Journal of Financial Economics 180 (2026) 104278

Contents lists available at ScienceDirect

Journal of Financial Economics

journal homepage: www.elsevier.com/locate/finec

Past is prologue: Inference from the cross section of returns around an event

Jonathan B. Cohn [a],*, Travis L. Johnson [a], Zack Liu [b], Malcolm I. Wardlaw [c]

[a] University of Texas at Austin, McCombs School of Business, Austin, TX, United States
[b] University of Houston, Bauer College of Business, Houston, TX, United States
[c] University of Georgia, Terry College of Business, Athens, GA, United States

ARTICLE INFO

Dataset link: Past is Prologue: Inference from the Cross Section of Returns Around an Event replication kit (Original data)

JEL classification:
G14
C12
C21

Keywords:
Event studies
Inference
Standard errors
Clustering

ABSTRACT

Confounding events can cause false positives when testing the relationship between short-term returns around a quasi-experimental event and firm characteristics. We show that this risk is severe in practice — return-characteristic relationships are often statistically significant at the 1% level on over 30% of all trading days. Benchmarking a relationship against the distribution of the same relationship on pre-event days is effective at addressing the problem. We introduce a novel GLS variation of this approach that achieves large gains in statistical power relative to OLS and provide Stata and Python modules that implement both procedures.

1. Introduction

A common empirical strategy in corporate finance is to test whether short-term stock returns around a quasi-exogenous event, such as a surprise election outcome, are correlated with a firm characteristic that proxies for exposure, such as a firm's effective tax rate.[1] One threat to causal inference in such tests is that contemporaneous events may affect firm values in ways that vary with the same characteristic. If unaccounted for, such confounding events may lead a researcher to conclude that the focal event causes a relationship between returns and the characteristic when, in reality, it does not. However, while the threat to causal inference from confounding events is clear in theory, its importance in real-world applications is unclear. Moreover, while approaches to addressing the problem have been proposed in the literature (e.g., Sefcik and Thompson, 1986), they have been largely ignored, and their effectiveness in practice remains unclear.

To fix ideas, consider three events from 2019: a tentative agreement on Phase 1 of a trade deal between the U.S. and China reached on October 11; a U.K. House vote against a no-deal Brexit on March 13; and a "Blood Moon" lunar eclipse, which many cultures regard as a bad omen, on January 21. The first two events plausibly portended a more favorable environment for investment and the third a less favorable environment. Supporting these hypotheses, the relationships between a firm's event-day return and its investment rate are positive for the first two events and negative for the third, all statistically significant at the 5% level based on standard t tests. This evidence would be more persuasive if not for the fact that the relationship between one-day returns and investment rate is significant at the 5% level on 48.4% of *all* trading days in 2019, and larger in magnitude than on any of the three event days more than 40% of the time (Fig. 1).

This example is not unique. Across a range of firm characteristics, including several studied in published work, we find that relationships between one-day returns and these characteristics are statistically significant at even the *1% level* on an average of 33.4% of trading days between 1991 and 2021. The prevalence of statistically significant relationships on otherwise "ordinary" days implies that a statistically significant relationship on the day of a given event is too low a bar for concluding that the event itself caused differences in returns. Matching the frequency of statistically significant relationships on ordinary days to nominal significance levels would require increasing standard errors by a factor of up to 4. Statistically significant relationships occur

* Corresponding author.
E-mail addresses: jonathan.cohn@mccombs.utexas.edu (J.B. Cohn), travis.johnson@mccombs.utexas.edu (T.L. Johnson), zliu@bauer.uh.edu (Z. Liu), malcolm.wardlaw@uga.edu (M.I. Wardlaw).

[1] This methodology dates back to the early 1980s (e.g., Leftwich, 1981). See Eisfeldt et al. (2025), Jiang et al. (2025), and Lu and Ye (2025) for recent examples.

https://doi.org/10.1016/j.jfineco.2026.104278
Received 1 June 2025; Received in revised form 25 February 2026; Accepted 6 March 2026
Available online 17 March 2026

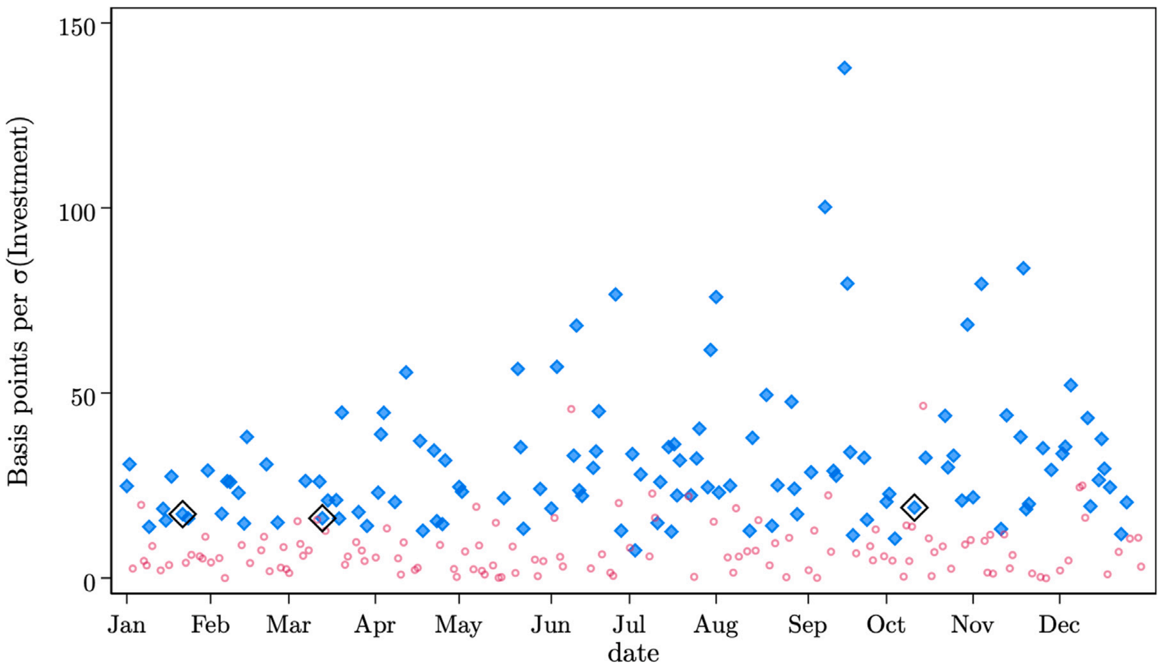

Fig. 1. Daily relationship between returns and investment in 2019.
This figure presents daily estimates of the cross-sectional relationship between firm-level returns and $Investment$, where $Investment$ is capital expenditures divided by total assets. For each trading day in 2019, we regress one-day returns on $Investment$ and compute the absolute value of the difference between the estimated daily coefficient and the average coefficient over the year. We then rescale this difference to basis points per standard deviation of $Investment$. A blue diamond indicates that the raw coefficient is statistically significant at the 5% level based on a t test using White (1980) robust standard errors, while a red circle indicates insignificance at the 5% level. The estimates for January 22 (the day after the Blood Moon, which occurred after trading hours), March 13, and October 11 are enclosed in larger diamonds for emphasis.

more frequently for defining characteristics such as size and book-to-market ratio and for five-day event return windows than for one-day windows — an important finding given the widespread use of multi-day event windows in the literature. Including controls, clustering standard errors by cross-sectional groups such as industries, and factor-adjusting returns all reduce statistical significance rates, though they still exceed 10% of days, on average, at the 1% significance level.

To shed further light on the nature of the confounding events problem, we build a simple conceptual framework in which both a focal event and "routine news" can affect the returns of firms that differ on a given characteristic differently on a given day. We show that the problem can be thought of as one of testing the wrong null hypothesis — that there is no relationship between event-day returns and a given characteristic. Because of differential exposure to routine news, there may be genuine return-characteristic relationships on the event day, even absent the focal event. We also show that clustering standard errors does not address the problem, though it may help with error correlations unrelated to the effects of confounding events. In contrast, an alternative approach that conducts statistical testing by comparing a relationship in the event period to the distribution of the same relationship in a time series of non-event periods does address the problem. Despite its intuitive appeal, we find that only 6% of papers conducting short-term event studies of returns around a quasi-experimental event use such a "time-series" approach.[2]

Further investigating this time-series approach empirically, we show that a t test comparing an event-day OLS regression coefficient to a series of non-event day OLS coefficients, as in the "portfolio OLS" approach of Sefcik and Thompson (1986), still produces statistical significance rates that are slightly too high due to fat tails in the distributions of relationships between returns and firm characteristics. In contrast, testing significance using a p value based on the empirical CDF corrects this problem, producing rates of significance at the 1% level on approximately 1% of days. A potentially greater concern with this OLS-based time-series approach is that it may have weak statistical power because it does not exploit information about return correlations that could allow more efficient estimates of both event- and non-event period relationships (Chandra and Balachandran, 1992). Power may be an especially important consideration in this setting since realistic effect sizes are often small relative to the overall noise in returns. We test for power by introducing relationships between short-term returns and a characteristic to create artificial event windows and determining how frequently this approach can detect these relationships. We find that detection rates are indeed low in many cases — for example, only 35.1% on average at the 5% significance level for a 50 bp increase in five-day returns per one-standard deviation change in a characteristic.

We next introduce a novel approach designed to address the power concern. Our approach continues to compare the event-period relationship to the distribution of non-event period relationships. However, it uses GLS rather than OLS to estimate relationships. Our key innovation is to borrow from the asset pricing literature (Giglio and Xiu, 2021; Lopez-Lira and Roussanov, 2023) and use principal component analysis (PCA) to encode the most important dimensions of return correlations based on recent past returns into a covariance matrix for each day, the inverse of which serves as the GLS weighting matrix for that day. We find that our "time-series GLS" approach typically detects artificial relationships two to three times as often as a "time-series OLS" approach, implying large gains in power. For example, it achieves detection rates of 61.2% on average at the 5% significance level for a 50 bp artificial increase in five-day returns per one-standard deviation change in a characteristic. To facilitate adoption of time-series approaches, we provide new Stata and Python modules that implement both OLS- and GLS-based procedures.[3]

[2] See Appendix A for a survey of 81 relevant papers published in 2014–2023.

[3] https://github.com/MalcolmWardlaw/csestudy.

To shed light on the economic drivers of the relationships between short-term returns and firm characteristics and the source of the GLS efficiency gains, we analyze the principal components (PCs) of returns. Each PC encodes an orthogonal factor explaining a portion of cross-sectional differences in realized returns. Like Lopez-Lira and Roussanov (2023), we find highly complex, multi-dimensional correlation patterns that are difficult to map to traditional factors related to expected returns. We additionally find that the nature of these correlations varies considerably over time and is often period-specific. For example, sensitivity to finance industry returns is important for explaining return differences in 2008, while sensitivity to meme stock returns is important in 2021. The complexity and instability of return correlations explain the limited effectiveness of both clustering and factor-adjusting returns in addressing the inference problem that confounding events cause.

Our paper contributes to the literature on event study analysis of returns. The literature has primarily explored the challenges that return correlations due to overlapping exposures create for inference in studies of abnormal *mean* event returns (Collins and Dent, 1984; Bernard, 1987; Lyon et al., 1999; Brav et al., 2000; Mitchell and Stafford, 2000; Jegadeesh and Karceski, 2009). Kolari and Pynnönen (2010) estimates Type I and II error rates for mean returns based on different approaches and recommends correcting standard errors for mean event-window returns using average cross-firm correlations. An approach like this is ineffective in our setting because it does not account for the dependence of return correlations on specific firm characteristics. We are not aware of any prior analysis of the practical challenge caused by return correlations for *cross-sectional* analysis of returns around an event – a more common methodology in modern corporate finance.[4] Sefcik and Thompson (1986) describes this challenge conceptually but does not assess its importance in real-world applications, which we find to be significant. The approach they propose using pre-event relationships to conduct statistical testing has largely been ignored. Our novel GLS variant of this approach offers substantial improvements in statistical power in real-world applications.

Our paper also contributes to the general literature on practical challenges in statistical inference based on regression analysis. Much of this work focuses on addressing correlation in regression errors – often via clustered standard errors – which, if ignored, can lead to misleading conclusions about statistical significance (Moulton, 1986, 1987; Bertrand et al., 2004; Petersen, 2009; Thompson, 2011; Cameron and Miller, 2015; Abadie et al., 2023). In the event-study setting we examine, however, over-rejection can occur even when regression errors are uncorrelated, because confounding events generate *genuine* cross-sectional return–characteristic relationships unrelated to the focal event. Standard clustering adjustments do not address this problem, since it arises from common shocks that shift the entire cross-sectional slope in a given period rather than from residual correlation conditional on the regressors. The time-series approaches we describe are conceptually similar to the Fama and MacBeth (1973) two-step procedure in that they base inference on the time-series variation in cross-sectional slope estimates to account for such shocks. The key difference is that our methods compare event-period relationships specifically to the distribution of non-event-period relationships, following the spirit of Sefcik and Thompson (1986), rather than averaging across all periods, and we introduce a GLS variant to improve efficiency.

2. Testing differences in returns around a quasi-experimental event

In this section, we present a framework for understanding the consequences of confounding events when testing hypotheses about the differential effect of a quasi-experimental event on the valuations of firms with different characteristics. We then explain why statistical tests based on standard errors from cross-sectional event-period return regressions are likely to produce frequent false positives. Finally, we describe different approaches that use the distribution of relationships in non-event periods to test for event-period significance, including a novel GLS-based approach, and explain why they sidestep the confounding events problem.

2.1. Conceptual framework

We begin by presenting a simple framework that models a discrete quasi-experimental event, where both the event itself and stochastic "routine news" can cause a dependence of event-period returns on a firm characteristic. We focus on a one-day event period and a single characteristic here for simplicity but extend the framework to multi-day events and multiple firm characteristics in Appendix B.

Let $r_{i,t}$ denote firm i's stock return on day t. Suppose that a researcher wishes to study differences in the effect of a quasi-experimental event occurring on day $t = \tau$, the event day, on firm value, as measured by stock return $r_{i,\tau}$, across firms that differ on characteristic $x_{i,\tau-1}$. In addition, suppose that firm i's return on day t depends on routine news events, n_t, which affects the returns of firms with different values of $x_{i,t-1}$ differently. n_t represents a composite of all background news events on day t with value implications that depend on x. It is observable to market participants on day t, but a researcher cannot measure it except through its effect on returns. It is drawn independently on each day t from a distribution with mean 0 and variance σ_n^2. Routine news arrives on the event day, $t = \tau$, just like it does on other days. So, on the event day, there are potentially two pieces of news that affect the returns of firms with different values of $x_{i,\tau-1}$ differently – the event that the researcher wishes to study and routine news.

Formally, the data generating process is:

$$r_{i,t} = m_t + \left(n_t + e\mathbb{1}(t=\tau)\right)(x_{i,t-1} - \bar{x}_{t-1}) + \delta_{i,t}, \tag{1}$$

where e is the differential effect of the quasi-experimental event that the researcher is studying and is the object of interest, $\mathbb{1}()$ is the indicator function, $\bar{x}_{t-1}$ is the cross-sectional mean of $x_{i,t-1}$, m_t is a mean-$\bar{m}$ random variable capturing the mean effect of news on day t on all firms and is independent of n_t, and $\delta_{i,t}$ is a mean-zero i.i.d. error term. We specify the demeaned value of x_{t-1} to ensure that the interaction term captures only cross-sectional differences in return responses and not any aggregate response.[5]

2.2. Testing for differences using cross-sectional regression standard errors

Because the researcher cannot observe routine news, and the returns on day τ reflect the effects of both routine news and the quasi-experimental event, the researcher cannot separately identify n_τ and e. The best the researcher can do is to estimate $n_\tau + e$. We rewrite Eq. (1) for day $t = \tau$ in its estimable form as:

$$r_{i,\tau} = a_\tau + b_\tau x_{i,\tau-1} + \delta_{i,\tau}, \tag{2}$$

where $a_\tau = m_\tau - (n_\tau + e)\bar{x}_{\tau-1}$ and $b_\tau = n_\tau + e$. This is the simplest version of the standard regression equation that researchers typically estimate to test for differences in short-term returns around a quasi-experimental event.

The possibility of routine news does not bias the event-period return relationship as an estimate of the differential event effect because routine news causes no differences in average returns as a function

[4] See Kothari and Warner (2007) for a survey of the literature on the econometrics of event study analysis.

[5] The general framework in Appendix B permits correlations among $\delta_{i,t}$. It also allows $x_{i,t-1}$ to affect expected returns $\mathbb{E}[r_{i,t}]$ outside of the event window, which we rule out here by assuming $\mathbb{E}(n_t) = 0$.

of x. Formally, let $\hat{b}_\tau$ denote a regression estimate of b_τ. Since $\delta_{i,\tau}$ is independent of $x_{i,\tau-1}$ and n_τ is mean-zero, $\hat{b}_\tau$ is an unbiased estimate of e.

While the standard error of $\hat{b}_\tau$ is valid for testing the null hypothesis that $b_\tau = 0$, routine news renders it invalid for testing the null hypothesis that $e = 0$. To see why, consider the effect of n_τ on the standard error of $\hat{b}_\tau$ as an estimate of e:

$$Var(\hat{b}_\tau - e) = Var(\hat{b}_\tau - b_\tau + b_\tau - e) = Var(\hat{b}_\tau - b_\tau + n_\tau) = Var(\hat{b}_\tau - b_\tau) + \sigma_n^2, \quad (3)$$

where the last step follows from the fact that the sampling error $\hat{b}_\tau - b_\tau$ is independent of the news n_τ. This is the key result from the analysis in this section.

Result 1. *The variance of estimation errors for b_τ as an estimate of e is $Var(\hat{b}_\tau - e) = Var(\hat{b}_\tau - b_\tau) + \sigma_n^2$.*

Note that the sampling error $Var(\hat{b}_\tau - b_\tau)$ would be present even absent routine news that affects firms with different values of $x_{i,\tau-1}$ differently. While sampling error tends towards 0 as the sample size grows, a larger sample does not reduce the additional variance from the impact of routine news, which drives a wedge between the true b_τ and e that would remain even in an infinite sample. Since the standard error of the estimate $\hat{b}_\tau$, $\sqrt{Var(\hat{b}_\tau - b_\tau)}$, overstates the precision of $\hat{b}_\tau$ as an estimate of e, a statistical test based on this standard error too frequently rejects the null hypothesis that the event has no differential value effects across firms differing on characteristic x. The severity of this problem scales with σ_n^2, which one can think of as the average volume of routine news arriving daily that affects the valuations of firms differing on x differently.

Intuitively, using the standard error of $\hat{b}_\tau$ obtained from estimating the regression Eq. (2) tests the wrong null hypothesis. This standard error would be appropriate for testing the null that there is no relationship between returns and characteristic x on day τ – i.e., that $b_\tau = 0$. However, the researcher's objective when estimating Eq. (2) is to test the null that there is no relationship between event-day returns associated with the focal event and x – i.e., that $e = 0$. These are different nulls since there may well be a genuine relationship between returns and x on the event day due to routine news, even if the focal event itself does not cause any differences in returns.

Another useful interpretation of this result is that routine news causes returns to have a rich cross-sectional correlation structure. To see this, rewrite Eq. (1) for $t \neq \tau$ as:

$$r_{i,t} = m_t + v_{i,t}, \quad (4)$$

where $v_{i,t} = n_t(x_{i,t-1} - \bar{x}_{t-1}) + \delta_{i,t}$, and observe that:

$$Cov(v_{i,t}, v_{j,t}) = E[v_{i,t}v_{j,t}] = (x_{i,t-1} - \bar{x}_{t-1})(x_{j,t-1} - \bar{x}_{t-1})\sigma_n^2, \quad (5)$$

which follows from the fact that $\mathbb{E}[\delta_{i,t}] = \mathbb{E}[\delta_{j,t}] = \mathbb{E}[\delta_{i,t}\delta_{j,t}] = 0$. Thus, the returns of firms with high values of x will tend to comove positively with each other, as will the returns of firms with low values of x. We can therefore use the covariance matrix of returns on days $t \neq \tau$, demeaned within each time period, to measure how routine news differentially affects firms.

2.3. Clustering standard errors

The correlation structure in Eq. (5) might suggest that clustering standard errors by groups such as industry or geographic location would help address the confounding events inference problem. However, while event-day routine news n_τ induces cross-sectional correlations in returns, it does not cause correlations in the return *errors* in regression Eq. (2). Rather, the correlations stem from variation in b_τ due to the unobservable realization of n_τ. While clustering may still be useful in correcting standard errors if the errors $\delta_{i,\tau}$ are correlated, doing so does not help address the confounding events problem at all since the residuals $\hat{\delta}_{i,\tau}$ contain no information about n_τ.

2.4. Using pre-event relationships to test significance

We next consider an alternative methodology to test for a characteristic-return relationship on the event day that involves comparing this relationship to the time-series distribution of the same relationship estimated on non-event days. We focus specifically on pre-event days. In principle, one could also use post-event days as well as or instead of pre-event days. However, post-announcement drift or follow-on events might muddy interpretation. Comparing an event-period relationship to a distribution of pre-event relationships effectively uses the pre-event distribution to bootstrap the distribution of the event-period relationship under the null hypothesis that the event itself has no differential effect. More precisely, this testing strategy uses the distribution of pre-event relationships to approximate the distribution of $\hat{b}$ under the null that $e = 0$. This approach sidesteps the problem caused by routine news, as well as any cross-correlations in $\delta_{i,t}$, since routine news also causes volatility in the benchmark pre-event relationship.

The procedure involves three steps.[6] The first two are:

Step 1: Estimate single-day cross-sectional regressions:

$$r_{i,t} = a_t + b_t x_{i,t-1} + \epsilon_{i,t}, \quad (6)$$

for the event day, $t = \tau$, and for a series of L pre-event days, $t = \tau_{-1}, \tau_{-2}, \ldots, \tau_{-L}$.[7]

Step 2: Estimate the event-specific effect as:

$$\hat{e} = \hat{b}_\tau - \mu(\hat{b}), \quad (7)$$

where $\hat{b}_{\tau_{-\ell}}$ is the estimated coefficient for pre-event day $\tau_{-\ell}$ and $\mu(\hat{b}) = \frac{1}{L}\sum_{\ell=1}^{L}\hat{b}_{\tau_{-\ell}}$ is the mean coefficient on pre-event days.[8]

The third and final step involves testing the statistical significance of $\hat{e}$. We describe two specific approaches to this step. Before doing so, we derive some useful properties of $\hat{e}$ (see Appendix B for the proof):

Result 2. *Assuming stationary distributions for sampling error $\hat{b}_{\tau_{-\ell}} - b_{\tau_{-\ell}}$ and news n_t, $\hat{e}$ has the following relations to the true event-specific effect e:*

$$\mathbb{E}(\hat{e} - e) = \underbrace{0}_{\textit{No bias}}, \quad (8)$$

$$Var(\hat{e} - e) = \underbrace{Var(\hat{b}_{\tau_{-\ell}})}_{\textit{Time-series sample variance}}, \quad (9)$$

$$\Phi(\hat{e} - e) = \underbrace{\Phi(\hat{b}_{\tau_{-\ell}})}_{\textit{Time-series sample CDF}}, \quad (10)$$

where $\Phi(x)$ is the cumulative distribution function evaluated at x.

First, $\hat{e}$ is an unbiased estimator of e, which is the object of interest. Second, the distribution of sampling error in $\hat{e}$ is simply the time-series distribution of the coefficients from the pre-event day regressions, which is useful for statistical testing. The first specific approach to testing the statistical significance of $\hat{e}$ leverages Eq. (9) and computes a t-statistic, using the standard deviation of the pre-event window coefficients as the standard error.

[6] We show how to extend this approach to a multi-day event period in Appendix B.

[7] A natural choice of pre-event days is the L days immediately preceding the event, $t = \tau - L, \tau - (L-1), \ldots, \tau - 1$. However, one could build in a lag $d > 0$ and use days $t = \tau - L - d, t = \tau - (L-1) - d, \ldots \tau - 1 - d$ if the risk of information leakage prior to the event is high.

[8] Subtracting $\mu(\hat{b})$ is only necessary when the characteristic-return relationship is non-zero unconditionally. In practice, subtracting $\mu(\hat{b})$ has little effect on the results that follow because the mean relationship between daily returns and characteristics is generally close to zero.

Step 3(t): Compute the t-statistic

$$\hat{t} \equiv \frac{\hat{e}}{\sqrt{\mathrm{Var}(\hat{e}-e)}} = \frac{\hat{b}_\tau - \mu(\hat{b})}{\sqrt{\mathrm{Var}(\hat{b}_{\tau_{-\ell}})}}, \tag{11}$$

and compare $\hat{t}$ to critical values from the standard normal, Student's T, or other cumulative distribution function.[9]

The second approach to testing the statistical significance of $\hat{e}$ is to compute a p-value based on the empirical cumulative distribution function (CDF) of the pre-event day coefficients - i.e., the empirical, model-free counterpart to Eq. (10).

Step 3(CDF): Compute the p-value

$$\hat{p} \equiv \frac{1}{L+1}\left[\sum_{\ell=1}^{L} \mathbb{1}\left(\left|\hat{b}_{\tau_{-\ell}} - \mu(\hat{b})\right| > \left|\hat{b}_\tau - \mu(\hat{b})\right|\right)\right], \tag{12}$$

and compare it to the significance threshold percentile p^*.

Note that this is a two-tailed test since it compares the absolute value of the (de-meaned) event-day coefficient to the distribution of absolute values on pre-event days. The advantage of this approach over the first is that it imposes no distributional assumptions on the time series of pre-event relationships that are used as a benchmark. As we show later, the distributions of relationships between short-term returns and firm characteristics tend to exhibit fat tails, which can cause excess false positives when using a t statistic for testing.[10]

The statistical significance threshold level p^* imposes a mild technical constraint on the number of pre-event periods L that the researcher can use with the CDF approach. Specifically, it requires that $L+1$ be divisible by $1/p^*$. Otherwise, no pre-event day absolute coefficient deviation falls exactly at the p^*th percentile of the distribution.[11] Because the critical values researchers care about in practice often include 1%, it is natural to choose L to be one less than a multiple of 100.

We consider two specific implementations of this overall approach – an OLS-based approach, which is akin to the Sefcik and Thompson (1986) portfolio OLS approach, and a novel GLS-based approach. The only difference between the two approaches is in Step 1 of the procedure – estimation of cross-sectional regressions for event and pre-event days. The OLS-based approach estimates the regressions in Step 1 using OLS. We refer to this approach as **time-series OLS** (TS-OLS for short).[12] The GLS approach, in contrast, estimates both the event-day and pre-event day regressions in Step 1 using GLS instead. We refer to the GLS-based approach as **time-series GLS** (TS-GLS for short).

2.4.1. Time-series GLS

The GLS approach requires estimating a covariance weighting matrix of returns on each day. The potential advantage of this approach relative to time-series OLS is that it increases the efficiency of both event-day and non-event day estimates, which should increase statistical power, making differential effects of an event easier to detect. Our approach borrows from the asset pricing literature (Giglio and Xiu, 2021; Lopez-Lira and Roussanov, 2023) and uses PCA to capture the most important information about return comovement in the estimated covariance matrix. It demeans returns on each day prior to conducting PCA so the resulting covariance estimates better represent the covariance of residuals in cross-sectional regressions that include a period-specific intercept. Our approach uses PCA to construct the covariance matrix rather than specifying factors *ex ante* because the objective is to estimate covariances as accurately as possible and not to explain the cross section of returns using factors with simple economic interpretations.

Writing $f_{k,t}$ as the realization of the kth PC of returns on day t, our approach specifies the covariance matrix of returns by assuming an arbitrage pricing theory (APT) structure:

$$r_{i,t} = \phi_i + \sum_{k=1}^{K} \lambda_{i,k} f_{k,t} + \epsilon_{i,t}, \tag{13}$$

$$\mathrm{Cov}(\epsilon_{i,t}, \epsilon_{j,t}) = \begin{cases} \sigma_i^2 & \text{when } i = j \\ 0 & \text{when } i \neq j. \end{cases} \tag{14}$$

Given this structure, the elements of the estimated covariance matrix of returns are:

$$\hat{\Omega}_{r,ij} \equiv \mathrm{Cov}(r_{i,t}, r_{j,t}) = \sum_{k=1}^{K} \lambda_{i,k}\lambda_{j,k}\mathrm{Var}(f_{k,t}) + \mathbb{1}(i=j)\sigma_i^2. \tag{15}$$

Note that our approach compares event-day and pre-event day GLS coefficients and does not simply estimate a single cross-sectional event-day regression using GLS. While a purely cross-sectional GLS event-day regression may have less sampling error and therefore be more efficient than a cross-sectional OLS event-day regression, Section 2.2 shows that cross-sectional standard errors underestimate the true estimation error for e even without sampling variation. Furthermore, implementations of purely cross-sectional GLS that are feasible in practice have estimation error in their covariance matrices. While GLS should make our time-series approach more efficient than an OLS-based time-series approach, it is the reliance on the time series of pre-event day relationships rather than the use of GLS that allows our approach to estimate a valid variance of the return-characteristic relationship under the null that $e = 0$.

2.4.2. Time-varying news arrival

One potential concern with both the TS-OLS and TS-GLS approaches is that the distribution of routine news may be time-varying and, in particular, may differ between the event period and pre-event periods. If the distribution in the event period is more variable than in pre-event periods, then both TS-OLS and TS-GLS could produce excess false positives.[13] This difference in variability could occur because the focal event is part of a set of connected events that affect different firms in ways similar to the focal event. One example is that of an unexpected change in corporate tax rates that is part of a broader package of tax reforms announced all at once, which would make it conceptually difficult to determine whether differences in event-period returns are attributable to the change in corporate tax rates or another part of the overall tax reform. Another example is that of a change to financial regulations that arrives during a time of broader market

[9] This approach is identical to estimating a second-stage regression of the time-series of coefficients $\hat{b}_{\tau_{-L}}, \ldots, \hat{b}_{\tau_{-1}}, \hat{b}_\tau$ on a constant and an indicator equal to 1 on day τ and 0 otherwise. The coefficient on the indicator equals $\hat{b}_\tau - \mu(\hat{b})$, while the standard error equals $\sqrt{\mathrm{Var}(\hat{b})}$ when a small-sample correction is applied to the variance. $\hat{t}_j$ is therefore the standard t-statistic for this second-stage regression. Note the similarity of this approach to the Fama and MacBeth (1973) methodology in asset pricing.

[10] While the p-value approach does not produce standard errors directly, standard errors can easily be inferred with the addition of a distributional assumption.

[11] Suppose $L = 252$. In this case, $\hat{p} \leq 1\%$ only when $\hat{b}_\tau$ is in the top three of the 253 estimated coefficients. This implies $\mathbb{P}(\hat{p} \leq 1\%) = 3/253 > 1\%$, and thus an event-day observation that is a random draw from the empirical pre-event distribution will be significant at the 1% level slightly more than 1% of the time.

[12] Sefcik and Thompson (1986) shows that this implementation is tantamount to forming portfolios with weights determined by the distribution of the explanatory characteristics and then comparing event and pre-event portfolio returns. Their approach maps into the t-statistic approach to implementing TS-OLS but is not amenable to computing a p value based on the empirical CDF of the coefficient distribution.

[13] Time variation in the distribution of n_t could also add noise to the estimated covariance matrix used in the TS-GLS approach, which would weaken the power of this approach.

turmoil, which would make it difficult to distinguish the impact of the specific regulation from contemporaneous shocks to the economy.

Although no calculation can fully address a potential shift in the distribution of routine news, we suggest some potential diagnostics for researchers concerned about this possibility. One is to examine the volatility of cross-sectional regression coefficients, $\hat{b}_t$, in rolling one-month windows. A large change in this volatility just prior to the event window may suggest a structural break not related to the event in question. More broadly, researchers should investigate the context in which an event occurs. If there are other atypical concurrent events or market conditions during the event period but not the pre-period, then the event is probably not a good candidate to use as a quasi-natural experiment.

2.5. Panel approaches

Finally, one might imagine employing approaches that treat stacked firm-day data for the event day and a series of pre-event days as a single panel. For example, a researcher could estimate:

$$r_{i,t} = a + b_1 x_{i,t-1} + b_2 x_{i,t-1} \mathbb{1}(t = \tau) + \epsilon_{i,t}, \tag{16}$$

where b_1 captures the average characteristic-return relationship in the panel and b_2 captures the differential event relationship. While $\hat{b}_2$ is an unbiased estimator of e, it captures both the event effect e and the realization of event-day routine news n_τ. As a result, like the cross-sectional regression in Eq. (2), its standard error is too small.

Another possible panel regression approach would be to use day fixed effects and allow the relation between $x_{i,t-1}$ and $r_{i,t}$ to vary across days, as in:

$$r_{i,t} = a_t + \sum_{\ell=1}^{L} b_\ell x_{i,t-1} \mathbb{1}(t = \tau_{-\ell}) + b_\tau x_{i,t-1} \mathbb{1}(t = \tau) + \epsilon_{i,t}. \tag{17}$$

However, this regression is algebraically equivalent to the cross-sectional OLS approach, as neither the coefficient estimate $\hat{b}_\tau$ nor the corresponding standard error uses any information from prior days.

3. Data and sample

Our analysis involves regressing firm-level stock returns over periods of one and five days on various firm characteristics. Our sample period is 1991–2021, which contains 7,811 trading days. This period is long enough to allow us to estimate the distributions of relationships between returns and characteristics accurately and to explore variation in these distributions over time. Our analysis uses daily firm-level stock return data from CRSP and annual financial statement data from Compustat. Applying standard data filters, we only include stocks traded on the NYSE, AMEX, and NASDAQ with share codes 10 or 11 and a share price greater than 3, and exclude financial and utility firms with a one-digit SIC code of 4 or 6.

We analyze eight firm characteristics. The first four are some of the most commonly studied characteristics in corporate finance and represent important dimensions of a firm's fundamental nature. These are $Size$, computed daily as the natural logarithm of market equity, which is the product of daily closing stock price and number of shares outstanding from CRSP; M/B (market-to-book ratio), computed monthly as the natural logarithm of the ratio of the market value of equity at the end of the prior month to book value (Compustat CEQ) measured at the end of the prior fiscal year (we exclude the small number of observations where CEQ is zero or negative); $Profit$, computed annually as gross profit (Compustat GP) divided by total assets (Compustat AT); and $Invest$, computed annually as capital expenditures (Compustat $CAPX$) divided by total assets. We winsorize each characteristic at the 1st and 99th percentiles to deal with possible outliers.

The other four characteristics we analyze are from recently published papers that focus on analyzing differences in stock returns around quasi-experimental events. We choose these specific characteristics because they are simple to replicate, ensuring that we should be able to match the variable values in the published papers. The characteristics, all measured annually, are $Cash/AT$, which is cash and short-term investments (Compustat CHE) divided by total assets; $Debt/AT$, which is the sum of long-term debt (Compustat $DLTT$) and debt in current liabilities (Compustat DLC), divided by total assets; $TaxRate$, which is 100 times income taxes paid (Compustat $TXDP$), divided by the difference between pre-tax income (Compustat PI) and special items (Compustat SPI), set to 0 if $PI < 0$; and $NYHQ$, which is an indicator variable equal to 1 if a firm is headquartered in New York (Compustat $STATE$ equal to "NY") and 0 otherwise.[14]

The primary unit of observation is a firm-day, though we also analyze returns over periods of five days. We associate with each observation the value of each characteristic as of the most recent available date prior to the date of the observation. The resulting sample consists of 21,071,829 firm-days belonging to 11,556 unique firms, as summarized in Table 1.

4. Cross-sectional approaches

As we formalize in Section 2, the possibility of other news differentially affecting firms with different characteristics can generate excess false positives when testing for return-characteristic relationships around a specific event. However, the practical importance of this problem is unknown. In this section, we assess the practical risk of false positives when using cross-sectional standard errors to test significance in real-world applications.

4.1. Frequency of statistically significant relationships

To quantify the frequency of false positives, we estimate long sequences of short-term return regressions. Specifically, for each of the eight characteristics we describe in Section 3 and each of the 7,811 trading days in our 1991–2021 sample period, we estimate OLS regressions of returns on the characteristic. We then compute the fraction of regressions in which the coefficient on the characteristic is statistically significant at the 1% and 5% levels based on its regression standard error.[15] The fraction of these coefficients that are statistically significant represents an estimate of the probability of a false positive – the more frequently the coefficient is statistically significant, the more likely the coefficient on the day of a specific event would appear statistically different from zero even absent the event. If this fraction is larger than the nominal significance level of the test, then the test is likely to produce excess false positives.

In practice, researchers sometimes compute White (1980) robust standard errors to account for heteroskedasticity or cluster standard errors to account for correlated regression errors when testing differences in event returns. The most common clustering dimension in event studies of returns is industry (see Appendix A). We therefore compute statistical significance rates using default standard errors, White (1980) robust standard errors, and standard errors clustered at the Fama–French 49-category industry level. In addition, we compute these rates based on industry-clustered standard errors where the regression includes Fama–French 49-category industry fixed effects, since the combination of industry fixed effects and industry clustering is sometimes used in practice. Fig. 2 presents the results from this analysis.

[14] Fahlenbrach, Rageth, and Stulz (2021) studies how $Cash/AT$ and $Debt/AT$ relate to returns around the arrival of the COVID-19 pandemic to test the effects of financial structure on resilience; Wagner, Zeckhauser, and Ziegler (2018) studies how $TaxRate$ relates to returns around the 2016 U.S. Presidential election to measure how much lower corporate tax rates benefit high-tax firms; and Acemoglu, Johnson, Kermani, Kwak, and Mitton (2016) studies how $NYHQ$ relates to returns around the nomination of Timothy Geithner as Treasury Secretary to test whether better-connected firms benefited from this appointment.

[15] We focus on 2-sided tests throughout the paper since these are the convention in finance research.

Table 1
Summary statistics.

Variable	Firm-Days	Firms	Firms/Day	Mean	Median	σ	Within-day σ
Return (%)	21,071,829	11,556	2698	0.12	0.00	4.00	3.82
Size	21,064,678	11,631	2697	6.14	6.00	1.96	1.81
M/B	20,487,758	11,487	2623	1.01	0.92	0.98	0.96
Profit.	21,084,670	11,650	2699	0.36	0.34	0.32	0.32
Invest.	20,911,833	11,613	2677	0.06	0.04	0.07	0.07
CASH/AT	21,084,729	11,651	2699	0.21	0.12	0.24	0.24
Debt/AT	21,010,836	11,645	2690	0.21	0.16	0.22	0.22
Tax Rate	14,393,605	8110	1843	23.96	23.64	16.65	16.36
NY HQ	20,170,658	11,000	2582	0.07	0.00	0.25	0.25

This table presents summary statistics for the sample of firm-day observations used in our analysis. $Return$ is the one-day stock return. $Size$ is the natural log of market equity, which is the product of the closing stock price on the previous day and the number of shares outstanding from CRSP. M/B is the log of market value of equity at the end of the prior month to book value (Compustat CEQ) measured at the end of the prior fiscal year (observations for which book value is zero or negative are excluded). $Profit.$ is gross profit (Compustat GP) divided by total assets (Compustat AT). $Invest.$ is capital expenditures (Compustat $CAPX$) divided by total assets. $Cash/AT$ is cash and short-term investments (Compustat CHE) divided by total assets. $Debt/AT$ is the sum of long-term debt (Compustat $DLTT$) and debt in current liabilities (Compustat DLC), divided by total assets. $Tax\ Rate$ is 100 times income taxes paid (Compustat $TXDP$), divided by the difference between pre-tax income (Compustat PI) and special items (Compustat SPI), set to 0 if $PI - SPI < 0$. $NY\ HQ$ is an indicator variable equal to 1 if a firm is headquartered in New York (Compustat $STATE$ equal to "NY") and 0 otherwise. Our sample is the 7,811 trading days in 1991 through 2021.

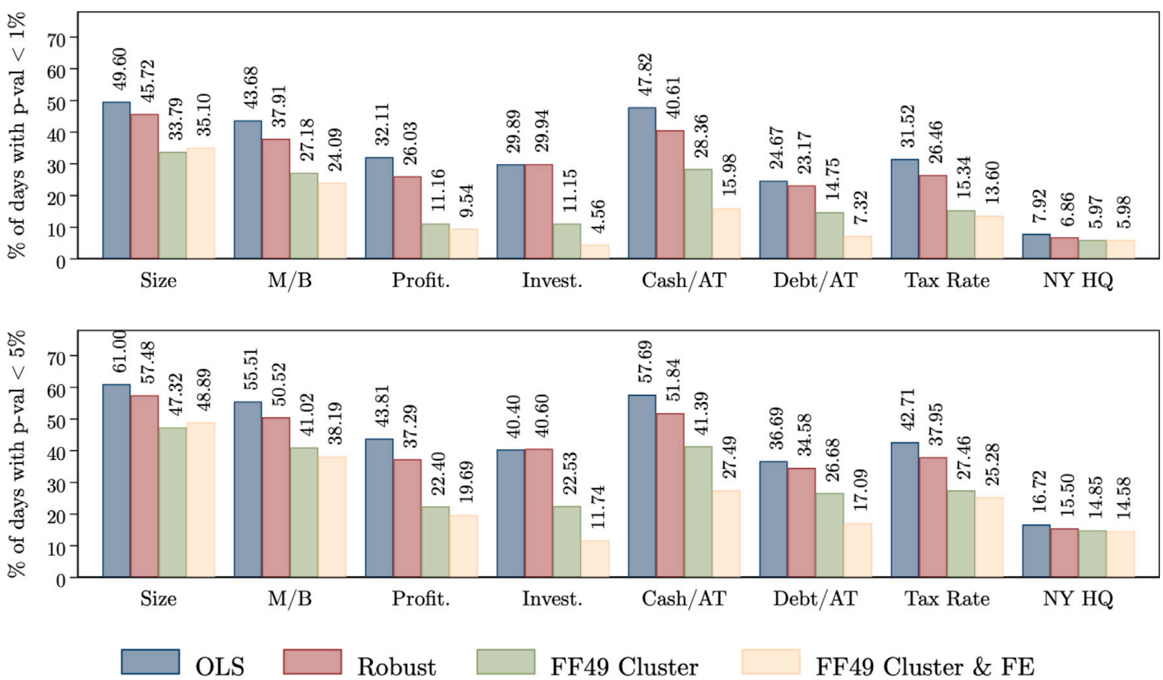

Fig. 2. Significance rates based on cross-sectional standard errors.
This figure depicts statistical significance rates from univariate cross-sectional OLS regressions of 1-day returns on each of eight characteristics. The bars show the percentage of days for which the coefficient from the regression is statistically significant based on default standard errors, White (1980) robust standard errors, Fama–French 49-category industry clustered standard errors, and industry clustered standard errors where we also include industry fixed effects. The top panel shows rates of significance at the 1% significance level, while the bottom panel shows rates of significance at the 5% significance level. Our sample is the 7,811 trading days in 1991 through 2021.

Statistical significance rates at the 1% and 5% significance levels based on default standard errors average 33.4% and 44.3%, respectively. These rates are greatest for $Size$, M/B, and $Cash/AT$, suggesting that these attributes capture important variation in exposure to day-to-day news. This pattern arises because day-to-day news is more likely to cause substantial valuation effects for firms of similar size or expected growth rates than for those headquartered in the same state, for example.[16]

Significance rates based on White (1980) robust standard errors are generally smaller than those based on default standard errors, and those based on industry-clustered standard errors smaller still. However, even significance rates based on industry-clustered standard errors average 18.5% and 30.5% at the 1% and 5% significance levels, respectively. Note that the lower frequency of statistically significant relationships when clustering standard errors does not *perforce* represent an "improvement", as clustering may also reduce the power of a test, making it difficult to detect a true effect (Abadie et al., 2023). Including industry fixed effects in addition to clustering at the industry level reduces significance rates further in most but not all cases.

The stark takeaway from Fig. 2 is that statistically significant relationships between returns on the day of a specific event and a

[16] In the framework of Section 2, $Size$, M/B, and $Cash/AT$ have higher σ_n^2 than $NYHQ$.

Table 2
Significance rates based on cross-sectional standard errors.

Panel A: Different Standard Errors

		1% Significance Level				5% Significance Level			
		Univariate		Multivariate		Univariate		Multivariate	
Return	Specification	1d	5d	1d	5d	1d	5d	1d	5d
Unadj.	Default SE	33.4%	40.8%	26.1%	33.5%	44.3%	51.3%	37.4%	44.4%
Unadj.	Robust SE	29.6%	37.1%	22.5%	29.7%	40.7%	47.9%	33.8%	41.0%
Unadj.	FF49 Cluster	18.5%	22.9%	16.3%	21.6%	30.5%	35.2%	28.0%	33.7%
Unadj.	FF49 Cl & FE	14.5%	19.6%	13.2%	20.0%	25.4%	30.9%	23.7%	31.6%

Panel B: Alternative Returns

		1% Significance Level				5% Significance Level			
		Univariate		Multivariate		Univariate		Multivariate	
Return	Specification	1d	5d	1d	5d	1d	5d	1d	5d
CAPM	FF49 Cl & FE	17.7%	21.4%	12.2%	17.1%	29.8%	33.7%	22.4%	28.1%
FF3	FF49 Cl & FE	14.0%	16.6%	10.1%	14.7%	25.2%	28.3%	20.0%	25.4%
FF4	FF49 Cl & FE	13.2%	16.0%	9.5%	13.9%	24.2%	27.5%	19.2%	24.5%
LogRet	FF49 Cl & FE	18.5%	22.9%	13.2%	18.1%	30.5%	35.0%	23.7%	29.5%

This table presents statistical significance rates for coefficients from cross-sectional OLS regressions of short-term stock returns on each of eight firm characteristics based on cross-sectional standard errors. Each cell reports the percentage of all one- or five-day periods in which the regression coefficient is statistically significant, averaged across the eight characteristics. Each row reports the percentages for different combinations of 1% and 5% significance, univariate and multivariate regressions (where all eight characteristics are included as explanatory variables), and one-day and five-day return periods. Panel A presents rates of significance based on default standard errors, White (1980) robust standard errors, and standard errors clustered by Fama–French 49-category industry, with the fourth row based on regressions that also include Fama–French 49-category industry fixed effects. Panel B presents rates of significance for alternative return measures, including CAPM, Fama–French 3-factor, and Fama–French 4-factor adjusted returns and logged returns ($log(1 + return)$). Our sample is the 7,811 trading days in 1991 through 2021.

firm characteristic are likely to be common, even if the event itself does not differentially affect the returns of firms that differ on that characteristic. Thus, tests of the significance of these relationships are likely to produce frequent false positives.

4.2. Alternative specifications

In the baseline analysis presented in Fig. 2, we estimate separate univariate regressions for each characteristic without controls, using daily raw returns. In practice, researchers often include control variables and/or analyze multiple characteristics of interest. We therefore extend our analysis to the cases of multivariate regressions. In addition, researchers sometimes measure returns over event periods longer than one day, because either the event spans multiple days or it is unclear when the market learns about the event and incorporates its implications into prices. We therefore also extend our analysis to returns measured over five days instead of one to understand the implications of using longer event periods. Finally, we extend our analysis to other measures of returns that are sometimes used in practice, including factor-adjusted and log returns.

Table 2 reports average statistical significance rates across the eight characteristics we study for all four combinations of univariate and multivariate regressions and one-day and five-day event periods. In the multivariate regressions, the other seven characteristics serve as control variables for the characteristic of interest. Panel A reports results using raw returns based on the different ways of computing standard errors shown in Fig. 2, while Panel B reports results where we use alternative return measures and compute standard errors clustered at the Fama–French 49-industry level, with industry fixed effects included.

Statistical significance rates are even higher for five-day event periods than for one-day event periods. While we assume that n_t is mean-zero and i.i.d. in our conceptual framework in Section 2, in practice there is a small amount of autocorrelation in return-characteristic relationships.[17] Thus, aggregate routine news over a five-day period will have greater variance, σ_n^2, resulting in more statistically significant relationships with five-day returns than with one-day returns.

Significance rates are generally lower for multivariate regressions than for univariate regressions. Intuitively, control variables account for common exposure to day-to-day news that overlaps to some degree with common exposure based on the characteristic of interest. This result suggests an unappreciated benefit of including control variables when analyzing differences in an outcome variable around a quasi-experimental event: doing so can reduce the risk of false positives due to contemporaneous events. Note that this rationale for including control variables is different than the standard rationales of either reducing bias by absorbing variation in regression errors that might otherwise be correlated with the characteristic of interest, or increasing efficiency by absorbing variation in the outcome variable unrelated to the characteristic of interest.

Panel B of Table 2 shows that factor-adjusting returns decreases significance rates. Intuitively, when factor loadings are correlated with firm characteristics, factor-adjusted returns will be less correlated with firm characteristics than unadjusted returns. However, the changes in statistical significance frequencies are small. As we show in Section 6, the factor structure of returns is complex and therefore difficult to summarize with a small number of pre-specified factors.

Ultimately, Table 2 shows that average significance rates remain stubbornly high, regardless of how returns are measured, how standard errors are computed, or whether control variables are included. Table 3 reports results for additional clustering approaches and shows that significance rates remain elevated for alternative clustering choices as well. The prevalence of statistically significant relationships would make it hard to conclude with confidence from a statistically significant relationship in the period around a specific event that the event causes differences in returns.

Another way to quantify the inference threat posed by false positives due to routine news in this setting is to determine by how much standard errors would need to be inflated to match the probability of statistical significance to its nominal level. We refrain from tabulating the required rates of inflation since the information is largely redundant with that in Table 2, with higher significance rates in that table translating into greater need to inflate standard errors. However,

[17] This autocorrelation manifests in a small positive autocorrelation in returns of characteristic-sorted portfolios, a phenomenon referred to as "factor momentum" (Gupta and Kelly, 2019; Arnott et al., 2023).

Table 3
Significance rates for alternative cross-sectional standard errors.

Panel A: Alternative Clustering

		1% Significance Level				5% Significance Level			
		Univariate		Multivariate		Univariate		Multivariate	
Return	Specification	1d	5d	1d	5d	1d	5d	1d	5d
Unadj.	SIC4 Cluster	22.2%	27.8%	18.7%	24.6%	33.9%	39.6%	30.0%	36.4%
Unadj.	FF49 Cluster	18.5%	22.9%	16.3%	21.6%	30.5%	35.2%	28.0%	33.7%
Unadj.	FF30 Cluster	21.3%	25.7%	18.5%	23.7%	30.7%	34.9%	28.3%	33.6%
Unadj.	State Cluster	30.2%	36.2%	25.7%	31.9%	42.3%	48.5%	37.6%	44.4%
Unadj.	5 Size x 5 MB	20.6%	25.7%	17.0%	23.0%	32.7%	38.6%	29.0%	35.5%
Unadj.	... x 5GP	24.2%	30.4%	19.7%	26.4%	35.9%	42.1%	31.1%	38.0%

Panel B: Two-way clustering

		1% Significance Level				5% Significance Level			
		Univariate		Multivariate		Univariate		Multivariate	
Return	Specification	1d	5d	1d	5d	1d	5d	1d	5d
Unadj.	FF49 & State	20.7%	24.6%	18.8%	23.5%	32.2%	36.2%	30.0%	35.4%
Unadj.	FF30 & State	20.1%	23.7%	16.7%	21.0%	31.7%	35.5%	27.6%	32.6%

This table presents mean statistical significance rates for coefficients from cross-sectional OLS regressions of short-term stock returns on each of eight firm characteristics based on alternative cross-sectional standard errors. Each cell reports the percentage of all 1- or 5-day periods from 1991 through 2021 in which the regression coefficient is statistically significant. Each row reports the percentages for different combinations of 1% and 5% significance, univariate and multivariate regressions (where all eight characteristics are included as explanatory variables), and one-day and five-day return periods. Panel A presents rates of significance based on standard errors clustered at different industry levels, at the state level, at the level of size quintile by book-to-market quintile bin, and at the level of size quintile by book-to-market quintile by gross profitability quintile bin. Panel B presents rates of significance based on standard errors clustered by both industry and size.

to provide some context, the average default standard error across all eight characteristics in a univariate regression with a one-day event window would need to be multiplied by 3.4 (4.0) to achieve the intended rate of false positives at the 5% (1%) level, while the average Fama–French 49 industry-clustered standard error would need to be multiplied by 2.2 (2.5).

5. Time-series approaches

In this section, we evaluate the time-series approaches described in Section 2.4, where statistical testing compares the event-day return–characteristic relationship to its distribution on pre-event days.

5.1. Practical considerations

Implementing either TS-OLS or TS-GLS requires two practical choices. The first is the number of pre-event days to use, L. A longer series of pre-event days affords a larger sample of pre-event coefficient observations to use in statistical testing. However, using pre-event days farther back in time from the event increases the likelihood that the distribution of routine news affecting firms has changed. In addition, as discussed in Section 2.4, using the CDF approach requires that L equal one less than a multiple of 100 (i.e., 99, 199, 299, etc.). In our analysis, we use $L = 199$ trading days, which seems like a reasonable compromise in terms of period length and satisfies the restriction for using the p-value approach.

The second practical choice is when to measure firm characteristics if these characteristics are time-varying. To avoid look-ahead bias, the characteristics should always be measured prior to a given day on which returns are measured. One option is to use the characteristics measured on the most recent available date prior to the earliest pre-event day in the comparison set. While this approach is simple, it may be inefficient if characteristics change over time. Another option is to measure a characteristic for each day as of the most recent prior date for which it is available. For example, if a characteristic is an annual financial variable, then $x_{i,t-1}$ would be the value of x for firm i as of the most recent fiscal year-end prior to day t. Given the potential efficiency gain, we adopt this latter approach in our analysis.

Implementing the TS-GLS approach requires two additional choices – K, the number of PCs to use in constructing the covariance matrix, and T_K, the number of days to use in computing the PCs. A larger number of days T_K allows for estimating more PCs and therefore a richer covariance matrix, but using data from farther in the past increases the likelihood that changes in covariances over time make estimates less comparable. We choose $T_K = 200$, a round number comparable to our choice of L that seems like a reasonable compromise relative to the tradeoff.

The number of PCs must be a whole number satisfying $0 \leq K \leq T_K$. Using more PCs allows for more precise estimates of the return covariances. This increased precision should initially increase the efficiency of cross-sectional estimates and hence of the statistical power of the TS-GLS approach. However, beyond a certain point, adding more PCs may result in overfitting, which reduces efficiency. We provide evidence in Appendix C that explanatory power continues to increase meaningfully with the number of PCs through at least $K = 50$ and that overfitting does not become a significant issue until above $K = 195$ PCs. Based on this analysis, we use $K = 100$ in the remainder of the analysis, though we also explore the sensitivity of results to using different numbers of PCs.

5.2. Frequency of statistically significant relationships on ordinary days

Mirroring our analysis of cross-sectional tests, we begin our analysis of the TS-OLS and TS-GLS approaches by applying these approaches for each characteristic to each one-day and five-day period in our sample period and computing the fraction of periods in which the estimated relationship is statistically significant at the 1% and 5% levels. We compute these fractions for both the t-statistic and CDF approaches to statistical testing described in Section 2.4. Fig. 3 presents the results.

As anticipated, statistical significance rates for a given nominal significance level using both time-series approaches are much closer to the nominal significance level than those based on cross-sectional regression standard errors. However, these rates are still slightly higher than the nominal significance level when we use the t-statistic approach, especially at the 1% significance level. These excess rates arise from fat tails in the distribution of coefficients that a t test fails to take into account. In contrast, statistical significance rates based on p values

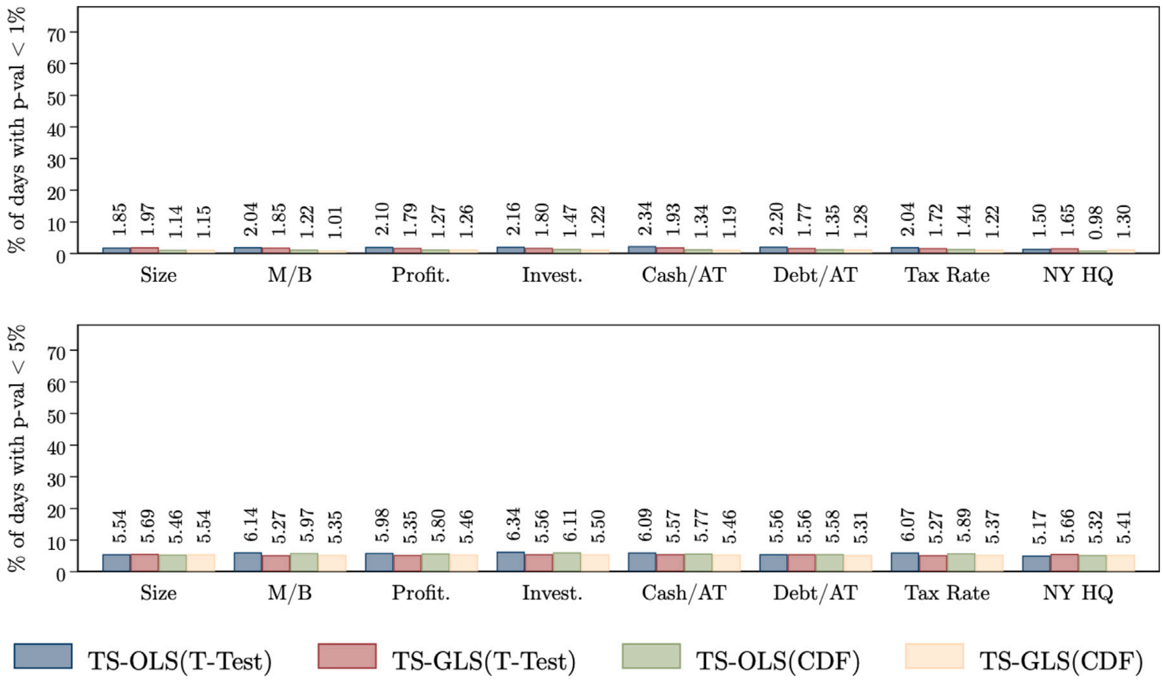

Fig. 3. Significance rates based on TS-OLS and TS-GLS.
This figure depicts statistical significance rates for the relationship between one-day returns and each of eight characteristics using the time-series OLS (TS-OLS) and time-series GLS (TS-GLS) approaches described in Section 2.4. We use 199 pre-event periods and, for each covariance matrix used in TS-GLS, 100 principal components extracted from the 200 preceding trading days. The bars show the percentage of days for which the coefficient from the regression is statistically significant based on t-statistics using the time-series of coefficients as standard errors ('T-Test') or based on a p value from comparison to the empirical time-series cumulative distribution function of coefficients ('CDF'). The top panel shows rates of significance at the 1% significance level, while the bottom panel shows rates of significance at the 5% significance level. Our sample is the 7,811 trading days in 1991 through 2021.

using the empirical CDF of coefficients are approximately equal to the nominal significance level. We therefore focus on the CDF approach for the remainder of our analysis.

The empirical success of TS-OLS and TS-GLS in rejecting at or near the nominal significance level is not mechanical or guaranteed by theory. As detailed in Section 2, it requires that the distribution of routine news, n_t, be stable over time. Significant changes in the volatility of n_t over time could make the pre-event distribution of the estimated relationship a poor proxy for the event-day distribution under the null, and we would again find frequent false positives. The empirical performance of TS-OLS and TS-GLS suggests that the distribution of n_t is generally relatively stable over the windows of around 200 trading days we use in our analysis.

5.3. *Frequency of statistically significant relationships on artificial event days*

We next assess the statistical power of the TS-OLS and TS-GLS approaches. One at a time, for each day in the period 1991–2021 and each characteristic, we create an artificial event effect by introducing a relationship between returns and the characteristic of 25 basis points (bp) per one-standard deviation increase in the characteristic. We do not add this effect on pre-event days relative to the artificial event day. We then use the TS-OLS and TS-GLS approaches to estimate and test the significance of differences in the return-characteristic relationship on the artificial event day at the 1% and 5% significance levels. Finally, for each characteristic, we compute the fraction of artificial event days on which the test detects the artificial event effect at each of the two significance levels. We repeat this process using an artificial event effect of 50 bp return per one-standard deviation increase in the characteristic. Fig. 4 presents detection rates based on these tests.

Using either time-series approach, detection rates vary widely across characteristics and are generally smaller for size, market-to-book ratio, and cash holdings. Note that these are the same characteristics for which tests based on standard errors from cross-sectional regressions produce the highest rates of statistically significant relationships absent an artificial event (Fig. 2). This is not a coincidence. As the variance in routine news, σ_n^2, increases, false positives in the cross-sectional regressions become more frequent, and detecting the differential effect of the artificial events becomes more difficult.

The main takeaway from Fig. 4 is that the TS-GLS approach detects artificial effects much more frequently than the TS-OLS approach. For example, TS-GLS detects 25 bp of additional return per one standard deviation change in a characteristic at the 5% significance level 1.2–2.7 times as often as TS-OLS, depending on the characteristic. The greatest detection rate gains occur for characteristics where TS-OLS detection rates are relatively low. To allow for a more comprehensive comparison, Table 4 reports average detection rates across all eight characteristics for TS-OLS and TS-GLS for the 25 and 50 bp artificial effects separately for one-day and five-day event periods.

Average detection rates across the eight characteristics are 1.3–1.9 times and 1.6–2.8 times as high for TS-GLS as for TS-OLS at the 5% and 1% significance levels, respectively. For both TS-GLS and TS-OLS, detection rates are naturally substantially lower for five-day event periods than for one-day event periods because five-day returns are noisier than one-day returns. The degree to which detection rates deteriorate with the length of the event period is a serious concern, since papers often analyze multi-day event periods. While not the focus of this paper, the results suggest there is a large benefit to precisely identifying when the market responds to an event, thereby allowing the use of a short event

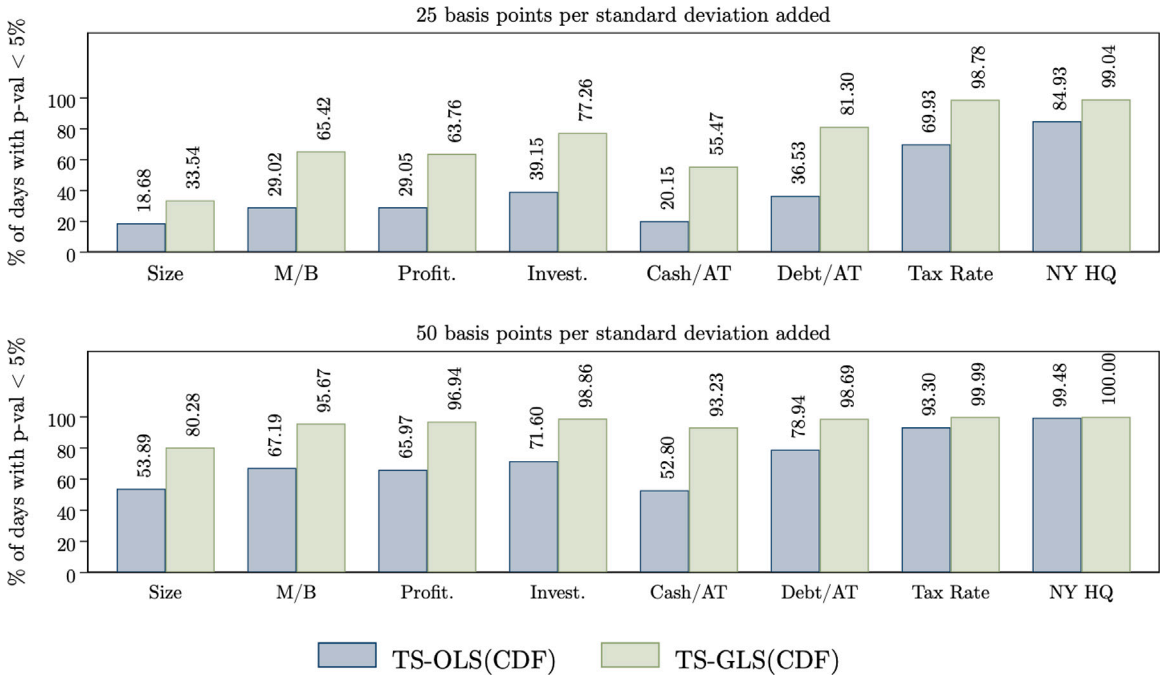

Fig. 4. Detection rates in artificial event periods: TS-OLS vs TS-GLS.
This figure depicts statistical significance rates for the relationship between one-day returns and each of eight characteristics using the time-series OLS (TS-OLS) and time-series GLS (TS-GLS) approaches described in Section 2.4. For each day, we create an artificial event by adding either 25 or 50 bp to a firm's return for each one standard deviation increase in the corresponding characteristic on the *x* axis. We use 199 pre-event days (with no added relationship) and, for each covariance matrix used in TS-GLS, 100 principal components extracted from the 200 preceding trading days. The bars show the percentage of days for which the coefficient from the regression is statistically significant at the 5% level based on a comparison to the empirical time-series cumulative distribution function of coefficients ('CDF'). The top panel shows results for a 25 bp added relationship, while the bottom panel shows results for a 50 bp added relationship. Our sample is the 7,811 trading days in 1991 through 2021.

Table 4
Detection rates in artificial event periods: TS-OLS vs TS-GLS.

Panel A: 1% Significance Level

Period	Effect (bp)	TS-OLS	TS-GLS	Difference	Ratio
1 day	25	20.76%	49.05%	28.29%	2.36
1 day	50	53.20%	85.94%	32.74%	1.62
5 day	25	5.20%	14.89%	9.69%	2.86
5 day	50	19.03%	42.34%	23.31%	2.22

Panel B: 5% Significance Level

Period	Effect (bp)	TS-OLS	TS-GLS	Difference	Ratio
1 day	25	40.93%	71.82%	30.89%	1.75
1 day	50	72.90%	95.46%	22.56%	1.31
5 day	25	15.29%	30.08%	14.79%	1.97
5 day	50	35.19%	61.24%	26.05%	1.74

This table presents statistical significance rates for the relationship between short-term (one- or five-day) returns and each of eight characteristics using the time-series OLS (TS-OLS) and time-series GLS (TS-GLS) approaches described in Section 2.4. For each one- or five-day period, we create an artificial event by adding either 25 or 50 bp to a firm's return for each one standard deviation increase in the characteristic in question, spreading the effect equally across days in the case of five-day periods. We use 199 pre-event periods (with no added relationship) and, for each covariance matrix used in TS-GLS, 100 principal components extracted from the 200 preceding trading days. Each row reports the percentage of all artificial event periods in which the estimated event-period relationship is statistically significant using TS-OLS and TS-GLS based on a p value from comparison to the empirical time-series cumulative distribution function of coefficients as well as the difference and ratio of these percentages. Different rows report these percentages for different combinations of event-period length (one- or five-day) and size of added return per one standard deviation increase in a characteristic (25 or 50 bp). The top panel shows significance rates at the 1% level, while the bottom panel shows significance rates at the 5% level. Our sample is the 7,811 trading days in 1991 through 2021.

Table 5
TS-GLS detection rates in artificial event periods: Robustness.

Panel A: 1% Significance Level

		1 day			5 day		
Effect (bp)	L	$K = 50$	$K = 100$	$K = 150$	$K = 50$	$K = 100$	$K = 150$
25	99	47.87%	48.91%	48.02%	16.12%	17.73%	17.62%
25	199	48.06%	49.05%	48.58%	13.46%	14.89%	14.71%
25	299	48.24%	49.40%	48.47%	12.54%	13.75%	13.69%
50	99	83.87%	84.74%	84.24%	43.55%	47.25%	47.77%
50	199	85.30%	85.94%	85.40%	39.04%	42.34%	42.84%
50	299	85.06%	85.64%	85.43%	37.83%	40.55%	41.12%

Panel B: 5% Significance Level

		1 day			5 day		
Effect (bp)	L	$K = 50$	$K = 100$	$K = 150$	$K = 50$	$K = 100$	$K = 150$
25	99	71.68%	72.49%	71.75%	30.01%	31.94%	32.37%
25	199	71.01%	71.82%	71.13%	28.31%	30.08%	30.48%
25	299	70.60%	71.34%	70.85%	27.74%	29.31%	29.54%
50	99	95.32%	95.51%	95.43%	60.05%	63.49%	64.34%
50	199	95.30%	95.46%	95.35%	57.94%	61.24%	62.23%
50	299	95.38%	95.57%	95.41%	57.34%	60.41%	61.38%

This table presents mean rates with which the time-series GLS (TS-GLS) approach to significance testing described in Section 2.4 successfully detects relationships between one-day returns and each of eight characteristics added to create artificial event days for different numbers of pre-event days (L) and principal components (K) used to construct covariance matrices. We use the preceding 200 trading days to compute principal components. Each row reports the percentage of all artificial event periods from 1991 to 2021 in which the estimated event-period relationship is statistically significant for different numbers of principal components. Different rows report these percentages for different combinations of number of pre-event days and size of added return per one standard deviation increase in a characteristic (25 or 50 bp). The top panel shows significance rates at the 1% level, while the bottom panel shows significance rates at the 5% level.

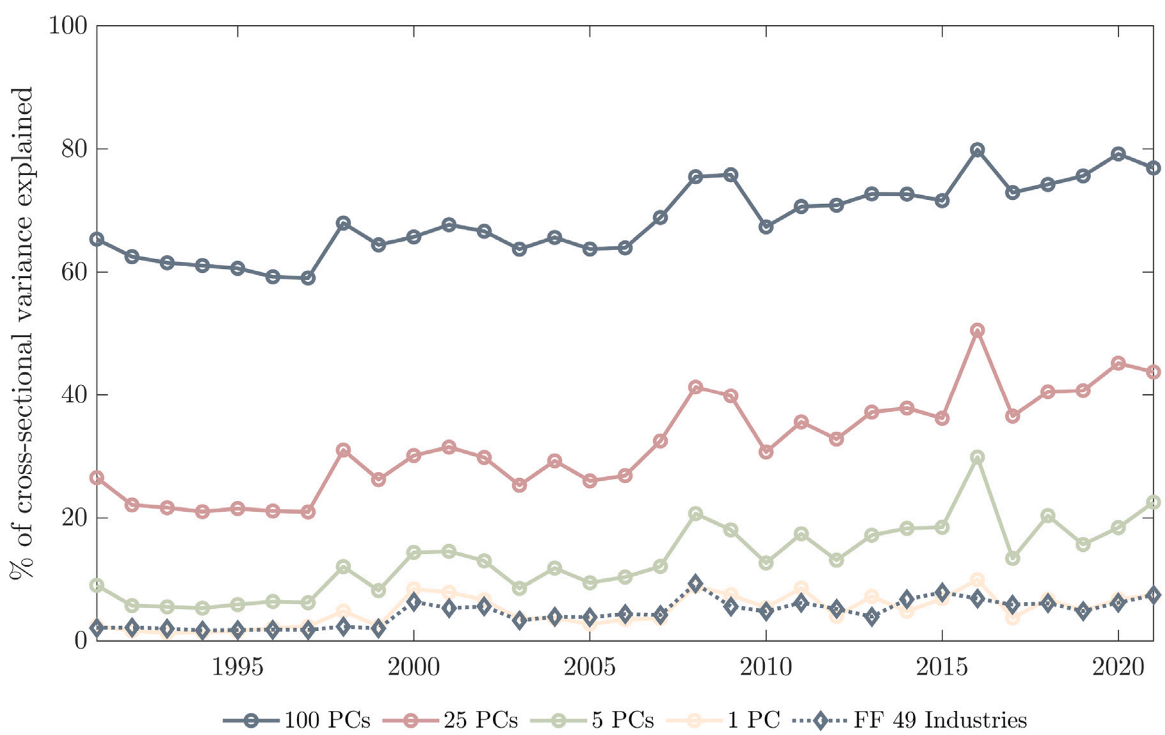

Fig. 5. Explanatory power of principal components of daily returns.
This figure plots the average percent of the daily cross-sectional variance of returns that the first K principal components (PCs) combined explain by calendar year, for $K = 1$, 5, 25, and 100. It also shows the average percent explained by Fama–French 49-category industry portfolios combined for comparison. Our sample is the 7,811 trading days in 1991 through 2021.

period. In Table 5, we show that detection rates are similar when using TS-GLS across various choices for L, the number of pre-event period windows, and K, the number of principal components.

We conclude from the analysis in this section that using either time-series approach instead of a cross-sectional regression standard error to test significance effectively addresses the threat of excessive

Table 6
Evolving correlations with principal components.

Panel A: 2008

	'Mkt'	?	'Crisis'	?	?
PC	1	2	3	4	5
% x-sectional var. explained	5.9%	3.6%	3.0%	2.7%	2.1%
$\lvert\rho(PC_i, MktRf)\rvert$	88%	4%	13%	2%	3%
$\lvert\rho(PC_i, SMB)\rvert$	28%	12%	41%	14%	20%
$\lvert\rho(PC_i, HML)\rvert$	4%	4%	46%	0%	30%
$\lvert\rho(PC_i, UMD)\rvert$	3%	8%	45%	2%	36%
$\lvert\rho(PC_i, Tech)\rvert$	7%	9%	5%	3%	6%
$\lvert\rho(PC_i, Finance)\rvert$	6%	4%	60%	4%	35%
$\lvert\rho(PC_i, Covid)\rvert$	5%	9%	27%	1%	2%
$\lvert\rho(PC_i, Memes)\rvert$	2%	4%	8%	6%	2%
R^2	88%	5%	55%	3%	19%

Panel B: 2010

	?	'Mkt'	?	?	?
PC	1	2	3	4	5
% x-sectional var. explained	5.2%	4.0%	1.6%	1.4%	1.2%
$\lvert\rho(PC_i, MktRf)\rvert$	1%	92%	10%	8%	5%
$\lvert\rho(PC_i, SMB)\rvert$	1%	30%	12%	23%	31%
$\lvert\rho(PC_i, HML)\rvert$	3%	3%	1%	26%	13%
$\lvert\rho(PC_i, UMD)\rvert$	3%	10%	15%	25%	22%
$\lvert\rho(PC_i, Tech)\rvert$	2%	0%	7%	6%	2%
$\lvert\rho(PC_i, Finance)\rvert$	9%	5%	7%	2%	6%
$\lvert\rho(PC_i, Covid)\rvert$	1%	0%	7%	22%	25%
$\lvert\rho(PC_i, Memes)\rvert$	1%	1%	14%	3%	1%
R^2	1%	95%	10%	21%	22%

Panel C: 2020

	'Mkt'	'Fin'	?	'Covid'	?
PC	1	2	3	4	5
% x-sectional var. explained	6.2%	4.7%	2.8%	2.5%	2.2%
$\lvert\rho(PC_i, MktRf)\rvert$	54%	24%	4%	7%	7%
$\lvert\rho(PC_i, SMB)\rvert$	17%	43%	1%	32%	13%
$\lvert\rho(PC_i, HML)\rvert$	36%	53%	9%	39%	0%
$\lvert\rho(PC_i, UMD)\rvert$	29%	55%	1%	45%	8%
$\lvert\rho(PC_i, Tech)\rvert$	21%	34%	9%	19%	15%
$\lvert\rho(PC_i, Finance)\rvert$	39%	49%	8%	19%	10%
$\lvert\rho(PC_i, Covid)\rvert$	24%	16%	11%	39%	19%
$\lvert\rho(PC_i, Memes)\rvert$	25%	32%	8%	9%	26%
R^2	49%	45%	6%	52%	16%

Panel D: 2021

	'Quant'	'Memes'	?	'Tech'	?
PC	1	2	3	4	5
% x-sectional var. explained	5.4%	5.1%	3.1%	2.3%	2.2%
$\lvert\rho(PC_i, MktRf)\rvert$	36%	11%	0%	17%	27%
$\lvert\rho(PC_i, SMB)\rvert$	36%	44%	2%	27%	29%
$\lvert\rho(PC_i, HML)\rvert$	45%	2%	37%	45%	48%
$\lvert\rho(PC_i, UMD)\rvert$	35%	21%	7%	8%	2%
$\lvert\rho(PC_i, Tech)\rvert$	9%	21%	17%	36%	30%
$\lvert\rho(PC_i, Finance)\rvert$	14%	5%	27%	46%	34%
$\lvert\rho(PC_i, Covid)\rvert$	14%	9%	14%	2%	4%
$\lvert\rho(PC_i, Memes)\rvert$	8%	93%	1%	4%	6%
R^2	59%	89%	15%	33%	40%

This table presents the correlations between portfolios representing the first five principal components of individual stock returns in a calendar year and a variety of other factors constructed from stock returns. The first four are the market ($MktRf$), size (SMB), value (HML) and momentum (UMD) factors. The second four are period-specific factors: $Tech$, the equal-weighted average return of Software, Hardware, and Chips Fama–French 49 industries; $Finance$, the average return of Banks, Real Estate, and Finance industries; $Covid$, the average return of Meals, Healthcare, and Drugs industries; and $Memes$, the average return of whichever subset of GameStop (GME), AMC (AMC), Bed Bath and Beyond (BBBY), and Blackberry (BB) stocks is available to trade on each day. We orthogonalize the non-$MktRf$ factors with respect to $MktRf$ in each calendar year. Panel A presents results for 2008, Panel B for 2010, Panel C for 2020, and Panel D for 2021.

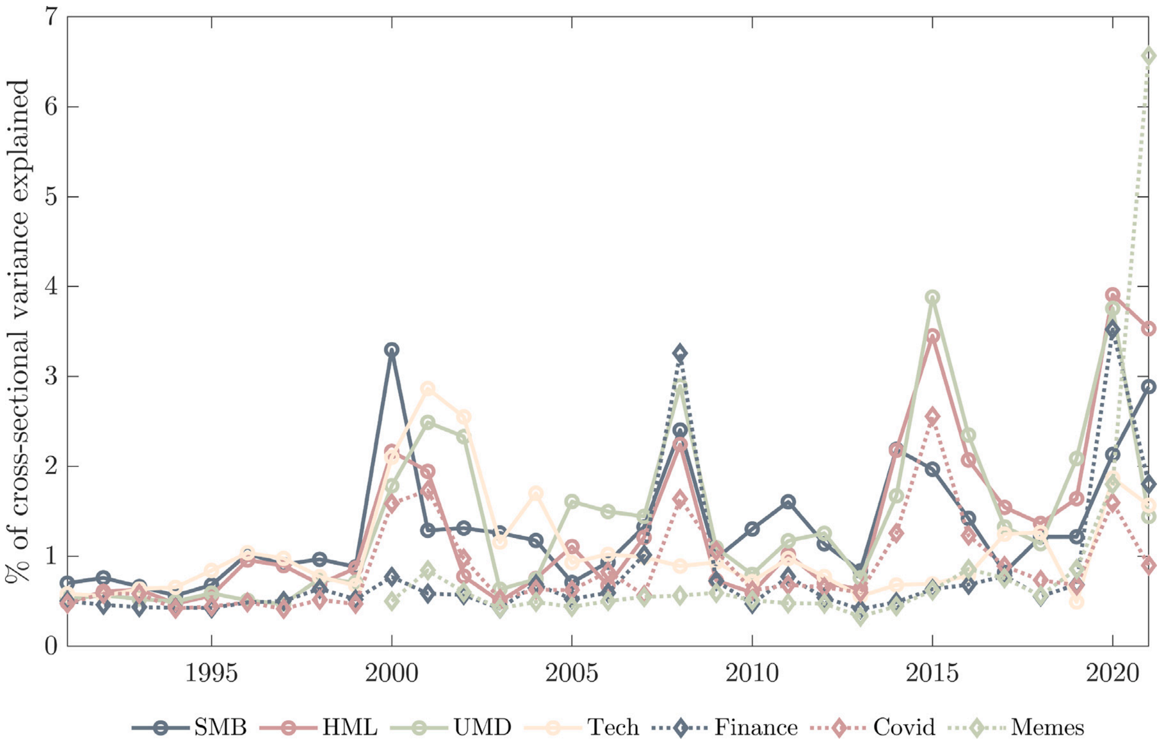

Fig. 6. Explanatory power of example factors for daily returns.
This figure plots the average percent of the daily cross-sectional variance of returns that seven example factors explain by calendar year. The factors are size (SMB), value (HML) and momentum (UMD) factors, as collected from Ken French's data library; $Tech$, the equal-weighted average return of Software, Hardware, and Chips Fama–French 49 industries; $Finance$, the average return of Banks, Real Estate, and Finance industries; $Covid$, the average return of Meals, Healthcare, and Drugs industries; and $Memes$, the average return of whichever subset of GameStop (GME), AMC (AMC), Bed Bath and Beyond (BBBY), and Blackberry (BB) stocks is available to trade on each day. All factors are orthogonalized with respect to the excess market return ($MktRf$) in each calendar year. Our sample is the 7,811 trading days in 1991 through 2021.

false positives caused by contemporaneous events. However, TS-GLS provides a big improvement in statistical power over TS-OLS. Our `csestudy` Stata and Python modules implement both approaches.

6. Analysis of return comovements

In this section, we further analyze the PCs employed in the TS-GLS approach to shed light on the structure of return covariances. As discussed in Section 2.2, return covariances reflect routine news that drives cross-sectional differences in returns. PCA therefore also offers insights into the economic drivers of the daily events that confound cross-sectional event studies. We begin by plotting the fraction of total variation in returns that the first one, five, 25, and 50 PCs explain by year. Fig. 5 presents these plots. It also shows the average explanatory power of the Fama–French 49-category industries for comparison.

Two observations are worth making. First, confirming the conclusions of Lopez-Lira and Roussanov (2023), it takes many PCs to explain the variation in returns. The PCs are essentially factors chosen sequentially to maximize incremental explanatory power. One might imagine then that only a few such factors would be needed to explain a large fraction of the variation in returns. Yet, the first 5 PCs explain less than 30% of the variation in returns in most years, and even the first 25 PCs explain less than 50% of the variation in most years. Going from 25 to 100 PCs nearly doubles the fraction of variation explained. It therefore appears that correlation patterns in returns are highly complex.

Second, the Fama–French 49-category industries together explain roughly the same fraction of cross-sectional return variation that the first PC does. In most years, just the first five PCs explain more than twice the variation that industry categories do. The limited ability of industry to summarize the cross section of returns explains why clustering by industry has limited ability to correct standard errors in purely cross-sectional tests. Return correlations are too complex to summarize well with virtually any set of pre-specified variables.

We next analyze the extent to which the first five PCs in each year map into identifiable economic factors that might plausibly drive return correlations. For the purposes of illustration, we present this analysis for three specific years – 2008, 2020, and 2021 – in which major market-moving events occurred, and for 2010 to provide a relatively quiescent year for comparison. We choose four well-known factors from the asset pricing literature: the equity market portfolio ($MktRf$), small-minus-big portfolio (SMB), high-minus-low portfolio (HML), and up-minus-down portfolio (UMD). We construct four other factors as long-only equal-weighted portfolios of stocks designed to capture period-specific conditions. The $Tech$ factor portfolio is constructed from the Software, Hardware, and Chips Fama–French 49 industries; the $Finance$ portfolio from Banks, Real Estate, and Finance industries; and the $Covid$ portfolio from Meals, Healthcare, and Drugs industries. The $Memes$ factor portfolio combines whichever subset of GameStop, AMC, Bed Bath & Beyond, and Blackberry stocks was available to trade on each day. We orthogonalize the non-$MktRf$ factors with respect to $MktRf$ within each calendar year.

For each of the first five PCs in each year, we compute the returns on a portfolio where the weights are the elements of the PC. We then compute the absolute values of the correlations between each of these returns and each of the pre-specified factors. Table 6 presents the results. Panels A, B, C, and D present the results for 2008, 2010, 2020, and 2021, respectively.

The first PC-weighted portfolio return is highly correlated with the equity market portfolio in three of the four years. Because we de-mean returns within each stock-day prior to conducting PCA, the first PC does not mechanically capture the simple average return on each day. Instead, the first PC is often, but not always, correlated with $MktRf$ because it captures differences in daily returns driven by cross-sectional differences in market betas. Furthermore, because PCA weights stocks equally, the resulting first PC is effectively equal-weighted, while $MktRf$ is value-weighted. For this reason, the first PC is also moderately correlated with the SMB factor portfolio return in all four years.

For 2008, the third PC-weighted portfolio return is highly correlated with the $Finance$, HML, and UMD factor portfolio returns. This PC appears to pick up common exposure to aspects of the financial crisis. For 2020, the fourth PC-weighted portfolio return is correlated with the $Covid$ factor portfolio return. For 2021, the second PC-weighted portfolio return is highly correlated with the $Memes$ factor portfolio return, while the fourth and fifth are correlated with the HML, $Finance$, and $Tech$ factor portfolio returns. For 2010, a relatively quiescent year, only the second PC-weighted portfolio return is strongly correlated with any of the factor portfolio returns.

We broaden our analysis by calculating the fraction of all return variances explained by the factors presented in Table 6 in each year of the sample period. The results, presented in Fig. 6, show that each factor is an important determinant of return variation in some years and largely irrelevant in others.[18] While this time variation is most dramatic for the $Finance$ and $Memes$ factors, whose explanatory powers peak in 2008 and 2021, respectively, even standard asset pricing factors such as SMB vary significantly over time in their importance. Fig. 6 also shows that there are many years, particularly early in our sample period, where none of the example factors is important in driving return variation.[19]

Overall, the analyses in Table 6 and Fig. 6 indicate that factors driving return comovement vary substantially from year to year and often represent factors unique to a period. Furthermore, it is often difficult to determine the precise economic drivers of comovement in any given period because correlation patterns are complex and time-varying. These patterns support the use of many statistically extracted components to power our GLS approach.

7. Conclusions

We show that traditional tests of the relationship between returns around a quasi-experimental event and firm characteristics produce too many false positives to allow for reliable inference about the differential treatment effect of the event. We also show that using the distribution of relationships on pre-event days to conduct statistical tests is an effective solution to this problem. An event-day relationship should stand out relative to the distribution on days when the event does not occur before we conclude that the event causes differences in returns. Our analysis shows that a GLS-based approach to estimating event and pre-event day relationships offers substantial improvements in statistical power over a more conventional OLS-based approach. It may therefore allow for the detection of even modest differential event effects that might otherwise be difficult to test.

CRediT authorship contribution statement

Jonathan B. Cohn: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Travis L. Johnson:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Zack Liu:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Malcolm I. Wardlaw:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Dimitris Papanikolaou was the editor for this article. We thank Roberto Gomez-Cram (EWFC Discussant), Paul Irvine (Lone Star Discussant), Tomasz Kaczmarek (Modern Finance Discussant), Minjoo Kim, Sheridan Titman, Steven Urry, Zirui Wang, James Weston (FRA Discussant), an anonymous referee, and seminar participants at Auburn University, Chinese University of Hong Kong, Clemson University, European Winter Finance Conference (EWFC), Financial Research Association (FRA) Conference, Hong Kong University, Hong Kong University of Science and Technology, Lone Star Finance Conference, Modern Finance Conference, Nanyang Technological University, National University of Singapore, Reichman University, UC-Davis FMA Napa Finance Conference, University of Houston, University of Massachusetts at Amherst, University of Oklahoma, University of St. Gallen, University of Texas at Austin, and University of Zurich for valuable feedback. We thank Christopher McDonald and George Rao for excellent research assistance.

Appendix A. Survey of recent papers using cross-sectional event studies

We survey recently published papers to assess how papers presenting results from cross-sectional event studies of returns around a single event typically calculate standard errors and other research design choices. The universe of papers we consider is those published in a set of top finance and accounting journals (*Journal of Accounting and Economics*, *Journal of Accounting Research*, *Journal of Banking & Finance*, *Journal of Corporate Finance*, *Journal of Finance*, *Journal of Financial Economics*, *Journal of Financial and Quantitative Analysis*, *Management Science*, *Review of Corporate Finance Studies*, *Review of Finance*, and *Review of Financial Studies*) from 2014 through 2023. We prompt version 3.5 of the large language model "Claude" to read all papers in this initial universe and identify papers that employ cross-sectional event studies of returns around a single event. To filter out other types of event studies, our prompt repeatedly emphasizes the importance of a single common event, provides an example regression specification, and gives multiple examples of what we are *not* looking for. We then manually review the papers Claude identifies to verify they actually use the methodology. This process yields 81 papers with 16,070 citations collectively on Google Scholar. For each paper in this set, we catalog:

1. Google Scholar citation count as of January 2025.
2. What type of returns the authors use (raw; market-adjusted; factor-adjusted; or log).
3. How long the event window is.
4. Whether fixed effects are used and what variety (industry or other).

[18] As a point of reference, SMB has the highest average percent explained, 2.7%, comparable to the average for the second PC. The other seven factors' average percent explained are all between 0.8% and 1.3%, comparable to the average for PCs 19 and 7, respectively.

[19] In untabulated analysis, we find that there are many years where at least one of the first five PCs does not strongly correlate with any of the 170+ factors from Chen et al. (2022).

5. How the authors compute standard errors (unadjusted or White (1980) robust; industry clustered; geography clustered; time-series; or other).[20]

Table A.1 provides details for these papers.

Panel A summarizes the frequency of methodological choices among these papers. Only five papers (6%) use a variant of the time-series methodology we describe in Section 2.4 to test the significance of event-period relationships. 25% use standard errors clustered on a cross-sectional dimension. The majority (64%) use unadjusted or White (1980) robust standard errors. Of secondary interest, simple returns are less common (25%) than returns adjusted for CAPM or Fama–French factors, and there is an approximately even split between using industry fixed effects (44%) and not (56%).

Panel B provides details for each of the 81 papers. Of the five papers that use a time-series approach, three (Akiyoshi, 2019; Borisov et al., 2016; Kolasinski and Yang, 2023) use the Sefcik and Thompson (1986) portfolio OLS approach. O'Donovan et al. (2019) also use a portfolio-based approach that appears similar to the Sefcik and Thompson (1986) approach. Cohn et al. (2016) use the t-statistic version of the time-series OLS approach that we describe in this paper.

Appendix B. General conceptual framework and proofs

In this appendix, we present a formal conceptual framework for our analysis that generalizes the framework in Section 2 to allow for J different firm characteristics; N-day event windows; non-zero relationships between firm characteristics and average returns; and a general covariance matrix for returns.

The general data generating process is:

$$r_{i,t} = a_t + \left[\lambda + n_t + \frac{e}{N}\mathbb{1}(t\in\tau)\right]'(x_{i,t-1} - \bar{x}_{t-1}) + \delta_{i,t}, \tag{B.1}$$

$$\mathbb{E}[n_t] = \mathbb{E}[\delta_{i,t}] = \mathbb{E}[x_{i,t-1}\delta_{i,t}] = \mathbb{E}[n_t\delta_{i,t}] = 0, \mathrm{Var}(\delta_t) = \Omega_\delta, \tag{B.2}$$

where a_t is a random variable capturing the aggregate return implications of any news on day t, λ is a $J\times 1$ vector containing each characteristic's relation with average returns, n_t is a $J\times 1$ vector containing mean-zero random variables expressing how the news arriving in period t affects returns as a function of firm characteristics, e is a $J\times 1$ vector containing the differential effects of the event on per-day stock returns, τ is the set of days in the event period, $x_{i,t}$ is a $J\times 1$ vector of firm characteristics, $\bar{x}_{t-1}$ is the mean $x_{i,t}$, $\delta_{i,t}$ is a mean-zero error term, and δ_t is an $M\times 1$ vector containing all $\delta_{i,t}$ with covariance matrix Ω_δ.

The general framework in Eq. (B.1) simplifies to the framework studied in the body of the paper when $N = J = 1$, $\lambda = 0$, and Ω_δ is a diagonal matrix with constant variance. Proofs in this general framework therefore apply to the corresponding results in the simplified model.

Suppose a researcher conducts a cross-sectional OLS regression of total event-period returns on a constant and pre-event firm characteristics $x_{i,\tau-1}$:

$$\sum_{t\in\tau} r_{i,t} = a_\tau + b_\tau' x_{i,\tau-1} + \epsilon_{i,t}, \tag{B.3}$$

and uses the resulting $\hat{b}_\tau$ as an estimate of e.

Result B.1 (*Variance of b as an Estimate of e*)**.** *Assuming characteristics $x_{i,t}$ are constant for each i during the event window, cross-sectional OLS estimates of b_τ in Eq. (B.3) have the following relationships to the true event-specific effect e:*

$$\mathbb{E}[\hat{b}_\tau - e] = \underbrace{N\lambda}_{bias}, \tag{B.4}$$

$$Var(\hat{b}_\tau - e) = \underbrace{Var(\hat{b}_\tau - b_\tau)}_{OLS\ estimation\ error} + \underbrace{N\sigma_n^2}_{Variance\ of\ routine\ news}. \tag{B.5}$$

where the $\mathbb{E}[\cdot]$ and $Var(\cdot)$ operators condition on $x_{i,\tau-1}$ and draw random $\delta_{i,t}$ and n_t for all $t\in\tau$.

Proof. Summing Eq. (B.1) across event window days yields:

$$\sum_{t\in\tau} r_{i,t} = \sum_{t\in\tau} a_t - \left(N\lambda + \sum_{t\in\tau} n_t + e\right)'\bar{x}_{t-1} + \left(N\lambda + \sum_{t\in\tau} n_t + e\right)' x_{i,\tau-1} + \sum_{t\in\tau}\delta_{i,t}, \tag{B.6}$$

which implies that in the population (with no estimation error):

$$a_\tau = \sum_{t\in\tau} a_t - \left(N\lambda + \sum_{t\in\tau} n_t + e\right)'\bar{x}_{t-1} \tag{B.7}$$

$$b_\tau = N\lambda + \sum_{t\in\tau} n_t + e, \tag{B.8}$$

$$\epsilon_{i,\tau} = \sum_{t\in\tau}\delta_{i,t}. \tag{B.9}$$

From this we get:

$$\mathrm{Var}(\hat{b}_\tau - e) = \mathrm{Var}(\hat{b}_\tau - b_\tau + b_\tau - e) = \mathrm{Var}\left(\hat{b}_\tau - b_\tau + N\lambda + \sum_{t\in\tau} n_t\right). \tag{B.10}$$

Because $\hat{b}_\tau$ is an OLS estimate,

$$\begin{bmatrix}\hat{a}_\tau - a_\tau\\ \hat{b}_\tau - b_\tau\end{bmatrix} = (X'X)^{-1}X'\epsilon_\tau, \tag{B.11}$$

where X is an $M\times(J+1)$ matrix containing a constant and $x_{i,\tau-1}$ for all i, and ϵ_τ is an $M\times 1$ vector containing all $\epsilon_{i,t}$. Using $\mathbb{E}[n_t\delta_{i,t}] = 0$, Eq. (B.11) therefore implies that $\hat{b}_\tau - b_\tau$ is uncorrelated with n_t, and Eq. (B.10) simplifies to:

$$\mathrm{Var}(\hat{b}_\tau - e) = \mathrm{Var}\left(\hat{b}_\tau - b_\tau\right) + N\sigma_n^2. \tag{B.12}$$

We complete the proof by noting, from Eq. (B.9) and the unbiasedness of OLS, that $\mathbb{E}[\hat{b}_\tau] - e = b_\tau - e = N\lambda$. □

We now turn to the time-series OLS approach articulated in Section 2.4.

Result B.2 (*Unbiasedness and Correct Variance of the Time-Series OLS Approach*)**.** *Write $b_{\tau_{-\ell}}$ for the vector of coefficients relating $\sum_{t\in\tau_{-\ell}} r_{i,t}$ to $x_{i,\tau_{-\ell}}$ in non-event window $\tau_{-\ell}$, and b_τ for the coefficients in the event window. The time-series OLS estimate of e is:*

$$\hat{e} = \hat{b}_\tau - \frac{1}{L}\sum_{\ell=1}^{L}\hat{b}_{\tau_{-\ell}}, \tag{7}$$

Assuming stationary distributions for estimation error $\hat{b}_{\tau_{-\ell}} - b_{\tau_{-\ell}}$ and news n_t, $\hat{e}_\tau$ has the following distribution across samples as $L\to\infty$ for a given e:

$$\mathbb{E}[\hat{e} - e] = \underbrace{0}_{No\ bias}, \tag{8}$$

$$\underbrace{Var(\hat{e} - e)}_{Sampling\ variance} = \underbrace{Var(\hat{b}_{\tau_{-\ell}})}_{Time\text{-}series\ sample\ variance}, \tag{9}$$

$$\underbrace{\Phi(\hat{e} - e)}_{Sampling\ CDF} = \underbrace{\Phi(\hat{b}_{\tau_{-\ell}})}_{Time\text{-}series\ sample\ CDF}. \tag{10}$$

Proof. Because e has no effect when $t\notin\tau$, the stationarity of $\hat{b}_{\tau_{-\ell}} - b_{\tau_{-\ell}}$ and news n_t imply:

$$\Phi(\hat{b}_\tau - e) = \Phi(\hat{b}_{\tau_{-\ell}}) \Rightarrow \tag{B.13}$$

$$\mathbb{E}[\hat{b}_\tau - e] = \mathbb{E}[\hat{b}_{\tau_{-\ell}}], \tag{B.14}$$

[20] For papers that do not report how they calculate standard errors, which is unfortunately common, we tabulate them in the unadjusted or White (1980) robust category.

Table A.1
Survey of cross-sectional event studies involving a quasi-experimental event.

Panel A: Summary Statistics

Total Papers:	81				
Total Google Cites:	16,070				
Return Types			Standard Errors		
Simple	20	25%	Simple	52	64%
CAPM Adjusted	40	49%	Industry Cluster	16	20%
FF Adjusted	16	20%	Geography Cluster	4	5%
Industry Adjusted	4	5%	Time-Series	5	6%
Log	1	1%	Other	4	5%
Fixed Effects					
None	45	56%			
Industry	30	37%			
Geography	1	1%			
Industry & Geography	5	6%			

Panel B: Survey of papers

Paper	Journal	Google Cites	Return Type	Fixed Effects	Standard Errors
Abudy et al. (2020)	*JBF*	36	Simple	None	Simple/Robust
Acemoglu et al. (2016)	*JFE*	844	CAPM Adj.	None	Simple/Robust
Acemoglu et al. (2018)	*RFS*	540	Simple	Industry	Simple/Robust
Acker et al. (2018)	*JBF*	11	Simple	Industry	Simple/Robust
Akey and Appel (2021)	*JF*	220	CAPM Adj.	None	Simple/Robust
Akiyoshi (2019)	*JFE*	19	CAPM Adj.	Industry	Time-Series
Albuquerque et al. (2020b)	*JCF*	14	CAPM Adj.	Industry	Ind. Cluster
Albuquerque et al. (2020a)	*RCFS*	1206	CAPM Adj.	Industry	Simple/Robust
Amihud and Stoyanov (2017)	*JFE*	76	FF Adj.	Industry	Simple/Robust
Amore et al. (2022)	*JBF*	151	CAPM Adj.	Ind. & Geog.	Simple/Robust
Andrade and Chhaochharia (2018)	*RFS*	36	FF Adj.	Industry	Ind. Cluster
Andrieş et al. (2020)	*JBF*	20	CAPM Adj.	None	Simple/Robust
Aragon and Li (2019)	*JFE*	39	Simple	None	Simple/Robust
Bae et al. (2021)	*JCF*	641	Simple	Industry	Simple/Robust
Berkowitz et al. (2015)	*JFE*	213	CAPM Adj.	Industry	Simple/Robust
Beuselinck et al. (2017)	*JCF*	174	Simple	Ind. & Geog.	Geo. Cluster
Bhojraj et al. (2017)	*JAE*	73	FF Adj.	Industry	Ind. Cluster
Bizjak et al. (2022)	*JF*	38	CAPM Adj.	Industry	Simple/Robust
Black et al. (2015)	*JBF*	194	CAPM Adj.	None	Ind. Cluster
Bongini et al. (2015)	*JBF*	140	CAPM Adj.	None	Simple/Robust
Borisov et al. (2016)	*RFS*	272	FF Adj.	Industry	Time-Series
Brown and Huang (2020)	*JFE*	198	CAPM Adj.	Industry	Ind. Cluster
Carboni et al. (2017)	*JBF*	71	CAPM Adj.	None	Simple/Robust
Chan and Kwok (2017)	*JBF*	74	Simple	Industry	Simple/Robust
Chang et al. (2014)	*JBF*	348	CAPM Adj.	None	Simple/Robust
Chen et al. (2018)	*JAE*	1384	FF Adj.	Industry	Simple/Robust
Chen et al. (2019)	*JCF*	63	FF Adj.	None	Simple/Robust
Child et al. (2021)	*JFE*	87	CAPM Adj.	Industry	Ind. Cluster
Choy et al. (2017)	*JCF*	50	CAPM Adj.	None	Other
Chung et al. (2020)	*JCF*	41	CAPM Adj.	None	Simple/Robust
Cohn et al. (2016)	*JF*	107	Simple	None	Time-Series
Cousins et al. (2020)	*MS*	91	CAPM Adj.	None	Ind. Cluster
Doidge and Dyck (2015)	*JF*	179	Simple	None	Ind. Cluster
Fahlenbrach et al. (2021)	*RFS*	658	Simple	Industry	Simple/Robust
Fang et al. (2023)	*JBF*	39	FF Adj.	Industry	Simple/Robust
Fisman et al. (2014)	*RFS*	188	FF Adj.	Industry	Simple/Robust
Foley-Fisher et al. (2016)	*JFE*	179	CAPM Adj.	None	Ind. Cluster
Frattaroli (2020)	*JFE*	50	CAPM Adj.	None	Simple/Robust
Gao et al. (2018)	*JBF*	71	Industry Adj.	None	Simple/Robust
Garel and Petit-Romec (2021)	*JCF*	280	Simple	Industry	Simple/Robust
Gilje and Taillard (2017)	*RFS*	133	Simple	None	Simple/Robust
Gómez-Cram and Olbert (2023)	*RFS*	16	Simple	None	Simple/Robust
Greene et al. (2020)	*JCF*	228	CAPM Adj.	Industry	Ind. Cluster
Grewal et al. (2019)	*MS*	566	Simple	None	Geo. Cluster
Hackbarth et al. (2015)	*RFS*	95	CAPM Adj.	None	Simple/Robust
Hill et al. (2019)	*JBF*	110	Simple	Industry	Simple/Robust
Huang et al. (2020)	*JCF*	19	CAPM Adj.	Geography	Simple/Robust
Igan et al. (2023)	*JBF*	55	Log	None	Geo. Cluster
Jain et al. (2019)	*JBF*	11	Simple	None	Simple/Robust
Kalcheva et al. (2020)	*JBF*	37	FF Adj.	None	Simple/Robust
Kim and Kim (2022)	*JCF*	2	CAPM Adj.	Industry	Other
Kolasinski and Yang (2023)	*MS*	0	FF Adj.	None	Time-Series
Kundu (2023)	*JFQA*	1	FF Adj.	Ind. & Geog.	Ind. Cluster

(*continued on next page*)

Table A.1 (*continued*).

Lewis and Verwijmeren (2014)	*JCF*	19	CAPM Adj.	Industry	Simple/Robust
Licht et al. (2018)	*JFE*	41	CAPM Adj.	None	Other C.S.
Lin et al. (2019)	*JAE*	42	Industry Adj.	None	Simple/Robust
Lins et al. (2017)	*JF*	2837	Simple	Industry	Simple/Robust
Liu et al. (2017)	*JFE*	365	CAPM Adj.	None	Ind. Cluster
Liu et al. (2021)	*JBF*	43	CAPM Adj.	None	Ind. Cluster
Liu et al. (2022)	*JBF*	19	CAPM Adj.	None	Simple/Robust
Loipersberger (2018)	*JBF*	67	CAPM Adj.	None	Other C.S.
Marsh (2023)	*JBF*	5	FF Adj.	None	Simple/Robust
May (2014)	*JCF*	18	FF Adj.	None	Simple/Robust
Moon and Schoenherr (2022)	*JFE*	15	Simple	None	Simple/Robust
Nesbitt et al. (2023)	*JAE*	18	CAPM Adj.	None	Simple/Robust
Neukirchen et al. (2023)	*JBF*	20	CAPM Adj.	Ind. & Geog.	Geo. Cluster
Nguyen and Phan (2020)	*JCF*	273	CAPM Adj.	None	Simple/Robust
O'Donovan et al. (2019)	*RFS*	239	CAPM Adj.	None	Time-Series
Pagano et al. (2023)	*JFE*	260	CAPM Adj.	None	Ind. Cluster
Ramelli et al. (2021b)	*RCFS*	131	CAPM Adj.	Industry	Simple/Robust
Ramelli et al. (2021a)	*JCF*	91	CAPM Adj.	Ind. & Geog.	Simple/Robust
Raykov and Silva-Buston (2020)	*JCF*	9	CAPM Adj.	None	Simple/Robust
Schoenherr (2019)	*JF*	325	Simple	Industry	Simple/Robust
Shan and Tang (2023)	*RoF*	139	FF Adj.	Industry	Ind. Cluster
Stahl (2023)	*RoF*	2	Industry Adj.	None	Simple/Robust
van Bekkum (2016)	*JBF*	7	Industry Adj.	None	Simple/Robust
Wagner et al. (2018)	*JFE*	337	Simple	Industry	Simple/Robust
Wang and Chou (2018)	*JBF*	120	FF Adj.	None	Simple/Robust
Wilson (2020)	*JCF*	5	FF Adj.	None	Simple/Robust
Zeume (2017)	*RFS*	217	CAPM Adj.	Industry	Ind. Cluster
Zhang (2015)	*JBF*	18	CAPM Adj.	None	Simple/Robust

This table lists and summarizes 81 papers published in top finance and accounting journals (*JAE*, *JAR*, *JBF*, *JCF*, *JF*, *JFE*, *JFQA*, *MS*, *RCFS*, *RFS*, *RoF*) from 2014–2023 using the cross-sectional event studies methodology around a quasi-experimental event. Panel A shows the total Google Scholar citation count across all papers, and provides paper counts and fractions of the total paper count for each category of return type, fixed effects, and standard errors. The five categories of returns used are simple or raw returns ("Simple"); CAPM or market adjusted ("CAPM Adj."); Fama French or factor adjusted ("FF Adj."); Industry adjusted ("Industry Adj."); Log ("Log"). The four categories of fixed effects used are None, Industry, Geography, and Industry & Geography; The five categories of standard error calculation are standard OLS or White (1980) robust ("Simple/Robust"); clustered by industry ("Ind. Cluster"); clustered by Geography ("Geo. Cluster"); time-series or portfolio-based ("Time-Series"); and bootstrapped or other cross-sectional approach ("Other C.S."). For each paper, Panel B reports the journal of publication, the number of Google Scholar citations as of January 2025, the return type, which fixed effects are used, and how standard errors are calculated.

$$\mathrm{Var}(\hat{b}_\tau - e) = \mathrm{Var}(\hat{b}_{\tau_{-\ell}}), \tag{B.15}$$

where all moments are across samples unless otherwise stated. We therefore have:

$$\mathbb{E}[\hat{e} - e] = \mathbb{E}\left[\hat{b}_\tau - \frac{1}{L}\sum_{\ell=1}^{L} \hat{b}_{\tau_{-\ell}} - e\right] = \mathbb{E}[\hat{b}_\tau - e] - \mathbb{E}[\hat{b}_{\tau_{-\ell}}] = 0. \tag{B.16}$$

As $L \to \infty$, time-series estimates of Φ and variance converge to their across-sample counterparts. Combined with Eqs. (B.15) and (B.13), this gives us Eqs. (9) and (10). □

Appendix C. Number of principal components

We use two approaches to gain insight into the optimal number of PCs to use. First, we analyze the relationship between the number of PCs K and the variance of the minimum variance portfolio constructed using an ex-ante estimate of the covariance matrix of returns Ω based on PCA, as specified in Eq. (15).[21] Given an estimated covariance matrix for returns on day t, $\hat{\Omega}_t$, the minimum variance portfolio's weights $w_{mvp,t}$ are

$$w_{mvp,t} = \underset{w \ \ s.t.\ w'\mathbf{1}=1}{\arg\min} \ Var_{t-1}(w'r_t) = \frac{\hat{\Omega}_t^{-1}\mathbf{1}}{\mathbf{1}'\hat{\Omega}_t^{-1}\mathbf{1}}, \tag{C.1}$$

where $\mathbf{1}$ is a vector of ones, and the denominator assures that the $w_{mvp,t}$ sum to one.

The relationship between K and the volatility of the minimum variance portfolio is informative about the incremental information content of each additional PC for forecasting the inverse of the next-day covariance matrix – exactly the object we use for GLS. Fig. C.1 plots this relationship. The variance of the minimum variance portfolio decreases sharply with the addition of the first several PCs. The variance flattens out around $K = 30$ and is largely invariant until $K = 180$.

Second, for different numbers of PCs, K, and different characteristics, we estimate daily return regressions using both GLS and OLS and then analyze the tightness of the GLS coefficient distribution relative to the OLS coefficient distribution. Since the time-series approaches involve comparing the event-day relationship to the pre-event relationship distribution, a tighter distribution implies greater ease in detecting an event-related effect. To quantify the relative tightness, we compute the ratio of the standard deviations of the daily GLS and OLS coefficients. Note that the OLS coefficient standard deviation is constant for a given characteristic in this calculation since we are only varying the number of PCs used to estimate GLS. We compute the ratio because it is informative about the increase in precision from using GLS instead of OLS. For each characteristic we study, Fig. C.2 plots this ratio against the number of PCs used.

As with the relationship between minimum variance and number of PCs, the ratio of GLS to OLS coefficient standard deviations declines sharply with the first twenty PCs and is essentially flat between 40 and 190 PCs. This conclusion adds to the evidence that the power gains from TS-GLS are largely insensitive to the number of PCs beyond the first twenty. In the spirit of choosing a round number, we use $K = 100$

[21] Clarke et al. (2006) shows that using PCA to estimate the covariance matrix and form a minimum-variance portfolio of US equities results in substantial risk reduction with little or no reduction in average returns.

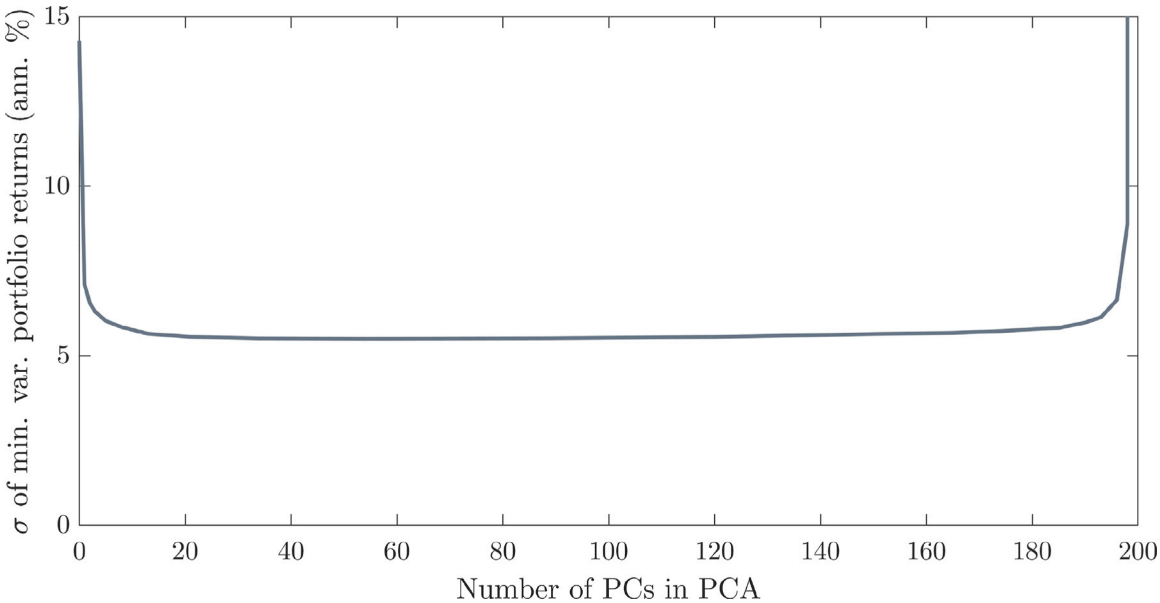

Fig. C.1. Minimum variance portfolio and # of principal components.
This figure depicts the average annualized standard deviation of daily returns for a minimum-variance portfolio of US equities as a function of the number of principal components (PCs), K, used to form an out-of-sample forecast for the covariance matrix of returns. Our covariance matrix forecasts are constructed using the first K PCs from implementing principal component analysis on a balanced panel of daily returns over the prior 200 trading days, assuming that, other than via these PCs, each stock's return is uncorrelated. See Section 2.4 for more details. Our sample is the 7,811 trading days in 1991 through 2021.

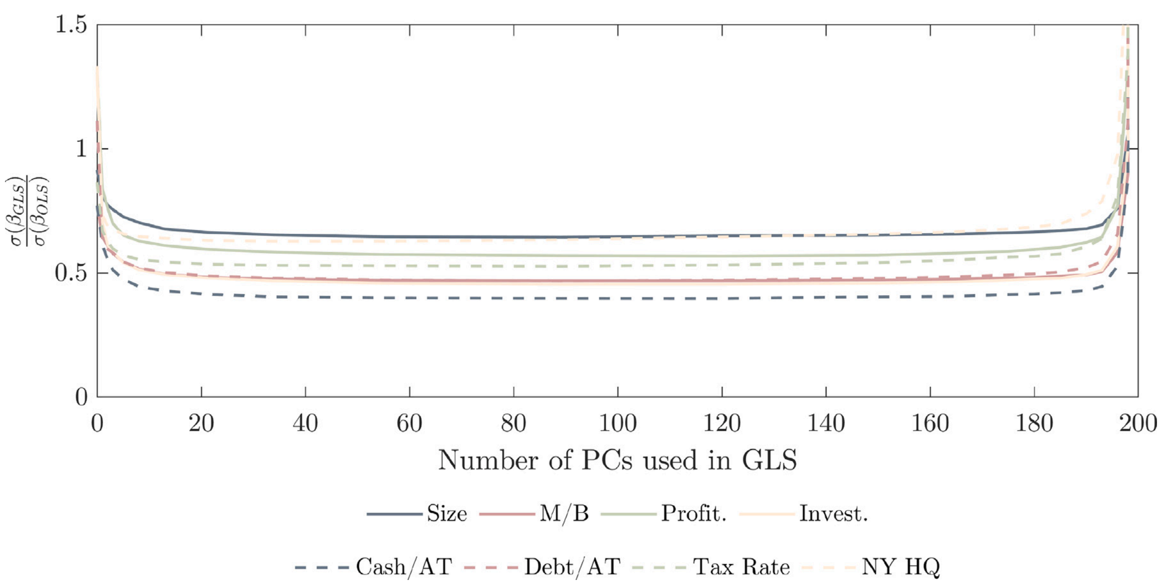

Fig. C.2. GLS/OLS coefficient volatility and # of principal components.
This figure plots the ratio of the time-series standard deviations of cross-sectional GLS and OLS coefficients from cross-sectional regressions of one-day returns on each of eight characteristics against the number of PCs we use to estimate the covariance matrix for the GLS regressions. See Section 2.4 for more details. Our sample is the 7,811 trading days in 1991 through 2021.

as our primary specification, though we also explore the sensitivity of results to using different numbers of PCs.

Data availability

Past is Prologue: Inference from the Cross Section of Returns Around a n Event replication kit (Original data) (Mendeley Data)

References

Abadie, A., Athey, S., Imbens, G.W., Wooldridge, J.M., 2023. When should you adjust standard errors for clustering? Q. J. Econ. 138 (1), 1–35. http://dx.doi.org/10.1093/qje/qjac038.

Abudy, M.M., Amiram, D., Rozenbaum, O., Shust, E., 2020. Do executive compensation contracts maximize firm value? Indications from a quasi-natural experiment. J. Bank. Financ. 114, 105787. http://dx.doi.org/10.1016/j.jbankfin.2020.105787.

Acemoglu, D., Hassan, T.A., Tahoun, A., 2018. The power of the street: Evidence from Egypt's Arab Spring. Rev. Financ. Stud. 31 (1), 1–42. http://dx.doi.org/10.1093/rfs/hhx086.

Acemoglu, D., Johnson, S., Kermani, A., Kwak, J., Mitton, T., 2016. The value of connections in turbulent times: Evidence from the United States. J. Financ. Econ. 121 (2), 368–391. http://dx.doi.org/10.1016/j.jfineco.2015.10.001.

Acker, D., Orujov, A., Simpson, H., 2018. Political donations and political risk in the UK: Evidence from a closely fought election. J. Bank. Financ. 92, 146–167. http://dx.doi.org/10.1016/j.jbankfin.2018.05.009.

Akey, P., Appel, I., 2021. The limits of limited liability: Evidence from industrial pollution. J. Financ. 76 (1), 5–55. http://dx.doi.org/10.1111/jofi.12978.

Akiyoshi, F., 2019. Effects of separating commercial and investment banking: Evidence from the dissolution of a joint venture investment bank. J. Financ. Econ. 134 (3), 703–714. http://dx.doi.org/10.1016/j.jfineco.2019.05.004.

Albuquerque, R., Koskinen, Y., Yang, S., Zhang, C., 2020a. Resiliency of environmental and social stocks: An analysis of the exogenous COVID-19 market crash. Rev. Corp. Financ. Stud. 9 (3), 593–621. http://dx.doi.org/10.1093/rcfs/cfaa011.

Albuquerque, R., Lei, Z., Rocholl, J., Zhang, C., 2020b. Citizens United vs. FEC and corporate political activism. J. Corp. Financ. 60, 101547. http://dx.doi.org/10.1016/j.jcorpfin.2019.101547.

Amihud, Y., Stoyanov, S., 2017. Do staggered boards harm shareholders? J. Financ. Econ. 123 (2), 432–439. http://dx.doi.org/10.1016/j.jfineco.2016.12.007.

Amore, M.D., Pelucco, V., Quarato, F., 2022. Family ownership during the Covid-19 pandemic. J. Bank. Financ. 135, 106385. http://dx.doi.org/10.1016/j.jbankfin.2021.106385.

Andrade, S.C., Chhaochharia, V., 2018. The costs of sovereign default: Evidence from the stock market. Rev. Financ. Stud. 31 (5), 1707–1751. http://dx.doi.org/10.1093/rfs/hhx067.

Andrieş, A.M., Nistor, S., Ongena, S., Sprincean, N., 2020. On becoming an O-SII ("other systemically important institution"). J. Bank. Financ. 111, 105723. http://dx.doi.org/10.1016/j.jbankfin.2019.105723.

Aragon, G.O., Li, L., 2019. The use of credit default swaps by bond mutual funds: Liquidity provision and counterparty risk. J. Financ. Econ. 131 (1), 168–185. http://dx.doi.org/10.1016/j.jfineco.2018.07.014.

Arnott, R.D., Kalesnik, V., Linnainmaa, J.T., 2023. Factor momentum. Rev. Financ. Stud. 36 (8), 3034–3070. http://dx.doi.org/10.1093/rfs/hhad013.

Bae, K.H., El Ghoul, S., Gong, Z.J., Guedhami, O., 2021. Does CSR matter in times of crisis? Evidence from the COVID-19 pandemic. J. Corp. Financ. 67, 101876. http://dx.doi.org/10.1016/j.jcorpfin.2020.101876.

van Bekkum, S., 2016. Ireland's 2010 EU/IMF intervention: Costs and benefits. J. Bank. Financ. 72, 175–183. http://dx.doi.org/10.1016/j.jbankfin.2016.06.012.

Berkowitz, D., Lin, C., Ma, Y., 2015. Do property rights matter? Evidence from a property law enactment. J. Financ. Econ. 116 (3), 583–593. http://dx.doi.org/10.1016/j.jfineco.2015.04.003.

Bernard, V.L., 1987. Cross-sectional dependence and problems in inference in market-based accounting research. J. Acc. Res. 25 (1), 1–48. http://dx.doi.org/10.2307/2491260.

Bertrand, M., Duflo, E., Mullainathan, S., 2004. How much should we trust differences-in-differences estimates? Q. J. Econ. 119 (1), 249–275. http://dx.doi.org/10.1162/003355304772839588.

Beuselinck, C., Cao, L., Deloof, M., Xia, X., 2017. The value of government ownership during the global financial crisis. J. Corp. Financ. 42, 481–493. http://dx.doi.org/10.1016/j.jcorpfin.2015.05.002.

Bhojraj, S., Sengupta, P., Zhang, S., 2017. Takeover defenses: Entrenchment and efficiency. J. Acc. Econ. 63 (1), 142–160. http://dx.doi.org/10.1016/j.jacceco.2016.12.002.

Bizjak, J.M., Kalpathy, S.L., Mihov, V.T., Ren, J., 2022. CEO political leanings and store-level economic activity during the Covid-19 crisis: effects on shareholder value and public health. J. Financ. 77 (5), 2949–2986. http://dx.doi.org/10.1111/jofi.13173.

Black, B.S., Kim, W., Jang, H., Park, K.S., 2015. How corporate governance affect firm value? Evidence on a self-dealing channel from a natural experiment in Korea. J. Bank. Financ. 51, 131–150. http://dx.doi.org/10.1016/j.jbankfin.2014.10.002.

Bongini, P., Nieri, L., Pelagatti, M., 2015. The importance of being systemically important financial institutions. J. Bank. Financ. 50, 562–574. http://dx.doi.org/10.1016/j.jbankfin.2014.09.006.

Borisov, A., Goldman, E., Gupta, N., 2016. The corporate value of (corrupt) lobbying. Rev. Financ. Stud. 29 (4), 1039–1071. http://dx.doi.org/10.1093/rfs/hhv048.

Brav, A., Geczy, C., Gompers, P.A., 2000. Is the abnormal return following equity issuances anomalous? J. Financ. Econ. 56 (2), 209–249. http://dx.doi.org/10.1016/S0304-405X(00)00040-4.

Brown, J.R., Huang, J., 2020. All the president's friends: Political access and firm value. J. Financ. Econ. 138 (2), 415–431. http://dx.doi.org/10.1016/j.jfineco.2020.05.004.

Cameron, A.C., Miller, D.L., 2015. A practitioner's guide to cluster-robust inference. J. Hum. Resour. 50 (2), 317–372. http://dx.doi.org/10.3368/jhr.50.2.317.

Carboni, M., Fiordelisi, F., Ricci, O., Lopes, F.S.S., 2017. Surprised or not surprised? The investors' reaction to the comprehensive assessment preceding the launch of the banking union. J. Bank. Financ. 74, 122–132. http://dx.doi.org/10.1016/j.jbankfin.2016.10.005.

Chan, M.K., Kwok, S., 2017. Risk-sharing, market imperfections, asset prices: Evidence from China's stock market liberalization. J. Bank. Financ. 84, 166–187. http://dx.doi.org/10.1016/j.jbankfin.2017.06.003.

Chandra, R., Balachandran, B.V., 1992. More powerful portfolio approaches to regressing abnormal returns on firm-specific variables for cross-sectional studies. J. Financ. 47 (5), 2055–2070. http://dx.doi.org/10.1111/j.1540-6261.1992.tb04678.x.

Chang, E.C., Luo, Y., Ren, J., 2014. Short-selling, margin-trading, and price efficiency: Evidence from the Chinese market. J. Bank. Financ. 48, 411–424. http://dx.doi.org/10.1016/j.jbankfin.2014.02.014.

Chen, J., Garel, A., Tourani-Rad, A., 2019. The value of academics: Evidence from academic independent director resignations in China. J. Corp. Financ. 58, 393–414. http://dx.doi.org/10.1016/j.jcorpfin.2019.06.003.

Chen, Y.C., Hung, M., Wang, Y., 2018. The effect of mandatory CSR disclosure on firm profitability and social externalities: Evidence from China. J. Acc. Econ. 65 (1), 169–190. http://dx.doi.org/10.1016/j.jacceco.2017.11.002.

Chen, A.Y., Zimmermann, T., et al., 2022. Open source cross-sectional asset pricing. Crit. Financ. Rev. 11 (2), 207–264. http://dx.doi.org/10.1561/104.00000057.

Child, T.B., Massoud, N., Schabus, M., Zhou, Y., 2021. Surprise election for Trump connections. J. Financ. Econ. 140 (2), 676–697. http://dx.doi.org/10.1016/j.jfineco.2021.04.009.

Choy, S.K., Lai, T.K., Ng, T., 2017. Do tax havens create firm value? J. Corp. Financ. 42, 198–220. http://dx.doi.org/10.1016/j.jcorpfin.2016.10.016.

Chung, C.Y., Kim, I., Rabarison, M.K., To, T.Y., Wu, E., 2020. Shareholder litigation rights and corporate acquisitions. J. Corp. Financ. 62, 101599. http://dx.doi.org/10.1016/j.jcorpfin.2020.101599.

Clarke, R.G., De Silva, H., Thorley, S., 2006. Minimum-variance portfolios in the US equity market. J. Portf. Manag. 33 (1), 10–24. http://dx.doi.org/10.3905/jpm.2006.661366.

Cohn, J.B., Gillan, S.L., Hartzell, J.C., 2016. On enhancing shareholder control: A (Dodd-) Frank assessment of proxy access. J. Financ. 71 (4), 1623–1668. http://dx.doi.org/10.1111/jofi.12402.

Collins, D.W., Dent, W.T., 1984. A comparison of alternative testing methodologies used in capital market research. J. Acc. Res. 22 (1), 48–84. http://dx.doi.org/10.2307/2490702.

Cousins, P., Dutordoir, M., Lawson, B., Neto, J.Q.F., 2020. Shareholder wealth effects of modern slavery regulation. Manag. Sci. 66 (11), 5265–5289. http://dx.doi.org/10.1287/mnsc.2019.3456.

Doidge, C., Dyck, A., 2015. Taxes and corporate policies: Evidence from a quasi natural experiment. J. Financ. 70 (1), 45–89. http://dx.doi.org/10.1111/jofi.12101.

Eisfeldt, A.L., Schubert, G., Zhang, M.B., Taska, B., 2025. Generative AI and firm values. Working paper.

Fahlenbrach, R., Rageth, K., Stulz, R.M., 2021. How valuable is financial flexibility when revenue stops? Evidence from the COVID-19 crisis. Rev. Financ. Stud. 34 (11), 5474–5521. http://dx.doi.org/10.1093/rfs/hhab059.

Fama, E.F., MacBeth, J.D., 1973. Risk, return, and equilibrium: Empirical tests. J. Polit. Econ. 81 (3), 607–636. http://dx.doi.org/10.1086/260061.

Fang, Y., Fiordelisi, F., Hasan, I., Leung, W.S., Wong, G., 2023. Corporate culture and firm value: Evidence from crisis. J. Bank. Financ. 146, 106710. http://dx.doi.org/10.1016/j.jbankfin.2022.106710.

Fisman, R., Hamao, Y., Wang, Y., 2014. Nationalism and economic exchange: Evidence from shocks to Sino-Japanese relations. Rev. Financ. Stud. 27 (9), 2626–2660. http://dx.doi.org/10.1093/rfs/hhu027.

Foley-Fisher, N., Ramcharan, R., Yu, E., 2016. The impact of unconventional monetary policy on firm financing constraints: Evidence from the maturity extension program. J. Financ. Econ. 122 (2), 409–429. http://dx.doi.org/10.1016/j.jfineco.2016.07.002.

Frattaroli, M., 2020. Does protectionist anti-takeover legislation lead to managerial entrenchment? J. Financ. Econ. 136 (1), 106–136. http://dx.doi.org/10.1016/j.jfineco.2019.03.014.

Gao, Y., Liao, S., Wang, X., 2018. Capital markets' assessment of the economic impact of the Dodd–Frank Act on systemically important financial firms. J. Bank. Financ. 86, 204–223. http://dx.doi.org/10.1016/j.jbankfin.2016.03.016.

Garel, A., Petit-Romec, A., 2021. Investor rewards to environmental responsibility: Evidence from the COVID-19 crisis. J. Corp. Financ. 68, 101948. http://dx.doi.org/10.1016/j.jcorpfin.2021.101948.

Giglio, S., Xiu, D., 2021. Asset pricing with omitted factors. J. Polit. Econ. 129 (7), 1947–1990. http://dx.doi.org/10.1086/714090.

Gilje, E.P., Taillard, J.P., 2017. Does hedging affect firm value? Evidence from a natural experiment. Rev. Financ. Stud. 30 (12), 4083–4132. http://dx.doi.org/10.1093/rfs/hhx069.

Gómez-Cram, R., Olbert, M., 2023. Measuring the expected effects of the global tax reform. Rev. Financ. Stud. 36 (12), 4965–5011. http://dx.doi.org/10.1093/rfs/hhad038.

Greene, D., Intintoli, V.J., Kahle, K.M., 2020. Do board gender quotas affect firm value? Evidence from California Senate Bill No. 826. J. Corp. Financ. 60, 101526. http://dx.doi.org/10.1016/j.jcorpfin.2019.101526.

Grewal, J., Riedl, E.J., Serafeim, G., 2019. Market reaction to mandatory nonfinancial disclosure. Manag. Sci. 65 (7), 3061–3084. http://dx.doi.org/10.1287/mnsc.2018.3099.

Gupta, T., Kelly, B., 2019. Factor momentum everywhere. J. Portf. Manag. 45 (3), 13–36. http://dx.doi.org/10.3905/jpm.2019.1.066.

Hackbarth, D., Haselmann, R., Schoenherr, D., 2015. Financial distress, stock returns, and the 1978 bankruptcy reform act. Rev. Financ. Stud. 28 (6), 1810–1847. http://dx.doi.org/10.1093/rfs/hhv009.

Hill, P., Korczak, A., Korczak, P., 2019. Political uncertainty exposure of individual companies: The case of the Brexit referendum. J. Bank. Financ. 100, 58–76. http://dx.doi.org/10.1016/j.jbankfin.2018.11.014.

Huang, H., Lee, E., Lyu, C., Zhao, Y., 2020. Bequest motive, information transparency, and family firm value: A natural experiment. J. Corp. Financ. 65, 101751. http://dx.doi.org/10.1016/j.jcorpfin.2020.101751.

Igan, D., Mirzaei, A., Moore, T., 2023. Does macroprudential policy alleviate the adverse impact of COVID-19 on the resilience of banks? J. Bank. Financ. 147, 106419. http://dx.doi.org/10.1016/j.jbankfin.2022.106419.

Jain, A., Tantri, P., Thirumalai, R.S., 2019. Demand curves for stocks do not slope down: Evidence using an exogenous supply shock. J. Bank. Financ. 104, 19–30. http://dx.doi.org/10.1016/j.jbankfin.2019.03.012.

Jegadeesh, N., Karceski, J., 2009. Long-run performance evaluation: Correlation and heteroskedasticity-consistent tests. J. Empir. Financ. 16 (1), 101–111. http://dx.doi.org/10.1016/j.jempfin.2008.11.001.

Jiang, H., Parsons, C., Sun, L., Titman, S., 2025. Skilled foreign labor, urban agglomeration, and value creation. Available at SSRN 4451552.

Kalcheva, I., Plefçnik, J.M., Tran, H., Turkiela, J., 2020. (Un)intended consequences? The impact of the 2017 tax cuts and jobs act on shareholder wealth. J. Bank. Financ. 118, 105860. http://dx.doi.org/10.1016/j.jbankfin.2020.105860.

Kim, S., Kim, W., 2022. Tunneling through trademarks. J. Corp. Financ. 76, 102274. http://dx.doi.org/10.1016/j.jcorpfin.2022.102274.

Kolari, J.W., Pynnönen, S., 2010. Event study testing with cross-sectional correlation of abnormal returns. Rev. Financ. Stud. 23 (11), 3996–4025. http://dx.doi.org/10.1093/rfs/hhq072.

Kolasinski, A., Yang, N., 2023. When myopic managers must mark to market. Manag. Sci. 6234–6254. http://dx.doi.org/10.1287/mnsc.2020.03249.

Kothari, S.P., Warner, J.B., 2007. Econometrics of event studies. In: Handbook of Empirical Corporate Finance. Elsevier, pp. 3–36. http://dx.doi.org/10.1016/B978-0-444-53265-7.50004-9.

Kundu, S., 2023. Impact of regulations on firm value: Evidence from the 2016 U.S. presidential election. J. Financ. Quant. Anal. 59 (4), 1659–1691. http://dx.doi.org/10.1017/s0022109023000510.

Leftwich, R., 1981. Evidence of the impact of mandatory changes in accounting principles on corporate loan agreements. J. Acc. Econ. 3 (1), 3–36. http://dx.doi.org/10.1016/0165-4101(81)90008-4.

Lewis, C., Verwijmeren, P., 2014. Cash-settled convertible bonds and the value relevance of their accounting treatment. J. Corp. Financ. 24, 101–111. http://dx.doi.org/10.1016/j.jcorpfin.2013.11.003.

Licht, A.N., Poliquin, C., Siegel, J.I., Li, X., 2018. What makes the bonding stick? A natural experiment testing the legal bonding hypothesis. J. Financ. Econ. 129 (2), 329–356. http://dx.doi.org/10.1016/j.jfineco.2018.05.001.

Lin, C., Officer, M.S., Schmid, T., Zou, H., 2019. Is skin in the game a game changer? Evidence from mandatory changes of D&o insurance policies. J. Acc. Econ. 68 (1), 101225. http://dx.doi.org/10.1016/j.jacceco.2019.101225.

Lins, K.V., Servaes, H., Tamayo, A., 2017. Social capital, trust, and firm performance: The value of corporate social responsibility during the financial crisis. J. Financ. 72 (4), 1785–1824. http://dx.doi.org/10.1111/jofi.12505.

Liu, X.S., Megginson, W.L., Xia, J., 2022. Industrial policy and asset prices: Evidence from the Made in China 2025 policy. J. Bank. Financ. 142, 106554. http://dx.doi.org/10.1016/j.jbankfin.2022.106554.

Liu, L.X., Shu, H., Wei, K.J., 2017. The impacts of political uncertainty on asset prices: Evidence from the Bo scandal in China. J. Financ. Econ. 125 (2), 286–310. http://dx.doi.org/10.1016/j.jfineco.2017.05.011.

Liu, C., Wang, S., Wei, K.J., 2021. Demand shock, speculative beta, and asset prices: Evidence from the Shanghai-Hong Kong stock connect program. J. Bank. Financ. 126, 106102. http://dx.doi.org/10.1016/j.jbankfin.2020.106102.

Loipersberger, F., 2018. The effect of supranational banking supervision on the financial sector: Event study evidence from Europe. J. Bank. Financ. 91, 34–48. http://dx.doi.org/10.1016/j.jbankfin.2018.03.002.

Lopez-Lira, A., Roussanov, N.L., 2023. Do common factors really explain the cross-section of stock returns? Working paper.

Lu, R., Ye, Z., 2025. Roe v. Rates: Reproductive healthcare and public financing costs. Working paper.

Lyon, J.D., Barber, B.M., Tsai, C.L., 1999. Improved methods for tests of long-run abnormal stock returns. J. Financ. 54 (1), 165–201. http://dx.doi.org/10.1111/0022-1082.00122.

Marsh, W.B., 2023. Supervisory stringency, payout restrictions, and bank equity prices. J. Bank. Financ. 154, 106936. http://dx.doi.org/10.1016/j.jbankfin.2023.106936.

May, A.D., 2014. Corporate liquidity and the contingent nature of bank credit lines: Evidence on the costs and consequences of bank default. J. Corp. Financ. 29, 410–429. http://dx.doi.org/10.1016/j.jcorpfin.2014.10.001.

Mitchell, M.L., Stafford, E., 2000. Managerial decisions and long-term stock price performance. J. Bus. 73 (3), 287–329. http://dx.doi.org/10.1086/209645.

Moon, T., Schoenherr, D., 2022. The rise of a network: Spillover of political patronage and cronyism to the private sector. J. Financ. Econ. 145 (3), 970–1005. http://dx.doi.org/10.1016/j.jfineco.2022.07.004.

Moulton, B.R., 1986. Random group effects and the precision of regression estimates. J. Econ. 32 (3), 385–397. http://dx.doi.org/10.1016/0304-4076(86)90021-7.

Moulton, B.R., 1987. Diagnostics for group effects in regression analysis. J. Bus. Econom. Statist. 5 (2), 275–282. http://dx.doi.org/10.1080/07350015.1987.10509506.

Nesbitt, W.L., Outslay, E., Persson, A.V., 2023. A reexamination of investors' reaction to tax shelter news: Evidence from the Luxembourg tax leaks. J. Acc. Econ. 75 (2–3), 101537. http://dx.doi.org/10.1016/j.jacceco.2022.101537.

Neukirchen, D., Engelhardt, N., Krause, M., Posch, P.N., 2023. The value of (private) investor relations during the COVID-19 crisis. J. Bank. Financ. 147, 106450. http://dx.doi.org/10.1016/j.jbankfin.2022.106450.

Nguyen, J.H., Phan, H.V., 2020. Carbon risk and corporate capital structure. J. Corp. Financ. 64, 101713. http://dx.doi.org/10.1016/j.jcorpfin.2020.101713.

O'Donovan, J., Wagner, H.F., Zeume, S., 2019. The value of offshore secrets: Evidence from the Panama Papers. Rev. Financ. Stud. 32 (11), 4117–4155. http://dx.doi.org/10.1093/rfs/hhz017.

Pagano, M., Wagner, C., Zechner, J., 2023. Disaster resilience and asset prices. J. Financ. Econ. 150 (2), 103712. http://dx.doi.org/10.1016/j.jfineco.2023.103712.

Petersen, M.A., 2009. Estimating standard errors in finance panel data sets: Comparing approaches. Rev. Financ. Stud. 22 (1), 435–480. http://dx.doi.org/10.1093/rfs/hhn053.

Ramelli, S., Ossola, E., Rancan, M., 2021a. Stock price effects of climate activism: Evidence from the first Global Climate Strike. J. Corp. Financ. 69, 102018. http://dx.doi.org/10.1016/j.jcorpfin.2021.102018.

Ramelli, S., Wagner, A.F., Zeckhauser, R.J., Ziegler, A., 2021b. Investor rewards to climate responsibility: Stock-price responses to the opposite shocks of the 2016 and 2020 U.S. elections. Rev. Corp. Financ. Stud. 10 (4), 748–787. http://dx.doi.org/10.1093/rcfs/cfab010.

Raykov, R., Silva-Buston, C., 2020. Holding company affiliation and bank stability: Evidence from the US banking sector. J. Corp. Financ. 65, 101739. http://dx.doi.org/10.1016/j.jcorpfin.2020.101739.

Schoenherr, D., 2019. Political connections and allocative distortions. J. Financ. 74 (2), 543–586. http://dx.doi.org/10.1111/jofi.12818.

Sefcik, S.E., Thompson, R., 1986. An approach to statistical inference in cross-sectional models with security abnormal returns as dependent variable. J. Acc. Res. 24 (2), 316–334. http://dx.doi.org/10.2307/2490725.

Shan, C., Tang, D.Y., 2023. The value of employee satisfaction in disastrous times: Evidence from COVID-19. Rev. Financ. 27 (3), 1027–1076. http://dx.doi.org/10.1093/rof/rfac055.

Stahl, J.R., 2023. Precious neighbors: The value of co-locating with the government. Rev. Financ. 27 (4), 1269–1296. http://dx.doi.org/10.1093/rof/rfac063.

Thompson, S.B., 2011. Simple formulas for standard errors that cluster by both firm and time. J. Financ. Econ. 99 (1), 1–10. http://dx.doi.org/10.1016/j.jfineco.2010.08.016.

Wagner, A.F., Zeckhauser, R.J., Ziegler, A., 2018. Company stock price reactions to the 2016 election shock: Trump, taxes, and trade. J. Financ. Econ. 130 (2), 428–451. http://dx.doi.org/10.1016/j.jfineco.2018.06.013.

Wang, Y.C., Chou, R.K., 2018. The impact of share pledging regulations on stock trading and firm valuation. J. Bank. Financ. 89, 1–13. http://dx.doi.org/10.1016/j.jbankfin.2018.01.016.

White, H., 1980. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. Econometrica 48 (4), 817–838. http://dx.doi.org/10.2307/1912934.

Wilson, J.I., 2020. The consequences of limiting shareholder litigation: Evidence from exclusive forum provisions. J. Corp. Financ. 64, 101712. http://dx.doi.org/10.1016/j.jcorpfin.2020.101712.

Zeume, S., 2017. Bribes and firm value. Rev. Financ. Stud. 30 (5), 1457–1489. http://dx.doi.org/10.1093/rfs/hhw108.

Zhang, Y., 2015. The securitization of gold and its potential impact on gold stocks. J. Bank. Financ. 58, 309–326. http://dx.doi.org/10.1016/j.jbankfin.2015.04.015.