationships by using an extended data horizon. Predictive power diminishes for $\alpha < 10\%$, in line with the marginal contribution of tail components to *OOS* predictability relative to the core.

We implement additional constraints to check robustness, as outlined by Campbell and Thompson (2008). Specifically, we impose a "Positive Slope" condition, ensuring that the predictor's coefficient aligns with theoretical expectations. Additionally, negative return forecasts are replaced with zero under the "Positive Forecast" restriction. These conditions prevent unrealistic predictions and improve model reliability. Including the Global Financial Crisis from the *OOS* period slightly reduces the R^2 values, though the pattern remains consistent, with predictive power peaking near the optimal α^*.

Figure 5 provides a graphical comparison of observed, forecasted, and mean monthly excess returns. The observed returns (blue line) show significant volatility, particularly during periods such as the dot-com bubble and the Global Financial Crisis. The forecasted returns (red line) capture the time-varying dynamics of market returns reasonably well on average. The

mean return (yellow line), computed using a rolling 15-year window, remains stable, offering a baseline for comparison. The forecasted returns often outperform the mean, particularly during periods of heightened market turbulence.

Table 5 benchmarks the *OOS* R^2 performance of the proposed risk measures against established predictors, including the variance risk premium (VRP) and the price-dividend (PD) ratio. Our model achieves the highest *OOS* R^2, with 4.92% for the fixed window and 4.47% for the expanding window. In contrast, the PD ratio exhibits limited predictive power, yielding a positive R^2 of 0.37% only in the expanding window, while the VRP shows negative R^2 values across both frameworks.

Our *OOS* analysis suggests the robustness and predictive strength of the proposed risk measures. By decomposing return distributions into core and tail components and optimizing the threshold α, the model offers significant improvements in short-term market return forecasting. These findings challenge traditional approaches that rely mostly on long-term predictors, such as the PD ratio, VRP, or tail risk measures and suggest to include the asymmetries in the core and tail of the distribution when thinking about the equity risk-return tradeoff.

5 Comparison with established predictors

5.1 Short-term performance

We evaluate the performance of our proposed risk measures by comparing them to established predictors, such as the variance risk premium (VRP) and the price-dividend (PD) ratio. These traditional predictors are widely recognized for their forecasting effectiveness over medium- and long-term horizons. Table 4 provides the coefficient estimates and adjusted R^2 values for eight regression models, highlighting the relative contribution of each predictor to explaining one-month-ahead excess returns.

The results suggest the superior performance of our proposed risk measures in capturing short-term market dynamics. When used independently, the VRP and PD ratio achieve adjusted R^2 values of 0.69% and 1.54%, respectively. Combining the two predictors yields an adjusted R^2 of 1.37%. In comparison, our risk measures yield an adjusted R^2 of 7.25%, significantly

outperforming traditional predictors at the one-month horizon. Adding the PD ratio to our risk measures marginally increases the adjusted R^2 to 7.39%, while incorporating the VRP slightly reduces it to 7.06%. Combining our predictors with both the PD ratio and VRP yields an adjusted R^2 that is slightly lower than that of our risk measures alone.

We also provide additional robustness checks in the Appendix that show the dominant contribution to predictability stems from core asymmetry. Models focusing solely on core asymmetries consistently outperform those that focus on tail asymmetries, suggesting that typical return variations within the core carry more predictive information than extreme events in the tails. Traditional predictors, while valuable for forecasting over longer horizons, provide limited incremental value in short-term settings.

Extending the analysis to longer horizons suggests an evolving hierarchy of predictors. At the one-month horizon, our risk measures significantly outperform both the VRP and PD ratio. However, as the horizon extends to 3–9 months, the VRP surpasses our predictors, reflecting its strength in medium-term forecasting. Beyond nine months, the PD ratio emerges as the most effective predictor.

To ensure robustness, we conducted several additional analyses (see Appendix). First, we tested alternative thresholds, such as median and zero, instead of the optimal α^*, and found qualitatively similar results. Second, we explored the impact of using alternative grids for five-minute returns and found no significant changes in the findings. Third, we implemented a simplified variance measure, inspired by Barndorff-Nielsen, Kinnebrock, and Shephard (2008), which involved splitting returns at zero and squaring them rather than calculating variance. This approach yielded an adjusted R^2 of approximately 3%, significantly lower than that of our proposed framework. Finally, regressions using only core or tail asymmetries confirmed that the bulk of predictability stems from core asymmetry, with tail asymmetry contributing little to the explanatory power.

5.2 Multiple-horizon performance

In this section we study the predictive power of different risk measures for various forecast horizons. Using both local projection and overlapping methodologies, we show that our measure performs best in the short term, and, in line with the literature, the variance risk

premium (VRP) dominates in the medium term, and the log price-dividend ratio (log(P/D)) becomes the strongest predictor over longer horizons.

We first outline the local projection and overlapping methodologies. Let v be a vector with the same dimension as the state vector z_t, and let V be a matrix with the number of columns equal to the dimension of z_t and the number of rows less than or equal to the dimension of z_t. We define the variables $y_t = v^\top z_t$ and $x_t = V z_t$ and estimate the predictive linear regression:

$$y_{t+1:t+h} = a_{mh} + b_{mh}^\top x_{t-m+1:t} + u_{t+h}^{(m)} \tag{7}$$

where

$$y_{t+1:t+h} = \frac{1}{h}\sum_{j=1}^{h} y_{t+j} \quad \text{and} \quad x_{t-m+1:t} = \frac{1}{m}\sum_{i=1}^{h} x_{t-i+1}. \tag{8}$$

The approach to estimate a_{mh} and b_{mh} in equation (8) is referred to as *overlapping* methodology. The gist consists of aggregating the predictors backwards and the forecasted variables forwards.

The estimation of a_{mh} and b_{mh} in equation (8) follows the "overlapping" methodology, which aggregates predictors backward and forecasted variables forward. However, such aggregation increases the persistence of variables, potentially inflating statistical significance and R^2, as noted by Stambaugh (1999) and Valkanov (2003). To address this issue, we employ local projections, which use the stationarity of the VAR to estimate a_{mh} and b_{mh} without artificially boosting predictability. Section E in the appendix provides the derivation of these coefficients in terms of VAR parameters:

$$a_{mh} = \left(v - V^\top b_{mh}\right)^\top (\mathbf{I} - \Theta)^{-1}\theta_0 \quad \text{and} \quad b_{mh} = \left(\frac{1}{m}\Sigma_{mm}^{xx}\right)^{-1}\left(\frac{1}{h}\sigma_{mh}^{xy}\right) \tag{9}$$

where $\mathbf{I}$ denotes the appropriate identity matrix,

$$\begin{aligned} \sigma_{mh}^{xy} &= V\Sigma^{zz}\left((\mathbf{I} - \Theta)^{-1}(\mathbf{I} - \Theta^m)\right)^\top \left((\mathbf{I} - \Theta)^{-1}\left(\Theta - \Theta^{h+1}\right)\right)^\top v \\ \Sigma_{mm}^{xx} &= V\Sigma_{mm}^{zz}V^\top, \end{aligned} \tag{10}$$

and where

$$\begin{aligned}\Sigma_{mm}^{zz} = m\Sigma^{zz} + \Sigma^{zz}\left(\left(\left(\mathbf{I}-\Theta\right)^2\right)^{-1}\left(\Theta^{m+1} - m\Theta^2 + \left(m-1\right)\Theta\right)\right)^{\top} \\ + \left(\left(\left(\mathbf{I}-\Theta\right)^2\right)^{-1}\left(\Theta^{m+1} - m\Theta^2 + \left(m-1\right)\Theta\right)\right)\Sigma^{zz},\end{aligned} \tag{11}$$

with $\text{vec}\left(\Sigma^{zz}\right) = \left(\mathbf{I} - \Theta \otimes \Theta\right)^{-1}\text{vec}\left(\Sigma\right)$.

Likewise, the R^2 of the predictive regression (7) is given by:

$$R_{mh}^2 = \frac{\sigma_{mh}^{xy\top}\left(\Sigma_{mm}^{xx}\right)^{-1}\sigma_{mh}^{xy}}{\left(\sigma_h^y\right)^2} \tag{12}$$

where σ_{mh}^{xy} and Σ_{mm}^{xx} are given by equation (10), and where

$$\left(\sigma_h^y\right)^2 = v^{\top}\Sigma_{hh}^{zz}v. \tag{13}$$

In Figure 6 we present the results of the multiple-horizon forecasting analysis, comparing our measure to alternative predictors across forecast horizons. Panel A gives results from the local projection methodology, while Panel B displays the overlapping methodology. Both approaches show how predictive power shifts as the forecast horizon extends from one to twenty-four months.

At the shortest horizon (one month), our proposed measure consistently outperforms traditional predictors such as the VRP and log(P/D). As such the risk measures introduced in this paper capture complementary information to the established predictors. Under the local projection approach, our predictor remains dominant for the first three months. Thereafter, the VRP becomes the strongest predictor, in line with prior research indicating that variance risk is particularly effective at forecasting medium-term equity returns due to its sensitivity to investor sentiment and risk appetite. Beyond five months, log(P/D) emerges as the most effective long-term predictor.

The overlapping methodology produces a similar pattern. Our measure delivers the highest predictive power at short horizons, while the VRP reaches its peak effectiveness around four months. As the forecast horizon extends beyond six months, log(P/D) overtakes the VRP.

Our results suggest that selecting the best measure depends on the forecast horizon. Our measure gives improved short-term return forecasting, particularly within a one-to-three-month window. Our finding suggests that investors should adjust their predictive models based on the investment horizon.

6 Conclusion

We introduce a novel approach to short-term equity return predictability by segmenting high-frequency return distributions into core and tail components. Our methodology identifies core asymmetry — representing differences in typical upside and downside variances — as the principal driver of predictive power, significantly outperforming traditional predictors such as the variance risk premium and valuation ratios. Our empirical results show an in-sample adjusted R^2 of 7.25% and an out-of-sample R^2 exceeding 3% for one-month-ahead market returns, with optimal thresholds for segmentation centered around 28.77

We challenge conventional views that prioritize extreme tail events by highlighting the additional informational value embedded within the core of the return distribution. The predictive power of core asymmetry highlights the importance of typical market fluctuations, which have been historically overlooked in favor of rare, high-impact events. These findings align with behavioral theories suggesting that investor sentiment and risk perception vary across segments of the return distribution.

We confirm the robustness of our results through sensitivity analyses, including alternative segmentation thresholds and cross-comparison with established predictors. Core asymmetry consistently exhibits significant explanatory power. Furthermore, combining core and tail measures with traditional predictors yields marginal gains, suggesting their complementary nature.

By taking a different approach through which we examine the risk-return tradeoff, we contribute to the ongoing dialogue on equity premium prediction. Our findings show the potential of using high-frequency data to uncover patterns that extend beyond traditional metrics.

References

Ang, Andrew, Robert J Hodrick, Yuhang Xing, and Xiaoyan Zhang. 2006. "The cross-section of volatility and expected returns." *The Journal of Finance* 61 (1):259–299.

Bakshi, Gurdip, Georgios Panayotov, and Georgios Skoulakis. 2011. "Improving the Predictability of Real Economic Activity and Asset Returns with Forward Variances." *Journal of Financial Economics* 100:475–495.

Bansal, Ravi and Dana Kiku. 2011. "Cointegration and Long-Run Asset Allocation." *Journal of Business & Economic Statistics* 29 (1):161–173.

Barndorff-Nielsen, Ole E, Silja Kinnebrock, and Neil Shephard. 2008. "Measuring downside risk-realised semivariance." *Working paper* .

Bekaert, Geert and Eric Engstrom. 2017. "Asset Return Dynamics under Habits and Bad Environment–Good Environment Fundamentals." *Journal of Political Economy* 125 (3):713–760.

Bekaert, Geert, Eric Engstrom, and Andrey Ermolov. 2015. "Bad Environments, Good Environments: A Non-Gaussian Asymmetric Volatility Model." *Journal of Econometrics* 186 (1):258–275.

Bekaert, Geert and Marie Hoerova. 2014. "The VIX, the Variance Premium and Stock Market Volatility." *Journal of Econometrics* 183:181–192.

Belo, Frederico and Xiaoliang Yu. 2013. "Household and Government Investment and the Stock Market." *Journal of Monetary Economics* 60 (3):361–378.

Bollerslev, Tim, George Tauchen, and Hao Zhou. 2009. "Expected Stock Returns and Variance Risk Premia." *Review of Financial Studies* 22:4463–4492.

Breckenfelder, Johannes and Roméo Tédongap. 2012. "Asymmetry Matters: A High-Frequency Risk-Reward Trade-Off." SSRN.

Campbell, John Y and Samuel B Thompson. 2008. "Predicting excess stock returns out of sample: Can anything beat the historical average?" *The Review of Financial Studies* 21 (4):1509–1531.

Colacito, Riccardo, Eric Ghysels, Jing Meng, and Wisit Siwasarit. 2016. "Skewness in Expected Macro Fundamentals and the Predictability of Equity Returns: Evidence and Theory." *Review of Financial Studies* 29:2069–2109.

Cooper, Ilan and Richard Priestley. 2009. "Time-Varying Risk Premiums and the Output Gap." *Review of Financial Studies* 22:2601–2633.

Drechsler, Itamar and Amir Yaron. 2010. "What's vol got to do with it." *The Review of Financial Studies* 24 (1):1–45.

Driesprong, Gerben, Ben Jacobsen, and Benjamin Maat. 2008. "Striking Oil: Another Puzzle?" *Journal of Financial Economics* 89:307–327.

Feunou, Bruno, Jean-Sébastien Fontaine, and Roméo Tédongap. 2013. "Modeling Market Downside Volatility." *Review of Finance* 17 (1):443–481.

Feunou, Bruno, Mohammad R Jahan-Parvar, and Cédric Okou. 2018. "Downside variance risk premium." *Journal of Financial Econometrics* 16 (3):341–383.

Feunou, Bruno, Mohammad R. Jahan-Parvar, and Roméo Tédongap. 2016. "Which Parametric Model for Conditional Skewness?" *The European Journal of Finance* 22 (13):1237–1271.

Goyal, Amit and Ivo Welch. 2008. "A Comprehensive Look at The Empirical Performance of Equity Premium Prediction." *Review of Financial Studies* 21 (4):1455–1508.

Goyal, Amit, Ivo Welch, and Athanasse Zafirov. 2024. "A Comprehensive 2022 Look at the Empirical Performance of Equity Premium Prediction." *Review of Financial Studies* 37 (11):3490–3557.

Hirshleifer, David, Kewei Hou, and Siew Hong Teoh. 2009. "Accruals, Cash Flows, and Aggregate Stock Returns." *Journal of Financial Economics* 91:389–406.

Huang, David, Fu Jiang, Jun Tu, and Guofu Zhou. 2015. "Investor Sentiment Aligned: A Powerful Predictor of Stock Returns." *Review of Financial Studies* 28 (3):791–837.

Huang, Tao and Junye Li. 2019. "Option-Implied Variance Asymmetry and the Cross-Section of Stock Returns." *Journal of Banking & Finance* 101:21–36.

Kelly, Bryan and Hao Jiang. 2014. "Tail Risk and Asset Prices." *Review of Financial Studies* 27:2841–2871.

Kilic, Mete and Ivan Shaliastovich. 2019. "Good and Bad Variance Premia and Expected Returns." *Management Science* 65 (6):2522–2544.

Maio, Paulo. 2016. "Cross-Sectional Return Dispersion and the Equity Premium." *Journal of Financial Markets* 31:1–20.

Merton, Robert C. 1973. "An intertemporal capital asset pricing model." *Econometrica: Journal of the Econometric Society* :867–887.

Møller, Stig and Jesper Rangvid. 2015. "End-of-the-Year Economic Growth and Time-Varying Expected Returns." *Journal of Financial Economics* 115:136–154.

Neely, Christopher J., David E. Rapach, Jun Tu, and Guofu Zhou. 2014. "Forecasting the Equity Risk Premium: The Role of Technical Indicators." *Management Science* 60:1772–1791.

Piazzesi, Monika, Martin Schneider, and Selale Tuzel. 2007. "Housing, Consumption, and Asset Pricing." *Journal of Financial Economics* 83:531–569.

Pollet, Joshua M. and M. Wilson. 2010. "Average Correlation and Stock Market Returns." *Journal of Financial Economics* 96:364–380.

Rapach, David E., Matthew C. Ringgenberg, and Guofu Zhou. 2016. "Short Interest and Aggregate Stock Returns." *Journal of Financial Economics* 121 (1):46–65.

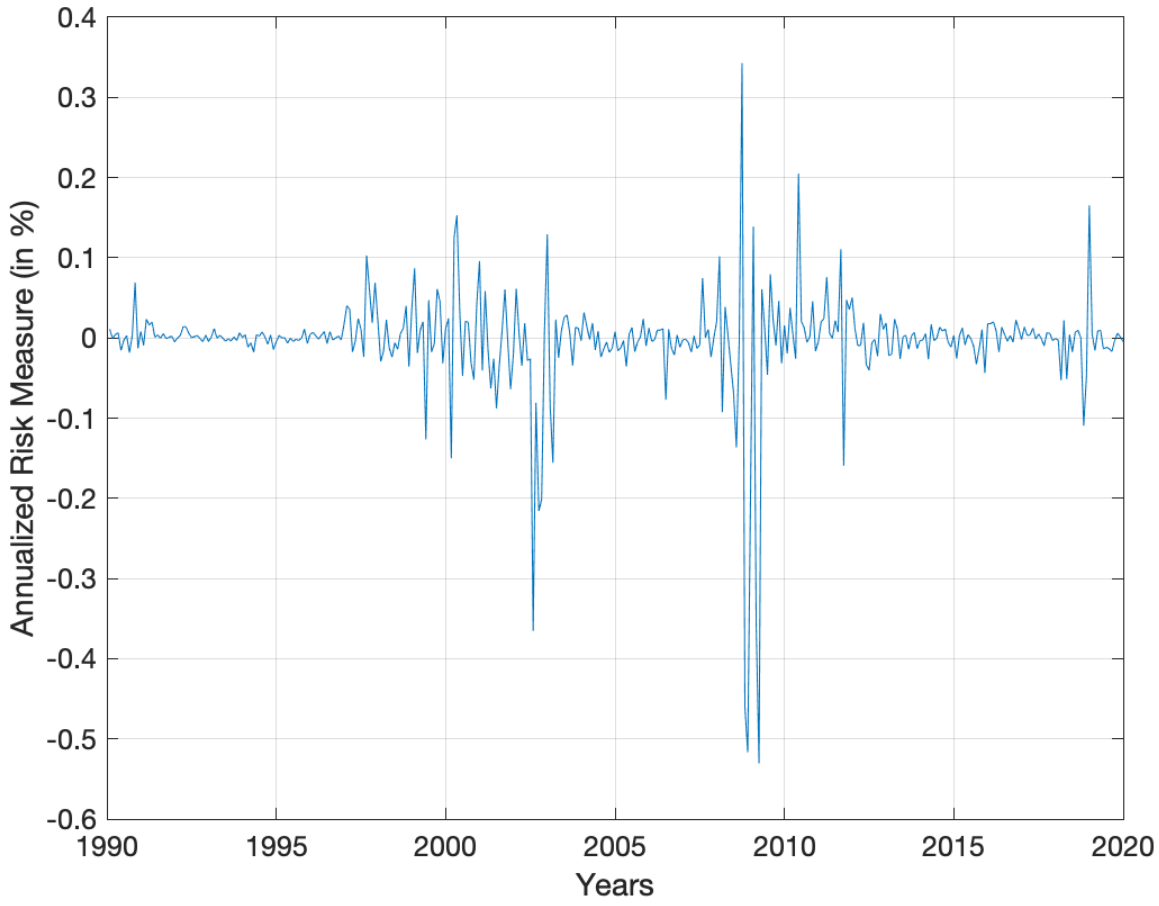
Stambaugh, Robert F. 1999. "Predictive regressions." *Journal of Financial Economics* 54 (3):375–421.

Valkanov, Rossen. 2003. "Long-horizon regressions: theoretical results and applications." *Journal of Financial Economics* 68 (2):201–232.

Figure 1: Asymmetry measures

The figure shows the monthly core asymmetry (Panel A), defined as $RV^{\Delta C}_{t-1:t}(\alpha) = RV^{C+}_{t-1:t}(\alpha) - RV^{C-}_{t-1:t}(\alpha)$, where $RV^{C+}_{t-1:t}(\alpha)$ and $RV^{C-}_{t-1:t}(\alpha)$ represent the upside and downside core volatilities, respectively. Panel B depicts the monthly tail asymmetry, defined as $RV^{\Delta T}_{t-1:t}(\alpha) = RV^{T+}_{t-1:t}(\alpha) - RV^{T-}_{t-1:t}(\alpha)$, where $RV^{T+}_{t-1:t}(\alpha)$ and $RV^{T-}_{t-1:t}(\alpha)$ capture the upside and downside tail volatilities, respectively. Here, α denotes the threshold that maximizes the in-sample predictability R-squared.The sample period spans from January 1990 to December 2019.

Panel A: Core asymmetry



Panel B: Tail asymmetry

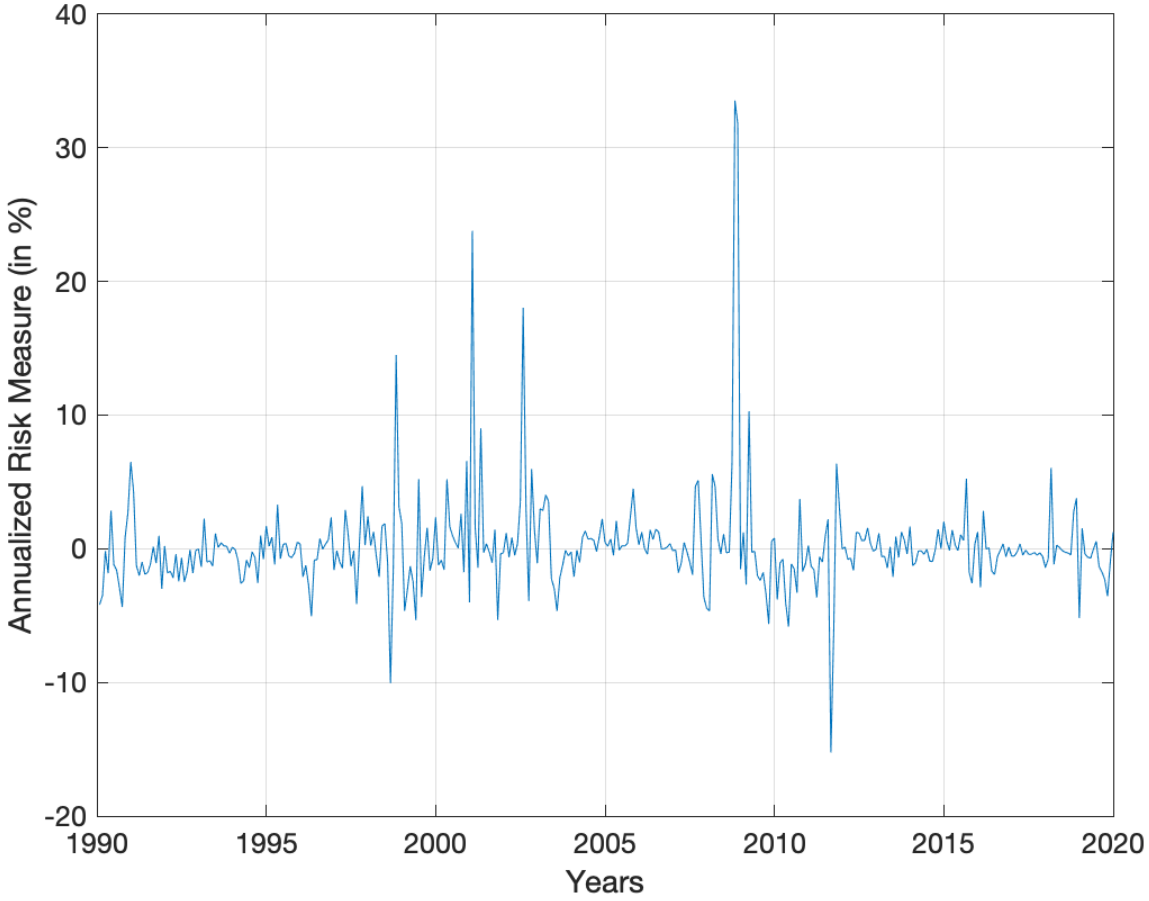


Figure 2: Within-sample predictability across thresholds: core versus tail

This figure represents the adjusted R-squared values for monthly VAR regressions of the excess market return. The blue (red) line captures the the explanatory power of the tails (cores) for various values alpha. Alpha defines the threshold between tail can core in the return distribution of minutes log returns. The sample period ranges from January 1990 to December 2019.

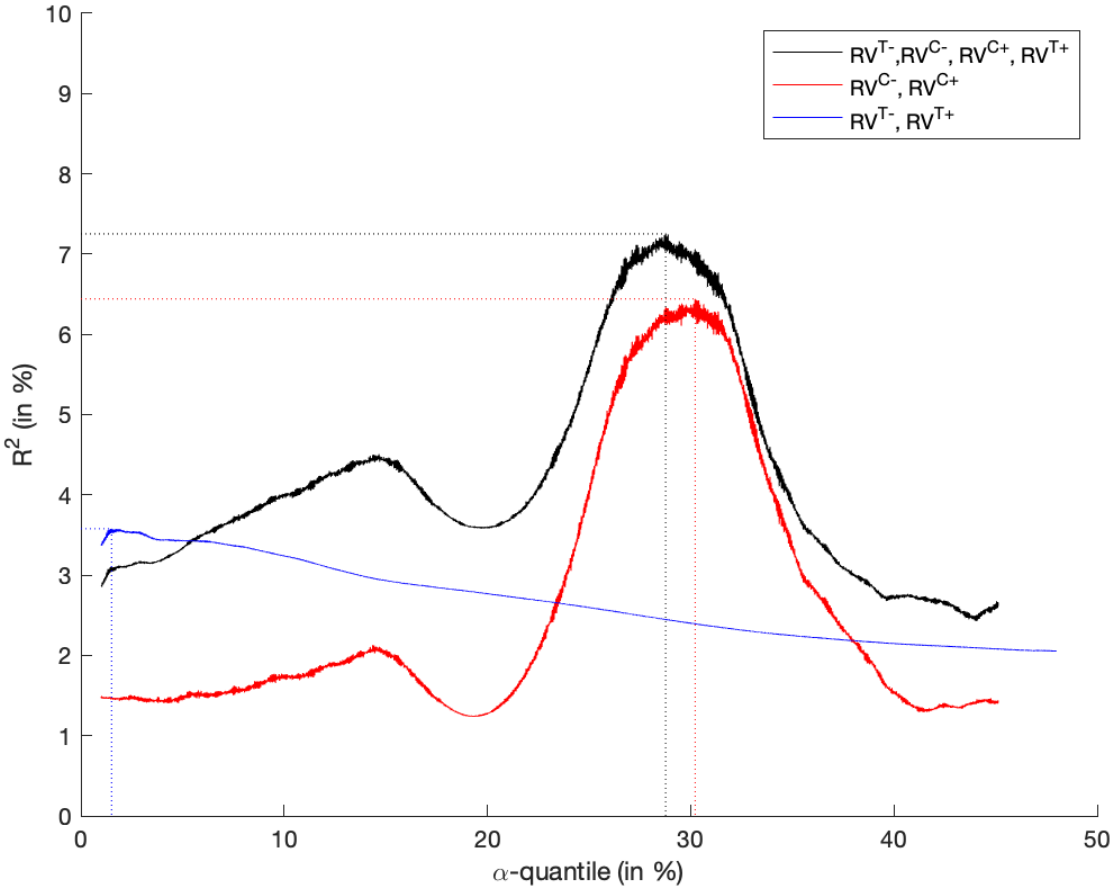


Figure 3: Within-sample predictability across thresholds: asymmetries

This figure represents the adjusted R-squared values for monthly VAR regressions of the excess market return. The light blue (light red) line captures the the explanatory power of the tail (core) asymmetry for various values alpha. The black line captures the same explanatory power as $[RV^{T-}, RV^{C-}, RV^{C+}, RV^{T+}]$. Finally, the light gray captures the explanatory power of the tail and core asymmetries combined. The sample period ranges from January 1990 to December 2019.

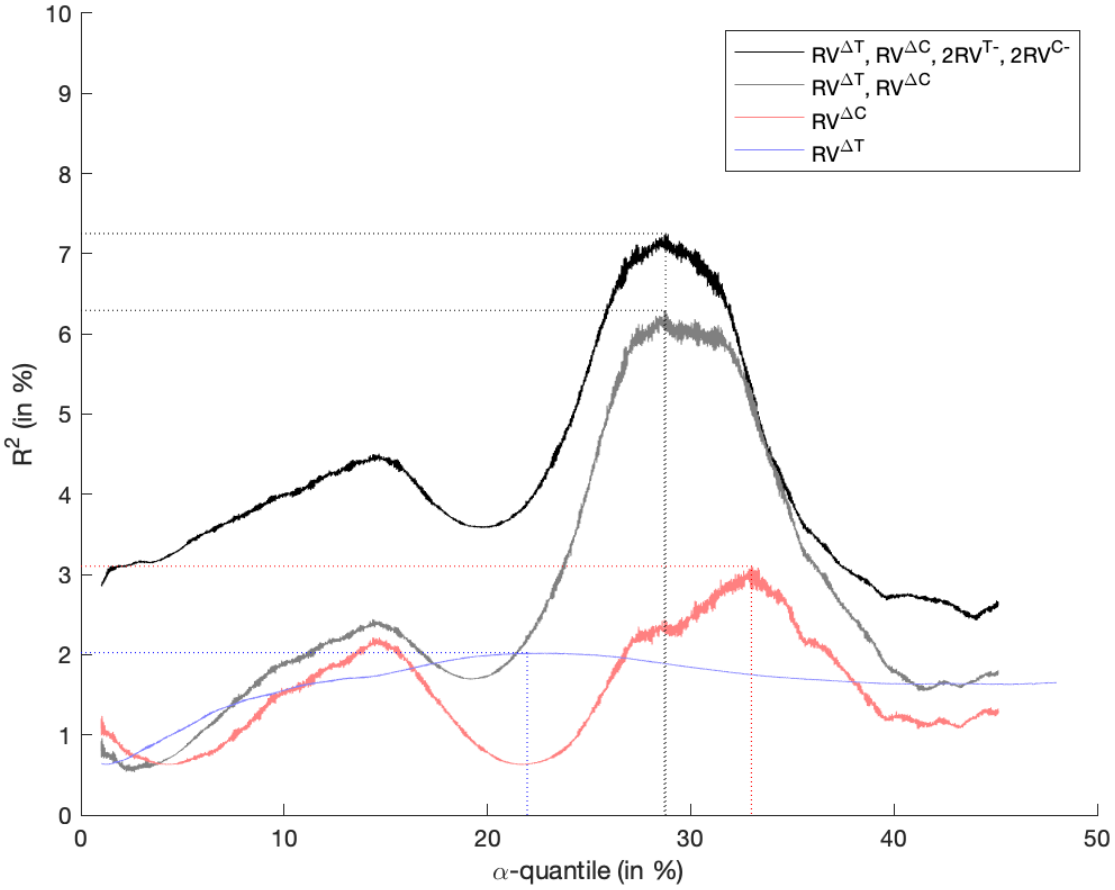


Figure 4: Out-of-sample predictability across thresholds

This figure represents the out-of-sample R-squared values for monthly VAR regressions of the excess market return. The window of analysis spans 15 years and the blue (red) line captures the the explanatory power of the expanding (fixed) window analysis for various values alpha. Alpha defines the threshold between tail can core in the return distribution of minutes log returns. The sample period ranges from January 1990 to December 2019.

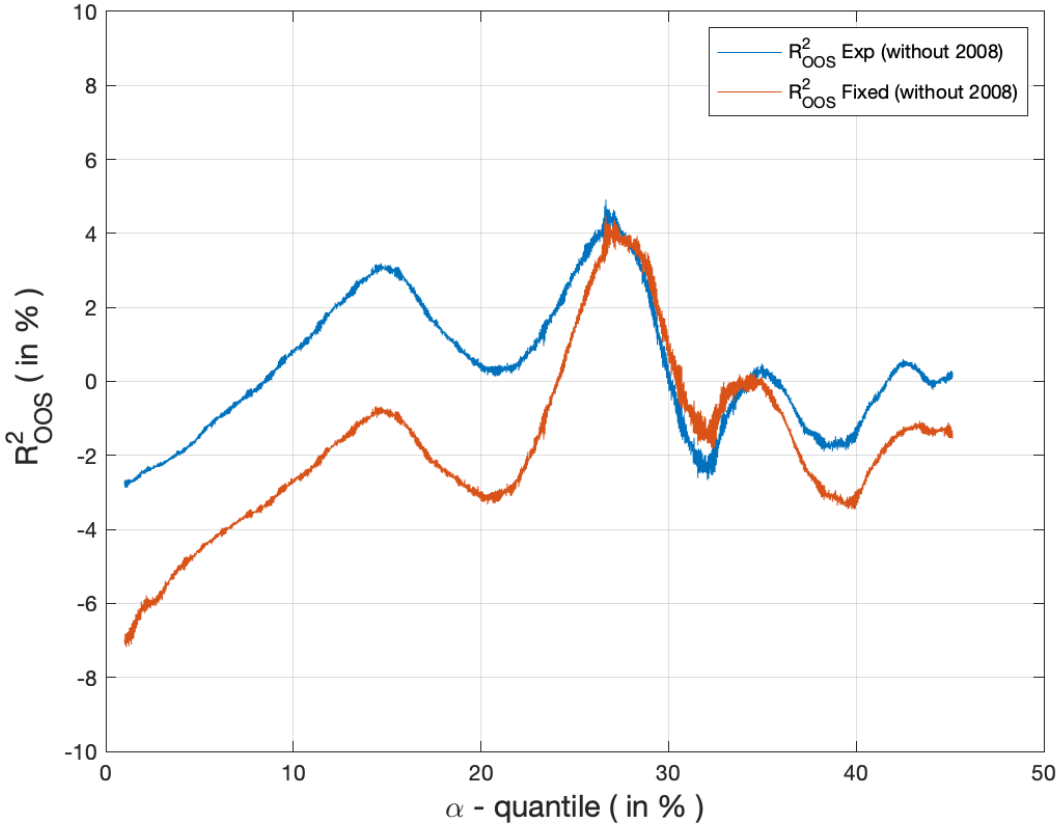


Figure 5: Monthly excess log return: actual, predicted and mean

This figure compares the observed monthly excess to forecasted ones. That is the blue line represents the observed monthly excess return, the red line captures our forecast and the yellow line is the average monthly excess return computed with a rolling window of 15 years.

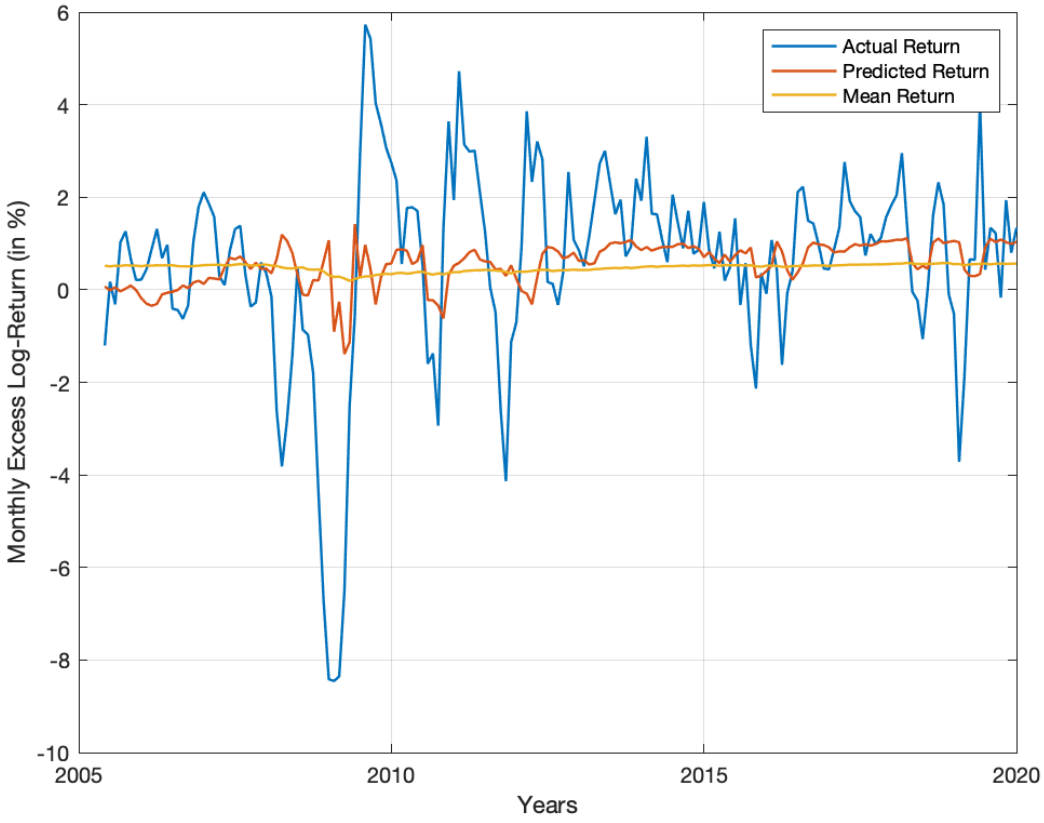
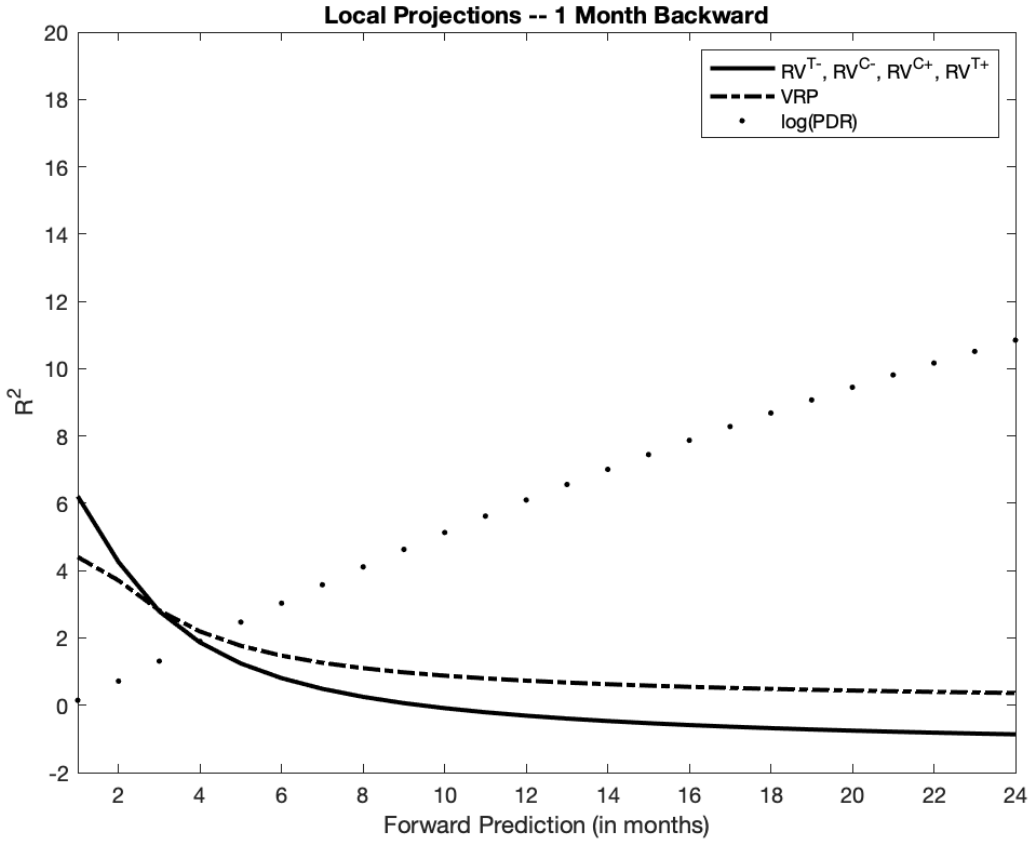


Figure 6: Multiple-Horizon Forecasting

This figure depicts the evolution of predictive power for different risk measures over forecast horizons ranging from one to twenty-four months. Panel A presents results using the local projection methodology, while Panel B displays the overlapping methodology. The asymmetry measure exhibits the strongest predictive power in the short term, the variance risk premium (VRP) dominates in the medium term, and the log price-dividend ratio (log(P/D)) becomes the most effective predictor over longer horizons.

Panel A: Local Projections



Panel B: Overlapping Methodology

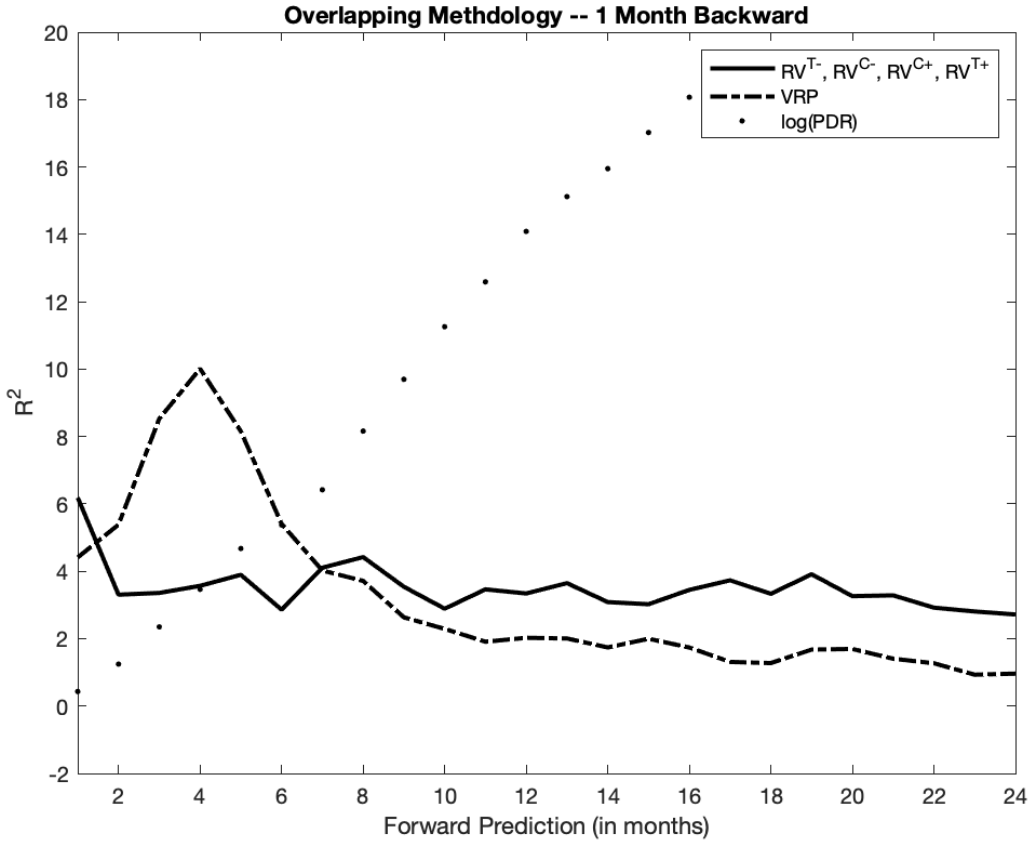


Table 1: Summary statistics

The upper panel of this figure contains the summary statistics for our VAR analysis. r denotes the annualized excess log return, RV^{T-}, RV^{C-}, RV^{C+} and RV^{T+}, denote the downside tail, downside core, upside core and upside tail realized variance. Similarly, $RV^{\Delta T}$ and $RV^{\Delta C}$ denote the tail and core asymmetry, respectively. Further $log(P/D)$ and VRP, denote the log price divide ratio and the variance risk premium, both of which are used in the robustness analysis. The measures of realized variances and the VRP are expressed in squared percentages. The lower panel captures the pairwise correlation between the variables of the analysis. The sample period ranges from January 1990 to December 2019.

	r	RV^{T-}	RV^{C-}	RV^{C+}	RV^{T+}	$RV^{\Delta T}$	$RV^{\Delta C}$	$\log(P/D)$	VRP
					Summary statistics				
$\mathbb{E}[\,\cdot\,]$	0.01	8.19	0.25	0.25	8.36	0.17	-0.00	0.68	19.82
$\mathbb{V}[\,\cdot\,]$	0.04	12.15	0.52	0.48	14.24	3.86	0.07	0.26	7.85
$\mathbb{S}[\,\cdot\,]$	-0.79	7.03	8.03	7.81	7.94	4.25	-3.45	-0.21	1.92
$\mathbb{K}[\,\cdot\,]$	4.74	76.04	88.34	86.57	90.97	35.53	28.46	3.19	9.09
ac_1	0.03	0.63	0.69	0.70	0.65	0.29	0.18	0.98	0.85
N	360	360	360	360	360	360	360	360	360
					Correlation matrix				
r	1.00	-0.13	-0.09	-0.12	-0.14	-0.11	-0.13	0.10	-0.02
RV^{T-}		1.00	0.91	0.93	0.97	0.43	-0.35	-0.01	-0.02
RV^{C-}			1.00	0.99	0.94	0.61	-0.59	-0.04	-0.05
RV^{C+}				1.00	0.95	0.59	-0.49	-0.06	-0.05
RV^{T+}					1.00	0.64	-0.43	-0.01	-0.04
$RV^{T\Delta}$						1.00	-0.48	-0.02	-0.06
$RV^{C\Delta}$							1.00	-0.11	0.05
$\log(P/D)$								1.00	0.09
VRP									1.00

Table 2: Within-sample predictability

This table contains the coefficients estimates and the t-statistics in brackets. The last column contains the adjusted R-squared. Further, α denotes the threshold that distinguishes between tail and core of the return monthly distribution of five-minute returns. Lastly, α^* denotes the threshold which maximizes the adjusted R-squared of the monthly VAR. The sample period ranges from January 1990 to December 2019.

Panel A

$\alpha^* = 28.77\%$

	$r_{t-1:t}$	$RV^{T-}_{t-1:t}$	$RV^{C-}_{t-1:t}$	$RV^{C+}_{t-1:t}$	$RV^{T+}_{t-1:t}$	R^2
$r_{t:t+1}$	-0.01	0.14	16.00*	-16.41*	-0.18†	7.25
	(-0.16)	(1.24)	(3.01)	(-2.59)	(-1.68)	
$RV^{T-}_{t:t+1}$	-0.26‡	0.25	-39.79	45.72	0.18	42.74
	(-2.03)	(1.15)	(-1.16)	(1.44)	(0.49)	
$RV^{C-}_{t:t+1}$	-0.01‡	0.00	-1.55	2.08	0.01	57.48
	(-2.54)	(0.38)	(-1.05)	(1.53)	(0.49)	
$RV^{C+}_{t:t+1}$	-0.01‡	0.00	-1.46	1.90	0.01	55.16
	(-2.01)	(0.49)	(-1.07)	(1.51)	(0.54)	
$RV^{T+}_{t:t+1}$	-0.33‡	0.19	-47.80	54.26	0.35	47.57
	(-2.23)	(0.68)	(-1.15)	(1.42)	(0.77)	

Panel B

$\alpha^* = 28.77\%$

	$r_{t-1:t}$	$2RV^{T-}_{t-1:t}$	$2RV^{C-}_{t-1:t}$	$RV^{\Delta T}_{t-1:t}$	$RV^{\Delta C}_{t-1:t}$	R^2
$r_{t:t+1}$	-0.01	-0.02	-0.21	-0.18†	-16.41*	7.25
	(-0.16)	(-0.54)	(-0.21)	(-1.68)	(-2.59)	
$2RV^{T-}_{t:t+1}$	-0.52‡	0.43	5.92	0.36	91.43	42.74
	(-2.03)	(1.51)	(1.22)	(0.49)	(1.44)	
$2RV^{C-}_{t:t+1}$	-0.03‡	0.01	0.53‡	0.02	4.16	57.48
	(-2.54)	(0.99)	(2.46)	(0.49)	(1.53)	
$RV^{T\Delta}_{t:t+1}$	-0.07	0.06	0.27	0.17	8.54	24.14
	(-1.21)	(1.18)	(0.23)	(1.38)	(1.03)	
$RV^{C\Delta}_{t:t+1}$	0.00*	0.00	-0.04‡	0.00	-0.18	28.84
	(2.94)	(0.59)	(-2.29)	(0.02)	(-1.01)	

Table 3: Return predictability with optimal thresholds

This table provides the optimal α and corresponding adjusted R-squared for various regressions. α denotes the threshold that distinguishes between tail and core of the return monthly distribution of five-minute returns. The optimal alpha captures the threshold which maximizes the R-squared for a given regression. The sample period ranges from January 1990 to December 2019.

Predictors	optimal α	R-squared
RV^{T-}	1.45	3.19
RV^{T+}	2.87	3.79
RV^{T-}, RV^{T+}	$\alpha^T = 1.51$	3.58
RV^{C-}	1.02	1.76
RV^{C+}	48.00	1.96
RV^{C-}, RV^{C+}	$\alpha^C = 30.26$	6.44
$RV^{T-}, RV^{C-}, RV^{C+}, RV^{T+}$	$\alpha^* = 28.77$	7.25
$2RV^{T-}, 2RV^{C-}, RV^{\Delta T}, RV^{\Delta C}$	$\alpha^* = 28.77$	7.25
$RV^{\Delta T}, RV^{\Delta C}$	28.71	6.29

Table 4: Within-sample predictability: alternative models

This table provides robustness check for our analysis and covers various alternative regressions. r denotes the monthly excess log return, RV^{T-}, RV^{C-}, RV^{C+} and RV^{T+}, denote the downside tail, downside core, upside core and upside tail realized variance. Further $log(P/D)$ and VRP, denote the log price divide ratio and the variance risk premium. The sample period ranges from January 1990 to December 2019.

	I	II	III	IV	V	VI	VII
r	-0.01	0.02	0.03	0.02	-0.02	0.01	0.00
	(-0.16)	(0.29)	(0.47)	(0.33)	(-0.26)	(0.17)	(0.05)
$RV^{T,-}$	0.14				0.13	0.10	0.08
	(1.24)				(1.12)	(0.96)	(0.77)
$RV^{C,-}$	16.00^{*}				14.99^{*}	15.18^{*}	$13.80^{\ddagger}$
	(3.01)				(2.61)	(2.76)	(2.23)
$RV^{C,+}$	-16.41^{*}				$-14.97^{\ddagger}$	$-15.57^{\ddagger}$	$-13.61^{\dagger}$
	(-2.59)				(-2.15)	(-2.31)	(-1.76)
$RV^{T,+}$	$-0.18^{\dagger}$				$-0.18^{\dagger}$	-0.13	-0.13
	(-1.68)				(-1.72)	(-1.19)	(-1.21)
$\log(P/D)$		0.02		$0.02^{\dagger}$	0.01		0.01
		(1.53)		(1.88)	(1.00)		(1.26)
VRP			0.04^{*}	0.05^{*}		$0.03^{\ddagger}$	0.04^{*}
			(3.71)	(3.83)		(2.56)	(2.60)
$adj.\ R^2$	7.25	1.54	4.95	6.05	7.39	9.06	9.48

Table 5: Out-of-sample predictability: alternative models

This table reports the OOS R-squared for our measures as well as established predictors (VRP, PDR). The analysis is conducted for both fixed and expanding window. The initial window spans 15 years, leaving another 15 years ofr the OOS analysis.

	Fixed	Expanding
$r,\ RV^{T-},\ RV^{C-},\ RV^{C+},\ RV^{T+}$	3.41	2.87
$r, \log(P/D)$	-0.51	0.37
$r,\ VRP$	-3.72	-1.53

Appendix

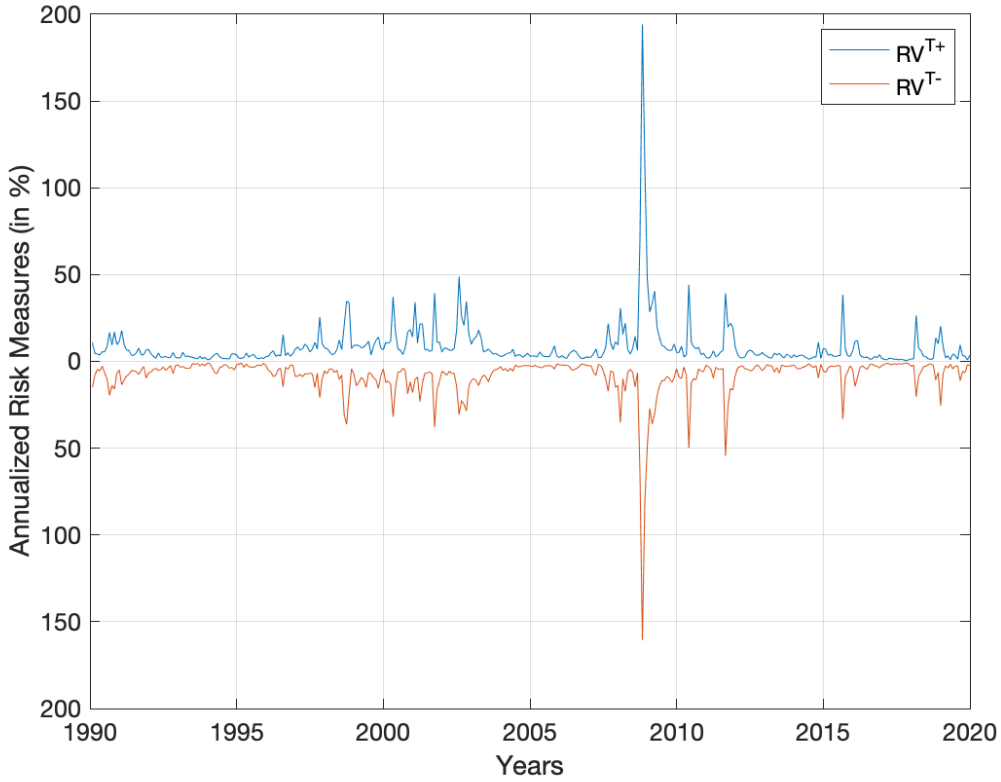
Asymmetries at the core of short-term return predictability

A Tail and core components

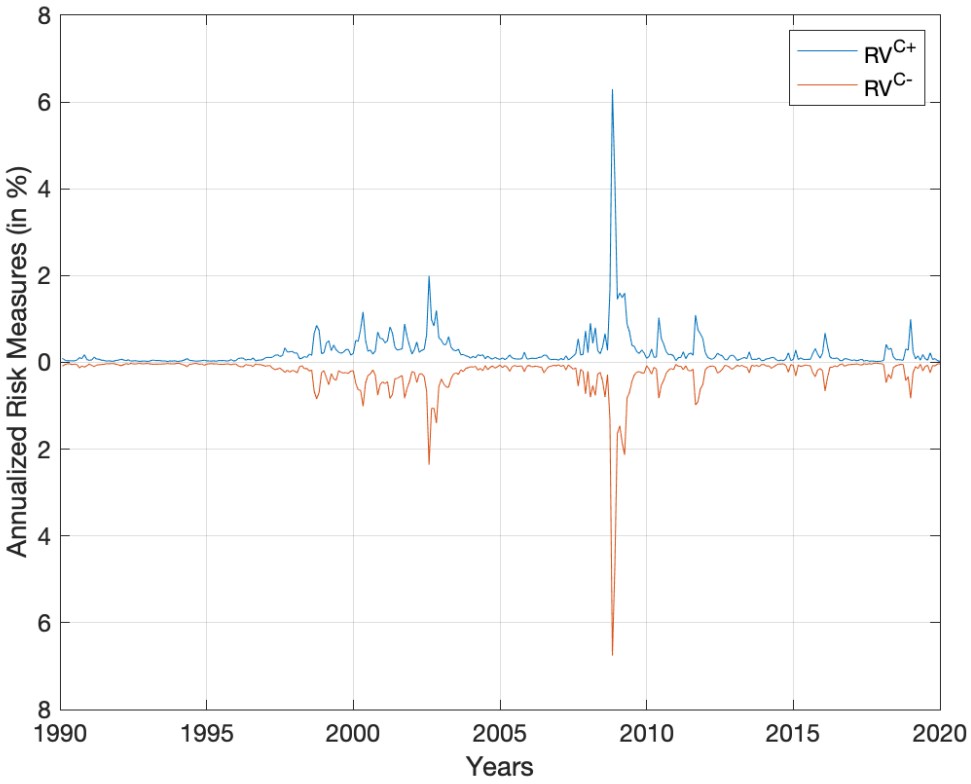
Figure 7: Core and Tail: upside and downside

Panel A and B represent of this figure, respectively, the tail and core components of the five-minute return distribution. The components are computed for optimal alpha, $\alpha^* = 28.77\%$ which maximizes the in-sample predicatibilty. Per definition the tail components are much more prone to outliers. The sample period ranges from January 1990 to December 2019.

Panel A: Tail components



Panel B: Core components



B Forecasting performance of tail and core components

Table 6: Detailed performance of tails and cores

This complements Table 3 by providing more details on the forecasting performance on the tails and core components separately. In panel A we note that when using the tails alone, the optimal is small. The predictors however are neither economically nor statistically significant. In panel B, we find that the core components are both economically and statistically significant. The resulting R-squared reaches almost the one with all four components jointly, showing that the cores provide the main forecasting power.

Panel A: Tail components

	$\alpha^T = 1.51\%$			
	$r_{t-1:t}$	$RV^{T-}_{t-1:t}$	$RV^{T+}_{t-1:t}$	R^2
$r_{t:t+1}$	-0.03	-0.01	-0.01	3.58
	(-0.41)	(-0.59)	(-1.15)	
$RV^{T-}_{t:t+1}$	-0.94*	0.19	0.08	7.91
	(-2.60)	(0.88)	(1.13)	
$RV^{T+}_{t:t+1}$	-0.76‡	0.30†	-0.00	13.45
	(-2.00)	(1.85)	(-0.06)	

Panel B: Core components

	$\alpha^C = 30.22\%$			
	$r_{t-1:t}$	$RV^{C-}_{t-1:t}$	$RV^{C+}_{t-1:t}$	R^2
$r_{t:t+1}$	-0.04	16.31*	-18.75*	6.44
	(-0.65)	(2.79)	(-2.91)	
$RV^{C-}_{t:t+1}$	-0.01*	-1.58	2.41	55.98
	(-2.82)	(-1.03)	(1.41)	
$RV^{C+}_{t:t+1}$	-0.01‡	-1.57	2.33	52.88
	(-2.45)	(-1.08)	(1.44)	

C Alternative downside-upside thresholds

Table 7: Within-sample predictability with alternative thresholds of downside and upside

This table contains the coefficients estimates and the t-statistics in brackets. The last column contains the adjusted R-squared. Further, α denotes the threshold that distinguishes between tail and core of the return monthly distribution of five-minute returns. Lastly, α^* denotes the threshold which maximizes the adjusted R-squared of the monthly VAR. The sample period ranges from January 1990 to December 2019.

Panel A: Threshold ZERO

	$\alpha^* = 29.49\%$					
	$r_{t-1:t}$	$RV^{T-}_{t-1:t}$	$RV^{C-}_{t-1:t}$	$RV^{C+}_{t-1:t}$	$RV^{T+}_{t-1:t}$	R^2
$r_{t:t+1}$	0.05	0.15	15.75*	-17.07*	-0.19†	7.21
	(0.63)	(1.33)	(3.03)	(-2.60)	(-1.75)	
$RV^{T-}_{t:t+1}$	-0.47‡	0.10	-50.82	62.60	0.24	45.76
	(-2.12)	(0.36)	(-1.34)	(1.63)	(0.63)	
$RV^{C-}_{t:t+1}$	-0.02‡	-0.00	-1.94	2.67†	0.01	59.86
	(-2.40)	(-0.18)	(-1.25)	(1.70)	(0.68)	
$RV^{C+}_{t:t+1}$	-0.02‡	0.00	-1.68	2.31†	0.01	58.17
	(-2.09)	(0.00)	(-1.25)	(1.70)	(0.58)	
$RV^{T+}_{t:t+1}$	-0.58‡	0.02	-60.71	73.79	0.42	50.03
	(-2.21)	(0.06)	(-1.31)	(1.57)	(0.89)	

Panel B: Threshold MEDIAN

	$\alpha^* = 28.65\%$					
	$r_{t-1:t}$	$RV^{T-}_{t-1:t}$	$RV^{C-}_{t-1:t}$	$RV^{C+}_{t-1:t}$	$RV^{T+}_{t-1:t}$	R^2
$r_{t:t+1}$	-0.02	0.07	11.32*	-10.60‡	-0.13	5.75
	(-0.27)	(0.67)	(2.63)	(-2.11)	(-1.18)	
$RV^{T-}_{t:t+1}$	-0.23†	0.52‡	-16.34	16.98	0.04	40.02
	(-1.95)	(2.31)	(-0.74)	(1.00)	(0.15)	
$RV^{C-}_{t:t+1}$	-0.01‡	0.02†	-0.53	0.85	-0.00	55.27
	(-2.41)	(1.67)	(-0.55)	(1.14)	(-0.03)	
$RV^{C+}_{t:t+1}$	-0.01†	0.02	-0.66	0.90	0.00	51.40
	(-1.90)	(1.49)	(-0.70)	(1.25)	(0.24)	
$RV^{T+}_{t:t+1}$	-0.30‡	0.51†	-23.71	24.85	0.16	45.46
	(-2.20)	(1.73)	(-0.88)	(1.21)	(0.46)	

D Measure development based on Barndorff-Nielsen, Kinnebrock, and Shephard (2008)

The realized variance (RV) is non-parametric measure of variance based on high frequency data, i.e. intra-daily data (Anderesen et al., 2001). The RV is defined as the sum of squared log returns. Barndorff-Nielsen, Kinnebrock, and Shephard (2008) decomposed the RV into downside and upside RV, denoted as RV^{-} and RV^{+}. To distinguish between downside and upside returns, Barndorff-Nielsen, Kinnebrock, and Shephard (2008) define a threshold of zero. Their approach conveniently implies that the sum of downside and upside RVs yield the total RV.

$$
\begin{aligned}
RV_{t-1:t} &= \sum_{i}^{n_t} \left(r_{i,t}\right)^2 \\
RV^{-}_{t-1:t} &= \sum_{i}^{n_t} \left(r_{i,t}\right)^2 \times \mathbb{I}[r_{i,t} < 0] \\
RV^{+}_{t-1:t} &= \sum_{i}^{n_t} \left(r_{i,t}\right)^2 \times \mathbb{I}[r_{i,t} > 0] \\
RV_{t-1:t} &= RV^{-}_{t-1:t} + RV^{+}_{t-1:t]}
\end{aligned}
$$

Table 8: Forecasting performance of realized variance

This table presents the forecasting performance of the realized variance. Somewhat counter-intuitively it shows that the higher the variance, the lower the return. This is due to the heterogeneous information (downside risk and upside potential) aggregated in to a single predictor.

	$r_{t-1:t}$	$RV_{t-1:t}$	R^2
$r_{t:t+1}$	-0.02	-0.02†	2.04
	(-0.31)	(-1.96)	
$RV_{t:t+1}$	-0.88‡	0.60*	44.16
	(-2.54)	(5.10)	

Table 9: Realized Semi-variances

This table shows how downside and upside realized variance, when computed as the sum of squared return below and above zero, predict future excess returns. While the sign of the coefficient estimates convey the expected economic intuition (compensation for downside risk and lower expected return for upside potential) they lack the statistical significance resulting in low R-squared

	$r_{t-1:t}$	$RV^{-}_{t-1:t}$	$RV^{+}_{t-1:t}$	R^2
$r_{t:t+1}$	-0.00	0.07	-0.10	1.81
	(-0.05)	(0.21)	(-0.33)	
$RV^{-}_{t:t+1}$	-0.35	0.71	-0.10	42.92
	(-1.64)	(1.58)	(-0.22)	
$RV^{+}_{t:t+1}$	-0.38†	0.64	-0.01	45.05
	(-1.75)	(1.37)	(-0.02)	

Table 10: Realized downside and upside for tail and core measures

This table presents the forecasting performance when decomposing the monthly return distribution in the logic inspired by Barndorff-Nielsen, Kinnebrock, and Shephard (2008) . The results are overall similar to our approach, although we note less statistical and economical significance of the predictores as well as a lower R-squared

	$\alpha^* = 27.30\%$					
	$r_{t-1:t}$	$RV^{T-}_{t-1:t}$	$RV^{C-}_{t-1:t}$	$RV^{C+}_{t-1:t}$	$RV^{T+}_{t-1:t}$	R^2
$r_{t:t+1}$	0.12	0.32	13.88‡	-14.29‡	-0.35	4.16
	(1.16)	(0.91)	(2.47)	(-2.10)	(-0.99)	
$RV^{T-}_{t:t+1}$	-0.51	0.49	-28.94	11.95	0.75	45.15
	(-1.60)	(0.88)	(-0.96)	(0.68)	(0.63)	
$RV^{C-}_{t:t+1}$	-0.03‡	0.00	-1.38	1.16†	0.03	50.17
	(-2.15)	(0.00)	(-1.22)	(1.77)	(0.75)	
$RV^{C+}_{t:t+1}$	-0.02†	0.01	-1.13	0.77	0.03	46.66
	(-1.72)	(0.35)	(-1.01)	(1.15)	(0.68)	
$RV^{T+}_{t:t+1}$	-0.52	0.46	-27.28	9.92	0.81	47.31
	(-1.64)	(0.79)	(-0.90)	(0.57)	(0.68)	

Table 11: Forecasting performance of tail and core components

This table provides more details on the forecasting performance on the tails and core components as computed in Barndorff-Nielsen, Kinnebrock, and Shephard (2008) fashion. In panel A we find that, similar as our predictors, tail predictors alone are neither economically nor statistically significant. In panel B, we find that the core components are both economically and statistically significant. For better comparability we chose the same optimal thresholds α^T and α^C which maximize the in-sample predictability.

Panel A: Tail components

	$\alpha^T = 1.51\%$			
	$r_{t-1:t}$	$RV^{T-}_{t-1:t}$	$RV^{T+}_{t-1:t}$	R^2
$r_{t:t+1}$	-0.02	0.07	-0.20	2.44
	(-0.27)	(0.18)	(-0.59)	
$RV^{T-}_{t:t+1}$	-0.10‡	0.52†	0.02	34.94
	(-2.56)	(1.86)	(0.11)	
$RV^{T+}_{t:t+1}$	-0.10‡	0.63‡	-0.01	39.32
	(-2.51)	(1.99)	(-0.04)	

Panel B: Core components

	$\alpha^C = 30.22\%$			
	$r_{t-1:t}$	$RV^{C-}_{t-1:t}$	$RV^{C+}_{t-1:t}$	R^2
$r_{t:t+1}$	0.03	13.51‡	-15.43‡	3.47
	(0.43)	(2.01)	(-2.24)	
$RV^{C-}_{t:t+1}$	-0.02‡	-1.05	1.70	47.73
	(-2.29)	(-1.12)	(1.64)	
$RV^{C+}_{t:t+1}$	-0.01†	-0.87	1.46	41.83
	(-1.90)	(-0.89)	(1.33)	

E Local projections

Before we derive the results of our parametric model it proves usefull to be familiar with some results of the VAR framework. Let z_t follow a VAR model

$$\underset{q\times 1}{z_t} = \underset{q\times 1}{\Psi} + \underset{q\times q}{\Phi}\underset{q\times 1}{z_{t-1}} + \underset{q\times 1}{\epsilon_t} \tag{14}$$

where $\mathbb{E}[\epsilon_t] = 0$ and $\mathbb{V}(\epsilon_t) = \underset{q\times q}{\Sigma}$.

Further, the stationarity of the VAR process let's solve for closed form solution of $\mathbb{E}[z_t]$. Rewriting equation (14) yields

$$\begin{aligned} \mathbb{E}[z_t] =& \Psi + \Phi\mathbb{E}[z_{t-1}] \\ =& \Psi + \Phi\mathbb{E}[z_t] \end{aligned}$$

$$\begin{aligned} \Rightarrow \qquad (I-\Phi)\mathbb{E}[z_t] =& \Psi \\ \mathbb{E}[z_t] =& (I-\Phi)^{-1}\Psi \end{aligned}$$

Similarily we can derive a closed form solution for expectation about future outcomes.

$$\begin{aligned} \mathbb{E}_t[z_{t+j}] =& \Psi + \Phi\mathbb{E}_t[z_{t+j-1}] + \underbrace{\mathbb{E}[\epsilon_{t+j}]}_{=0} \\ =& \Psi + \Phi\mathbb{E}[\Psi + \Phi\mathbb{E}_t[z_{t+j-2}]] \\ =& \Psi + \Phi\Psi_{j-1} + \Phi\cdot\Phi_{j-1}\mathbb{E}_t[z_{t+j-2}]] \end{aligned} \tag{15}$$

for t large enough, we can state that $\mathbb{E}_t[z_{t+j}]] = \mathbb{E}_t[z_{t+j-2}]]$. We can hence derive a system of equation for equation (15) such that we can postulate

$$\mathbb{E}[z_{t+j}] = \Psi_j + \Phi_j z_t \qquad , j \geq 0$$

where

$$\begin{cases} \Psi_j = \Psi + \Phi\Psi_{j-1} \\ \Phi_j = \Phi\Phi_{j-1} \end{cases}$$
$$\Leftrightarrow \begin{cases} \Psi_j - (I-\Phi)^{-1}\Psi = \Phi\big(\Psi_{j-1} - (I-\Phi)^{-1}\Psi\big) \\ \Phi_j = \Phi^j \end{cases}$$

We can solve for Ψ_j explicitely

$$\Psi_j = (I-\Phi)^{-1}\Psi + \Phi^{j-1}\big(\Psi_1 - (I-\Phi)^{-1}\Psi\big)$$

$$
\begin{aligned}
&= (I-\Phi)^{-1}\Psi + \Phi^{j-1}\big(\Psi - (I-\Phi)^{-1}\Psi\big) \\
&= (I-\Phi)^{-1}\Psi + \Phi^{j-1}(I-\Phi)^{-1}\big((I-\Phi)\Psi - \Psi\big) \\
&= (I-\Phi)^{-1}\Psi + \Phi^{j-1}(I-\Phi)^{-1}\Phi\Psi \\
&= (I-\Phi)^{-1}\Psi + \Phi^{j}(I-\Phi)^{-1}\Psi \\
&= (I-\Phi^{j})(I-\Phi)^{-1}\Psi
\end{aligned}
$$

We can hence write

$$
\begin{aligned}
\mathbb{E}_t[z_{t+j}] &= \Psi_j + \Phi_j z_t \\
&= (I-\Phi^j)(I-\Phi)^{-1}\Psi + \Phi^j z_t
\end{aligned} \tag{16}
$$

F Parametric estimation as an alternative to aggregation

We define forward and backwards aggregated variables respectively as

$$
\begin{aligned}
z_{t+1:t+h} &= \sum_{j=1}^{k} z_{t+j} = z_{t+1} + z_{t+2} + ... + z_{t+h} \\
z_{t-m+1:t} &= \sum_{i=1}^{m} z_{t-i+1} = z_{t-m+1} + z_{t-m+2} + ... + z_t
\end{aligned}
$$

We further define a selection mechanism

$$
\begin{aligned}
y_t &= \iota^\top z_t \qquad \text{where } \iota \text{ is a selection vector} \\
x_t &= \Lambda z_t \qquad \text{where } \Lambda \text{ is a selection matrix}
\end{aligned}
$$

So we can rewrite the regression for the selected variables as follows:

$$
y_{t+1:t+h} = \psi_{m,h} + \Phi_{m,h}^\top x_{t-m+1:t} + u_{t+h}^{(m)} \tag{17}
$$

and solve for the estimates

$$
\begin{aligned}
\psi_{m,h} &= \mathbb{E}[y_{t+1:t+h}] - \Phi_{m,h}^\top \mathbb{E}[x_{t-m+1:t}] \\
\Phi_{m,h} &= \left(\Sigma_{mm}^{xx}\right)^{-1}\Sigma_{m,h}^{xy}
\end{aligned}
$$

We now want to explicitly express these parameters as functions of the VAR parameters as well as the corresponding periods of forward and backward aggregation, denoted by h and m respectively. The gist lies in the stationarity of the underlying process.

Let's start with $\psi_{m,h}$:

$$\begin{aligned}\mathbb{E}[y_{t+1:t+h}] &= \sum_{j=1}^{h} \mathbb{E}[y_{t+j}] = \sum_{j=1}^{h} \mathbb{E}[y_t] \\ &= h\mathbb{E}[y_t] = h\mathbb{E}[\iota^\top z_t] = h\iota^\top \mathbb{E}[z_t] \\ &= h\iota^\top (I - \Phi)^{-1}\Psi\end{aligned}$$

$$\begin{aligned}\mathbb{E}[x_{t-m+1:t}] &= \sum_{i=1}^{m} \mathbb{E}[x_{t-i+1}] = \sum_{i=1}^{m} \mathbb{E}[x_t] \\ &= m\mathbb{E}[x_t] = m\mathbb{E}[\Lambda z_t] = m\Lambda\mathbb{E}[z_t] \\ &= m\Lambda(I - \Phi)^{-1}\Psi\end{aligned}$$

Hence, we can rewrite $\psi_{m,h}$ as

$$\begin{aligned}\psi_{m,h} &= \mathbb{E}[y_{t+1:t+h}] - \Phi_{m,h}^\top \mathbb{E}[x_{t-m+1:t}] \\ &= h\iota^\top (I - \Phi)^{-1}\Psi - \Phi_{m,h}^\top m\Lambda(I - \Phi)^{-1}\Psi\end{aligned} \tag{18}$$

We now want to express $\Phi_{m,h}$ in a similiar fashion. To do so we use the same approach to solve for $\Sigma_{m,h}^{xy}$ and $\Sigma_{m,m}^{xx}$. It proves useful to introduce the following theorem

Theorem

$$\mathbb{C}(X, Y) = \mathbb{C}\big[\mathbb{E}_c[X], \mathbb{E}_c[Y]\big] + \mathbb{E}\big[\mathbb{C}_c(X, Y)\big] \tag{19}$$

$$\begin{aligned}\Sigma_{m,h}^{xy} &= \mathbb{C}[x_{t-m+1:t}, y_{t+1:t+h}] \\ &= \mathbb{C}\big[\mathbb{E}_t[x_{t-m+1:t}], \mathbb{E}_t[y_{t+1:t+h}]\big] + \underbrace{\mathbb{E}\big[\mathbb{C}_t[x_{t-m+1:t}], y_{t+1:t+h}]\big]}_{=0} \\ &= \mathbb{C}\big[x_{t-m+1:t}, \mathbb{E}_t[y_{t+1:t+h}]\big] \\ &= \mathbb{C}\left[\sum_{i=1}^{m} x_{t-i+1}, \sum_{j=1}^{h} \mathbb{E}_t[y_{t+j}]\right] \\ &= \mathbb{C}\left[\sum_{i=1}^{m} \Lambda z_{t-i+1}, \sum_{j=1}^{h} \mathbb{E}_t[\iota^\top z_{t+j}]\right] = \mathbb{C}\left[\Lambda \sum_{i=1}^{m} z_{t-i+1}, \iota^\top \sum_{j=1}^{h} \left(\Psi_j + \Phi_j z_t\right)\right] \\ &= \Lambda \sum_{i=1}^{m} \mathbb{C}\left[z_{t-i+1}, , \iota^\top \sum_{j=1}^{h} \left(\Phi^j\right) z_t\right]\end{aligned}$$

$$
\begin{aligned}
&= \Lambda \sum_{i=1}^{m} \mathbb{C}[z_{t-i+1}, z_t] \left(\sum_{j=1}^{h} \Phi^j \right)^\top \iota \\
&= \Lambda \left\{ \mathbb{C}[z_t, z_t] + \sum_{i=2}^{m} \mathbb{C}[z_{t-i+1},, z_t] \right\} \left(\sum_{j=1}^{h} \Phi^j \right)^\top \iota \\
&= \Lambda \left\{ \Sigma^{zz} + \sum_{i=2}^{m} \mathbb{C}\left[z_{t-i+1}, \mathbb{E}_{t-i+1}[z_t] \right] \right\} \left(\sum_{j=1}^{h} \Phi^j \right)^\top \iota \\
&= \Lambda \left\{ \Sigma^{zz} + \sum_{i=2}^{m} \mathbb{C}\left[z_{t-i+1}, \Psi_{i-1} + \Phi_{i-1} z_{t-i+1} \right] \right\} \left(\sum_{j=1}^{h} \Phi^j \right)^\top \iota \\
&= \Lambda \left\{ \Sigma^{zz} + \sum_{i=2}^{m} \mathbb{C}\left[z_{t-i+1}, \Phi^{i-1} z_{t-i+1} \right] \right\} \left(\sum_{j=1}^{h} \Phi^j \right)^\top \iota \\
&= \Lambda \left\{ \Sigma^{zz} + \sum_{i=2}^{m} \left(\mathbb{C}\left[z_{t-i+1}, z_{t-i+1} \right] \left(\Phi^{i-1} \right)^\top \right) \right\} \left(\sum_{j=1}^{h} \Phi^j \right)^\top \iota \\
&= \Lambda \left\{ \Sigma^{zz} + \sum_{i=2}^{m} \left(\Sigma^{zz} \left(\Phi^{i-1} \right)^\top \right) \right\} \left(\sum_{j=1}^{h} \Phi^j \right)^\top \iota \\
&= \Lambda \sum_{i=1}^{m} \left(\Sigma^{zz} \left(\Phi^{i-1} \right)^\top \right) \left(\sum_{j=1}^{h} \Phi^j \right)^\top \iota \\
\Sigma_{m,h}^{xy} &= \Lambda \Sigma^{zz} \left(\sum_{i=1}^{m} \Phi^{i-1} \right)^\top \left(\sum_{j=1}^{h} \Phi^j \right)^\top \iota \qquad (20) \\
\Sigma_{m,h}^{xy} &= \Lambda \Sigma^{zz} \left((I - \Phi)^{-1} (I - \Phi^m) \right)^\top \left((I - \Phi)^{-1} (\Phi - \Phi^{h+1}) \right)^\top \iota \qquad (21)
\end{aligned}
$$

where from equation (14)

$$
\begin{aligned}
\Sigma^{zz} &= \mathbb{V}[z_t] = \Phi \mathbb{V}[z_{t-1}] \Phi^\top + Var[\epsilon_t] \\
&= \Phi \Sigma^{zz} \Phi^\top + \Sigma \\
\Rightarrow vec(\Sigma^{zz}) &= vec(\Phi \Sigma^{zz} \Phi^\top) + vec(\Sigma) \\
&= (\Phi \otimes \Phi) vec(\Sigma^{zz}) + vec(\Sigma) \\
&= [I - \Phi \otimes \Phi]^{-1} vec(\Sigma) \qquad (22)
\end{aligned}
$$

and, more generally for variance covariance between z_{t-i} and z_t is given by

$$
\begin{aligned}
\Sigma_i^{zz} &= \mathbb{C}[z_{t-i}, z_t] = \mathbb{C}\left[z_{t-i}, \mathbb{E}[z_t] \right] \\
&= \mathbb{C}[z_{t-i}, \Psi_i + \Phi_i z_{t-i}] = \Sigma^{zz} \left(\Phi^i \right)^\top \qquad (23)
\end{aligned}
$$

Now, let's solve for the backward aggregated variable.

$$
\begin{aligned}
\Sigma^{xx}_{m,m} =& \mathbb{V}[x_{t-m+1:t}] = \mathbb{V}\left[\sum_{i=1}^{m} x_{t-i+1}\right] \\
=& \mathbb{V}\left[\Lambda \sum_{i=1}^{m} z_{t-i+1}\right] = \Lambda \mathbb{V}\left[\sum_{i=1}^{m} z_{t-i+1}\right]\Lambda^\top \\
=& \Lambda\left[m\Sigma^{zz} + \sum_{i=1}^{m-1}(m-i)\left(\Sigma^{zz}_i + (\Sigma^{zz}_i)^\top\right)\right]\Lambda^\top \\
=& \Lambda\left[m\Sigma^{zz} + \sum_{i=1}^{m-1}(m-i)\left(\Sigma^{zz}(\Phi^i)^\top + (\Phi^i)\Sigma^{zz}\right)\right]\Lambda^\top \\
=& \Lambda\left[m\Sigma^{zz} + \Sigma^{zz}\left(\sum_{i=1}^{m-1}(m-i)(\Phi^i)\right)^\top + \left(\sum_{i=1}^{m-1}(m-i)(\Phi^i)\right)\Sigma^{zz}\right)\right]\Lambda^\top \\
=& \Lambda\left[m\Sigma^{zz} + \Sigma^{zz}\left(\left[(I-\Phi)^2\right]^{-1}\left[\Phi^{m+1} - m\Phi^2 + (m-1)\Phi\right]\right)^\top \right. \\
& \qquad \left. + \left[(I-\Phi)^2\right]^{-1}\left[\Phi^{m+1} - m\Phi^2 + (m-1)\Phi\right]\Sigma^{zz}\right)\Big]\Lambda^\top
\end{aligned}
\tag{24}
$$

Proof

We define the following function

$$
\begin{aligned}
f(x) =& \sum_{i=1}^{m-1}(m-i)x^i = \sum_{i=1}^{m-1} i x^{m-i} = x^{m-1}\sum_{i=1}^{m-1} x^{1-i} = x^{m-1}\left(\sum_{i=1}^{m-1} y^i\right)'\Big|_{y=x^{-1}} \\
=& x^{m-1}\left(\frac{y-y^m}{1-y}\right)'\Big|_{y=x^{-1}} \\
=& x^{m-1}\frac{(1-my^{m-1})(1-y)+(y-y^m)}{(1-y)^2}\Big|_{y=x^{-1}} \\
=& x^{m-1}\frac{1-my^{m-1}+(m-1)y^m}{(1-y)^2}\Big|_{y=x^{-1}} \\
=& x^{m-1}\frac{1-mx^{1-m}+(m-1)x^{-m}}{(1-x^{-1})^2} = \frac{x^{m-1}-m+(m-1)x^{-1}}{x^{-2}(1-x)^2} \\
=& \frac{x^{m+1}-mx^2+(m-1)x}{(1-x)^2}
\end{aligned}
\tag{25}
$$

Equivalently, if x is matrix, we can rewrite equation (25) as

$$
\sum_{i=1}^{m-1}(m-1)\Phi^i = \left[(I-\Phi)^2\right]^{-1}\left[\Phi^{m+1} - m\Phi^2 + (m-1)\Phi\right] \qquad Q.E.D.
$$

Finally we can rewrite $\Phi_{m,h}$ as

$$\begin{aligned}\Phi_{m,h} =& \left(\Sigma^{xx}_{m,m}\right)^{-1}\Sigma^{xy}_{m,h} \\ =& \Bigg(\Lambda\Bigg[m\Sigma^{zz} + \Sigma^{zz}\Bigg(\Big[(I-\Phi)^2\Big]^{-1}\Big[\Phi^{m+1} - m\Phi^2 + (m-1)\Phi\Big]\Bigg)^\top \\ &\qquad + \Big[(I-\Phi)^2\Big]^{-1}\Big[\Phi^{m+1} - m\Phi^2 + (m-1)\Phi\Big]\Sigma^{zz}\Bigg)\Bigg]\Lambda^\top\Bigg)^{-1} \times \\ &\Lambda\Sigma^{zz}\Big((I-\Phi)^{-1}(I-\Phi^m)\Big)^\top\Big((I-\Phi)^{-1}(\Phi-\Phi^{h+1})\Big)^\top\iota\end{aligned}$$

with $vec(\Sigma^{zz}) = [I - \Phi\otimes\Phi]^{-1}vec(\Sigma)$.

And recall $\psi_{m,h}$

$$\psi_{m,h} = h\iota^\top(I-\Phi)^{-1}\Psi - \Phi^\top_{m,h}m\Lambda(I-\Phi)^{-1}\Psi$$

Recall the regression

$$y_{t+1:t+h} = \psi_{m,h} + \Phi^\top_{m,h}x_{t-m+1:t} + u^{(m)}_{t+h}$$

The $R^2_{m,h}$ is given by

$$\begin{aligned}R^2_{m,h} =& \frac{\Phi^\top_{m,h}\Sigma^{xx}_{m,m}\Phi_{m,h}}{\Sigma^{yy}_{h,h}} = \frac{\Sigma^{xy}_{m,h}\left(\Sigma^{xx}_{m,m}\right)^{-1}\Sigma^{xx}_{m,m}\left(\Sigma^{xx}_{m,m}\right)^{-1}\Sigma^{xy}_{m,h}}{\Sigma^{yy}_{h,h}} \\ =& \frac{\Sigma^{xy}_{m,h}\left(\Sigma^{xx}_{m,m}\right)^{-1}\Sigma^{xy}_{m,h}}{\Sigma^{yy}_{h,h}}\end{aligned} \tag{26}$$