Are Conditional Factors Priced? Characterizing Risk Premia of Conditional Systematic Risk Factors with Staggered Regressions

F. Douglas Foster[1], Anna Lopatnikova*[1], and Stephen Satchell[2]

[1]Discipline of Finance, The University of Sydney, NSW 2006, Australia
[2]Trinity College, University of Cambridge, CB2 1TQ, United Kingdom

February 26, 2021

Abstract

We introduce the staggered Fama-MacBeth regression method and use it to evaluate economic significance of popular conditional systematic risk factors. Prior literature demonstrated that exposure to these factors is rewarded in the cross-section of stock returns. Much of the evidence comes from contemporaneous regressions; however, in predictive regressions, the evidence tends to disappear. The proposed staggered regression method combines the benefits of contemporaneous and predictive regressions while eliminating critical shortcomings. Using the method, we confirm the economic significance of downside risk, β^-, relative downside risk, $\beta^- - \beta^+$, and coskewness, but not of exceedance correlations and related measures of asymmetric dependence.

Keywords: staggered regression; systematic risk factors; asymmetric dependence; downside risk; conditional skewness; asset pricing

*Corresponding author. *E-mail:* anna.lopatnikova@sydney.edu.au
Author *E-mail* addresses: douglas.foster@sydney.edu.au, anna.lopatnikova@sydney.edu.au, ses11@cam.ac.uk
We are grateful for the insightful comments of Jamie Alcock, Petra Andrlikova, Guanglian Hu, Andrew Patton, Richard Philip, Buhui Qiu, Tom Smith, Xue Wang (discussant), and participants at the University of Sydney seminar, the Australasian Finance and Banking Conference, and the University of Sydney-Zhejiang University symposium. This research was supported fully by the Australian Government through the Australian Research Council's Discovery Projects funding scheme (project DP180104120). The views expressed herein are those of the authors and are not necessarily those of the Australian Government or Australian Research Council.

1 Introduction

When the joint distribution of stock returns deviates from multivariate normal and investors possess a plausible utility function, the CAPM (Sharpe, 1964; Lintner, 1965; Black, 1972) does not hold and risk factors outside the linear market model may explain the cross section of stock returns. Candidate systematic risk factors to supplement the linear market model have included measures of stock return deviations from multivariate normality: conditional and unconditional skewness and kurtosis (Kraus and Litzenberger, 1976; Harvey and Siddique, 2000; Dittmar, 2002; Albuquerque, 2012; Conrad, Dittmar, and Ghysels, 2013), downside and upside risk(Bawa and Lindenberg, 1977; Ang, Chen, and Xing, 2006), and asymmetric tail risk (Longin and Solnik, 2001; Ang and Chen, 2002; Hong, Tu, and Zhou, 2006; Alcock and Hatherley, 2016). Unlike many fundamental variables, such as the book-to-market ratio or profitability (see e.g., Hou, Xue, and Zhang, 2015, and references therein), that have been found to explain the cross-section of stock returns, conditional systematic risk factors provide *direct* measures of risk, which gives them greater theoretical appeal.

A shortcoming of systematic risk factors is that they are not directly observable. As a result, tests of economic significance of these factors face econometric difficulties. One of the most famous examples is the economic significance of CAPM beta. Even though classic studies, such as Gibbons (1982), found statistically significant premia associated with CAPM beta using tests where CAPM betas and average returns are estimated *contemporaneously*, tests where lagged CAPM betas are used *predictively* (e.g. Fama and French, 1992) show no statistically significant premium associated with CAPM beta. Downside risk, coskewness, and asymmetric tail risk have faced similar challenges: their economic significance has been established using contemporaneous regressions (Harvey and Siddique, 2000; Ang, Chen, and Xing, 2006; Alcock and Hatherley, 2016; Jiang, Wu, and Zhou, 2018) – where the factor sensitivities and average returns were determined using data from the same time period; in predictive regressions, the estimated risk premia suffer a dramatic decline and their statistical significance disappears.

Prior literature has relied on contemporaneous regressions to assess the economic significance of systematic factors because predictive regressions can suffer from "false negatives" – instances where underlying risk premia can be masked by factor time-variation and errors in estimation of factor loadings.[1] This problem is particularly pronounced for conditional systematic factors: On the one hand, a long interval of historical return time series is required to estimate each factor loading at a single point in time, in order to capture deviations from multivariate normality with sufficient accuracy. On the other hand, a long estimation interval means that relevant factor innovations are averaged out, if exposure to the factor is time-varying – as Fama and French (1997), Ghysels (1998), and Lewellen and Nagel (2006), among others, have demonstrated to be the case for CAPM beta. When predictive regressions are not able to capture underlying risk premia, contemporaneous methods can help. For example, Ghysels (1998) argues, in-sample regression can be an informative test of a model, especially if factors have structural breaks. Motivated by this argument, Harvey and Siddique (2000) establish economic significance of conditional coskewness using the full information maximum likelihood method, which uses all available data simultaneously to estimate factor loadings and average returns. Ang, Chen, and Xing (2006) report returns of portfolios formed on contemporaneous CAPM betas, downside betas and upside betas. Alcock and Hatherley (2016) use contemporaneous Fama-MacBeth regressions to study the economic significance of asymmetric correlations of stock returns. Jiang, Wu, and Zhou (2018) study asymmetry in stock co-movements using contemporaneous portfolio sorts.

But contemporaneous regressions can be subject to "false positives," because they, in effect, use the same data on both sides of the regression: To estimate return premia, realized average returns are regressed on estimated factor sensitivities that are, in turn, estimated using the same data series (as represented schematically in Figure 1, panel (a)). As a consequence, contemporaneous regressions can yield biased results from confounding effects, simultaneity, endogeneity, and reverse causality. When the results of contemporaneous and predictive regressions are in conflict – as is the case for all conditional systematic risk factors – it is not clear *a priori* whether it is because of the "false negatives" of predictive regressions or "false positives" of contemporaneous regressions.

In this paper, we focus on one of the most common methods used to test the economic signifance of conditional systematic risk factors, the Fama-MacBeth regression – a two-pass regression method

[1] Methodologies to reduce the impact of the errors-in-variables bias, such the correction proposed by Litzenberger and Ramaswamy (1979) and extended by Shanken (1992), and more recent works such as Gagliardini et al. (2016); Chordia et al. (2017); Kim and Skoulakis (2018), affirm risk premia associated by CAPM beta, but may not eliminate the problem for other conditional factors particularly in the presence of a time-varying factor.

first introduced by Black, Jensen, and Scholes (1972) and refined by Fama and MacBeth (1973). We focus on the modern incarnation of the Fama-MacBeth method, motivated by results by Lewellen and Nagel (2005), Ang and Chen (2005), and Ang, Chen, Xing (2006), where factor loadings are estimated using shorter estimation periods (4 to 12 months of daily data) and applied cross-sectionally to individual stocks rather than portfolios; this choice improves sensitivity of the method in the presence of time-varying factors despite the increased downward bias from errors in variables.

We propose a modification to the Fama-MacBeth regression technique to address the challenges of predictive and contemporaneous Fama-MacBeth regressions. We apply it to characterize the most popular conditional systematic risk factors and make inferences about their economic significance. Our contribution is three-fold:

Our first contribution is to introduce *staggered* regressions – a simple extension of the Fama-MacBeth regression. In the staggered method, factors loadings and average returns are estimated using alternating months (as demonstrated schematically in Figures 1 and 2), thus reducing the potential for "false positives" found in contemporaneous regressions. By bringing the factors and returns closer together in time, the method also results in fewer "false negatives" than do predictive regressions.[2] The method addresses many of the challenges researchers have faced with factor premium estimation: aging of factors in long-window estimates, errors-in-variables of short-window estimates, low power of portfolio sorts, and confounding factors in contemporaneous regressions. We demonstrate the statistical properties (bias and standard error) of staggered regressions using a parsimonious theoretical model with a time-varying factor. As a straightforward modification of the traditional Fama-MacBeth regression, the method provides estimation benefits at nearly no increase in estimation difficulty.[3] Inspiration for staggered-month estimation at the core of the method come from the instrumental-variables technique recently reported by Jegadeesh, Noh, Pukthuanthong, Roll, and Wang (2019), in which the authors use factors estimated on staggered

[2]A shortcoming of the staggered Fama-MacBeth regression method is that it is not predictive – it is not designed to forecast future returns based on lagged factor loadings. In effect, the staggered regression method is an improved version of the contemporaneous regression method, to be used when predictive regressions fail. The staggered regression, just like contemporaneous regression, provides information about *whether* a factor is priced, but it reduces the risk of "false positive" results. Once a factor has been demonstrated to have a risk premium, forecasting methods can be used to predict future returns based on the factor.

[3]Being so closely related to Fama-MacBeth regressions, the method is not a fully optimized econometric technique, as critiqued, for example, by Lewellen, Nagel, and Shanken (2010) or Kan, Robotti, and Shanken (2013).

months as instruments. (Staggered estimation has established predecessors, such as Ball, Brown, and Officer, 1976; Scholes and Williams, 1977; or Mankiw and Shapiro, 1986.)

Our second contribution is to characterize the most popular conditional systematic factors using the proposed staggered regression method and, in the process, extend prior studies by 15-18 years of more recent data. We demonstrate that downside risk remains a robust factor in explaining the cross-section of stock returns. In past studies, exposure to downside risk, β^-, has been associated with a material and statistically significant risk premium only in contemporaneous regressions; the risk premium was immaterial in predictive regressions. Staggered regressions confirm that the premium is not due to direct confounding factors and it is likely that investors exposed to downside risk have been rewarded by the market. Similarly, we find that coskewness also represents a priced factor and negative coskewness is associated with a statistically significant premium. We also find that, based on our analysis, it is not likely that investors have been rewarded for exposure to asymmetric tail dependence as measured by exceedance correlations. Metrics based on exceedance correlations lose their explanatory power outside of contemporaneous regressions.

As our third contribution, we demonstrate a mechanism by which exceedance correlations are strongly associated with an apparent risk premium in contemporaneous regressions. Using a market model with idiosyncratic jumps, we derive the sensitivity of exceedance correlations to the presence of jumps and crashes (i.e. negative jumps) in time series of stock returns. We confirm our analytical results in Monte-Carlo simulations. Exceedance correlations were applied initially to study tail dependence (e.g. Longin and Solnik, 2001) between stock market indices and were later used to measure asymmetric dependence of individual stock returns on market returns (e.g. Ang and Chen, 2002). Two studies, Cizeau, Potters, and Bouchaud (2001) and Campbell, Forbes, Koedijk, and Kofman (2008), have pointed out that caution is needed in applying exceedance correlations to studies of extreme tail dependence. We add to the understanding of exceedance correlations by considering their behavior during market moves of arbitrary magnitude. Because exceedance correlations are highly sensitive to deviations of stock returns from (multivariate) normality, we propose ways they can be re-purposed.

This paper is organized as follows: Section 2 introduces staggered regressions. Section 3 reports the results of our analysis of conditional systematic factors using staggered regressions. Section 4 compares econometric properties of staggered regressions with those of contemporaneous and predictive regressions. Section 5 concludes. Appendix C details how false positives can occur in contemporaneous regressions using exceedance correlations as an example.

2 Staggered Regressions

In this section, we introduce the staggered cross-sectional regression. Even though it is a simple modification of the traditional Fama-MacBeth method – as represented graphically in Figure 1 – the staggered regression addresses many of the challenges outlined in the introduction. The construction helps to find a better balance between the trade-offs inherent in Fama-MacBeth regression and achieve a number of desirable outcomes simultaneously: (1) retain the power and information contained in the time series of individual stock returns; (2) have sufficient data for estimation of factor loadings to reduce errors-in-variables, particularly for metrics conditional on market returns; (3) reduce the wash-out effect from the time lag between the factor estimate and the return estimate; (4) reduce the impact of confounding factors, direct endogeneity, and reverse causality. We provide a comparison of the econometric properties of staggered, contemporaneous, and predictive regressions in Section 4.

The idea is simple: During the first step of the Fama-MacBeth method, we split the estimation window into months of daily data and then stagger the months on which we estimate realize returns and the months on which we estimate the factors.[4] To promote the flow of causality from factors to returns and to reduce the potential impact of short-term reversals, we refine this staggered construction in three ways: First, we ensure that the last month used for estimation of factor loadings precedes the last month used for the calculation of average returns. Second, we introduce a skipped month, which follows each month used for return estimation. Effectively, we split the

[4] When the factors are time-varying and the errors of factor exposure and average return estimation are uncorrelated, the shortest possible staggered estimation subperiod provides the best results (as we show in Section 4) – it minimizes bias in factor premium estimation. We choose a month-long estimation subperiod, because it is sufficiently short to capture the underlying return-factor correlations and also sufficiently long to reduce correlations between errors in estimation of factor loadings and average returns. We explore how the factor premium estimates depend on the length of the staggered estimation subperiod empirically in Section 3.4.

estimation period into equal periods, each equal to a quarter of a year, and, in each quarter, the most recent month is used for return estimation, the middle month – for estimation of factor loadings, and the remaining month is skipped, as shown in Figure 2. The skipped month allows us to reduce the impact of reverse causality (particularly for shorter-lived factors). Third, we also skip the last two days of each month used for estimation of factor loadings, to reduce the impact of returns reversals, documented, for example, by Jegadeesh and Titman (1995) or Huang, Liu, Rhee, and Zhang (2009). Statistical significance in the model is assessed in the usual way, by estimating standard error of risk premia estimated in cross-sectional regressions, with adjustments for overlapping data with automatic lag selection, as proposed by Newey and West (1987).[5]

In this paper, we report the results of staggered regressions based on 18-month long estimation windows ($\tilde{T}_1 = 18$). Within each estimation window, six months of data are used to run factor exposure estimation and a *different* – but nearly contemporaneous – six months of data are used to estimate average returns. We compare the results of 18-month staggered regressions to the results of contemporaneous and predictive regressions, where factors and returns are estimated using 6-month long windows of continuous data. Our choice of 18-month-long staggered estimation and 6-month long contiguous estimation periods was informed by related empirical studies: Ang, Chen, and Xing (2006) use 12 contiguous months of daily data to study downside risk; Lewellen and Nagel (2006) use three months of daily data or weekly data to study conditional CAPM; Fama and MacBeth (1973) use five years of monthly data to test the CAPM. As a robustness check, we perform our analyses with shorter and longer estimation periods.

The method does not overcome all the shortcomings of the Fama-MacBeth method – or regression methods in general (e.g. data snooping per Lo and MacKinlay, 1990), but it achieves a better balance of benefits and shortcomings compared with contemporaneous and predictive regressions, as demonstrated using a theoretical model in Section 4. It provides an additional lens for regression analysis. As we show in Section 3.3, staggered regressions can resolve conflicting evidence from contemporaneous and predictive regressions and shed additional light on underlying structure and

[5] As a robustness check, we run standard error estimates with Newey and West (1987) corrections using specified lags, with the lag equal to T_1 for contemporaneous regressions, $2 \times T_1$ for staggered regressions, and $T_1 + T_2$ for predictive regressions. The resulting standard error estimates are sightly lower (but are within 7%) of standard errors estimated using automatic lag selection.

behavior of the analyzed factors.

3 Economic Significance of Conditional Systematic Risk Factors

In this section, we present our empirical analysis of economic significance of conditional systematic risk factors. We start with an overview of the factors, briefly outline the data set used in the study, and then discuss regression results. We conclude with a simulation analysis of downside risk and upside risk factors, β^- and β^+.

3.1 Conditional systematic factors

We focus on the most popular conditional systematic factors: downside risk (β^-), upside risk (β^+), coskewness, asymmetric tail risk (exceedance correlations ρ^{exc}, and a related measure of asymmetric dependence of stock returns on market returns J^{adj}). Figure 3 provides a "taxonomy" for conditional systematic risk factors.

To estimate factor loadings, we use *excess* stock and market returns – returns above the risk-free rate. In the remainder of this text, we use stock (market) returns as a shorthand for excess stock (market) return and drop the word "excess."

We start with the popular measure of downside risk, β^-, first proposed by Bawa and Lindenberg (1977). It is defined as the β of stock returns on market returns conditioned on negative market moves. The companion metric, upside beta, β^+, is conditioned on positive market moves:

$$\beta_i^- = \frac{\text{cov}(r_i, r_m | r_m < \mu_m)}{\text{var}(r_m | r_m < \mu_m)} \tag{1}$$

$$\beta_i^+ = \frac{\text{cov}(r_i, r_m | r_m > \mu_m)}{\text{var}(r_m | r_m > \mu_m)} \tag{2}$$

where r_i is the security i's (excess) return, r_m is the (excess) market return and μ_m is the average excess market return.

We also measure the economic significance of the difference between β_i^- and β_i^+, $\beta_i^- - \beta_i^+$. The corresponding factor serves as a non-linear correction to the market model, and, in regressions, we run $\beta_i^- - \beta_i^+$ alongside CAPM β. According to Ang, Chen, and Xing (2006), such kinked non-linearity may arise if investors are disappointment-averse with Gul (1991) utility. Disappointment-averse investors would not only require remuneration for greater exposure to shares with greater relative downside risk but would also construct their portfolios to take these exposures into account (Dahlquist, Farago, and Tédongap, 2016).

Another non-linear correction to the market model we consider is conditional coskewness. Such a correction may arise if the the pricing kernel is non-linear, but this non-linearity is smooth (Harvey and Siddique, 2000; Dittmar, 2002). Following Harvey and Siddique (2000), we use conditional coskewness, estimated based on idiosyncratic stock returns $\epsilon_i = r_i - \hat{\alpha}_i - \hat{\beta}_i r_m$. Conditional coskewness separates the non-linear dependence of stock returns from the linear dependence reflected in CAPM β:

$$\text{Coskew}_i = \frac{E[\epsilon_i (r_m - \mu_m)^2)]}{\sqrt{\text{var}(\epsilon_i)}\text{var}(r_m)}. \tag{3}$$

Whereas coskewness arises from a quadratic correction to the linear pricing kernel, cokurtosis arises from the third-order correction (Dittmar, 2002). We do not report our results for cokurtosis in this paper for the sake of brevity, but they are available upon request. We define conditional cokurtosis in a manner analogous to conditional coskewness:

$$\text{Cokurt}_i = \frac{E[\epsilon_i (r_m - \mu_m)^3)]}{\sqrt{\text{var}(\epsilon_i)\text{var}(r_m)^3}}. \tag{4}$$

In addition to corrections of the market model for non-linear dependence of stock returns on market returns, researchers have found empirical evidence that asymmetric tail risk, measured using conditional correlations, may also explain the cross-section of stock returns and be economically significant. The study by Longin and Solnik (2001) found that returns of different countries' market indices have higher correlations during extreme market moves than would be expected if the joint distribution of these returns was multivariate normal. Econometric tests by Ang and Chen (2002) and Hong, Tu, and Zhou (2006) implied that individual equities have a higher correlation with the market during downward moves than during upward moves. Studies by Alcock and Hatherley (2016) and Jiang, Wu, and Zhou (2018) found strong evidence to support economic significance of asymmetric correlations and their role in explaining the cross-section of stock returns.

We consider two popular metrics of asymmetric correlations. The first is the exceedance correlation, ρ^{exc} , introduced by Longin and Solnik (1995), a measure of the difference in correlations between stock and market returns when both stock and market returns rise above a threshold c and when they both fall below a threshold $-c$. The second metric is a compound metric of exceedance correlations for a set of thresholds $C = \{c_n\}$, J^{adj}, introduced by Hong, Tu, and Zhou (2006) and further refined by Alcock and Hatherley (2016). As prescribed by Alcock and Hatherley (2016), we remove β-dependence of individual stocks through a transformation: $\bar{r}_i = r_i - \hat{\beta}_i r_m + r_m$, which results in every stock having effectively $\beta = 1$. We then standardize both stock returns and market returns so that $\tilde{r}_i$ and $\tilde{r}_m$ have a mean 0 and variance 1.

Exceedance correlations ρ_c^{exc} for a cutoff c then equal:

$$\rho_c^{exc} = \rho(\tilde{r}_i, \tilde{r}_m | \tilde{r}_i > c, \tilde{r}_m > c) - \rho(\tilde{r}_i, \tilde{r}_m | \tilde{r}_i < -c, \tilde{r}_m < -c), \tag{5}$$

where ρ is the Pearson correlation function. For future use in Appendix C, where we discuss the properties of exceedance correlations, we also define ρ^+ and ρ^- as

$$\rho^+ = \rho(\tilde{r}_i, \tilde{r}_m | \tilde{r}_i > 0, \tilde{r}_m > 0) \tag{6}$$

$$\rho^- = \rho(\tilde{r}_i, \tilde{r}_m | \tilde{r}_i < 0, \tilde{r}_m < 0). \tag{7}$$

The J^{adj} metric combines exceedance correlations for a set of cutoffs. Following Alcock and Hatherley (2016), we use $C = [0, 0.2, 0.4, 0.6, 0.8, 1]$. We then define a vector, $\boldsymbol{\rho}_C^{exc} = \{\rho_c^{exc}\}_{c \in C}$, and construct J^{adj} as:

$$J^{adj} = T_1 \text{sign}(\sum_C \boldsymbol{\rho}_C^{exc})(\boldsymbol{\rho}_C^{exc})' \Omega_C^{-1} \boldsymbol{\rho}_C^{exc}, \tag{8}$$

where Ω_C is the variance-covariance matrix of $\boldsymbol{\rho}_C^{exc}$ constructed as in Hong, Tu, and Zhou (2006) using the Andrews (1991) method with a Bartlett kernel to ensure that the matrix is almost surely invertable.

Lastly, in our discussion of exceedance correlations and in robustness checks, we also consider skewness and kurtosis. Even though these metrics do not represent systematic risk factors, they provide information about deviations of the distributions of stock returns from normal.

3.2 Data

We use data provided by the Center for Research in Securities Prices (CRSP) and Compustat for the period from 1 January 1963 to 31 December 2018, updating previously published results by approximately 15 years (30-40% increase in sample size). We use daily adjusted stock returns and the value weighted index provided by CRSP. To compute excess stock returns, we use monthly risk-free rate data provided by the Fama-French database within CRSP. We select stocks that traded on NYSE/NASDAQ/AMEX during the period, with share codes 10 and 11. We eliminate shares without trading on over 30 percent of trading days. We also omit shares with missing book values in the Compustat database or book value records interrupted for over four years. There are 3133 individual stocks in our sample.

To estimate factor loadings, we use the value-weighted index provided by CRSP as a proxy for the market. For robustness checks, we also run our analyses using the equally weighted index provided by CRSP. We also run checks with with equally weighted and (approximately) value weighted indices we construct from our sample of stocks. We run additional robustness checks, reported in Appendix A and the Online Appendix, with size, the book-to-market ratio, and momentum factors (short- and medium-term) as controls in regressions. For these checks, we use book values provided by Compustat. Following Fama and French (1992), we assume book values become available to market participants with a six month lag. We do not winsorize factor loadings; winsorizing at [1, 99] level makes very little discernible difference.

3.3 Regression results: contemporaneous, staggered, and predictive

We apply staggered regressions to analyze whether conditional systematic factors are priced cross-sectionally. We confirm that exposure to downside risk β^-, the difference between downside and upside risk betas $\beta^- - \beta^+$, and coskewness may be rewarded by the market. We also show that it is less likely that investors have been rewarded for exposure to tail risk as measured by exceedance correlations, ρ^{exc} and J^{adj}, despite the highly statistically significant premia for these factors found in contemporaneous regressions.

We use a six-month rolling window ($T_1 = 6$) of contiguous return data for contemporaneous and predictive regressions, and an 18-month rolling window with six one-month long estimation subperiods ($\tilde{T}_1 = 18, T_1 = 6, T_S = 1$) for staggered regressions, as schematically represented in Fig. 2. Because of the construction of staggered regressions, the number of daily returns used for estimation of factor loadings and average returns over the 18 months period is equivalent to 6 months. We chose a contiguous 6 rolling months window and an 18 months staggered window to bracket the 12-months window used in previous empirical studies of conditional systematic factors (as cited in the previous Section). We conducted robustness checks with a $T_1 = 4, \tilde{T}_1 = 12$ window and a $T_1 = 12, \tilde{T}_1 = 36$ window and found qualitatively similar results.

Table 2 summarizes the regression results.

Particularly striking is the decline in statistical significance of estimated premia associated with exceedance correlation, ρ^{exc} and its related metric J^{adj}, in panels (7) and (8), when there is no overlap between data used for factor and realized return estimation. In staggered regressions and predictive regressions with a 1-month return estimation period ($T_2 = 1$), statistical significance of these factors disappears. For the predictive regressions with a 6-month return estimation period ($T_2 = 6$), measured risk premia associated with both ρ^{exc} and J^{adj} are small, but statistically significant in regressions controlled only for CAPM beta. However, this significance falls when size and book-to-market factors are added as controls (Table OA3) and disappears when momentum factors are added (Table A1).

This result points to a confounding factor responsible for the high statistical significance seen in contemporaneous regressions. In Appendix C, we show analytically and in Monte-Carlo simulations how such linkage arises in the presence of stock price jumps and, for predictive regressions, persistent skewness of individual stock returns.

Staggered regressions also capture statistically significant links between individual stock returns and CAPM β, in panel (1); downside risk β^-, in panel (2); $\beta^- - \beta^+$, in panel (5); and coskewness, in panel (6), in a way that predictive regressions fail to capture. Even though the economic significance of these factors declines dramatically in predictive regressions, it is strong in staggered regressions. This behavior would be consistent with relatively short-term responses – on the scale of weeks – to innovations in these factors.

We have structured the regressions as single factor, as in panels (1)-(4), and "market-model-plus-factor" two-factor regressions, as in panels (5)-(8), because of the significant potential for bias from errors-in-variables in multiple regressions (Jagannathan and Wang, 1998; Jegadeesh, Noh, Pukthuanthong, Roll, and Wang, 2019). Relatively high intercept values in Table 2 indicate that errors-in-variables is a valid concern. We run single factor regressions on factor exposures estimated using individual stock returns rather than idiosyncratic returns – β, β^- and β^+. We use "market-model-plus-factor" regressions on factors that are meant to complement the market model and/or are estimated using idiosyncratic returns. These factors are non-linear corrections to the market model, represented by $\beta^- - \beta^+$, coskewness, cokurtosis, exceedance correlations ρ^{exc}, and J^{adj}. In the presence of errors-in-variables, single factor regressions also provide biased estimates, but this bias is of known direction – downward in absolute value and significance. Panels of multiple regressions with size, book-to-market and momentum factors are reported in Appendix A and the Online Appendix.

All regressions in this Section are equal-weighted for ease of comparison with prior literature. Value-weighted regressions result in similar outcomes.

A few notes on the statistical summary of factor estimates provided in Table 1: The first observation is that mean factor loading estimates and their standard deviations in staggered and contiguous estimation are within a few percentage points of each other. Mean β, β^-, and β^+ values are less than 1 due to equally weighed averaging (the average values are close to those reported by, e.g. Ang, Chen, and Xing, 2006 or Alcock and Hatherley, 2016). Value-weighted averages of these quantities are equal to 1, within standard error.

3.4 Staggered regressions with different estimation subperiods

In this section, we provide results of staggered regressions with a range of estimation subperiods, T_S (Table 3). We use subperiods of $T_S = 1$ month, 2 months, 3 months, and 6 months, within a 18-month estimation period $\tilde{T}_1 = 18$ (T_S and $\tilde{T}_1$ are defined, e.g. in Fig. 2). Staggered regressions reported in the previous section used an estimation subperiod $T_S = 1$, as shown in Fig. 2. For these regressions, the 18-month estimation period was effectively split into 6 equal periods, each equal to a quarter of a year, and, in each quarter, one month was used to estimate factor loadings and one month was used for to estimate average returns. In this section, we use three additional alternative staggered configurations: First, $T_S = 2$, where the the 18 month estimation is split into 3 six-month periods; in each six-month period a 2-month period is used to estimate factor loadings and a 2 month period is used to estimate average returns. Second, $T_S = 3$, where the the 18 month estimation is split into 2 nine-month periods; in each nine-month period a 3-month period is used to estimate factor loadings and a 3 month period is used to estimate average returns. Third, $T_S = 6$, which is equivalent to the predictive regression with 6-month realized return estimation.

A longer estimation subperiod implies a greater elapsed time between factor loading and return estimation. In predictive regressions with a six month return estimation period, the corresponding returns are, on average, six months "older" than estimated factor loadings. Note that, because of the two-day skip at the beginning and end of each estimation period, the $T_S = 1$ staggered regression uses 10% fewer days of data than the $T_S = 2$ staggered regression, which affects the regression results slightly.

By comparing results of the $T_S = 3$ and predictive regressions with the results of the $T_S = 1$ and $T_S = 2$ staggered regressions, we observe that the economic and statistical significance of market risk, measured by CAPM β, panel (1), and downside risk, measured by β^-, panel (2), decline gradually with lengthening estimation period (when one takes into account the fact that the step from a three-month to a six-month estimation period is greater than that from two-month to three-month estimation period).

However, the economic and statistical significance of measures of non-linearity, $\beta^+ - \beta^-$, panel (5), and coskewness, panel (6), take a step down as the estimation subperiod moves from two-months to three months. This behavior again points to short-term drivers of non-linear dependence between individual stock returns and market returns, that take place on the time scale of a few weeks.[6]

Statistical significance of risk premia associated with measures of asymmetric dependence, ρ^{exc} and J^{adj}, is highly non-linear, rising dramatically for the $T_S = 6$ regressions. This effect arises because, as we will discuss in detail in Appendix C, these factors are highly sensitive to jumps – positive (negative) jumps are associated with negative (positive) ρ^{exc} and J^{adj}. Additionally, as discussed in Section 3.3, the economic significance of these factors remains small relatively to their significance in contemporaneous regressions and declines further, together with their statistical significance, when size, book-to-market and momentum factors are added to regressions as controls.

3.5 Staggered regressions with factor loading estimation periods following return estimation periods and vice versa.

In order to shed further light on the impact of shorter term innovations in factor loadings on returns, we compare the results of staggered regressions with factor loading estimation periods *preceding* return estimation periods (Table 4, column Staggered - A) and vice versa – with factor loading estimation periods *following* return estimation periods (Table 4, column Staggered - B).

Counterintuitively, the statistical link between average returns and factor loading estimates appears stronger for all factors, except CAPM β and downside risk β^-, when factor loading estimation follows return estimation. Both the measured risk premium and its statistical significance are higher when factor loading estimation follows return estimation, particularly for $\beta^+ - \beta^-$ and for coskewness - measures of non-linear dependence of individual stock returns on market returns.

This is not a statistically significant, but qualitatively interesting finding pointing to an effect not addressed in the literature to date: that asymmetry between β^- and β^+ is a result of an

[6]As we note in the next Section, we explore this phenomenon further in Foster et al. (2020).

asymmetric response to a systematic shock. We explore this insight further in Foster et al. (2020), where we establish a strong link between price delay and beta asymmetry.

3.6 Simulations to explore differences in β^- and β^+

In this section we explore the difference in the risk premia of downside risk beta, β^-, and upside risk, β^+, and their statistical significance. In our regressions on these factors, summarized in Table 2, panels (2) and (3), exposure to downside risk, β^-, earns a statistically significant risk premium, but exposure to β^+ does not. Empirical tests in Ang, Chen, and Xing (2006) and Alcock and Hatherley (2016) find a discount associated with β^+. We reproduce this discount when we run multiple regressions that include both β^- and β^+, as shown in panel (5).

A natural question to ask is whether the difference between β^- and β^+ is a result of errors in variables, due to higher estimation error for β^+ than for β^-, but our evidence is not consistent with this hypothesis. First, the standard error for β^+ is only around 10% higher than that for β^- and the difference is too small to create a gap in statistical significance of observed magnitude. Second, we have constructed simulations in attempt to replicate the empirical result. We tested a number of hypotheses that may have caused a difference in β^- and β^+ and the difference in estimated risk premia: errors-in-variables from differences in idiosyncratic risk (variance, skewness, kurtosis) observed during positive and negative market moves; effects of fat-tailed idiosyncratic risk; effects of indexing and compounding; effects of next-day reversals; effects of a greater concentration of idiosyncratic jumps during moderate market moves. Of the tested hypotheses, only the last produced results somewhat similar to those found empirically, but with much lower magnitude and statistical significance; moreover, this hypothesis is not consistent with results of staggered regressions.

We report the results of a subset of these simulations in the Online Appendix.

4 Staggered, Contemporaneous, and Predictive Regressions: A Comparison

In this section, we compare the bias and standard error of factor premium estimation across the three types of Fama-MacBeth regressions used in this paper: contemporaneous, predictive, and staggered.

4.1 Sources of Bias in Premium Estimation

We start by considering how the choice of contemporaneous, predictive, or staggered regression affects the premium estimator bias. We use a parsimonious single-factor model, with a time-varying factor f, which earns a premium γ, to capture the salient features of the premium-estimation problem: factor variation across time, errors in estimation of factor loadings, correlated errors in estimation of factor loadings and average returns, and correlations in factor and its premium across time. We start with a general formulation of factor time-variation and then, in Section 4.3, assume it takes the form of a mean-reverting AR(1) process to obtain a closed-form solution for factor premium estimation bias in contemporaneous, predictive, and staggered regression.

We consider the individual excess stock returns r_{it} of N stocks, $i = 1..N$, such that:

$$r_{it} = \gamma_t f_{it} + \varepsilon_{it}, \tag{9}$$

where f_{it} is the factor loading of stock i at time t, γ_t is the corresponding risk premium at time t, and ε_{it} is the disturbance (aka noise). We aim to estimate the factor premium γ_t and, using the Fama-MacBeth two-pass method, proceed in two steps, as described in Section 2.[7]

As the first pass, we select a set of estimation points t and estimate factor loadings and average returns for each of these points. The three types of Fama-MacBeth regressions compared in this paper – contemporaneous, predictive, and staggered – differ in what data intervals (sections of time

[7] Although the risk premium in our model can vary over time, the focus of the model is the time variation of factor loading. For the impact of time variation of risk premium, see, e.g. Gagliardini et al. (2016) and references therein.

series of stock returns) are used to estimate returns and factor loadings at each t, as schematically represented in Figure 1. We call the interval used for factor loading estimation I_{1t} and the interval used for return estimation I_{2t}. In contemporaneous regressions, $I_{1t} = I_{2t} = (t - T_1, t]$, where $T_1 > 0$. In predictive regressions, $I_{1t} = (t - T_1, t]$ and $I_{2t} = (t, t + T_2]$, with $T_1, T_2 > 0$. In staggered regressions, I_{1t} and I_{2t} is are split into staggered subperiods as described in Section 2 and illustrated in Figure 2. The number of data points in I_{1t} and I_{2t} are $\mathcal{T}_1$ and $\mathcal{T}_2$ respectively. We follow notation prevalent in the literature, where t is expressed in months when daily data are used for estimation. The units of t, T_1, and T_2 are months; the units of $\mathcal{T}_1$ and $\mathcal{T}_2$ (and also τ, s, and u, introduced below) are days. With this notation, we write factor loadings and returns estimated at each t as $\hat{f}_{it,I_1}$ and $\hat{r}_{it,I_2}$ respectively.

As the second pass, for each estimation point t, we regress the estimated average returns $\hat{r}_{it,I_2}$ against the estimated factor loadings $\hat{f}_{it,I_1}$ cross-sectionally to obtain the estimated factor risk premium, $\hat{\gamma}_t$:

$$\hat{\gamma}_t = \frac{\sum_{i=1}^{N}(\hat{r}_{it,I_2} - \frac{1}{N}\sum_{j=1}^{N}\hat{r}_{jt,I_2})(\hat{f}_{it,I_1} - \frac{1}{N}\sum_{j=1}^{N}\hat{f}_{jt,I_1})}{\sum_{i=1}^{N}(\hat{f}_{it,I_1} - \frac{1}{N}\sum_{j=1}^{N}\hat{f}_{jt,I_1})^2}. \tag{10}$$

In Fama-MacBeth regressions, the longitudinal average across all estimation periods is then used as the average premium estimate $\hat{\gamma}$, and the distribution of $\hat{\gamma}_t$ is used to run statistical tests.

Because the estimates $\hat{f}_{it,I_1}$ and $\hat{r}_{it,I_2}$ are generally not equal to f_{it} and r_{it}, the premium estimate can be biased, i.e. $E[\hat{\gamma}_t] \neq \gamma_t$. To evaluate the magnitude of the bias in Fama-MacBeth regressions, we start by evaluating $\hat{f}_{it,I_1}$ and $\hat{r}_{it,I_2}$ and then use Eq. (10) to estimate $E[\hat{\gamma}_t]$, under a number of (realistic) simplifying assumptions described below.

We start with the factor loading estimate $\hat{f}_{it,I_1}$. We can define $\bar{f}_{it,I_1} \equiv E[\hat{f}_{it,I_1} | f_{i\tau}, \tau \in I_1]$ – expected value of the factor loading estimator conditional upon a realization of $f_{i\tau}$ – so that

$$\hat{f}_{it,I_1} = \bar{f}_{it,I_1} + w_{it,I_1}, \tag{11}$$

where w_{it,I_1} is a random variable measuring variation in estimated factor loading, such that $E[w_{it,I_1}] = 0$ and $\sigma^2_{wi,I_1} \equiv \mathrm{Var}(w_{it,I_1})$.

The precise form of $E[\hat{f}_{it,I_1}|f_{i\tau}, \tau \in I_1]$ is unknown; however, a plausible simplifying assumption is that it is equal to the mean of $f_{i\tau}$ over the interval I_{1t}, so that we have for $\bar{f}_{it,I_1}$ (which we defined above to represent $E[\hat{f}_{it,I_1}|f_{i\tau}, \tau \in I_1]$):

$$\bar{f}_{it,I_1} = \frac{1}{\mathcal{T}_1} \sum_{\tau \in I_1} f_{i\tau}. \tag{12}$$

Under this assumption, $\hat{f}_{it,I_1}$ is an unbiased estimator of the factor loading mean across interval I_{1t}, with estimation error equal to w_{it,I_1}.

In a similar fashion, we can express the return estimator as:

$$\hat{r}_{it,I_2} = \bar{\gamma}_{t,I_2} \bar{f}_{it,I_2} + \mathrm{Cov}_{I_2}(\gamma_t, f_{it}) + e_{it,I_2}, \tag{13}$$

where $\bar{\gamma}_{t,I_2} = \frac{1}{\mathcal{T}_2} \sum_{\tau \in I_2} \gamma_{i\tau}$; Cov_{I_2} is a longitudinal covariance that reflects co-variation of the factor loading and the factor premium over the interval I_{2t}; and e_{it,I_2} is a random variable measuring variation in return estimation, such that $E[e_{it,I_2}] = 0$ and $\mathrm{Var}(e_{it,I_2})$ is the square error of $\hat{r}_{it,I_2}$ estimation.

The covariance between the factor loading and the factor premium can affect factor premium estimation. Even though its impact is the same for contemporaneous, predictive, and staggered regressions, we can include a simple form of factor-premium co-variation into our analysis. We can assume that, to lowest two orders in f_{it}, $\mathrm{Cov}_{I_2}(\gamma_t, f_{it}) = a_t + b_t \bar{f}_{it,I_2}$.[8] Under this assumption we can write:

$$\hat{r}_{it,I_2} = a_t + (\bar{\gamma}_{t,I_2} + b_t) \bar{f}_{it,I_2} + e_{it,I_2}, \tag{14}$$

Now we can use the expressions for factor loading and return estimates in Eqs. (11) and (14) to expand the premium estimate in Eq. (10). Under the assumptions described above and an additional assumption that the error terms w_{it,I_m} and e_{it,I_m} are independent of the underlying

[8]Strictly speaking, the assumption is $E[\mathrm{Cov}_{I_2}(\gamma_t, f_{it})|f_{i\tau}, \tau \in I_2] = a_t + b_t \bar{f}_{it,I_2}$.

factor structure, we have:

$$E[\hat{\gamma}_t] = \frac{\mathrm{Cov}_C(\hat{r}_{it,I_2}, \hat{f}_{it,I_1})}{\mathrm{Var}_C(\hat{f}_{it,I_1})} = \frac{(\bar{\gamma}_{t,I_2} + b_t)\mathrm{Cov}_C(\bar{f}_{it,I_2}, \bar{f}_{it,I_1}) + \mathrm{Cov}_C(e_{it,I_2}, w_{it,I_1})}{\mathrm{Var}_C(\bar{f}_{it,I_1}) + \sigma^2_{wi,I_1}}, \tag{15}$$

where the subscript C on Cov_C and Var_C indicates that the covariance and variance are cross-sectional; σ^2_{wi,I_1} is the variance of the factor loading estimation error over period I_{1t} (as defined after Eq. 11).

We can use Eq. (15) as a framework to analyze sources of bias in estimation of γ_t. In addition to the fact that we estimate averaged $\bar{\gamma}_{t,I_2}$ rather than instantaneous γ_t, there are four other sources of bias.

Importantly, two of the sources of bias in $\hat{\gamma}_t$ can bias the estimate upward, that is to say create an appearance of a larger factor premium, or worse, a statistically significant estimated premium in the absence of an underlying factor premium. First, an apparent estimated premium can result from cross-sectionally correlated estimation residuals, $\mathrm{Cov}_C(e_{it,I_2}, w_{it,I_1}) \neq 0$. This can happen in the presence of confounding factors, correlated with both factors and returns. Second, factor loadings correlated with factor premia can result in upward-biased premium estimation, as reflected by b_t, defined in Eq. (14).

The two other sources of bias present in our model can push the premium estimate downward. The first source of bias is time variation of the factor. When the factor is time-varying, we have $\mathrm{Cov}_C(\bar{f}_{it,I_2}, \bar{f}_{it,I_1}) < \mathrm{Var}_C(\bar{f}_{it,I_1})$, particularly if the estimation periods used to estimate returns and factor loadings are different, $I_{1t} \neq I_{2t}$. The second source of downward bias is the errors-in-variables bias, arising from the error in estimating $\bar{f}_{it,I_1}$. This error is reflected in Eq. (15) as $\sigma^2_{wi,I_1} > 0$ in the denominator.

4.2 Premium Estimation Error

In this Section, we estimate the standard error of Fama-MacBeth regressions, used in testing the statistical significance of non-zero factor premium estimates. In Fama-MacBeth regressions, standard error is estimated as square root of the variance of $\hat{\gamma}_t$ estimates across the entire sample, $t = 1..T_T$:

$$\hat{s}^2 = \frac{1}{T_T}\sum_{t=1}^{T_T}(\hat{\gamma}_t - \frac{1}{T_T}\sum_{t'=1}^{T_T}\hat{\gamma}_{t'})^2. \tag{16}$$

The variance $\hat{s}^2$ depends on underlying variation of the risk premium across time and the error of risk premium estimation at each estimation point t. If these two drivers of $\hat{s}^2$ are uncorrelated, the expectation of the error $E[\hat{s}^2]$ comprises two components:

$$E[\hat{s}^2] = \sigma_\gamma^2 + E[\hat{s}_\gamma^2] \tag{17}$$

$$\sigma_\gamma^2 \equiv E[\frac{1}{T_T}\sum_{t=1}^{T_T}(\gamma_t - \frac{1}{T_T}\sum_{t'=1}^{T_T}\gamma_{t'})^2] \tag{18}$$

$$\hat{s}_\gamma^2 \equiv \frac{1}{T_T}\sum_{t=1}^{T_T}(\hat{\gamma}_t - \gamma_t - \frac{1}{T_T}\sum_{t'=1}^{T_T}(\hat{\gamma}_{t'} - \gamma_{t'}))^2, \tag{19}$$

where σ_γ^2 reflects underlying variation of the risk premium across time and $\hat{s}_\gamma^2$ is the average squared premium estimation error.

In Equation (17), the $\hat{s}_\gamma^2$ term is the only term that depends on regression architecture (σ_γ^2 is driven by the underlying process). We can make the regression-dependence explicit and write:

$$\hat{s}_{\gamma I_1,I_2}^2 \approx \frac{1}{T_T(N-2)}\Big[\frac{\sigma_{r,I_2}^2}{\sigma_{f,I_1}^2} - E[\gamma_t^2]\Big], \tag{20}$$

where N is the number of stocks in the sample; σ_{r,I_2}^2 is the cross-sectional variance of stock returns estimated over I_{2t} interval; σ_{f,I_1}^2 is the cross-sectional variance of factor loadings estimated over I_{1t} interval; we ignored effects of errors in estimation of factor loadings and average returns, assuming that they are not correlated across stocks and, therefore, their impact is smaller than that of σ_{r,I_2}^2 and σ_{f,I_1}^2 by $\mathcal{O}(1/N)$.[9]

[9] Asymptotic and small sample properties (out of the scope of this paper) can be established by extension of these properties for contemporaneous and predictive Fama-MacBeth regressions, as in Shanken (1992), Jagannathan and Wang (1998), and Shanken and Zhou (2007).

4.3 Premium Estimation in Contemporaneous, Predictive, and Staggered Regressions

Having set up general expressions for the estimated risk premium (Eq. 15) and its standard error (Eqs. 16 to 20), we can now illustrate the differences in premium estimation between contemporaneous, predictive, and staggered regressions.

For the purposes of this comparison, we ignore sources of bias common across the three estimation methods – errors-in-variables and co-variation between factor and its premium – and focus on the sources of bias that stem from the geometry of estimation intervals. We treat confounding factors separately, as a source of bias most likely to affect contemporaneous regressions, in which returns and factor loadings are estimated using the same data. Under these assumptions, which can be written as $b_t = 0$, $\sigma^2_{w,I_1} = 0$, and $\text{Cov}_C(e_{it,I_2}, w_{it,I_1}) = 0$, in this section we focus on a simplified version of Eq. (15):

$$\hat{\gamma}/\gamma \equiv \frac{E[\hat{\gamma}_t]}{\bar{\gamma}_{t,I_2}} = \frac{\text{Cov}_C(\bar{f}_{it,I_2}, \bar{f}_{it,I_1})}{\text{Var}_C(\bar{f}_{it,I_1})}. \tag{21}$$

To demonstrate how differences in $\hat{\gamma}/\gamma$ arise across contemporaneous, predictive, and staggered Fama-MacBeth regressions, we model factor time-variation using a simple mean-reverting AR(1) process:

$$f_{i\tau} = f_{i\tau-1} + (1-\phi)(f_i - f_{i\tau-1}) + \epsilon_{i\tau}, \tag{22}$$

where $0 < \phi < 1$. The disturbance term $\epsilon_{i\tau}$ has a defined variance, which we assume to be time-invariant and equal to σ^2_i (additionally, we assume that $\sigma^2_i = \sigma^2$ to simplify the algebra).

For this process, $\hat{\gamma}/\gamma$ can be written as (as shown in Appendix B):

$$\hat{\gamma}/\gamma = \frac{\mathcal{T}_1}{\mathcal{T}_2} \frac{\sum_{s\in I_2}\sum_{u\in I_1} e^{-|s-u|/T_B} + \eta^2}{\sum_{s,u\in I_1} e^{-|s-u|/T_B} + \eta^2}, \tag{23}$$

with the time constant T_B defined as

$$T_B \equiv -1/\log\phi > 0, \tag{24}$$

and the constant $\eta^2 \equiv \frac{\sigma_{f0}^2}{\sigma^2}(1-e^{-2/T_B})$, where $\sigma_{f0}^2 = \text{Var}_C(f_i)$.

Evaluating the sums over I_{1t} and I_{2t} using a continuous approximation, valid when the sampling period is much shorter than T_B, we get for $\hat{\gamma}/\gamma$ in contemporaneous, staggered, and predictive regressions ($\hat{\gamma}/\gamma_{ctp}$, $\hat{\gamma}/\gamma_{stg}$, and $\hat{\gamma}/\gamma_{prd}$, respectively):

$$\begin{aligned}
\hat{\gamma}/\gamma_{ctp} =& 1, \quad \hat{\gamma}/\gamma_{prd} = \frac{\frac{T_1^2 T_B}{T_1 T_2}(1-e^{-T_1/T_B})(1-e^{-T_2/T_B}) + \eta^2}{2\frac{T_B}{T_1}\left[1-\frac{T_B}{T_1}(1-e^{-T_1/T_B})\right] + \eta^2} \\
\hat{\gamma}/\gamma_{stg} =& \frac{\frac{T_B^2}{T_1^2}\left[\frac{T_1}{T_S} + (1+e^{2T_S/T_B})\sum_{k=1}^{T_1/T_S-1}(\frac{T_1}{T_S}-k)e^{-3kT_S/T_B}\right]\left(1-e^{-T_S/T_B}\right)^2 + \eta^2}{2\frac{T_B^2}{T_1^2}\left[\left(\frac{T_1}{T_B}[1-\frac{T_B}{T_S}(1-e^{-T_S/T_B})] + e^{T_S/T_B}(1-e^{-T_S/T_B})^2\sum_{k=1}^{T_1/T_S-1}(\frac{T_1}{T_S}-k)e^{-3kT_S/T_B}\right] + \eta^2}.
\end{aligned} \tag{25}$$

Note, η^2 contains information – the ratio of cross-sectional volatility of the long term factor mean to the short-term volatility σ_{f0}^2/σ^2 – that is difficult to measure. When η^2 is large, driven either by a strong and persistent cross-sectional variation in long-term factor mean or a persistence in innovations ($T_B \gg 1$), factor premium estimates in all the three types of Fama-MacBeth regressions are (nearly) unbiased. However, for all the factors we consider in this paper, empirical results show (such as those in Table 2) a strong possibility that premium estimates are biased downward in predictive regressions. It is therefore likely that, for the factors considered in this paper, η^2 is small. For market beta, this conclusion is consistent with results of Ghysels (1998), which demonstrate that beta estimates have structural breaks, and, therefore, the ratio σ_{f0}^2/σ^2 is low. The AR(1) model is likely only a crude approximation to the underlying factor dynamics, but it captures the critical feature of time variation – factor autocorrelations declining with time.

Figure 4 provides a visual representation of the results in Eq. (25). The left column presents the ratio of estimated premium to the true premium, $\hat{\gamma}/\gamma$. The right column presents the t statistic, which we calculate as the ratio of premium estimate to the standard error of regression: $\hat{\gamma}/\hat{s}_\gamma$. For all regressions in Figure 4, the effective estimation period is $T_1 = 6$ months (which, for staggered regressions, is stretched over $\tilde{T}_1 = 18$ months).

The first column, panels (a) and (b), presents the modeled estimated premium and t statistic as a function of the time constant of factor f, T_B. The dashed line represents the results for contemporaneous regressions; the solid line – staggered regressions with subperiod $T_S = 1$; and the dashed-dotted and dotted lines represents predictive regressions with $T_2 = 1$ and $T_2 = 6$ respectively. In the simplified model considered in this section, the factor premium estimated in contemporaneous regressions is unbiased: $\hat{\gamma}/\gamma_{ctp} = 1$. Premium estimates in staggered and predictive regressions are biased downward – $\hat{\gamma}/\gamma_{stg} < 1$ and $\hat{\gamma}/\gamma_{prd} < 1$ – particularly when the underlying factor varies rapidly, i.e. T_B is small. But staggered regressions result in significantly less biased premium estimates than predictive regressions. Staggered regressions also result in lower estimation error than predictive regressions (because they use a longer time series to estimate returns). In the presence of an underlying factor premium, staggered regressions provide a more sensitive tool for estimating factor loadings.

Figure 4, panels (a) and (b), show that, in the *absence* of confounding factors – i.e. when $\text{Cov}_C(e_{it,I_2}, w_{it,I_1}) = 0$ in Eq. (15) – contemporaneous regressions provide the best results. However, because of confounding factors – $\text{Cov}_C(e_{it,I_2}, w_{it,I_1}) \neq 0$ – arising as a result of using the same time series of returns on both sides of the regressions, contemporaneous regressions are prone to "false positives." When confounding factors may be present, staggered regressions provide the best balance between estimation bias and error.

Panels (c) and (d) compare the results of staggered regressions with different staggered subperiods T_S. In these panels, the time constant $T_B = 2$. All other things being equal, smaller estimation subperiods work better in staggered regressions than larger subperiods. However, staggered regressions with subperiods that are too small may be affected by short-term reversals and other confounding factors. A month-long staggered estimation subperiod $T_S = 1$ provides a good balance.

Panels (e) and (f) compare the results of predictive regressions with different return estimation periods, T_2. In these panels, as in panels (c) and (d), $T_B = 2$. As the estimation period T_2 increases, premium estimation bias grows and $\hat{\gamma}/\gamma_{prd}$ falls. For small T_2 standard error falls faster than premium estimate, resulting in increasing statistical significance of the estimate, which peaks around $T_2 \approx T_B$. For $T_2 > T_B$, estimation bias increases and statistical significance of estimated premium falls.

5 Conclusions

We have characterized conditional systematic factors of stock returns using a new econometric technique – the staggered regression. The economic significance of these factors had been characterized primarily using contemporaneous regressions (using individual stocks and portfolios). And even though these factors are associated with highly statistically significant risk premia in contemporaneous regressions, statistical significance is lost in predictive regressions. This disconnect invites the question: are the results of contemporaneous regressions "false positives" or the results of predictive regressions "false negatives"?

The staggered regression method helps to resolve this dilemma by combining the benefits of contemporaneous and predictive regressions, while avoiding some of the weaknesses. The power of contemporaneous regressions is in their ability to capture return responses to shorter-term innovations in the factors because both sides of the regression – the average returns and the estimated factors – are created from the same data set. But the data overlap results in a key shortcoming: contemporaneous regressions are vulnerable to spurious significance, driven by reverse-causality and confounding factors, such as outliers. Predictive regressions avoid this problem by separating the data set used to estimate average returns and the set used to estimate factors. But estimation of factor loadings typically requires many months of daily data (or years of weekly or monthly data), which means that average lag between the factor loading estimate and the corresponding average returns can be months or years. Like predictive regressions, staggered regressions separate the data sets used for factor and return estimation, but the staggered construction allows to capture the effect of shorter-term innovations in the factors on returns (and the skipped month reduces "reverse-causality").

Staggered regression analysis of β, β^-, β^+, $\beta^- - \beta^+$, and coskewness demonstrates that the economic significance of these factors found by other researchers through contemporaneous regressions is robust. We conclude that it is *not* due to direct simultaneity or other confounding factors stemming from overlapping data.

The economic significance of exceedance correlations ρ^{exc} and the related metric J^{adj} is not robust to staggered regressions. When the data used to estimate the factors and the returns are separated, the statistical significance of risk premia associated with these metrics is lost.

We create a simple model to demonstrate how jumps in return series drive the linkage between exceedance correlations and contemporaneous returns. We confirm the results in simulations. Further, we demonstrate that exceedance correlations may result in apparent statistically significant asymmetries as a result of non-normal idiosyncratic returns, without any underlying asymmetric dependence.

Because exceedance correlations are a sensitive measure of deviations of stock returns from multivariate normality that include not only asymmetric dependence, but also skewness and kurtosis of idiosyncratic returns without underlying asymmetry, we propose that these correlations can be used in other applications where a one-stop metric of the impact of jumps on stock returns is required.

References

Albuquerque, R., 2012. Skewness in stock returns: reconciling the evidence on firm versus aggregate returns. The Review of Financial Studies 25, 1630–1673.

Alcock, J., Hatherley, A., 2016. Characterizing the asymmetric dependence premium. Review of Finance 21, 1701–1737.

Andrews, D., 1991. Heteroskedasticity and autocorrelation consistent covariant matrix estimation. Econometrica 59, 817–858.

Ang, A., Chen, J., 2002. Asymmetric correlations of equity portfolios. Journal of Financial Economics 63, 443–494.

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

Ball, R., Brown, P., Officer, R., 1976. Asset pricing in the australian industrial equity market. Australian Journal of Management 1, 1–32.

Bawa, V. S., Lindenberg, E. B., 1977. Capital market equilibrium in a mean-lower partial moment framework. Journal of Financial Economics 5, 189–200.

Black, F., 1972. Capital market equilibrium with restricted borrowing. The Journal of Business 45, 444–455.

Black, F., Jensen, M. C., Scholes, M., 1972. The capital asset pricing model: Some empirical tests. Studies in the theory of capital markets 81, 79–121.

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

Campbell, R. A., Forbes, C. S., Koedijk, K. G., Kofman, P., 2008. Increasing correlations or just fat tails? Journal of Empirical Finance 15, 287–309.

Chordia, T., Goyal, A., Shanken, J. A., 2017. Cross-sectional asset pricing with individual stocks: betas versus characteristics. Available at SSRN 2549578 .

Cizeau, P., Potters, M., Bouchaud, J.-P., 2001. Correlation structure of extreme stock returns. Quantitative Finance 1, 217–222.

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

Dahlquist, M., Farago, A., Tédongap, R., 2016. Asymmetries and portfolio choice. The Review of Financial Studies 30, 667–702.

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

Fama, E. F., French, K. R., 1992. The cross-section of expected stock returns. The Journal of Finance 47, 427–465.

Fama, E. F., French, K. R., 1997. Industry costs of equity. Journal of Financial Economics 43, 153–193.

Fama, E. F., MacBeth, J. D., 1973. Risk, return, and equilibrium: Empirical tests. Journal of Political Economy 81, 607–636.

Foster, F. D., Lopatnikova, A., Satchell, S., 2020. Price delay and asymmetric dependence of stock returns. Working Paper .

Gagliardini, P., Ossola, E., Scaillet, O., 2016. Time-varying risk premium in large cross-sectional equity data sets. Econometrica 84, 985–1046.

Ghysels, E., 1998. On stable factor structures in the pricing of risk: do time-varying betas help or hurt? The Journal of Finance 53, 549–573.

Gibbons, M. R., 1982. Multivariate tests of financial models: A new approach. Journal of Financial Economics 10, 3–27.

Gul, F., 1991. A theory of disappointment aversion. Econometrica 59, 667–686.

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

Hong, Y., Tu, J., Zhou, G., 2006. Asymmetries in stock returns: Statistical tests and economic evaluation. The Review of Financial Studies 20, 1547–1581.

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

Huang, W., Liu, Q., Rhee, S. G., Zhang, L., 2009. Return reversals, idiosyncratic risk, and expected returns. The Review of Financial Studies 23, 147–168.

Jagannathan, R., Wang, Z., 1998. An asymptotic theory for estimating beta-pricing models using cross-sectional regression. The Journal of Finance 53, 1285–1309.

Jegadeesh, N., Noh, J., Pukthuanthong, K., Roll, R., Wang, J., 2019. Empirical tests of asset pricing models with individual assets: Resolving the errors-in-variables bias in risk premium estimation. Journal of Financial Economics 133, 273–298.

Jegadeesh, N., Titman, S., 1995. Overreaction, delayed reaction, and contrarian profits. The Review of Financial Studies 8, 973–993.

Jiang, L., Wu, K., Zhou, G., 2018. Asymmetry in stock comovements: an entropy approach. Journal of Financial and Quantitative Analysis 53, 1479–1507.

Kan, R., Robotti, C., Shanken, J., 2013. Pricing model performance and the two-pass cross-sectional regression methodology. The Journal of Finance 68, 2617–2649.

Kim, S., Skoulakis, G., 2018. Ex-post risk premia estimation and asset pricing tests using large cross sections: The regression-calibration approach. Journal of econometrics 204, 159–188.

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

Lewellen, J., Nagel, S., 2006. The conditional capm does not explain asset-pricing anomalies. Journal of Financial Economics 82, 289–314.

Lewellen, J., Nagel, S., Shanken, J., 2010. A skeptical appraisal of asset pricing tests. Journal of Financial Economics 96, 175–194.

Lintner, J., 1965. Security prices, risk, and maximal gains from diversification. The Journal of Finance 20, 587–615.

Litzenberger, R. H., Ramaswamy, K., 1979. The effect of personal taxes and dividends on capital asset prices: Theory and empirical evidence. Journal of financial economics 7, 163–195.

Lo, A. W., MacKinlay, A. C., 1990. Data-snooping biases in tests of financial asset pricing models. The Review of Financial Studies 3, 431–467.

Longin, F., Solnik, B., 1995. Is the correlation in international equity returns constant: 1960–1990? Journal of International Money and Finance 14, 3–26.

Longin, F., Solnik, B., 2001. Extreme correlation of international equity markets. The Journal of Finance 56, 649–676.

Mankiw, N. G., Shapiro, M. D., 1986. Risk and return: Consumption beta versus market beta. The Review of Economics and Statistics 68, 452–459.

Newey, W. K., West, K. D., 1987. Hypothesis testing with efficient method of moments estimation. International Economic Review pp. 777–787.

Scholes, M., Williams, J., 1977. Estimating betas from nonsynchronous data. Journal of Financial Economics 5, 309–327.

Shanken, J., 1992. On the estimation of beta-pricing models. The Review of Financial Studies 5, 1–33.

Shanken, J., Zhou, G., 2007. Estimating and testing beta pricing models: Alternative methods and their performance in simulations. Journal of Financial Economics 84, 40–86.

Sharpe, W. F., 1964. Capital asset prices: A theory of market equilibrium under conditions of risk. The Journal of Finance 19, 425–442.

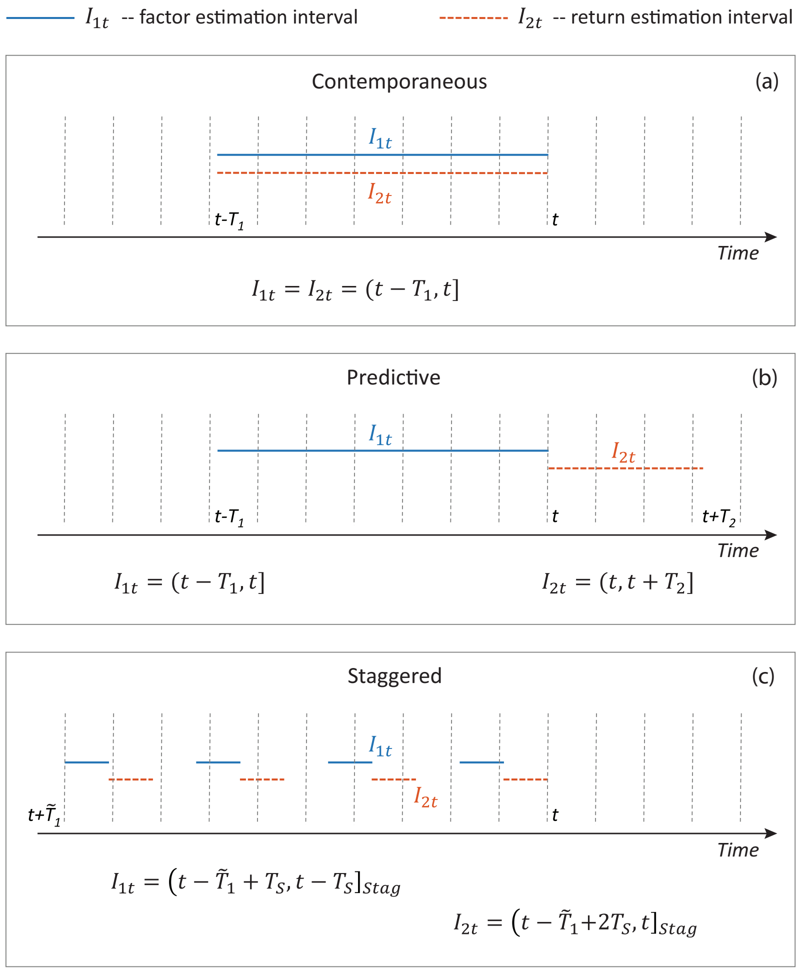

Figure 1: Contemporaneous, predictive, and staggered Fama-MacBeth regressions.
This figure provides a visual comparison of estimation periods used in (a) contemporaneous, (b) predictive, and (c) staggered regressions. Horizontal lines represent estimation periods used for estimation of factor loadings and average returns (solid and dashed lines, respectively), spaced out vertically to avoid visually overlapping lines. Throughout the text, we denote factor loading estimation periods I_{1t} and average return estimation periods I_{2t}, often omitting the subscript t to streamline notation when this omission would not cause confusion. The horizontal axis represents time, and the vertical dashed lines represent estimation points (such as $t-1$, t, or $t+1$). In *contemporaneous* regressions, schematically represented in panel (a), factor loadings and average returns for each estimation point are calculated using the same estimation period, $I_{1t} = I_{2t} = (t - T_1, t]$. In *predictive* regressions, panel (b), the return estimation period $I_{2t} = (t, t + T_2]$ for each estimation point t follows the factor loading estimation period $I_{1t} = (t - T_1, t]$ for that estimation point. In *staggered* regressions, panel (c), the estimation periods used for factor loadings and returns are staggered, so that $I_{1t} = (t - \tilde{T}_1 + T_S, t - T_S]_{Stag} \equiv (t - \tilde{T}_1 + T_S, t - \tilde{T}_1 + 2T_S] \cup \ldots \cup (t - 5T_S, t - 4T_S] \cup (t - 2T_S, t - T_S]$ and $I_{2t} = (t - \tilde{T}_1 + 2T_S, t]_{Stag} \equiv (t - \tilde{T}_1 + 2T_S, t - \tilde{T}_1 + 3T_S] \cup \ldots \cup (t - 4T_S, t - 3T_S] \cup (t - T_S, t]$, where $\tilde{T}_1$ is the total length of the staggered estimation period and T_S is the staggered estimation subperiod, as described in Section 2 and presented in more detail in Figure 2.

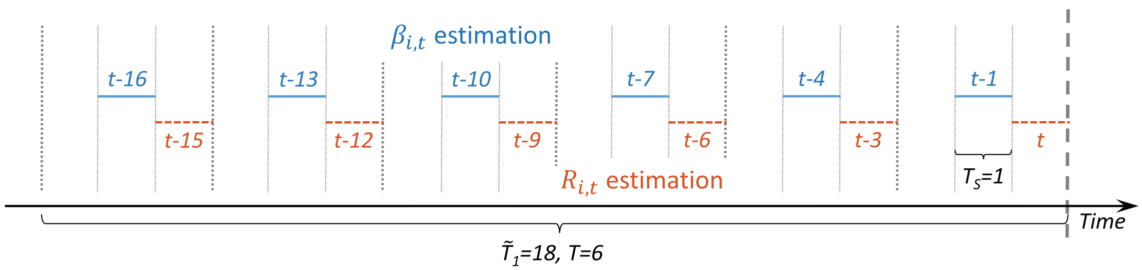

Figure 2: Staggered estimation for Fama-MacBeth regressions.

This figure presents a schematic of staggered estimation used in this paper. For each month t, an estimation period of 18 months was split into six quarters. The latest months of each quarter – months t, $t-3$, $t-6$, $t-9$, $t-12$, $t-15$, and $t-18$ – were used for realized return estimation. The middle months of each quarter – months $t-1$, $t-4$, $t-7$, $t-10$, $t-13$, and $t-16$ – were used for factor loading estimation. The last months of each quarter were skipped to promote the flow of causality from risk factors to returns. Additionally (not represented on this figure), two days were skipped at the beginning and end of each month to reduce impact of short-term reversals and non-synchronous trading on regression results.

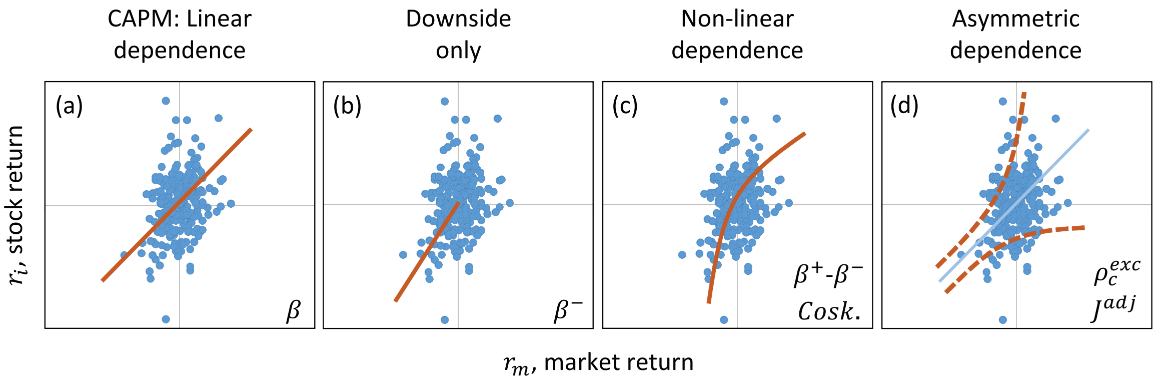

Figure 3: Types of conditional systematic risk factors.

This figure presents a schematic representation of conditional systematic risk factors considered in this paper, using simulated individual stock return and market return time series for a given estimation period. Panel (a) represents the CAPM, and the linear dependence of individual stock returns on market returns implied by the multivariate normal joint distribution of stock returns. Conditional systematic risk factors measure deviations of joint distribution of stock returns from multivariate normal. Measure of downside risk, β^-, is represented in panel (b); measures of nonlinear dependence of stock returns on market returns, $\beta^+ - \beta^-$ and coskewness, are presented in panel (c); measures of asymmetric dependence ρ^{exc} and J^{adj} are presented in panel (d). In panel (d), the dashed lines represent heteroskedasticity of idiosyncratic risk conditional on market returns – lower idiosyncratic risk when the market return is negative and higher idiosyncratic risk when it is positive.

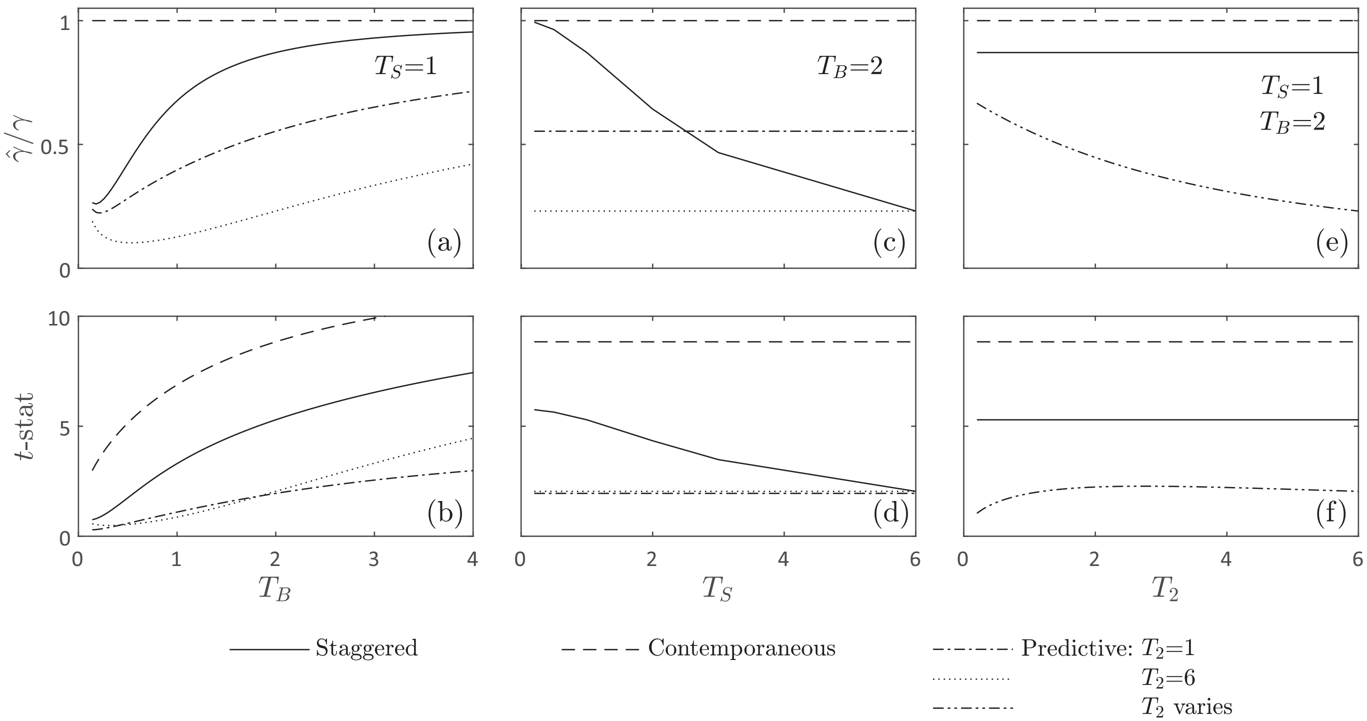

Figure 4: Bias and error in contemporaneous, staggered, and predictive regressions: a theoretical model.
This figure compares the relative estimated premium (the ratio of estimated premium to true premium, $\hat{\gamma}/\gamma$) and t-statistic (which we use as a relative measure of standard error) for the three types of Fama-MacBeth regressions across a range of parameters: T_B the time constant of the factor, T_S the staggered estimation subperiod, and T_2 the return estimation period. The factor loading estimation period is set to $T_1 = 6$ months. Solid lines represent staggered regressions, dashed lines – contemporaneous regressions, dashed-dotted, dotted, and dash-dot-dotted lines - predictive regressions with return estimation period $T_2 = 1$ months, $T_2 = 6$ months, and – in panels (e) and (f) – T_2 varied, respectively. In panels (c)-(f), the time constant $T_B = 2$. In panels, (a), (b), (e), and (f), the estimation subperiod for the staggered regression is $T_S = 1$ – the same as used in empirical investigations in Section 3.

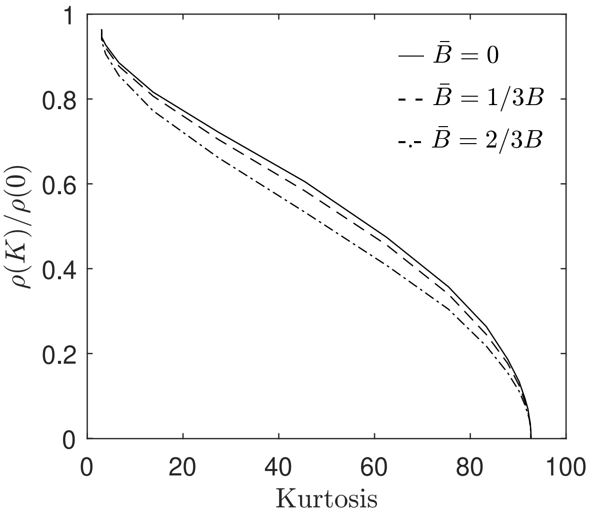

Figure 5: Dependence of Pearson correlation on kurtosis of underlying data.
We estimate the Pearson correlation function for model data with normally distributed noise $N(0,\sigma_0)$ and outliers of density q, mean $\bar{B}$, and variance B. As $B \to \infty$, $K \to 3/q$.

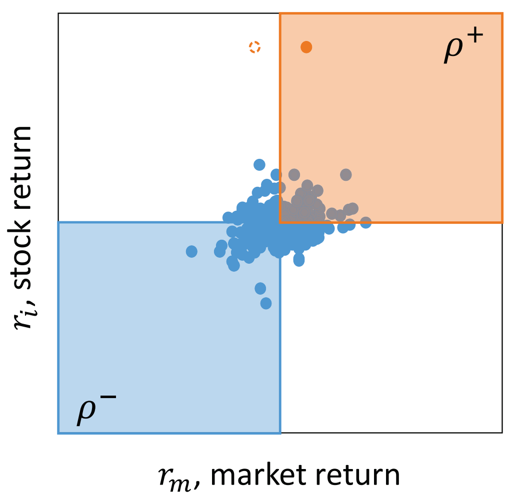

Figure 6: Impact of a positive idiosyncratic jump on exceedance correlations: a schematic representation. If a positive stock "jump" happens on a day when the market return was positive (filled circle), it suppresses ρ^+. A positive "jump" never affects ρ^-, even if it happens on a day when the market return was negative (open circle), because of the geometry of exceedance correlations. On average, positive idiosyncratic jumps lead to negative exceedance correlations, $\rho^{exc} = \rho^+ - \rho^-$.

Table 1: Descriptive factor statistics. This table presents statistics of estimated factor loadings. Factor loadings are estimated in two ways. For Contemporaneous & Predictive regressions, factor loadings are estimated using contiguous periods $(t - T_1, t]$ of daily return data preceding the estimation month t, where $T_1 = 6$ months. For Staggered regressions, factor loadings are estimated using staggered months $t-16$, $t-13$, $t-10$, $t-7$, $t-4$, and $t-1$, as described in Section 2 (and Figure 2). The table provides the mean (Mean), standard deviation (Std), and quantile breakpoints at 25%, 50%, and 75%. The study used daily return time series of stocks traded on NYSE/Amex/Nasdaq during the period between 1963 and 2018, with sharecodes 10 and 11.

Factors	Regression Types									
	Contemporaneous & Predictive					Staggered				
	Mean	Std	25%	50%	75%	Mean	Std	25%	50%	75%
β	0.917	0.59	0.503	0.860	1.255	0.902	0.57	0.499	0.847	1.237
β^-	0.975	0.78	0.481	0.904	1.381	0.962	0.78	0.470	0.894	1.371
β^+	0.857	0.86	0.328	0.801	1.318	0.847	0.86	0.315	0.790	1.308
$\beta^+ - \beta^-$	-0.118	0.92	-0.553	-0.074	0.349	-0.115	0.94	-0.564	-0.072	0.365
Cosk	-0.019	0.20	-0.130	-0.019	0.090	-0.020	0.21	-0.141	-0.020	0.099
Ckrt	0.016	0.72	-0.203	0.000	0.218	0.029	0.73	-0.216	0.007	0.250
ρ^{exc}	-0.125	0.28	-0.317	-0.125	0.061	-0.104	0.29	-0.302	-0.101	0.091
J^{adj}	-2.632	7.72	-7.536	-3.985	3.418	-2.210	8.01	-7.392	-3.598	4.024
Skew	0.306	1.11	-0.098	0.282	0.725	0.296	1.02	-0.101	0.270	0.697
Kurt	7.144	7.16	3.831	4.930	7.332	6.612	6.13	3.777	4.796	6.882

Table 2: Contemporaneous, staggered, and predictive cross-sectional regressions of average returns on estimated risk factor loadings. This table compares the results of three types of Fama-MacBeth regressions: Contemporaneous, in which returns and factor loadings for each month t are estimated using the same data periods $(t-T,t]$, where $T=6$ months; Staggered regressions, in which returns and factor loadings are estimated on staggered months as described in Section 2 (where, in each estimation period $(t-\tilde{T}_1,t]$ for each month t, where $\tilde{T}_1$ is 18 months, months $t-15$, $t-12$, $t-9$, $t-6$, $t-3$, and t are used for realized return estimation and months $t-16$, $t-13$, $t-10$, $t-7$, $t-4$, and $t-1$ are used for factor loading estimation); and Predictive regressions, where, for each month t, factor loadings are estimated using return time series $(t-T_1,t]$ and average returns are estimated either for a single month $t+1$ or for the period $(t,t+T_1]$, where T_1 is six months. Each panel reports results of a separate single- or multiple regression model. Reported t-statistics are adjusted for overlapping periods using the Newey-West method. Average $AdjR^2$ of cross-sectional regressions are reported for each estimation model. The study used daily return time series of stocks traded on NYSE/Amex/Nasdaq during the period between 1963 and 2018, with sharecodes 10 and 11.

	Regression Types							
Models	Contemporaneous		Staggered		Predictive – 1mo		Predictive – 6mo	
	Coefficient	t-Statistic	Coefficient	t-Statistic	Coefficient	t-Statistic	Coefficient	t-Statistic
Panel (1)								
Int	0.067	4.06	0.048	3.24	0.096	4.20	0.106	4.32
β	0.053	2.90	0.032	2.23	0.011	0.72	0.007	0.45
Adj R^2		4.9 %		3.8 %		2.9 %		2.6 %
Panel (2)								
Int	0.064	3.63	0.044	2.94	0.094	4.06	0.101	4.23
β^-	0.048	3.81	0.031	3.19	0.013	1.17	0.012	1.22
Adj R^2		3.8 %		2.7 %		2.0 %		2.0 %
Panel (3)								
Int	0.106	5.33	0.073	4.16	0.102	4.05	0.111	4.32
β^+	0.008	0.86	0.007	0.93	0.009	1.14	0.004	0.64
Adj R^2		2.4 %		2.1 %		1.6 %		1.3 %
Panel (4)								
Int	0.068	4.05	0.047	3.20	0.096	4.22	0.103	4.36
β^-	0.053	4.72	0.033	3.96	0.010	1.05	0.014	1.53
β^+	-0.010	-1.52	-0.005	-0.94	0.001	0.13	-0.006	-1.15
Adj R^2		4.9 %		3.6 %		2.7 %		2.4 %
Panel (5)								
Int	0.060	3.72	0.044	3.00	0.097	4.28	0.105	4.34

Table 2: (Continued.)

Models	Regression Types							
	Contemporaneous		Staggered		Predictive – 1mo		Predictive – 6mo	
	Coefficient	t-Statistic	Coefficient	t-Statistic	Coefficient	t-Statistic	Coefficient	t-Statistic
β	0.056	3.11	0.034	2.40	0.010	0.60	0.006	0.38
$\beta^{+} - \beta^{-}$	-0.026	-4.95	-0.016	-4.31	-0.003	-0.74	-0.008	-1.82
Adj R^2		5.8 %		4.3 %		3.4 %		3.0 %
Panel (6)								
Int	0.061	3.74	0.043	2.98	0.096	4.21	0.105	4.31
β	0.056	3.10	0.035	2.40	0.011	0.68	0.006	0.43
Cosk	-0.126	-6.42	-0.069	-4.48	-0.017	-0.90	-0.034	-2.20
Adj R^2		5.5 %		4.2 %		3.2 %		2.9 %
Panel (7)								
Int	0.026	1.73	0.049	3.27	0.094	4.13	0.104	4.29
β	0.050	2.85	0.032	2.24	0.011	0.69	0.006	0.41
ρ^{exc}	-0.348	-9.83	0.005	0.58	-0.022	-1.88	-0.047	-3.81
Adj R^2		8.0 %		4.0 %		3.1 %		2.8 %
Panel (8)								
Int	0.046	2.98	0.047	3.20	0.095	4.15	0.105	4.30
β	0.052	2.89	0.032	2.25	0.011	0.69	0.006	0.42
J^{adj}	-0.008	-9.65	0.000	0.29	-0.001	-1.82	-0.001	-3.50
Adj R^2		6.5 %		4.0 %		3.0 %		2.7 %

Table 3: Comparison of staggered regressions with 1-month, 2-month, 3-month, and 6-month estimation subperiods. This table presents results of staggered Fama-MacBeth regressions: In column Staggered - 1mo, factor loadings and average returns are estimated as in Table 2. In column Staggered - 2mo, for each month t, factor loadings are estimated using months $t-15$, $t-14$, $t-9$, $t-8$, $t-3$, and $t-2$ – i.e. three 2-month-long continuous periods. Average returns are estimated using months $t-13$, $t-12$, $t-7$, $t-6$, $t-1$, and t - using 2-month continuous periods that follow the 2-month periods used for factor loading estimation. In column Staggered - 3mo, for each month t factor loadings and average returns are estimated using 3-month continuous periods: $t-14$, $t-13$, $t-12$, $t-5$, $t-4$, $t-3$ and $t-11$, $t-10$, $t-9$, $t-2$, $t-1$, t respectively. Column Predictive - 6mo reports results of predictive regressions, as in Table 2 (last column). Each panel reports results of a separate single- or multiple regression model. Reported t-statistics are adjusted for overlapping periods using the Newey-West method. Average AdjR^2 of cross-sectional regressions are reported for each estimation model. The study used daily return time series of stocks traded on NYSE/Amex/Nasdaq during the period between 1963 and 2018, with sharecodes 10 and 11.

Models	Regression Types							
	Staggered - 1mo		Staggered - 2mo		Staggered - 3mo		Predictive - 6mo	
	Coefficient	t-Statistic	Coefficient	t-Statistic	Coefficient	t-Statistic	Coefficient	t-Statistic
Panel (1)								
Int	0.048	3.24	0.050	3.06	0.050	2.91	0.106	4.32
β	0.032	2.23	0.028	2.00	0.026	1.91	0.007	0.45
Adj R^2		3.8 %		4.2 %		3.9 %		2.6 %
Panel (2)								
Int	0.044	2.94	0.043	2.63	0.050	2.86	0.101	4.23
β^-	0.031	3.19	0.032	3.36	0.024	2.67	0.012	1.22
Adj R^2		2.7 %		2.8 %		2.8 %		2.0 %
Panel (3)								
Int	0.073	4.16	0.069	3.81	0.065	3.44	0.111	4.32
β^+	0.007	0.93	0.008	1.14	0.011	1.60	0.004	0.64
Adj R^2		2.1 %		2.3 %		2.1 %		1.3 %
Panel (4)								
Int	0.047	3.20	0.048	2.90	0.050	2.93	0.103	4.36
β^-	0.033	3.96	0.034	4.08	0.023	3.13	0.014	1.53
β^+	-0.005	-0.94	-0.005	-1.02	0.001	0.15	-0.006	-1.15
Adj R^2		3.6 %		3.9 %		3.6 %		2.4 %
Panel (5)								
Int	0.044	3.00	0.047	2.88	0.048	2.81	0.105	4.34

Table 3: (Continued.)

Models	Regression Types							
	Staggered - 1mo		Staggered - 2mo		Staggered - 3mo		Predictive - 6mo	
	Coefficient	t-Statistic	Coefficient	t-Statistic	Coefficient	t-Statistic	Coefficient	t-Statistic
β	0.034	2.40	0.030	2.14	0.027	1.98	0.006	0.38
$\beta^{+} - \beta^{-}$	-0.016	-4.31	-0.017	-4.27	-0.009	-2.60	-0.008	-1.82
Adj R^2		4.3 %		4.7 %		4.4 %		3.0 %
Panel (6)								
Int	0.043	2.98	0.047	2.86	0.048	2.83	0.105	4.31
β	0.035	2.40	0.031	2.13	0.027	1.98	0.006	0.43
Cosk	-0.069	-4.48	-0.067	-4.81	-0.033	-2.58	-0.034	-2.20
Adj R^2		4.2 %		4.5 %		4.2 %		2.9 %
Panel (7)								
Int	0.049	3.27	0.048	2.94	0.050	2.91	0.104	4.29
β	0.032	2.24	0.028	2.00	0.026	1.92	0.006	0.41
ρ^{exc}	0.005	0.58	-0.008	-0.98	-0.002	-0.29	-0.047	-3.81
Adj R^2		4.0 %		4.4 %		4.2 %		2.8 %
Panel (8)								
Int	0.047	3.20	0.048	2.94	0.050	2.90	0.105	4.30
β	0.032	2.25	0.028	2.00	0.026	1.92	0.006	0.42
J^{adj}	0.000	0.29	-0.000	-1.46	0.000	0.15	-0.001	-3.50
Adj R^2		4.0 %		4.3 %		4.1 %		2.7 %

Table 4: Comparison of staggered cross-sectional regressions of with factor loading estimation periods preceding (A) and following (B) periods used to estimate average returns. This table compares the results of staggered Fama-MacBeth regressions: In Staggered A, for each month t, factor loadings are estimated using months $t-16$, $t-13$, $t-10$, $t-7$, $t-4$, and $t-1$, and average returns are estimated using months $t-15$, $t-12$, $t-9$, $t-6$, $t-3$, and t, so that return estimation periods effectively follow factor loading estimation periods. In Staggered B, the return and factor loading estimation periods are swapped. Each panel reports results of a separate single- or multiple regression model. Reported t-statistics are adjusted for overlapping periods using the Newey-West method. Average AdjR^2 of cross-sectional regressions are reported for each model. The study used daily return time series of stocks traded on NYSE/Amex/Nasdaq during the period between 1963 and 2018, with sharecodes 10 and 11.

Models	Regression Types			
	Staggered - A		Staggered - B	
	Coefficient	t-Statistic	Coefficient	t-Statistic
Panel (1)				
Int	0.048	3.24	0.059	3.75
β	0.032	2.23	0.024	1.73
Adj R^2		3.8 %		3.5 %
Panel (2)				
Int	0.044	2.94	0.051	3.25
β^-	0.031	3.19	0.028	3.10
Adj R^2		2.7 %		2.5 %
Panel (3)				
Int	0.073	4.16	0.085	4.60
β^+	0.007	0.93	-0.004	-0.63
Adj R^2		2.1 %		1.9 %
Panel (4)				
Int	0.047	3.20	0.058	3.73
β^-	0.033	3.96	0.038	4.44
β^+	-0.005	-0.94	-0.017	-2.77
Adj R^2		3.6 %		3.4 %
Panel (5)				
Int	0.044	3.00	0.053	3.47
β	0.034	2.40	0.028	2.02
$\beta^+ - \beta^-$	-0.016	-4.31	-0.024	-4.87
Adj R^2		4.3 %		4.1 %
Panel (6)				
Int	0.043	2.98	0.054	3.53
β	0.035	2.40	0.027	1.94
Cosk	-0.069	-4.48	-0.092	-5.55
Adj R^2		4.2 %		4.0 %
Panel (7)				
Int	0.049	3.27	0.056	3.61

Table 4: (Continued.)

Models	Regression Types			
	Staggered - A		Staggered - B	
	Coefficient	t-Statistic	Coefficient	t-Statistic
β	0.032	2.24	0.024	1.74
ρ^{exc}	0.005	0.58	-0.011	-1.38
Adj R^2		4.0 %		3.7 %
Panel (8)				
Int	0.047	3.20	0.057	3.66
β	0.032	2.25	0.024	1.74
J^{adj}	0.000	0.29	-0.000	-1.77
Adj R^2		4.0 %		3.6 %

A Robustness Checks

In this section, we report regression results controlled for size, book-to-market ratio, and momentum factors (short- and medium-term). Table A1 summarizes the results. Additional robustness checks are reported in the Online Appendix.

B Estimated Premium of a Factor with an AR(1) Process

In this Section, we derive the relative premium estimate $\hat{\gamma}/\gamma$ for a factor evolving in time according to a mean-reverting AR(1) process introduced in Section 4.3, for each stock i:

$$f_{i\tau} = f_{i\tau-1} + (1-\phi)(f_i - f_{i\tau-1}) + \epsilon_{i\tau}. \tag{26}$$

The process reverts to the mean $f_i = E[f_{it}]$ on a time scale defined by the time constant, $T_B = -1/\log\phi$. As in Section 4.3, $\mathrm{Var}_C(f_i) = \sigma_{f0}^2$ and $\mathrm{Var}(\epsilon_{i\tau}) = \sigma^2$, which we assume to be independent of time and asset.

For this process, we have:

$$\frac{1}{\mathcal{T}_1\mathcal{T}_2}\sum_{s\in I_2}\sum_{u\in I_1}\mathrm{Cov}_C(f_{is},f_{iu}) = \frac{1}{N}\frac{1}{\mathcal{T}_1\mathcal{T}_2}\sum_{s\in I_2,u\in I_1}\sum_{i=1}^{N}\int (f_{is}-f)(f_{iu}-f)p(f_{is},f_{iu})\,d\Omega_{is}d\Omega_{iu}, \quad (27)$$

where f is the cross-sectional population mean, $f \equiv E[E_C[f_{i\tau}]]$; $\Omega_{i\tau}$ represents all the possible realizations of factor f for stock i at time τ; and $p(f_{is}, f_{iu})$ is the joint pdf of the factor realizations for stock i at times s and u, such that $\int p(f_{is}, f_{iu})\, d\Omega_{is}d\Omega_{iu} = 1$.

We can decompose the cross-sectional covariance in Eq. (27) into two parts:

$$\frac{1}{\mathcal{T}_1\mathcal{T}_2}\sum_{s\in I_2}\sum_{u\in I_1}\mathrm{Cov}_C(f_{is},f_{iu}) =$$

$$= \frac{1}{N}\frac{1}{\mathcal{T}_1\mathcal{T}_2}\sum_{s\in I_2,u\in I_1}\sum_{i=1}^{N}\int (f_{is}-f_i+f_i-f)(f_{iu}-f_i+f_i-f)p(f_{is},f_{iu})\,d\Omega_{is}d\Omega_{iu} \quad (28)$$

$$= \frac{1}{N}\frac{1}{\mathcal{T}_1\mathcal{T}_2}\sum_{s\in I_2,u\in I_1}\sum_{i=1}^{N}\Big[\int (f_{is}-f_i)(f_{iu}-f_i)p(f_{is},f_{iu})\,d\Omega_{is}d\Omega_{iu}+$$

$$+\int (f_i-f)(f_i-f)p(f_{is},f_{iu})\,d\Omega_{is}d\Omega_{iu}\Big], \quad (29)$$

where, since variation in factor loading of a specific stock i is independent of cross-sectional variation, the cross-terms of the form $\int (f_{is}-f_i)(f_i-f)p(f_{is},f_{iu})\,d\Omega_{is}d\Omega_{iu}$ cancel out.

In Equation (29), the first term is the sum of individual stock autocovariances from time s to time u; the second term is the cross-sectional variance of long-term factor means, σ^2_{f0}:

$$\frac{1}{\mathcal{T}_1\mathcal{T}_2}\sum_{s\in I_2}\sum_{u\in I_1}\mathrm{Cov}_C(f_{is},f_{iu}) =$$

$$= \frac{1}{\mathcal{T}_1\mathcal{T}_2}\sum_{s\in I_2}\sum_{u\in I_1}\mathrm{Cov}(f_{is},f_{iu}) + \sigma^2_{f0} \quad (30)$$

Similarly,

$$\frac{1}{\mathcal{T}_1^2}\sum_{s,u\in I_1}\mathrm{Cov}_C(f_{is},f_{iu}) = = \frac{1}{\mathcal{T}_1^2}\sum_{s,u\in I_1}\mathrm{Cov}(f_{is},f_{iu}) + \sigma^2_{f0} \quad (31)$$

For the AR(1) process in Eq. (26), the factor autocovariance $\mathrm{Cov}(f_{is}, f_{iu})$ takes the form:

$$\mathrm{Cov}(f_{is}, f_{iu}) = \frac{\phi^{|s-u|}\sigma^2}{1-\phi^2}, \tag{32}$$

where σ^2 is the variance of the disturbance term $\omega_{i\tau}$.

Using the time constant, $T_B \equiv -\frac{1}{\log \phi}$, we can write:

$$\mathrm{Cov}(f_{is}, f_{iu}) = \frac{e^{-|s-u|/T_B}}{1-e^{-2/T_B}}\sigma^2. \tag{33}$$

When the time-variation of factor f follows an AR(1) process, the autocovariance of the factor takes an exponentially-declining form, where the time constant T_B reflects the rate of decline – slower decline if T_B is larger and faster decline if T_B is smaller.

We therefore have:

$$\hat{\gamma}/\gamma = \frac{\frac{1}{\mathcal{T}_1\mathcal{T}_2}\sum_{s\in I_2}\sum_{u\in I_1}\mathrm{Cov}(f_{is}, f_{iu}) + \sigma^2_{f0}}{\frac{1}{\mathcal{T}_1^2}\sum_{s,u\in I_1}\mathrm{Cov}(f_{is}, f_{iu}) + \sigma^2_{f0}} = \frac{\mathcal{T}_1}{\mathcal{T}_2}\frac{\sum_{s\in I_2}\sum_{u\in I_1} e^{-|s-u|/T_B} + \eta^2}{\sum_{s,u\in I_1} e^{-|s-u|/T_B} + \eta^2}, \tag{34}$$

where

$$\eta^2 \equiv \frac{\sigma^2_{f0}}{\sigma^2}(1-e^{-2/T_B}). \tag{35}$$

We can now use Eq. (34) to derive the $\hat{\gamma}/\gamma$ for contemporaneous, predictive, and staggered regressions.

For contemporaneous regressions, under our simplifying assumptions outlined in Section 4, we have:

$$\hat{\gamma}/\gamma_{ctp} = 1. \tag{36}$$

In contemporaneous regressions, the estimation periods for factors and returns are the same, $I_1 = I_2$, and, in the expression for $\hat{\gamma}/\gamma$ in Eq. (23), the numerator and the denominator are equal.

For predictive and staggered regressions we can also derive closed-form expressions for $\hat{\gamma}/\gamma$, if we assume that the sampling time step in the time series of stock returns is much smaller than T_B (a realistic scenario). Under this assumption, we can apply a continuous approximation to sums over I_1 and I_2, which allows us to evaluate these sums analytically, as shown in the Online Appendix.

C Exceedance Correlations and the Distribution of Stock Returns

Exceedance correlations, defined in Section 3.1 (Eq. 5), are based on the Pearson correlation function, highly sensitive to outliers in the data. When an outlier is added to an otherwise jointly normal distribution of two random variables, as in Figure 5, their Pearson correlation is suppressed.

In this section, we demonstrate how the presence of outlier stock returns – jumps and crashes (i.e. negative jumps) – can produce the appearance of asymmetric dependence in exceedance correlations. Jumps and crashes are common in stock return time series; they are what makes the joint distribution of stock return leptokurtic and contribute to its non-zero skewness.

Consider a situation where the estimation window for exceedance correlations captures a single positive jump and no negative jumps; in this case, the expected value of the exceedance correlation will be negative, even if the jumps are uncorrelated with market moves. The geometry of exceedance correlation induces the linkage, as demonstrated in Figure 6: if the positive jump coincides with a negative market move, it falls outside the cutoff boundary and does not affect exceedance correlations. But if the positive jump coincides with a positive market move, it falls within the cutoff boundary and reduces ρ^+. If, in the *absence* of an outlier (a state of the world we denote as $\mathcal{N}$), the expected value of ρ^+ equals the expected value of ρ^-, $E[\rho^+|\mathcal{N}] = E[\rho^-|\mathcal{N}]$, in the presence of a positive outlier $\mathcal{O}+$, we have $E[\rho^+|\mathcal{O}+] < E[\rho^-|\mathcal{O}+]$ and $E[\rho^{exc}|\mathcal{O}+] = E[\rho^+ - \rho^-|\mathcal{O}+] < 0$. At the same time, the presence of a positive outlier in a time series of stock returns makes it more likely that the realized total return during the estimation window was positive, creating a statistical link between positive average returns and negative exceedance correlations – in the absence of any underlying asymmetric dependence. Similar logic leads to a similar conclusion for a negative outlier

$\mathcal{O}-$, and $E[\rho^{exc}|\mathcal{O}-] > 0$, while the returns for the period are likely to be negative.

In this section, we demonstrate how the statistical linkage between exceedance correlations and returns arises in the presence of jumps, resulting in an apparent premium associated with this factor in contemporaneous regressions. We use a simple market model of individual stock returns with jumps, in which the dependence of the stock returns on the market does not exhibit any asymmetric dependence. We confirm the results of the theoretical model using simulations in Appendix C.4.

C.1 Exceedance correlations in the presence of jumps

To model the impact of large idiosyncratic moves – or jumps – on exceedance correlations, we start with the market model, where the excess return of stock i, r_i, has a linear dependence on excess market return r_m:

$$r_{i\tau} = \beta_i r_{m\tau} + \omega_{i\tau}. \tag{37}$$

The disturbance term, $\omega_{i\tau}$, represents idiosyncratic noise. Each idiosyncratic move can be either a "regular" move with variance σ_{i0}^2 or – with probability q – a "jump" with variance B_i^2, such that $B_i^2 \gg \sigma_{i0}^2$. The idiosyncratic move process $\omega_{i\tau}$ can be written as:

$$\omega_{i\tau} = y_{i\tau}(1 - b_{i\tau}) + z_{i\tau} b_{i\tau}, \tag{38}$$

where $b_{i\tau}$ is a Bernoulli process (where $b_{i\tau} = 1$ with probability q and $b_{i\tau} = 0$ with probability $1-q$) that determines whether the idiosyncratic move of stock i at time t is a "regular" move $y_{i\tau}$ or a jump $z_{i\tau}$. The idiosyncratic process $\omega_{i\tau}$ is uncorrelated with the market $r_{m\tau}$ and the dependence of stock returns on market returns in this model is linear, with no asymmetric dependence.

The variances of random variables $y_{i,t}$ and $z_{i,t}$ are finite and equal $\mathrm{Var}(y_{i,t}) = \sigma_{i0}^2$ and $\mathrm{Var}(z_{i,t}) = B_i^2$. The jump process can have a non-zero mean, such that $E[z_{i,t}] = \bar{B}_i \geq 0$. The combined process is leptokurtic, with kurtosis $K > 3$ (for $q > 0$). If the mean jump is non-zero

($\bar{B}_i \neq 0$), the combined process is also skewed, with a non-zero skewness of the same sign as $\bar{B}_i$.[10]

To evaluate the impact of jumps on exceedance correlations, we first consider the impact of jumps on the Pearson correlation function between individual stock returns and the market return. For the process in Eq. (37), the Pearson correlation function is equal to:

$$\rho_{iB} = \beta_i \frac{\sigma_m}{\sigma_{iB}} = \frac{\sigma_{i0}}{\sigma_{iB}} \rho_{i0}, \tag{39}$$

where the subscript B indicates the presence of jumps and the subscript 0 indicates no jumps (i.e. $q = 0$); σ_m^2 is the variance of market returns, and σ_{iB}^2 is the variance of the individual stock returns in the presence of jumps. As shown in Appendix C.2, we can derive an expression for σ_{iB} as a function of σ_{i0}, q, $\bar{B}_i$ and B_i^2, so that we have (dropping the stock index i to streamline notation):

$$\rho_B = \frac{1}{\sqrt{(1-q) + q(1-q)\frac{\bar{B}^2}{\sigma_0^2} + q\frac{B^2}{\sigma_0^2}}} \rho_0 \equiv D_B \rho_0 < \rho_0, \tag{40}$$

When $B > \sigma_0$, the denominator in the fraction above is greater than 1 and, therefore, $\rho_B < \rho_0$. To simplify notation, we introduced D_B the fractional reduction of correlation in the presence of outliers. Figure 5 demonstrates the rapid decline of the Pearson correlation function as kurtosis of returns increases due to jumps.

With the expression for the Pearson correlation function in the presence of outliers, we quantify the impact of outliers on exceedance correlation estimates and contemporaneous estimates of average returns. We show that, in the presence of a (net) positive outlier, expected exceedance correlations are negative, but expected idiosyncratic returns are positive; in the presence of a negative outlier, the expected exceedance correlations are positive, and the expected idiosyncratic returns are negative. We then apply Bayes theorem to estimate expected exceedance correlations conditioned on positive and negative returns. Appendix C.3 provides the details of the analysis.

In Appendix C.3, we show that expected exceedance correlations given the sign of average

[10] If $\bar{B}_i > 0$, then the market model in Eq. (37) needs to include $\alpha_i = -q\bar{B}_i$: $r_{i,t} = \alpha_i + \beta_i r_{m,t} + \omega_{i,t}$, to compensate for the net positive impact of jumps. The addition of the constant term α_i does not affect the Pearson correlation function or exceedance correlations.

returns (p or n) are

$$E[\rho^{exc}|p] = -\rho_0(1 - D_B)\delta a < 0$$
$$E[\rho^{exc}|n] = \rho_0(1 - D_B)\delta a > 0, \tag{41}$$

where a represents the increased probability of a positive (negative) return in the presence of a positive (negative) outlier and δ is the number of outliers expected during estimation period T.

Because $D_B < 1$ and $\delta, a, \rho_0 > 0$, in the presence of jumps, exceedance correlations conditioned on positive returns are negative and vice versa.

The spread between exceedance correlations conditioned on positive and negative returns can be material: The probability δ of capturing an outlier in six months of daily data can be of the order unity (we assume it is 50%). The constant a measuring the likelihood that a total return during a period with a net positive jump is positive and vice versa can also be quite large. Stocks, particularly small stocks, experience 10-20% jumps quite often. If stock returns are normally distributed otherwise, with a 2% standard deviation, a 20% jump during a half-year measurement period makes it approximately 70% likely that the total return will be positive. In this case, $a = 0.4$. The decline in correlation due to kurtosis and skewness, $(1-D_B)$, can be 10-20%, as can be inferred from Figure 5 and empirical measures of kurtosis (Table 1). Putting these numbers together in Eq. (41), we get a 6% difference in ρ^{exc} conditioned on positive returns vs. negative idiosyncratic returns.

It is important to note that, if the distribution of idiosyncratic stock returns is persistently skewed and leptokurtic (as shown to be true for individual stock returns by, e.g. Boyer, Mitton, and Vorkink, 2009), the link between exceedance correlation and returns can persist even if the factor loadings and the average returns are estimated using non-overlapping periods. To understand this mechanical effect, consider a period with a realized positive idiosyncratic return. This period is much more likely to contain a positive jump than a negative jump. A positive jump is more likely to happen to a stock with a persistently positively skewed and leptokurtic distribution. The stock with a persistently positively skewed and leptokurtic distribution is likely to have a negative exceedance correlation. Through this Bayesian chain, a link between realized positive (negative) returns and negative (positive) exceedance correlation can persist even when the returns and the correlation are measured using non-overlapping periods.

C.2 Variance in the Presence of Outliers

In this section, we calculate the variance of the process described in Section C.1, Eqs. (37) and (38), where the "regular" process y_τ (we consider a single asset and drop the asset index i to streamline notation) has a mean y_0, variance σ_0^2, and the jump process has a mean $\bar{B}$, jump variance B^2. The probability that an idiosyncratic move is a jump is q.

The mean of the combined process is

$$\bar{y} = (1-q)y_0 + q\bar{B}. \tag{42}$$

The variance of $\mathcal{T}$ steps ($\mathcal{T} \gg 1$) of the combined process is:

$$\sigma^2 = \frac{1}{\mathcal{T}} \sum_{\tau=1}^{\mathcal{T}} (y_\tau - (1-q)y_0 - q\bar{B})^2 = \tag{43}$$

$$= \frac{1}{\mathcal{T}} \sum_{\tau \in 1-q} (y_\tau - y_0 + qy_0 - q\bar{B})^2 + $$
$$+ \frac{1}{\mathcal{T}} \sum_{\tau \in q} (y_\tau - y_0 + qy_0 - q\bar{B})^2, \tag{44}$$

where $\tau \in q$ (and $\tau \in 1-q$) is a shorthand for "step τ is (isn't) a jump."

We regroup the expressions in parentheses in different ways for regular steps and jump steps (the expressions in the square brackets below are equivalent):

$$\sigma^2 = \frac{1}{\mathcal{T}} \sum_{\tau \in 1-q} \left[(y_\tau - y_0) + q(y_0 - \bar{B})\right]^2 +$$
$$+ \frac{1}{\mathcal{T}} \sum_{\tau \in q} \left[(y_\tau - \bar{B}) + (1-q)(\bar{B} - y_0)\right]^2. \tag{45}$$

We now expand the quadratic expressions in brackets:

$$\sigma^2 = \frac{1}{\mathcal{T}} \sum_{\tau \in 1-q} \left[(y_\tau - y_0)^2 + 2q(y_0 - \bar{B})(y_\tau - y_0) + q^2(\bar{B} - y_0)^2 \right] +$$
$$+ \frac{1}{\mathcal{T}} \sum_{\tau \in q} \left[(y_\tau - \bar{B})^2 + 2(1-q)(y_\tau - \bar{B})(\bar{B} - y_0) + (1-q)^2(\bar{B} - y_0)^2 \right] \tag{46}$$

For a large enough $\mathcal{T}$, the second term in each bracket tends to 0. The first term in the brackets is the variance of the "regular" process and the jump process respectively. The last term in the brackets is a constant. Thus, we have:

$$\sigma^2 = (1-q)\left[\sigma_0^2 + q^2(\bar{B} - y_0)^2\right] +$$
$$+ q\left[B^2 + (1-q)^2(\bar{B} - y_0)^2\right]. \tag{47}$$

Simple manipulation leads to:

$$\sigma^2 = (1-q)\sigma_0^2 + q(1-q)(\bar{B} - y_0)^2 + qB^2. \tag{48}$$

Because $y_0 << \bar{B}$, we can simplify the expression further if we assume $y_0 = 0$:

$$\sigma^2 = (1-q)\sigma_0^2 + q(1-q)\bar{B}^2 + qB^2. \tag{49}$$

C.3 Expected Exceedance Correlations Given the Sign of Realized Returns

In this section, we derive the impact of jumps on exceedance correlations.

Exceedance correlations are affected if the jumps are within the boundaries of $\tilde{r}_i > c, \tilde{r}_m > c$ and $\tilde{r}_i < c, \tilde{r}_m < c$. In what follows, we consider the case of $c = 0$ for simplicity, and it is straight predictive to extend the analysis to $c > 0$.

In the presence of a positive jump (and no negative jumps), expected exceedance correlation is negative and equal to:

$$
\begin{aligned}
E[\rho^{exc}|O+] &= \frac{1}{2}(\rho_0^+ - \rho_0^-) + \frac{1}{2}\Big(\rho_0^+ \frac{1}{\sqrt{(1-q) + q(1-q)\frac{\tilde{B}^2}{4} + q\tilde{B}^2}} - \rho_0^-\Big) = \\
&= -\frac{1}{2}\Big(1 - \frac{1}{\sqrt{(1-q) + q(1-q)\frac{\tilde{B}^2}{4} + q\tilde{B}^2}}\Big)\rho_0^- = \\
&= -\frac{1}{2}(1 - D_B)\rho_0^- < 0, \qquad (50)
\end{aligned}
$$

where $O+$ denotes a scenario a positive jump; ρ_0^+ and ρ_0^- are the positive and negative exceedance correlations for the model with no jumps, and $\rho_0^+ = \rho_0^-$; $\tilde{B}$ is the size of the jump in units of standard deviation of the model without jumps; for a single jump, $q = 1/T$, where T is the number of return data points in the estimation window.

Similarly, in the presence of a negative jump (and no positive jumps), expected exceedance correlation is positive:

$$
\begin{aligned}
E[\rho^{exc}|O-] &= \frac{1}{2}\Big(1 - \frac{1}{\sqrt{(1-q) + q(1-q)\frac{\tilde{B}^2}{4} + q\tilde{B}^2}}\Big)\rho_0^- = \qquad (51) \\
&= \frac{1}{2}(1 - D_B)\rho_0^- > 0. \qquad (52)
\end{aligned}
$$

Since the distribution of idiosyncratic stock returns is positively skewed – i.e. tends to have more positive outliers – it is not surprising that the mean exceedance correlation estimated for a sample of stock returns is negative, as can be seen in the Statistical summary of factors provided in Table 1.

The geometry of exceedance correlation not only affects exceedance correlation estimation, but also the estimation of premia associated with this correlation. This effect arises because jumps affect not only exceedance correlations, but also average returns.

To demonstrate this linkage, we estimate the expected value of exceedance correlation during periods of *positive* returns, denoted by p, and during periods of *negative* returns, denoted by n.

For the state of the world s with a realized positive p or negative n return, we have:

$$E[\rho^{exc}|s] = E[\rho^{exc}|O+]\,P(O+|s) + E[\rho^{exc}|O-]\,P(O-|s) + \rho_0^{exc}P(N|s) \tag{53}$$

where $P(\bullet|s)$ is the probability of capturing a positive ($O+$), negative ($O-$), or no (N) outlier in the estimation window conditioned on s.

To estimate the probabilities of having captured a positive or negative jump when the realized returns were positive or negative, we use the following assumptions:

$$P(O+) = P(O-) = \delta \approx qT \tag{54}$$

$$P(p) = P(n) = \frac{1}{2} \tag{55}$$

$$P(p|N) = P(n|N) = \frac{1}{2} \tag{56}$$

$$P(p|O+) = P(n|O-) = \frac{1+a}{2} \tag{57}$$

$$P(p|O-) = P(n|O+) = \frac{1-a}{2} \tag{58}$$

We assume that the probability of capturing a jump in the estimation window is equal to δ, which is approximately equal to qT when $qT \ll 1$. We also assume that a realized positive return and a realized negative return are *unconditionally* equally likely and are equally likely if no jump is captured. Lastly, a positive jump makes a positive realized return more likely and a negative return makes a negative return more likely ($a > 0$). The converse is true as well.

We apply Bayes theorem to estimate probabilities of having captured a jump in the estimation window given a positive or negative return:

$$P(O+|p) = P(O-|n) = (1+a)\delta \tag{59}$$

$$P(O-|p) = P(O+|n) = (1-a)\delta. \tag{60}$$

The expected exceedance correlations given the sign of realized returns (p or n) then are

$$E[\rho^{exc}|p] = -\rho_0(1 - D_B)\delta a < 0$$
$$E[\rho^{exc}|n] = \rho_0(1 - D_B)\delta a > 0. \tag{61}$$

C.4 Impact of skewness and kurtosis on ρ^{exc} and J^{adj}: Simulations

To generalize the relationship between exceedance correlations and average returns in the presence of jumps and other deviations of the stock return distribution from multivariate normality (Eq. 41), we conduct a series of simulations. The simulations do not build in any asymmetric dependence (idiosyncratic risk distributions are the same for positive and negative market moves) or any exogenous risk premia.

To perform the simulations, we created a data-generating process for $N = 500$ shares with randomly assigned CAPM βs for each stock i, with $\beta_i \sim N(1, 0.5)$, consistently with empirical results. We used an unobserved macroeconomic variable r_0 to create a set of time series of individual stock returns $\{r_{i,t}\}_{t=1}^{T}$, where $T = 132$, to represent six months of daily data, such that

$$r_{i,t} = \beta_i r_0 + \varepsilon_{i,t}, \tag{62}$$

where $\varepsilon_{i,t}$ is an idiosyncratic noise process. The noise process $\varepsilon_{i,t}$ samples from a non-central t distribution with a variety of degrees of freedom from $\nu = 2$ to 100 and noncentrality parameters from $d = 0$ to 1 to match skewness and kurtosis levels in idiosyncratic returns observed empirically. We then created an initially equally-weighted index of stocks in our sample and used this index without rebalancing as the basis for estimating conditional systematic factor loadings.

To model individual stock returns, we used a continuously compounded process without conversion to daily compounding (in contrast to the β^- and β^+ simulations described in the Online Appendix), because daily compounding of a symmetric continuously compounded process induces a positively skewed return profile. In exceedance correlation simulations we aim to demonstrate the impact of long-tailed skewness and therefore use a process that allows us to turn skewness down to zero.

Table A2 summarizes the results. We used the non-central t distribution to model the stylized features of individual stock returns such as fat tails and positive skewness. We used a non-central t distribution with degrees of freedom $\nu = 4$ and non-centrality parameter $d = 0.3$ to match summary statistical features of return time series. Using the simulated return time series, we estimated factor loadings and average returns, and ran Fama-MacBeth regressions to estimate "premia" associated with factors in the simulation. We then compared the results of simulated regressions for the realistic $\nu = 4$ ($d = 0.3$) return distribution with those for a symmetric leptokurtic distribution $\nu = 4$ ($d = 0$) and the normal distribution.

When the distribution of idiosyncratic returns is leptokurtic ($\nu = 4$), there is a large and highly statistically significant negative slope between exceedance correlations (ρ^{exc} or J^{adj}) and average returns [Table A2, panel (a)]. When the distribution of idiosyncratic return is not only leptokurtic, but also skewed ($d = 0.3$), the mean values of ρ^{exc} or J^{adj} are negative [Table A2, panel (b)], because the positively skewed distribution of returns contains positive jumps that suppress the values of ρ^{exc} or J^{adj}. For the (approximately) symmetric ($d = 0$) idiosyncratic return distribution, the mean values of ρ^{exc} or J^{adj} are close to zero, because positive and negative jumps are now in balance. However, the highly statistically significant slope between returns and ρ^{exc} or J^{adj} remains, as would be expected from theory developed in Section C.1. When the distribution of idiosyncratic returns is normal – i.e. there are no jumps – the coefficient between returns and exceedance correlation goes to zero.

Lastly, even though the regression coefficient between exceedance correlations and returns is artificially induced by outliers, controlling for kurtosis and skewness in regressions is insufficient to eliminate it.

C.5 Exceedance correlation as a measure of deviation of stock return distribution from normal

Exceedance correlations have been used widely to study asymmetric dependence phenomena, but because of their sensitivity to other deviations of stock return distributions from multivariate normality, they could potentially be re-purposed for other applications. Investment managers could use exceedance correlations to screen past stock returns for influence of large jumps; corporate boards and CFOs could diagnose whether issued option-like instruments would be accurately priced by Black-Scholes equation.

Table A1: Contemporaneous, staggered, and predictive cross-sectional regressions of average returns on estimated risk, size, book-to-market, and momentum factor loadings. This table compares the results of three types of Fama-MacBeth regressions: Contemporaneous, in which returns and factor loadings for each month t are estimated using the same data periods $(t - T_1, t]$, where $T = 6$ months; Staggered regressions, in which returns and factor loadings are estimated on staggered months as described in Section 2 (where, in each estimation period $(t - \tilde{T}_1, t]$ for each t, where $\tilde{T}_1$ is 18 months, months $t - 15$, $t - 12$, $t - 9$, $t - 6$, $t - 3$, and t are used for realized return estimation and months $t - 16$, $t - 13$, $t - 10$, $t - 7$, $t - 4$, and $t - 1$ are used for factor loading estimation); and Predictive regressions, where, for each time t, factor loadings are estimated using return time series $[t - T_1, t)$ and average returns are estimated for either a single month $t + 1$ or for the period $(t, t + T_1]$, where T_1 is six months. Size and book-to-market factors are estimated using classic Fama and French (1992) method. Size factor is the logarithm of issuing firm market capitalization measured in December of year preceding the first date of return estimation period. Book-to-market factor is based on latest reported annual book equity preceding the first date of return estimation period. If return estimation period starts before July 1 of year Y_0, then size and book-to-market are estimated using data from $Y_0 - 2$. Medium term momentum factor, Past Ret_{-T}, is based on the period prior and equal in length to the return estimation period. Short-term momentum, Ret_{-1}, is based on one month preceding the estimation period (for Staggered regressions, Ret_{-1} is estimated on staggered basis, contemporaneously to systematic risk factors). Panels (1) through (8) report results of separate multiple regression models. Reported t-statistics are adjusted for overlapping periods using the Newey-West method. Average $Adj R^2$ of cross-sectional regressions are reported for each estimation model. The study used daily return time series of stocks traded on NYSE/Amex/Nasdaq during the period between 1963 and 2018, with sharecodes 10 and 11.

Models	Regression Types							
	Contemporaneous		Staggered		Predictive – 1mo		Predictive – 6mo	
	Coefficient	t-Statistic	Coefficient	t-Statistic	Coefficient	t-Statistic	Coefficient	t-Statistic
Panel (1)								
Int	0.191	5.33	0.113	3.65	0.217	4.66	0.228	4.74
β	0.055	3.04	0.027	2.09	0.011	0.71	0.008	0.68
Mcap	-0.027	-5.71	-0.014	-3.79	-0.025	-4.63	-0.027	-4.60
B/M	0.020	2.98	0.015	2.57	0.012	1.40	0.017	2.08
Past Ret_{-T}	0.029	2.73	0.014	2.35	0.043	2.90	0.050	3.41
Ret_{-1}	-0.009	-5.41	-0.024	-2.97	-0.049	-7.09	-0.008	-4.04
Adj R^2		18.7 %		31.1 %		15.4 %		16.2 %
Panel (2)								
Int	0.171	4.74	0.104	3.36	0.207	4.40	0.219	4.56
β^-	0.040	3.57	0.023	2.91	0.006	0.58	0.006	0.80
Mcap	-0.023	-5.49	-0.012	-3.50	-0.023	-4.63	-0.025	-4.59
B/M	0.023	3.17	0.015	2.52	0.013	1.46	0.019	2.21
Past Ret_{-T}	0.031	2.80	0.014	2.37	0.041	2.65	0.048	3.25
Ret_{-1}	-0.009	-5.14	-0.022	-2.74	-0.048	-7.01	-0.007	-4.00
Adj R^2		17.5 %		30.4 %		14.6 %		15.6 %

Table A1: (Continued.)

Models	Regression Types							
	Contemporaneous		Staggered		Predictive – 1mo		Predictive – 6mo	
	Coefficient	t-Statistic	Coefficient	t-Statistic	Coefficient	t-Statistic	Coefficient	t-Statistic
Panel (3)								
Int	0.213	5.33	0.131	3.86	0.214	4.32	0.224	4.47
β^+	0.010	1.19	0.007	1.10	0.009	1.16	0.007	1.25
Mcap	-0.025	-5.54	-0.014	-3.65	-0.024	-4.68	-0.026	-4.57
B/M	0.019	2.68	0.014	2.33	0.014	1.49	0.019	2.23
Past Ret_{-T}	0.035	3.08	0.015	2.26	0.039	2.53	0.050	3.25
Ret_{-1}	-0.009	-5.01	-0.022	-2.67	-0.046	-6.80	-0.008	-3.97
Adj R^2		16.8 %		30.3 %		14.5 %		15.4 %
Panel (4)								
Int	0.177	4.92	0.108	3.51	0.210	4.49	0.221	4.61
β^-	0.042	4.16	0.023	3.57	0.003	0.36	0.004	0.71
β^+	-0.000	-0.08	0.000	0.01	0.005	0.93	0.003	0.72
Mcap	-0.023	-5.45	-0.013	-3.59	-0.024	-4.62	-0.026	-4.59
B/M	0.022	3.13	0.014	2.43	0.013	1.48	0.018	2.19
Past Ret_{-T}	0.028	2.66	0.014	2.36	0.041	2.71	0.049	3.31
Ret_{-1}	-0.009	-5.34	-0.023	-2.92	-0.048	-7.03	-0.008	-4.06
Adj R^2		18.4 %		31.0 %		15.1 %		16.0 %
Panel (5)								
Int	0.181	5.12	0.108	3.53	0.219	4.72	0.229	4.78
β	0.058	3.22	0.029	2.28	0.010	0.63	0.008	0.64
$\beta^+ - \beta^-$	-0.017	-3.60	-0.009	-3.06	0.002	0.54	0.000	0.13
Mcap	-0.025	-5.63	-0.013	-3.72	-0.025	-4.68	-0.027	-4.64
B/M	0.021	3.09	0.015	2.50	0.012	1.40	0.017	2.07
Past Ret_{-T}	0.026	2.48	0.013	2.33	0.044	2.97	0.051	3.48
Ret_{-1}	-0.009	-5.57	-0.024	-3.01	-0.050	-7.15	-0.008	-4.14
Adj R^2		19.3 %		31.5 %		15.6 %		16.5 %
Panel (6)								
Int	0.186	5.19	0.110	3.58	0.219	4.71	0.229	4.77
β	0.058	3.23	0.030	2.24	0.011	0.70	0.008	0.65

Table A1: (Continued.)

Models	Regression Types							
	Contemporaneous		Staggered		Predictive – 1mo		Predictive – 6mo	
	Coefficient	t-Statistic	Coefficient	t-Statistic	Coefficient	t-Statistic	Coefficient	t-Statistic
Cosk	-0.084	-5.31	-0.040	-3.34	0.006	0.43	0.004	0.38
Mcap	-0.026	-5.67	-0.013	-3.76	-0.025	-4.69	-0.027	-4.64
B/M	0.021	2.99	0.014	2.44	0.012	1.39	0.017	2.07
Past Ret_{-T}	0.027	2.56	0.013	2.33	0.043	2.88	0.050	3.42
Ret_{-1}	-0.009	-5.51	-0.024	-3.00	-0.049	-7.13	-0.007	-4.03
Adj R^2		19.1 %		31.3 %		15.5 %		16.4 %
Panel (7)								
Int	0.133	4.06	0.115	3.71	0.219	4.69	0.231	4.80
β	0.052	2.94	0.028	2.13	0.011	0.68	0.008	0.66
ρ^{exc}	-0.304	-9.65	0.011	1.70	0.000	0.03	-0.001	-0.19
Mcap	-0.022	-5.07	-0.014	-3.84	-0.025	-4.66	-0.027	-4.63
B/M	0.020	2.98	0.015	2.54	0.012	1.40	0.017	2.06
Past Ret_{-T}	0.032	3.02	0.013	2.31	0.044	2.92	0.051	3.40
Ret_{-1}	-0.009	-5.43	-0.023	-2.92	-0.050	-7.14	-0.008	-3.99
Adj R^2		21.1 %		31.2 %		15.4 %		16.3 %
Panel (8)								
Int	0.163	4.76	0.114	3.67	0.219	4.69	0.230	4.77
β	0.054	3.01	0.028	2.12	0.011	0.69	0.008	0.66
J^{adj}	-0.007	-9.51	0.000	1.19	0.000	0.23	0.000	0.08
Mcap	-0.024	-5.42	-0.014	-3.82	-0.025	-4.65	-0.027	-4.62
B/M	0.020	2.97	0.015	2.55	0.012	1.40	0.017	2.08
Past Ret_{-T}	0.030	2.88	0.014	2.33	0.044	2.94	0.051	3.41
Ret_{-1}	-0.009	-5.43	-0.024	-2.97	-0.050	-7.11	-0.008	-4.00
Adj R^2		19.9 %		31.2 %		15.4 %		16.3 %
Panel (9)								
Int	0.160	4.70	0.108	3.56	0.219	4.72	0.230	4.80
β	0.056	3.12	0.030	2.31	0.010	0.61	0.007	0.62
$\beta^{+} - \beta^{-}$	-0.000	-0.02	-0.011	-3.36	0.002	0.58	0.001	0.23
J^{adj}	-0.007	-9.41	0.001	2.91	0.000	0.05	-0.000	-0.02

Table A1: (Continued.)

Models	Regression Types							
	Contemporaneous		Staggered		Predictive – 1mo		Predictive – 6mo	
	Coefficient	t-Statistic	Coefficient	t-Statistic	Coefficient	t-Statistic	Coefficient	t-Statistic
Mcap	-0.024	-5.43	-0.013	-3.75	-0.025	-4.68	-0.027	-4.65
B/M	0.021	3.08	0.015	2.49	0.012	1.41	0.017	2.07
Past Ret$_{-T}$	0.028	2.69	0.013	2.29	0.045	3.01	0.051	3.49
Ret$_{-1}$	-0.009	-5.52	-0.024	-2.92	-0.050	-7.16	-0.008	-4.12
Adj R^2		20.4 %		31.5 %		15.6 %		16.5 %
Panel (10)								
Int	0.165	4.82	0.111	3.61	0.219	4.71	0.230	4.79
β	0.054	3.05	0.030	2.28	0.011	0.70	0.008	0.64
Cosk	0.062	3.93	-0.052	-3.92	0.004	0.25	0.003	0.33
J^{adj}	-0.007	-9.50	0.001	3.74	0.000	0.32	0.000	0.36
Mcap	-0.025	-5.50	-0.014	-3.78	-0.025	-4.69	-0.027	-4.64
B/M	0.020	2.96	0.014	2.42	0.012	1.40	0.017	2.07
Past Ret$_{-T}$	0.030	2.83	0.013	2.30	0.044	2.92	0.051	3.42
Ret$_{-1}$	-0.009	-5.41	-0.023	-2.84	-0.049	-7.13	-0.007	-4.01
Adj R^2		20.3 %		31.4 %		15.5 %		16.4 %

Table A2: Apparent risk premia of exceedance correlations: Simulation results. This table presents the results of contemporaneous regressions of average returns on estimated factor loadings performed on simulated data. Part (a) summarizes regression results; part (b) presents a summary of factor estimates. To demonstrate the impact of outlier stock returns on exceedance correlations and measured risk premia, we model idiosyncratic risk with "long tails" using the non-central t distribution with degrees of freedom ν and noncentrality parameter d. There is no asymmetric dependence built into the model. We run panels of 200 market simulations for 500 stocks.

(a) Regression results

		$\nu = 4 (d = 0.3)$		$\nu = 4 (d = 0)$		Normal	
Panel		Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat
I	Int	-0.057	-9.1	-0.003	-0.4	0.003	0.45
	β	0.107	21.2	0.085	17.0	0.074	14.6
	ρ^{exc}	-0.204	-35.8	-0.224	-39.7	-0.002	-0.3
II	Int	-0.053	-8.5	-0.003	-0.5	0.003	0.5
	β	0.109	21.5	0.086	17.1	0.074	14.6
	J^{adj}	-0.014	-29.6	-0.016	-33.8	-0.000	-0.4
III	Int	-0.042	-6.7	-0.003	0.0	0.003	0.4
	β	0.115	22.5	0.087	17.2	0.074	14.6
	Cosk	0.055	1.7	0.056	1.8	-0.010	-0.3
IV	Int	-0.106	-15.7	0.007	1.1	0.046	1.6
	β	0.088	18.6	0.075	16.0	0.074	14.6
	ρ^{exc}	-0.084	-15.3	-0.105	-19.2	-0.002	-0.3
	Skew	0.308	82.5	0.292	85.2	-0.001	-0.0
	Kurt	-0.006	-11.1	0.000	0.3	-0.014	-1.5
V	Int	-0.105	-15.5	0.007	1.1	0.046	1.6
	β	0.089	18.7	0.076	16.0	0.074	14.6
	J^{adj}	-0.005	-11.5	-0.007	-15.5	-0.000	-0.4
	Skew	0.313	84.3	0.297	87.2	-0.001	-0.1
	Kurt	-0.006	-11.3	0.000	0.2	-0.014	-1.5

Table A2: Simulation results for exceedance correlations (Continued).

(b) Statistical summary

	$\nu = 4(d = 0.3)$		$\nu = 4(d = 0)$		Normal	
	Mean	Std	Mean	Std	Mean	Std
Ret	0.068	0.89	0.081	0.89	0.075	0.88
β	0.95	0.77	0.96	0.78	0.97	0.77
Idio	0.03	0.02	0.03	0.01	0.03	0.02
Skew	0.34	1.06	0.00	1.10	0.00	0.21
Kurt	7.4	7.2	7.4	7.1	3.0	0.4
ρ^{exc}	-0.12	0.69	-0.01	0.7	0.03	0.68
J^{adj}	-1.21	8.1	-0.04	8.1	0.24	7.7
Cosk	-0.00	0.13	0.00	0.13	0.00	0.12

OA "Are Conditional Factors Priced? Characterizing Risk Premia of Conditional Systematic Risk Factors with Staggered Regressions": Online Appendix

OA.1 Monte-Carlo Simulations

OA.1.1 Construction

To perform the simulations described in Section 3.6, we created a data-generating process for $N = 500$ shares with randomly assigned CAPM βs for each stock i, as described in Appendix C.4.

Because non-linearity is a potentially important driver of regression results for factors measuring non-linearity of the dependence of stock returns and market returns, we experimented with different modes of compounding: continuous compounding, daily compounding, and compounding where the data generating process uses continuous compounding, but the returns are converted into daily compounding for estimation of factor loadings. None of these modifications had a material impact on factor estimates.

The continuously compounded return process combined with conversion into daily returns before beta estimation produced the closest match to empirical findings. With continuous compounding alone, our simulations resulted in a dispersion of β^- and of β^+ much higher than that seen empirically (Table OA2), even when levels of idiosyncratic risk, skewness and kurtosis were closely matched. Daily compounding resulted in a number problems at the data generation step (outside of the scope of this paper). We therefore recommend the combined method of compounding for simulations.

To match levels of kurtosis and skewness in empirical data, we used the non-central t distribution with degrees of freedom $\nu = 4$ and noncentrality parameter $d = 0.3$ for simulations with continuous compounding and $d = 0.1$ for those with combined compounding. In Tables OA1 and OA2 we use the shorthand D.C. for combined compounding to indicate that daily compounded returns are used for estimation of factor loadings. In column (e) of these tables, we also introduce heteroskedasticity and make idiosyncratic risk slightly larger conditionally on positive market moves than on negative market moves (with realized idiosyncratic risk circa 0.025 and 0.027 respectively).

Other than the natural non-linearity arising from compounding and a slight asymmetry in idiosyncratic risk introduced in column (e) only, no other asymmetry is modeled in the reported simulations. The dependence of simulated stock returns on the market is noisy, but perfectly linear and symmetric.

OA.1.2 Results

We report simulation results in Table OA1. The table provides a summary of mean factor premia and their Newey-West adjusted t-stats – the latter, in parentheses. A quick scan through the table reveals that, in simulations, β^- and β^+ are associated with similar risk premia, even when the premia are estimated in a multiple regression. Not only is the premium associated with β^- not greater than that for β^+, but, in simulation, the situation is the reverse: β^+ "earns" a higher premium than β^- and at a higher level of statistical significance.

As may be expected from the theoretical model in Section C.1, simulation parameters, particularly those related to skewness and kurtosis of idiosyncratic risk, did affect the estimate and apparent economic significance of exceedance correlations and the related metric J^{adj}.

OA.2 Continuous Approximation Integrals to Estimate Factor Covariances

Starting with predictive regressions, we first evaluate the sum in the numerator of Eq. (23):

$$\frac{1}{\mathcal{T}_1\mathcal{T}_2}\sum_{s\in I_2}\sum_{u\in I_1} e^{-|s-u|/T_B} \approx \frac{1}{TT_1}\int_{t-T}^{t}\int_{t}^{t+T_1} e^{-|s-u|/T_B}\,du\,ds \tag{63}$$

$$= \frac{1}{TT_1}\int_{-T}^{0}\int_{0}^{T_1} e^{-|x-y|/T_B}\,dx\,dy \tag{64}$$

$$= \frac{1}{TT_1}\int_{-T}^{0} e^{x/T_B}dx\int_{0}^{T_1} e^{-y/T_B}dy \tag{65}$$

$$= \frac{T_B^2}{TT_1}(1-e^{-T/T_B})(1-e^{-T_1/T_B}) \tag{66}$$

We then evaluate the sum in the denominator:

$$\frac{1}{\mathcal{T}_1^2}\sum_{s,u\in I_1} e^{-|s-u|/T_B} \approx \frac{1}{T^2}\int_{t-T}^{t}\int_{t-T}^{t} e^{-|s-u|/T_B}\,du\,ds \tag{67}$$

$$= \frac{1}{T^2}\int_0^T\int_0^T e^{-|x-y|/T_B}\,dx\,dy \tag{68}$$

$$= \frac{1}{T^2}\Big[\int_0^T dx\int_0^x e^{-(x-y)/T_B} dy + \int_0^T dx\int_x^T e^{(x-y)/T_B} dy\Big] \tag{69}$$

$$= 2\frac{T_B}{T}\big[1 - \frac{T_B}{T}(1-e^{-T/T_B})\big] \tag{70}$$

Thus, for predictive regressions we have:

$$\hat{\gamma}/\gamma_{prd} = \frac{\frac{T_B^2}{TT_1}(1-e^{-T/T_B})(1-e^{-T_1/T_B}) + \eta^2}{2\frac{T_B}{T}\big[1-\frac{T_B}{T}(1-e^{-T/T_B})\big] + \eta^2}. \tag{71}$$

For staggered regressions, we have, in the numerator:

$$\frac{1}{\mathcal{T}_1\mathcal{T}_2}\sum_{s\in I_2}\sum_{u\in I_1} e^{-|s-u|/T_B} \approx \frac{1}{T^2}\sum_{k_1=1}^{K}\sum_{k_2=1}^{K}\int_{t-3k_1T_S+T_S}^{t-3k_1T_S+2T_S}\int_{t-3k_2T_S+2T_S}^{t-3k_2T_S+3T_S} e^{-|s-u|/T_B}\,du\,ds \tag{72}$$

$$= \frac{1}{T^2}\sum_{k_1=1}^{K}\sum_{k_2=1}^{K}\int_0^{T_S}\int_0^{T_S} e^{-|x-y+3(k_1-k_2)T_S+T_S|/T_B}\,dx\,dy \tag{73}$$

$$= \frac{1}{T^2}\Big[\sum_{k_1=1}^{K}\sum_{k_2=1}^{k_1} e^{-3(k_1-k_2)T_S/T_B}\int_0^{T_S}\int_0^{T_S} e^{-(x-y+T_S)/T_B}\,dx\,dy +$$

$$+\sum_{k_1=1}^{K-1}\sum_{k_2=k_1+1}^{K} e^{3(k_1-k_2)T_S/T_B}\int_0^{T_S}\int_0^{T_S} e^{(x-y+T_S)/T_B}\,dx\,dy\Big] \tag{74}$$

$$= \frac{T_B^2}{T^2}\Big[\big(K + \sum_{k=1}^{K-1}(K-k)e^{-3kT_S/T_B}\big)\big(1-e^{-T_S/T_B}\big)^2 +$$

$$+\sum_{k=1}^{K-1}(K-k)e^{-3kT_S/T_B}\big(e^{T_S/T_B}-1\big)^2\Big] \tag{75}$$

$$= \frac{T_B^2}{T^2}\Big[K + (1+e^{2T_S/T_B})\sum_{k=1}^{K-1}(K-k)e^{-3kT_S/T_B}\Big]\big(1-e^{-T_S/T_B}\big)^2 \tag{76}$$

where $K \equiv T/T_S$ and T_S is the staggered estimation subperiod.

Similarly, in the denominator, we have:

$$\frac{1}{\mathcal{T}_1^2}\sum_{s,u\in I_1} e^{-|s-u|/T_B} \approx \frac{1}{T^2}\sum_{k_1=1}^{K}\sum_{k_2=1}^{K}\int_{t-3k_1T_S+T_S}^{t-3k_1T_S+2T_S}\int_{t-3k_2T_S+T_S}^{t-3k_2T_S+2T_S} e^{-|s-u|/T_B}\,du\,ds \tag{77}$$

$$= \frac{1}{T^2}\sum_{k_1=1}^{K}\sum_{k_2=1}^{K}\int_0^{T_S}\int_0^{T_S} e^{-|x-y+3(k_1-k_2)T_S|/T_B}\,dx\,dy \tag{78}$$

$$\begin{aligned}&= \frac{1}{T^2}\Big[K\int_0^{T_S}\int_0^{T_S} e^{-|x-y|/T_B}\,dx\,dy + \\ &+ \sum_{k=1}^{K-1}(K-k)\int_0^{T_S}\int_0^{T_S} e^{-(x-y+3kT_S)/T_B}\,dx\,dy + \\ &+ \sum_{k=1}^{K-1}(K-k)\int_0^{T_S}\int_0^{T_S} e^{(x-y-3kT_S)/T_B}\,dx\,dy\Big]\end{aligned} \tag{79}$$

$$\begin{aligned}&= \frac{T_B^2}{T^2}\Big[\big(2K\frac{T_S}{T_B}[1-\frac{T_B}{T_S}(1-e^{-T_S/T_B})]+ \\ &+ 2e^{T_S/T_B}(1-e^{-T_S/T_B})^2\sum_{k=1}^{K-1}(K-k)e^{-3kT_S/T_B}\Big]\end{aligned} \tag{80}$$

For staggered regressions, we have for $\hat{\gamma}/\gamma$:

$$\hat{\gamma}/\gamma_{stg} = \frac{\frac{T_B^2}{T^2}\Big[K+(1+e^{2T_S/T_B})\sum_{k=1}^{K-1}(K-k)e^{-3kT_S/T_B}\Big]\big(1-e^{-T_S/T_B}\big)^2+\eta^2}{\frac{T_B^2}{T^2}\Big[\big(2K\frac{T_S}{T_B}[1-\frac{T_B}{T_S}(1-e^{-T_S/T_B})]+2e^{T_S/T_B}(1-e^{-T_S/T_B})^2\sum_{k=1}^{K-1}(K-k)e^{-3kT_S/T_B}\Big]+\eta^2}. \tag{81}$$

OA Robustness Checks

In this section, we report regression results controlled for size, book-to-market ratio, momentum factors (short- and medium-term), skewness, and kurtosis.

A note on the skewness premium observed in Table OA4. The premium not only becomes statistically less significant in the absence of data overlap, but also reverses sign in staggered regressions. Staggered regressions appear to capture a reversal effect, which takes places on the order of a few days to a few weeks (the two-day skip between months used for factor and return estimation ensures the effect is not due to very near-term reversals or non-synchronous trading), again pointing to interesting dynamics at this time scale.

Table OA1: Simulation results for β^- and β^+

This table presents results of two types of Monte-Carlo simulations: panels (a)-(c) use continuously compounded returns to create return time series and to estimate β's; in (d) and (e), simulated continuously compounded returns are converted into daily compounded returns before β estimation. We model idiosyncratic risk using the non-central t distribution with degrees of freedom ν and noncentrality parameter d. We run panels of 200 market simulations for 500 stocks.

Panel		(a) $\nu = 4$ $(d = 0.3)$	(b) $\nu = 4$ $(d = 0)$	(c) Normal	(d) $\nu = 4$ D.C.	(e) $\nu = 4$ D.C/A.D.
I	Int	-0.010	0.000	-0.005	0.066	0.017
		(-1.6)	(0.0)	(-0.8)	(10.6)	(2.6)
	β	0.110	0.080	0.088	0.079	0.161
		(21.8)	(16.1)	(17.1)	(14.6)	(28.5)
II	Int	0.082	0.065	0.059	0.104	0.103
		(18.1)	(14.5)	(12.9)	(20.5)	(19.1)
	β^-	0.013	0.013	0.022	0.041	0.074
		(5.9)	(5.7)	(9.3)	(10.2)	(17.3)
III	Int	0.070	0.056	0.066	0.098	0.071
		(15.6)	(12.6)	(14.5)	(19.0)	(13.4)
	β^+	0.026	0.022	0.014	0.047	0.1056
		(11.8)	(9.7)	(6.3)	(11.7)	(25.3)
IV	Int	0.058	0.046	0.046	0.066	0.018
		(11.6)	(9.2)	(9.1)	(10.5)	(2.7)
	β^-	0.013	0.012	0.021	0.036	0.062
		(5.6)	(5.4)	(9.2)	(9.0)	(14.3)
	β^+	0.026	0.022	0.014	0.043	0.098
		(11.7)	(9.5)	(6.1)	(10.6)	(23.3)
V	Int	-0.010	0.000	-0.005	0.066	0.019
		(-1.5)	(0.0)	(-0.8)	(10.6)	(2.9)
	β	0.110	0.081	0.088	0.079	0.159
		(21.6)	(16.1)	(17.1)	(14.6)	(28.1)
	$\beta^- - \beta^+$	-0.006	-0.004	0.004	-0.002	-0.013
		(-3.9)	(-2.6)	(2.1)	(-0.7)	(-3.9)

Table OA2: Simulation results for β^- and β^+: Statistical summary

Means and standard deviations (the latter – in brackets) of βs estimated in simulations described in Table OA1

	(a) $\nu = 4$ ($d = 0.3$)	(b) $\nu = 4$ ($d = 0$)	(c) Normal	(d) $\nu = 4$ D.C.	(e) $\nu = 4$ D.C/A.D.
β	0.96	0.97	0.97	1.00	1.00
	[0.78]	[0.78]	[0.77]	[0.58]	[0.58]
β^-	0.96	0.96	0.97	0.99	1.00
	[1.74]	[1.76]	[1.75]	[0.79]	[0.77]
β^+	0.95	0.96	0.96	1.00	1.00
	[1.78]	[1.75]	[1.76]	[0.78]	[0.79]

Table OA3: Contemporaneous, staggered, and predictive cross-sectional regressions of average returns on estimated risk, size, and book-to-market factor loadings. This table compares the results of three types of Fama-MacBeth regressions: Contemporaneous, in which returns and factor loadings for each month t are estimated using the same data periods $[t-T_1, t)$, where $T_1 = 6$ months; Staggered regressions, in which returns and factor loadings are estimated on staggered months as described in Section 2 (where, in each estimation period $[t-\tilde{T}_1, t)$ for each t, where $\tilde{T}_1$ is 18 months, months $t-16$, $t-13$, $t-10$, $t-7$, $t-4$, and $t-1$ are used for realized return estimation and months $t-17$, $t-14$, $t-11$, $t-8$, $t-5$, and $t-2$ are used for factor loading estimation); and Predictive regressions, where, for each time t, factor loadings are estimated using return time series $[t-T_1, t)$ and average returns are estimated either for a single month $t+1$ or for the period $(t, t+T_1]$, where T_1 is six months. Panels (1) through (8) report results of separate single- or multiple regression models. Reported t-statistics are adjusted for overlapping periods using the Newey-West method. Average AdjR^2 of cross-sectional regressions are reported for each estimation model. The study used daily return time series of stocks traded on NYSE/Amex/Nasdaq during the period between 1963 and 2018, with sharecodes 10 and 11.

	Regression Types							
Models	Contemporaneous		Staggered		Predictive – 1mo		Predictive – 6mo	
	Coefficient	t-Statistic	Coefficient	t-Statistic	Coefficient	t-Statistic	Coefficient	t-Statistic
Panel (1)								
Int	0.193	5.22	0.130	3.91	0.229	4.91	0.243	4.74
β	0.064	3.40	0.040	2.73	0.023	1.42	0.021	1.43
Mcap	-0.027	-5.60	-0.018	-4.27	-0.026	-4.85	-0.029	-4.55
B/M	0.022	3.12	0.023	3.51	0.012	1.36	0.019	2.10
Adj R^2		18.3 %		11.2 %		13.5 %		14.7 %
Panel (2)								
Int	0.176	4.65	0.120	3.56	0.222	4.67	0.234	4.56
β^-	0.048	3.97	0.031	3.41	0.016	1.53	0.016	1.74
Mcap	-0.023	-5.36	-0.016	-3.98	-0.025	-4.83	-0.027	-4.52
B/M	0.024	3.25	0.024	3.53	0.014	1.50	0.021	2.23
Adj R^2		16.9 %		10.0 %		12.6 %		14.0 %
Panel (3)								
Int	0.230	5.42	0.157	4.10	0.236	4.66	0.248	4.52
β^+	0.016	1.74	0.012	1.72	0.015	1.91	0.014	2.02
Mcap	-0.026	-5.49	-0.019	-4.11	-0.027	-4.87	-0.028	-4.50
B/M	0.021	2.79	0.022	3.37	0.014	1.50	0.021	2.20
Adj R^2		16.2 %		9.8 %		12.5 %		13.7 %
Panel (4)								

Table OA3: (Continued.)

Models	Contemporaneous		Staggered		Predictive – 1mo		Predictive – 6mo	
	Coefficient	t-Statistic	Coefficient	t-Statistic	Coefficient	t-Statistic	Coefficient	t-Statistic
Int	0.181	4.85	0.125	3.75	0.224	4.76	0.237	4.63
β^-	0.049	4.59	0.031	4.04	0.011	1.20	0.013	1.65
β^+	0.001	0.16	0.002	0.32	0.008	1.32	0.005	1.07
Mcap	-0.024	-5.34	-0.017	-4.09	-0.025	-4.82	-0.028	-4.54
B/M	0.024	3.28	0.023	3.54	0.014	1.49	0.020	2.20
Adj R^2		18.0 %		10.9 %		13.3 %		14.4 %
Panel (5)								
Int	0.181	4.98	0.123	3.79	0.230	4.94	0.244	4.78
β	0.067	3.61	0.042	2.91	0.022	1.33	0.021	1.40
$\beta^+ - \beta^-$	-0.019	-3.95	-0.012	-3.53	-0.000	-0.05	-0.003	-0.70
Mcap	-0.026	-5.50	-0.018	-4.24	-0.026	-4.88	-0.029	-4.59
B/M	0.023	3.28	0.023	3.52	0.012	1.36	0.018	2.08
Adj R^2		19.0 %		11.6 %		13.8 %		15.0 %
Panel (6)								
Int	0.186	5.06	0.126	3.85	0.229	4.92	0.243	4.76
β	0.067	3.59	0.042	2.88	0.023	1.41	0.021	1.41
Cosk	-0.090	-5.40	-0.047	-3.59	-0.002	-0.13	-0.007	-0.66
Mcap	-0.026	-5.54	-0.018	-4.28	-0.026	-4.88	-0.029	-4.57
B/M	0.022	3.14	0.023	3.43	0.012	1.36	0.019	2.08
Adj R^2		18.7 %		11.5 %		13.7 %		14.9 %
Panel (7)								
Int	0.134	3.96	0.134	4.01	0.231	4.94	0.243	4.75
β	0.060	3.29	0.040	2.76	0.023	1.42	0.021	1.41
ρ^{exc}	-0.306	-9.63	0.023	2.72	0.002	0.17	-0.018	-2.19
Mcap	-0.022	-4.95	-0.019	-4.34	-0.027	-4.89	-0.028	-4.55
B/M	0.022	3.11	0.023	3.51	0.012	1.35	0.018	2.08
Adj R^2		20.7 %		11.4 %		13.7 %		14.8 %
Panel (8)								

Table OA3: (Continued.)

Models	Contemporaneous		Staggered		Predictive – 1mo		Predictive – 6mo	
	Coefficient	t-Statistic	Coefficient	t-Statistic	Coefficient	t-Statistic	Coefficient	t-Statistic
Int	0.164	4.65	0.131	3.95	0.231	4.93	0.243	4.75
β	0.063	3.37	0.040	2.76	0.023	1.41	0.021	1.41
J^{adj}	-0.007	-9.45	0.000	2.18	0.000	0.23	-0.000	-1.70
Mcap	-0.024	-5.29	-0.019	-4.31	-0.027	-4.87	-0.028	-4.55
B/M	0.022	3.10	0.023	3.51	0.012	1.35	0.019	2.09
Adj R^2		19.5 %		11.3 %		13.6 %		14.8 %
Panel (9)								
Int	0.159	4.56	0.125	3.83	0.230	4.94	0.243	4.78
β	0.065	3.51	0.042	2.94	0.022	1.33	0.020	1.38
$\beta^+ - \beta^-$	-0.002	-0.43	-0.014	-4.11	-0.001	-0.17	-0.002	-0.52
J^{adj}	-0.007	-9.34	0.001	4.06	0.000	0.51	-0.000	-1.59
Mcap	-0.024	-5.29	-0.018	-4.26	-0.027	-4.89	-0.029	-4.59
B/M	0.023	3.26	0.023	3.54	0.012	1.38	0.018	2.09
Adj R^2		20.1 %		11.7 %		13.9 %		15.0 %
Panel (10)								
Int	0.165	4.69	0.127	3.88	0.229	4.91	0.243	4.76
β	0.063	3.42	0.043	2.93	0.023	1.43	0.021	1.40
Cosk	0.056	3.47	-0.067	-4.63	-0.009	-0.52	-0.002	-0.21
J^{adj}	-0.007	-9.43	0.001	4.83	0.000	0.82	-0.000	-1.15
Mcap	-0.025	-5.36	-0.018	-4.29	-0.027	-4.88	-0.028	-4.57
B/M	0.022	3.11	0.023	3.45	0.012	1.38	0.019	2.09
Adj R^2		19.9 %		11.6 %		13.8 %		14.9 %

Table OA4: Contemporaneous, staggered, and predictive cross-sectional regressions of average returns on estimated risk factor loadings, controlled for skewness, kurtosis, book-to-market ratio, and size. This table compares the results of three types of Fama-MacBeth regressions: Contemporaneous, in which returns and factor loadings for each month t are estimated using the same data periods $(t-T, t]$, where $T=6$ months; Staggered regressions, in which returns and factor loadings are estimated on staggered months as described in Section 2 (where, in each estimation period $(t-\tilde{T}_1, t]$ for each t, where $\tilde{T}_1$ is 18 months, months $t-15$, $t-12$, $t-9$, $t-6$, $t-3$, and t are used for realized return estimation and months $t-16$, $t-13$, $t-10$, $t-7$, $t-4$, and $t-1$ are used for factor loading estimation); and Predictive regressions, where, for each time t, factor loadings are estimated using return time series $[t-T_1, t)$ and average returns are estimated for either a single month $t+1$ or for the period $(t, t+T_1]$, where T_1 is six months. Size and Book-to-Market factors are estimated using classic Fama and French (1992) method. Size factor is the logarithm of issuing firm market capitalization measured in December of year preceding the first date of return estimation period. Book-to-market factor is based on latest reported annual book equity preceding the first date of return estimation period. If return estimation period starts before July 1 of year Y_0, then size and book-to-market are estimated using data from Y_0-2. Medium term momentum factor, Past Ret_{-T}, is based on the period prior and equal in length to the return estimation period. Short-term momentum, Ret_{-1}, is based on one month preceding the estimation period (for Staggered regressions, Ret_{-1} is estimated on staggered basis, contemporaneously to systematic risk factors). Panels (1) *through* (8) *report results of separate multiple regression models. Reported t-statistics are adjusted for overlapping periods using the Newey-West method. Average $AdjR^2$ of cross-sectional regressions are reported for each estimation model. The study used daily return time series of stocks traded on NYSE/Amex/Nasdaq during the period between 1963 and 2018, with sharecodes 10 and 11.*

Models	Regression Types							
	Contemporaneous		Staggered		Predictive – 1mo		Predictive – 6mo	
	Coefficient	t-Statistic	Coefficient	t-Statistic	Coefficient	t-Statistic	Coefficient	t-Statistic
Panel (1)								
Int	0.066	2.97	0.081	3.97	0.109	3.83	0.111	4.02
Skew	0.146	9.95	-0.007	-1.98	0.008	1.99	0.019	3.57
Adj R^2		10.6 %		0.5 %		0.5 %		0.6 %
Panel (2)								
Int	0.027	1.78	0.053	3.56	0.093	4.12	0.101	4.26
β	0.048	2.69	0.029	2.11	0.012	0.78	0.007	0.45
Skew	0.144	10.03	-0.010	-3.01	0.007	1.82	0.016	3.63
Adj R^2		15.0 %		3.9 %		3.2 %		2.9 %
Panel (3)								
Int	0.061	4.06	0.057	3.92	0.099	4.35	0.100	4.19
β	0.044	2.52	0.029	2.10	0.011	0.72	0.006	0.43
Skew	0.156	9.98	-0.009	-2.87	0.011	2.57	0.018	3.95
Kurt	-0.006	-7.53	-0.001	-1.92	-0.001	-2.68	-0.000	-0.24
Adj R^2		15.6 %		4.0 %		3.3 %		3.0 %

Table OA4: (Continued.)

Models	Regression Types							
	Contemporaneous		Staggered		Predictive – 1mo		Predictive – 6mo	
	Coefficient	t-Statistic	Coefficient	t-Statistic	Coefficient	t-Statistic	Coefficient	t-Statistic
Panel (4)								
Int	0.054	3.66	0.055	3.86	0.098	4.33	0.101	4.20
β	0.044	2.52	0.029	2.13	0.011	0.69	0.006	0.40
J^{adj}	-0.003	-8.24	-0.000	-0.67	-0.000	-1.21	-0.001	-2.79
Skew	0.150	9.99	-0.010	-3.05	0.010	2.51	0.017	3.96
Kurt	-0.006	-7.53	-0.001	-1.89	-0.001	-2.66	-0.000	-0.21
Adj R^2		16.0 %		4.1 %		3.3 %		3.1 %
Panel (5)								
Int	0.194	5.23	0.112	3.73	0.229	4.91	0.243	4.74
β	0.064	3.40	0.028	2.12	0.023	1.42	0.021	1.43
Mcap	-0.027	-5.60	-0.014	-3.86	-0.026	-4.85	-0.029	-4.55
B/M	0.022	3.11	0.015	2.62	0.012	1.36	0.019	2.10
Adj R^2		17.2 %		30.2 %		13.5 %		14.7 %
Panel (6)								
Int	0.141	3.44	0.144	3.87	0.248	4.58	0.251	4.40
Skew	0.134	9.93	-0.014	-5.27	-0.003	-1.13	0.007	2.17
Mcap	-0.015	-3.72	-0.014	-3.70	-0.026	-4.93	-0.027	-4.47
B/M	0.014	1.70	0.016	2.58	0.015	1.50	0.022	2.16
Adj R^2		21.7 %		28.0 %		11.4 %		12.8 %
Panel (7)								
Int	0.109	3.41	0.119	3.88	0.230	4.91	0.240	4.71
β	0.056	3.05	0.029	2.22	0.024	1.47	0.021	1.43
Skew	0.132	10.00	-0.016	-5.74	-0.004	-1.63	0.005	1.89
Mcap	-0.017	-4.03	-0.014	-3.99	-0.027	-4.87	-0.028	-4.52
B/M	0.013	1.94	0.016	2.75	0.013	1.44	0.019	2.10
Adj R^2		25.6 %		30.4 %		13.7 %		14.8 %
Panel (8)								

Table OA4: (Continued.)

Models	Regression Types							
	Contemporaneous		Staggered		Predictive – 1mo		Predictive – 6mo	
	Coefficient	t-Statistic	Coefficient	t-Statistic	Coefficient	t-Statistic	Coefficient	t-Statistic
Int	0.152	4.58	0.127	4.16	0.242	5.12	0.244	4.76
β	0.053	2.91	0.029	2.20	0.022	1.40	0.020	1.38
Skew	0.142	10.02	-0.015	-5.58	-0.000	-0.02	0.007	3.06
Kurt	-0.006	-7.65	-0.001	-3.08	-0.002	-3.78	-0.001	-1.25
Mcap	-0.018	-4.27	-0.015	-4.06	-0.027	-4.92	-0.028	-4.54
B/M	0.012	1.71	0.016	2.72	0.012	1.37	0.018	2.06
Adj R^2		26.2 %		30.5 %		13.7 %		14.9 %
Panel (9)								
Int	0.143	4.37	0.126	4.15	0.243	5.14	0.245	4.77
β	0.053	2.91	0.029	2.23	0.022	1.38	0.020	1.37
J^{adj}	-0.003	-8.05	-0.000	-0.34	0.000	0.05	-0.000	-1.12
Skew	0.137	10.02	-0.015	-5.62	0.000	0.05	0.007	3.03
Kurt	-0.006	-7.66	-0.001	-3.08	-0.002	-3.79	-0.001	-1.25
Mcap	-0.017	-4.15	-0.015	-4.07	-0.027	-4.94	-0.028	-4.54
B/M	0.012	1.73	0.016	2.69	0.012	1.36	0.018	2.06
Adj R^2		26.5 %		30.5 %		13.8 %		15.0 %
Panel (10)								
Int	0.191	5.33	0.113	3.65	0.217	4.66	0.228	4.74
β	0.055	3.04	0.027	2.09	0.011	0.71	0.008	0.68
Mcap	-0.027	-5.71	-0.014	-3.79	-0.025	-4.63	-0.027	-4.60
B/M	0.020	2.98	0.015	2.57	0.012	1.40	0.017	2.08
Past Ret_{-T}	0.029	2.73	0.014	2.35	0.043	2.90	0.050	3.41
Ret_{-1}	-0.009	-5.41	-0.024	-2.97	-0.049	-7.09	-0.008	-4.04
Adj R^2		18.7 %		31.1 %		15.4 %		16.2 %
Panel (11)								
Int	0.113	3.06	0.138	3.84	0.216	4.16	0.225	4.38
Skew	0.134	10.00	-0.012	-4.47	-0.000	-0.02	0.001	0.26
Mcap	-0.013	-3.64	-0.013	-3.65	-0.023	-4.69	-0.026	-4.60
B/M	0.012	1.57	0.015	2.49	0.015	1.46	0.020	2.25

Table OA4: (Continued.)

Models	Contemporaneous		Staggered		Predictive – 1mo		Predictive – 6mo	
	Coefficient	t-Statistic	Coefficient	t-Statistic	Coefficient	t-Statistic	Coefficient	t-Statistic
Past Ret_{-T}	0.055	4.36	0.014	2.05	0.037	2.29	0.050	3.15
Ret_{-1}	-0.008	-4.24	-0.010	-1.10	-0.046	-6.72	-0.008	-3.96
Adj R^2		23.9 %		29.3 %		13.7 %		14.7 %
Panel (12)								
Int	0.105	3.45	0.118	3.81	0.216	4.64	0.228	4.75
β	0.047	2.63	0.028	2.20	0.011	0.72	0.007	0.60
Skew	0.133	10.03	-0.013	-4.86	-0.000	-0.07	-0.000	-0.10
Mcap	-0.016	-4.09	-0.014	-3.93	-0.024	-4.60	-0.027	-4.61
B/M	0.011	1.74	0.016	2.70	0.013	1.44	0.017	2.13
Past Ret_{-T}	0.048	4.33	0.013	2.17	0.044	2.80	0.053	3.35
Ret_{-1}	-0.008	-5.00	-0.014	-1.68	-0.050	-7.10	-0.008	-4.06
Adj R^2		27.2 %		31.3 %		15.5 %		16.4 %
Panel (13)								
Int	0.147	4.63	0.127	4.09	0.229	4.86	0.230	4.78
β	0.044	2.49	0.028	2.17	0.010	0.66	0.007	0.59
Skew	0.143	10.05	-0.012	-4.80	0.004	1.48	0.002	0.58
Kurt	-0.006	-7.63	-0.001	-3.36	-0.002	-3.99	-0.001	-1.00
Mcap	-0.017	-4.34	-0.015	-4.00	-0.025	-4.66	-0.027	-4.63
B/M	0.010	1.51	0.016	2.66	0.012	1.38	0.017	2.11
Past Ret_{-T}	0.047	4.27	0.013	2.18	0.042	2.69	0.052	3.35
Ret_{-1}	-0.008	-5.02	-0.014	-1.73	-0.050	-7.11	-0.008	-4.08
Adj R^2		27.7 %		31.4 %		15.6 %		16.4 %
Panel (14)								
Int	0.138	4.43	0.126	4.09	0.231	4.89	0.232	4.81
β	0.043	2.50	0.028	2.20	0.010	0.65	0.007	0.59
J^{adj}	-0.003	-8.08	-0.000	-0.59	0.000	0.32	0.000	0.00
Skew	0.138	10.05	-0.012	-4.87	0.005	1.55	0.002	0.55
Kurt	-0.006	-7.64	-0.001	-3.35	-0.002	-4.00	-0.001	-1.01
Mcap	-0.017	-4.22	-0.015	-4.01	-0.025	-4.67	-0.027	-4.64

Table OA4: (Continued.)

Models	Regression Types							
	Contemporaneous		Staggered		Predictive – 1mo		Predictive – 6mo	
	Coefficient	t-Statistic	Coefficient	t-Statistic	Coefficient	t-Statistic	Coefficient	t-Statistic
B/M	0.010	1.53	0.015	2.64	0.012	1.37	0.017	2.11
Past Ret_{-T}	0.047	4.27	0.012	2.17	0.042	2.73	0.052	3.34
Ret_{-1}	-0.008	-5.06	-0.015	-1.78	-0.050	-7.13	-0.008	-4.02
Adj R^2		28.0 %		31.4 %		15.6 %		16.5 %