Market Events and Variation in Factor Structure

Yury Olshanskiy and Roman Sigalov[1]

June 13, 2024

Abstract

We study the stability of factor structure by analyzing its variation on different market events. We start by documenting variation in distributions, means, volatilities, and correlations, in a set of characteristics managed long-short portfolios on the weeks with large market movements, leading earnings announcements, and FOMC announcements with unexpected shocks to interest rates. This variation manifests in differences in factors extracted using characteristics based on statistical methods that we document using Instrumented PCA. The factor structure shows variation in the factor loadings and in the distribution of factors itself. We propose two ways of capturing event-specific variation in the factor structure. The first method, Treatment-IPCA, estimates orthogonal factors specific to the events we consider. We find significant premia associated with the treatment factors. The second method, Boosted-IPCA allows us to test the differential importance of firm characteristics in describing the cross-section of stock returns on market events relative to base periods.

1 Introduction

In this paper, we study a classic linear cross-sectional relationship among stock returns,

$$r_{i,t+1} - r_{f,t} = \alpha_t + \beta_{i,t}^{\top} f_{t+1} + \varepsilon_{i,t+1}, \tag{1}$$

with a focus on the time-variation in risk-exposures β and the distribution of factors f. We contribute to the literature in two ways. We first document that significant market events such

[1]Olshanskiy: MIT Sloan School of Managements, email: ols@mit.edu. Sigalov: Harvard Economics and Harvard Business School, email: rsigalov@g.harvard.edu. We thank John Y. Campbell, Hui Chen, Joanne Im, Leonid Kogan, Maria Turina, Adrien Verdelhan, Robbie Minton, Emil N. Siriwardane, Adi Sunderam and Markus Schwedeler for helpful comments and discussions.

as macroeconomic and earnings announcements, periods of large market fluctuations exhibit a different factor structure either through factor composition or factors' risk premia. Second, we propose a methodology of capturing the difference in the factor structure on such market events by extracting the event specific orthogonal factors as well as differences in the effect of firm characteristics in explaining the factor structure.

The study of the cross-section of stock returns lies in the core of financial economics. Identifying common variation or common factors in the movements of stock prices not only reveals information about fundamental economic drivers but also points to systematic sources of risk premia as first outlined by Arbitrage Pricing Theory (APT) in Ross (1976). APT postulates that any source excess returns that is not associated with a common factor in a large cross-section of stock returns will be arbitraged away by creating well diversified portfolios.

Many studies, most prominently Fama and French (1992), identified factors — in their case tradable portfolios based on fundamental characteristics sorts — that have a high explanatory power for the cross section. They further found that portfolios formed based on book-to-market and size carry significant risk premia. While the source of these premia, whether rational or behavioral, have been widely debated in the literature (Lakonishok et al., 1994; Zhang, 2005), Kozak et al. (2018) have shown that APT logic still holds regardless of the source of mispricing. Factors identified in Fama and French (1992) have been refined by considering intangible capital in for constructing book-to-market Eisfeldt et al. (2022) and expanded to include five and then six factors (Fama and French, 2015; Barillas and Shanken, 2018).

On the other hand, a separate strand of literature emerged that argues for a more statistical approach in identifying the factor structure in the cross section of stock returns. In one way or another, these papers are built on Principal Component Analysis (PCA), a statistical tool aimed to extracting factors, linear combinations of return series, that explain the most variation in the panel setting. The standard PCA targets the second moments of returns while economic theories such as APT relate the second moments to the first moment, namely the risk premia such factors should earn. Lettau and Pelger (2020) along with Bryzgalova et al. (2023) impose additional first moment restrictions to make the PCA approach more robust and aligned with such economic theories. Using PCA also leads to poorly estimated factors and loadings when the sample of test assets is large. This problem was addressed by Kozak et al. (2020).

The purely statistical approach has a solid theoretical foundation due to results in Kozak et al. (2018) showing that principal component factors should be associated with higher average returns to avoid near arbitrage opportunities in the spirit of Ross's APT. On the other hand, it loses useful information about firms' fundamental. When a panel of stock returns is considered, the standard PCA approach doesn't leave room for the dependence of firms' loadings on their fundamental characteristics such as profitability, book-to-market, etc. PCA requires having a full panel of stocks that limits the cross-section and leads to a survivorship bias. Moreover, it imposes restrictions by not allowing assets to change their loadings over time. This drawback has been recognized in the literature and an alternative formulation was proposed in Kelly et al. (2019) (KPS, henceforth). They use Instrumented PCA (IPCA) where they allow firms' loadings to flexibly depend on the fundamentals of the firm.

While there has been a significant progress in improving the baseline PCA to account for firm fundamentals and asset pricing restrictions, little is known about the variation in the factor structure. If we assume that financial markets are, at least in part, driven by information flows, then we can argue that these flows are time varying. There are important market events that reveal information about fundamentals of the economy. Recently, elevated inflation brought back the relevance of CPI announcements. We observe that different types of firms such as value and growth react differentially to CPI surprises. Beside book-to-market, there are other firm fundamentals that are particularly sensitive to news about inflation and/or expectation of Fed monetary policy. In general, the informational content of macroeconomic announcements may be substantially different as it reveals information about the state of the economy. Prior research identified that macroeconomic announcements look very different from other days when it comes to compensation for factor exposure. The Security Market Line, while flat on days outside of macroeconomic announcements, has a significant positive slope on announcements (Savor and Wilson, 2014). A similar dichotomy has been identified for leading earnings announcements (Chan and Marsh, 2022).

While we also consider similar events in our paper, we do not focus on a particular anomaly, nor do we analyze any firm characteristics and portfolios sorted based on certain characteristics. We take a more holistic approach and bring the tools developed in the statistical factor identification literature to study the differences. We aim to systematically identify the differences in factor structure on the events that we deem to be important for financial markets.

We build on IPCA from KPS to compare the factor structure on events such as macroeconomic announcements, leading earnings weeks, and periods with unusual market dynamics. We find that there are indeed differences that are either related to the firm characteristics' loadings on the factors that described the most variation in the cross-section, or in the risk premia the factors themselves earn even if the loadings on events stay similar to loadings outside of these events. We then introduce a new approach that we call Treatment IPCA that introduces an event specific factor that is orthogonal to other factors. Such additional factor allows us to isolate the additional variation that is present on certain market events. Importantly, this factor introduces little to no change to the outside-of-event factors and loadings. We further introduce a Boosted IPCA, that refines the non-event factor structure and allows us to test the change in the role of firm characteristics in explaining the factor structure on market events.

2 Is the Factor Structure Stable?

Since development of APT models and seminal Merton's ICAPM and CCAPM which provided a theoretical explanation for the empirical failure of the CAPM in describing risk compensation for diversified portfolios, the large strand of asset pricing literature has been dedicated to proposing reasonable measurements for both exposures and factors that drive returns in (1).

The central question is finding the estimation procedure, i.e., mappings ($\mathcal{G}$, $\mathcal{F}$) from a set of observables to the vector spaces of risk exposures β_t and factors f_{t+1} respectively. Fama and French (1992) directly constructs $\mathcal{F}$ by utilizing their prior knowledge on portfolios that drive the cross-section such as Market, HML and SMB. This approach was subsequently improved with the introduction of new factors: Fama and French (2015) incorporated operating profitability and investment information, while Barillas and Shanken (2018) argued for adding momentum factor. Once $\mathcal{F}$ is set, $\mathcal{G}$ is constructed via a time series regression of an asset's return on the factors.

A statistical approach such as Principal Component Analysis (Korajczyk and Connor, 1988; Connor and Korajczyk, 1993) define $\mathcal{G}$ and $\mathcal{F}$ simultaneously using the panel of asset or portfolio returns in isolation without using any other information beyond realized returns. To achieve this, it imposes a strict assumption of constant loadings, $\beta_t = \beta$. Lettau and Pelger (2020), Bryzgalova et al. (2023), Kozak et al. (2020) follow the same approach and improve the estimation by proposing a set of economically motivated restrictions such as penalty for lack of factors' premia targeting.

$$\mathbb{E}[r_{t+1}^n] = r_{f,t} + \beta_t^\top \lambda_t. \tag{2}$$

More recent improvements such as Kelly et al. (2019) and **?** incorporate data on firm fundamentals to construct ($\mathcal{G}$, $\mathcal{F}$) that result in a time invariant mapping between firm characteristics and factor loadings $G : Z_t \to \beta$[1]. That allows to improve performance, both in-sample and out-of-sample, when fitting (1) and (2), and also achieve higher tractability of factors. Incorporating similar data into portfolio classification problem Bryzgalova et al. (2019) help to improve factor construction, $\mathcal{F}$.

Typically, the recommended methods for estimating $(\mathcal{G}, \mathcal{F})$ assume a time-invariant framework where each period is treated symmetrically. For example, portfolio sorting procedure will be same when constructing f_t for every t or PCA-based method will treat all periods similarly in minimization problem. Consequently, any changes in the loadings β and factors f over time are solely attributed to variations in the company's characteristics and realized returns. It should be noted that it does not necessarily imply that various time periods make an equal contribution to the estimated factors and loadings: the estimation methods may implicitly weigh periods differently due to the heteroskedasticity of data.

Furthermore, a vast amount of research indicate that stocks exhibit significant variations in performance during certain periods, such as macroeconomic and earnings announcements, compared to other periods. That may include substantial variation in risk-premia (Savor and Wilson, 2014) and return anomalies. The literature suggests that there are spikes in information arrival and investor attention, which are associated with specific market and world events (Wachter and Zhu, 2022). When investigating such periods, researchers typically establish a control period in which the market fluctuates in a conventional manner, such as pre-announcement times, and a treated period in which the market or a particular asset class deviates from the benchmark. The literature's connection to the factor structure is usually limited by applying some form of risk adjustment based on the most widely used models such as the three factor Fama-French model.

Though it is still plausible that the fundamental factor structure for many of these events is similar to generally studied ($\mathcal{G}$, $\mathcal{F}$), this assumption necessitates additional investigation. Our suggestion is to integrate these two strands of literature by jointly estimating the factor structure and studying pre-selected market events. Conceptually, we introduce a certain degree

[1]Note that $\mathcal{G}$ is more comprehensive than G as it relies on the complete dataset employed for estimation.

of flexibility in the mappings $(\mathcal{G}, \mathcal{F})$, thereby breaking the previously discussed symmetry. The estimates β_t and f_t are supposed to be driven not only by time dependent variation in a set of observables but also by changing $(\mathcal{G}(t), \mathcal{F}(t))$. The value of this approach is that if one imposes restrictions on how $(\mathcal{G}(t), \mathcal{F}(t))$ are related with each other, it gives an interpretation of the structural differences between the periods. For instance, suppose $\mathcal{G}(t_1)$ includes information on stock momentum characteristics[2], while $\mathcal{G}(t_2)$ comprises information on operational leverage. In that case, it could be argued that the news disclosed at time t_2 could be more relevant for levered firms. Subsequently, one could examine whether any interest rate-related news was disclosed during that period.

Rather than assuming a completely general specification that permits $\mathcal{G}(t)$ to vary unrestrictedly with t, this paper concentrates on a simpler method. We manually categorize time periods by identifying some as events and others as non-events. Next, we present two collections of mappings, namely $(\mathcal{G}^E, \mathcal{F}^E)$ for events and $(\mathcal{G}^{NE}, \mathcal{F}^{NE})$ for non-events. Our initial step is to estimate the mappings separately for events and non-events. This will help emphasize that the parameters responsible for driving the cross-sectional relationship (1) differ in the examined time periods.

We chose IPCA from KPS as the workhorse model for our paper because of its high interpretability and good fitting properties. However, the initial analysis could also be performed using a different framework. We begin by estimating the model separately for events and control periods, treating them as distinct samples. Upon analyzing the dissimilarities that manifest in the mapping from firm characteristics to factor exposures and the characteristics of estimated factors, we deduce that significant distinctions exist in both respects for the days and weeks surrounding macro and earnings announcements, as well as during periods of substantial market fluctuations.

Once we establish the presence of heterogeneity, we propose a method for measuring the additional structure that arises during predefined events. To achieve this, we suggest two strategies: Boosted-IPCA and Treatment-IPCA. The underlying idea for both estimates is that $(\mathcal{G}^E, \mathcal{F}^E)$ must share some similarities with $(\mathcal{G}^{NE}, \mathcal{F}^{NE})$ and cannot be entirely structurally independent of non-events. Boosted-IPCA leverages the estimates derived from IPCA and updates them

[2]To put it more formally, the information is contained within an intermediate interpretable mapping, denoted as G. In the case of IPCA, for instance, the mapping $\Gamma : Z_t \to \beta$ is predominantly influenced by momentum characteristic. Further details on this will be elaborated upon in the paper.

to better align with event data. Treatment-IPCA follows the same idea but performs a joint estimation of $(\mathcal{G}^E, \mathcal{F}^E)$ and $(\mathcal{G}^{NE}, \mathcal{F}^{NE})$.

2.1 Events Definitions

We consider two types of events. The first type is chosen with *ex-ante* attributes. Those are weeks of FOMC announcements and leading earnings announcements (LEAD) weeks. The second type is determined *ex-post*. These are the "events" chosen based on realized market moves or changes in interest rates. Both events and non-events are defined as weeks, and, accordingly, our analysis is on the weekly return frequency.

Predetermined events The first group of events is predetermined, meaning they do not depend on the realization of market returns, volatility, or other variables. First are LEAD weeks, where a large share of S&P 500 firms release quarterly results ahead of other companies. These announcements are believed to carry new information about economic fundamentals and Chan and Marsh (2022) found that similar to macroeconomic announcements (Savor and Wilson, 2014) these days feature a meaningful risk-return tradeoff. We follow the procedure described in Chan and Marsh (2022) to identify LEAD week as the first week for a given earnings report season where at least 50 S&P 500 firms report their results. These weeks usually fall on the last week of January, April, July, and October. We split the sample into LEAD and non-LEAD weeks and label LEAD as events and non-LEAD weeks as non-events. Finally, following (Savor and Wilson, 2014) , we use the weeks of FOMC announcements from 1970 to 2022 as events and label other weeks as non-events.

Ex-post events The second category of events we consider is the ones that depend on the realization of market returns and interest rates. The first event type is the weeks where the absolute return on the market is above the 90th percentile, i.e., the weeks of large market moves. The second event type is the subsample of FOMC announcements with the largest (Nakamura and Steinsson, 2018, NS, henceforth) shocks. Note that NS shocks start only in 1995 and end in 2019. We, therefore, extend the sample through 2022 to include the most recent years, where we find a substantial variation in interest rate shocks.

Placebo To ensure that a randomness in realized returns does not drive the results we find, we contrast them with a placebo event when we report results for particular events. Specifically, we identify each odd week as an event and even as a non-event.

2.2 Change in the Distribution of Characteristic-Managed Portoflios

We follow Kelly et al. (2019) in analyzing a set of firm characteristics and, in particular, returns of portfolios formed based on these characteristics. As these portfolios constitute the building block of Instrumented IPCA introduced in KPS, we first analyze their response to events. We use the same characteristics as KPS, as outlined in section 4.1 of the paper. We report the complete set of characteristics we employ in Appendix A. We exclude several variables that are highly correlated with others (assets, a2me, q) as well as suv as it is very sensitive to the exact definition one uses. Following KPS, we use fundamental variables' definitions from the online appendix to Freyberger et al. (2020).

Each unit of analysis is defined by an $N_{t+1} \times 1$ vector of returns r_{t+1} where N_{t+1} is the number of firms over a given period $[t, t+1)$. We assign to this period an $N_{t+1} \times L$ characteristic matrix Z_t. Each column of matrix Z_t contains a vector of a particular firm characteristic, for example, book-to-market. All the characteristics are cross-sectionally ranked and normalized to belong from -0.5 to 0.5. In this way, the inner product of a column l of Z_t and the vector of realized returns

$$h_{l,t+1} \equiv (Z_t)_{\cdot l} \cdot r_{t+1} \tag{3}$$

is the return on a long-short or arbitrage portfolio of stocks formed based on characteristic l. This portfolio assigns a greater positive weight to firms that exhibit higher values of certain characteristics, such as a higher book-to-market ratio, while assigning a lower negative weight to firms with lower values of those characteristics.

The long-short potfolios form a $L \times 1$ vector,

$$h_{t+1} \equiv Z_t^\top r_{t+1} \tag{4}$$

where the l-th element of h_{t+1} is equal to a portfolio formed based on characteristic l. These portfolios are the main building block of IPCA in KPS. Therefore, we start by analyzing the factor structure of these portfolios and investigating their variation on the market events we identify.

PCA based methods are estimated through the covariance matrix of characteristic based portfolio returns of characteristic based portfolio returns. Therefore, as the starting point of our analysis we document changes in volatilities and correlations among these portfolios on events vs non-events.

Portfolio Volatilities We calculate the standard deviation of each portfolio in h on events and non-events and plot them against one another in panel (a) of Figure 1. The leftmost plot shows the placebo specification, with little difference between events and non-events. It is worth noticing, though, that the weekly volatility of the long-short portfolios varies substantially, increasing almost four times from 0.5% (annualized) to 2% for the price relative to the 52-week high (w52h)-based portfolio.

The following plot shows that for weeks with significant market moves, there is a general increase in the volatility of each portfolio. Though it is not purely a mechanical observation, since we consider a long-short portfolio that might be market neutral, the sharp increase of portfolio volatility is still expected because the portfolios generally have non-zero market beta and since large market moves are associated with an increase in volatility. The third plot for Leading Earnings Announcement weeks shows that portfolio volatilities do not change significantly on events with a few exceptions, such as price relative to 52-week high (w52h), momentum (mom), and turnover (turn).

The fourth panel reveals that portfolio volatilities do not change significantly for the sample of all FOMC announcement weeks, and all portfolios lie tightly along the 45-degree line. The picture looks markedly different when we narrow the sample of all FOMC announcements to the ones where interest rate shocks accompany the FOMC announcements. For example, subsetting the weeks to the ones with the absolute value of Nakamura-Steinsson shocks above the 80th percentile, several portfolios, for instance, market cap-based and investments-based portfolios, are significantly more volatile than on other weeks.

Overall, among many specifications we test,the portfolio's variance seems relatively stable unless one studies events of extreme market volatility. Nevertheless, specific characteristic portfolios may show the variation of the second moment depending on the asset. This may reflect how some events are more indicative of a particular company's underlying fundamentals or characteristics compared to others.

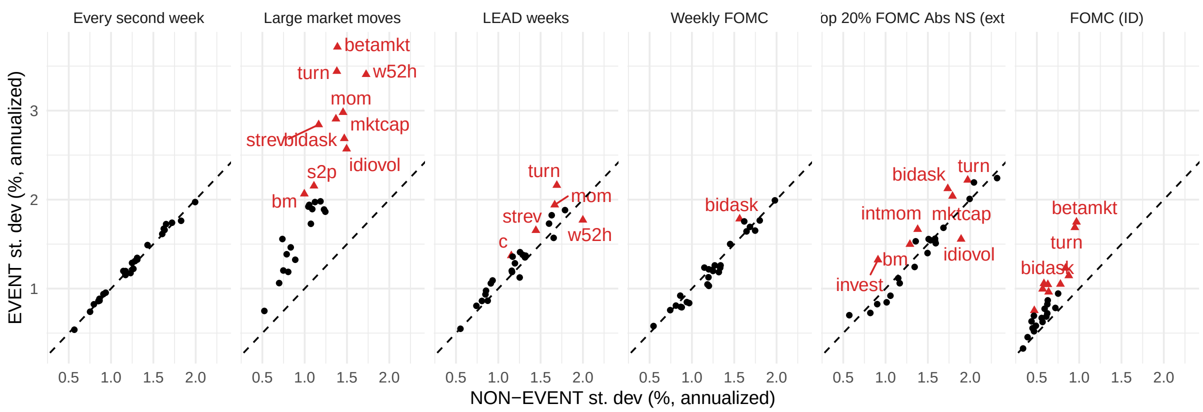
Portfolio Correlations To conclude the second-moment discussion of the long-short portfolio, we estimate the correlation matrix of the portfolio cross-section and plot the off-diagonal entries on events and non-events against each other in panel (b) of Figure 1. The leftmost plot showing the placebo event shows a small dispersion around the diagonal, indicating that the whole second-moment matrix seems robustly estimated across the times. It makes it more intriguing that the matrix varies substantially across the events.

The large market moves weeks exhibit a significant dispersion of correlations. Though the correlations appear around 45-degree line, their overall dispersion suggests possible variability in the factor structure underlying the characteristic portfolios.

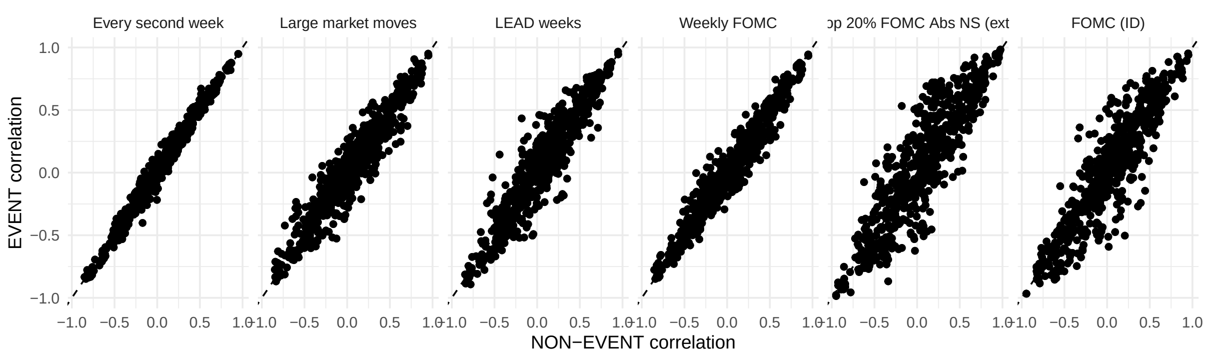
Similarly, for LEAD weeks, the correlation matrix differs from its non-event counterpart. In this case, however, the dispersion in correlations is not accompanied by an increase in volatility, as we have seen above in panel (a). We document a wider distribution of correlations around 45-degree line for the largest NS shocks on FOMC announcements. Since PCA identifies familiar sources of variation from the covariance matrix, a difference in correlations between events and non-events is the first indication of a potentially different factor structure.

Portfolio average returns Finally, it is important to note that the realized portfolio premia vary across events. Figure 2 compares non-event average portfolio returns to those on events. The placebo specification again shows no differences in average returns. For large market moves, we see that two portfolios, price relative to 52-week high and short-term reversal, have very different average returns, and the average return on these portfolios is much lower on weeks with large market moves. We will later find that these characteristics contribute to an additional orthogonal variation and generate a substantial premium.

In the third panel, we do not find large noticeable deviations in average returns of characteristic portfolios earned on LEAD weeks relative to other weeks. Weeks of FOMC announcements feature a broader dispersion in average returns, but it does not feature strong outliers similar to events with large market moves. In the last panel, FOMC weeks with the largest Nakamura-Steinsson shocks have a wider dispersion in average returns, including significantly lower average returns for portfolios formed based on short-term reversal, bid-ask spread, book-to-market, price-relative to 52 weeks high, leverage. We observe a larger presence of fundamental firm characteristics among the outliers. Similarly, we will find that fundamental characteristics



(a) Standard deviation



(b) Correlations

Figure 1: Change in Second Moments of Characteristic-Managed Portfolios on Events vs Non-Events.

change their role in the factor structure when we analyze Boosted-IPCA.

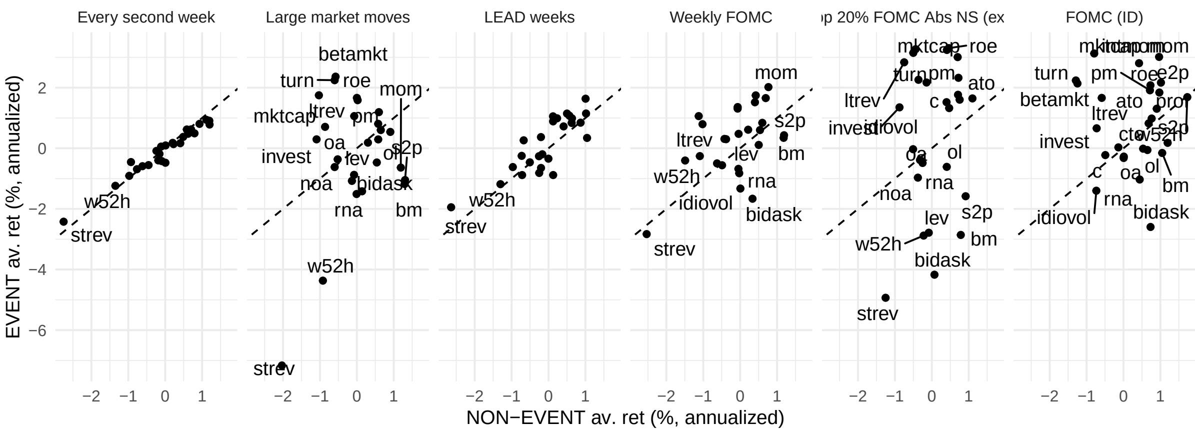


Figure 2: Average Returns on Characteristic-Managed Portfolios

2.3 Instrumented Principal Components Analysis

The presented descriptive facts suggest that the factor structure might differ for certain events but does not provide evidence that the equation (1) is subject to change. To make the discussion on the difference in factor structure more concrete and highlight the different ways the factor structure may be different, we now turn to explaining the primary method we use to document the differences in factor structure. Kelly et al. (2019) consider a factor structure with latent factors and loadings where the loadings depend on firm characteristics. We start by comparing the estimates from IPCA on events and non-events. Similar to standard PCA, it posits that returns are linear in latent factors

$$r_{i,t+1} = \alpha_{i,t} + \beta_{i,t}^T f_{t+1} + \varepsilon_{i,t+1}, \tag{5}$$

where f_{t+1} is a $K \times 1$ vector of factors, and $\beta_{i,t}$ is a $K \times 1$ vector of factor loadings. Unlike standard PCA, it assumes that the intercept α and loadings β depend on firm characteristics

$$\alpha_{i,t} = z_{i,t}^T \Gamma_\alpha + u_{\alpha,i,t}, \;\; \beta_i^T = z_{i,t}^T \Gamma_\beta + u_{\beta,i,t}, \tag{6}$$

where $z_{i,t}$ is an $L \times 1$ vector of firm i's characteristics and Γ_α $(L \times 1)$ and Γ_β $(L \times K)$ are common among firms. In this way, even for the same firm the intercept α and loadings β may change if its characteristics change over time. The full expression for returns is, therefore,

$$r_{i,t+1} = z_{i,t}^T \Gamma_\alpha + z_{i,t}^T \Gamma_\beta f_{t+1} + \varepsilon_{i,t+1} = z_{i,t}^T \left(\Gamma_\alpha + \Gamma_\beta f_{t+1}\right) + \varepsilon_{i,t+1}^*. \tag{7}$$

IPCA solves the following minimization problem

$$\min_{\Gamma_\alpha,\Gamma_\beta,f} \sum_{t=1} ||r_{t+1} - Z_t(\Gamma_\alpha + \Gamma_\beta f_{t+1})||_2{}^2. \tag{8}$$

Our analysis in this paper is centered around the return cross-section. Hence, unlike the original paper, we start by cross-sectionally demeaning the returns over a given period. As the loadings and factors are identified up to a rotation, meaning that for an orthogonal $K \times K$ matrix R we have $\Gamma_\beta f_{t+1} = \Gamma_\beta R^{-1} R f_{t+1}$, and intercepts up to subtracting a constant vector ξ from factors and adding $\Gamma_\beta \xi$ to Γ_α, we employ the same normalization as in KPS. In particular, we impose $\Gamma_\beta^T \Gamma_\beta = \mathbf{I}$, $f^T f$ is a diagonal matrix with decreasing entries, each factor in f has a positive mean, and $\Gamma_\alpha^T \Gamma_\beta = \mathbf{0}$.

The IPCA minimization problem features two main first-order conditions for loadings and factors reflecting time-series and cross-section, respectively. We will use the first order condition for factors extensively in this paper as it allows us to project returns on loadings estimated from an arbitrary subsample to obtain the factors

$$f_{t+1}(\hat{\Gamma}) = \left(\hat{\Gamma}_\beta^T (Z_t^T Z_t) \hat{\Gamma}_\beta\right)^{-1} \hat{\Gamma}_\beta^T \left(Z_t^T r_{t+1} - (Z_t^T Z_t)\hat{\Gamma}_\alpha\right) \tag{9}$$

for the subsample of interest $\{t+1\}$.

2.4 Baseline Factor Structure

Before analyzing the differences in factor structure, we will begin by presenting the baseline factor structure through the estimation of IPCA with four factors on the entire sample. We report the loadings on the each factor in Figure 3 and their correlation with the standard set of factors from French's website in Table 1. The first, most volatile, latent factor has significant loadings on the market beta, the price relative to 52-week high, momentum, and share turnover. Like the value-weighted market return, it exhibits a negative correlation with the value, profitability, momentum, and investment factors and a positive correlation with SMB and short-term reversal factors. As we will see, the first factor's loadings are generally the most robust factor across event studies revealing the primary importance of the market for the cross-section of returns for all types of events. We must notice that the factor is not exactly market and it is slightly more prone to momentum and short-term reversal factors. The large w52h and mom magnitude in the loading do reflect the bias to the factors.

The second factor has large negative loadings on several momentum-related characteristics such as price relative to 52-week high, short-term and long-term reversal, and intermediate momentum. Accordingly, its closest standard counterpart is the negative momentum factor. Relative to the first two factors, the third factor has larger loadings on fundamental firm characteristics shown at the bottom of the figure.

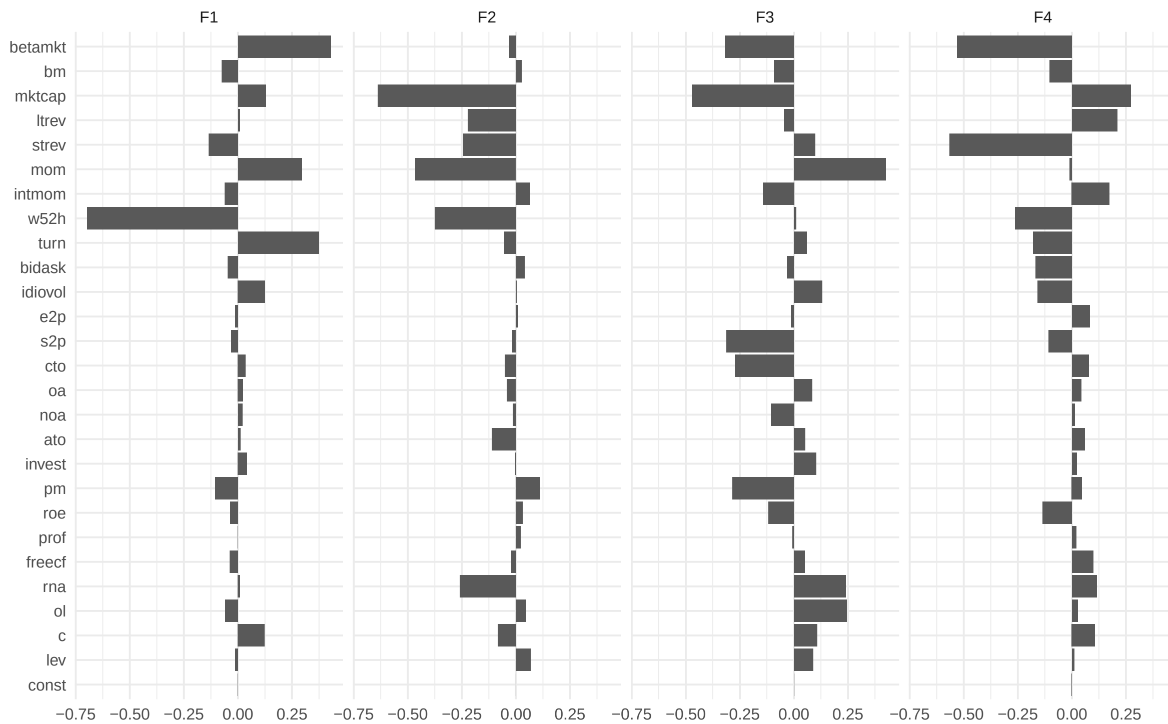


Figure 3: IPCA Factor Loadings

2.5 Difference in Factor Structure: Loadings

The first natural step to studying events' factor structure is to look at loadings on the events. We estimate IPCA separately on non-events and events for each specification described above. This generates two sets of coefficients Γ_β defined in equation (7). Each column of Γ_β corresponds to a different factor. If the factor structure is similar across the two subperiods, we expect the two Γ_β's estimated from different subsamples to line up column by column.

The normalization procedure in KPS that we follow sets the sign of the average factor return to be positive. To facilitate a more straightforward comparison of the factor structure between

	F1	F2	F3	F4
Mkt	74.52	-18.24	-26.51	-14.93
SMB	36.02	23.68	19.29	-31.04
HML	-29.42	42.4	-33.98	-20.29
RMW	-41.72	-4.65	-34.13	19.71
CMA	-46.81	22.81	-6.04	-21.46
Mom	-26.21	-53.89	39.51	-3.93
STrev	45.92	15.46	-14.73	40.26
LTrev	-7.08	26.51	-3.65	-39.6

Table 1: IPCA Factors Correlation with Fama-French Factors

events and non-events, we have slightly deviated from the conventional approach. Specifically, we have imposed a requirement that the loadings on events and non-events are positively correlated with each other, instead of anchoring the sign of factors to positive values. Hence, we first follow the normalization procedure from KPS for events and non-events separately. Second, we calculate a cross-characteristic correlation between loadings from kth (corresponding to the kth factor) column of Γ_β estimated on events and kth column of Γ_β estimated on non-events. Finally, we multiply the loadings in the kth column of events-based Γ_β by the sign of that correlation.

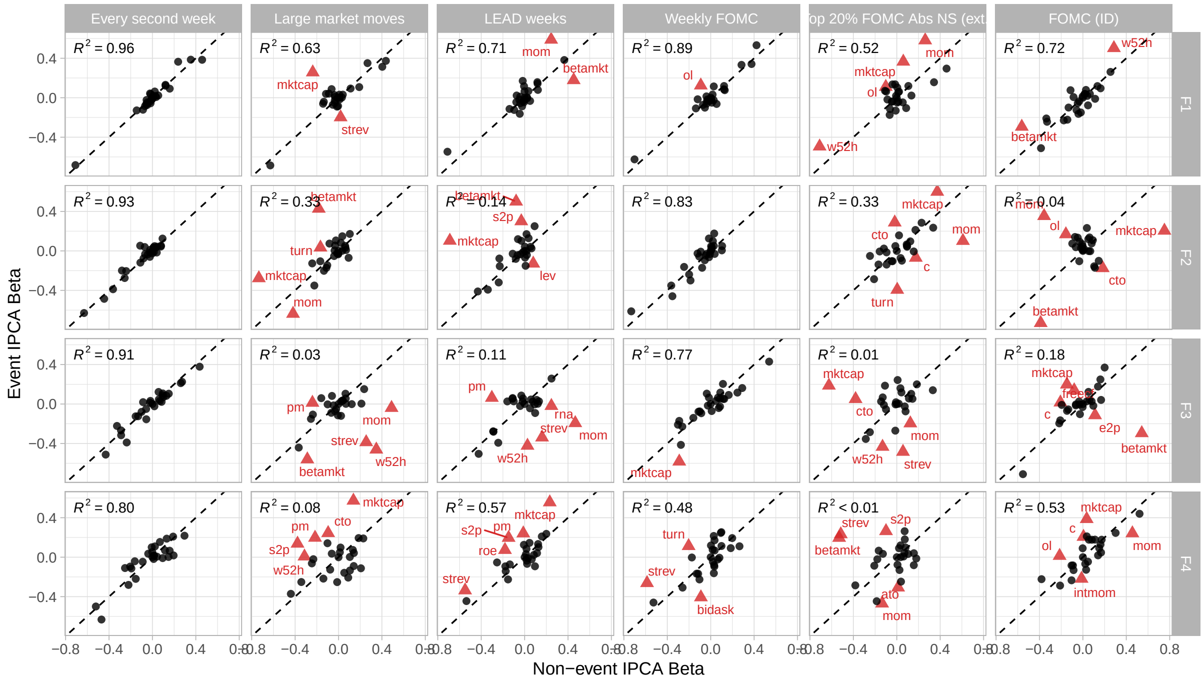
We compare the Γ_β columns in Figure 4 column by column. To understand the figure, consider the top left panel, which compares the loadings for placebo events where we label even weeks as events. Since the odd/even week split is not based on any economic rationale, we do not anticipate any differences in the factor structure. This panel plots the first column of Γ_β corresponding to the first factor estimated on non-events on the x-axis against the first column of Γ_β estimated on events on the y-axis. We observe that all points align along the 45-degree line, indicated as the upward-sloping dashed line. This is true for all factors (i.e., all rows) in the first column. The level of similarity is measured by the R^2 of the regression of events' Γ_β vs. non-events' Γ_β.

Let us instead consider the second column, where events are defined as top-10% of weeks by the absolute return on the market. We see significant deviations from the 45-degree line highlighted by red triangles in Figure 4 for the characteristics with the most notable difference. We observe that market capitalization has a close to zero loading in the first factor outside the events but a large positive coefficient on events. The opposite is true for the second factor. Market capitalization moves from the second factor to the first, i.e., the factor with relatively higher volatility. The observations indicate an increased role of market capitalization during weeks of large market movements.

The third column compares loadings on LEAD weeks. There, we also observe a substantial dispersion of loadings similar to weeks of significant market movements. However, when considering the sample of all FOMC weeks in the fourth column, the loadings appear to be much more similar, except for the fourth factor. However, filtering the FOMC announcements by the magnitude of the Nakamura-Steinsson shock reveals more dispersion in the loadings, particularly for the third and fourth factors.

It is worth discussing why there is a similarity in FOMC meetings when we include all of them, and why we observe more dispersion when selecting only the ones with shocks. While we might expect that frequently happening FOMC announcements reveal new information about the fundamental of the macroeconomy, most FOMC announcements do not feature any significant news or moves in interest rates. For instance, if we assume that the factor structure differs only when significant interest rate information is released, then the proportion of such events will be relatively small, say 20%. In such a scenario, since IPCA minimizes quadratic variance, it would assign only 0.04 weight to the days when the "informative" announcements occur. Thus, the estimated factor loadings will be havily tilted to a more standard factor structure. Therefore, it is essential to introduce additional filters that will help us to identify announcements with significant news flow.

Are event loadings spanned by non-event loadings? Figure 4 discussed above compares the factors sorted based on their standard deviation in descending order. At the same time, the loadings on event factors may be spanned by loadings on non-event factors, thus also implying a similar factor structure. We address this by estimating a regression of a given column of Γ_β estimated on events against all columns of Γ_β estimated on non-events. We repeat the same procedure for all columns of Γ_β estimated on events. We report R^2 from these estimates in



Note: we change the loadings so that cross-characteristic slope is positive

Figure 4: Comparing Loadings on Events and Non-Events

Figure 5 where an R^2 close to 100% indicates loadings on non-events can span the loadings on events.

In line with findings from Figure 4, the loadings for even weeks can be generated by the loadings for odd weeks, reinforcing the notion that the factor structure remains highly stable in the absence of significant market events. This indicates that employing factor estimates from odd weeks to fit factors on even weeks would yield almost identical results to using in-sample estimates.

The R^2 value for large market movements is slightly smaller, indicating a difference in factor structure compared to other weeks. There is a larger share of independent variation in factor loadings, and the increased R^2 compared to Figure 4 indicates that some rotation occurred. This redistribution of factor loadings among other factors will be discussed in the next section as it is one of the natural consequences of changes in the correlation of characteristic-based portfolios.

While one can find a similar picture for LEAD events, the larger differences appear for weeks of FOMC announcements in the fourth panel of Figure 5. The fourth factor has much smaller R^2 relative to other factors and previous event types. When we use Nakamura-Steinsson shocks to identify FOMC weeks with interest rate news we see a very different picture in the right-most panel of Figure 4. The R^2 is substantially lower for all factors compared to other specifications. The lower value of R^2 for all Γ-s indicates a difference in the factor structure. However, we cannot yet determine precisely what distinguishes these events from non-events and will develop tools in the next sections.

Finally, in the simulation exercise we perform in the Appendix Section B, we show that in the presence of strong factor structure, even in purely independent samples, we will "detect" a certain degree of correlation between loadings across the two subsamples. Therefore, the baseline level that we should compare the R^2 of cross-characteristic regressions is above zero with the 95% confidence level under the null, reaching 40%. Therefore, the R^2 on the order of 40% to 75% are consistent with little commonality between the loadings across events and non-events. Consistent with this evidence, we will find that firm characteristics, especially fundamentals, change their effect on the factor structure on the FOMC announcement weeks filtered by Nakamura-Steinsson shocks when we estimate Boosted IPCA.

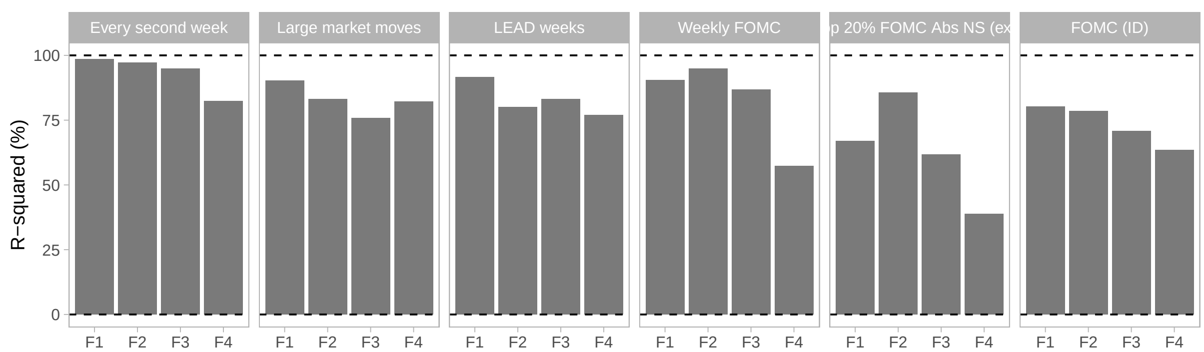


Figure 5: Do Non-Event Loadings Span Event Loadings?

2.6 Sources of Variation in Factor Structure

In this section, we estimated the latent factor structure using Instrumented PCA for events and non-events separately. We documented that there are significant differences in loadings that exist for periods of large market moves, LEAD weeks and FOMC announcements with large interest rates shocks. At the same time, the exact source of variations requires further investigation. The same differences in the loadings we observed may be the consequences of different sources of variation in the factor structure. The changing distribution of factors on events when combined with the standard normalization procedure of orthogonalizing the factors may generate variation in loadings. An additional factor, specific to the event, may be absorbed into the other factors thus altering both events factors as well as event loadings. Finally, a change in the effect of a firm characteristic on events most straightforwardly leads to a difference in loadings.

In the following three sections, we will address the three sources of variation in factor structure separately. First, in Section **??**, we will consider the variation in the factor distribution by extrapolating the factor structure estimated from non-events with IPCA to events. Second, in Section 3, we will introduce Treatment IPCA that will estimate an orthogonal factor specific to events. Third, in Section 4, we introduce Boosted IPCA that allows us to estimate and test the variation in the effect of a particular firm characteristics on the factor structure.

3 Treatment IPCA

In the previous section, we used the subsample of non-events to estimate the factor structure $\mathcal{G}^{NE}$ and extrapolated it to the subsample of events to obtain $\mathcal{F}^{E}$. We further documented that the distribution of event factors $\mathcal{F}^{E}$ is different from the distribution of non-event factors $\mathcal{F}^{NE}$. In this section, we jointly estimate $(\mathcal{G}^{E}, \mathcal{F}^{E})$ and $(\mathcal{G}^{NE}, \mathcal{F}^{NE})$ using the entire sample of returns and characteristics. Certain characteristics may not be important for the variation in the cross-section of equity returns but may generate an independent source of variation on market events, thus, generating additional events on events. To capture this source of variation in the factor structure, we introduce Treatment-IPCA. It posits the existence of additional factors $\tilde{f}$, that we call treatment factors, that are identically zero outside market events and are orthogonal to other factors, that we call base factors, on events.

3.1 Formulation of Treatment IPCA

To capture the difference in factor structure on specific periods, we introduce Treatment-IPCA, which allows us to incorporate additional factors on pre-specified events. In particular, we assume that the return has a latent factor structure

$$r_{i,t+1} = z_{i,t}^T(\Gamma_\alpha + \Gamma_\beta f_{t+1}) + \mathbb{I}_{t+1}(A) z_{i,t}^T(\tilde{\Gamma}_\alpha + \tilde{\Gamma}_\beta \tilde{f}_{t+1}) + \varepsilon_{i,t+1}, \tag{10}$$

where the indicator $\mathbb{I}_{t+1}(A)$ is equal to one when $t+1$ belongs to set A of pre-specified events and zero otherwise, f_{t+1} is a $K_1 \times 1$ vector of *base* factors, $\tilde{f}_{t+1}$ is a $K_2 \times 1$ vector of *treatment* (or *tilde*) factors, Γ_β ($L \times K_1$) and $\tilde{\Gamma}_\beta$ ($L \times K_2$) are matrices of *base* and *treatment* loadings respectively, Γ_α ($L \times 1$) and $\tilde{\Gamma}_\alpha$ ($L \times 1$) are matrices of *base* and *treatment* intercepts respectively.

By formulating the problem in this way, we, first, allow the intercept Γ_α to differ on specific dates and, second, allow additional factors $\tilde{f}$ to emerge only on these dates. The minimization problem is similar to IPCA in KPS

$$\min_{\substack{\{\Gamma_\alpha, \tilde{\Gamma}_\alpha, \Gamma_\beta, \tilde{\Gamma}_\beta\} \\ \{f_{t+1}, \tilde{f}_{t+1}\}_{t\in\mathcal{T}}}} \sum_{t\in\mathcal{T}} \left\| r_{t+1} - Z_t \left(\Gamma_\alpha + \Gamma_\beta f_{t+1}\right) - \mathbb{I}_{t+1}(A) Z_t \left(\tilde{\Gamma}_\alpha + \tilde{\Gamma}_\beta \tilde{f}_{t+1}\right) \right\|^2, \tag{11}$$

where r_{t+1} stacks individual stock returns $r_{i,t+1}$ and Z_t stacks firm characteristics $z_{i,t}^T$. The minimization problem is solved similarly to KPS with alternating least squares that estimates cross-sectional regressions to obtain factors given intercepts $(\Gamma_\alpha, \tilde{\Gamma}_\alpha)$ and loadings $(\Gamma_\beta, \tilde{\Gamma}_\beta)$ and time-series regression to obtain intercepts and loadings given factors $(f, \tilde{f})$ iteratively until the estimates converge. We outline the procedure and modifications that we make relative to IPCA in the Appendix Section D.

Normalization Similar to IPCA, we need to impose some restrictions on intercepts, loadings, and factors to achieve identification. We first do the same orthogonalization as in the IPCA described in the previous section for base factors and treatment factors separately: $\hat{E}[f^T f|\text{non-events}]$ and $\hat{E}[\tilde{f}^T \tilde{f}|\text{events}]$ are $K2 \times K2$ and $K1 \times K1$ diagonal matrices respectively with decreasing entries. We then make the treatment factors orthogonal to base factors on events $\hat{E}[f^T \tilde{f}|\text{events}] = 0$. The loadings and intercept mappings for base and treatment factors restricted by the following set of conditions: $\Gamma_\beta^T \Gamma_\beta = \mathbf{I_{K_1 \times K_1}}$, $\tilde{\Gamma}_\beta^T \tilde{\Gamma}_\beta = \mathbf{I_{K_2 \times K_2}}$, $\Gamma_\beta^T \Gamma_\alpha = \mathbf{0_{K_1 \times 1}}$, $\tilde{\Gamma}_\beta^T \tilde{\Gamma}_\alpha = \mathbf{0_{K_2 \times 1}}$, $\Gamma_\beta^T \tilde{\Gamma}_\alpha = \mathbf{0_{K_2 \times 1}}$.

3.2 Treatment Factor Premium

The normalization described in the previous section makes treatment factors orthogonal to base factors allowing us to compare their returns. We plot average returns on base factors (in black) over the whole sample and treatment factors (in red) on events in Figure 6 along with 95% confidence bands for the same event types we considered previously. There is a question of how many base and treatment factors to include. To determine this, we consider the placebo specification with odd and even weeks. In IPCA, including more factors will reduce the relevance of the marginal factor and will drive the premium of this factor to zero. We already determined that the odd and even weeks factor structure is very similar. Therefore, as we include more base factors and estimate the treatment factor, its return declines the more base factors we include. When we include four base factors, additional factors earn returns statistically indistinguishable from zero, as does the treatment factor. We illustrate this in the first panel of Figure 6. We see that the returns on base factors decline with the factor number, albeit not monotonically. The treatment factor has average returns that are statistically indistinguishable from zero.

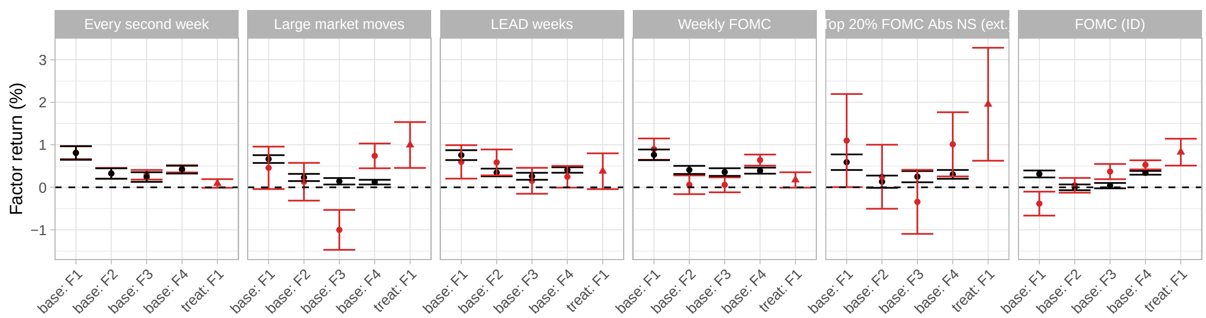
When we consider other event types, the picture looks different. When the event is defined as a top-10% of weeks by absolute market moves shown in the second panel of Figure 6, the returns on the base factors also decline with the factor number. The return on the treatment factor is, on the other side, large and statistically different from zero. Moreover, its magnitude is similar to that of the first base factor.

The treatment factor earns positive returns on LEAD weeks and the sample of all FOMC announcements. However, we cannot reject the hypothesis that these returns are statistically different from zero. When we narrow the sample of FOMC announcements to the ones with the largest NS shocks, the treatment factor earns a significant premium that is considerably larger in magnitude than the first base factor. While we documented previously in Figure **??** that the first factor earns larger returns on events, the treatment factor is orthogonal to the first factor on events, thus generating a distinct source of returns.

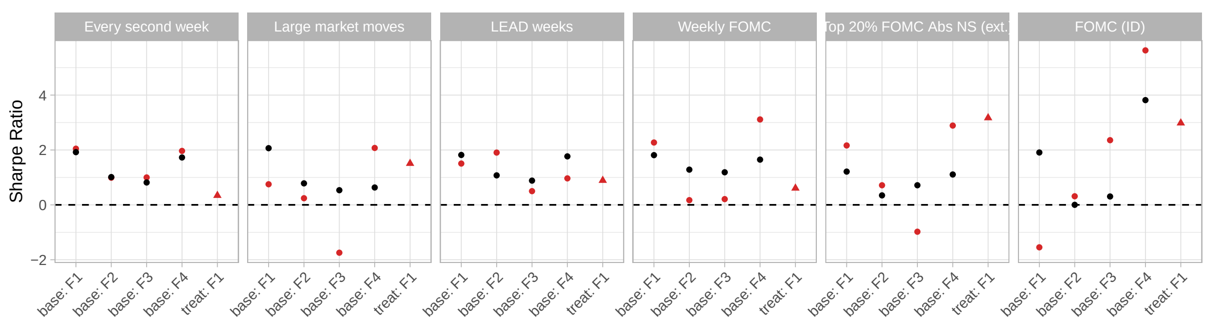
3.3 Treatment Factor Loadings

We present the loadings on the treatment factor in Figure 7. For large market moves, the treatment factor loads on short-term reversal and price relative to 52-week. These coefficients are larger in magnitude than for the placebo specification. While the estimation procedure does not consider the returns of the characteristics managed portfolios, we note that these portfolios have the most extreme returns during large market moves as indicated in Figure 2. This significant difference in portfolio returns is reflected in the average return of this treatment factor in the second panel of Figure 6. As a result, for large market moves, the characteristics that explain additional **orthogonal** variation in the factor structure are the ones that have significantly different returns.

For the largest Nakamura-Steinsson shocks FOMC announcements, the treatment factor has



(a) Factor Returns (over the period of interest)



(b) Sharpe Ratio (annualized)

Figure 6: Treatment Factor Returns

a large loading on price relative to a 52-week high. The return on the corresponding portfolio is significantly different on these events as indicated in Figure 2 generating a large premium on this factor as indicated in the last panel of Figure 6. At the same time, we see in Figure 2 a handful of other portfolios with average returns that are significantly affected by the event: short-term reversal (strev), leverage (lev) and others. These characteristics do not appear in the treatment factor with a magnitude comparable to momentum and price relative to a 52-week high.

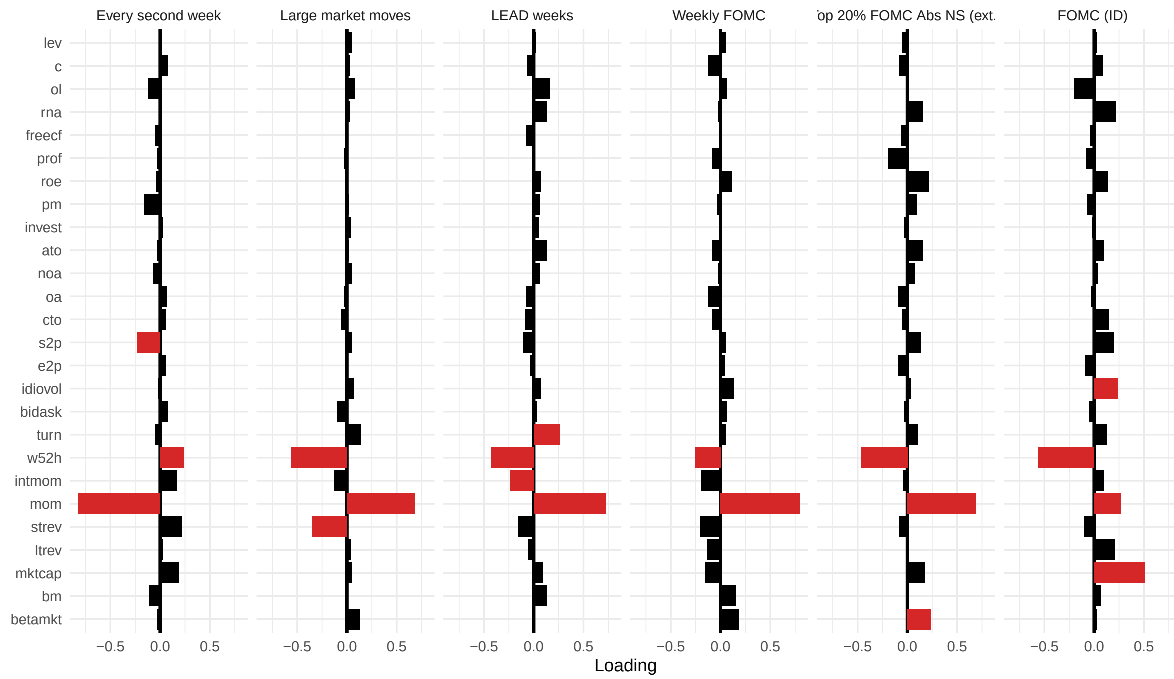


Figure 7: Treatment Factor Loadings

4 Boosted IPCA

By design, the Treatment-IPCA estimates orthogonal factors specific to event: $t+1 \in A$. At the same time, the factor structure may vary in the way that changes the role of characteristics on a given event. For example, firms with high leverage, either financial or operational, may be particularly susceptible to changes in discount rates as the value of their cash flows net of costs of financing the debt or fixed costs will exhibit higher fluctuations. For such scenario we expect the effect of leverage and profitability characteristics to be different during periods when interest rates exhibit a lot of variation, for example, during the FOMC announcement and particularly when the shocks to interest rates are large. Similarly, momentum factor is known to exhibit sharp drawdown when the market experiences large moves as documented by **?**. Thus, the role of a momentum characteristic may differ during such period.

For a particular characteristic to change its role, an additional factor estimated on events should be both volatile and correlated with the base factors. To capture these properties, we introduce Boosted-IPCA that, in contrast to Treatment IPCA, does not impose the orthogonality restriction. Moreover, the second step estimation does not rely on any information about the first step factors including their loadings and volatilities. This allows the boosted factors to be correlated with the base factors and to adjust or even cancel the effect of characteristics that are important to the factor structure outside the events but are less or not important on the events. We use this approach to test the changing role of characteristics on specified events.

The idea of boosting comes from the statistical theory of function approximations and was first introduced in Friedman (2001). Its main idea is to approximate a function with a series of simple functions. The errors from the previous step are being used to fit a new simple function at each step. This procedure includes many steps where each simple model brings an incremental improvement to the overall model. While the procedure features two steps in our case, we use the boosting approach of refining residuals from the non-event-based estimates.

4.1 Boosted IPCA Optimization Problem

Boosted IPCA approaches the problem of capturing the cross-sectional variation specific to events differently. Instead of estimating the factor structure jointly, it proceeds in two steps. First, we estimate a standard IPCA on the sample of non-events. Namely, we take the estimates $\hat{\Gamma}_\alpha, \hat{\Gamma}_\beta, \hat{f}_{t+1}$ of mapping $(\mathcal{G}^{NE}, \mathcal{F}^{NE})$. Then the factors are estimated on events as in equation (9), and residuals on events are obtained

$$\hat{\epsilon}_{t+1} \equiv r_{t+1} - Z_t \left(\hat{\Gamma}_\alpha + \hat{\Gamma}_\beta \hat{f}_{t+1}\right) \ \forall \ t+1 \in \text{ events} \tag{12}$$

Residuals $\hat{\epsilon}_{t+1}$ are then used in the second step, where we estimate an IPCA on the sample of events

$$\breve{\Gamma}_\alpha, \breve{\Gamma}_\beta, \{\breve{f}\} = \arg \min_{\breve{\Gamma}_\alpha, \breve{\Gamma}_\beta, \{\breve{f}_{t+1}\}} \sum_{t \in \text{events}} \left\| \hat{\varepsilon}_{t+1} - Z_t \left(\breve{\Gamma}_\alpha + \breve{\Gamma}_\beta \breve{f}_{t+1}\right) \right\|^2 \tag{13}$$

4.2 Boosted Factor Distribution

Boosted IPCA does not impose any orthogonality restrictions between the boosted and the base factors allowing it to alter the influence of firm characteristics on market events. The boosted factor will have a large impact on the existing factor structure outside of events if it has both large volatilities relative to the base factors on events and if it is correlated with the base factors.

Volatility We compare the volatility of non-event implied base factors on events and boosted factors in Figure 8. The volatility of the boosted factor in the placebo specification is similar to the volatility of the fourth base factor. This is also the case for LEAD weeks and the sample of all FOMC announcements. For large market moves, the volatility of the boosted factor is above the volatility of the fourth factor but is still below the volatility of other factors. For these event types, the boosted factor is not likely to have a large influence on the existing factor structure.

The sample of the largest NS shocks FOMC announcements reveals that the volatility of the boosted factor is above the second, third, and fourth factors and is on par with the 1st. While this can indicate the presence of a strong fifth orthogonal factor, we will show by comparing the correlation that, in this case, the boosted factor aims to "correct" the existing factor structure created by the base factors.

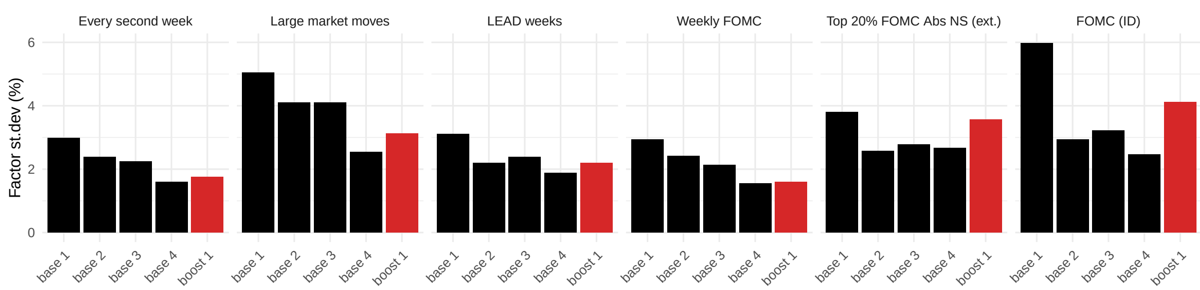


Figure 8: Comparing Volatility of Boosted and Base Factors

Correlation with the base factors We present the correlations of the boosted factor with implied base factors on events in Figure 9. In the first panel, placebo specification, the boosted factor has small correlations with the implied base factors. In the following panels we see much larger correlations. In particular, the boosted factor in the largest NS shocks FOMC announcements has significant correlations with all factors and especially with the third factor. For this event, the boosted factor both has a very large volatility and is significantly correlated with the base factors.

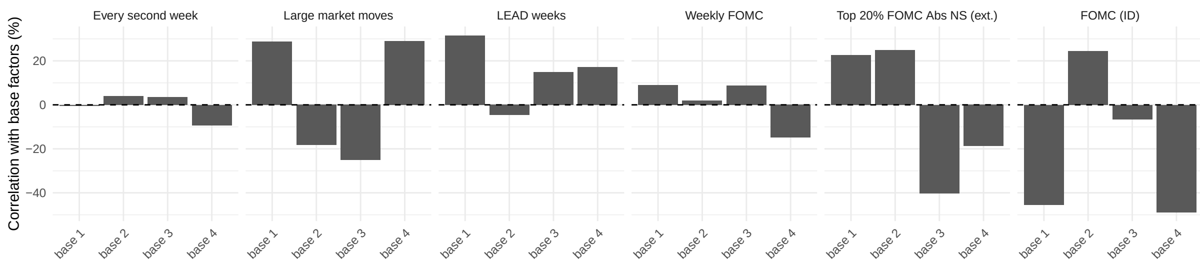


Figure 9: Correlation of Boosted Factor with Implied Base Factors on Events

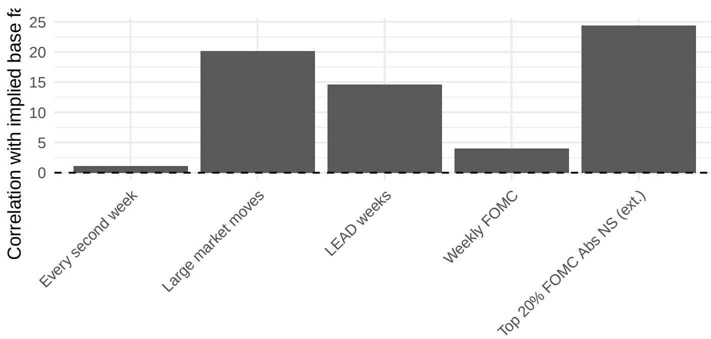


Figure 10: R^2 from Regression of Boosted Factor on Implied Base Factors on Events

4.3 Testing the Boosted Factors

We use Boosted IPCA to test the changes in the factor structure on events.

Booststrap procedure We follow the bootstrap procedure outlined in KPS to test the boosted factor. However, we introduce an important modification that allows us to test whether the boosted factor is different from an additional factor that may also appear outside the events. We test the loadings on the boosted factor against the null hypothesis that the loadings are equal to the loadings on the boosted factor estimated outside the events. Denote the time period in the subsample of events as $\mathcal{T}^E \subset \mathcal{T}$ and the subsample of non-events as $\mathcal{T}^{NE} \equiv \mathcal{T} \setminus \mathcal{T}^E$.

Boosted IPCA described above estimates IPCA using $\mathcal{T}^{NE}$, calculates implied factors on $\mathcal{T}^E$, and uses the residuals on $\mathcal{T}^E$, $\widehat{\varepsilon}$, to estimate a single factor IPCA. This results in loadings $\hat{\Gamma}_\alpha, \hat{\Gamma}_\beta$

for base factors $\hat{f}_t$, loadings $\breve{\Gamma}_\alpha, \breve{\Gamma}_\beta$ for boosted factor $\breve{f}_t$. The null hypothesis, on the other hand, is obtained by estimating IPCA on $\mathcal{T}^{NE}$ and calculating residuals on the same subsample $\mathcal{T}^{NE}$. These residuals are then used to estimate a single factor IPCA to obtain loadings $\ddot{\Gamma}_\alpha, \ddot{\Gamma}_\beta$.

We use the estimate $\ddot{\Gamma}_\beta$ as the null hypothesis for our bootstrap procedure. The first order conditions in IPCA do not directly rely on the equity returns but rather on the return of characteristic managed portfolios. This makes the bootstrap procedure both faster and easier to implement. We calculate the residuals on the portfolios after including the boosted factor on subsample $\mathcal{T}^E$

$$\begin{aligned} Z_t^T u_{t+1} &\equiv Z_t^T \left(\hat{\varepsilon}_{t+1} - Z_t \left(\breve{\Gamma}_\alpha + \breve{\Gamma}_\beta \breve{f}_{t+1} \right) \right) \\ &= Z_t^T \hat{\varepsilon}_{t+1} - Z_t^T Z_t \left(\breve{\Gamma}_\alpha + \breve{\Gamma}_\beta \breve{f}_{t+1} \right) \ \forall \ t+1 \in \mathcal{T}^E. \end{aligned} \tag{14}$$

We construct the bootstrapped version of portfolios $\widehat{\varepsilon}_{t+1}$ for testing a characteristic l as

$$Z_t^T \widehat{\varepsilon}_{t+1}{}^{BS(l)} \equiv Z_t^T Z_t \left(\breve{\Gamma}_\alpha + \breve{\Gamma}_\beta{}^{BS(l)} \breve{f}_{t+1} \right) + \chi \cdot Z_t^T u_{unif(\mathcal{T}^E)} \ \forall \ t \in \mathcal{T}^E, \tag{15}$$

where

$$\breve{\Gamma}_\beta{}^{BS(l)} \equiv (\breve{\Gamma}_{\beta,1}, ..., \breve{\Gamma}_{\beta,l-1}, \ddot{\Gamma}_{\beta,l}, \breve{\Gamma}_{\beta,l+1}, \dots)^T \tag{16}$$

is the vector of loadings where we replace the loading on the characteristic of interest with its value from non-event boosted IPCA. Multiplier χ is a student-t random variable with unit standard deviation and five degrees of freedom that allows accounting for heteroskedasticity in the data, and $u_{unif(\mathcal{T}^E)}$ is the resampled residual from the uniform distribution over time periods $t+1 \in \mathcal{T}^E$. Note that this residual is resampled for the whole time period, similar to a block bootstrap that accounts for clustering of errors within the same time period.

We use residuals $\widehat{\varepsilon}_{t+1}{}^{BS(l)}$ to estimate a single factor boosted IPCA and to obtain loadings $\breve{\Gamma}_\beta{}^{BS(l)}$. We repeat the bootstrap procedure 1000 times and calculate the p-value of coefficient $\breve{\Gamma}_{\beta,l}$ as the share of samples where $\left(\breve{\Gamma}_\beta{}^{BS(l)}\right)^2$ is above $\breve{\Gamma}_{\beta,l}$.

Bootstrap estimation results We present the estimation results for the event types we considered previously in Figure 11. It shows estimates for loadings on the boosted factor where the grey bars indicate no significance and different colors indicate different significance levels: green at 10% level, yellow at 5%, and red at 1%. Both large market moves and the sample of all FOMC announcements reveal that the effect of momentum and short-term reversal characteristics is different on events and non-events.

When we consider the largest NS shocks FOMC announcements in the last panel of Figure 11, we note that the momentum loading is insignificant even though its point estimate is large. This is consistent with this characteristic being a part of the fifth factor common across all time periods and its effect on returns not significantly changing on such events. At the same time, we see significant loading on profitability. Differences in profitability indicate differences in operational leverage that will affect the sensitivity of firm value to unexpected changes in interest rates.

Figure 13 presents the distribution of bootstrapped loadings for momentum (mom), profitability (prof), and price relative to 52-week high (w52h). The solid red vertical line shows the true estimated loading, while the dashed black line shows the null hypothesis for the boosted loading estimated on non-events. The distribution for momentum has a non-normal shape that appears due to the normalization of the magnitude of loadings $\breve{\Gamma}$ and the large magnitude of the estimated coefficient on momentum. The estimated loading falls within this distribution, reflected in its insignificance. Profitability and price relative to 52-week high have a more normal distribution. We can see that the estimated coefficients fall into the tails of the bootstrapped distribution, indicating their significance.

We compare the positive and negative NS shocks in Figure 12. The first panel repeats the estimates for the largest NS shocks FOMC announcements. The second panel presents the estimates for the largest positive NS shocks. We document that leverage and sales-to-price ratio have significantly different loadings from the non-event estimated boosted factor. Firms with higher leverage will be affected disproportionately more by an unexpected increase in interest rates and the path of interest rates due to the need to potentially refinance their mostly nominal debt at a higher rate, as well as due to their higher sensitivity of movements in interest rates and discount rates more generally. Firms with higher sales generate higher cash flows that let them weather increases in interest rates. These are also the firms with lower valuations, i.e., more value firms, that are less sensitive to movements in interest rates.

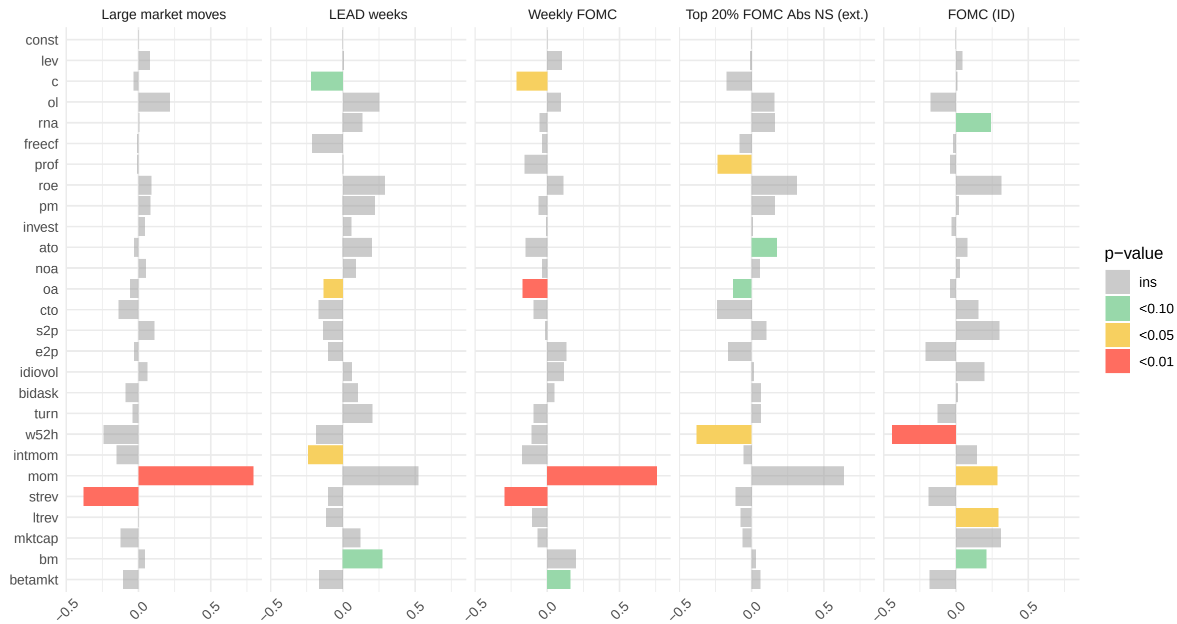


Figure 11: Boosted Factor Loadings with Bootstrap Significance Levels

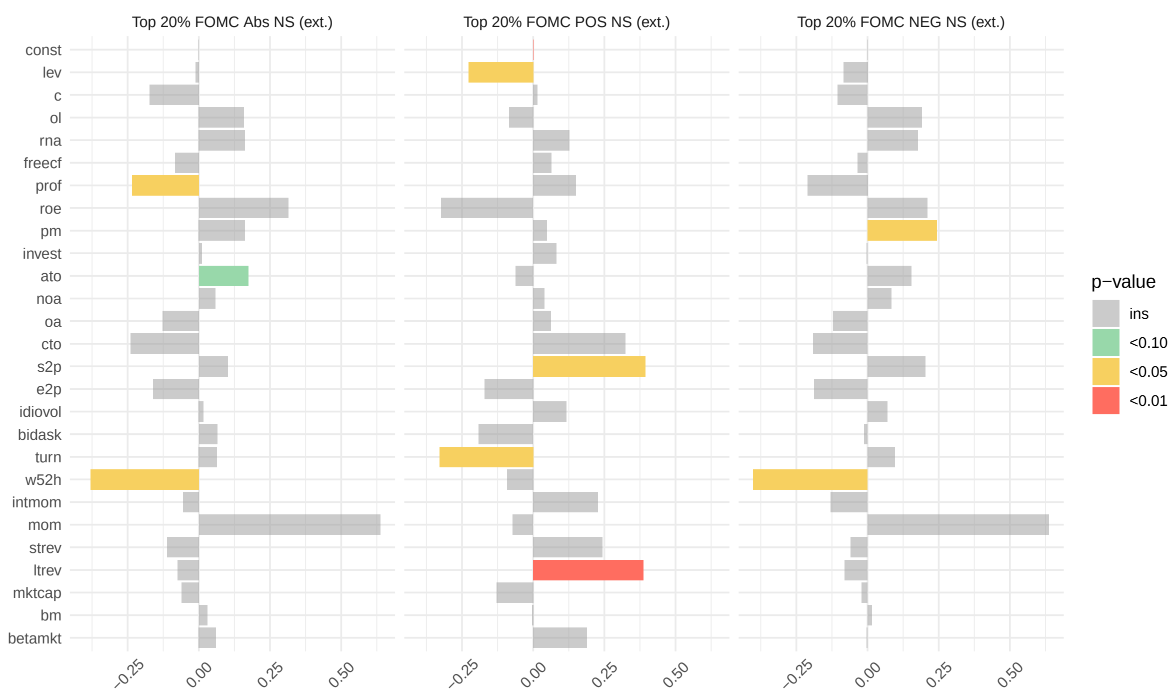


Figure 12: Boosted Factor Loadings: Event Identified with Nakamura-Steinsson Shocks

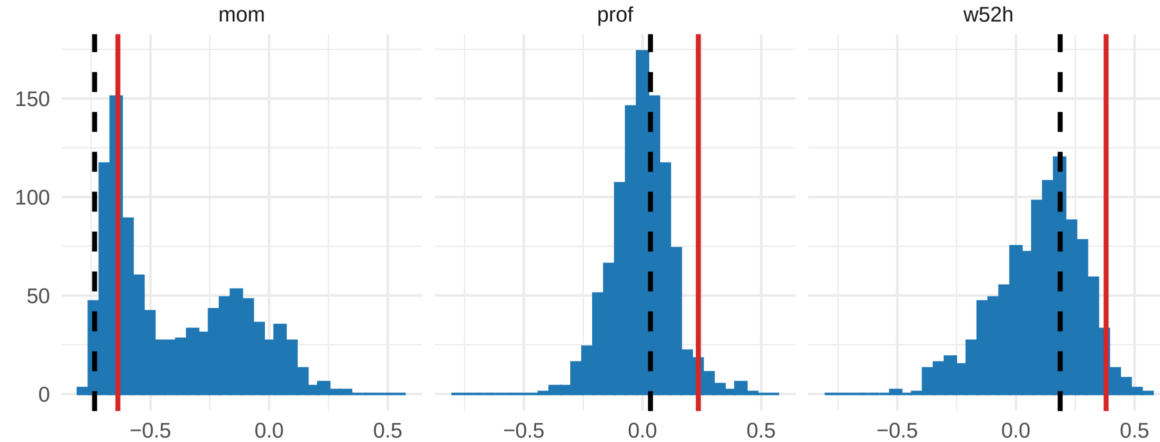


Figure 13: Examples of Bootstrapped Distributions

5 Simulations

The discussion of Boosted and Treatment IPCA is impossible without checking its robustness to misspecification of the defined model. One must recognize that a particular choice of the underlying firm characteristics, the number of factors for both base and treatment procedure ultimately will disturb the final loading picture together with the factor behavior. A researcher must hope that if some firm characteristic has a relative importance on the event then one or another way it must show up in the ultimate answer even if the model is not perfectly specified.

In this section, we do a simulation exercise that will test an environment with the misspecification. We simulate an event factor structure that cannot be perfectly captured by either treatment IPCA specification (10) or boosted structure to test if the estimated treatment and boosted factors reveal some useful patterns in the data. Nevertheless, both BIPCA and TIPCA estimated factors capture the major difference in the tractable way.

To be concrete, consider an environment where returns exhibit a four-factor structure on both event and non-event periods. However, the loadings for all four factors vary between these two periods. To make the example tractable imagine that the first two factors mainly capture size and book-to-market placing different exposure to the characteristics while third and fourth factor capture momentum and profitability. The loadings on the factors are captured by top panel of Figure 14. The other 23 loadings are chosen to be random and relatively small for the example.[3]

[3]The other base-factor loadings are chosen to be randomly distributed on the interval [-0.05,0.05].

The event four factor loadings are all chosen to have some small noise in addition to the standard loadings.[4] The major difference between the event factors appears in their correlation with the short-term reversal and turnover characteristics reported by red columns in bottom panel of Figure 14. The deliberate design of the loading choice was made to ensure two factors to be positively correlated to short-term reversal while only one to be correlated with turnover.

The choice of events and non-events, as well as the volatility of the factors, does not significantly affect the results we report in the following sections. However, for the sake of tractability, we will use the actual firm characteristics to generate returns, along with the number of periods and event labels specified in the "large market movements". The distribution of the factors on events and non-events is chosen to mimic the distribution for the estimated factors in Section 2.5.[5] Finally, the idiosyncratic component of firm returns is bootstrapped from the residuals of the IPCA procedure.

Let us apply Boosted-IPCA and Treatment-IPCA to simulated data and see if the procedures capable to capture the difference in the exposures. Note that since the estimated model is misspecified, it is impossible to perfectly uncover the factor structure on events by both procedures. At the same time, the first stage of Boosted-IPCA reveals the true factor structure quite well (see 15). The BIPCA-estimated first and second (third and forth) factor loadings capture size and book-to-market (momentum and profitability) in almost identical proportions to the one used for the underlying process. We noticed a slight contamination in the estimates. This contamination mainly arises from characteristics of one factor showing up in others. This contamination results from the strong correlation between actual firm characteristics and estimation error in the factors. Nevertheless, by examining the estimates, we can confidently identify which firm characteristics are truly relevant for non-events.

The base loadings of the Treatment-IPCA are mostly similar to the ones of the underlying processes. However, there is one noticeable difference: the third factor mixes loadings for both events and non-events, reflecting that unlike BIPCA the TIPCA-base factors attempt to fit both events and non-events. It is important to note that the captured turnover is smaller than the modeled loading on turnover for the event. It is worth noting that the short-term reversal was not observed in the base structure, which can only be considered satisfactory if it is identified in the treated factor.

[4] They are chosen to be randomly distributed on the interval [-0.05,0.05].

[5] specifically, it is bootstrapped from the separate event/non-event IPCA-estimates of the factors.

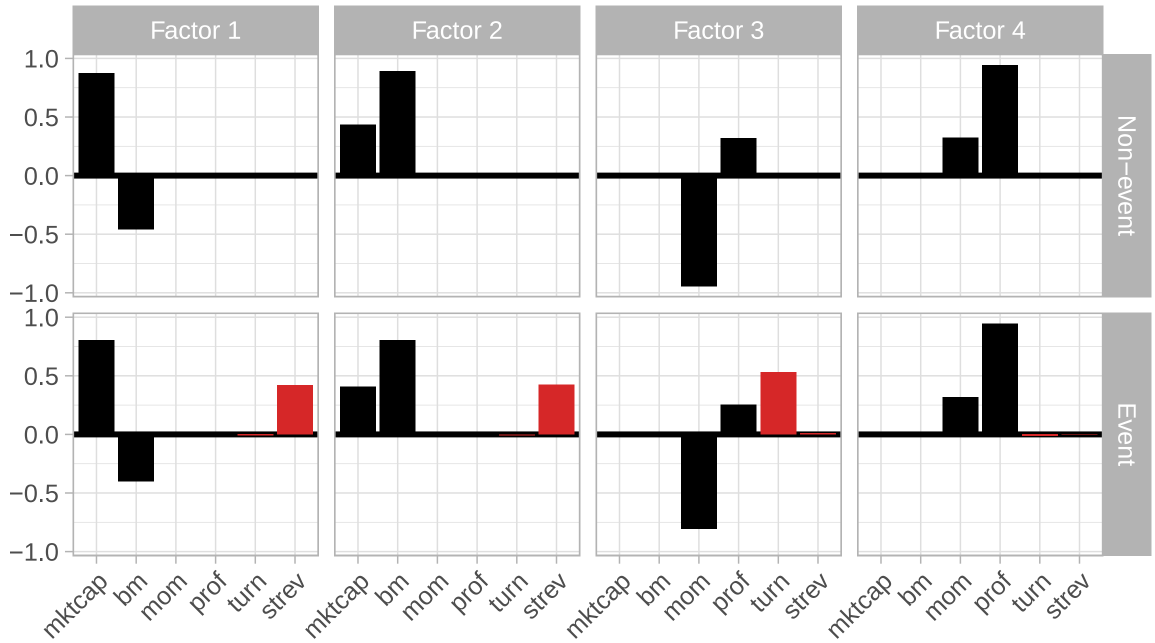


Figure 14: Loadings for Simulation Example

The figure represents the largest loadings in the simulated example. The top (bottom) panel corresponds to the underlying Γ used to produce returns on non-event (event) periods by four factors. All other firm characteristics can be found in Figure E.7.

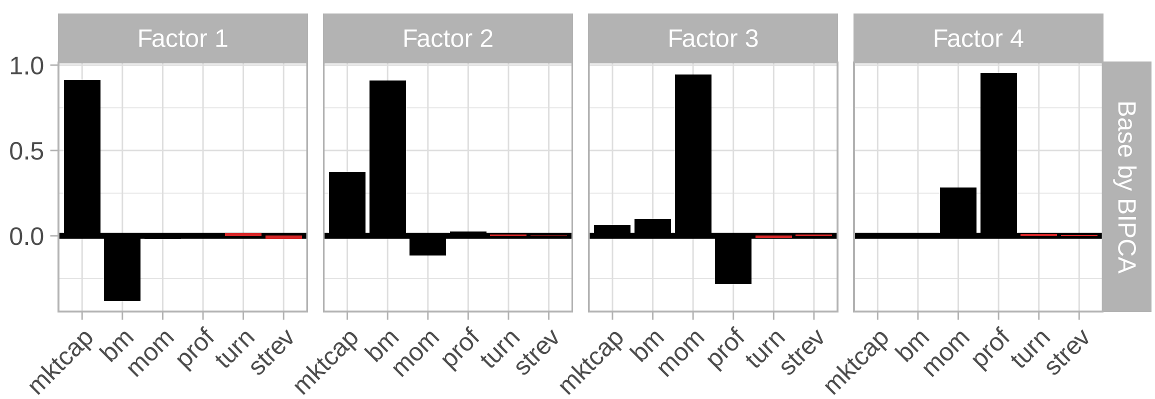


Figure 15: Simulation: BIPCA-Estimated Base (Non-Event) Loadings

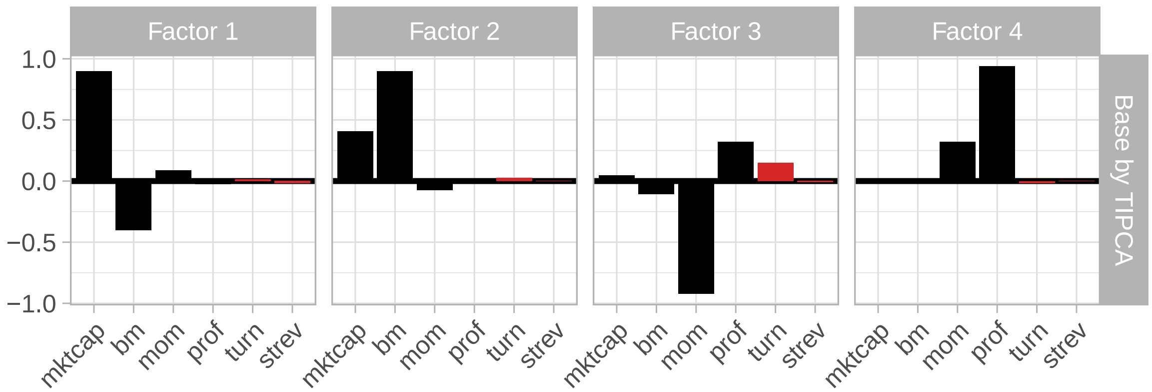


Figure 16: Simulation: TIPCA-Estimated Base (Non-Event) Loadings

We wish to have most striking difference to be concentrated in the boosted and treatment factors. When we estimate with only one factor, we limit the amount of information it can convey. Therefore, we want this factor to convey the most valuable information. The left panel of Figure 17 demonstrates that the boosted factor successfully extracts the information about short-term reversal that was added to the two most volatile factors. It also captures some changes in turnover, although the loading is small due to the weaker relative importance of turnover. It is worth noting that the boosted factor places minimal weight on other firm characteristics and primarily captures size and momentum that have relatively lower absolute value during events.

The treatment factor alone cannot fully account for short-term reversal. It considers the correlation between short-term reversal and other components. One of these components is market capitalization, which has a strong positive correlation with short-term reversal in both factors. On the other hand, book-to-market has a weaker presence in the boosted factor because it has an opposite sign correlation with short-term reversal in the factors. The example shows limits of interpretability for a one treatment factor. On the one hand, it shows that the difference appears somewhere in connection to short-term reversal and market capitalization that is true. On the other hand, it misses a third importantly related characteristic (bm).

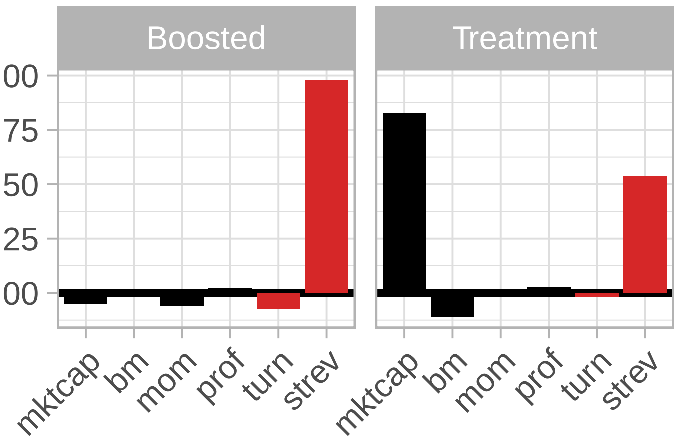


Figure 17: Simulation: Loadings of Boosted- and Treatment- Factors

6 Conclusion

This paper studies the classic linear cross-sectional relationship among stock returns, emphasizing time-varying risk-exposures and the distribution of factors. Two primary contributions stand out. Firstly, we unveil that significant market events, encompassing macroeconomic announcements and periods of pronounced market fluctuations, induce distinct factor structures, evident either through alterations in factor composition or shifts in factors' risk premia. Secondly, we propose a novel methodology to capture these event-specific differences, involving the extraction of orthogonal factors and exploring variations in the impact of firm characteristics on explaining the factor structure.

References

Barillas, F. and J. Shanken (2018, April). Comparing Asset Pricing Models: Comparing Asset Pricing Models. The Journal of Finance 73(2), 715–754.

Bryzgalova, S., V. DeMiguel, S. Li, and M. Pelger (2023). Asset-Pricing Factors with Economic Targets. SSRN Electronic Journal.

Bryzgalova, S., M. Pelger, and J. Zhu (2019). Forest Through the Trees: Building Cross-Sections of Stock Returns. SSRN Electronic Journal.

Chan, K. F. and T. Marsh (2022, June). Asset pricing on earnings announcement days. Journal of Financial Economics 144(3), 1022–1042.

Connor, G. and R. A. Korajczyk (1993). A Test for the Number of Factors in an Approximate Factor Model. Journal of Finance 48(4), 1263–1291.

Eisfeldt, A. L., E. T. Kim, and D. Papanikolaou (2022). Intangible Value. Critical Finance Review (11), 299–332.

Fama, E. F. and K. R. French (1992, June). The Cross-Section of Expected Stock Returns. The Journal of Finance 47(2), 427–465.

Fama, E. F. and K. R. French (2015, April). A five-factor asset pricing model. Journal of Financial Economics 116(1), 1–22.

Freyberger, J., A. Neuhierl, and M. Weber (2020, May). Dissecting Characteristics Nonparametrically. The Review of Financial Studies 33(5), 2326–2377.

Friedman, J. H. (2001, October). Greedy function approximation: A gradient boosting machine. The Annals of Statistics 29(5).

Kelly, B. T., S. Pruitt, and Y. Su (2019, December). Characteristics are covariances: A unified model of risk and return. Journal of Financial Economics 134(3), 501–524.

Korajczyk, R. A. and G. Connor (1988, September). Risk and return in an equilibrium APT: Application of a new test methodology. Journal of Financial Economics 21(2), 255–289.

Kozak, S., S. Nagel, and S. Santosh (2018, June). Interpreting Factor Models. The Journal of Finance 73(3), 1183–1223.

Kozak, S., S. Nagel, and S. Santosh (2020, February). Shrinking the cross-section. Journal of Financial Economics 135(2), 271–292.

Lakonishok, J., A. Shleifer, and R. W. Vishny (1994, December). Contrarian Investment, Extrapolation, and Risk. Journal of Finance 49(5).

Lettau, M. and M. Pelger (2020, September). Estimating latent asset-pricing factors. Journal of Econometrics 218(1), 1–31.

Nakamura, E. and J. Steinsson (2018, August). High-Frequency Identification of Monetary Non-Neutrality: The Information Effect*. The Quarterly Journal of Economics 133(3), 1283–1330.

Ross, S. A. (1976). The Arbitrage Theory of Capital Asset Pricing. Journal of Economic Theory 13, 341–360.

Savor, P. and M. Wilson (2014, August). Asset pricing: A tale of two days. Journal of Financial Economics 113(2), 171–201.

Wachter, J. A. and Y. Zhu (2022, April). A Model of Two Days: Discrete News and Asset Prices. The Review of Financial Studies 35(5), 2246–2307.

Zhang, L. (2005, February). The Value Premium. The Journal of Finance 60(1), 67–103.

A Characteristics

We use the following characteristics to construct the cross-sectionally normalized characteristics matrix Z_t: market beta (beta), sales-to- assets (ato), book-to-market (bm), cash-to-short-term-investment (c), capital turnover (cto), capital intensity (d2a), earnings-to-price (e2p), cash flow-to-book (freecf), idiosyncratic volatility with respect to the FF3 model (idiovol), investment (invest), leverage (lev), market capitalization (mktcap), turnover (turn), net operating assets (noa), operating accruals (oa), operating leverage (ol), profit margin (pm), gross profitability (prof), price relative to its 52-week high (w52h), return on net operating assets (rna), return on assets (roa), return on equity (roe), momentum (mom), intermediate momentum (intmom), short-term reversal (strev), long-term reversal (ltrev), sales-to-price (s2p), the ratio of sales and general administrative costs to sales (sga2s), bid-ask spread (bidask).

We use fundamental variables' definitions from the online appendix to Freyberger et al. (2020).

B PCA Simulations

Sample generation In this section, we show that that a strong factor structure will imply a string correlation between principal factors even if the samples used to extract the principal components are completely independent from one another. Suppose that there are two independent sample $j = N, E$ that will represent non-events and events, respectively. For each j we **independently** sample (1) $T \times K$ matrix F^j of factors where each entry is $\mathcal{N}(0, 1)$, (2) $N \times K$ matrix B^j of factor loadings where each entry is $\mathcal{N}(0, 1)$ and (3) $T \times N$ matrix of disturbances E^j where each entry is $\mathcal{N}(0, 3)$. The $T \times N$ matrix of returns is then

$$R^j = F^j(B^j)^{+}E^j. \tag{B.1}$$

We then calculate the $N \times N$ covariance matrix of returns and extract the first K principal components to obtain $N \times 1$ vectors of weights w_k^j for $k = 1, \ldots, K$. The first K principal components can be then calculated as

$$PC_k^{j \to j} = R^j w_k^j \text{ for } k = 1, \ldots, K. \tag{B.2}$$

With this notation the non-event factors on events can be calculated as

$$PC_k^{N \to E} = R^E w_k^N \text{ for } k = 1, \ldots, K. \tag{B.3}$$

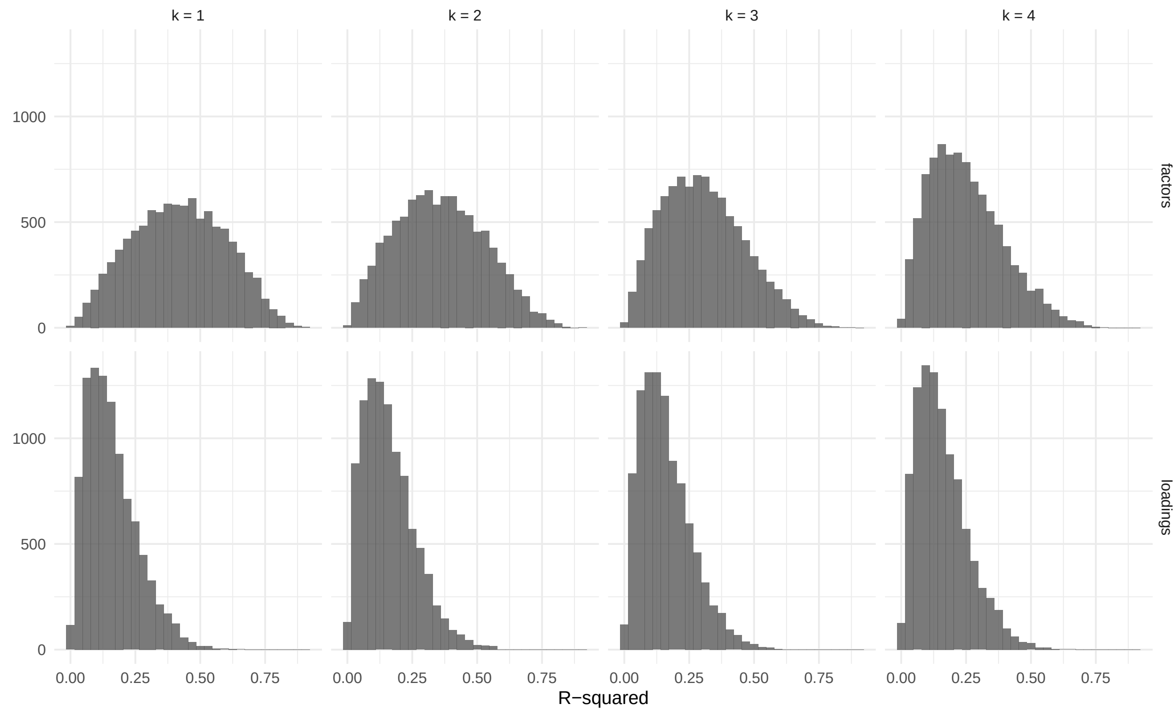


Figure B.1: How Well do Factors Span Each Other in Independent Samples?

We are interested in how well do these factors $\{PC_k^{N\to E}\}_{k=1,\ldots,K}$ span the actual event factors $\{PC_k^E\}_{k=1,\ldots,K}$. To do this we estimate a regression of PC_k^E on the full set of $\{PC_k^{N\to E}\}_{k=1,\ldots,K}$ and report the R^2.

Simulation results We repeat this full process 10,000 times and report the distribution of R^2 for each factor in the top row of Figure B.1. Contrary to the prior, we observe that the distribution of R^2 does not bunch close to zero, but instead is wide, centered at around 40% with the 95% confidence interval spanning 5% to 75%. This example shows that when the factor structure is strong, i.e. there are a few factors that describe a large share of variation in the cross-section, and there are few test portfolios, we tend to find similar factor structure in two samples even when it is not common.

To understand this result, consider a corner case where we have 4 factors and 4 assets. When we use PCA to extract 4 principal components from this cross-section, the eigenvectors in one subsample will necessarily span the eigenvectors in the other subsample and, as a result, so will the factors. In the language of our simulations above this will imply $R^2 = 1$. When we increase the number of assets while leaving the number of factors unchanged we introduce additional

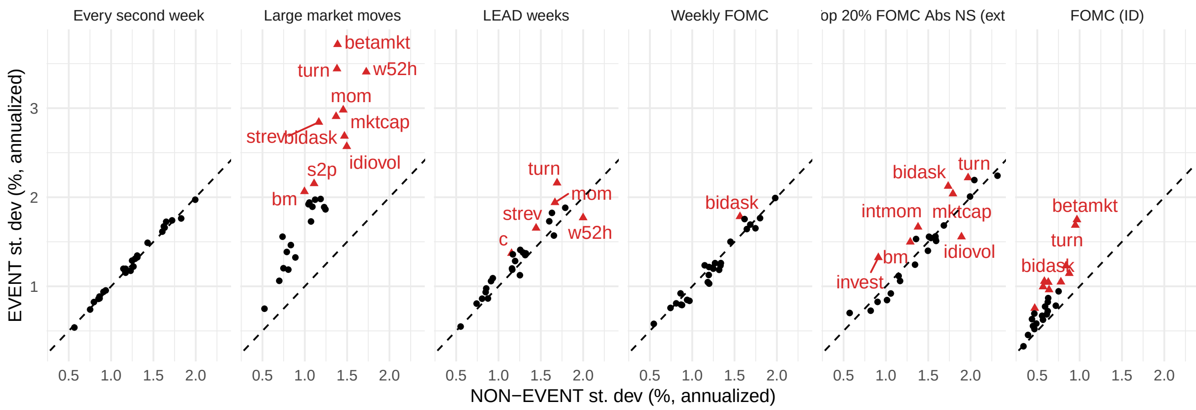


Figure B.2: Average Returns on Characteristic Portfolios (Columns)

variation that leads to imperfect spanning and resulting in $R^2 < 1$. However, the effective dimensionality still remains small since the number of factors is large relative to the size of the cross-section resulting in relatively large R^2.

C Change in Distribution of Fama-French Factors

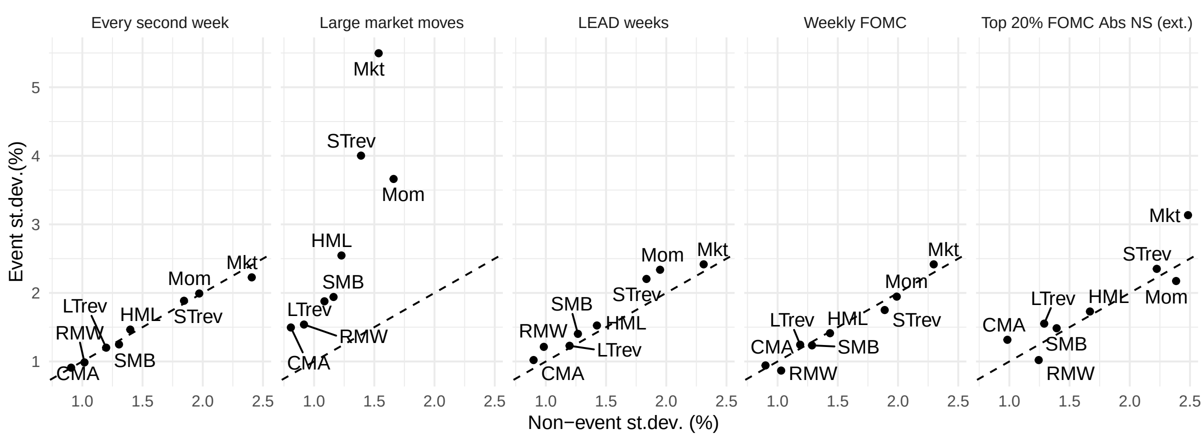


Figure C.3: Change in Volatility of Fama-French Factors

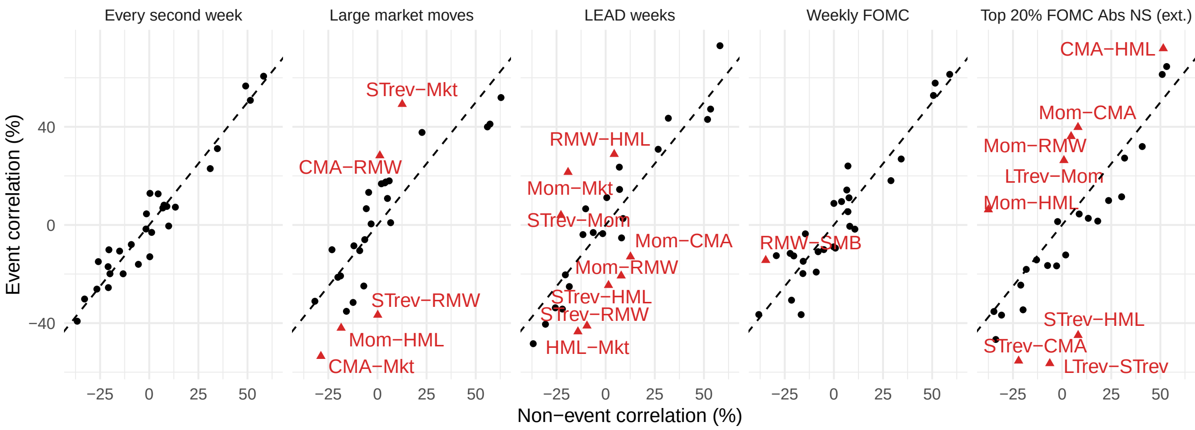


Figure C.4: Change in Correlation of Fama-French Factors

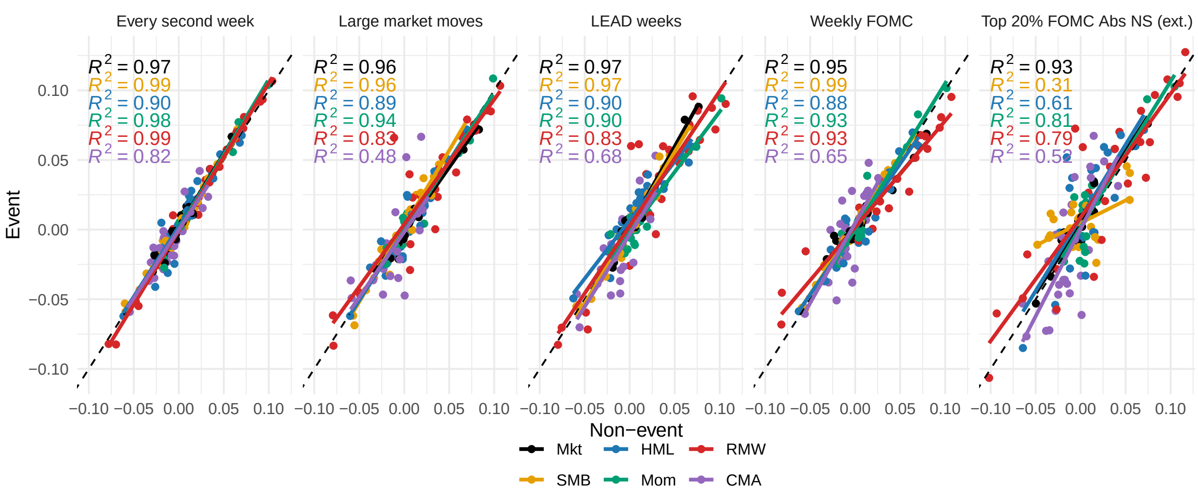


Figure C.5: Change in Fama-French Betas of Characteristic Managed Portfolios

D Treatment-IPCA: Implementation

D.1 Notation

Let us introduce notation and clarify dimensions of the used objects. Given that at time t, we observe cross-section of N firms, (K_1,K_2)-factor L-characteristic analysis uses the following vectors and matrices:

$$
\begin{array}{rrcl}
\text{returns} & r_{t+1} & : & N \times 1 \\
\text{firm chars} & z_t & : & N \times L \\
\text{char-alphas} & \Gamma_\alpha,\ \widetilde{\Gamma}_\alpha & : & L \times 1 \\
\text{char-exposures} & \Gamma_\beta,\ \widetilde{\Gamma}_\beta & : & L \times K_1 \text{ and } L \times K_2 \\
\text{factors} & f_{t+1},\ \widetilde{f}_{t+1} & : & K_1 \times 1 \text{ and } K_2 \times 1 \\
\text{augmented factors} & \check{f}_{t+1},\ \check{\widetilde{f}}_{t+1} & : & (K_1+1) \times 1 \text{ and } (K_2+1) \times 1 \\
\text{vectorized exposures} & V_\beta,\ \widetilde{V}_\beta & : & LK_1 \times 1 \text{ and } LK_2 \times 1
\end{array}
\tag{D.4}
$$

Let us also denote the indicator of event $\mathcal{A}$ at time t by $\mathbb{I}_t(\mathcal{A}) = I(t \in \mathcal{A})$.

The circle will stand for joining of standard elements and the tilde indicators:

$$
\mathring{\Gamma}_\beta = (\Gamma_\beta, \widetilde{\Gamma}_\beta), \quad \mathring{f}^{\top}_{t+1} = (f^{\top}_{t+1}, \widetilde{f}^{\top}_{t+1}) \tag{D.5}
$$

The global minimization problem is

$$
\min_{\substack{\{\Gamma_\alpha, \Gamma_\alpha, \Gamma_\beta, \Gamma_\beta\} \\ \{f_{t+1}, \widetilde{f}_{t+1}\}_{t \in \mathcal{T}}}} \sum_{t \in \mathcal{T}} \left\| r_{t+1} - z_t \left(\Gamma_\alpha + \Gamma_\beta f_{t+1} \right) - z_t \left(\widetilde{\Gamma}_\alpha + \widetilde{\Gamma}_\beta \widetilde{f}_{t+1} \right) \mathbb{I}_t(\mathcal{A}) \right\|^2 \tag{D.6}
$$

Without loss of generality assume that $\widetilde{f}_{t+1}$ is zero-vector for $t \in \mathcal{A}$.

The necessary and sufficient condition for global minimum of the problem is to form solution of *cross-section regression* and *time-Series regression*. Numerically, the solution is achieved by iteratively applying cross-section and time-series regression from an arbitrary starting point. The procedure is known as *altering least squares*. We begin with the estimates Γ_α and Γ_β obtained from the standard K_1-factor IPCA procedure applied to the sample, and $\tilde{\Gamma}_\alpha$ and $\tilde{\Gamma}_\beta$ simulated as random standard Gaussian matrices.

D.2 Cross-Section Regression

The cross-section regression targets obtaining f_{t+1} and $\tilde{f}_{t+1}$ by fitting

$$\min_{\{f_{t+1}, \widetilde{f}_{t+1}\}} \left\| r_{t+1} - z_t \left(\Gamma_\alpha + \Gamma_\beta f_{t+1}\right) - z_t \left(\widetilde{\Gamma}_\alpha + \widetilde{\Gamma}_\beta \widetilde{f}_{t+1}\right) \mathbb{I}_t(\mathcal{A}) \right\|^2 \qquad t+1 \in \mathcal{T}, \tag{D.7}$$

for given set $\{\Gamma_\alpha,\ \Gamma_\alpha,\ \Gamma_\beta,\ \Gamma_\beta\}$.

D.2.1 Normal Period (like in KPS)

If $t+1 \notin \mathcal{A}$, we solve:

$$\min_{\{f_{t+1}\}} \left(r_{t+1} - z_t \check{\Gamma} \check{f}_{t+1}\right)^\top \left(r_{t+1} - z_t \check{\Gamma} \check{f}_{t+1}\right), \tag{D.8}$$

where $\check{f}_{t+1}^\top = (1, f_{t+1}^\top)$ and $\check{\Gamma} = (\Gamma_\alpha, \Gamma_\beta)$. The latter is $L \times (1 + K_1)$ matrix.

The first order condition is then

$$\Gamma_\beta^\top z_t^\top \left(r_{t+1} - z_t \check{\Gamma} \check{f}_{t+1}\right) = 0, \tag{D.9}$$

which provides closed form solution for factor:

$$f_{t+1} = \left(\Gamma_\beta^\top z_t^\top z_t \Gamma_\beta\right)^{-1} \left(\Gamma_\beta^\top z_t^\top r_{t+1} - \Gamma_\beta^\top z_t^\top z_t \Gamma_\alpha\right). \tag{D.10}$$

D.2.2 Model with indicator (KPS)

If $t+1 \in \mathcal{A}$, we solve:

$$\min_{\{f_{t+1}, \widetilde{f}_{t+1}\}} \left(r_{t+1} - z_t \check{\Gamma} \check{f}_{t+1}\right)^\top \left(r_{t+1} - z_t \check{\Gamma} \check{f}_{t+1}\right), \tag{D.11}$$

where $\check{f}_{t+1}^\top = (1, f_{t+1}, \widetilde{f}_{t+1}^\top)$ and $\check{\Gamma} = (\Gamma_\alpha + \widetilde{\Gamma}_\alpha, \mathring{\Gamma}_\beta)$, where $\mathring{\Gamma}_\beta = (\Gamma_\beta, \widetilde{\Gamma}_\beta)$. The latter is $L \times (1 + K_1 + K_2)$ matrix. The solution is

$$(f_{t+1}, \widetilde{f}_{t+1}^\top)^\top = \left(\mathring{\Gamma}_\beta^\top z_t^\top z_t \mathring{\Gamma}_\beta\right)^{-1} \left(\mathring{\Gamma}_\beta^\top z_t^\top r_{t+1} - \mathring{\Gamma}_\beta^\top z_t^\top z_t \check{\Gamma}_\alpha\right). \tag{D.12}$$

D.3 Time-Series Regression

Given set of factors $\{f_t, \tilde{f}_t\}_{t\in\mathcal{T}}$, we find loadings using time-series regression:

$$\min_{\gamma:=\{\Gamma_\alpha,\Gamma_\alpha,\Gamma_\beta,\Gamma_\beta\}} \sum_t \left(r_{t+1} - z_t\check{\Gamma}\check{f}_{t+1}\right)^\top \left(r_{t+1} - z_t\check{\Gamma}\check{f}_{t+1}\right). \tag{D.13}$$

Denote vectorized parameters:

$$\left.\begin{aligned} V_\beta &\equiv vec\left(\Gamma_\beta^\top\right) \\ \widetilde{V}_\beta &\equiv vec\left(\widetilde{\Gamma}_\beta^\top\right) \end{aligned}\right\}$$

Rewrite objective function:

$$\sum_t \|r_{t+1} - \left(z_t\Gamma_\alpha + \underbrace{z_t\left(\underbrace{I_L\otimes f_{t+1}}_{LK\times L}\right)^\top}_{L\times 1} V_\beta + \mathbb{I}_t(\mathcal{A})\cdot z_t\widetilde{\Gamma}_\alpha + \mathbb{I}_t(\mathcal{A})\cdot z_t\left(I_L\otimes\widetilde{f}_{t+1}\right)^\top \widetilde{V}_\beta\right)\|^2 \tag{D.14}$$

$$\sum_t \left\| r_{t+1} - \underbrace{\left(z_t,\ z_t\left(I_L\otimes f_{t+1}\right)^\top,\ \mathbb{I}_t(\mathcal{A})\cdot z_t,\ \mathbb{I}_t(\mathcal{A})\cdot z_t\left(I_L\otimes\tilde{f}_{t+1}\right)^\top\right)}_{\equiv H_t\ :\ N\times(L+LK_1+L+LK_2)} \times \underbrace{\begin{pmatrix}\Gamma_\alpha\\ V_\beta\\ \widetilde{\Gamma}_\alpha\\ \widetilde{V}_\beta\end{pmatrix}}_{\gamma\ :\ (L+LK_1+L+LK_2)\times 1} \right\|^2 \tag{D.15}$$

$$H_t^\top = \begin{pmatrix} z_t^\top \\ \left(I_L\otimes f_{t+1}\right)z_t^\top \\ \mathbb{I}_t(\mathcal{A})\cdot z_t^\top \\ \mathbb{I}_t(\mathcal{A})\cdot\left(I_L\otimes\tilde{f}_{t+1}\right)z_t^\top \end{pmatrix} \tag{D.16}$$

The problem simplifies to

$$\frac{1}{2}\sum_t \|r_{t+1} - H_t\gamma\|^2 \to \min_\gamma \tag{D.17}$$

FOC:

$$\sum_t H_t^\top\left(r_{t+1} - H_t\gamma\right) = 0 \tag{D.18}$$

$$\sum_t H_t^\top r_{t+1} = \left(\sum M_t^\top H_t\right)\gamma \Rightarrow \gamma = \left(\Sigma_t H_t^\top H_t\right)^{-1} \sum_t H_t^\top r_{t+1} \tag{D.19}$$

To simplify the notation, let us omit the subscript t and consider only the upper diagonal elements of $H_t^\top H_t$[6]:

$$\begin{bmatrix} z^\top z & z^\top z \times (I_L \otimes f)^\top & \mathbb{I}_t(\mathcal{A}) z^\top z & \mathbb{I}_t(\mathcal{A}) \ \times z^\top z \ \times (I_L \otimes \widetilde{f})^\top \\ & z^\top z \otimes f f^\top & \mathbb{I}_t(\mathcal{A}) \left(z^\top z \otimes f\right) & \mathbb{I}_t(\mathcal{A}) \left(z^\top z \otimes f \widetilde{f}^\top\right) \\ & & \mathbb{I}_t(\mathcal{A}) z^\top z & \mathbb{I}_t(\mathcal{A}) \left(z^\top z \otimes \widetilde{f}^\top\right) \\ & & & \mathbb{I}_t(\mathcal{A}) \left(z^\top z \otimes \widetilde{f}\widetilde{f}^\top\right) \end{bmatrix} \tag{D.22}$$

Similarly,

$$H_t^\top r_{t+1} = \begin{bmatrix} z^\top r \\ (z^\top r \otimes f) \\ \mathbb{I}_t(\mathcal{A}) \cdot z^\top r \\ \mathbb{I}_t(\mathcal{A}) \cdot (z^\top r \otimes \widetilde{f}) \end{bmatrix} \tag{D.23}$$

D.4 Normalization

The main targeted normalization conditions are $\Gamma_\beta^T \Gamma_\beta = \mathbf{I}_{\mathbf{K_1}\times\mathbf{K_1}}$, $\tilde{\Gamma}_\beta^T \tilde{\Gamma}_\beta = \mathbf{I}_{\mathbf{K_2}\times\mathbf{K_2}}$, $\Gamma_\beta^T \Gamma_\alpha = \mathbf{0}_{\mathbf{K_1}\times\mathbf{1}}$, $\tilde{\Gamma}_\beta^T \tilde{\Gamma}_\alpha = \mathbf{0}_{\mathbf{K_2}\times\mathbf{1}}$, $\Gamma_\beta^T \tilde{\Gamma}_\alpha = \mathbf{0}_{\mathbf{K_2}\times\mathbf{1}}$ and factor orthogonality conditions: $\hat{E}[f^T \tilde{f}|\text{events}] = 0$, $\hat{E}[f^T f|\text{non-events}]$ and $\hat{E}[\tilde{f}^T \tilde{f}|\text{events}]$ are $K2 \times K2$ and $K1 \times K1$ diagonal matrices respectively with decreasing entries.

Firstly, we change the base factors and Γ-s such that $\Gamma_\alpha^{new} \times \Gamma_\beta^{new} = 0$, $\Gamma_\beta^{new} \times \Gamma_\beta^{new} = I$, $f_{t+1}^\top f_{t+1}$ is diagonal, using PKY rotation procedure:

$$A := \left(\Gamma_\beta^\top \Gamma_\beta\right)^{-1} \Gamma_\beta^\top \Gamma_\alpha$$

$$\Gamma_\alpha^{new} := \Gamma_\alpha - \Gamma_\beta A$$

[6]One must use the properties of Kronecker product:

$$\left(I_L \otimes f_{t+1}\right) z_t^\top z_t \left(I_L \otimes f_{t+1}\right)^\top = z^\top z \otimes f f^\top \tag{D.20}$$

$$\left(I_L \otimes f_{t+1}\right) z_t^\top z_t = z_t^\top z_t \otimes f_{t+1} \tag{D.21}$$

$$f_t^{new,0} := f_t + A, \quad t \in \mathcal{T}$$

$$f^{new} = R^{-1} \times f_t^{new,0}, \quad \Gamma_\beta^{new} = \Gamma_\beta \times R,$$

where rotation matrix R can be found through the Cholesky factorization $\Gamma_\beta^\top \Gamma_\beta = L^\top L$ and eigenvectors matrix of $|\mathcal{T} \backslash \mathcal{A}|^{-1} \sum_{t \notin A} f_t^{new,0} f_t^{new,0\top}$, V:

$$R = L^{-1} \times V.$$

The base factors are ordered based on their standard deviation based on $t \in \mathcal{T} \backslash \mathcal{A}$.

Assume that base loadings and factors Γ_β, Γ_α, and f are updated accordingly, and hence the entry of $\mathring{\Gamma}_\beta$ are also updated. Then, the next step is

$$\tilde{A} := \left(\mathring{\Gamma}_\beta^\top \mathring{\Gamma}_\beta \right)^{-1} \mathring{\Gamma}_\beta^\top \tilde{\Gamma}_\alpha$$

$$\tilde{\Gamma}_\alpha^{new} := \tilde{\Gamma}_\alpha - \mathring{\Gamma}_\beta \tilde{A}$$

$$(f_t^{new\top}, \tilde{f}_t^{new\top}) := (f_t^{new\top}, \tilde{f}_t^\top) + \tilde{A}^\top, \quad t \in \mathcal{A}$$

Next, we adjust $\tilde{\Gamma}_\beta$ to orthogonalize tilde factors to base factors.

$$\tilde{R} := \left(\sum_{t \in A} f_t^{new\top} \tilde{f}_t^{new} \right)^\top \left(\sum_{t \in A} \tilde{f}_t^{new} \tilde{f}_t^{new\top} \right)^{-1}$$

$$f_t^{new} := f_t^{new} - \tilde{f}_t^{new} \times \tilde{R}^\top, \quad t \in \mathcal{A}$$

$$\tilde{\Gamma}_\beta^{new} := \tilde{\Gamma}_\beta - \Gamma_\beta^{new} \tilde{R}$$

Next, the same rotation procedure with Cholesky factorization and eigenvector decomposition is applied to $\tilde{f}_t^{new}$ and $\tilde{\Gamma}_\beta^{new}$ to make the ultimate rotation updates of the matrices. Finally, the columns can be reordered to guarantee a decreasing standard deviation of the treatment factor. The signs of columns are selected to ensure the positive average outcome of base and treatment factors on non-event and events, respectively.

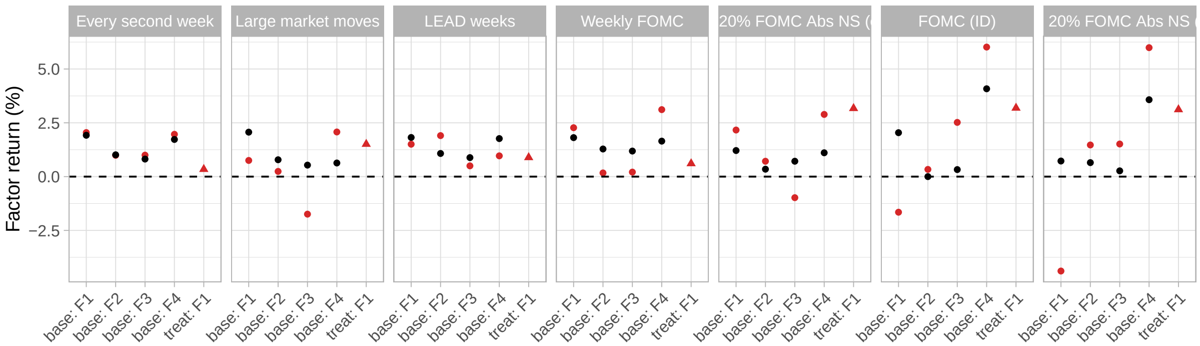


Figure D.6: Sharpe Ratio of Treatment and Base Factors

E Simulations

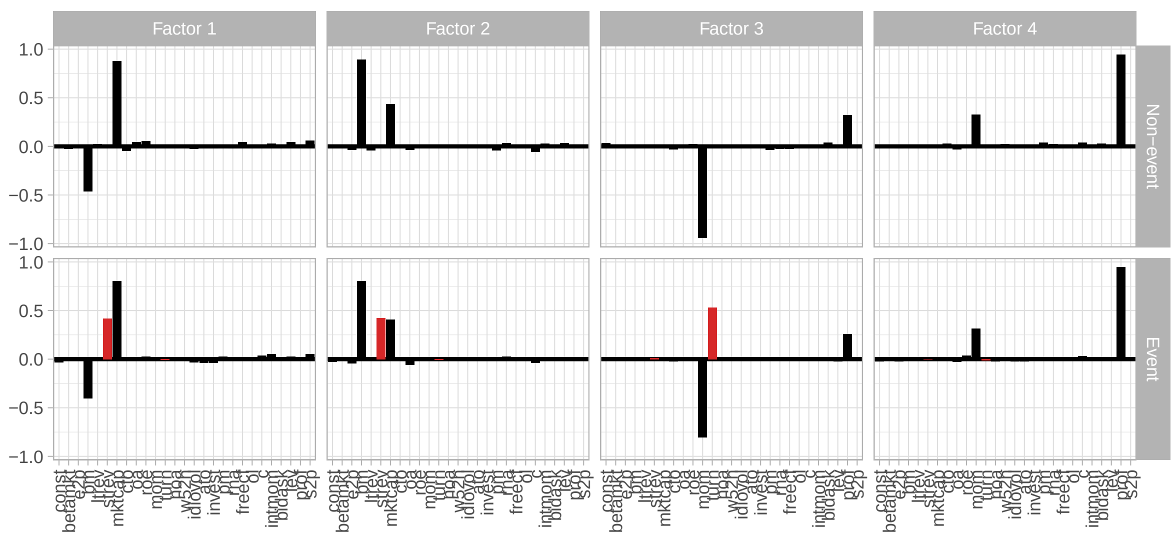


Figure E.7: Loadings for Simulation Example

F Possible Extension: High-Frequency Identification

The advantage of data stored in stock markets for assessing macro-events is that the structure is richer and potentially carries more information than just the shift in a bunch of indices and interest rates. The current progress of the paper is to develop additional measures of FOMC assessment based on the treatment and boosted factors estimated from the cross-section of stocks.

We provide a motivational picture that emphasizes the potential of this approach, highlighting the variability in characteristic managed portfolios.

Figure F.8 aggregates the information on how the cross-section of stock returns reacted to big FOMC announcements pre-selected based on NS shocks. For every characteristic portfolio, we plot the cumulative squared return normalized by predicted squared return on any other day. The latter must grow linearly over a day[7] and estimated via linear regression on data for non-FOMC days. The kink around 14:00 reflects the informational arrival and overall reaction to the monetary policy news. The high dispersion of the portfolio returns reflects that cross-sectionally different stocks are exposed differently to the monetary shocks. The adjustment of portfolios in the market according to their market exposure does not eliminate the cross-sectional difference in response to the shocks.

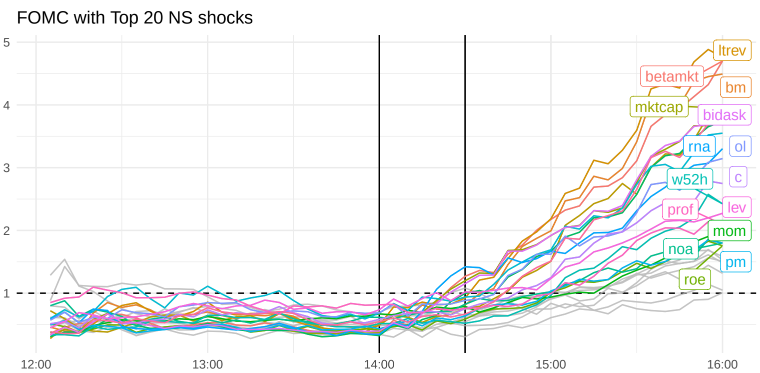


Figure F.8: Normalized Portfolio Cumulative Squared Log Return on FOMC Events

The plot reports cumulative squared log return for characteristic portfolios. The value is normalized by $\beta \cdot t$, where t is the time between 12:00 and reported time, and β is coefficient from the linear regression of the average cumulative squared log return of the same portfolio on time, $ret_p = \beta_p t + \epsilon$, for non-fomc dates.[a]

[a]usage of linear regression approximation gives almost same results but avoids a contamination when one normalizes the curve by

[7]If the log-returns had an expected return of zero and were independent and equally distributed time-series within a day, then this statement would be completely accurate.

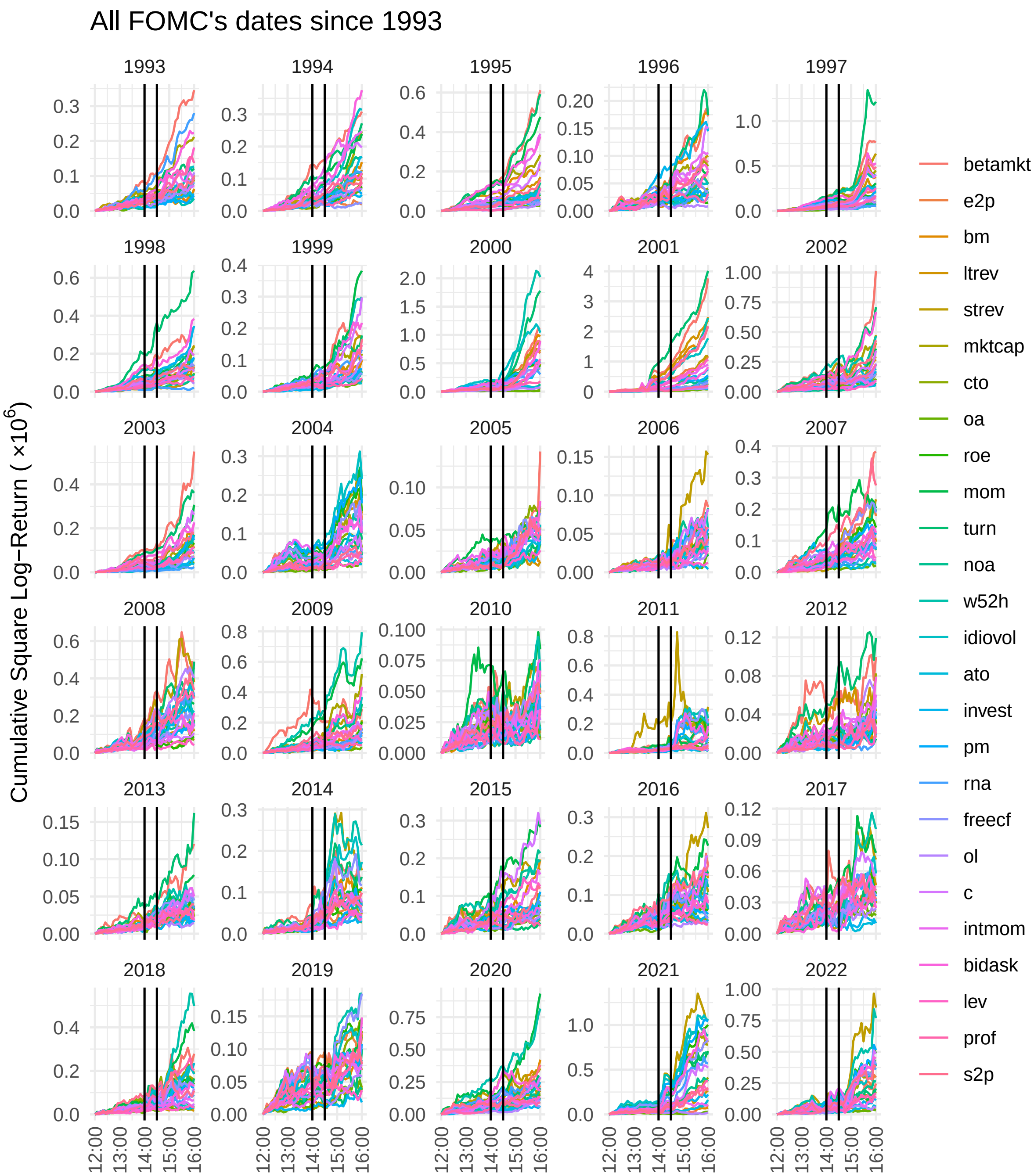


Figure F.9: Intraday Stock Market Reaction on FOMC Announcements