Horizon-specific risks, higher moments, and asset prices.

Jozef BARUNÍK[a]*, and Josef KURKA[b]

[a] Institute of Economic Studies, Charles University, Opletalova 26, 110 00, Prague, Czech Republic

[b] Department of Econometrics, IITA, The Czech Academy of Sciences, Pod Vodarenskou Vezi 4, 182 00, Prague, Czech Republic

January 15, 2020

This is a preliminary draft, please, do not cite, copy or distribute without permission of the authors.

Abstract

Asset pricing traditionally works with information aggregated over horizons, however investors' preferences are horizon-specific. Decomposing returns, and risk factors to components representing individual horizons may hence provide valuable insights into pricing mechanisms of investors. With increasing size of factor-investing literature, the number of factors approximating risk, and possibly explaining the cross section of returns is growing rapidly. However, most of the factors perform poorly in subsequent out-of-sample testing. Therefore, attention should be turned to theory-based factors approximating the risks such as moments of the return distribution that are found to be priced empirically. We derive an asset pricing model that contains second, third and fourth centralized moments of returns on aggregate wealth decomposed to short-run and long-run components. Empirical results show that horizon-specific risk from higher moments is priced, and uncovers different effects of the moment-based risk factors in short-run, and long-run.

1 Introduction

2 Relevant literature

Factor-based models approximate the pricing kernel using single or multiple risk factors. The cornerstone of factor investing is capital asset pricing model (CAPM) developed by Sharpe (1964), Lintner (1965), and Black (1972). The simplicity of CAPM was appealing, however numerous problems in terms of empirical validity appeared; e.g. discrepancies in significance and magnitude of estimated coefficients in different subsamples, or inconsistency of the beta-return relationship. In addition, other factors besides beta were found to significantly influence

*Corresponding author, Tel. +420(776)259273, Email address: barunik@fsv.cuni.cz

excess returns. Failures of CAPM led to emergence of alternative models trying to build on its foundations, e.g. arbitrage pricing theory (APT) developed by Ross (1976). Main idea of APT is that such a complex problem as systemic risk and asset returns cannot be explained by a single factor, but rather by a combination of factors.

Since then, many multi-factor models have been proposed, fitting various potential pricing factors from different classes (statistical factors, macroeconomic factors, or market factors) into the equation of excess returns. One of the most pronounced multi-factor models is the Fama-French three factor model (FF3), which contains the market beta analogously to CAPM, and additionally accounts for the role of market capitalization and book-to-market ratio. The amount of attention a model receives, however, is not an indicator of its quality, which shall be assessed rather based on model's out-of-sample performance. Out-of-sample performance not only provides a better notion about sources of cross-sectional returns predictability, but also helps uncover which relationships were false discoveries due to spurious correlations or misused methodology. There is a large stream of recent literature approaching the validity of risk factors (McLean and Pontiff, 2016; Hou et al., 2017; Harvey et al., 2016), and the joint conclusion is insignificance of majority of factors in out-of-sample testing adjusted for pronounced faults in research design.

2.1 Higher moments and asset pricing

Poor out-of-sample performance of empirical factors suggests, that attention should be turned to risk factors capturing the fundaments of asset returns, such as moments of distribution. A lot of emphasis has been placed on the first two moments in terms of risk-return relationships (e.g., Bakshi and Kapadia, 2003; Ang et al., 2006), whereas higher moments are regularly omitted. Low amount of attention higher moments receive is not proportional to magnitude of their importance, since their relevance to asset pricing is theoretically and empirically documented, their absence stems rather from drive for model simplicity. Asset pricing models are almost exclusively derived assuming risk aversion of investors, which implies a negative relationship between risk and investors' utility. Asset riskiness consists of several aspects, although it is often proxied solely by volatility. Volatility expresses level of uncertainty, while higher moments contain information about probability and sign of extreme events. Intuitivelly, investors should not neglect any of these information in the decision making process.

Why have then higher moments not been receiving corresponding amount of attention? CAPM, leading benchmark model does not hold unless returns are normally distributed (Chung et al., 2006), and under normally distributed returns the data are fully described by the first two moments, hence higher moments become unimportant. However in real world financial data, departures from normal distribution are documented (e.g., Fama, 1976), and the fat-tailedness or asymmetry in distribution become even more significant when returns are compounded over multiple periods (Fama, 1996), or grouped into a stock index (Bakshi et al., 2003). Thus, problems that have arisen in standard asset pricing and portfolio selection models, e.g. equity premium puzzle [1] or deliberate underdiversification [2], may be explainable through tails of distribution, i.e. higher moments. Introduction of rapid downward jumps helps explaining the equity premium puzzle (Rietz, 1988; Longstaff and Piazzesi, 2004), even if the jumps are offset by positive jumps of the same magnitude to make the distribution symmetric (Barro,

[1] Mehra and Prescott (1985) noted that class of general equilibrium models is not able to explain large average equity risk premia and low risk-free rate observed on US markets.

[2] Investors deliberately hold insufficiently diversified portfolios, although they would be capable of obtaining a sufficient number of assets to fully diversify away the idiosyncratic risk.

2006). Barberis and Huang (2008) observe that investors deciding under risk often depart from the expected utility framework. They suggest the decision-making is better modelled under cumulative prospect theory, whose overweighting of tales better coincides with preference for positively skewed assets displayed by investors [3]. Other type of irrational behaviour in form of deliberate underdiversification may also be explained by preference for "lottery-like" (positively skewed) assets, since investors fear losing the possibility of extremely high returns if they become completely diversified (Simkowitz and Beedles, 1978; Mitton and Vorkink, 2007).

Given the role of tail events in explaining financial market puzzles, common sense suggests they should also be reflected in asset returns. Asset pricing models rely on an approximation of the stochastic discount factor (SDF), i.e. the pricing kernel, which can be derived either assuming a particular form of utility function, or in a model-free manner using Taylor expansion. Dittmar (2002) argues that results of many multi-factor or nonlinear pricing kernel models stem from arbitrary assumptions about utility function, hence the model-free approach should be favoured. Even with the model-free approximation of the pricing kernel, certain cut-off point must be chosen to attain the best trade-off between simplicity and precision. The cut-off is frequently made before elements representing higher moments than variance, due to belief of low economic significance of such elements, and drive for the most simplistic models possible (Jean, 1971).

However, model quality should not be assessed based on simplicity, but rather based on in- and out-of-sample return predictability. Hence models should include all elements improving their precision. As noted above, indicators of tail events belong among such elements. Preference for skewness and aversion to kurtosis can be motivated by already existing concepts, while sticking with the model-free approach. Arditti (1967) motivates preference for positive skewness via decreasing risk premia in wealth, Kraus and Litzenberger (1976) formulate the three-moment CAPM based on properties of theoretically feasible utility functions [4], Fang and Lai (1997) add cokurtosis to formulate four-moment CAPM. Dittmar (2002) derives set of conditions under which counterintuitive risk taking is eliminated. Risk aversion, decreasing absolute risk aversion and decreasing absolute prudence must be assumed to accomplish that, which requires expansion up to element representing the fourth moment. The resulting pricing kernel implies that investors have preferences over both skewness and kurtosis.

Harvey and Siddique (2000) note that failures of CAPM are most significant for assets in the lowest deciles of market-cap, i.e. for the most significantly skewed assets. In their model based on pricing kernel quadratic in market returns[5], excess returns are dependent on covariance and coskewness with the market. Chabi-Yo (2012) derives a pricing kernel with stochastic volatility, skewness and kurtosis risk using an alternative model-free approach; small-noise expansion. Maheu et al. (2013) build on the theoretical base provided by Chabi-Yo (2012), and derive an autoregressive asset pricing model containing jumps stemming from higher moment risk. One of the intermediate steps in the derivation relates excess returns to second, third and fourth centralized moment of returns on aggregate wealth.

Majority of empirical results support strong performance of models containing higher-moment risk. Significant role of idiosyncratic skewness is widely confirmed, either when using model-free implied measures (Conrad et al., 2013), ex post realized measures (Amaya et al.,

[3] Barberis and Huang (2008) claim their model could explain e.g. poor performance of IPOs or success of momentum strategies.

[4] Desirable properties of utility function according to Arrow (1970): i) positive marginal utility with respect to wealth, ii) decreasing marginal utility in wealth, i.e. risk aversion, iii) non-increasing absolute risk aversion (Kraus and Litzenberger, 1976).

[5] Such pricing kernel is derived by Taylor expansion cut-off before the fourth derivative.

2015), or idiosyncratic skewness forecasted by a time-series model (Boyer et al., 2009). Important role of idiosyncratic skewness indicates that investors are willing to accept low returns and high volatility, if they are compensated by positive skewness. Such phenomenon is closely connected to deliberate underdiversification (Simkowitz and Beedles, 1978; Mitton and Vorkink, 2007), which may be caused by presence of "lotto investors"; i.e. those demanding "lottery-like" assets with high upside potential. Their presence on financial markets is supported by strong predictive power of maximum past returns in Bali et al. (2011). Preference for "lottery-like" assets also plays a central role in explaining the idiosyncratic volatility puzzle (Hou and Loh, 2016)[6].

Three-moment CAPM of Kraus and Litzenberger (1976) produces mixed results. There is evidence that investors pay positive risk premia for skewness, however the results are not very robust to model specification (Friend and Westerfield, 1980). Four-moment CAPM is preferred over classical CAPM, either when tested on futures markets data (Christie-David and Chaudhry, 2001), or in the initial testing on NYSE data (Fang and Lai, 1997). Significant exposure to skewness risk is reported also for option implied returns of stocks included in the Center for Research in Securities Prices (CRSP) NYSE/Amex/Nasdaq daily stock file (Chang et al., 2013), exposure to skewness and kurtosis risk is also documented for hedge funds (Agarwal et al., 2009). It shall be noted that it is crucial to check robustness of the results obtained, since they could be distorted by grouping of data into portfolios; either by canceling out individual departures from linearity (Friend and Westerfield, 1980), or by magnifying the effects of skewness (Boyer et al., 2009).

2.2 Horizon-specific risk

Apart from non-robust significance of risk factors, failures of asset pricing and portfolio selection models are related to specification of investment horizon. It was noted way back, that horizon choice significantly affects model outcomes in terms of asset pricing (Levhari and Levy, 1977), portfolio selection (Tobin, 1965), and portfolio performance (Levy, 1972). In light of such discoveries, models able to capture heterogeneous preferences across investment horizons started to emerge (Gressis et al., 1976; Lee et al., 1990).

A natural way to explicitly model heterogeneous investment horizons is frequency domain analysis (see e.g., Ramsey and Lampart, 1998; Rua and Nunes, 2009). Although the origins of the method date to the distant past, increased number of its economic applications is a phenomenon of the most recent years. Heterogeneity of investment horizons directly enters the model of Crouzet et al. (2016) with frequency-specific traders who concentrate on different trading horizons due to high cost of acquiring complete information.

Returns, risk, and other economic determinants can also be decomposed to elements operating on various levels of persistence (Adrian and Rosenberg, 2008). In their seminal work, Bansal and Yaron (2004) suggest such decomposition of consumption and dividend growth processes. Shocks to consumption at different frequencies have different implications for model outcomes; they enter the pricing kernel with different weights (Dew-Becker and Giglio, 2016), have varying effects on asset returns (Ortu et al., 2013; Yu, 2012), and lifetime utility Bidder and Dew-Becker (2016), exposure of firms' cash flow to shocks with different persistence also

[6] Idiosyncratic volatility puzzle is a phenomenon observed by Ang et al. (2006), who documented a negative relationship between idiosyncratic volatility and returns. This is very puzzling as investors should require positive risk premia, if any, for idiosyncratic volatility. However, high idiosyncratic volatility indicates possible high future exposure to idiosyncratic skewness (Boyer et al., 2009). Preference for right skewed assets, along with market frictions, holds a prominent place amongst explanations of idiosyncratic volatility puzzle (Hou and Loh, 2016).

varies (Li and Zhang, 2016). Bandi and Tamoni (2017) decompose the betas in consumption CAPM model, thus disentangle the effect of exposure to market risk for various horizons. Other literature dealing with risk factors decomposition includes Kamara et al. (2016) or Boons and Tamoni (2015).

Varying preferences of investors over different horizons justify some degree of horizon-dependence in their risk attitudes. Such behaviour is acceptable e.g. under myopic loss aversion, under which participation of an individual on an experiment (investment, bet, game, etc.) depends on horizon of evaluation (for details see Benartzi and Thaler, 1995). Horizon-dependent risk aversion requires adjustments of standardly used preferences, e.g. Epstein-Zin (Epstein and Zin, 2013), described by a discount factor and risk aversion parameter. Under horizon-dependence, true representation of preferences requires addition of a patience coefficient (Gonzalo and Olmo, 2016). This is not of primary concern to our research, since we employ the model-free approach (e.g., Dittmar, 2002; Chabi-Yo, 2012), however using the resulting coefficients of our model, we will be able to disentangle short-run and long-run characteristics of investors' risk attitude. Moreover, our model allows us to separately examine the effects of volatility, skewness and kurtosis in different horizons.

3 Methodology

We derive our model building on the approach of Maheu et al. (2013) and Chabi-Yo (2012), using an unspecified utility function of wealth $U(W_{t+1})$, i.e. outcome of our model is not driven by assumptions on particular design of the utility function.

Proposition 3.1. *Consider return on asset i R^i_{t+1} as an outcome of the Euler equation*

$$E_t(R^i_{t+1}M_{t+1} \mid \Omega_t) = 1, \tag{1}$$

and the pricing kernel M_{t+1} derived from the general utility function of wealth

$$\begin{aligned} M_{t+1} &\approx \sum_{n=0}^{3} \frac{U^{(n+1)}(1+C_t)}{U'(1)n!}(R^w_{t+1} - C_t)^n \\ &= g_{0,t} + g_{1,t}(R^w_{t+1} - C_t) + g_{2,t}(R^w_{t+1} - C_t)^2 + g_{3,t}(R^w_{t+1} - C_t)^3, \end{aligned} \tag{2}$$

where $g_{n,t} = [U^{(n+1)}(1+C_t)/U'(1)][1/n!] = [U^{(n+1)}(1+C_t)/U'(1+C_t)n!][U'(1+C_t)/U'(1)]$.

Then, we are able to express excess returns of asset i as follows

$$\begin{aligned} E_t(R^{i,e}_{t+1}) =& \theta^{(s)}_{1,t}E_t[(\epsilon^{(s)}_{t+1})^2] + \theta^{(l)}_{1,t}E_t[(\epsilon^{(l)}_{t+1})^2] + \theta^{(s)}_{2,t}E_t[(\epsilon^{(s)}_{t+1})^3] + \theta^{(l)}_{2,t}E_t[(\epsilon^{(l)}_{t+1})^3] + \\ & \theta^{(s)}_{3,t}E_t[(\epsilon^{(s)}_{t+1})^4] + \theta^{(l)}_{3,t}E_t[(\epsilon^{(l)}_{t+1})^4], \end{aligned} \tag{3}$$

where $E_t[(\epsilon_{t+1})^n]$ denotes n-th centralized moment of return on aggregate wealth, $\epsilon^{(s)}_{t+1}$ represents the short-run component of returns on aggregate wealth, and $\epsilon^{(l)}_{t+1}$ represents the long-run component of returns on aggregate wealth.

Proposition 3.1 assumes a general utility function of wealth can be accurately approximated by taking a Taylor expansion up to the fourth order, which is justified in Dittmar (2002). The expansion point $W_t(1 + C_t)$ is specified by generalizing C_t as

$$C_t = E_t(R^w_{t+1}). \tag{4}$$

The crucial assumption for derivation of Proposition 3.1 is orthogonality of components at individual scales, i.e.

$$\begin{aligned} (\epsilon_{t+1}^{(s)})^n(\epsilon_{t+1}^{(l)})^n &= 0, \\ (\epsilon_{t+1}^{(s)})^n R_{t+1}^{(l)} &= 0, \\ (\epsilon_{t+1}^{(l)})^n R_{t+1}^{(s)} &= 0, \end{aligned} \tag{5}$$

for $n = \{1, 2, 3\}$[7].

4 Model and data

Equation (32), and its variations including more horizons, can be estimated using the Fama-Macbeth two stage regression. Centralized moments of returns on aggregate wealth are approximated by returns on a market index, and decomposed to short-run, medium-run, and long-run components. Firstly, we regress returns of the individual assets on horizon-specific centralized moments of market returns

$$\begin{aligned} r_{i,t+1} =& \alpha_{1,i} + \beta_i^{(s)}\sigma_{m,t}^{(s)} + \beta_i^{(m)}\sigma_{m,t}^{(m)} + \beta_i^{(l)}\sigma_{m,t}^{(l)} + \gamma_i^{(s)}s_{m,t}^{(s)} + \gamma_i^{(m)}s_{m,t}^{(m)} + \gamma_i^{(l)}s_{m,t}^{(l)} + \\ & \mathcal{K}_i^{(s)}k_{m,t}^{(s)} + \mathcal{K}_i^{(m)}k_{m,t}^{(m)} + \mathcal{K}_i^{(l)}k_{m,t}^{(l)} + \epsilon_{t+1}, \end{aligned} \tag{6}$$

where $r_{i,t+1}$ is log-return of asset i on day $t+1$, $i = \{1, ..., n\}$, $t = \{1, ..., T\}$, ϵ_{t+1} is the error term at $t + 1$, and construction of the risk factors on the right hand side is shown in the next subsection. Secondly, mean return of each asset i through the given time period is regressed on the factor loadings obtained in Equation (6)

$$\begin{aligned} \bar{r}_i =& \alpha_2 + \lambda_\sigma^{(s)}\beta_i^{(s)} + \lambda_\sigma^{(m)}\beta_i^{(m)} + \lambda_\sigma^{(l)}\beta_i^{(l)} + \lambda_s^{(s)}\gamma_i^{(s)} + \lambda_s^{(m)}\gamma_i^{(m)} + \lambda_s^{(l)}\gamma_i^{(l)} + \\ & \lambda_k^{(s)}\mathcal{K}_i^{(s)} + \lambda_k^{(m)}\mathcal{K}_i^{(m)} + \lambda_k^{(l)}\mathcal{K}_i^{(l)} + \mathcal{E}_i, \end{aligned} \tag{7}$$

where $\bar{r}_i = 1/T\sum_{t=1}^T r_{t,i}$, $i = \{1, ..., n\}$, $\mathcal{E}_i$ is the error term for asset i, and $\beta_i^{(s)}$, $\beta_i^{(m)}$, $\beta_i^{(l)}$, $\gamma_i^{(s)}$, ..., $\mathcal{K}_i^{(l)}$ are the coefficients from Equation (6).

4.1 Market data

The market is approximated by S&P 500 index. We compute realized measures representing second, third, and fourth moments of returns distribution; Realized Variance (RV), Realized Skeweness (RS), and Realized Kurtosis (RK); from 5-minute prices of the index. RV is a non-parametric measure of variance proposed by Andersen et al. (2001), and Barndorff-Nielsen and Shephard (2002). It is constructed using the high-frequency intraday prices

$$RV_t = \sum_{i=1}^{N} r_{t,i}^2, \tag{8}$$

[7] Complete derivation of Proposition 3.1 is presented in Section A.1.

where $r_{t,i} = p_{t,i/N} - p_{t,(i-1)/N}$, $p_{t,i/N}$ is natural logarithm of i-th intraday price at day t. Amaya et al. (2015) proposed that the notion of measuring moments of distribution non-parametrically using intraday prices can be extended to higher moments. Specifically, measure of the third moment, Realized Skewness (RS)

$$RS_t = \frac{\sqrt{N}\sum_{i=1}^{N} r_{t,i}^3}{RV_t^{3/2}}, \tag{9}$$

and measure of the fourth moment, Realized Kurtosis (RK)

$$RK_t = \frac{N\sum_{i=1}^{N} r_{t,i}^4}{RV_t^2}. \tag{10}$$

4.1.1 Market moments decomposition

Let's consider a time series of a general risk factor $\{x_t\}_{t\in\mathbb{Z}}$. The process $\{x_t\}_{t\in\mathbb{Z}}$ can be decomposed in a following way:

$$x_t = \sum_{j=1}^{\infty} x_t^{(j)},$$

where $x_t^{(j)}$ represents component of x_t operating at the j-th scale. The individual components $x_t^{(j)}$ are orthogonal to each other.[8] Alternatively, to avoid the infinite sum, the risk factor x can be represented as

$$x_t = \sum_{j=1}^{J} x_t^{(j)} + \pi_{t,x}^{(J)}, \tag{11}$$

for $J \geq 1$. Equation (11) decomposes x_t into sum of components $x_t^{(j)}$ composed of fluctuations with half-life within the interval $[2^{j-1}, 2^j)$, and the long-run term $\pi_{t,x}^{(J)}$ involving fluctuations with half-life over 2^J. In this manner, we can decompose our risk factors into components operating over different horizons. As Bandi and Tamoni (2017) point out, choice of J depends on practical considerations like sample length, and economic rationale. In our case

$$\begin{aligned}
RV_t &= RV_t^{(s)} + RV_t^{(m)} + RV_t^{(l)}, \\
RV_t^{(s)} &= \sum_{j=1}^{3} RV_t^{(j)}, \\
RV_t^{(m)} &= \sum_{j=4}^{5} RV_t^{(j)}, \\
RV_t^{(l)} &= \pi_{t,RV}^{(5)},
\end{aligned}$$

where $RV_t^{(s)}$ denotes short-run realized variance, $RV_t^{(m)}$ denotes medium-run realized variance, $RV_t^{(l)}$ denotes long-run realized variance.[9] RS, and RK are decomposed analogically.

[8] For details, see Bandi and Tamoni (2017)

[9] Short-run corresponds approximately to less than 3 weeks, medium-run approximately to between 3 weeks and 3 months, long-run approximately to more than 3 months

4.2 Individual stocks data

Main empirical testing of our model is done using two datasets. Firstly, we attempt to explain returns of individual stocks present in the S&P 500, secondly returns of Exchange Traded Funds (ETF henceforth). S&P dataset consists of 456 stocks for period from 2000 to 2018. ETF dataset consists of 233 symbols for period from 01/2010 to 11/2018. We construct daily returns from intraday dividends and splits adjusted prices such that

$$r_{t,i} = p_{t,N/N} - p_{t,1/N},$$

where $pt, 1/N$ is natural logarithm of first intraday price at day t, and $p_{t,N/N}$ is the last intraday price at day t.

?DESCRIPTIVE STATISTICS?

5 Results

In this section, we describe the empirical results of our model from several perspectives. Firstly, it is crucial that the decomposition of higher moment risk to heterogeneous horizons contributes vastly to the explanation of mean returns of individual assets. We compare explanatory power of our model to CAPM, and Four-Moment CAPM (FM-CAPM). Due to known poor performance of CAPM in empirical tests, it is used as a base model, while the FM-CAPM is the benchmark model. We compare the ability to explain returns on underlying assets both in-sample, and out-of-sample.

Secondly, the results reveal which of our risk factors are priced in the cross-section of returns. We observe, firstly, which of the moments play the largest role in approximating the pricing kernel, secondly, which horizons are most important for investors. Apart from observing the role of these two aspects alone, our model could uncover strong horizon-specific effects of certain risk factor that seems unimportant on aggregate. Lastly, the signs of individual coefficients are intuitive for the aggregate FM-CAPM; investors require higher risk-premia for increases in volatility, and kurtosis, and lower risk premia for increases in skewness. Such intuition holds for the aggregate across horizons, however, our model may uncover departures for some specific horizons.

Thirdly, thanks to general and complete derivation of the model, we are able to quantify investors' risk tastes (risk aversion, absolute risk aversion, absolut prudence). More importantly, we are able to track the changes of such tastes across investment horizons. Heterogeneity of risk tastes for different horizons may be a large step towards explaining the cross-section of stocks (and other assets) returns.

5.1 S&P 500 data

Figure 1 plots actual daily returns of the S&P 500 stocks against the predicted values from respective models. If the model explains the cross-section of returns well the dots should be aligned close to the 45 degree line. CAPM is able to explain the part of the returns near median value, but fails to explain returns in the upper, and especially the lower quantiles. Overall, the cluster of points faces the horizontal direction rather than 45 degrees. Addition of skewness and kurtosis in the FM-CAPM adds significantly to explaining the cross-section of S&P 500 returns. The unexplained points from the lower quantiles are driven closer to the "perfect prediction" line, and the cluster of dots gets closer to the 45 degree slope. The improvement from FM-CAPM to horizon-specific model (HSM) is not as notable, as the difference between

Table 1: We report coefficients from the second stage Fama-Macbeth regression, i.e. Equation (7). First column displays results of CAPM model, second column results of FM-CAPM model, third column results Horizon-specific FM-CAPM. Daily S&P 500 returns.

	Model			
	CAPM	Four Moment	Fama French	Horizon-specific
Constant	0.00049***	0.00050***	0.00056***	0.00053***
	(0.00007)	(0.00008)	(0.00009)	(0.00008)
Volatility Short			0.00017**	0.00016**
			(0.00007)	(0.00007)
Volatility Medium			0.00021***	0.00020***
			(0.00005)	(0.00005)
Volatility Long			0.00067***	0.00068***
			(0.00018)	(0.00018)
Skewness Short			−0.10590***	−0.07231***
			(0.03016)	(0.01313)
Skewness Medium			0.01211	0.01411*
			(0.00762)	(0.00734)
Skewness Long			−0.02320	−0.02128
			(0.03836)	(0.03665)
Kurtosis Short			0.21194	0.21324
			(0.14653)	(0.14487)
Kurtosis Medium			0.04854	0.05295*
			(0.03244)	(0.03187)
Kurtosis Long			0.06255	0.05679
			(0.07317)	(0.07214)
Market	−0.10413***		−0.08342***	
	(0.00874)		(0.01162)	
Volatility		0.00115***		
		(0.00015)		
Skewness		−0.08444***		
		(0.01742)		
Kurtosis		0.34048*		
		(0.17612)		
SMB			−0.00437	
			(0.00869)	
HML			−0.00526	
			(0.00474)	
Observations	443	443	443	443
R^2	0.24368	0.32983	0.36502	0.36266
Adjusted R^2	0.24196	0.32526	0.34730	0.34941
Note:			*p<0.1; **p<0.05; ***p<0.01	

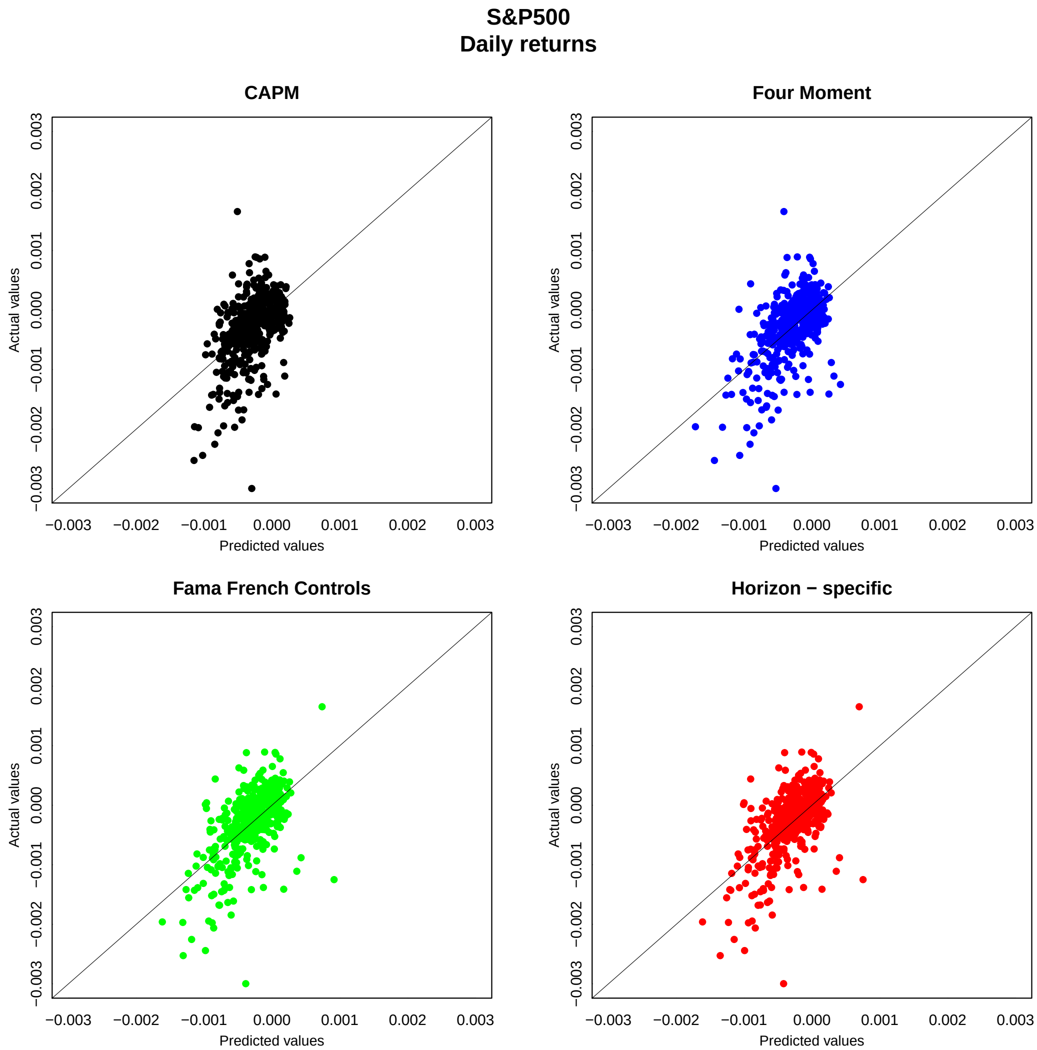

Figure 1: Daily returns predicted by the respective models plotted against actual returns S&P 500 stocks. CAPM in the top left corner (black dots), FM-CAPM in the top right corner (blue dots), Horizon-specific FM-CAPM (HS FM CAPM) in the bottom left corner (red dots).

CAPM, and FM-CAPM. Yet it is evident that decomposition of the across horizons aggregated information contributes to precision of the model.

The contribution to model precision can be quantified by adjusted R-squared, which is over 0.32 for FM-CAPM, and almost 0.36 for HSM. Therefore, uncovering the information hidden by aggregation across investment horizons brings over 10 percent increase in Adjusted R-Squared. Most of the outliers in terms of magnitude of the prediction error lie under the 45 degree line. Prevalence of under-predictions can be illustrated by studying the residuals; mean value of residuals is virtually zero, whereas the median value is positive, and it is not a property exclusive to the HSM as the same phenomenon is present in CAPM, and FM-CAPM.

The results from FM-CAPM in Table 1 indicate that the highest explanatory power is shared by volatility and skewness. High power of skewness in explaining the cross-section of stock returns is in line with previous research (Harvey and Siddique, 2000; Amaya et al., 2015; Chang et al., 2013), along with the lower risk-premia required for higher skewness. All three models also show positive relationship between relized volatility and stock returns. Both relationships remain consistent after the decomposition into individual investment horizons.

The decomposition into short-run, medium-run, and long-run has two important implications from the empirical perspective. Firstly, it uncovers relationships that stay hidden on aggregate like the investors' preference for excess kurtosis in the long-run. Such phenomenon may be, however driven by demand of positive skewness in the corresponding horizon. Secondly, we can uncover the discrepancies in the persistence of individual risk attributes. Higher required risk-premia for assets with higher volatility (driven by risk aversion) is strongest in magnitude in the long-run, while the medium-run component is also priced in the cross-section of S&P 500 returns. On the other hand preference for positive skewness, while also priced in long-run, is strongest in the short-run. Such investor behaviour may be driven by speculators seeking substantial returns on short-term investments. Overall, the highest predictive power is shared by the long-run components. This result fits well into the recent strand of literature pronouncing the importance of long-run risk in investor's decision making (e.g., Bidder and Dew-Becker, 2016; Ortu et al., 2013).

5.2 ETF data

Figure 2 plots actual daily returns of the ETF stocks against the predicted values from respective models. In case of ETFs, there is no notable difference between CAPM and FM-CAPM. For both models, the clusters of data take a horizontal shape, therefore, they explain virtually none of the returns. Poor empirical performance of both models is illustrated also by low adjusted R-squared, approximately 0.08 for both models. HSM shows significant improvement in empirical performance compared to the aforementioned models. Based on Figure 2, HSM is able to explain returns across majority of the distribution. Similarly as in the S&P dataset, the worst performance is displayed in the lowest quantiles leading to prevalence of under-predicting. The increase in adjusted R-squared compared to CAPM, and FM-CAPM is also substantial; from 0.08 to 0.19.

Table 2 shows negative slope coefficient in CAPM, which goes against economic logic. Therefore, it is not surprising that the model explains virtually none of the variation in ETF returns. The addition of skewness and kurtosis has no positive impact on model performance, moreover, standard errors in FM-CAPM get inflated, thus we identify only a significant relationship between volatility and ETF returns. Again, the decomposition into individual horizons helps uncover relationships that are not revealed when looking on the aggregates. Consistently with the model estimated on S&P 500 data, there is a positive relationship between long-run volatility and ETF returns, and strong negative effect of long-run skewness. On the other hand, insignificance of skewness in all three horizons contrasts on S&P 500 data, although the short-term demand for positive skewness may be priced through short-run volatility.[10] In addition kurtosis is positively priced in the short-run, i.e. higher required returns for excess kurtosis, and negatively priced in the medium-run, and long-run.

6 Robustness checks

Above, we show that the HSM is able to outperform both CAPM, and FM-CAPM. In light of inconsistent performance of majority factor-based models discussed in Section 2, we need to

[10]Chabi-Yo (2009) suggests that portfolios with high idiosyncratic volatility have high non-systematic coskewness.

Table 2: We report coefficients from the second stage Fama-Macbeth regression, i.e. Equation (7). First column displays results of CAPM model, second column results of FM-CAPM model, third column results Horizon-specific FM-CAPM. Daily ETF returns.

	Model			
	CAPM	Four Moment	Fama French	Horizon-specific
Constant	-0.00026^{***}	-0.00027^{***}	-0.00030^{***}	-0.00028^{***}
	(0.00003)	(0.00003)	(0.00003)	(0.00004)
Volatility Short			-0.00019^{**}	-0.00021^{**}
			(0.00009)	(0.00009)
Volatility Medium			-0.00014	-0.00013
			(0.00011)	(0.00011)
Volatility Long			0.00054^{**}	0.00046^{**}
			(0.00024)	(0.00023)
Skewness Short			-0.03826	-0.02140
			(0.04403)	(0.02591)
Skewness Medium			0.00138	0.00720
			(0.02175)	(0.02265)
Skewness Long			0.07203	0.01280
			(0.06866)	(0.07108)
Kurtosis Short			0.51306^{*}	0.48989^{*}
			(0.27944)	(0.29309)
Kurtosis Medium			-0.10012	-0.15965^{*}
			(0.07951)	(0.08210)
Kurtosis Long			-0.00109	-0.03523
			(0.07803)	(0.08174)
Market	0.01432^{***}		0.05472^{***}	
	(0.00351)		(0.01422)	
Volatility		-0.00035^{*}		
		(0.00018)		
Skewness		-0.02601		
		(0.02528)		
Kurtosis		0.49680^{*}		
		(0.28864)		
SMB			-0.00664	
			(0.00813)	
HML			-0.03880^{***}	
			(0.00834)	
Observations	233	233	233	233
R^2	0.06739	0.08952	0.25491	0.16288
Adjusted R^2	0.06336	0.07759	0.21426	0.12909
Note:			$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$	

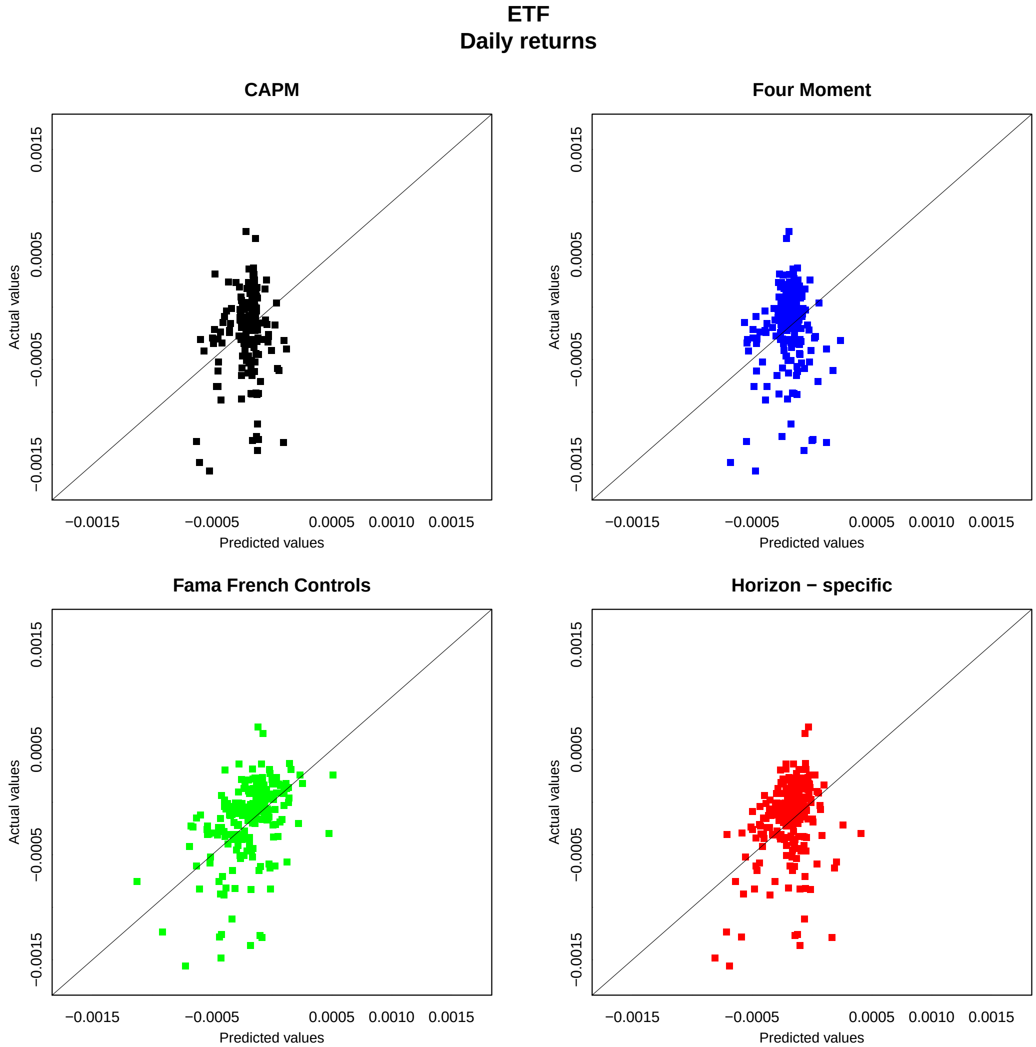

Figure 2: Daily returns predicted by the respective models plotted against actual returns ETF stocks. CAPM in the top left corner (black dots), FM-CAPM in the top right corner (blue dots), Horizon-specific FM-CAPM (HS FM CAPM) in the bottom left corner (red dots).

conduct additional tests to check robustness of the discovered relationships. Specifically, we test the ability of all three models to predict the returns out-of-sample, we estimate the models on monthly data, and we estimate single-horizon, and horizon-specific models with a single factor to account for possible multicollinearity. We report only the main results in the section below, additional figures and tables related to robustness checks can be found in Appendix B.

6.1 Portfolio sorts

6.2 Out of sample performance

We test the out of sample performance using rolling windows, i.e. we fit Equation (6), and Equation (7) using in-sample data, then estimate Equation (6) on out-of-sample data, and use the combination of $\beta_i^{(s)}, ..., \beta_i^{(l)}, \gamma_i^{(s)}, ..., \mathcal{K}_i^{(m)}$ coefficients from out-of-sample Equation (6),

Table 3: Time-series mean of the mean cross-sectional errors for S&P 500 data. RMSE in the top row, MAE in the bottom row.

Error type	In-sample	Out-of-sample	CAPM	FM-CAPM	HS FM-CAPM
RMSE	60	24	1.532	1.325	1.277
	60	36	1.278	1.084	1.047
MAE	60	24	1.156	0.955	0.923
	60	36	0.963	0.787	0.766

Table 4: Time-series mean of the mean cross-sectional errors for ETF data. RMSE in the top row, MAE in the bottom row.

Error type	In-sample	Out-of-sample	CAPM	FM-CAPM	HS FM-CAPM
RMSE	60	24	0.648	0.681	0.786
	60	36	0.511	0.533	0.634
MAE	60	24	0.479	0.495	0.544
	60	36	0.384	0.388	0.455

and $\lambda_\sigma^{(s)}, ..., \lambda_k^{(l)}$ from in-sample Equation (6) to explain the out-of-sample returns. We then roll the window by one month and repeat the estimation iteratively. Choice of the window length for in-sample, and out-of-sample data is limited by the sample size, and driven by the necessity to include sufficient portion of data for the set to be informative. We report results from models with in-sample window of length 60 months, and out-of-sample window of length 24, or 36. Mean of the cross-sectional Root Mean Square Error (RMSE), and Mean Absolute Error (MAE), throughout all the iterations are reported.

Mean out-of-sample error is 3 to 4 percent lower for HSM compared to FM-CAPM in the S&P 500 dataset, regardless of out-of-sample window length and the error type considered. HSM outperforms CAPM by over 17 percent in RMSE, and by around 20 percent when considering MAE. ETF dataset displays opposite out-of-sample results as CAPM outperforms both models. HSM has RMSE more than 20 percent higher than CAPM, MAE 14 percent higher for 24 months of out-of-sample, and 18 percent higher for 36 months of out-of-sample. FM-CAPM also shows superior performance to HSM although by a slightly smaller margin. Our model focuses on explaining the in-sample returns, therefore the fact that it outperforms both models for the S&P 500 dataset supports the superiority in explaining the in-sample returns. On the other hand, HSM displays worse out-of-sample performance than both models in the ETF dataset. However, due to lower sample length we are predicting returns only for period 2015 - 2016. This decreases the weight of this result, firstly, because the 2-year (1-year) mean is simply less informative than 11-year (12-year) mean, secondly because the HSM is not superior even in S&P 500 data in the period 2015-2016 (see Figures (3 - 6) in Appendix B).

6.3 Monthly data

In this section we reestimate the models from Section 5 on monthly data. Predicted against actual values for S&P 500 stocks are plotted in Figure (7), Appendix B. Although the superior performance of HSM is not that visible graphically, Table (5) reveals more than two percentage points increase in Adjusted R-squared from FM-CAPM to HSM. Interpretation of the horizon-specific coefficients is different than with daily data, as the intervals are altered by using monthly data. [11] The short-run horizon in Table 5 corresponds to short-run, medium-run, and part of the long-run horizon from Table 1. Therefore significant effect of short-run skewness in the HSM estimated on monthly data is in accordance to significance of short-run, and long-run skewness on daily data. The same applies to significant effect of short-run kurtosis on monthly data, and long-run kurtosis on daily data. We also find volatility in all horizons being priced in the cross-section of S&P 500 monthly returns with a positive sign.

The fit of all three models on ETF monthly returns is shown graphically in Figure 8, regression results are displayed in Table 6. Figure 8 suggests that HSM outperforms both more simplistic models, which is confirmed by substantial dominance in terms of Adjusted R-Squared. Volatility coefficients in the HSM are consistent with the daily model, although the interpretation of long-run is different. Given the ambiguous effect of kurtosis in Table 2, it is not surprising we find no significant effect of short-run kurtosis in Table 6. On the other hand we reveal that medium-run, long-run kurtosis, and long-run skewness are priced in the monthly ETF returns, however all the signs are reversed compared to Table 2. Overall, while the results for S&P stocks are quite consistent through daily and monthly data, there are inconsistencies in the regressions on ETF stocks, especially regarding direction of the effects.

6.4 Single-horizon, and single-moment models

Tables 7-12 show regression results for both datasets using only risk-factors from a single horizon, Tables 13-18 show results of regressions using all horizons of a single moment. We do these exercises to check the robustness of results under different specifications, and examine the effects of possible multicollinearity (mainly correlation between skewness and kurtosis). The results for both sets of assets correspond to the results in the core models. Adjusted R-Squared remains on a reasonable level, given the fact that we leave out significant part of predictors. Statistical significance mostly corresponds to the full models, and the signs remain unchanged. An exception in this respect is the long-run model for ETF data. Long-run components of all three moments are statistically significant in the full model (Table 2), while only volatility is statistically significant in the long-run model (Table 12), moreover, with a sign opposite to the full model.

7 Conclusion

[11] Here, short-run corresponds to less than 16 months, medium-run to between 16 and 64 months, and long-run to more than 64 months.

References

Adrian, T. and J. Rosenberg (2008). Stock returns and volatility: Pricing the short-run and long-run components of market risk. *The journal of Finance 63*(6), 2997–3030.

Agarwal, V., G. Bakshi, and J. Huij (2009). Do higher-moment equity risks explain hedge fund returns?

Amaya, D., P. Christoffersen, K. Jacobs, and A. Vasquez (2015). Does realized skewness predict the cross-section of equity returns? *Journal of Financial Economics 118*(1), 135–167.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2001). The distribution of realized exchange rate volatility. *Journal of the American statistical association 96*(453).

Ang, A., R. J. Hodrick, Y. Xing, and X. Zhang (2006). The cross-section of volatility and expected returns. *The Journal of Finance 61*(1), 259–299.

Arditti, F. D. (1967). Risk and the required return on equity. *The Journal of Finance 22*(1), 19–36.

Arrow, K. J. (1970). Essays in the theory of risk-bearing. Technical report.

Bakshi, G. and N. Kapadia (2003). Delta-hedged gains and the negative market volatility risk premium. *The Review of Financial Studies 16*(2), 527–566.

Bakshi, G., N. Kapadia, and D. Madan (2003). Stock return characteristics, skew laws, and the differential pricing of individual equity options. *The Review of Financial Studies 16*(1), 101–143.

Bali, T. G., N. Cakici, and R. F. Whitelaw (2011). Maxing out: Stocks as lotteries and the cross-section of expected returns. *Journal of Financial Economics 99*(2), 427–446.

Bandi, F. M. and A. Tamoni (2017). The horizon of systematic risk: a new beta representation.

Bansal, R. and A. Yaron (2004). Risks for the long run: A potential resolution of asset pricing puzzles. *The journal of Finance 59*(4), 1481–1509.

Barberis, N. and M. Huang (2008). Stocks as lotteries: The implications of probability weighting for security prices. *American Economic Review 98*(5), 2066–2100.

Barndorff-Nielsen, O. E. and N. Shephard (2002). Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 64*(2), 253–280.

Barro, R. J. (2006). Rare disasters and asset markets in the twentieth century. *The Quarterly Journal of Economics 121*(3), 823–866.

Benartzi, S. and R. H. Thaler (1995). Myopic loss aversion and the equity premium puzzle. *The quarterly journal of Economics 110*(1), 73–92.

Bidder, R. and I. Dew-Becker (2016). Long-run risk is the worst-case scenario. *American Economic Review 106*(9), 2494–2527.

Black, F. (1972). Capital market equilibrium with restricted borrowing. *The Journal of business 45*(3), 444–455.

Boons, M. and A. Tamoni (2015). Horizon-specific macroeconomic risks and the cross-section of expected returns.

Boyer, B., T. Mitton, and K. Vorkink (2009). Expected idiosyncratic skewness. *The Review of Financial Studies 23*(1), 169–202.

Chabi-Yo, F. (2009). Pricing kernels with coskewness and volatility risk.

Chabi-Yo, F. (2012). Pricing kernels with stochastic skewness and volatility risk. *Management Science 58*(3), 624–640.

Chang, B. Y., P. Christoffersen, and K. Jacobs (2013). Market skewness risk and the cross section of stock returns. *Journal of Financial Economics 107*(1), 46–68.

Christie-David, R. and M. Chaudhry (2001). Coskewness and cokurtosis in futures markets. *Journal of empirical finance 8*(1), 55–81.

Chung, Y. P., H. Johnson, and M. J. Schill (2006). Asset pricing when returns are nonnormal: Fama-french factors versus higher-order systematic comoments. *The Journal of Business 79*(2), 923–940.

Conrad, J., R. F. Dittmar, and E. Ghysels (2013). Ex ante skewness and expected stock returns. *The Journal of Finance 68*(1), 85–124.

Crouzet, N., I. Dew-Becker, and C. Nathanson (2016). A model of multi-frequency trade. Technical report, Northwestern University Working Paper.

Dew-Becker, I. and S. Giglio (2016). Asset pricing in the frequency domain: theory and empirics. *Review of*

Financial Studies 29(8), 2029–2068.

Dittmar, R. F. (2002). Nonlinear pricing kernels, kurtosis preference, and evidence from the cross section of equity returns. *The Journal of Finance 57*(1), 369–403.

Epstein, L. G. and S. E. Zin (2013). Substitution, risk aversion and the temporal behavior of consumption and asset returns: A theoretical framework. In *HANDBOOK OF THE FUNDAMENTALS OF FINANCIAL DECISION MAKING: Part I*, pp. 207–239. World Scientific.

Fama, E. F. (1976). *Foundations of finance: portfolio decisions and securities prices*. Basic Books (AZ).

Fama, E. F. (1996). Discounting under uncertainty. *Journal of Business*, 415–428.

Fang, H. and T.-Y. Lai (1997). Co-kurtosis and capital asset pricing. *Financial Review 32*(2), 293–307.

Friend, I. and R. Westerfield (1980). Co-skewness and capital asset pricing. *The Journal of Finance 35*(4), 897–913.

Gonzalo, J. and J. Olmo (2016). Long-term optimal portfolio allocation under dynamic horizon-specific risk aversion. Technical report, Universidad Carlos III de Madrid. Departamento de Economía.

Gressis, N., G. C. Philippatos, and J. Hayya (1976). Multiperiod portfolio analysis and the inefficiency of the market portfolio. *The Journal of Finance 31*(4), 1115–1126.

Hansen, L. P. and R. Jagannathan (1991). Implications of security market data for models of dynamic economies. *Journal of political economy 99*(2), 225–262.

Harvey, C. R., Y. Liu, and H. Zhu (2016). . . . and the cross-section of expected returns. *The Review of Financial Studies 29*(1), 5–68.

Harvey, C. R. and A. Siddique (2000). Conditional skewness in asset pricing tests. *The Journal of Finance 55*(3), 1263–1295.

Hou, K. and R. K. Loh (2016). Have we solved the idiosyncratic volatility puzzle? *Journal of Financial Economics 121*(1), 167–194.

Hou, K., C. Xue, and L. Zhang (2017). Replicating anomalies. Technical report, National Bureau of Economic Research.

Jean, W. H. (1971). The extension of portfolio analysis to three or more parameters. *Journal of financial and Quantitative Analysis 6*(1), 505–515.

Kamara, A., R. A. Korajczyk, X. Lou, and R. Sadka (2016). Horizon pricing. *Journal of Financial and Quantitative Analysis 51*(6), 1769–1793.

Kraus, A. and R. H. Litzenberger (1976). Skewness preference and the valuation of risk assets. *The Journal of Finance 31*(4), 1085–1100.

Lee, C. F., C. Wu, and K. J. Wei (1990). The heterogeneous investment horizon and the capital asset pricing model: theory and implications. *Journal of Financial and Quantitative Analysis 25*(3), 361–376.

Levhari, D. and H. Levy (1977). The capital asset pricing model and the investment horizon. *The Review of Economics and Statistics*, 92–104.

Levy, H. (1972). Portfolio performance and the investment horizon. *Management Science 18*(12), B–645.

Li, J. and H. H. Zhang (2016). Short-run and long-run consumption risks, dividend processes, and asset returns. *The Review of Financial Studies 30*(2), 588–630.

Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *The Review of Economics and Statistics 47*(1), 13–37.

Longstaff, F. A. and M. Piazzesi (2004). Corporate earnings and the equity premium. *Journal of financial Economics 74*(3), 401–421.

Maheu, J. M., T. H. McCurdy, and X. Zhao (2013). Do jumps contribute to the dynamics of the equity premium? *Journal of Financial Economics 110*(2), 457–477.

McLean, R. D. and J. Pontiff (2016). Does academic research destroy stock return predictability? *The Journal of Finance 71*(1), 5–32.

Mehra, R. and E. C. Prescott (1985). The equity premium: A puzzle. *Journal of monetary Economics 15*(2), 145–161.

Mitton, T. and K. Vorkink (2007). Equilibrium underdiversification and the preference for skewness. *The Review of Financial Studies 20*(4), 1255–1288.

Ortu, F., A. Tamoni, and C. Tebaldi (2013). Long-run risk and the persistence of consumption shocks. *Review* *of Financial Studies 26*(11), 2876–2915.

Ramsey, J. B. and C. Lampart (1998). Decomposition of economic relationships by timescale using wavelets. *Macroeconomic dynamics 2*(1), 49–71.

Rietz, T. A. (1988). The equity risk premium a solution. *Journal of monetary Economics 22*(1), 117–131.

Ross, S. A. (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory 13*(3), 341 – 360.

Rua, A. and L. C. Nunes (2009). International comovement of stock market returns: A wavelet analysis. *Journal of Empirical Finance 16*(4), 632–639.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The journal of finance 19*(3), 425–442.

Simkowitz, M. A. and W. L. Beedles (1978). Diversification in a three-moment world. *Journal of Financial and Quantitative Analysis 13*(5), 927–941.

Tobin, J. (1965). Money and economic growth. *Econometrica: Journal of the Econometric Society*, 671–684.

Yu, J. (2012). Using long-run consumption-return correlations to test asset pricing models. *Review of Economic Dynamics 15*(3), 317–335.

A Technical appendix

A.1 Derivation of Proposition 3.1

Hansen and Jagannathan (1991) show that solution to portfolio choice problem can be expressed in terms of Euler equation (Dittmar, 2002)

$$E_t(R^i_{t+1}M_{t+1} \mid \Omega_t) = 1, \tag{12}$$

where R^i_{t+1} is return of asset i, and M_{t+1} is the pricing kernel, also called stochastic discount factor. If we assume the investor is deciding between a pool of risky assets, which yields the return on aggregate wealth R^w_{t+1}, and the risk-free asset, the Euler equation is

$$E_t(R^w_{t+1}M_{t+1} \mid \Omega_t) = 1. \tag{13}$$

In what follows, we build on the approach of Maheu et al. (2013) and Chabi-Yo (2012). Our model is derived without explicitly assuming any form of utility function. We begin by taking a Taylor expansion of an unspecified utility function $U(W_{t+1})$ up to the fourth order, expanding the series up to the fourth order is justified in Dittmar (2002). Aggregate wealth in time $t+1$ is determined in a typical manner $W_{t+1} = W_t(1 + R^w_{t+1})$, where R^w_{t+1} is net return on aggregate wealth. $U(W_{t+1})$ is expanded around $W_t(1 + C_t)$, where C_t is an arbitrary return

$$\begin{aligned} U(W_{t+1}) &\approx \sum_{n=0}^{4} \frac{U^{(n)}(W_t(1+C_t))}{n!}(W_{t+1} - W_t(1+C_t))^n \\ &= \sum_{n=0}^{4} \frac{U^{(n)}(W_t(1+C_t))}{n!}(W_t(R^w_{t+1} - C_t))^n. \end{aligned} \tag{14}$$

Without loss of generality we can assume the initial wealth is equal to 1. Taking a derivative of the sum in Equation (14) yields

$$U'(W_{t+1}) \approx \sum_{n=0}^{3} \frac{U^{(n+1)}(1+C_t)}{n!}(R^w_{t+1} - C_t)^n. \tag{15}$$

Note that we can interpret $U'(W_{t+1})$ as the marginal utility of wealth at time $t+1$. The pricing kernel represents investors' discounting between periods, it corresponds to changes in marginal utility given the time period in which wealth is received. The pricing kernel, $M_{t+1} \equiv U'(W_{t+1})/U'(W_t)$, can be approximated as

$$\begin{aligned} M_{t+1} &\approx \sum_{n=0}^{3} \frac{U^{(n+1)}(1+C_t)}{U'(1)n!}(R^w_{t+1} - C_t)^n \\ &= g_{0,t} + g_{1,t}(R^w_{t+1} - C_t) + g_{2,t}(R^w_{t+1} - C_t)^2 + g_{3,t}(R^w_{t+1} - C_t)^3, \end{aligned} \tag{16}$$

where $g_{n,t} = [U^{(n+1)}(1+C_t)/U'(1)][1/n!] = [U^{(n+1)}(1+C_t)/U'(1+C_t)n!][U'(1+C_t)/U'(1)]$. Solution to investors' portfolio choice between a risk-free asset and a pool of risky assets can be expressed in terms of R^w_{t+1} and M_{t+1} as; $E_t[(1 + R^w_{t+1})M_{t+1} \mid \Omega_t] = 1$. Using the pricing kernel from Equation (16) and rearranging, we obtain

$$E_t(R^w_{t+1}) - R^f_t = \theta_{1,t} cov(R^w_{t+1}, R^w_{t+1} - C_t) + \theta_{2,t} cov(R^w_{t+1}, (R^w_{t+1} - C_t)^2) + \theta_{3,t} cov(R^w_{t+1}, (R^w_{t+1} - C_t)^3), \tag{17}$$

where R^f_t is the risk-free rate, and $\theta_{n,t} = -g_{n,t} R^f_t$. At this point, it is convenient to specify the expansion point. A natural choice would be $C_t = 0$ (e.g., Harvey and Siddique, 2000; Dittmar, 2002), an alternative approach is to set $C_t = E_t(R^w_{t+1})$. We employ the latter specification also used by Chabi-Yo (2012) or Maheu et al. (2013). Chabi-Yo (2012) shows that this approach is equivalent to small noise expansion, if we write

$$R^w_{t+1} - E_t(R^w_{t+1}) = \epsilon Y_{t+1}, \tag{18}$$

then driving ϵ towards zero causes R^w_{t+1} to approach $E_t(R^w_{t+1})$. Let us simplify notation by denoting for each $i = \{1, ..., n\}$

$$\begin{aligned} R^{i,e}_{t+1} &= R^i_{t+1} - R^f_t, \\ R_{t+1} &= R^w_{t+1} - R^f_t, \end{aligned} \tag{19}$$

$$\epsilon_{t+1} = R^w_{t+1} - E_t(R^w_{t+1}) = (R^w_{t+1} - R^f_t) - (E_t(R^w_{t+1}) - R^f_t) = R_{t+1} - E_t(R_{t+1}). \tag{20}$$

Thus far we have assumed that investors' decision making depends solely on the aggregate representation of returns; R^w_t. However, as noted in the previous section, investment horizons are heterogeneous, and returns consist of components operating at different frequencies. Hence, entering the aggregate representation R^w_t into the model comes at a cost of accuracy. We are able to decompose returns to the individual components with different levels of persistence

$$R_{t+1} \equiv \sum_{j=1}^{N} R^{(j)}_{t+1} + R^{(\infty)}_{t+1} = R^{(short)}_{t+1} + R^{(long)}_{t+1}, \tag{21}$$

where $R^{(short)}_{t+1} = \sum_{j=1}^{N} R^{(j)}_{t+1}, R^{(long)}_{t+1} = R^{(\infty)}_{t+1} = R^{(n>N)}_{t+1}$, and choice of N is restricted by the economic meaning of short-term and long-term fluctuations [12]. Substituting the decomposition from Equation (21 to Equation (20), we obtain

$$\epsilon_{t+1} = [R^{(short)}_{t+1} - E_t(R^{(short)}_{t+1})] + [R^{(long)}_{t+1} - E_t(R^{(long)}_{t+1})] = \epsilon^{(s)}_{t+1} + \epsilon^{(l)}_{t+1}. \tag{22}$$

Let us assume, that for each i, investors invest their whole wealth into asset i, hence $R^{i,e}_{t+1}$ and R_{t+1} can be treated as interchangeable in the equation below. This allows us to express risk premium of asset i

$$\begin{aligned} E_t(R^{i,e}_{t+1}) =& \theta_{1,t} cov_t(R^{i,e}_{t+1}, \epsilon_{t+1}) + \theta_{2,t} cov_t(R^{i,e}_{t+1}, \epsilon^2_{t+1}) + \theta_{3,t} cov_t(R^{i,e}_{t+1}, \epsilon^3_{t+1}) \\ =& \theta_{1,t} cov_t(R^{i,e}_{t+1}, \epsilon^{(s)}_{t+1} + \epsilon^{(l)}_{t+1}) + \theta_{2,t} cov_t(R^{i,e}_{t+1}, (\epsilon^{(s)}_{t+1} + \epsilon^{(l)}_{t+1})^2) + \\ & \theta_{3,t} cov_t(R^{i,e}_{t+1}, (\epsilon^{(s)}_{t+1} + \epsilon^{(l)}_{t+1})^3) \\ =& \theta_{1,t} cov_t(R_{t+1}, \epsilon^{(s)}_{t+1} + \epsilon^{(l)}_{t+1}) + \theta_{2,t} cov_t(R_{t+1}, (\epsilon^{(s)}_{t+1} + \epsilon^{(l)}_{t+1})^2) + \\ & \theta_{3,t} cov_t(R_{t+1}, (\epsilon^{(s)}_{t+1} + \epsilon^{(l)}_{t+1})^3) \end{aligned} \tag{23}$$

[12] Note that due to the equivalence in Equation (21, the decomposition is not restricted to two horizons. In fact, we are able to construct components from arbitrary number of horizons by splitting the sum in intermediate points.

Returns inside the covariances in Equation (23) can be decomposed using $R_{t+1} = R^{(s)}_{t+1} + R^{(l)}_{t+1}$. Given orthogonality of components at individual scales, i.e. $(\epsilon^{(s)}_{t+1})^n(\epsilon^{(l)}_{t+1})^n = 0$ for $n = \{1, 2\}$, we can rewrite Equation (23) as sum of covariances of elements operating at matching horizons

$$
\begin{aligned}
E_t(R^{i,e}_{t+1}) =& \sum_{n=1}^{3} \theta_{n,t} cov_t(R^{(s)}_{t+1} + R^{(l)}_{t+1}, (\epsilon^{(s)}_{t+1})^n + (\epsilon^{(l)}_{t+1})^n) \\
=& \theta^{(s)}_{1,t} cov_t(R^{(s)}_{t+1}, \epsilon^{(s)}_{t+1}) + \theta^{(l)}_{1,t} cov_t(R^{(l)}_{t+1}, \epsilon^{(l)}_{t+1}) + \theta^{(s)}_{2,t} cov_t(R^{(s)}_{t+1}, (\epsilon^{(s)}_{t+1})^2) + \\
& \theta^{(l)}_{2,t} cov_t(R^{(l)}_{t+1}, (\epsilon^{(l)}_{t+1})^2) + \theta^{(s)}_{3,t} cov_t(R^{(s)}_{t+1}, (\epsilon^{(s)}_{t+1})^3) + \theta^{(l)}_{3,t} cov_t(R^{(l)}_{t+1}, (\epsilon^{(l)}_{t+1})^3),
\end{aligned}
\tag{24}
$$

where the second equality stems from assumption that $(\epsilon^{(s)}_{t+1})^n R^{(l)}_{t+1} = 0, (\epsilon^{(l)}_{t+1})^n R^{(s)}_{t+1} = 0$, for $n = \{1, 2, 3\}$. Note that $cov_t(R_{t+1}, \epsilon^n_{t+1}) = E(\epsilon^{n+1}_{t+1})$ for $n = \{1, 2, 3\}$, hence;

$$
\begin{aligned}
E_t(R^{i,e}_{t+1}) =& \theta^{(s)}_{1,t} var_t(\epsilon^{(s)}_{t+1}) + \theta^{(l)}_{1,t} var_t(\epsilon^{(l)}_{t+1}) + \theta^{(s)}_{2,t} E_t[(\epsilon^{(s)}_{t+1})^3] + \theta^{(l)}_{2,t} E_t[(\epsilon^{(l)}_{t+1})^3] + \\
& \theta^{(s)}_{3,t} E_t[(\epsilon^{(s)}_{t+1})^4] + \theta^{(l)}_{3,t} E_t[(\epsilon^{(l)}_{t+1})^4],
\end{aligned}
\tag{25}
$$

where

$$\theta^{(s)}_{1,t} = -\frac{U^{(2)}(1 + C_t)}{U'(1 + C_t)} \frac{U'(1 + C_t)R^f_t}{U'(1)} w^{(s)}_{1,t}, \tag{26}$$

$$\theta^{(l)}_{1,t} = -\frac{U^{(2)}(1 + C_t)}{U'(1 + C_t)} \frac{U'(1 + C_t)R^f_t}{U'(1)} w^{(l)}_{1,t}, \tag{27}$$

$$\theta^{(s)}_{2,t} = -\frac{U^{(3)}(1 + C_t)}{U'(1 + C_t)2!} \frac{U'(1 + C_t)R^f_t}{U'(1)} w^{(s)}_{2,t}, \tag{28}$$

$$\theta^{(l)}_{2,t} = -\frac{U^{(3)}(1 + C_t)}{U'(1 + C_t)2!} \frac{U'(1 + C_t)R^f_t}{U'(1)} w^{(l)}_{2,t}, \tag{29}$$

$$\theta^{(s)}_{3,t} = -\frac{U^{(4)}(1 + C_t)}{U'(1 + C_t)3!} \frac{U'(1 + C_t)R^f_t}{U'(1)} w^{(s)}_{3,t}, \tag{30}$$

$$\theta^{(l)}_{3,t} = -\frac{U^{(4)}(1 + C_t)}{U'(1 + C_t)3!} \frac{U'(1 + C_t)R^f_t}{U'(1)} w^{(l)}_{3,t}, \tag{31}$$

where $w^{(j)}_{i,t}$; $i = \{1, 2, 3\}$, and $j = \{s, l\}$; are the spectral weights. For sake of clarity, we define $v^{(s)}_t = var_t(\epsilon^{(s)}_{t+1})$; $v^{(l)}_t = var_t(\epsilon^{(l)}_{t+1})$; $s^{(s)}_t = E_t[(\epsilon^{(s)}_{t+1})^3]/(v^{(s)}_t)^{3/2}$; $s^{(l)}_t = E_t[(\epsilon^{(l)}_{t+1})^3]/(v^{(l)}_t)^{3/2}$; $k^{(s)}_t = E_t[(\epsilon^{(s)}_{t+1})^4]/(v^{(s)}_t)^2$; $k^{(l)}_t = E_t[(\epsilon^{(l)}_{t+1})^4]/(v^{(l)}_t)^2$. Then Equation (25) becomes

$$E_t(R^{i,e}_{t+1}) = \beta^{(s)}_t v^{(s)}_t + \beta^{(l)}_t v^{(l)}_t + \gamma^{(s)}_t s^{(s)}_t + \gamma^{(l)}_t s^{(l)}_t + \mathcal{K}^{(s)}_t k^{(s)}_t + \mathcal{K}^{(l)}_t k^{(l)}_t, \tag{32}$$

where

$$
\begin{aligned}
\beta^{(s)}_t &= \theta^{(s)}_{1,t}, \\
\beta^{(l)}_t &= \theta^{(l)}_{1,t},
\end{aligned}
\tag{33}
$$

$$
\begin{aligned}
\gamma_t^{(s)} &= \theta_{2,t}^{(s)} (v_t^{(s)})^{3/2}, \\
\gamma_t^{(l)} &= \theta_{2,t}^{(l)} (v_t^{(l)})^{3/2},
\end{aligned} \tag{34}
$$

$$
\begin{aligned}
\mathcal{K}_t^{(s)} &= \theta_{3,t}^{(s)} (v_t^{(s)})^{2}, \\
\mathcal{K}_t^{(l)} &= \theta_{3,t}^{(l)} (v_t^{(l)})^{2}.
\end{aligned} \tag{35}
$$

A.2 Euler equation

Covariance of variables X and Y can be decomposed as, $cov(X,Y) = E(XY) - E(X)E(Y)$, hence $E(XY) = cov(X,Y) + E(X)E(Y)$. Following this decomposition, Equation (13) can be rewritten as

$$
\begin{aligned}
cov_t(R^w_{t+1} M_{t+1}) + E_t(R^w_{t+1}) E_t(M_{t+1}) &= 1 \\
E_t(R^w_{t+1}) E_t(M_{t+1}) &= 1 - cov_t(R^w_{t+1} M_{t+1}) \\
E_t(R^w_{t+1}) &= \frac{1}{E_t(M_{t+1})} - \frac{cov_t(R^w_{t+1} M_{t+1})}{E_t(M_{t+1})}.
\end{aligned} \tag{36}
$$

Recall that $1/E_t(M_{t+1}) = R^f_t$, then Equation (36) becomes

$$
E_t(R^w_{t+1}) - R^f_t = -R^f_t \, cov_t(R^w_{t+1} M_{t+1}), \tag{37}
$$

and we have

$$
M_{t+1} \approx g_{0,t} + g_{1,t}(R^w_{t+1} - C_t) + g_{2,t}(R^w_{t+1} - C_t)^2 + g_{3,t}(R^w_{t+1} - C_t)^3. \tag{38}
$$

Substituting Equation (38) into Equation (37) yields

$$
\begin{aligned}
E_t(R^w_{t+1}) - R^f_t =& \theta_{1,t} cov(R^w_{t+1}, R^w_{t+1} - C_t) + \theta_{2,t} cov(R^w_{t+1}, (R^w_{t+1} - C_t)^2) \\
&+ \theta_{3,t} cov(R^w_{t+1}, (R^w_{t+1} - C_t)^3),
\end{aligned} \tag{39}
$$

$\theta_{n,t} = -g_{n,t} R^f_t$.

A.3 Orthogonality

Breaking down Equation (23, we get

$$
\begin{aligned}
E_t(R^{i,e}_{t+1}) =& \theta_{1,t} cov_t(R_{t+1}, \epsilon^{(s)}_{t+1} + \epsilon^{(l)}_{t+1}) + \theta_{2,t} cov_t(R_{t+1}, (\epsilon^{(s)}_{t+1} + \epsilon^{(l)}_{t+1})^2) \\
&+ \theta_{3,t} cov_t(R_{t+1}, (\epsilon^{(s)}_{t+1} + \epsilon^{(l)}_{t+1})^3) \\
=& \theta_{1,t} cov_t(R_{t+1}, \epsilon^{(s)}_{t+1} + \epsilon^{(l)}_{t+1}) + \theta_{2,t} cov_t(R_{t+1}, (\epsilon^{(s)}_{t+1})^2 + (\epsilon^{(l)}_{t+1})^2 + 2\epsilon^{(s)}_{t+1}\epsilon^{(l)}_{t+1}) \\
&+ \theta_{3,t} cov_t(R_{t+1}, (\epsilon^{(s)}_{t+1})^3 + (\epsilon^{(l)}_{t+1})^3 + 3(\epsilon^{(s)}_{t+1})^2 \epsilon^{(l)}_{t+1} + 3\epsilon^{(s)}_{t+1} (\epsilon^{(l)}_{t+1})^2).
\end{aligned} \tag{40}
$$

Assuming orthogonality of elements from different scales, i.e. $(\epsilon^{(s)}_{t+1})^n (\epsilon^{(l)}_{t+1})^n = 0$ for $n = \{1, 2\}$, eliminates the components displayed in red, hence we can write

$$E_t(R^{i,e}_{t+1}) = \sum_{n=1}^{3} \theta_{n,t} cov_t(R_{t+1}, (\epsilon^{(s)}_{t+1})^n + (\epsilon^{(l)}_{t+1})^n). \tag{41}$$

In the next part, we use the decomposition $R_{t+1} = R^{(s)}_{t+1} + R^{(l)}_{t+1}$, and the fact that $cov(X + Y, U + Z) = cov(X, U) + cov(X, Z) + cov(Y, U) + cov(Y, Z)$.

$$\begin{aligned} E_t(R^{i,e}_{t+1}) = & \sum_{n=1}^{3} \theta_{n,t} cov_t(R^{(s)}_{t+1} + R^{(l)}_{t+1}, (\epsilon^{(s)}_{t+1})^n + (\epsilon^{(l)}_{t+1})^n) \\ & = \theta^{(s)}_{1,t} cov_t(R^{(s)}_{t+1}, \epsilon^{(s)}_{t+1}) + \theta^{(l)}_{1,t} cov_t(R^{(l)}_{t+1}, \epsilon^{(l)}_{t+1}) + \theta^{(s)}_{2,t} cov_t(R^{(s)}_{t+1}, (\epsilon^{(s)}_{t+1})^2) \\ & + \theta^{(l)}_{2,t} cov_t(R^{(l)}_{t+1}, (\epsilon^{(l)}_{t+1})^2) + \theta^{(s)}_{3,t} cov_t(R^{(s)}_{t+1}, (\epsilon^{(s)}_{t+1})^3) + \theta^{(l)}_{3,t} cov_t(R^{(l)}_{t+1}, (\epsilon^{(l)}_{t+1})^3) \end{aligned} \tag{42}$$

assuming $(\epsilon^{(s)}_{t+1})^n R^{(l)}_{t+1} = 0, (\epsilon^{(l)}_{t+1})^n R^{(s)}_{t+1} = 0$, for $n = \{1, 2, 3\}$, again eliminates the red elements and leaves us with Equation (24). Validity of these assumptions was tested by simulations with supportive results.

A.4 Centralized moments

Excess returns as a function of centralized moments appear in Equation (25), which is derived from Equation (24). For sake of clarity, we leave out the horizon decomposition from the derivation, however, the extension is trivial. We have

$$E_t(R^{i,e}_{t+1}) = \theta_{1,t} cov_t(R_{t+1}, \epsilon_{t+1}) + \theta_{2,t} cov_t(R_{t+1}, \epsilon^2_{t+1}) + \theta_{3,t} cov_t(R_{t+1}, \epsilon^3_{t+1}). \tag{43}$$

Recall that $R_{t+1} - E_t(R_{t+1}) = \epsilon_{t+1}$, then

$$\begin{aligned} cov_t(R_{t+1}, \epsilon_{t+1}) = & E_t(R_{t+1}\epsilon_{t+1}) - E_t(R_{t+1})E_t(\epsilon_{t+1}) = E_t([\epsilon_{t+1} + E_t(R_{t+1})]\epsilon_{t+1}) \\ & - E_t(R_{t+1})E_t(\epsilon_{t+1}) \\ = & E_t(\epsilon^2_{t+1}) + E_t(\epsilon_{t+1}E_t(R_{t+1})) - E_t(R_{t+1})E_t(\epsilon_{t+1}) \\ = & E_t(\epsilon^2_{t+1}) = var_t(\epsilon_{t+1}), \end{aligned} \tag{44}$$

since $E_t(\epsilon_{t+1}E_t(R_{t+1})) = E_t(R_{t+1})E_t(\epsilon_{t+1})$, and $E_t(\epsilon_{t+1}) = 0$. Analogously, we can derive that

$$\begin{aligned} cov_t(R_{t+1}, \epsilon^2_{t+1}) &= E_t(\epsilon^3_{t+1}) \\ cov_t(R_{t+1}, \epsilon^3_{t+1}) &= E_t(\epsilon^4_{t+1}). \end{aligned} \tag{45}$$

Equation (43) then becomes

$$E_t(R^{i,e}_{t+1}) = \theta_{1,t} var_t(\epsilon_{t+1}) + \theta_{2,t} E_t(\epsilon^3_{t+1}) + \theta_{3,t} E_t(\epsilon^4_{t+1}). \tag{46}$$

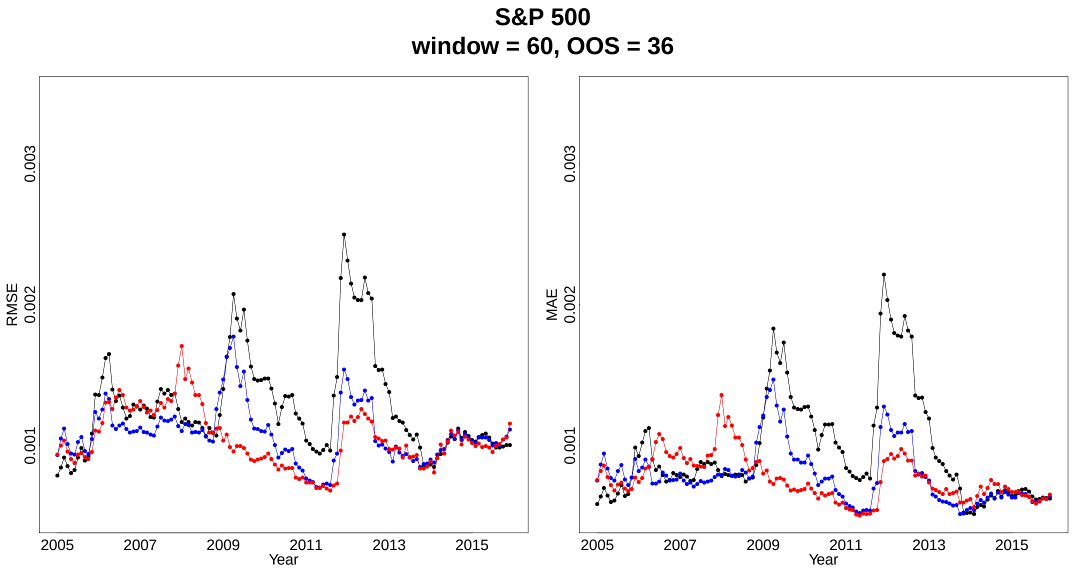

Figure 3: Out of sample prediction errors. RMSE in the left panel, MAE in the right panel S&P 500 data, window length 60 months, out-of-sample length 36 months.

B Additional figures and tables

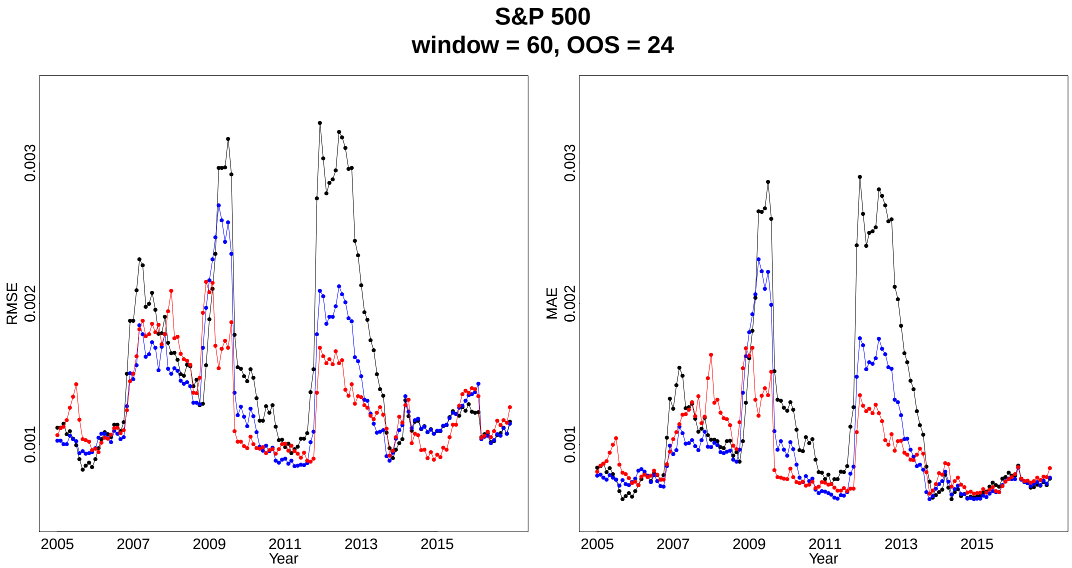

Figure 4: Out of sample prediction errors. RMSE in the left panel, MAE in the right panel S&P 500 data, window length 60 months, out-of-sample length 24 months.

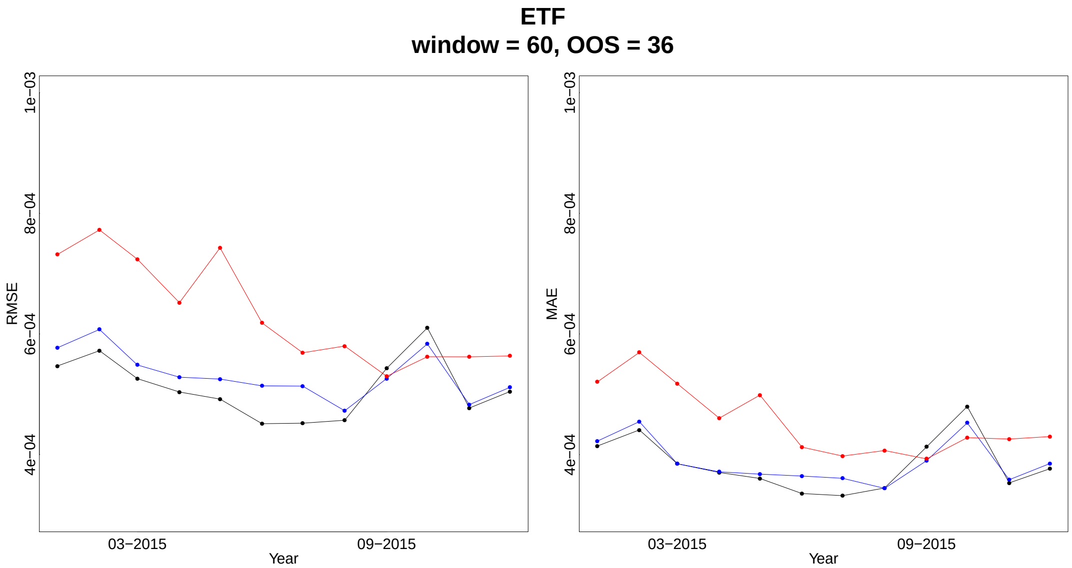

Figure 5: Out of sample prediction errors. RMSE in the left panel, MAE in the right panel ETF, window length 60 months, out-of-sample length 36 months.

Table 5: We report coefficients from the second stage Fama-Macbeth regression, i.e. Equation (7). First column displays results of CAPM model, second column results of FM-CAPM model, third column results Horizon-specific FM-CAPM. Monthly S&P 500 returns.

	Model			
	CAPM	Four Moment	Fama French	Horizon-specific
Constant	0.00878***	0.00893***	0.01020***	0.01116***
	(0.00087)	(0.00054)	(0.00087)	(0.00059)
Volatility Short			−0.00218**	−0.00122
			(0.00085)	(0.00078)
Volatility Medium			0.00369**	0.00374**
			(0.00166)	(0.00177)
Volatility Long			0.00992***	0.00793***
			(0.00176)	(0.00185)
Skewness Short			0.02660	0.02953
			(0.01819)	(0.01928)
Skewness Medium			0.07924***	0.10823***
			(0.01162)	(0.01064)
Skewness Long			−0.02162***	−0.01142**
			(0.00562)	(0.00567)
Kurtosis Short			−0.05903	−0.10202
			(0.15024)	(0.15753)
Kurtosis Medium			0.05434	0.14509**
			(0.05940)	(0.06077)
Kurtosis Long			−0.43810***	−0.33906***
			(0.07323)	(0.07664)
Market	−0.14411*		−0.07905	
	(0.07484)		(0.08740)	
Volatility		0.00730**		
		(0.00320)		
Skewness		0.15718***		
		(0.02777)		
Kurtosis		−0.15192		
		(0.22108)		
SMB			0.12014	
			(0.08077)	
HML			−0.69414***	
			(0.06455)	
Observations	443	443	443	443
R^2	0.00834	0.11347	0.38253	0.28935
Adjusted R^2	0.00609	0.10742	0.36530	0.27458
Note:			*p<0.1; **p<0.05; ***p<0.01	

Table 6: We report coefficients from the second stage Fama-Macbeth regression, i.e. Equation (7). First column displays results of CAPM model, second column results of FM-CAPM model, third column results Horizon-specific FM-CAPM. Monthly ETF returns.

	Model			
	CAPM	Four Moment	Fama French	Horizon-specific
Constant	0.00251***	0.00224***	0.00224***	0.00245***
	(0.00057)	(0.00061)	(0.00059)	(0.00055)
Volatility Short			−0.00154	−0.00328***
			(0.00133)	(0.00089)
Volatility Medium			0.00620***	0.00619***
			(0.00204)	(0.00178)
Volatility Long			0.00489*	0.00576**
			(0.00281)	(0.00272)
Skewness Short			0.03651**	0.02743
			(0.01737)	(0.01710)
Skewness Medium			0.04250***	0.03304***
			(0.01327)	(0.01187)
Skewness Long			0.00278	−0.00170
			(0.00805)	(0.00789)
Kurtosis Short			0.02938	−0.02068
			(0.16126)	(0.15776)
Kurtosis Medium			0.32718***	0.28592***
			(0.06465)	(0.06087)
Kurtosis Long			−0.04347	−0.09446
			(0.11157)	(0.10947)
Market	−0.10465**		0.00969	
	(0.04200)		(0.05035)	
Volatility		0.00200		
		(0.00417)		
Skewness		0.00243		
		(0.02548)		
Kurtosis		0.36524		
		(0.23318)		
SMB			0.19699**	
			(0.07582)	
HML			−0.04760	
			(0.06787)	
Observations	233	233	233	233
R^2	0.02617	0.08295	0.32299	0.30369
Adjusted R^2	0.02196	0.07093	0.28606	0.27559
Note:				*p<0.1; **p<0.05; ***p<0.01

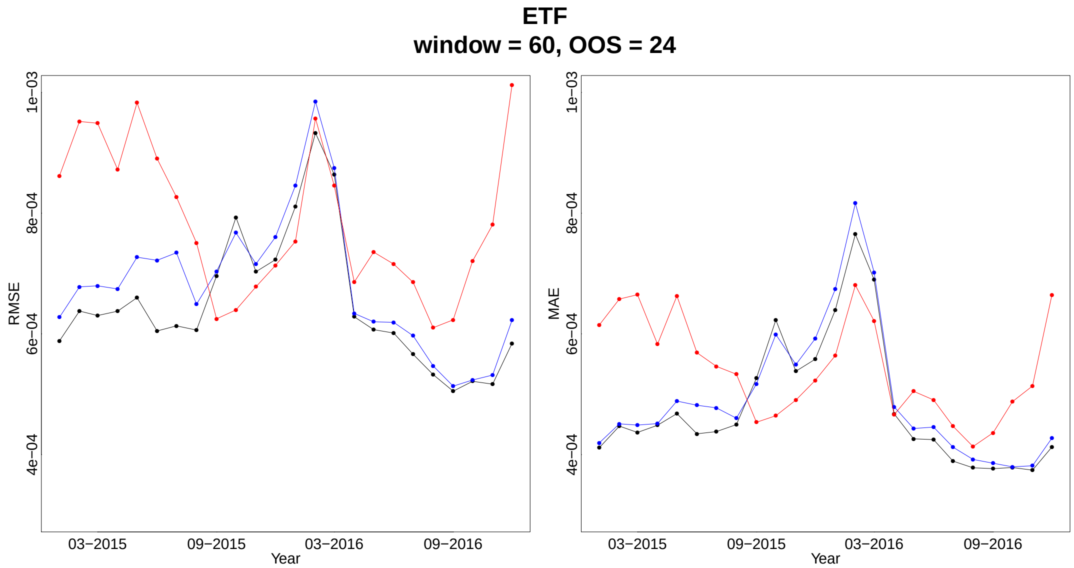

Figure 6: Out of sample prediction errors. RMSE in the left panel, MAE in the right panel ETF data, window length 60 months, out-of-sample length 24 months.

Table 7: S&P 500, short-run

	Model		
	CAPM	Four Moment	Horizon-specific
Constant	0.00027***	0.00076***	0.00088***
	(0.00003)	(0.00008)	(0.00008)
Volatility	0.00159***	0.00113***	
	(0.00013)	(0.00014)	
Skewness		−0.08197***	
		(0.01205)	
Kurtosis		0.16501	
		(0.18209)	
Volatility Short			0.00034***
			(0.00007)
Skewness Short			−0.11598***
			(0.01014)
Kurtosis Short			0.27574*
			(0.16361)
Observations	456	456	456
R^2	0.25802	0.32998	0.26745
Adjusted R^2	0.25639	0.32554	0.26259
Note:			*p<0.1; **p<0.05; ***p<0.01

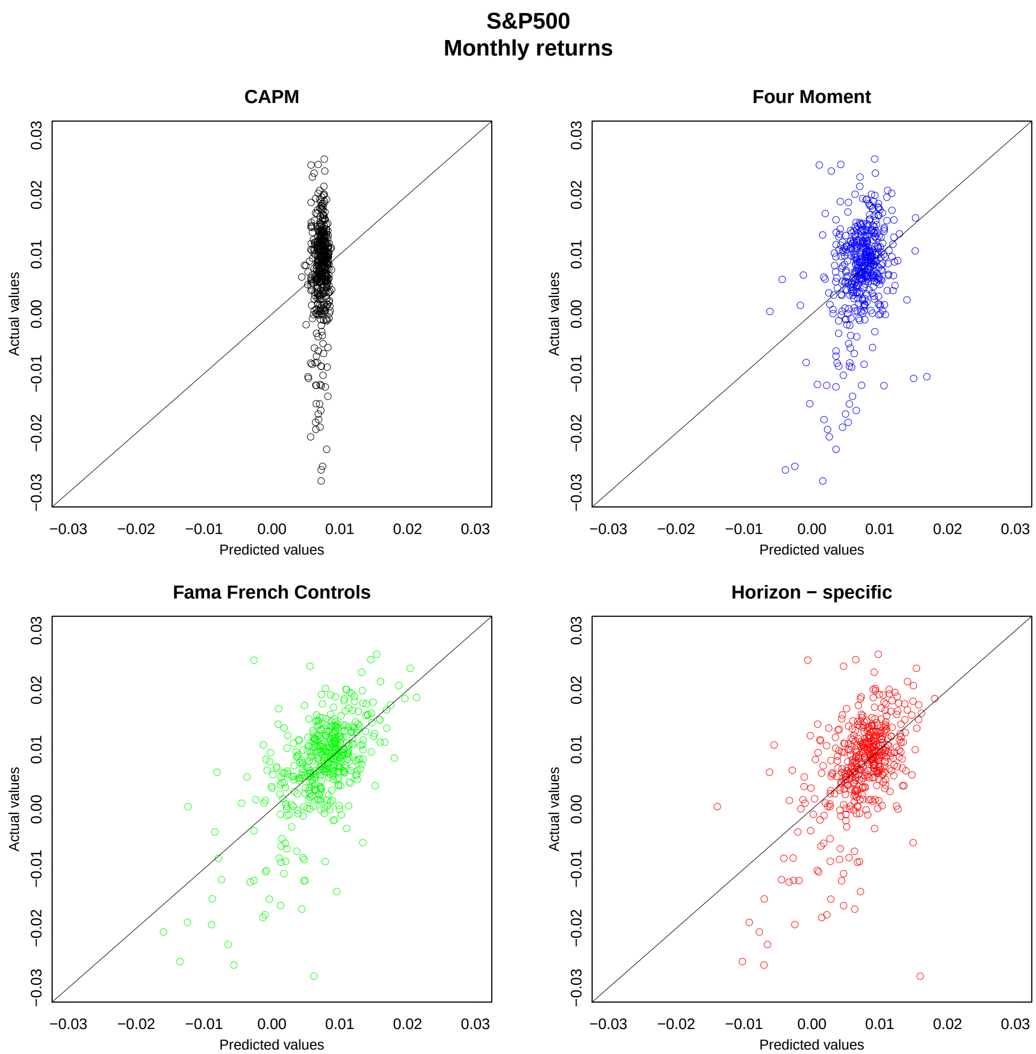

Figure 7: Monthly returns predicted by the respective models plotted against actual returns S&P 500 stocks. CAPM in the top left corner (black dots), FM-CAPM in the top right corner (blue dots), Horizon-specific FM-CAPM (HS FM CAPM) in the bottom left corner (red dots).

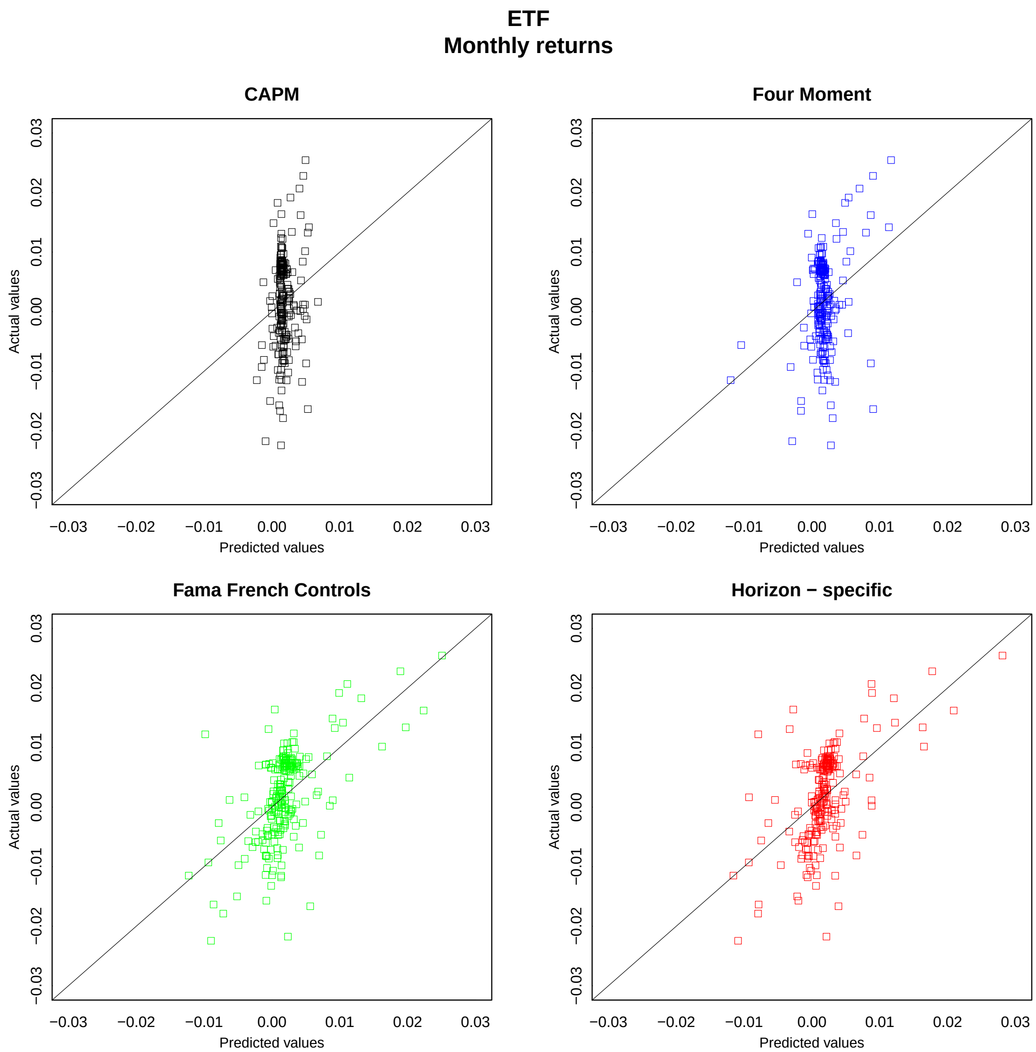

Figure 8: Monthly returns predicted by the respective models plotted against actual returns ETFs. CAPM in the top left corner (black dots), FM-CAPM in the top right corner (blue dots), Horizon-specific FM-CAPM (HS FM CAPM) in the bottom left corner (red dots).

Table 8: S&P 500, medium-run

	Model		
	CAPM	Four Moment	Horizon-specific
Constant	0.00027***	0.00076***	0.00062***
	(0.00003)	(0.00008)	(0.00006)
Volatility	0.00159***	0.00113***	
	(0.00013)	(0.00014)	
Skewness		−0.08197***	
		(0.01205)	
Kurtosis		0.16501	
		(0.18209)	
Volatility Medium			0.00037***
			(0.00005)
Skewness Medium			−0.02526***
			(0.00466)
Kurtosis Medium			0.05624
			(0.03716)
Observations	456	456	456
R^2	0.25802	0.32998	0.28705
Adjusted R^2	0.25639	0.32554	0.28232
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 9: S&P 500, long-run

	Model		
	CAPM	Four Moment	Horizon-specific
Constant	0.00027***	0.00076***	0.00043***
	(0.00003)	(0.00008)	(0.00005)
Volatility	0.00159***	0.00113***	
	(0.00013)	(0.00014)	
Skewness		−0.08197***	
		(0.01205)	
Kurtosis		0.16501	
		(0.18209)	
Volatility Long			0.00136***
			(0.00013)
Skewness Long			−0.03648***
			(0.00772)
Kurtosis Long			−0.21632***
			(0.03656)
Observations	456	456	456
R^2	0.25802	0.32998	0.30393
Adjusted R^2	0.25639	0.32554	0.29931
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 10: ETF, short-run

	Model		
	CAPM	Four Moment	Horizon-specific
Constant	−0.00020***	−0.00021***	−0.00023***
	(0.00003)	(0.00003)	(0.00003)
Volatility	−0.00025***	−0.00030*	
	(0.00005)	(0.00018)	
Skewness		−0.00372	
		(0.01549)	
Kurtosis		0.29211	
		(0.30202)	
Volatility Short			−0.00019**
			(0.00008)
Skewness Short			−0.01822
			(0.01170)
Kurtosis Short			0.64293**
			(0.27589)
Observations	233	233	233
R^2	0.08483	0.08836	0.12973
Adjusted R^2	0.08087	0.07642	0.11833
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 11: ETF, medium-run

	Model		
	CAPM	Four Moment	Horizon-specific
Constant	−0.00020***	−0.00021***	−0.00019***
	(0.00003)	(0.00003)	(0.00003)
Volatility	−0.00025***	−0.00030*	
	(0.00005)	(0.00018)	
Skewness		−0.00372	
		(0.01549)	
Kurtosis		0.29211	
		(0.30202)	
Volatility Medium			−0.00012
			(0.00008)
Skewness Medium			0.00370
			(0.00705)
Kurtosis Medium			−0.16554**
			(0.06810)
Observations	233	233	233
R^2	0.08483	0.08836	0.10983
Adjusted R^2	0.08087	0.07642	0.09817
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 12: ETF, long-run

	Model		
	CAPM	Four Moment	Horizon-specific
Constant	-0.00020^{***}	-0.00021^{***}	-0.00016^{***}
	(0.00003)	(0.00003)	(0.00003)
Volatility	-0.00025^{***}	-0.00030^{*}	
	(0.00005)	(0.00018)	
Skewness		−0.00372	
		(0.01549)	
Kurtosis		0.29211	
		(0.30202)	
Volatility Long			-0.00035^{*}
			(0.00019)
Skewness Long			−0.01122
			(0.01278)
Kurtosis Long			−0.04193
			(0.06341)
Observations	233	233	233
R^2	0.08483	0.08836	0.04977
Adjusted R^2	0.08087	0.07642	0.03732
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 13: S&P 500, volatility

	Model		
	CAPM	Four Moment	Horizon-specific
Constant	0.00027***	0.00076***	0.00034***
	(0.00003)	(0.00008)	(0.00004)
Volatility	0.00159***	0.00113***	
	(0.00013)	(0.00014)	
Skewness		−0.08197***	
		(0.01205)	
Kurtosis		0.16501	
		(0.18209)	
Volatility Short			0.00007
			(0.00007)
Volatility Medium			0.00035***
			(0.00005)
Volatility Long			0.00108***
			(0.00014)
Observations	456	456	456
R^2	0.25802	0.32998	0.29416
Adjusted R^2	0.25639	0.32554	0.28948
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 14: S&P 500, skewness

	Model		
	CAPM	Four Moment	Horizon-specific
Constant	0.00027***	0.00076***	0.00092***
	(0.00003)	(0.00008)	(0.00008)
Volatility	0.00159***	0.00113***	
	(0.00013)	(0.00014)	
Skewness		−0.08197***	
		(0.01205)	
Kurtosis		0.16501	
		(0.18209)	
Skewness Short			−0.11436***
			(0.01397)
Skewness Medium			−0.00647
			(0.00809)
Skewness Long			−0.02621***
			(0.00826)
Observations	456	456	456
R^2	0.25802	0.32998	0.24668
Adjusted R^2	0.25639	0.32554	0.24168

Note: *p<0.1; **p<0.05; ***p<0.01

Table 15: S&P 500, kurtosis

	Model		
	CAPM	Four Moment	Horizon-specific
Constant	0.00027***	0.00076***	0.00033***
	(0.00003)	(0.00008)	(0.00005)
Volatility	0.00159***	0.00113***	
	(0.00013)	(0.00014)	
Skewness		−0.08197***	
		(0.01205)	
Kurtosis		0.16501	
		(0.18209)	
Kurtosis Short			−0.46458***
			(0.11115)
Kurtosis Medium			−0.09444***
			(0.03144)
Kurtosis Long			−0.28201***
			(0.02879)
Observations	456	456	456
R^2	0.25802	0.32998	0.26357
Adjusted R^2	0.25639	0.32554	0.25868
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 16: ETF, volatility

	Model		
	CAPM	Four Moment	Horizon-specific
Constant	−0.00020***	−0.00021***	−0.00022***
	(0.00003)	(0.00003)	(0.00003)
Volatility	−0.00025***	−0.00030*	
	(0.00005)	(0.00018)	
Skewness		−0.00372	
		(0.01549)	
Kurtosis		0.29211	
		(0.30202)	
Volatility Short			−0.00024***
			(0.00007)
Volatility Medium			−0.000001
			(0.00008)
Volatility Long			0.00046**
			(0.00020)
Observations	233	233	233
R^2	0.08483	0.08836	0.12201
Adjusted R^2	0.08087	0.07642	0.11051

Note: *p<0.1; **p<0.05; ***p<0.01

Table 17: ETF, skewness

	Model		
	CAPM	Four Moment	Horizon-specific
Constant	−0.00020***	−0.00021***	−0.00019***
	(0.00003)	(0.00003)	(0.00003)
Volatility	−0.00025***	−0.00030*	
	(0.00005)	(0.00018)	
Skewness		−0.00372	
		(0.01549)	
Kurtosis		0.29211	
		(0.30202)	
Skewness Short			−0.00772
			(0.02165)
Skewness Medium			0.03200
			(0.02095)
Skewness Long			−0.04443***
			(0.01352)
Observations	233	233	233
R^2	0.08483	0.08836	0.12734
Adjusted R^2	0.08087	0.07642	0.11590
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 18: ETF, kurtosis

	Model		
	CAPM	Four Moment	Horizon-specific
Constant	-0.00020^{***}	-0.00021^{***}	-0.00017^{***}
	(0.00003)	(0.00003)	(0.00003)
Volatility	-0.00025^{***}	-0.00030^{*}	
	(0.00005)	(0.00018)	
Skewness		-0.00372	
		(0.01549)	
Kurtosis		0.29211	
		(0.30202)	
Kurtosis Short			0.87068^{***}
			(0.15349)
Kurtosis Medium			-0.26625^{***}
			(0.07481)
Kurtosis Long			-0.11944^{**}
			(0.05300)
Observations	233	233	233
R^2	0.08483	0.08836	0.14523
Adjusted R^2	0.08087	0.07642	0.13403
Note:			*p<0.1; **p<0.05; ***p<0.01