Abstract

The important role event studies play in securities fraud litigation will only increase in the wake of Halliburton II, as proving price impact has become a virtual requirement to successfully invoke the Basic presumption of reliance and secure class certification. This Article explores an array of considerations related to the use of event studies in such litigation. These considerations span the space of substantive securities law; civil procedure; financial economics; and statistical methodology. A central thesis of the statistical discussion is that the proper use of event studies will require experts and courts to pay attention to a number of important ways in which the securities fraud litigation context differs from the empirical context of many academic event studies.

One set of issues has to do with the problem of disentangling the effects of multiple pieces of news released on the same date or, worse still, at the exact same moment. Another set of issues involves the question of whether an event study could even be probative as to the date of an alleged misrepresentation, when the alleged effect is to confirm rather than improve market expectations.
Another important point of this Article is that courts and experts using event studies in securities fraud litigation should pay more attention to four issues related to statistical methodology. We use the example of six dates in the Halliburton litigation to illustrate these points:

- First, because a litigation-relevant event study typically involves only a single firm, important issues related to non-normality of a stock’s returns arise.
- Second, because the plaintiff must show either that price dropped or rose, but will never carry its burden if the opposite happened, experts should unquestionably be using one-sided significance testing rather than the conventionally deployed two-sided approach.
- Third, securities fraud litigation often involves multiple test dates, which has important and tricky implications for the appropriate level of date-specific confidence levels if the goal is an overall confidence level equal to the 95% level that courts and experts say it is.
- Finally, event studies must be modified appropriately to account for the possibility that stock price volatility varies across time.

Failure to address these issues should render testimony concerning event studies inadmissible as a matter of law under Daubert and Rule 702. Happily, each of these issues can be addressed using straightforward and simple techniques that do not increase the difficulty of conducting event studies. We explain these techniques and illustrate their use with our Halliburton application.
# Table of Contents

Introduction .................................................................................................................. 1

I. Event Studies’ Role in the Substance and Procedure of Securities Litigation .................. 4
   A. Event Studies and the Merits .............................................................. 4
      1. Reliance ......................................................................................... 5
      2. Materiality ..................................................................................... 9
      3. Loss causation, economic loss, and damages .............................. 9
      4. Summing up the Role of Event Studies vis-à-vis Merits Elements ........................................................................... 12
   B. Event Studies and Pleading .................................................................... 15
   C. Event Studies and Class Certification .................................................. 16
      1. Implications of Halliburton I for class certification ................. 17
      2. Amgen and class certification .......................................................... 18
      3. Halliburton strikes back: Halliburton II and class certification 21
      4. Developments related to class certification since Halliburton II 24
   D. Summing Up: Substance, Procedure, Economics, and Event Studies in Securities Fraud Litigation ........................................................................ 27

II. The Theory of Financial Economics and the Practice of Event Studies: A View From 10,000 Feet ........................................................................ 28

III. Using the *Halliburton* Litigation to Illustrate Event Study Practice, and Problems .......................................................................... 35
   A. The Standard Approach to Testing for Statistical Significance in An Event Study, as Applied to Our Halliburton Results ....................... 40
   B. Violations of the Four Statistical Assumptions Undergirding the Standard Approach to Statistical Significance Testing 45
      1. Non-normality in Excess Returns ..................................................... 45
      2. The Inappropriateness of Two-Sided Tests (With or Without the Normality Assumption) ............................................................ 52
      3. Multiple event dates of interest ...................................................... 58
4. Instability Over Time in the Distribution of the Excess Return

C. Best Practices for Experts Conducting Reliable Event Studies for Use in Securities Fraud Litigation ........................................... 70

D. Dealing with Multiple Pieces of News on an Event Date .... 71

E. Power in Event Studies Used in Securities Fraud Litigation ... 74

Conclusion ........................................................................................................ 74

Appendix A: Allowing for Heterogeneity in the Distribution of Excess Returns ....................................................................... 76

A. A Model that Allows for Heterogeneity in the Variance of Excess Returns ........................................................................... 76

B. Estimation Results .................................................................................. 78

C. Statistical Significance Testing Methodology ............................ 81
INTRODUCTION

In June, 2014, on its second trip to the US Supreme Court in connection with its defense against allegations of federal securities fraud, Halliburton scored a partial victory. Halliburton failed to persuade the Supreme Court to overrule Basic, Inc. v. Levinson, which had approved the fraud-on-the-market (FOTM) presumption of reliance in private securities fraud litigation. But Halliburton did persuade the Court to allow defendants to introduce evidence of lack of price impact at class certification. As the Court explained, Basic “does not require courts to ignore a defendant’s direct . . . salient evidence showing that the alleged misrepresentation did not actually affect the stock’s market price and, consequently, that the Basic presumption does not apply.”\(^1\) A critical component of the Court’s reasoning was the perception that defendants could readily marshal such direct evidence in the form of event studies.\(^2\)

Prior to Halliburton II, litigants in securities fraud cases had used event studies in two ways. The first was to address the requirement that plaintiffs prove loss causation established by the Court’s decision in Dura Pharmaceuticals.\(^3\) The second was to demonstrate market efficiency in order to obtain Basic’s FOTM presumption.\(^4\) Halliburton II opened the door to yet a third role for event studies in federal securities fraud litigation – to prove lack of price impact in order to defeat class certification.

As a result, event studies have become increasingly important in federal securities fraud litigation. Yet judicial reliance on event studies entails a number of challenges. For one thing, it is unclear that courts

\(^1\) Halliburton II, 134 S. Ct. at 2416

\(^2\) See slip op. at 19-20 (describing event study looking at the impact on stock price from the misrepresentation as direct evidence capable of “show[ing] no price impact with respect to the specific misrepresentation challenged in the suit”).

\(^3\) See In re Vivendi Universal, S.A. Securities Litigation, No. 02 Civ. 5571, 2009 WL 1066254, 10 (S.D.N.Y. April 21, 2009), (describing the use of an event study to establish loss causation as “almost obligatory.”).

\(^4\) See, e.g., Halliburton II, slip op. at 19 (explaining use of event studies to demonstrate market efficiency); Teamsters Local 445 Freight Div. Pension Fund v. Bombardier Inc., 546 F.3d 196, 207 (2d Cir. 2008) (explaining that the fifth Cammer factor, which requires evidence tending to demonstrate that unexpected corporate events or financial releases cause an immediate response in the price of a security, is the most important indicator of market efficiency).
fully understand the event study methodology that is being employed in these cases and, in particular, the limits of that methodology. For example, Justice Alito asked counsel for the petitioner at oral argument in Halliburton II: “Can I ask you a question about these event studies to which you referred. How accurately can they distinguish between the effect on price of the facts contained in a disclosure and an irrational reaction by the market, at least temporarily, to the facts contained in the disclosure?” By themselves, event studies can do no more than demonstrate significant price changes, over and above what otherwise would have been expected given market conditions, on dates when information is released to the public. And event studies cannot determine the rationality of that reaction. Yet counsel responded to the Court that: “Event studies are very effective at making that sort of determination.”

Second, event studies plausibly measure only one type of price impact – the movement of a stock price to the release of unanticipated material information. In circumstances in which fraudulent statements falsely confirm prior statements, price stability rather than a price change would constitute evidence of price impact, yet event studies have at most very limited capacity to test for price stability.5 Similarly, in cases involving multiple “bundled” disclosures, event studies have limited capacity to identify the relative contribution of each piece of information or the degree to which the effects of multiple disclosures may offset each other.6

Third, there are some statistically important differences between the empirical context of securities fraud litigation and the scholarly contexts for which event studies were originally designed. For example, event study methodology was originally designed to measure the significance of a single event across multiple firms, the significance of

---

5 As we discuss in a number of places below, courts have responded to this limitation by allowing plaintiffs to show price impact indirectly via event studies that show a price drop on the date of an alleged corrective disclosure.

6 This sort of problem, which we discuss below, has arisen in cases; see, e.g., [Oscar or Green; check 2008 Halliburton class cert denial order ___. Cf. Alex Rinaudo & Atanu Saha, An intraday event study methodology for determining loss causation, 2 J. Fin. Persp. 161 (2014) (explaining how problem of multiple disclosures can be partially addressed by using an intraday event methodology); Esther Bruegger & Frederick C. Dunbar, Estimating Financial Fraud Damages with Response Coefficients, 35 J. Corp. L. 11, 25 (2009) (explaining that “content analysis’ is now part of the tool kit for determining which among a number of simultaneous news events had effects on the stock price.”).
multiple events at a single firm, or the significance of multiple events at multiple firms. By contrast, event studies used to establish price impact and loss causation in securities fraud litigation attempt to measure the impact of a single event at a single firm. This application has important methodological implications for the proper use of event studies.

Event studies now have a potential role in numerous aspects of securities litigation and at multiple procedural points. Understanding the scope for their use, and their proper deployment, implicates an array of considerations arising from substantive securities law; civil procedure; the theory of financial economics; the theory of statistics; and the application of statistical theory to stock behavior in particular situations that are distinct from applications typical of academic scholarship on event studies. Many articles and judicial opinions have been written about the state of substantive securities litigation; procedure as it relates to class action (and other) litigation; and the theory and practice of event studies. However, we are aware of no work that has sought to treat these different aspects of securities litigation holistically in a way that will provide scholars, practitioners, and judges an outlet for one-stop-shopping concerning the range of major issues. A key objective of this Article is to fill that gap by discussing the interaction of substantive law, procedure, and applied statistics together in one place. We believe this contribution will be especially valuable because of the very different types of expertise necessary to navigate all these issues.

Part I of this Article explores the substantive elements and procedural milestones of securities fraud litigation, with a focus on the role event studies play at these key moments. Part II provides a bird’s eye view of how financial economics theory, via the efficient market hypothesis, suggests that event studies can provide probative evidence regarding the effects of alleged misrepresentations on stock price. Part III conducts a stylized event study using data related to the ongoing Halliburton litigation. An important focus of this Part is on several assumptions that are critical to justify the use of standard event study methodology in securities fraud litigation. As we discuss, these assumptions are often violated, which renders standard event study methodology unreliable and therefore inadmissible under Rule 702 and Daubert. We also discuss simple improvements to event study methodology that cure these problems. Courts should insist that experts employ these cures.
I. EVENT STUDIES’ ROLE IN THE SUBSTANCE AND PROCEDURE OF SECURITIES LITIGATION

In this Part we take a systematic look at the different questions that event studies might answer in a securities fraud case. We begin by imagining a case has reached the merits stage. We then consider the pleading stage. Finally, we address the procedural moment that has been the focus of several recent Supreme Court cases, class certification. Our focus in this Part is on the interlocking relationship between concerns of substantive securities law, procedure, and financial economics; we leave issues of statistical methodology for Part II.

A. Event Studies and the Merits

As the Supreme Court explained in Dura, a securities fraud action has six basic elements. First, there must have been “a material misrepresentation (or omission).” Second, there must have been “scienter, i.e., a wrongful state of mind.” Third, the material misrepresentation or omission must have occurred in “connection with the purchase or sale of a security.” Fourth, plaintiff purchasers must have undertaken reliance, which is “often referred to … as transaction causation.” Fifth, there must have been some economic loss on the part of plaintiffs. And sixth, there must be loss causation, “i.e., a causal connection between the material misrepresentation and the loss.”

In addition, damages must be calculated in the event that plaintiffs win a judgment. For those class action cases that are settled, as the overwhelming majority of cases that aren’t dismissed will be, settlement brings oversight by the trial court pursuant to Rule 23(e)(2)’s requirement that the court hold a hearing and find that the settlement is

---

8 Id. at 341.
9 Id. at 341.
10 Id. at 341-342 (describing the “nonconclusive[]” presumption that “the price of a publicly traded share reflects a material misrepresentation and that plaintiffs have relied upon that misrepresentation as long as they would not have bought the share in its absence”) (citing Basic Inc. v. Levinson, 108 S.Ct. 978).
11 Id. at 341 (internal quotation marks omitted).
12 Id. at 342 (citing 15 U.S.C. § 78u–4(b)(4)).
13 Id.
“fair, reasonable, and adequate.” Such a finding would usually require the court to consider the damages that plaintiffs might win in the event of a trial, so evidence concerning damage determination is broadly relevant to settlement.

Neither the scienter nor the in-connection-with-purchase-or-sale element is likely to involve an event study. For example, proof of the knowledge of falsity is proof of scienter, and usually will require direct evidence of knowledge and intent. For purposes of discussion, we will simply assume that a plaintiff can meet the relevant pleading standard and can meet the burden of production, so that scienter will be a question for the jury. We will also assume the plaintiff can prove it bought stock between an alleged misrepresentation and an alleged corrective disclosure date, meeting the element requiring that the fraud in question be in connection with the purchase or sale of securities.

Event studies as deployed in securities litigation can address the other elements, and damages, in overlapping ways, as we now discuss.

1. **Reliance**

Under traditional rules, each plaintiff would have to show transaction causation, i.e., that it would not have purchased the stock but for the misrepresentation. In modern securities fraud litigation, that would be prohibitively difficult to do, since most purchasers don’t actually follow all news related to stocks they buy. However, the *Basic* fraud on the market theory, when it is operative, bestows on plaintiffs the presumption of reliance. The *Basic* Court determined that reliance could be presumed if (i) the misrepresentations were publicly known; (ii) the misrepresentations were material; (iii) the stock traded in an efficient market; and (iv) the plaintiff traded between the time the misrepresentations were made and when the truth was revealed. *Basic* reasoned that, in an efficient market, information is incorporated into stock prices and that the typical investor relies on the integrity of the market price. Importantly, the *Basic* presumption was rebuttable.

---

15 See 15 U.S.C. 78u-4(b)(2)(A) (“the complaint shall, with respect to each act or omission alleged to violate this chapter, state with particularity facts giving rise to a strong inference that the defendant acted with the required state of mind”).
16 *Halliburton II*, slip op. at 6-7.
Following *Basic*, the use of class actions to litigate claims of federal securities fraud expanded greatly. As noted, one of the necessary predicates to obtain the presumption of reliance is proof that the stock traded in an efficient market. Early cases such as *Cammer v. Bloom* described the market efficiency inquiry in terms of a multi-factor test. The *Cammer* test involved five factors: “(1) the stock’s average weekly trading volume; (2) the number of securities analysts that followed and reported on the stock; (3) the presence of market makers and arbitrageurs; (4) the company’s eligibility to file a Form S-3 Registration Statement; and (5) a cause-and-effect relationship, over time, between unexpected corporate events or financial releases and an immediate response in stock price.” Economists serving as expert witnesses generally use event studies to address the fifth *Cammer* factor. In this context, the event study is used to determine the extent to which the market for a particular stock responds to new information. Experts will generally look at multiple information or news events—some relevant to the litigation in question and some not—for a stock and evaluate the extent to which these events are associated with price changes in the expected directions. Interestingly, though, financial economists virtually never use the *Cammer* factors in academic scholarship.

Figure 1 illustrates the connection between the efficient market hypothesis and the presumption of reliance. When a firm makes a representation, be it an affirmative statement or omission, it will either affect the perceived value of the firm or not. Suppose it does affect the

---

17 See, e.g., Paul G. Mahoney, Precaution Costs and the Law of Fraud in Impersonal Markets, 78 VA. L. REV. 623, 663 (1992) (stating that “the rate at which securities fraud class action suits were filed nearly tripled between April 1988, just after *Basic* was decided, and June 1991”). We discuss the critical role the *Basic* presumption of reliance plays in class certification analysis in Part I.C, infra.


19 See also Teamsters Local 445 Freight Div. Pension Fund v. Bombardier Inc., 546 F.3d 196, 207 (2d Cir. 2008) (explaining that the fifth *Cammer* factor, which requires evidence tending to demonstrate that unexpected corporate events or financial releases cause an immediate response in the price of a security, is the most important indicator of market efficiency). But see Do Courts Count *Cammer* Factors? (suggesting that courts simply count the five *Cammer* factors).

firm’s perceived value, so that we move to the top arm of the figure. Under the efficient market hypothesis, market perceptions are quickly incorporated into market prices. Consequently, the change in the firm’s perceived value flowing from the firm’s representation will lead to a change in its stock price—i.e., the representation will have a price impact. The role of the efficient market hypothesis is indicated in Figure 1 using dashed arrows. On the other hand, if the representation does not affect the firm’s perceived value, then there is no reason for the market to react, and there will be no change in the firm’s stock price. Accordingly, representations that do not change market participants’ perceptions of firm value will have no price impact.

Finally, a purchaser of stock operating in an efficient market will reasonably believe that the market price incorporates all relevant public information. Thus, under the fraud on the market theory purchasers are presumed to have relied—albeit indirectly—on any representation that makes its way into market price.21

The presumption is not absolute, however. In *Halliburton II*, the Supreme Court enumerated two ways that a defendant could rebut the presumption. If a defendant could show “that a plaintiff would have bought or sold the stock even had he been aware that the stock's price was tainted by fraud,” that would vitiate the presumption.22 Such an individualized showing seems unlikely in typical cases. More relevant for our purposes, a defendant could rebut the presumption by “show[ing] that the alleged misrepresentation did not, for whatever reason, actually affect the market price.”23

---

21 The Supreme Court succinctly summed up the fraud on the market theory in *Halliburton I*, citing to *Basic*: “Because the market ‘transmits information to the investor in the processed form of a market price,’ we can assume … that an investor relies on public misstatements whenever he ‘buys or sells stock at the price set by the market.’” *Erica P. John Fund, Inc. v. Halliburton Co.*, 131 S.Ct. 2179, 2185 (U.S., 2011) (citing *Basic*, at 244, 247).

22 *Halliburton II*, 134 S.Ct. at 2408.

23 *Halliburton II*, 134 S.Ct. at 2408.
Figure 1: How the Efficient Market Hypothesis Supports the Presumption of Reliance

Figure 1 illustrates that when the efficient market hypothesis holds, a lack of price impact is inconsistent with the hypothesis that the firm’s value changed following the representation in question. Accordingly lack of price impact vitiates the logical foundation for the allegation that plaintiffs relied on a misrepresentation: if (1) the extent of plaintiffs’ demonstrated reliance is on the integrity of market price, and (2) the market price did not react to the alleged misrepresentation, then plaintiffs can’t be presumed to have relied on the alleged misrepresentation.

This discussion concerns the dates of alleged misrepresentations. As we have mentioned, misrepresentations that falsely confirm market expectations will not lead to an observable change in price. Note, though, that that does not mean they have no price impact. The relevant price impact is simply counterfactual: price would have fallen had there not been fraud. Price impact ultimately does occur when the fraud is revealed via corrective disclosures. That is why it is appropriate to allow plaintiffs to use event studies concerning dates of alleged corrective disclosure to establish price impact for cases involving confirmatory alleged misrepresentations.

In the Halliburton litigation, the statements of interest were made publicly. The parties did not dispute the efficiency of the market in
Halliburton’s stock. And the complaint focused on purchases made between the alleged misrepresentations and the alleged corrective disclosures. Thus, all these prerequisites for the presumption of reliance were met. All that remained for plaintiffs to establish in order to meet the Basic presumption of reliance, per Halliburton II, then, is proof that alleged misrepresentations had price impact as depicted in Figure 1.

2. Materiality

To prevail on the merits, the plaintiff must establish that alleged misrepresentations were material. Otherwise, we would never get to box (1) in Figure 1, because there would be no reason for market perceptions of the firm’s stock price to change. Thus a failure to show price impact would imply a lack of materiality: price impact is a necessary condition for materiality. A defendant’s showing that there was no price impact on an alleged misrepresentation or alleged corrective disclosure date should thus be dispositive of a case. However, the converse does not hold—a plaintiff’s showing that there was price impact is not dispositive—because an event study cannot tell us why there was an unusually large price change on a particular date. In sum, price impact is a necessary but not a sufficient condition for materiality.

3. Loss causation, economic loss, and damages

To prove loss causation, plaintiffs just establish a causal connection between the alleged fraud and their economic harm. Purchasing at an inflated price might be related to a loss, but the Dura Court held that by itself, such a purchase it is insufficient to prove economic harm. The Dura Court held that loss causation, which had been developed by numerous lower court decisions prior to Basic, generally may be met by showing that the stock price declined in response to a “materialization of the risk” that had been concealed by the fraud. Practically speaking, plaintiffs need to prove both price impact on the alleged corrective disclosure date and a causal link between this price impact and the alleged misrepresentations (with exceptions made in cases where the

---

24 Archdiocese of Milwaukee Supporting Fund, Inc. v. Halliburton Co., 2008 WL 4791492, at *1 (N.D.Tex., 2008) (“The parties do not dispute that there was an efficient market in this case.”).
alleged misrepresentations were confirmatory; see discussion ___, infra).

In practice, plaintiffs have addressed *Dura*’s loss causation requirement by presenting event studies demonstrating a decline in stock price in response to an issuer’s corrective disclosure. As the First Circuit recently explained: “The usual—it is fair to say ‘preferred’—method of proving loss causation in a securities fraud case is through an event study….” Because plaintiffs have the burden of proving loss causation, the methodology and validity of the event studies through which they attempt to establish the causal connection between the fraud and stock price is hotly contested. Many courts, applying *Daubert*, have given careful scrutiny to proffered expert testimony. Indeed, many courts have refused to admit expert testimony with respect to loss causation that did not include an event study.

Proof of economic loss and damages also overlaps proof of loss causation. For plaintiffs to recover damages, they must show that they suffered an economic loss that was caused by the alleged fraud. The 1934 Act provides that plaintiffs may recover actual damages, which must be proved. A plaintiff who can prove damages has obviously proved she sustained an economic loss. At the same time, a plaintiff who

---

25 *Halliburton I* was spawned because the district court had denied class certification on the ground that plaintiffs had failed to persuade the court that there was such a causal link (even though plaintiffs had presented an event study showing there was price impact). Under pre-*Halliburton I* Fifth Circuit precedent, such a showing was necessary at the class certification stage; *Oscar Private Equity Investments v. Allegiance Telecom, Inc.*, 487 F.3d 261, 269 (C.A.5 (Tex.), 2007) (“loss causation must be established at the class certification stage by a preponderance of all admissible evidence”).

26 *Bricklayers & Trowel Trades Int’l Pension Fund v. Credit Suisse Secs.*, 752 F.3d 82, (1st Cir. 2014).


28 See *Bricklayers*, 752 F.3d at 82.

29 See *In re Imperial Credit Industries, Inc. Securities Litigation*, 252 F. Supp. 2d 1005, 1015 (C.D.Cal. 2003) (“Because of the need to distinguish between the fraud-related and non-fraud related influences of [sic] the stock's price behavior, a number of courts have rejected or refused to admit into evidence damages reports or testimony by damages experts in securities cases which fail to include event studies or something similar.”)

30 15 U.S.C. § 78u-4(b)(4) places the burden of establishing loss causation on the plaintiffs in any private securities fraud action brought under Chapter 2B of Title 15. *See also Dura*.

cannot prove damages cannot prove she suffered an economic loss. Accordingly, any evidence that would tend to prove the magnitude of damages also will be probative as to whether the plaintiff suffered economic loss, since economic loss generally is just the amount of damages. Thus the economic loss and damages determination elements merge into one.

Under the efficient market hypothesis, a fraudster firm’s stock price will fall by an amount that reflects the market’s assessment of the reduction in the firm’s value due to the revelation of fraud. This drop in stock price is one way to measure the market’s valuation of the magnitude of the fraud. Consider a plaintiff who bought a thousand shares of stock at $100 after a misrepresentation. If the stock later drops from $100 to $90 following a corrective disclosure, that indicates that the stock’s price was inflated by $10 per share as a result of the fraud. A plaintiff who bought a thousand shares after the misrepresentation and still owns those shares at the time of the corrective disclosure will have lost $10,000—$10 each on a thousand shares. This idea is the germ from which a reasonable damage methodology springs.

As with other elements, when calculating damages it is necessary to take into account any effects of other factors that typically are associated with a move in the firm’s stock price. Rather than look directly at the change in the firm’s price, an expert should be expected to use a market model to predict what would have happened to the firm’s stock price on the date in question in the absence of the corrective disclosure. Damages can then be calculated by computing the estimated excess return for the event date and then multiplying this estimated excess return by the firm’s market capitalization at the beginning of the day. For example, if the market model indicates that the firm’s share price could have been expected to drop $4 on the day in question due to other market conditions, then the damages would be appropriately calculated as $6 per share rather than $10. A plaintiff who had purchased 1,000 shares would then be due $6,000 in damages. Thus, if the firm is found liable, the damages it should pay may be calculated from event study results for the alleged corrective disclosure date.

In sum, to prove damages a plaintiff would typically (i) use the estimated price drop after accounting for market-wide factors (i.e. the estimated excess return), on the date of a corrective disclosure as the best estimate of the impact of the fraud on the firm’s price, and (ii) multiply it by the product of (a) the number of shares that the plaintiff acquired after
the misrepresentation and still owned on the date of the corrective disclosure and (b) the stock price just before the corrective disclosure occurred. To prove economic loss, the damages figure just described must be positive, which we can reasonably assume to be true in any case that makes it to the merits stage.\footnote{The only way economic loss would be non-positive is if the estimated excess return described in (i) indicates the stock price actually \textit{rose} an unusual amount on the alleged corrective disclosure date, after adjusting for other market factors on the corrective disclosure date. A plaintiff who pleaded such facts would surely lose on a Rule 12(b)(6) motion to dismiss for failure to state a claim. A plaintiff who failed to plead enough facts to allege a price drop on the date of corrective disclosures likely would, too; \textit{see}, \textit{e.g., Dura}.} To prove loss causation, the plaintiff would have to show that the estimated excess return used to estimate the price drop in (i) above was statistically significant, and also that there was reason to think the significant price drop was the result of fraud, rather than some alternative cause.

4. \textit{Summing up the Role of Event Studies vis-à-vis Merits Elements}

The Court’s decision in \textit{Basic} was influenced by a law review article by Professor Daniel Fischel of the University of Chicago Law School.\footnote{Daniel R. Fischel, \textit{Use of Modern Finance Theory in Securities Fraud Cases Involving Actively Traded Securities}, 38 BUS. LAW. 1 (1982).} Fischel argued that the FOTM theory offered a more coherent approach to securities fraud than then-existing practice because it recognized the market model of the investment decision.\footnote{Id. at 9-10.} Fischel argued that the only relevant inquiry was the extent to which market prices were distorted by fraudulent information—it was unnecessary for the court to make separate inquiries into materiality, reliance, causation and damages.\footnote{Id. at 13.} Importantly, Fischel stated that the effect of fraudulent conduct on market price could be determined through a blend of financial economics and applied statistics, and in his article described the event study methodology (though he did not use the term “event study”).\footnote{Id. at 17-18.}

While subsequent Supreme Court decisions have reinforced the distinction between the elements, the discussion in this section shows, in broad terms, how event studies may be used to establish various
elements at issue in securities fraud litigation. In the spirit of Fischel’s argument, we illustrate the multi-element relevance of an event study-based finding of price impact in Figure 2. As the upper part of this figure indicates, a showing that there was price impact can be used to support three distinct elements: the presumption of reliance, materiality and loss causation. Similarly, a finding of materiality helps support the elements of reliance (via the Basic presumption) and loss causation. Further, the magnitude of the event study supports economic loss and damage determination.

**Figure 2: The Multi-Elemental Role of Price Impact**
We use dashed arrows for this part of the figure in order to emphasize that price impact on the date of an important event might not be sufficient to establish these elements. For example, suppose the stock price drops an unusual amount on the date of an alleged corrective disclosure, relative to the market model used in an expert’s event study, but suppose that defendants identify a different source of bad news for the firm that could also explain the stock price drop. Then an event study’s negative price impact on the alleged corrective disclosure date would support, but not by itself prove, the existence of loss causation, materiality, or the preconditions for the Basic presumption of reliance.

Alternatively, suppose defendants provide evidence that all information in the alleged misrepresentations was already in the public sphere. Then under the efficient market hypothesis, there can’t have been a price impact caused by the alleged misrepresentations, because any such impact would have been immediately factored into price (this is the work of the word “quick” in box (2) of Figure 1). Event study evidence of a price response would merely indicate that the price changed for some unusual reason on the event date; there would be no basis for ascribing causality to the alleged misrepresentations in particular. In this sense, a plaintiff’s ability to prove that the Basic presumption of reliance is applicable is logically dependent on the plaintiff’s ability to prove materiality.37

Relatedly, what if an event study does not support a statistically significant finding? Since price impact is a necessary predicate to the Basic presumption of reliance and the materiality and loss causation elements, a plaintiff’s failure to show price impact negates each of these three elements; it also vitiates any possibility of proving damages or economic loss. While event studies are not formally required, it is difficult to see how a plaintiff would demonstrate price impact without one.38

In sum, a single event study may provide evidence relating to materiality, reliance, loss causation, economic loss, and damages. While

37 This is so even though Amgen held that such proof cannot be required at class certification. The Court made clear in Amgen that class certification on the back of the Basic presumption of reliance does not resolve the issue of reliance in favor of the plaintiffs for any other purpose; plaintiffs might subsequently fail to prove materiality at summary judgment or trial, which would “end the case for one and for all.” Amgen, 133 S.Ct. 1184, 1196 (U.S.,2013).
38 But see Local 703???
such evidence might be insufficient on its own to prove one or more of these elements, event study evidence that negates any of the first three elements will imply the plaintiff’s inability to establish entitlement to damages. These observations explain why event studies play such a central role in securities fraud litigation.\(^\text{39}\)

### B. Event Studies and Pleading

As one would expect, there is a tight relationship between what plaintiffs must prove to win on the merits and what they must plead. However, due to the PSLRA,\(^\text{40}\) there are some important substantive-specific pleading details. One is that in pleading scienter, a complaint must “state with particularity facts giving rise to a strong inference that the defendant acted with the required state of mind.”\(^\text{41}\) This requirement must be met with respect to every act or omission at issue.

Beyond the heightened pleading standard for scienter elements, the relevant pleading standard is the plausibility standard introduced in *Bell Atl. Co. v. Twombly*\(^\text{42}\) and extended to “all civil actions” in *Ashcroft v. Iqbal*.\(^\text{43}\) One way a complaint might run afoul of the plausibility standard would be to plead that the stock price dropped without alleging that the price drop was more than one would have expected in the absence of the corrective disclosure.\(^\text{44}\)

Another problem is the one that arose in *Dura*. There, the plaintiffs pleaded that the stock price was inflated as a result of an alleged misrepresentation, but they failed to plead that the stock price dropped as a result of any corrective disclosure. The Supreme Court held that such a complaint fails to show entitlement to relief as required under Rule

\(^{39}\) Given the critical role of event studies, securities cases are plagued by the battle-of-the-experts problem, with opposing experts reaching opposite conclusions about whether there was a statistically significant price impact on relevant dates, and with each side’s expert opining that the other’s event study is fatally flawed. CITES ___.


\(^{41}\) ___

\(^{42}\) ___

\(^{43}\) ___

\(^{44}\) Such a complaint would at least be vulnerable to the possibility that there exists a non-culpable explanation for the drop in the price on the date in question—namely that the overall market or the firm’s industry suffered losses on that date. This argument would parallel the Supreme Court’s reasoning in *Twombly*, ___, and in *Iqbal*, ___. See *Local 703??*, ___.

15
8(a)(2), since, *inter alia*, a plaintiff might have sold its holdings before any corrective disclosure occurred.\(^\text{46}\)

In the post-*Dura* state of affairs, plaintiffs must identify both alleged misrepresentation and corrective disclosure dates to adequately plead loss causation. They would also be well advised to allege that an expert-run event study establishes materiality, reliance, loss causation, economic loss, and damages. Failure to do so would not necessarily be fatal, but it would leave plaintiffs vulnerable to a Rule 12(b)(6) motion to dismiss for failure to state a claim on alternative-explanation grounds.\(^\text{47}\)

Further, even if they do not do their own event study, plaintiffs might have to contend with a defendant’s event study at the merits stages—i.e., summary judgment or trial.

**C. Event Studies and Class Certification**

Plaintiffs typically move for class certification in securities fraud cases. Assuming that Rule 23(a)’s numerosity, commonality, typicality, and adequacy of representation requirements are met (and they have been uncontroverted in recent Supreme Court securities fraud cases), the relevant hurdle facing plaintiffs is to convince the district court that the action satisfies Rule 23(b)(3)’s predominance and superiority requirements—“that the questions of law or fact common to class members predominate over any questions affecting only individual members, and that a class action is superior to other available methods for fairly and efficiently adjudicating the controversy.”\(^\text{48}\)

The common questions in a typical case include whether there was scienter on the defendant’s part; whether the statements in question were material to the stock’s price; whether they caused the stock price to be inflated; how much it was inflated; and so on. All of these questions concern either the actions or state of mind of the defendant, or the market response to these actions. It takes no more work and costs no more to answer these questions in a case with a class of a thousand plaintiffs than in a case with a single plaintiff. As to these questions, then, a class action

\(^{46}\) *Dura*, ___. We note as well that under the PSLRA, “all discovery and other proceedings shall be stayed during the pendency of any motion to dismiss” subject to narrow exceptions. 15 U.S.C. § 78u-4(b)(3)(B).  
\(^{47}\) *See* note 44, *supra*.  
clearly would be superior to using Rule 20 joinder or separate actions, because it would consume fewer resources and take less time to complete.

The hurdle to class certification arises because a win on the merits would require each plaintiff to establish reliance on the firm’s alleged misrepresentations. If each plaintiff in the proposed class had to come to court to establish reliance individually, there would have to be an inquiry into whether each plaintiff knew of the alleged misrepresentation at the time it purchased the stock and whether each plaintiff would have purchased in the absence of the misrepresentation. There is little debate that these individualized inquiries would predominate over the common questions identified above. So if reliance must be established individually, then a class action would not meet the predominance and superiority requirements. For this reason, the Basic presumption of reliance is the glue binding together the typical securities fraud class action vis-à-vis Rule 23(b)(3).

In the balance of this section, we discuss the implications for class certification analysis of the Supreme Court’s decisions and reasoning in Halliburton I, Amgen, and Halliburton II.

1. Implications of Halliburton I for class certification

The evidence at issue in Halliburton I went not directly to the question whether there had been price impact, but rather to the question of whether any demonstrated price impact was causally related to alleged misrepresentations. The district court had denied class certification because, it found, plaintiffs had failed to prove that the negative price movements on alleged corrective disclosure dates were more likely the result of fraudulent representations by Halliburton than the result of generally negative news that might well have surprised Halliburton itself as much as it surprised the investing public. The district court’s denial

---

49 Indeed, the phrase “price impact” appears nowhere in the district court’s opinion explaining the denial of class certification, and the plaintiffs had submitted an expert report based on an event study, asserting that the alleged corrective disclosures had “resulted in a company-specific decline in the stock price that cannot be attributed to general market trends or other external factors.” Archdiocese of Milwaukee Supporting Fund, Inc. v. Halliburton Co., 2008 WL 4791492, at *1 (N.D.Tex., 2008).

50 For instance, with respect to the alleged asbestos claims magnitude misrepresentation referenced in Part II, supra, which was allegedly the subject of a corrective disclosure
of class certification was thus compelled by Fifth Circuit precedent that made proof of loss causation a prerequisite to class certification.\footnote{Archdiocese of Milwaukee Supporting Fund, Inc. v. Halliburton Co., 2008 WL 4791492, at *2 (N.D.Tex., 2008) (explaining that “the Fifth Circuit requires loss causation to be established at the class certification stage by a preponderance of all admissible evidence.”) (quotation marks and footnote omitted).}

When the case reached the Supreme Court, Halliburton argued that the lower courts’ analysis of loss causation—in the form of failure to link demonstrated price impact on alleged corrective disclosure dates to alleged misrepresentations—should be understood instead as going to the question of whether the plaintiffs had shown price impact of the alleged misrepresentations.\footnote{Erica P. John Fund, Inc. v. Halliburton Co., 131 S.Ct. 2179, 2187 (U.S., 2011).} But the Supreme Court unanimously rejected this argument. It pointed out that the Court of Appeals had repeatedly used the phrase “loss causation” in explaining the rationale for its affirmandom of the district court’s denial of class certification.\footnote{Erica P. John Fund, Inc. v. Halliburton Co., 131 S.Ct. 2179, 2187 (U.S., 2011).} The Supreme Court emphasized that it was taking “the Court of Appeals at its word,” and also that the loss causation and reliance elements were distinct. Thus, the Court held, indirect evidence attacking loss causation could not defeat class certification. The Court vacated the judgment affirming denial of class certification, and remanded. The practical effect of the result in \textit{Halliburton I} is that at the class certification stage, a district court must give plaintiffs the benefit of an unrebuttable presumption that demonstrated price changes were the result of alleged fraud, rather than possibly resulting from some other factor.\footnote{The district court in the \textit{Halliburton} litigation recently took exactly that tack in ruling on a new class certification motion, as we discuss in Part I.C.4, \textit{infra}.}

\section*{2. Amgen and class certification}

That brings us to \textit{Amgen}. The parties in \textit{Amgen} had agreed that the market for Amgen’s stock was efficient and that the statements in question were public (as the parties in \textit{Halliburton} had also agreed). The dispute at class certification concerned the reasons why Amgen’s stock

on December 7, 2001, the district court explained that “[p]laintiffs make no effort to distinguish any corrective effects from the effects of new negative information” which itself would not have been the subject of any fraudulent representation. \textit{Archdiocese of Milwaukee Supporting Fund, Inc. v. Halliburton Co.}, 2008 WL 4791492, at *11 (N.D.Tex., 2008).
price had dropped on alleged corrective disclosure dates, not whether it had dropped.\textsuperscript{55} Defendants argued that to obtain class certification via the Basic presumption of reliance, plaintiffs should be required to prove—not just plead—the materiality of alleged misrepresentations. Further, they argued that defendants should be allowed to rebut such materiality by showing that the truth regarding alleged misrepresentations was publicly known before plaintiffs purchased their shares.\textsuperscript{56}

This latter argument is known as the “truth on the market” defense to Basic’s fraud on the market-based presumption of reliance, since the argument is that the market already knew the truth. Under the efficient market hypothesis, a defendant’s proof that the price drop on an alleged corrective disclosure date was due to generally negative news, rather than to alleged misrepresentations, severs the causal link between alleged misrepresentations and price impact. In terms of Figure 1, an alleged misrepresentation that occurs after the market already knows the truth cannot change market perceptions of firm value (box (1)), because any effect of the truth will already have been incorporated into the market price (box (2)). The Amgen Court held that the truth-on-the-market defense goes only to materiality. Because materiality itself involves a common question to the class, the Court held that lack of materiality is irrelevant to the class certification inquiry,\textsuperscript{57} and therefore the truth-on-the-market defense could not defeat class certification.

It is true that Amgen had asserted that plaintiffs failed to prove materiality. And it is also true that materiality, \textit{qua} element of a securities fraud action, is analytically distinct from reliance. But

\begin{itemize}
\item \textsuperscript{55} See, e.g., the defendants’ brief in opposition to class certification at the district court, \textit{In Re AMGEN INC. Securities Litigation.}, 2009 WL 1406107 (C.D.Cal.) (“Defendants have made a showing both that information was publicly available and that the market drops that Plaintiff relies on to establish loss causation were not caused by the revelation of any allegedly concealed information. Rather, as Defendants have shown, the market was privy to the truth, and the price drops were the result of third-parties' reactions to public information.”) (quotation marks removed).
\item \textsuperscript{57} Id. at 1204 (“even a definitive rebuttal on the issue of materiality would not undermine the predominance of questions common to the class”).
\end{itemize}
Amgen’s argument can also be viewed as attacking what our discussion in Part I.A.1, supra, shows is a key logical prerequisite of the fraud on the market theory underlying the Basic presumption of reliance (see also discussion of Figure 1). Under the efficient market hypothesis, a plaintiff cannot have relied on price changes caused by alleged representations if these representations were not material and thus did not move the market price. Given the efficient market hypothesis, movements in the stock will happen as a result of representations only if those representations are material.

Since Rule 23(b)(3)’s predominance requirement cannot be met without recourse to the Basic presumption, it would seem that predominance cannot be met by plaintiffs who fail to prove all the logical prerequisites of the Basic presumption—including materiality of the misrepresentations alleged to have caused the price change on an alleged corrective disclosure date. A convincing assertion of the truth-on-the-market defense would then seem to be an appropriate way to rebut the Basic presumption, and therefore to foreclose class certification pursuant to Rule 23(b)(3). That is the way dissenting Justices Scalia, Thomas and Kennedy saw things in Amgen.58

But a six-justice majority rejected this argument. Justice Ginsburg wrote for the Court that for class certification purposes, materiality differed importantly from the market efficiency and publicity prerequisites for invoking the Basic presumption:

A failure of proof on the issue of materiality … not only precludes a plaintiff from invoking the fraud-on-the-market presumption of classwide reliance; it also establishes as a matter of law that the plaintiff cannot prevail on the merits of her Rule 10b–5 claim…. While the failure of common, classwide proof on the issues of market efficiency and publicity leaves open the prospect of individualized proof of reliance, the failure of common proof on the issue of materiality ends the case for the class and for all individuals alleged to compose the class.59

Thus, the Amgen majority appears to have erected a wall that precludes considering, at the class certification stage, any question that


59
would be merits-determinative as to the entire putative class. Amgen did offer defendants a ray of hope. Four Justices, writing in separate opinions, indicated a willingness to reexamine Basic’s FOTM presumption. Accordingly, the following term, Halliburton petitioned for certiorari in the same litigation that had generated Halliburton I, asking the Court to overrule Basic. In the alternative, Halliburton argued that plaintiffs seeking to benefit from the Basic presumption of reliance should have the burden of proving price impact at the motion for class certification.

3. Halliburton strikes back: Halliburton II and class certification

Less than 18 months after Amgen, the Court punched a big hole in the wall between class certification and merits-relevant issues. In Halliburton II, the Court held that evidence of lack of price impact—its own unquestionably sufficient to defeat the loss causation and materiality elements—could be used to defeat the Basic presumption at class certification. Chief Justice Roberts argued that price impact and materiality are not the same thing. Without proof of publicity and market efficiency, he argued, “the fraud-on-the-market theory underlying the presumption completely collapses” such that class certification cannot be allowed.

Since these indirect forms of evidence are necessary to establish price impact, in their absence there would be no basis to believe that market price would reflect the effects of alleged misrepresentations. Thus there could be no reason to adopt Basic’s presumption of reliance on market price as the reflector of fraud. Embedded in this argument is

60 This wall looms a bit tall in light of the Court’s earlier declaration in Wal-Mart v. Dukes that determinations overlapping merits issues will sometimes need to be made at the class certification stage. See Wal-Mart Stores, Inc. v. Dukes, 131 S.Ct. 2541, 2551-52 (U.S.,2011) (explaining that the “rigorous analysis” necessary to justify class certification will “[f]requently … entail some overlap with the merits of the plaintiff’s underlying claim,” explaining that “[t]hat cannot be helped,” and pointing out that “[t]he necessity of touching aspects of the merits in order to resolve preliminary matters, e.g., jurisdiction and venue, is a familiar feature of litigation.”).

61 See Cert petition: at __ (arguing that the “court of appeals' previous opinion found no evidence that Halliburton's alleged misrepresentations moved the market price… Without price impact, there is no reason to presume classwide reliance on the misrepresentations by relying on the market price, because nothing establishes that the misrepresentations were even reflected in that price.”).

the financial economics-fueled claim that if the stock price was not affected by an alleged corrective disclosure, there can be no basis for believing that the alleged misrepresentations were reflected in the firm’s price; thus there can be no basis for presuming that a purchaser relied on such misrepresentations via price. In short, any time the Basic presumption of reliance is invoked, price impact is at stake one way or another.

The same is not true of, say, the truth-on-the-market defense. This defense takes it as given that plaintiffs have demonstrated a significant price change on the event date in question but challenges the cause of such price change. The truth-on-the-market defense therefore requires the court to consider additional evidence over and above the event study evidence likely to be in the record already. This practical reality, together with the existence of the very recent Halliburton I and Amgen precedents might rationalize the Halliburton II Court’s willingness to allow courts to consider price impact, but not other necessary conditions for the Basic presumption of reliance, at class certification.

Chief Justice Roberts suggested as much when he emphasized in Halliburton II that the Court’s choice was not whether to consider price impact evidence at class certification or instead wait to evaluate it at the merits. The Chief Justice explained that price impact evidence from event studies was often already before the court at the class certification stage, because plaintiffs were using event studies to demonstrate market efficiency and defendants were using event studies to counter this evidence. Under these circumstances, prohibiting a court from relying on this same evidence to evaluate whether the fraud affected stock price “makes no sense.”

Because the question of price impact itself is unavoidably on the docket for consideration at class certification, the Chief Justice explained that the Court’s actual choice concerned merely the type of evidence it would allow parties to use in demonstrating price impact on the dates of alleged misrepresentations or alleged corrective disclosures. “The choice … is between limiting the price impact inquiry before class certification to indirect evidence”—directed at establishing market efficiency in

---

63 The Halliburton litigation provides an odd factual context to make this determination, since Halliburton had actually not disputed the efficiency of the public market in its stock. Archdiocese of Milwaukee Supporting Fund, Inc. v. Halliburton Co., 2008 WL 4791492, at *1 (N.D.Tex., 2008).
The direct evidence that the Court’s majority determined to allow, concerning price impact on dates of alleged misrepresentations and alleged corrective disclosures, will typically be provided in the form of event studies.

Taken together, *Halliburton I*, *Amgen*, and *Halliburton II* evince a bright-line two-pronged rule: (1) Defendants may rebut the Basic presumption of reliance—and thus defeat class certification—using event study evidence on price impact; but (2) no other type of evidence may be used to rebut the Basic presumption of reliance at class certification. On remand, the district court in the *Halliburton* litigation has extracted and applied this very rule:

> Basic presupposes that a misrepresentation is reflected in the market price at the time of the transaction. See *Halliburton II*, 134 S.Ct. at 2416. Thus, at this stage of the proceedings, the Court concludes that the asserted misrepresentations were, in fact, misrepresentations, and assumes that the asserted corrective disclosures were corrective of the alleged misrepresentations.... While it may be true that a finding that a particular disclosure was not corrective as a matter of law would “sever the link between the alleged misrepresentation and ... the price received (or paid) by the plaintiff ....” the Court is unable to unravel such a finding from the materiality inquiry. See *Halliburton II*, 134 S.Ct. at 2415–16.

Therefore, the district court determined that whether alleged misrepresentations actually misrepresented, and whether disclosures that plaintiffs alleged corrected these alleged misrepresentations really were corrective, “is not a proper inquiry at the certification stage.”

A key remaining open question is how the burden of persuasion will be allocated when defendants seek to use event studies related to price impact to rebut the Basic presumption of reliance. Will defendants have to provide statistically significant evidence proving that there was no price drop following a corrective disclosure? Or will they have to prove

---

only that plaintiffs cannot prove that there was such a drop? As Professor Merritt Fox has explained, the distinction may not be trivial, for reasons related to the way statistical significance testing works.\textsuperscript{67} We discuss this issue further in Part ___, infra [STILL NEED TO ADD THIS].

4. Developments related to class certification since Halliburton II

Early returns suggest that defendants are responding to the Supreme Court’s invitation and seeking to defeat class certification by showing the absence of price impact.\textsuperscript{68} In addition to subsequent developments in the Halliburton litigation itself, discussed below [STILL TO BE ADDED ___], two early opinions have considered this question most carefully.

In \textit{Local 703 v. Regions Fin. Corp.},\textsuperscript{69} the Court of Appeals concluded that the defendant had provided evidence that the stock price did not change in light of the misrepresentations and that the trial court, acting prior to \textit{Halliburton II} “did not fully consider this evidence.”\textsuperscript{70} Accordingly, the court vacated and remanded “for fuller consideration…of all the price-impact evidence submitted below.”\textsuperscript{71} On remand, defendants argued that they had successfully rebutted the Basic presumption by providing evidence of no price impact on both the misrepresentation date and the date of the corrective disclosure.\textsuperscript{72} But the trial court disagreed. The court reasoned that the defendant’s own expert conceded that the 24% decline in the issuer’s stock on the date of the corrective disclosure was far greater than the NYSE’s 6.1% decline that day, and that, given this discrepancy, the defense had not shown the

\textsuperscript{69} Local 703, I.B. of T. Grocery & Food Employees Welfare Fund v. Regions Fin. Corp., 762 F.3d 1248 (11th Cir. Ala. 2014)
\textsuperscript{70} Id. at 1258.
\textsuperscript{71} Id.
\textsuperscript{72} Id. at *18.
absence of price impact.73 This decision places the burden of persuasion concerning price impact squarely on defendants.

In Aranaz v. Catalyst Pharmaceutical Partners Inc.,74 the district court permitted defendant an opportunity to rebut price impact at class certification. The Aranaz court explained, however, that defendant was limited to direct evidence that the alleged misrepresentations had no impact on stock price.75 The defendants conceded that the stock price rose by 42% on the date of the allegedly misleading press release and fell by 42% on the date of the corrective disclosure76 but argued that other statements in the two publications caused the “drastic changes in stock price.”77 The court concluded that, because the defendant had the burden of proving that “price impact is inconsistent with the results of their analysis,”78 their evidence was not sufficient to show an absence of price impact. This determination as to the burden of persuasion tracks the approach taken by the Local 703 court, discussed above. Further, following Amgen, the Aranaz court ruled that the truth-on-the-market defense would not defeat class certification because it concerns materiality and not price impact.79

DISCUSS JULY 2015 DISTRICT COURT CERTIFICATION ORDER IN Halliburton

73 Defendants argued that their expert’s event study “conclusively finds no price impact on January 20, 2009 [the date of the corrective disclosure].”
75 Under Halliburton I and Amgen, this limit is appropriate. The district court in Halliburton took the same approach on remand following Halliburton II; see also Erica P. John Fund, Inc. v. Halliburton Co., 309 F.R.D. 251, 261-62 (N.D. Texas, 2015) (“This Court holds that Amgen and Halliburton I strongly suggest that the issue of whether disclosures are corrective is not a proper inquiry at the certification stage. Basic presupposes that a misrepresentation is reflected in the market price at the time of the transaction…. [A]t this stage of the proceedings, the Court concludes that the asserted misrepresentations were, in fact, misrepresentations, and assumes that the asserted corrective disclosures were corrective of the alleged misrepresentations…. While it may be true that a finding that a particular disclosure was not corrective as a matter of law would ‘sever the link between the alleged misrepresentation and … the price received (or paid) by the plaintiff ….,’ the Court is unable to unravel such a finding from the materiality inquiry.”).
76 Aranaz at *22.
77 Id. at *29.
78 Id. at 32.
79 Id. at *27.
These early cases highlight two points about the scope of *Halliburton II*. First, while the district courts have understood the decision in *Halliburton II* as providing the defendant with a right to rebuttal, they have not read it to increase the proof required by plaintiffs to obtain class certification. This was a point that Justice Ginsburg wrote separately to emphasize in *Halliburton II*.81

Second, courts have properly recognized both the limits and the evidentiary value of event studies in the context of alleged fraudulent confirmatory statements.82 As one court explained: “A material misstatement can impact a stock's value either by improperly causing the value to increase or by improperly maintaining the existing stock price.”83 In addition, the courts noted that a showing that a stock was inflated prior to the defendants’ alleged misrepresentation does not *per se* demonstrate the absence of price impact resulting from the representations themselves.84 While event study evidence related to alleged misrepresentation dates might therefore have limited probative value, courts have noted that event studies as to corrective disclosure dates can still be probative in such cases. That is because a showing that the stock price responded to a subsequent corrective disclosure can

---

80 See, e.g., Local 703, I.B. of T. Grocery & Food Emples. Welfare Fund v. Regions Fin. Corp., 2014 U.S. Dist. LEXIS 162403 (N.D. Ala. Nov. 19, 2014) (explaining that while *Halliburton II* “recognized that an event study can show the reaction of market price to corrective disclosures, nothing in *Halliburton II* requires the plaintiffs to produce an event study in opposition to defendants' event study on a class certification motion”) (citation omitted). See also McIntire v. China MediaExpress Holdings, Inc., 2014 U.S. Dist. LEXIS 113446, *39 (S.D.N.Y. Aug. 15, 2014) (observing that *Halliburton II* did not change prior Second Circuit case law which permitted a defendant to rebut the Basic presumption by showing a lack of price impact).

81 See *Halliburton II*, at __ (Ginsberg, J., concurring).

82 This principle had been recognized by some lower court decisions prior to *Halliburton II*, e.g., FindWhat, 658 F.3d at 1310 ("A corollary of the efficient market hypothesis is that disclosure of confirmatory information — or information already known by the market — will not cause a change in the stock price. This is so because the market has already digested that information and incorporated it into the price.").


84 See, e.g., IBEW Local 98 Pension Fund v. Best Buy Co., 2014 U.S. Dist. LEXIS 108409, *19 (D. Minn. Aug. 6, 2014) (“Even though the stock price may have been inflated prior to the earnings phone conference, the alleged misrepresentations could have further inflated the price, prolonged the inflation of the price, or slowed the rate of fall.”).
provide indirect evidence of price impact.85 Such a conclusion opens the door to consideration with respect to price impact of the type of event study conducted for purposes of loss causation, as we discuss below.

D. Summing Up: Substance, Procedure, Economics, and Event Studies in Securities Fraud Litigation

Plaintiffs in securities fraud litigation must prove six elements (plus damages). Several of these elements—reliance, materiality, and loss causation—are susceptible of at least partial proof based on an event study, as are damage levels. Event studies can play this role because courts have largely adopted the efficient market hypothesis of financial economics, under which the price of a stock traded on public markets can usually be expected to respond quickly to material new information concerning the value of a firm.

None of that is to say that event studies are infallible in the securities fraud litigation context. It is important to recognize that at best, event studies tell us whether price changed by an unusual amount on a particular day. They cannot tell us the specific reason such a change occurred, so they cannot rule out the possibility that a stock price dropped on negative news that is unrelated to any actual fraud. Depending on the context, event studies may need to be coupled with some analysis sufficient to rule out alternative explanations. And of course, event studies cannot separately identify the relative contribution to a price impact of multiple, simultaneously released disclosures; doing that requires additional evidence.86

Still, event studies are in many cases at least partially probative as to multiple elements of a securities fraud action, and following Halliburton II they will be especially powerful at the class certification stage. Developments in case law indicate that event studies will become even more central in securities litigation, as loss causation and price impact have taken center stage at the pleading and class certification stages.

85 See id. at 20 (“price impact can be shown by a decrease in price following a revelation of the fraud”); In re Bank of Am. Corp., 281 F.R.D. at 143 (finding that stock price's negative reaction to corrective disclosure served to defeat defendant's argument of lack of price impact).
86 See discussion at ___, infra.
II. THE THEORY OF FINANCIAL ECONOMICS AND THE PRACTICE OF EVENT STUDIES: A VIEW FROM 10,000 FEET

The theory of financial economics adopted by courts for purposes of securities litigation is based on a central premise: Publicly released information concerning a security’s price will be quickly incorporated into market price, provided that the market in question functions efficiently. This premise is known in financial economics as the semi-strong form of the efficient market hypothesis, but we will refer to it simply as the efficient market hypothesis. Under the efficient market hypothesis, statements that overstate a firm’s value will lead to an inflation of the firm’s stock price over the level that true conditions warrant. Conversely, statements that correct such inflationary misrepresentations will lead the stock price to fall.

To measure how much stock prices respond to various types of news, financial economists began using event studies. The simplest form of an event study would be a comparison of a stock’s percentage change in price—its “return”—on a day when news of interest hits the market, to the range of returns typically observed. If a stock’s return on the date a planned merger is announced is 3%, that might constitute evidence that the stock market viewed the merger as a good thing for the firm, provided that 3% is atypical enough.

No real world event study is this simple, however. A first necessary modification is to account for changes in market conditions that might be expected to affect the firm’s price even in the absence of the merger announcement. For example, if the government happened to release a strong jobs report on the same date, many companies’ stocks would likely have increased in value anyway. This sort of effect may be accounted for by controlling statistically for an index of other stocks’ returns on each date considered in the event study. An especially convincing version of this idea is to control for the stock performance of other firms engaged in broadly similar lines of business to those of the firm in question. Event studies are therefore conducted by using

---

87 There are also strong and weak forms. The strong form of the efficient market hypothesis holds that even information that is held only privately is reflected in stock prices, since those with the information can be expected to trade on it. The weak form holds only that “historical price data are efficiently digested and, therefore, are useless for predicting subsequent stock price changes.” Robert Hagin, MODERN PORTFOLIO THEORY at 11, Dow Jones-Irwin (1979).
regression analysis to account for market- and/or industry-level conditions. As we discuss below, the appropriate measure of price impact flowing from these studies is the stock’s so-called “excess return”—the difference between the actual return on the date in question and the return that would be predicted for the day in question based on the regression model’s estimates.

Figure 3 illustrates the calculation of excess returns from actual returns and expected returns. The figure plots the stock’s actual daily return on the vertical axis, and its expected daily return on the horizontal axis, where the expected return is based on estimates from a regression model of the type described in the previous paragraph. The upwardly sloped straight line represents the collection of points where the actual and expected returns are equal. The magnitude of the excess return at a given point is the height between that point and the upwardly sloped straight line. The point plotted with a circle indicates an example in which the actual return is above the expected one, so the excess return is positive. By contrast, at the point plotted with a square, the actual return is below the expected one, so the excess return is negative.

A second modification to the simplistic approach described above is to require statistical significance of the estimated excess return before concluding that an event of interest actually had an effect on the stock price. The use of statistical significance testing guards against the possibility that researchers or experts will determine that an event had an impact on price, when the observed price change was really just the result of other factors unrelated to the alleged corrective disclosure. Statistical significance testing ensures that there is a low probability—typically intended to be 0.05, or a 5% chance—of determining that an event caused a price impact when the event really did not.

As one pair of recent commentators have noted, “failure to make adjustments for the effect of market and industry moves nearly always dooms an analysis of securities prices in litigation.” Brav and Heaton, footnote 20, supra, at 8.

This kind of mistake of statistical inference—rejecting a hypothesis that is actually true—is known as a Type I error. Another type of mistake is to fail to reject an actually false hypothesis. This second type of mistake is known as a Type II error. A test with a high Type II error rate is also said to have low power, because statistical power is the probability of correctly rejecting a false hypothesis; low power and a high Type II error rate are thus two sides of the coin. It has recently been pointed out that single-firm event studies like those used in securities fraud litigation can be expected to have relatively low power. The details of this potentially very important issue are largely
The proper approach to testing for statistical significance involves some subtle methodological details, which we defer to Part II, infra. However, every method of testing for statistical significance will involve the use of the regression model estimates discussed above to estimate a “critical value” for the excess return. The estimated excess return of interest will be considered statistically significant if and only if (i) its magnitude exceeds the critical value, and (ii) its sign is the one that would be expected based on the claimed direction of price impact on the date in question.  

---

90 As we discuss in Part III.B.1, infra, event studies used in securities fraud litigation typically use an unreliable method to estimate the critical value. We discuss the proper way to do so, based on recent work by GHK, note 123, supra, in Part III.B.1, infra.  
91 In academic studies, it is sometimes the case that theory does not make a clear prediction as to the sign of the event date excess return. In such cases statistical
Consider again our example from above, in which a planned merger is announced, and the firm’s stock price rises 3% on that day. Conducting an event study designed to detect the price impact of the announcement would first require determining the excess return for the date in question. If market conditions on the event date were such that the regression model underlying the event study predicts that the firm’s stock price should have been expected to rise even in the absence of the merger, then the estimated excess return will be less than 3%; if market conditions would have predicted a price drop on the announcement date, then the opposite is true. With the estimated excess return in hand, one would then test for statistical significance by comparing the estimated excess return to the critical value implied by the regression model’s estimates. If the estimated excess return is larger in magnitude than the critical value, and has the right sign, then we would determine that the merger announcement had a statistically significant price impact. The estimated excess return itself can be used to estimate damages if such estimation turns out to be necessary.

Figure 4 illustrates the key issues just discussed. As in Figure 3, the upwardly sloped line indicates the set of points where the actual and excess returns are equal. The shaded area depicts the set of points where the actual return is far enough below the expected return—i.e., where the excess return is sufficiently negative—so that the excess return indicates a statistically significant price drop on the date in question.

---

significance is appropriately tested by checking whether the estimated excess return is large in magnitude regardless of its sign; such tests are “two-sided” tests of statistical significance. In securities fraud litigation, by contrast, there will virtually always be a clear direction in which price must move if the plaintiff is to carry its burden of persuasion. That means that tests of statistical significance based on event study results must be conducted in a “one-sided” way, so that an estimated excess return is considered statistically significant only if it moves in the alleged direction. The one-sided/two-sided distinction is an important one, and one that courts and expert witnesses regularly miss. We discuss it in more detail in ___, infra.
The graph also depicts three points that help clarify the statistical significance analysis. The circle depicts a point that has a positive excess return, which is the wrong sign for an excess return to have on an alleged corrective disclosure date. The triangle depicts a point that has a negative excess return, consistent with an alleged corrective disclosure, but for which the excess return is too close to the expected return to be statistically significant. The square, however, depicts an excess return that is both negative and sufficiently far below the expected return such that we conclude there was a statistically significant price drop—as would be necessary for a plaintiff alleging a corrective disclosure. Thus, for an alleged corrective disclosure date, only the square would demonstrate the plaintiff’s required proof.

A number of substantive economic and methodological statistical assumptions are embedded in the foregoing discussion. First, there is the assumption that markets react rationally in incorporating bad news. While financial economists continue to debate the degree of rationality in financial markets, as a practical matter this assumption seems
unproblematic in the securities fraud litigation context. \textsuperscript{92} Reasonable investors might disagree as to the reduction in firm value associated with unfavorable news about the failed restructuring, but it seems unreasonable to imagine that such news would lead anyone to increase the value she puts on the firm.

To briefly illustrate the conduct of event studies in securities fraud litigation, consider the following facts taken from the \textit{Halliburton} litigation:

\textbf{(1)} Plaintiffs alleged that in its 1998 10-K filing, Halliburton failed to disclose that it faced the risk of having to “shoulder the responsibility” for certain asbestos claims filed against other companies; further, plaintiffs alleged that Halliburton failed to correctly account for this risk. \textsuperscript{93} The 10-K report was filed on March 23, 1999, which is thus a date on which plaintiffs alleged Halliburton made a misrepresentation. On June 28, 2001, Halliburton disclosed that one of the companies in question had requested Halliburton’s assistance in paying for asbestos claims. \textsuperscript{94}

\textbf{(2)} On November 8, 2001, plaintiffs alleged that Halliburton stated in its Form 10-Q filing for the third quarter of 2001 that the company had an accrued liability of $125 million related to asbestos claims, and that “[W]e believe that open asbestos claims will be resolved without a material adverse effect on our financial position or the results of operations.” \textsuperscript{95} Plaintiffs also alleged that this representation was false and misleading. \textsuperscript{96} They further alleged that a (partial) corrective disclosure occurred on December 7, 2001, when Halliburton “issued a release detailing all the recent asbestos verdicts against it, which made it clear that Halliburton’s existing $125 million accrued liability for its asbestos exposure was grossly inadequate and that the Company’s

\textsuperscript{92} The Supreme Court concluded as much in \textit{Halliburton II}: ___.
\textsuperscript{93} FCAC at 58, ¶ 74.
\textsuperscript{94} FCAC at 116-117, ¶ 170 (citing a Halliburton press release issued on June 28, 2001).
\textsuperscript{95} FCAC at 126, ¶ 189.
\textsuperscript{96} FCAC at 126-28, ¶ 190.
prior reassurances regarding the financial impact of its exposure to asbestos liabilities were false.\textsuperscript{97}

The alleged misrepresentations described above were “confirmatory”, in the sense that the plaintiffs alleged that Halliburton failed to inform the market of negative news. The alleged result was that Halliburton’s stock price was inflated because it remained at a higher level than it would have had the alleged misrepresentations not been publicly made. Thus, in considering the price impact of the alleged misrepresentations in the \textit{Halliburton} litigation, the district court allowed plaintiffs to focus on whether the respective alleged corrective disclosures on June 28, 2001, and December 7, 2001, were associated with reductions in Halliburton’s stock price.\textsuperscript{98}

The court concluded that the estimated excess return was negative on June 28, 2001, as would have to be true for plaintiffs to benefit from the \textit{Basic} presumption of reliance. But it also determined, after weighing two competing expert reports, that the defendants had carried their burden of

\textsuperscript{97} FCAC at 128, ¶ 191. The district court has characterized the December 7, 2001, issue only as involving a Baltimore jury’s verdict, 309 F.R.D. 251, 276 (N.D.Tex., 2015). The distinction is immaterial for our purposes.

\textsuperscript{98} \textit{Erica P. John Fund v. Halliburton Co.}, 309 F.R.D. 251 ___ (N.D.Tex., 2015). As discussed in Part ___, this is not a novel approach. For example, one Court of Appeals has explained that public statements falsely stating information which is important to the value of a company’s stock traded on an efficient market may affect the price of the stock even though the stock’s market price does not soon thereafter change. For example, if the market believes the company will earn $1.00 per share and this belief is reflected in the share price, then the share price may well not change when the company reports that it has indeed earned $1.00 a share even though the report is false in that the company has actually lost money (presumably when that loss is disclosed the share price will fall).

\textit{Nathenson v. Zonagen Inc.}, 267 F.3d 400, 419 (C.A.5 (Tex.), 2001). However, by its very nature a corrective disclosure cannot be confirmatory: for the alleged corrective disclosure to be truly corrective, it must really be new news. Thus evidence concerning the stock price change on the date of an alleged corrective disclosure will always be probative. For simplicity, we will generally focus on the case in which alleged misrepresentations were confirmatory, so that we may focus our discussion on the corrective disclosure date. \textit{But see} Part ___, \textit{infra}, which considers the situation when plaintiffs must establish price impact on both an alleged misrepresentation date and an alleged corrective disclosure date.
proving that the alleged corrective disclosure did not have a statistically significant price impact. For that reason, the district court denied class certification with respect to the first alleged misrepresentation. For the second alleged misrepresentation, however, the district court found that the alleged corrective disclosure on December 7 was associated with a statistically significant price impact in the direction necessary for plaintiffs to benefit from the Basic presumption. The court therefore certified a class action with respect to the alleged misrepresentations associated with December 7, 2001.99

These developments illustrate the use of event studies at the cutting edge of contemporary securities fraud litigation. In the next Part, we present our own stylized event study of dates involved in the ongoing Halliburton litigation to illustrate both typical practice and some important refinements that experts and courts should make.

III. USING THE HALLIBURTON LITIGATION TO ILLUSTRATE EVENT STUDY PRACTICE, AND PROBLEMS

This Part uses data and methods from the expert reports in the Halliburton case to illustrate a number of important methodological issues related to measuring price impact. Our objective is partly to provide a basic illustration for those readers having limited familiarity with the broad operational details. While we do not shy away from important methodological points, we attempt to keep the level of discussion as accessible as possible to those without formal training in statistical methodology.

Plaintiffs alleged that between the middle of 1999 and the latter part of 2001, Halliburton and various of the company’s officers—collectively referred to here as simply “Halliburton”—employed a “scheme to manipulate and falsify Halliburton’s [1998-2001] financial results and statements” via false and misleading statements about various aspects of

99 Halliburton subsequently requested and received permission to pursue an interlocutory appeal of the class certification order, pursuant to Rule 23(f); see ___. At the time of writing, briefing had not yet occurred. In the meantime, the district court has denied both Halliburton’s motion to stay discovery pending resolution of that appeal and its motion in the alternative for summary judgment on the basis that its indirect evidence severs the link between price impact and any alleged misrepresentation. See No. 637 [order denying stay, etc.].
the company’s business. The class period on which we will focus is the period between July 22, 1999, and December 7, 2001.

The operative complaint, together with the report filed by plaintiffs’ experts, named a total of 35 dates on which either misrepresenting statements or corrective disclosures (or both) allegedly occurred. The parties sparred over the statistical facts regarding Halliburton’s stock performance on these key dates, with much disagreement evident from the different conclusions drawn in the parties’ experts’ reports. Our objective in this Part is not to take sides on these disagreements, as such, but instead to use the facts of the Halliburton litigation erect a foundation on which judges, legal scholars, and testifying experts may build in the future.

On July 25, 2015, the district court issued its most recent order and memorandum opinion concerning class certification. By this point of a litigation that had been ongoing for more than 13 years, the aspects of the controversy for which event studies could be probative revolved around six dates on which plaintiffs Halliburton had issued alleged corrective disclosures: December 21, 2000; June 28, 2001; August 9, 2001; October 30, 2001; December 4, 2001; and December 7,
Since the possibility of unusual stock return behavior is the object of event study in the case, these dates should be removed from the set used in estimating the market model, and we do exclude them. To account for factors outside the litigation likely associated with Halliburton’s stock performance, we followed the parties’ experts and estimated a market model with multiple reference indexes. The first such index, introduced by the defendants’ expert, is intended to track the performance of the S&P 500 Energy Index during the class period. The plaintiffs’ expert pointed out that this index is dominated by “petroleum refining companies, not energy services companies like Halliburton.” In his own market model, he therefore added a second index intended to more closely reflect the performance of Halliburton’s industry peers, which we also included. Third, we included an index constructed to mimic the one the defendants’ expert constructed to reflect the engineering and construction aspects of Halliburton’s business. Because we found that the return on the S&P 500 overall index added no meaningful explanatory power to the model, we did not include it.

The resulting market model estimates are set forth in Table 1. These estimates indicate that Halliburton’s daily stock return moves

---

109 On this date, “Halliburton announced Maryland verdict.” Coffman Report at 3, ¶ 8 (citing FCAC at ¶ 191).
110 This issue was controverted between the parties, with plaintiffs’ expert Coffman excluding all 35 of the dates identified in either the complaint or an earlier expert’s report. The district court accepted the argument that dates not identified as alleged corrective disclosure dates should be included in the event study, as defendants’ expert had argued. 309 F.R.D. 251, 265 (2015).
111 The defendants’ expert used this index in the market model she described in several reports. We obtained a list of companies represented in this index during the class period from Exhibit 1 of the report of the plaintiffs’ expert. We then calculated the return on a value-weighted index based on these firms by calculating the daily percentage change in total market capitalization of these firms.
113 This index is composed “of the companies cited by analysts as Halliburton’s peers at least three times during the Class Period and with a market cap of at least $1 billion at the end of the Class Period.” Coffman Report at 19, ¶ 33.
114 We calculated the return on this index in the same way as the return on the energy index described in 111, supra; we took the list of included companies from Exhibit 3b of the Coffman Report.
115 We took the list of companies for this index from the Allen Report at 12, n. 20.
116 We used simple daily returns to estimate this model. We found nearly identical results when we entered all return variables in this model in terms of the natural
nearly one-for-one with the industry peer index constructed from analyst reports: a one percentage point increase in the industry peer index return is associated with roughly a 0.9-point increase in Halliburton’s return. The energy index return is much less correlated with Halliburton’s stock return, with a coefficient of only about 0.2. Both the energy and industry peer index coefficients are highly statistically significant, with each being many multiples of its estimated standard error. By contrast, the return on the energy and construction index has essentially no association with Halliburton’s stock return and is statistically insignificant.

Table 1: Market Model Regression Estimates

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient Estimate</th>
<th>Estimated standard error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Industry Peer Index</td>
<td>0.903</td>
<td>0.031</td>
</tr>
<tr>
<td>Energy Index</td>
<td>0.210</td>
<td>0.048</td>
</tr>
<tr>
<td>E&amp;C Index</td>
<td>0.033</td>
<td>0.036</td>
</tr>
<tr>
<td>Intercept</td>
<td>-0.001</td>
<td>0.001</td>
</tr>
<tr>
<td>Root mean squared error</td>
<td>0.01745</td>
<td></td>
</tr>
<tr>
<td>Number of dates</td>
<td>593</td>
<td></td>
</tr>
</tbody>
</table>

We used the market model estimates to calculate daily estimated excess returns for the six event dates excluded from estimation of the model. For each date, we first calculated the expected return. To do so, we calculated the contribution of each index by multiplying the index’s Table 1 coefficient estimate by the observed value of the index on the event date in question. Then we summed up the three index-specific products just created, and added the intercept (which is so low as to be effectively zero). The result is the event-date expected return based on the market model (the variable plotted on the horizontal axis of Figure 3 and Figure 4). The excess return for each event date is then found by subtracting each date’s estimated expected return from its actual return.

Table 2 reports the actual, estimated expected, and estimated excess returns for each of the six alleged corrective disclosure dates in the *Halliburton* litigation, sorted in descending order of magnitude, i.e., from logarithm of one plus the daily return, as experts sometimes do. For simplicity we therefore decided to stick with the raw daily return.
most negative to least negative. The actual returns are all negative, indicating that Halliburton’s stock price dropped on each of the alleged corrective disclosure dates. In half of the three cases, the estimated expected return was also negative, indicating that typical market factors would be expected to cause Halliburton’s stock price to fall on those three dates, even in the absence of any unusual event; for these dates, the excess return will be less negative than the actual return. For the other three dates, market developments would have been expected to cause an increase in Halliburton’s stock price, which means the estimated excess returns on those dates will imply larger price drops than are reflected in the actual returns.

Table 2: Actual, Expected, and Excess Returns for Event Dates

<table>
<thead>
<tr>
<th>Event Date</th>
<th>Actual Return</th>
<th>Estimated Expected Return</th>
<th>Estimated Excess Return</th>
</tr>
</thead>
<tbody>
<tr>
<td>December 7, 2001</td>
<td>-0.424</td>
<td>0.03</td>
<td>-0.427</td>
</tr>
<tr>
<td>August 9, 2001</td>
<td>-0.045</td>
<td>0.06</td>
<td>-0.051</td>
</tr>
<tr>
<td>December 4, 2001</td>
<td>-0.007</td>
<td>0.029</td>
<td>-0.036</td>
</tr>
<tr>
<td>December 21, 2000</td>
<td>-0.020</td>
<td>-0.008</td>
<td>-0.012</td>
</tr>
<tr>
<td>October 30, 2001</td>
<td>-0.052</td>
<td>-0.043</td>
<td>-0.009</td>
</tr>
<tr>
<td>June 28, 2001</td>
<td>-0.038</td>
<td>-0.031</td>
<td>-0.008</td>
</tr>
</tbody>
</table>

Finally, the estimated excess return column in Table 2 shows that estimated excess returns were negative on all six dates. Even on dates when Halliburton’s stock price would have been expected to fall based on market developments, it fell more than it would have been expected to. Because these estimated excess returns account for the effect of market developments on Halliburton’s stock price, they embody the first basic contribution of an event study: controlling for factors that would typically be expected to cause variation in Halliburton’s stock price. What remains to be done is to test for statistical significance in order to determine whether the event date estimated excess returns are large enough such that we should reject the possibility that they might simply reflect essentially random movements in price.

The next section discusses how such a determination is typically made. Section B discusses four problems with this typical approach. Section C discusses best practices in response to these four problems. Section D discusses the issues that arise when multiple pieces of news
materialize on the same date, with some not relevant to alleged fraud. Finally, in section E we discuss a number of issues related to statistical power—i.e., how effective event studies used in securities fraud litigation are at detecting material fraud in the event that it actually did occur.

A. The Standard Approach to Testing for Statistical Significance in An Event Study, as Applied to Our Halliburton Results

As discussed in Part II, supra, tests of statistical significance all boil down to asking whether some statistic’s observed value is far enough away from some baseline level one would expect that statistic to take. For example, if one flips a fair coin 100 times, one should expect to see heads come up on roughly 50% of the flips, so the baseline level of the heads share is 50%. But of course one can also expect that there will be random variation in the heads share, such that it would be unreasonable to reject the null hypothesis that the coin is fair simply because one observes a heads share of, say, 49% or 51%. These results would not equal exactly the baseline level, but they are close enough such that the evidence against the hypothesis that the coin is fair is far too weak to reject the working hypothesis that the coin is unfair. On the other hand, common sense (and statistical methodology!) suggests that if 89% of 100 tosses yielded a heads, that would be strong evidence that the coin is biased toward heads. Accordingly, when the heads share is far enough from 50%, we say that there is statistically significant evidence against the null hypothesis that the coin is fair.

Event study tests of whether a stock price moved function very similarly to the coin-toss example. They all revolve around whether the stock’s excess return was more extreme than would reasonably be expected due to chance variation. What counts as a reasonable expectation? Typically courts and experts have assumed that there is statistically significant evidence of an event-date effect if the event-date’s excess return is among the 5% most extreme values one would expect to observe in the absence of any fraudulent activity. In this situation, experts might say that there is statistically significant evidence at the 5% level or, equivalently, “at level 0.05”, or “with 95% confidence.”

---

117 CITE BRAV & HEATON STUFF QUESTIONING THAT PRACTICE ___
Suppose that in the absence of any fraud-related event, it is known that Halliburton’s stock’s excess returns will follow a normal distribution. For a random variable that follows a normal distribution, 95% of realizations of that variable will take on a value that is within 1.96 standard deviations of zero, where the standard deviation is a measure of how spread out a large random sample of the variable is likely to be.\textsuperscript{118} The standard deviation of a firm’s excess returns is often estimated using the root mean squared error, a statistic that is usually reported by statistical software.\textsuperscript{119} Our Table 1 above reports that the root mean squared error for our Halliburton market model was 0.01745. Expressed in percentage terms, this means that under the assumptions supporting the model (discussed further below), the standard deviation of Halliburton’s excess stock return on a typical date within the class period is 1.745%.

Thus, on a typical date in the class period, the normality assumption for excess returns implies that, 95% of the time, Halliburton’s excess returns can be expected to fall between -0.0342 and 0.0342—in percentage terms, between -3.42% and 3.42%.\textsuperscript{120} Experts and courts will typically adopt a rule of regarding an event date estimated excess return as statistically significant at the 95% level if it both is further from zero than -3.42% or 3.42% and has the sign necessary to be consistent with the plaintiff’s allegations. For an alleged corrective disclosure, the estimated excess return would have to be negative and imply a price drop of more than 3.42%. That is, the estimated excess return would have to be less than -0.0342 for the court to find it statistically significant; thus,

\begin{itemize}
\item \textsuperscript{118} Formally, the standard deviation of random variable X is the square root of the variance, and the variance is the average value of the square of the deviation of the realized value of X from the average value of X: Variance equals the expected value of \((X-\mu)^2\), where \(\mu\) is the average value of X.
\item \textsuperscript{119} Excess returns in a market model have a mean of zero by definition, because the market model equation is allowed to have a non-zero intercept (this can be proved mathematically). Using the formula in footnote 118, supra, the variance of excess returns therefore equals the average value of the squared excess return from the market model. The root mean squared error reported by statistical software is the square root of an estimate of this average.
\item \textsuperscript{120} This follows because 1.96 times 1.745 equals 3.4202.
\end{itemize}
0.0342 is the critical value a court would adopt when faced with our event study results.\textsuperscript{121}

The logic of this inferential approach may be summed up using Figure 5 just below. The horizontal axis represents different values that a firm’s event-date estimated excess return might take on. When the estimated excess return is in the middle area, labeled “Not Very Unusual Value”, one does not reject the null hypothesis that the event date was typical of other dates—that is, the event date estimated excess return is treated as statistically insignificant. As we discuss more below, the standard approach to using event studies in securities fraud litigation has been to declare the event date estimated excess return statistically significant only if it lies in the “Very Unusual Value” shaded area that corresponds to the direction in which price should move given plaintiffs’ allegations. Thus, on an alleged corrective disclosure date, the standard approach entails finding statistical significance only if the estimated excess return is negative enough to fall in the “Very Unusual Negative Value” area.

\textbf{Figure 5: The Standard Approach to Testing Whether the Event-Date Effect is Positive}

\begin{center}
\begin{tikzpicture}
\draw[->, thick] (-1.5,0) -- (1.5,0);
\draw (0,0) -- (0,-0.5) node[below] {0};
\draw[fill=blue!20] (-1.5,0) rectangle (-0.5,0.5);
\node at (-1,-0.5) {Very Unusual Negative Value};
\draw[fill=blue!20] (0.5,0) rectangle (1.5,0.5);
\node at (1,-0.5) {Very Unusual Positive Value};
\draw[fill=blue!10] (-0.5,0) rectangle (0.5,0.5);
\node at (0,-0.5) {Not Very Unusual Value};
\end{tikzpicture}
\end{center}

In the first column of Table 3, we present the estimated excess returns from Table 2. The second column reports whether the estimated excess return is statistically significant based on the standard approach to testing described above. The event date estimated excess returns are statistically significant for December 7, 2001; August 9, 2001; and

\textsuperscript{121} As we discuss in ___ \textit{infra}, this two-sided approach to statistical significance testing is erroneous, because it amounts to the imposition of a confidence level of 97.5\% rather than the stated level of 95\%. 

42
December 4, 2001; they are statistically insignificant for the other three dates.

Table 3: Standard Significance Testing for Event Dates
(sorted by magnitude of estimated excess return)

<table>
<thead>
<tr>
<th>Event Date</th>
<th>Estimated Excess Return</th>
<th>Statistically Significant Using Standard Approach?</th>
</tr>
</thead>
<tbody>
<tr>
<td>December 7, 2001</td>
<td>-0.427</td>
<td>Yes</td>
</tr>
<tr>
<td>August 9, 2001</td>
<td>-0.051</td>
<td>Yes</td>
</tr>
<tr>
<td>December 4, 2001</td>
<td>-0.036</td>
<td>Yes</td>
</tr>
<tr>
<td>December 21, 2000</td>
<td>-0.012</td>
<td>No</td>
</tr>
<tr>
<td>October 30, 2001</td>
<td>-0.009</td>
<td>No</td>
</tr>
<tr>
<td>June 28, 2001</td>
<td>-0.008</td>
<td></td>
</tr>
</tbody>
</table>

We can illustrate the standard approach by again using a graph that relates actual and expected returns. As in earlier figures, Figure 6 again plots the actual return on the vertical axis and the expected return on the horizontal axis (with the set of points where these variables are equal indicated using an upwardly sloped straight line). This figure also includes dots indicating the expected and actual return for each day in the estimation period—these are the dots that cluster around the upwardly sloped line.

In addition, the figure includes three larger circles and three larger squares. The three circles indicate the alleged corrective disclosure dates for December 31, 2000; October 30, 2001; and June 28, 2001—the alleged corrective disclosure dates on which Table 3 tells us estimated excess returns were negative (below the upwardly sloped line), but were not statistically significant according to the standard approach. The three squares indicate the alleged corrective disclosure dates for which estimated excess returns were both negative and statistically significant. These are the three dates in the top three rows of Table 3—December 7, 2001; August 9, 2001; and December 4, 2001. We can tell that these dates have statistically significant drops in price because they appear in the shaded region of the graph; as discussed in relation to Figure 4, supra, points in this region have statistically significant price drops according to the standard approach.
The validity of the standard approach testing procedure in this context relies importantly on four assumptions:

1. Halliburton’s excess returns actually follow a normal distribution—that assumption is the source of the 1.96 multiplier for the standard deviation of Halliburton’s estimated excess returns in estimating the critical value.

2. It is appropriate to use a multiplier that is derived by considering what would constitute an unusual excess return in either the positive or negative direction—i.e., an unusually large
unexpected movement of the stock in either the direction of increase or the direction of decrease.

3. It is appropriate to analyze each event-date test in isolation, without taking into account the fact that multiple tests (six in our Halliburton example) are being conducted.

4. Halliburton’s excess returns have the same distribution on each date; under the first assumption, of normality, this is equivalent to assuming that the standard deviation of Halliburton’s excess returns is the same on every date.

As it happens, all of these assumptions are false in the context of the Halliburton litigation. The court did take appropriate account of the falsity of the third assumption, involving multiple comparisons, but it did not even address the other three. Yet violations of any of these assumptions will render the standard approach to testing for statistically significant unreliable as a matter of method. That is true even if in some cases these violations do not cause the standard approach and reliable approaches to yield different conclusions concerning statistical significance. Just as a stopped clock is right twice a day, an unreliable statistical method will yield the right answer sometimes. But the law demands more—it demands a method that yields the right answer as often as asserted by those using the method.

B. Violations of the Four Statistical Assumptions Undergirding the Standard Approach to Statistical Significance Testing

Here we consider violations of the four assumptions noted at the end of section A, supra.

1. Non-normality in Excess Returns

The market model may be written as

\[ H_t = \alpha + M_t' \beta + \varepsilon_t, \]

\[ 122 \text{ 309 F.R.D. 251, } ___ \text{ (2015).} \]
where \( t \) is the date; \( H_t \) is Halliburton’s daily stock return; \( \alpha \) is the intercept in the market model; \( M_t \) is a (column) vector of market index returns (such as the three included in our market model, discussed supra); \( \beta \) is a vector of coefficients that link these indexes’ returns to Halliburton’s return; and \( \epsilon_t \) is the excess return for date \( t \).

Gelbach, Helland and Klick (“GHK”) show that under the null hypothesis that nothing unusual happened on the event date, the estimated excess return for a single event date will have the same statistical properties as the actual excess return for that date.\(^{123}\) This result has two critical implications for present purposes. First, ignoring non-normality will yield unreliable statistical inference. Second, this disease has a simple cure.

We consider the disease first. If Halliburton’s excess returns are not normally distributed, then the estimated excess return for the event date will also not be normally distributed. This means that there is no justification to use a critical value based on the normal distribution for testing statistical significance.

And there is no good reason to assume that excess stock returns are actually normally distributed. There has long been considerable evidence against normality.\(^{124}\) Stocks’ excess returns often exhibit empirical evidence of skewness, “fat tails”, or both; and neither of these features would occur if excess returns were actually normal.\(^{125}\) Using the estimated model described above, we found that the same is true for Halliburton in our class period: the excess returns distribution is highly non-normal over this period. To illustrate, consider Figure 7. The figure’s solid line is a plot of the estimated probability density curve for

---

\(^{123}\) Jonah B. Gelbach, Eric Helland, and Jonathan Klick, *Valid inference in single-firm, single-event studies*, ___ Am. L. Econ. Rev. ___ (2013). This result holds provided that the number of dates used to estimate the market model is large; in the present case, there are 593 such dates, which is surely large in the relevant sense.


\(^{125}\) The existence of skewness indicates, roughly speaking, that the distribution of returns is weighted more heavily to one side of the mean than the other; the existence of fat tails—formally known as kurtosis—indicates that extreme values of the excess return are more likely, in either direction, than they would be under a normal distribution.
Halliburton’s estimated excess returns, while the dashed line is a plot of the density function for the normal distribution that has mean zero and the same standard deviation as Halliburton’s estimated excess returns over the estimation period.

The figure shows that the point with the highest density occurs where Halliburton’s excess return is greater than 0, which indicates negative skewness; summary statistics indicate likewise, so that the distribution of Halliburton’s excess returns has a longer “left tail”, with more probability concentrated among positive excess returns than among negative ones. Further, the distribution exhibits excess kurtosis, which means that values far from the distribution’s center are more likely, in both directions, than would be the case if excess returns were normally distributed. Formal statistical tests reinforce the story told by the graph above: Halliburton’s estimated excess returns systematically fail to follow a normal distribution over the estimation period.

Given the strong evidence against normality, both in general and in this instance, it is inappropriate to admit into evidence event studies that rely on normality. Federal Rule of Evidence 702(c) requires that expert testimony be “the product of reliable principles and methods.” And the Supreme Court stated in Daubert that when a court considers expert testimony using a scientific technique, “the court ordinarily should consider the known or potential rate of error” of that technique.127

---

126 To test for normality, we used tests discussed by Ralph B. D’Agostino, Albert Belanger and Ralph B. D’Agostino, Jr., A Suggestion for Using Powerful and Informative Tests of Normality, 44 The American Statistician 316 (Nov., 1990), implemented by the statistical software Stata via the “sktest” command. This test rejected normality with a confidence of 99.98%, due primarily to the distribution’s excess kurtosis.

In light of both known theoretical statistical results and empirical evidence, the standard approach of using tests that assume excess returns are normally distributed fails both of these criteria. Thus the standard approach is an unreliable method for conducting a test of statistical significance when estimated excess returns might not be normal, violating Rule 702(c). Moreover, without an assessment of whether estimated excess returns are non-normally distributed, and the extent to which such non-normality matters empirically, an expert who uses the standard approach cannot determine the approach’s rate of error, which contravenes Daubert. Empirical evidence shows that the theoretical concerns related to assuming normality are practically relevant.

GHK’s second result is that there is a simple and reliable alternative that (i) does not rely on normality, and (ii) requires no assumptions in addition to those made in standard event studies.\textsuperscript{128} The alternative is

\textsuperscript{128} GHK, note 123, \textit{supra}, at ___.

Figure 7: Estimated Density of Halliburton Excess Returns, Compared to Normal Distribution
called the “SQ” test, since it uses a statistic called the sample quantile. Before we describe the SQ test procedure’s steps, we will provide a simple intuitive explanation that helps illustrate the test and explain why it works.

Given the asymmetry in the distribution of excess returns apparent in Figure 7, finding critical values that delineate the middle 95% of the excess returns distribution is a bit more complicated than it would be if the distribution were symmetric, as when we inappropriately assume normality. We will need to separately find a “left tail” critical value and a “right tail” critical value, such that (i) 2.5 percent of the estimated excess returns in the class period are to the left of the left tail critical value and (ii) 2.5 percent of the estimated excess returns are to the right of the right tail critical value. The result will be that 95 percent of estimated excess returns lie between these two critical values. Thus, among event dates typical of the dates included in the estimation of the market model, estimated excess returns will lie in between the left and right tail critical values 95 percent of the time. Our left tail critical value is \(-0.0376\).

In Figure 8, we plot the estimated excess returns for the estimation period on the vertical axis. The figure’s horizontal axis shows each such value’s percentile rank among all trading days in the estimation period. The two values where we have drawn vertical lines are the 2.5th and 97.5th percentiles of the distribution, i.e., the respective values such that 2.5 percent of estimated excess returns lie to the left and 97.5 percent lie to the right. We choose these values since 95 percent of estimated excess returns fall between them (there is equal probability of falling either to the left of the left tail critical value or to the right of the right tail critical value, as with the standard approach assuming normality).

There are 593 days in the estimation period; 14/593=0.0236, and 15/593=0.0253. Thus the 2.5th percentile is somewhere between the 14th and 15th most negative estimated excess returns. However with such a large number of dates in the estimation period, these values (\(-0.03759558\) and \(-0.03758176\)) are so close that at any reasonable level of precision we would obtain the same left tail critical value using either of them, or using any value between them. Our right tail critical value, which plays no role in the rest of our analysis, is 0.0347 (the 14th and 15th greatest estimated excess returns are 0.03478797 and 0.03456537, so to four digits of precision, the 97.5th percentile is between 0.0348 and 0.0346).
Thus, when we drop the normality assumption and instead allow the distribution of estimated excess returns to drive our choice of critical values directly, we should conclude that an alleged corrective disclosure date’s estimated excess return is statistically significant if it is less than -0.0376 (i.e., indicates a price drop of greater than 3.76%). Recall that the standard approach critical value for an alleged corrective disclosure date was -0.0342. Thus, for Halliburton, the SQ test will yield a finding of statistical significance less often than would the standard approach based on the assumption of normality: an event date excess return of -3.43% is sufficient for a finding of statistical significance using the standard approach, but only an excess return of at least -3.77% will be sufficient when using the corresponding SQ test. This difference of 0.3 percentage points is not as small as it might seem at first blush. Halliburton’s median daily market value was $17.6 billion over the estimation period, so 0.34% of the firm’s median value amounts to roughly $60 million: the difference in critical values when we account for non-normality via the
SQ test amounts to requiring Halliburton’s market value to have dropped by $60 million more on an alleged corrective disclosure date to find a significant effect.

Table 4 repeats the estimated excess returns for the six event dates in the *Halliburton* litigation, this time using the SQ test’s left tail critical value of -0.0376 in place of the standard approach critical value of -0.0342. The second column of the table shows that the statistical significance result for December 4, 2001, is overturned as a result. This happens because the December 4, 2001, estimated excess return of -0.036 lies in between the standard approach and SQ test left tail critical values. This example illustrates that dropping the normality assumption might well matter in practice.

<table>
<thead>
<tr>
<th>Event Date</th>
<th>Estimated Excess Return</th>
<th>Statistically Significant Using SQ Test?</th>
</tr>
</thead>
<tbody>
<tr>
<td>December 7, 2001</td>
<td>-.427</td>
<td>Yes</td>
</tr>
<tr>
<td>August 9, 2001</td>
<td>-.051</td>
<td>Yes</td>
</tr>
<tr>
<td>December 4, 2001</td>
<td>-.036</td>
<td><strong>NO</strong></td>
</tr>
<tr>
<td>December 21, 2000</td>
<td>-.012</td>
<td>No</td>
</tr>
<tr>
<td>October 30, 2001</td>
<td>-.009</td>
<td>No</td>
</tr>
<tr>
<td>June 28, 2001</td>
<td>-.008</td>
<td>No</td>
</tr>
</tbody>
</table>

It is useful to observe that the SQ test involves estimating the exact same market model as the standard approach does. It requires only the trivial additional step of sorting the estimated excess return values for the class period in order to find the critical value—something that statistical software packages can do in one easy step in any case. Finally, the SQ test adds no assumptions not relied on by the standard approach. Thus any time the standard approach is reliable—i.e., when normality and the other assumptions addressed here are all satisfied—the SQ test is reliable as well.

---

130 See GHK, note 123, *supra*, at ___.

---

51
In sum, (i) the standard approach is generally not reliable due to the non-normality of excess returns, but (ii) the SQ test is reliable even with non-normality, provided the other assumptions necessary for the standard approach are satisfied, and (iii) the SQ test adds virtually no additional work for an expert or court to understand.

There is simply no reason to use the standard approach, but not to use the SQ test. And there is good reason to use the SQ test.

2. The Inappropriateness of Two-Sided Tests (With or Without the Normality Assumption)

The approach described so far requires a critical value such that 95% of all excess returns can be expected to be closer to zero than the critical value, regardless of whether these excess returns are positive or negative. This approach makes sense when an expert would conclude that an effect of the alleged type exists on a date when a firm’s stock either rose much more than would be expected or dropped much more than would be expected.

But despite the fact that it is widely used by experts, this two-sided testing makes no sense in the securities litigation context. A disclosure counts as a corrective disclosure for legally relevant purposes only if it causes the firm’s stock price to fall. It would be bizarre to reject the null hypothesis that a disclosure had no legally relevant impact on Halliburton’s stock if the disclosure occurred on a date when Halliburton’s stock shot up by an unusually high amount. No reasonable person would regard an unusual increase in firm value as evidence that the market had reacted negatively to newly disclosed information.  

Consider again Figure 5, repeated here as Figure 9 for convenience.

---

131 Obviously it would make just as little sense to conclude that Halliburton’s stock rose as a result of a fraudulent misrepresentation on a date when the company’s stock dropped by a surprising amount.
Suppose we are testing the statistical significance of the estimated excess return on an alleged corrective disclosure date. On such a date, the plaintiff’s allegation is that price fell due to the revelation of earlier fraud. As noted, a finding that the date had an unusually large and positive excess return on that date would certainly not be credited to the plaintiff by the court. That is why only estimated excess returns that are large and negative are treated as statistically significant for proving price impact on an alleged corrective disclosure date.

But recall that the critical value was chosen so that on a typical date, the excess return would fall in the “Not Very Unusual Value” range 95 percent of the time, and in each “Very Unusual Value” area 2.5 percent of the time. Since only negative estimated excess returns will be treated as supporting an allegation of a corrective disclosure, a finding of statistical significance would occur only 2.5 percent of the time when there was actually not an unusual drop in price—not the 5 percent that courts and experts say they are attempting to enforce. That means that the confidence level demanded of plaintiffs is 97.5 percent—not 95 percent, as stated. To put it differently, this is equivalent to requiring that Type I errors happen half as often as courts have stated they want to require.

While a reduction in Type I errors is desirable with all else held equal, other things are decidedly unequal. That is because the power of a test falls—possibly precipitously—as the Type I error rate demanded of it is reduced. In other words, using a confidence level of 97.5 percent in deploying event studies may induce many more false negatives than using a confidence level of 95 percent with an otherwise equivalent test.
of statistical significance.\footnote{We discuss power implications of flubbing the one-sided/two-sided issue in Part III.E, infra \_\_\_.} A method that delivers many more false negatives than the method that courts say they require surely raises important Daubert and Rule 702 concerns.\footnote{The fact that two-tailed tests are erroneous has been noted in recent literature. \textit{See}, \textit{e.g.}, Fox \_\_\_ and Brav & Heaton \_\_. Those authors seem to take for granted that courts will continue to use a method that is twice as demanding of plaintiffs as the method that courts say they require. We see no reason why courts should allow such a state of affairs to continue, especially one that is so easy to remedy.}

Fortunately, it is trivially easy to correct this mistake. Rather than use the 2.5\textsuperscript{th} percentile of estimated excess returns to test for statistical significance on an alleged corrective disclosure date, one simply uses the 5\textsuperscript{th} percentile of this distribution.\footnote{And if one were testing for the statistical significance on the date of an alleged misrepresentation that would be expected to move the market, one would use the 95\textsuperscript{th} percentile of the estimated excess returns distribution rather than the 97.5\textsuperscript{th}.} In our application, the 5\textsuperscript{th} percentile is -0.0308 (an excess return equivalent to a price drop of 3.08\%).\footnote{The 29\textsuperscript{th} and 30\textsuperscript{th} most negative estimated excess returns are -0.03089066 and -0.03074954, whose midpoint is -0.0308201; the shares of estimated excess returns less than or equal to these values are 4.89\% and 5.06\%, so their midpoint is an appropriate estimate of the 5\textsuperscript{th} percentile.} This is less demanding than the left tail critical value of -0.0376 we calculated when using the two-sided SQ approach. Consequently, an appropriate switch to the one-sided test will cause us to correct an erroneous finding of no statistical significance whenever the estimated excess return is between -0.0376 and -0.0309; this is a range of Halliburton market value of nearly $120 million, which is obviously of substantive interest.

In the last column of Table 5, we report the results of applying the one-sided SQ statistical significance test to our six \textit{Halliburton} alleged corrective disclosure dates. The first two columns repeat the results of statistical significance testing using the (inappropriate) standard approach and the (inappropriate) two-sided SQ test, both taken from Table 4, \textit{supra}. Once again there is no difference in the significance conclusions we draw for five of the six dates. For the December 4, 2001, event however, the switch to one-sided SQ testing would lead us to conclude that there was a statistically significant price impact, whereas the two-sided SQ test approach would have missed this effect. As with the switch from the (two-sided) standard approach to the two-sided SQ
test, this example shows that correcting an erroneous approach to statistical significance testing can make a practical difference.

**Table 5: Significance Testing for Event Dates Using Standard Approach and Both Two-Sided and One-Sided SQ Test**

<table>
<thead>
<tr>
<th>Event Date</th>
<th>Estimated Excess Return</th>
<th>Statistically Significant Using:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Standard Approach</td>
</tr>
<tr>
<td>December 7, 2001</td>
<td>-.427</td>
<td>Yes</td>
</tr>
<tr>
<td>August 9, 2001</td>
<td>-.051</td>
<td>Yes</td>
</tr>
<tr>
<td>December 4, 2001</td>
<td>-.036</td>
<td>YES</td>
</tr>
<tr>
<td>December 21, 2000</td>
<td>-.012</td>
<td>No</td>
</tr>
<tr>
<td>October 30, 2001</td>
<td>-.009</td>
<td>No</td>
</tr>
<tr>
<td>June 28, 2001</td>
<td>-.008</td>
<td>No</td>
</tr>
</tbody>
</table>

It is true that the standard approach and the one-sided SQ test arrive at the same conclusions for all six dates. However, that is merely happenstance. It is easy to imagine examples in which that would not occur. Moreover, we deliberately chose to discuss the erroneous assumptions of non-normality and two-sided testing in that order so that we could illustrate that each may be of practical relevance. In other situations—with other firms, having other alleged corrective disclosure date price impacts—the standard approach might yield quite different conclusions from the one-sided SQ test approach.

For convenience and comparison’s sake, Table 6 summarizes the steps necessary to implement the standard approach and the one-sided SQ test approach. As the table shows, the only difference between the two methods is in step 3, concerning how the test’s critical value is calculated. The standard approach relies on the assumed normality of the estimated excess returns distribution to determine this critical value, and uses the wrong level of confidence. By contrast, the SQ test approach makes no assumptions concerning the structure—i.e., normality or lack
thereof—of this distribution, and the one-sided version of the test uses the appropriate level of confidence by construction.\footnote{The reason the SQ test works—the reason that the standard approach’s normality assumption may be jettisoned—is that the critical value necessary for testing the null hypothesis of no event date effect is simply the 95\textsuperscript{th} percentile of the true excess returns distribution. Due to an advanced statistics result known as the Glivenko-Cantelli theorem, the percentiles of this distribution—also known as quantiles—may be appropriately estimated using the sorting method described in footnote \_\_, supra. See GHK, note 123, \textit{supra}, for details.}

Table 6: Comparing the Standard Approach and the SQ Test Approach for Testing the Null Hypothesis of No Event Date Effect (When the Alternative Hypothesis is a Positive Effect)

<table>
<thead>
<tr>
<th>Step</th>
<th>Standard Approach</th>
<th>SQ Test Approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Estimate market model</td>
<td>Use ordinarily least squares estimate of regression of Halliburton’s daily stock return on daily return of market index variable(s). (Use only data for a period that does not include any event dates in question.)</td>
<td>Same as standard approach.</td>
</tr>
<tr>
<td>2. Calculate estimated excess return for event date</td>
<td>The event date’s estimated excess return equals the actual event-date return for Halliburton minus the predicted value based on the coefficients from the market model from step 1.</td>
<td>Same as standard approach.</td>
</tr>
<tr>
<td>3. Calculate critical value</td>
<td>Critical value is product of -1.96 and standard error of regression from market model.</td>
<td>Using estimated coefficients from the market model, calculate the estimated excess return for all dates in the estimation period from step 1. Then find the 5\textsuperscript{th} percentile of these estimated excess returns. This number is the critical value.</td>
</tr>
<tr>
<td>4. Determine whether to reject null hypothesis if event-</td>
<td>Reject null hypothesis if event-</td>
<td>Same as standard approach,</td>
</tr>
</tbody>
</table>
Leaving aside the other two assumptions to be discussed momentarily, having to do with multiple comparisons and the stability of the excess returns distribution, the one-sided SQ test has important reliability characteristics. First, provided that the number of observations in the estimation period is reasonably large, the SQ test will erroneously reject a true null hypothesis 5 percent of the time, as intended. 137 This property of the SQ test satisfies the reliability requirements in both Rule 702(c) and Daubert: the principles and methods underlying it are based on widely known results in statistical theory, and according to these same theoretical results, the error rate of the SQ test is known to equal the intended rate of 5%. 138 Second, by avoiding the mistake of using a critical value based on two-sided testing logic, the SQ test generally will have greater power than the standard approach. Third, even if the standard approach were modified to use a one-sided critical value, the one-sided SQ would have the same power, when the comparison is done properly. 139 Thus we emphasize that there are circumstances in which the

137 To be entirely precise, the SQ test approach described in the text will erroneously reject a true null hypothesis with probability that becomes ever closer to 0.05 as the number of observations in the estimation period grows. This is an example of an asymptotic result, according to which the probability limit of the erroneous rejection probability precisely equals 0.05. Contemporary econometrics is dominated by a focus on such asymptotic results; see ___. GHK show that the SQ test performs extremely well even when using estimation period sample sizes considerably lower than the 250 days used here; see GHK, note 123, supra.

138 See footnote 137, supra, for the technical qualifier to this statement.

139 Statistical power is the probability of correctly rejecting a false null hypothesis—here, correctly rejecting the null hypothesis that Halliburton’s stock did nothing unusual on the event date. We have already considered an alternative measure of a test’s reliability, namely the probability that the test will erroneously reject a true null hypothesis. That measure, known as the size of a statistical test, is the one on which we focused in the text supra (because it is the focus of most discussions in the statistical event study literature, as well as in court). To compare two testing procedures’ power appropriately, one must first ensure that the comparison is done in a way that accounts for any differences in test size. To put things simply, tests that tend to reject true null
one-sided SQ test will perform much better than the standard approach, and there are no circumstances under which it will do worse.

3. **Multiple event dates of interest**

As we have seen, there are six alleged corrective disclosure dates at issue in the *Halliburton* litigation. This fact raises important questions, because the approach to testing for statistical significance discussed above is tailored to the case in which a statistical significance judgment must be made about only one event. When there are multiple events at issue statistical significance testing methodology must be appropriately modified. How to modify it depends on how the multiple events relate to the question of interest in litigation.

We first address the actual situation in *Halliburton*, where a finding of statistical significance on any one of the alleged corrective disclosure dates would be sufficient\textsuperscript{140} to certify a class at least as to issues related to that date. We then consider a different situation, in which plaintiffs would have to show price impact both on an alleged corrective disclosure date and on the date of a non-confirmatory alleged misrepresentation that could be expected to move the market price. These two situations turn out to be polar opposites, in the sense that the first requires using more demanding critical values, while the second requires using less demanding ones.

a. **When the question of interest is whether any event date had an unusual effect**

Suppose that a plaintiff alleges two separate corrective disclosure dates, and that the court will certify a class action if the plaintiff is able

\begin{itemize}
  \item \textbf{hypotheses a lot generally will also reject false null hypotheses a lot. Therefore, power tends to be greater when size is greater, and vice-versa.}
  \item The discussion \textit{supra} shows that even when the desired test size is 0.05—a 5\% chance of erroneously rejecting a true null hypothesis—the actual size of the standard approach will vary with the shape of the distribution of Halliburton’s excess returns. By contrast, we have seen that the actual size of the SQ test is 0.05, as intended. Thus it does not make sense to compare the power of the standard approach and SQ test without first accounting for the potentially different size of the standard approach.
  \item When one does that, one can show that the size-corrected power of the two tests is the same; \textit{see} GHK, note 123, \textit{supra}.
\end{itemize}

\textsuperscript{140} And was found to be sufficient; \textit{see} 309 F.R.D. 251 (2015).
to show statistically significant price impact on either or both of those dates. By design, a test that has 95% confidence will yield an erroneous finding of statistical significance 5% of the time when the event date was actually a typical one. Suppose we use the same critical value for each of two dates that we would use if there were only one date, concluding that either event date price impact is statistically significant if that date’s estimated excess return is more negative than the one-sided critical value based on the SQ test, as described in part III.B.2, supra. The multiple testing problem arises because with this approach, the probability that at least one of two tests will yield a finding of statistical significance, when the two dates were actually typical, is quite different from the intended probability of 0.05.

To see why, let A and B be two possible occurrences, each occurring with probability $p$. The probability that at least one of A and B actually occurs equals the probability that A occurs, plus the probability that B occurs, minus the probability that both occur at once.\(^{141}\) Thus, the probability that at least one of A and B occurs equals $2p$ minus the probability that both occur. If A and B are statistically independent, then the probability that both occur will equal $p^2$. Thus, the probability that at least one of them occurs is $2p - p^2$. Now observe that if A and B are the potential occurrences that our statistical significance test will erroneously find that there were significant price impacts on each of two dates, when there really was no impact on either, then $p$ equals 0.05. Assuming that the two event dates’ estimated excess returns are statistically independent of each other,\(^{142}\) the probability that at least one of date 1’s estimated excess return will be found statistically significant is 0.0975—nearly twice the intended probability of 0.05.\(^{143}\)

This is the multiple testing problem in a nut shell: the more tests one does while using the same critical value, the more likely it is that at least one test will yield a finding of significance even when there truly was no price impact. More event dates simply means more bites at the same apple, and the odds the apple will be eaten up rise with the number of

\(^{141}\) One must subtract this last probability to avoid double-counting such instances.

\(^{142}\) Because the estimated excess returns are computed using coefficients estimated from the same market model, this assumption will not generally be true. However, if each date’s true excess return is independent of other dates’ returns, as is usually assume, the amount of dependence will usually be extremely small. (But see section ___. infra, discussing the possibility of dependence in excess returns.)

\(^{143}\) This is simply the result of computing $2p-p^2=2\times0.05-0.05^2$. 
bites. With six dates, the effective confidence of testing this way falls to less than 74 percent, because there is more than a 26 percent chance that at least one date will have a price impact declared statistically significant even when there is no price impact on any of the dates. In the *Halliburton* litigation, defendants raised this issue, and it has played a substantial role.

Various approaches exist to correcting for this problem. Some of them solve the Type I error rate problem at the cost of substantially increasing the Type II error probability—i.e., substantially reducing the power of the test to detect price impact where it actually occurred. As it involves some fairly technical mathematical details, we will not discuss multiple testing methodology in any detail. Instead, we will simply describe a method that was developed to mitigate the low-power problem. This method, called the Holm-Bonferroni p-value correction, was adopted by the district court in *Halliburton*.

We must first explain what a p-value is for a single test, i.e., the type of test discussed before we began discussing the multiple testing problem in this section. Let \( \hat{\epsilon} \) be the estimated excess return for a given alleged corrective disclosure date. What we will call the “usual p-value” associated with \( \hat{\epsilon} \) is the probability that one would observe an excess return at least as negative as \( \hat{\epsilon} \) on a typical date (i.e., one with no actual price impact). There is a close relationship between the usual p-value and the result of statistical significance testing of the form described above: if the estimated excess return was found statistically significant at the 0.05 level, then the usual p-value for that date must be less than or equal to 0.05; if the estimated excess return was found not statistically significant, then the usual p-value must be above 0.05.

It is very simple to calculate the usual p-value for an alleged corrective disclosure date when using the one-sided SQ test. To do so, one counts up the number of estimated excess returns from the market

\[144\] With six dates, as in the *Halliburton* litigation, it can be shown that even when there is no price impact on any date, a collection of independent tests intended to have a Type I error rate of 0.05 will yield at least one finding of statistical significance more than 26 percent of the time. Considering all six tests together as a single “family-wise” test, this is a radical inflation of the Type I error rate.


model estimation period that are more negative than the estimated excess return on the event date, and then one divides by the number of dates included when estimating the market model (593 in our Halliburton example). We report the usual p-value for each alleged corrective disclosure date in the second column of Table 7; the third column reports whether price impact was found statistically significant using the one-sided SQ test. As claimed, the usual p-value is less than 0.05 for all three dates with statistically significant price impacts, and greater than 0.05 for the other three.

Table 7: Controlling for Multiple Testing
Using the Holm-Šidák Approach

<table>
<thead>
<tr>
<th>Event Date</th>
<th>Excess Return</th>
<th>p-value</th>
<th>Statistically Significant?*</th>
<th>One-Sided SQ Approach, Ignoring Multiple Testing Issue</th>
<th>One-Sided SQ Approach, With Holm-Šidák Correction for Multiple Testing</th>
</tr>
</thead>
<tbody>
<tr>
<td>December 7, 2001</td>
<td>-.427</td>
<td>0</td>
<td>Yes</td>
<td>0</td>
<td>Yes</td>
</tr>
<tr>
<td>August 9, 2001</td>
<td>-.051</td>
<td>0.0017</td>
<td>Yes</td>
<td>0.0034</td>
<td>Yes</td>
</tr>
<tr>
<td>December 4, 2001</td>
<td>-.036</td>
<td>0.0269</td>
<td>Yes</td>
<td>0.0787</td>
<td>NO</td>
</tr>
<tr>
<td>December 21, 2000</td>
<td>-.012</td>
<td>0.2222</td>
<td>No</td>
<td>0.6340</td>
<td>No</td>
</tr>
<tr>
<td>October 30, 2001</td>
<td>-.009</td>
<td>0.2609</td>
<td>No</td>
<td>0.7795</td>
<td>No</td>
</tr>
<tr>
<td>June 28, 2001</td>
<td>-.008</td>
<td>0.3013</td>
<td>No</td>
<td>0.8837</td>
<td>No</td>
</tr>
</tbody>
</table>

The fourth column of the table reports Holm-Šidák p-values, which are corrected for multiple testing. The final column reports whether the Holm-Šidák p-value is less than 0.05, in which case there is statistically significant price impact, provided that the p-values for other

148 Let date \( m \) be the event date whose usual p-value is lower than exactly \( m \) other dates' usual p-values. Thus, December 7, 2001, is date 5; December 4, 2001, is date 4, and so on, with June 28, 2001, being date 0. Let \( p_m \) be the usual p-value for date \( m \). Then the Šidák p-value equals \( 1 - (1 - p_m)^{6-m} \).

We note also that for small values of \( p_m \) and small values of the exponent \( 6-m \), Šidák p-values are well approximated by \( (6 - m)p_m \), which is known as the Bonferroni p-value (subject to the caveat that it is set to 1 if the calculation would otherwise yield a value greater than 1). In our application it does not matter which we use, though in general the Šidák p-value is more accurate than the Bonferroni p-value.
Correcting for multiple testing via the Holm-Šidák procedure yields significant price impacts for December 7, 2001, and August 9, 2001, but not for the other four dates. Thus, relative to the one-sided SQ test that does not correct for multiple tests, the effect of correcting for multiple tests is to convert the finding of statistical significance for December 4, 2001, to a finding of insignificance.

b. When the question of interest is whether both of two event dates had an effect of known sign

Now consider the situation in which the plaintiff alleges that the defendant made a misrepresentation involving non-confirmatory information on Date 1, and then issued a corrective disclosure on Date 2. A court will require the plaintiff to show price impact for both Date 1 and Date 2. This situation differs from the one just considered, because there it was sufficient for the plaintiff to show price impact as to any of multiple dates; here the plaintiff must show price impact for both dates. In this situation, the appropriate statistical threshold for finding price impact statistically significant is less demanding, since it must be met for both dates.

First consider what would happen if we used a traditional one-sided test for each date separately, using the 95% level of confidence for each. For each day, we have seen that the probability of finding statistical significance when there was no actual price impact is 0.05—1 in 20. Because these significance tests are roughly independent, the probability that both tests will reject when each null hypothesis is true is only 1 in 400, i.e., one-quarter of one percent. To put it differently,

---

149 That is, we consider date 5’s price impact to be statistically significant if its Šidák p-value is less than 0.05. If date 5’s price impact is not statistically significant, then we consider all dates’ price impacts to be insignificant. If date 5’s price impact is significant, then we turn to considering date 4’s price impact, considering it significant if date 4’s Šidák p-value is less than 0.05; if not, we stop, but if so, we turn to date 3’s price impact, and so on.

150 They have a bit of dependence given that excess returns are calculated from the same market model. However, this dependence can be shown to vanish as the number of dates in the estimation period grows, and with 593 we would expect very little to persist.

151 This is the case because 0.05×0.05=0.0025; to put it differently, 0.05 equals 1/20, and 1/20 times itself is 1/400, which is one-fourth of 1/100—or a quarter of a percent.
requiring each date to separately meet the 95% confidence level for a finding of statistical significance is equivalent to requiring 99.75% statistical confidence overall.\textsuperscript{152} This is obviously a much more demanding standard than the 95% level that courts and experts say they are using. Thus this approach is an unreliable way to test with 95% confidence, and the approach should fail both Rule 702(c) and Daubert.

To develop a reliable test, we can again work with the usual \(p\)-values. For an overall \(p\)-value equal to 0.05—again, corresponding to the standard that experts and courts say they want to apply—we should determine that price impact is significant on both days if each date has a usual \(p\)-value of less than roughly 0.2236.\textsuperscript{153} Using the one-sided SQ approach, as we have argued one should, this means that

\begin{itemize}
  \item the estimated price impact for the alleged corrective disclosure date is statistically significant if the estimated excess return for that date is not greater than more than 22.36% of estimated excess returns in the estimation period; and
  \item the estimated price impact for the alleged misrepresentation date is statistically significant if the estimated excess return for the alleged corrective disclosure date is not less than at least 22.36% of estimated excess returns in the estimation period.
\end{itemize}

To put it differently, [INSERT AND DISCUSS NEW VERSION OF Figure 8 BUT WITH VERTICAL LINES AT 22.36% AND 77.64%]. This test has probability 0.05—confidence level 95%—of erroneously making a finding of statistical significance as to both dates considered together.

[MAKE POINT THAT DEC 21, 2000, EXCESS RETURN WOULD JUST MAKES THE CUTOFF, IF THERE WERE NO OTHER ALLEGED CORRECTIVE DISCLOSURE DATES. SO NEED TO

\textsuperscript{152} This is true because statistical confidence, as this term is usually used, equals the percentage equivalent of one minus the probability of falsely rejecting a true null hypothesis. Here that probability is 0.0025, so in percentage terms we have 100\% \times (1-0.0025), which is 99.75%.

\textsuperscript{153} This is true because the probability of finding that both tests have a usual \(p\)-value of \(q\) is \(q^2\). Setting this equal to 0.05 and solving for \(q\) yields \(q = 0.2236068\). Thus, we should declare the pair of price impact estimates if each has a usual \(p\)-value less than this level.
ASK READER TO IMAGINE WE HAVE A NON-CONFIRMATORY ALLEGED MISREPRESENTATION DATE WITH SUFFICIENTLY POSITIVE EXCESS RETURN]

Since both arms of our test would be met, the appropriate conclusion would be to find that there is statistically significant evidence of price impact for the alleged misrepresentation and alleged corrective disclosure involving December 21, 2000. This conclusion follows even though we would not find statistically significant evidence of price impact if December 21, 2000, were the only date of interest. This example illustrates the consequences of the appropriate loosening of the threshold for finding statistical significance when a party must demonstrate that something unusual happened on each of multiple dates.\textsuperscript{154}

c. What if there are multiple pairs of alleged misrepresentation and alleged corrective disclosure dates, with plaintiffs prevailing if they can show that any one pair has statistically significant evidence?

[THIS DISCUSSION INVOLVES THE FUN TASK OF COMBINING THE HOLM-Šidák APPROACH FROM PART a AND THE APPROACH FROM PART b. FOR EACH PAIR OF DATES, WE USE THE APPROACH FROM PART b, BUT ACROSS PAIRS, WE USE THE HOLM-Šidák APPROACH.]

\textsuperscript{154} We emphasize that this discussion does not necessarily carry the day as a matter of evidence law. Evidence law could take the view that for any single piece of statistical evidence to be credited, that single piece must meet the 95\% confidence standard even if that means that a party who must show 2 pieces of evidence is as a result actually held to a standard of 99.75\% confidence. We think such a view is indefensible, especially if courts choose to correct for the other type of multiple tests that we have discussed. But of course that position would be one of law rather than one related to the proper conduct of statistics as such. Our primary objective here is to ensure that judges, advocates, litigants, and experts understand the true content of such an approach. This is especially important for those situations in which there might be many more than just 2 dates in question. For example, if there were 5 dates, for example, then the true overall level of confidence required when a court requires the plaintiff to meet the 95\% confidence level separately for each date would be more than 999,999 in a million.
4. *Instability Over Time in the Distribution of the Excess Return*

For a typical event study to be probative, the behavior of the stock in question must be stable over the market model’s estimation period. For example, it must be true that, aside from the alleged events under study, the association between Halliburton’s stock and the broader market during the class period is similar to the relationship for the estimation period. If, for example, Halliburton’s association with its industry peers, or other firms in the broader market, differed substantially in the two periods, then it goes without saying that the market model would not be a reliable tool for predicting the performance of Halliburton’s stock on event dates, even in the absence of any actual misrepresentations or corrective disclosures.

A second requirement is that, aside from any effects of the alleged misrepresentations or corrective disclosures, excess returns on event dates must have the same probability distribution as they do during the estimation period. As we discussed in section 1, *supra*, the standard approach to estimating the critical value for use in statistical significance testing is based on the assumption that, aside from the effects of any fraud or corrective disclosure, all excess returns come from a normal distribution with the same standard deviation. Imagine that the date of an alleged corrective disclosure happens to fall during a time of unusually high volatility in the firm’s stock price—say, due to a spike in market uncertainty about demand in the firm’s principle industry. In that case, even typical excess returns will be unusually dispersed—and thus unusually likely to fall far from zero. Failing to account for this fact would lead an event study to find statistically significant price impact on too many dates, simply due to the increase in volatility.155

Event studies can be adjusted to deal with this problem, too. To do so, one must use a model that is capable of estimating the volatility of the event-date excess return both for dates used in the estimation period and for dates that are the object of the price impact inquiry. The details of doing so are fairly involved; they require a substantial amount of mathematical notation, as well as a discussion of some technical econometric issues. Accordingly, we relegate these details to Appendix A, which appears at the end of this Article. Here we give only the most

155 CITE BAKER PAPER ON FINANCIAL CRISIS. ___ ALSO ALLEN REPORT IN *Halliburton* LITIGATION AS TO DECEMBER 7, 2001.
general discussion of the methodology, together with the empirical results for our Halliburton application.

A notationally simple way to allow for differences across time in Halliburton’s excess return is to write it as follows:

$$\varepsilon_t = \sigma_t u_t,$$

where $\sigma_t$ is the standard deviation of the excess return on date-$t$—which might change across dates—and $u_t$ is known as “white noise”, because it is assumed to be independently and identically distributed (“iid”) across dates. In other words, regardless of how much one knows about the date-$t$ value of the standard deviation $\sigma_t$, one won’t know anything more about $u_t$. We further assume that $u_t$ has mean 0 and variance 1.156

The basic idea of our approach to dealing with heterogeneity in the distribution of $\varepsilon_t$ is simple. In Appendix A, we show how to estimate the date-$t$ variance, $\sigma_t$, as well as the excess return $\varepsilon_t$—even when the variance differs across dates. With estimates of both these variables in hand for each date, we then divide our estimate of $\varepsilon_t$ by our estimate of $\sigma_t$, which yields an estimate of $u_t$. We can think of $u_t$ as a standardized version of our excess return $\varepsilon_t$, because by definition each $u_t$ has standard deviation equal to 1, and since $u_t$ is iid, each standardized excess return may be viewed as a random draw from the same distribution.

Further, if an alleged corrective disclosure had no impact on Halliburton’s stock price, the estimated standardized excess return observed on that date—the estimated $u_t$ value—must have been drawn from the same distribution from which all other $u_t$ values were drawn. Consequently, we can use the SQ testing approach developed supra,157 but applying it to the estimated standardized excess return $\hat{u}_t$ rather than the estimated excess return $\hat{\varepsilon}_t$. In words, we will simply be asking whether Halliburton’s stock price fell by a surprisingly large number of units of standard deviation on alleged corrective disclosure dates, rather than by a surprisingly large amount. The only difference is that here, we

156 The mean of 0 actually follows as a consequence of the inclusion of an intercept in the market model. The variance of 1 is not a restrictive assumption; we could allow variance to be any fixed positive constant without altering any of the substance of our analysis.

157 See Part II.B.1, as well as refinements to accommodate one-sided testing and multiple inferences in Parts II.B.2 and II.B.3, respectively.
allow for the possibility that the standard deviation of Halliburton’s excess returns, $\sigma_t$, is not the same on every date $t$.

[INSERT FIGURE AND DISCUSSION ILLUSTRATING]

In Appendix A, we report estimates from an econometric model allowing for variation over time in the variance of Halliburton’s excess returns. The results indicate that this variance indeed varies over time, in at least three important ways. First, Halliburton’s excess returns have greater variance on days when the industry peer index returns have greater variance. Second, Halliburton’s excess returns are more variable on days when a measure of overall stock market volatility suggests this volatility is greater.\(^{158}\) Third, variance in Halliburton’s excess returns tends to be greater on days when it was greater the day before, and when Halliburton’s actual excess return was further from zero (whether positive or negative). Thus, it is potentially important to account for the evolving nature of variance in Halliburton’s excess returns.

Using the model estimates described in Appendix A, we estimated the standardized estimated excess returns described just above—i.e., the $u_t$ variables—for each date in our data. We repeated the test for normality that we carried out with respect to $\epsilon_t$ in Part II.B.1, and we found that the data resoundingly reject the null hypothesis that $u_t$ is

\(^{158}\) This market-level measure is known as the VIX and is published by the Chicago Board Options Exchange. It uses data on options prices, together with certain assumptions about the behavior of securities prices, to back out an estimate of the variance of stock returns for the day in question. It has been used in at least one event study in securities fraud litigation, see ___. Its use as a variance forecasting tool has recently been advocated in a student note by Andrew Baker, ___; we discuss Baker’s approach, and its implicit assumption that standardized excess returns are normally distributed, in Appendix A. Finally, we note that another recent paper suggests that when the assumptions about the behavior of securities prices, referred to above, are incorrect, the VIX index does not directly measure the variance of the market return; see Victor Chow, Wanjun Jiang, and Jingrui Li, Does VIX Truly Measure Return Volatility? (August 30, 2014), available at http://ssrn.com/abstract=2489345 (proving that the VIX index reliably measures the variance of the stock market only under certain assumptions, and offering a generalized alternative for use in its place). Because our mission here is illustrative only, however, there is no harm in using the VIX index itself; we note in addition that it appears to be several times less important in explaining the variance of Halliburton’s excess returns than is volatility in the industry peer index.
distributed normally. Accordingly, it is unreliable to base a test for statistical significance on the assumption that $u_t$ follows a normal distribution. Instead, we use the SQ test approach described in the previous sections.

In the first column of Table 8, we report a set of $p$-values for each of the six alleged corrective disclosure dates in the Halliburton litigation. These $p$-values were computed using the one-sided SQ test approach with no correction for multiple testing, but with our allowance for the possibility that the variance of Halliburton’s excess returns varies across dates. In other words, these $p$-values are comparable to those reported in the second column of Table 7—except that the ones in the first column of Table 8 take account of the evolving variance in Halliburton’s excess returns.

<table>
<thead>
<tr>
<th>Event Date</th>
<th>Assuming Constant Variance**</th>
<th>Assuming Constant Variance**</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Allowing for Possibility</td>
<td></td>
</tr>
<tr>
<td></td>
<td>of Evolving Variance*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$p$-value</td>
<td>Statistically Significant?</td>
</tr>
<tr>
<td></td>
<td>Single Test</td>
<td>Holm-Šidák</td>
</tr>
<tr>
<td>December 7, 2001</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>August 9, 2001</td>
<td>.002</td>
<td>.003</td>
</tr>
<tr>
<td>December 4, 2001</td>
<td>.010</td>
<td>.030</td>
</tr>
<tr>
<td>December 21, 2000</td>
<td>.256</td>
<td>.694</td>
</tr>
<tr>
<td>October 30, 2001</td>
<td>.317</td>
<td>.881</td>
</tr>
<tr>
<td>June 28, 2001</td>
<td>.298</td>
<td>.851</td>
</tr>
</tbody>
</table>

* Using GARCH model described in Appendix A.
** Repeated from Table 7.

While there is a bit of negative skew in the standardized estimated excess return, the test rejects normality primarily because of excess kurtosis—i.e., fat tails—in the standardized excess return distribution.

Baker, ___ does exactly this in his simulation study.
There is one date for which there is a notable difference: December 4, 2001. On this date, the single-test p-value that accounts for evolving variance is 0.010—considerably lower than the single-test p-value of 0.027 we reported in Table 7. The next column of Table 8 reports the Holm-Šidák p-value corrected for multiple testing, discussed in Part II.B.3, *infra*. Again the corrected p-values are generally quite similar to those in Table 7 for all but December 4, 2001. For that day, the corrected p-value—is just 0.030 when we account for evolving variance.\(^{161}\) Since this is less than 0.05, we conclude that the price impact on December 4, 2001, is statistically significant at the 5% level—with 95% confidence—once one appropriately accounts for evolving variance.

What drives this potentially important reversal? On December 4, 2001, our model yields an estimated \(\sigma_t\) value of 0.015—i.e., a standard deviation of 1.5 percent. This is lower than the root mean square error of 0.017 in the constant-variance model underlying Table 7, and that is part of the story. But there is more to it. Of the 16 estimation period dates that had a smaller estimated excess return than the one for December 4, 2001, when we assumed constant variance across dates, *all but one* had an estimated \(\sigma_t\) greater than 0.015 once we allowed for the possibility that \(\sigma_t\) might evolve over time. In some cases, the difference was quite substantial, and this is what is driving the very large change in the p-value for December 4, 2001.\(^{162}\)

Thus, our variance-estimation model tells us that the alleged corrective disclosure date of December 4, 2001, is a date on which the

---

\(^{161}\) The corresponding figure in Table 7 is 0.079—much greater.

\(^{162}\) For example, 5 of the 16 dates had estimated values of \(\sigma_t\) in excess of 0.023. While this might not seem like much of a difference, it is, because the standardized estimated excess return \(u_t\) is the ratio of the estimated excess return \(\varepsilon_t\) to the estimate of \(\sigma_t\). Dividing the December 4, 2001, estimated excess return by 0.015 while dividing these other 5 dates’ estimated excess returns by 0.023 is the same as raising the December 4, 2001, estimated excess return by a factor of more than 50%. To see this, observe that since \(u_t = \varepsilon_t / \sigma_t\) we have

\[
\frac{\varepsilon_{4\text{Dec2001}}}{\varepsilon_t} = \frac{u_{4\text{Dec2001}}}{u_t} \times \frac{0.023}{0.015} = 1.53 \times \frac{u_{4\text{Dec2001}}}{u_t},
\]

so that this constellation of estimated values of \(\sigma_t\) makes a very large difference in the relative value of the December 4, 2001, alleged corrective disclosure date’s standardized estimated excess return, by comparison to dates with very negative non-standardized estimated excess returns.
Halliburton excess return could be expected to have relatively low variance—not to bounce around too much for essentially random reasons. Consequently, even the relatively large price drop for that date—an estimated excess return of -0.0365—understates the extent of the effective price impact. Once we account for the low expected variability on December 4, 2001, we see that the price impact on that date was large enough so that it should have been regarded as statistically significant.

C. Best Practices for Experts Conducting Reliable Event Studies for Use in Securities Fraud Litigation

Taken together, the discussion in Part II.B illustrates how four oft-ignored statistical issues can lead experts—and courts relying on them—to reach unreliable conclusions when considering event study evidence as to price impact. Failure to account for non-normality of excess returns, adopting the illogical approach of two-sided statistical significance testing, failing to account for multiple inferences, and failing to account for evolution over time in the variance of excess returns—each of these mistakes can lead to consequential errors in determining whether there has been statistically significant price impact. To guard against these mistakes, we provide the following suggested approach to conducting event studies.

1. One or more relevant comparison indexes should be constructed for a security.

2. A regression model should be estimated that relates the daily return for the security in question to the comparison indexes, with estimated excess returns calculated from this model.

3. Any expert who assumes normality should be expected to conduct a formal statistical test of normality. If the test rejects normality, then the court should not admit any testimony related to statistical significance testing that relies on normality. Better yet, experts should simply use an appropriate version of the SQ test to determine critical values, since the SQ approach does not depend on normality—and is nevertheless justifiable in the presence of non-normality.
4. Significance testing should appropriately address multiple testing concerns. When multiple dates are alleged to involve corrective disclosures, the effective critical values for finding a significant price impact will be more demanding. On the other hand, when a plaintiff is expected to demonstrate price impact for both an alleged corrective disclosure date and for an alleged misrepresentation date, the effective critical values will often be less demanding. When both types of multiple inferences are involved, the net effect of correcting for multiple inferences will depend on the number of inferences involved.

5. The regression model should allow for heterogeneity in the variance of excess returns. The expert should allow for the possibility that such heterogeneity arises through multiple channels: observed volatility in any comparison indexes; the VIX index, which a measure of overall volatility based on option prices; and links across time in the variance and observed volatility of the security in question, which can be allowed via the inclusion of generalized autoregressive conditional heteroskedasticity components.

The discussion necessary to elucidate these suggested practices is unavoidably mathematical in places. However, each of these suggestions is simple to implement for anyone with the most basic statistical programming experience and a copy of statistical software such as the industry-standard package Stata. In other words, anyone remotely qualified to serve as an expert in securities fraud litigation involving millions—or even billions—of dollars should be able to implement these practices without breaking a sweat. Further, all of them involve familiar statistical principles. Finally, courts should find that failure to implement these practices constitutes a failure to use reliable principles and methods. Courts simply should not admit expert testimony, or consider reports, that do not follow these practices.

D. Dealing with Multiple Pieces of News on an Event Date

Statistical significance testing answers a very specific question in an event study: If nothing unusual had happened on an event date in question, would it have been extremely unlikely to observe an estimated
excess return as different from typical values as the event date’s is? There are (at least) two ways in which this question differs from the legally relevant question. First, the question event studies answer is not the same as asking whether the event in question is the cause of any unusualness in the stock’s estimated excess return. Thus, it is possible that (i) the stock did move an unusual amount on the date in question, but that (ii) some factor other than the event in question was the cause of that move. For example, suppose that on the same day that Halliburton made an alleged corrective disclosure, one of its major customers announced for the first time that despite Halliburton’s excellent performance, the customer was terminating activity in one of the regions where it uses Halliburton’s services.

Second, it is possible that the event in question did cause a change in stock price in the hypothesized direction, even when the estimated excess return on the event date of interest was not particularly unusual, because some other factor operated in the opposite direction. For an example of this situation, suppose that Halliburton made an alleged corrective disclosure on the same date that a major customer announced good news for the company. It is possible that the two pieces of news have roughly similar effects on Halliburton’s stock price at least up to the precision that appropriate statistical tests can provide. In that case, there will be no especially unusual change in the firm’s stock price—no unusual estimated excess return—even though, the corrective disclosure reduced Halliburton’s stock price ex hypothesi.

Both of these problems arise because an additional factor occurs at the same time as the legally relevant alleged event. If multiple unusual events—events that would affect the stock price even aside from any industry-wide developments—occurred on the event date, then even an event study that controls for market- or industry-level factors will still be problematic. Suppose our firm announced both favorable restructuring news and a big jury verdict against it on the same day. All a traditional event study can tell us is the net market response to these two developments. Without further refinement, it would not distinguish the sources of this response.

However, further refinements might still make an event study useful in such a situation. For example, if the two pieces of information were announced at different times on the same day, one might be able to use intra-day price changes to parse the separate impacts of the two
Here is where both the theory of and empirical evidence related to financial economics are especially important. Theory suggests that stock prices should respond rapidly in a public market with many traders paying attention to a well-known firm with many shares outstanding. After all, no one wants to be left holding a bag of bad news, and everyone can be expected to want to buy a stock whose issuer’s good news has yet to be reflected in price. These standard market factors can be expected to put immediate pressure on a firm’s stock price to move up in response to good news and move down in response to bad news. Empirical evidence suggests that financial economics theory is correct on this point: one widely cited study indicates that prices react within just a few minutes to public news related to stock earnings and dividends.

When multiple sources of news are released at the same time, no event study can, by itself, separate out the effects of the different news. The event study can tell us whether the net effect of all the news was associated with an unusually large price drop or rise, which could be useful if there is some way to disentangle the expected effects of the other types of news involved. For example, suppose that a firm announces bad regulatory news on the same day that it announces bad earnings news, with plaintiffs alleging only that the regulatory news constitutes a corrective disclosure. Experts might be able to use historical price and earnings data for the firm to estimate the relationship between earnings news and the firm’s stock price. If this study controlled appropriately for market expectations concerning the firm’s earnings—say, using analysts’ predictions—it might provide a plausible way to separate out the component of the event date’s estimated excess return that could reasonably be attributed to the earnings news, with the rest being due to the alleged corrective disclosure related to regulatory news.

In sum, while the release of multiple pieces of news on the same date certainly complicates the use of event studies to measure price impact, event studies might play a useful role even in those cases.

---

163 Alternatively, ___ HEATON & ___.
164 James M. Patell & Mark A. Wofson, The Intraday Speed of Adjustment of Stock Prices to Earnings and Dividend Announcements, 13 J. Fin. Econ. 223 (1984). This study is cited, for example, in the report of Halliburton’s expert witness Lucy Allen; see ___, Allen Report, No. 572-2, n. 93 at 43 (September 10, 2014).
E. Power in Event Studies Used in Securities Fraud Litigation

STILL TO ADD, BUT BASIC POINTS TO COVER:

- Discuss the fact that mistaken use of two-sided testing can make a large difference—larger than one would expect in comparison 0.05 and 0.025 (or 95% and 97.5%), because of the important differences between the relevant states of the world in Type I and Type II error analysis (i.e., whether the null is true or false).

- Brav & Heaton point that single-firm event studies can be expected to have low power, since they don’t involve averaging away noise.

- Related issue concerning bias and damages calculation (including from the ex ante perspective, which Brav & Heaton do not address).

- Broader issue related to the tradeoff between significance level and power. Make the point that this is an unavoidable fact of life. (An alternative would be Bayesian estimation, though that is difficult to explain, unconventional, sensitive to priors, and therefore might be unlikely to be adopted—though there are pro-Bayesian counterarguments.)

CONCLUSION

Event studies play an important role in securities fraud litigation. In the wake of Halliburton II, that role will only increase, as proving price impact has become a virtual requirement to successfully invoke the Basic presumption of reliance and secure class certification. This Article has explored a complex of considerations related to the use of event studies in such litigation. These considerations span the space of substantive securities law; civil procedure; financial economics; and statistical methodology. The proper use of event studies will require experts and courts to pay attention to a number of important ways in which the securities fraud litigation context differs from the empirical context of many academic event studies.

One set of issues has to do with the problem of disentangling the effects of multiple pieces of news released on the same date or, worse still, at the exact same moment. When such issues are present, all an
expert can hope to do with an event study is determine that something caused an unusual move in price on the date in question. Identifying the contribution of alleged misrepresentations will require additional evidence, perhaps from econometric studies using different types of data. Another set of issues involves the question of whether an event study could even be probative as to the date of an alleged misrepresentation, when the alleged effect is to confirm rather than improve market expectations. In such cases, courts have appropriately allowed plaintiffs to focus on showing price impact of the alleged misrepresentations indirectly via evidence relating to alleged corrective disclosure dates.

Another important point of this Article is that courts and experts using event studies in securities fraud litigation should pay more attention to a number of issues related to statistical methodology. First, because a litigation-relevant event study typically involves only a single firm, issues related to non-normality of a stock’s returns arise. Second, because the plaintiff must show either that price dropped or rose, but will never carry its burden if the opposite happened, experts should unquestionably be using one-sided significance testing rather than the conventionally deployed two-sided approach. Third, securities fraud litigation often involves multiple test dates, which has important and tricky implications for the appropriate level of date-specific confidence levels if the goal is an overall confidence level equal to the 95% level that courts and experts say it is. Finally, event studies must be modified appropriately to account for the possibility that stock price volatility varies across time.

Failure to address these issues should render event studies inadmissible as a matter of law under Daubert and Rule 702. Happily, each of these issues can be addressed using straightforward and simple techniques that do not increase the difficulty of conducting event studies.
APPENDIX A: ALLOWING FOR HETEROGENEITY IN THE DISTRIBUTION OF EXCESS RETURNS

In this appendix we relax the assumption that excess returns have the same distribution on every date. In doing so, we allow for two possible reasons why the variance of daily excess returns might vary across dates. The first is that variance of daily excess returns might be associated with contemporaneous and lagged variance in the daily returns of the industry index variables we use in the market model described in Part III, as well as an index of overall volatility in the stock market.\textsuperscript{165} The second source of heterogeneity follows a long literature that models the variance of securities using a model known as generalized autoregressive conditional heteroskedasticity ("GARCH").\textsuperscript{166} The GARCH model class allows the variance of a security’s return on date $t$ to be related both to the variance on earlier dates (this is the autoregressive part) and to the magnitude of earlier dates’ realized values of squared excess returns (this is the conditional heteroskedasticity part).

A. A Model that Allows for Heterogeneity in the Variance of Excess Returns

To accommodate these possible sources of variance heterogeneity, we write the daily excess return $\varepsilon_t$ as the product of two components:

$$\varepsilon_t = \sigma_t u_t,$$

\textsuperscript{165} This index is the VIX index published by the Chicago Board Options Exchange. In a forthcoming student note, Andrew Baker has shown that this index, sometimes known as the "fear index", is associated with financial sector firms’ daily returns, and that modeling variance to take account of movements in the VIX shows there is significant daily heterogeneity in variance for these firms. See Andrew Baker, ___. We include this index here to account for such concerns.

where \( u_t \) is a random variable that has mean 0 and variance 1. This variable might or might not be normally distributed, but by assumption it is independently and identically distributed (“iid”); importantly, the independence assumed here is universal, in the sense that we assume \( u_t \) has no association with any other variable in the model—it is pure “white noise”.

Let \( I_t \) denote the set of all information available on date \( t \). The variable \( \sigma_t \) is known as the conditional variance of the daily excess return, because the variance of the daily excess return on date \( t \)—given all information available as of that date—is \( E[\varepsilon_t^2 | I_t] = E[(\sigma_t u_t)^2 | I_t] \),\(^{167}\) which can be shown to equal \( E[\sigma_t^2 | I_t] \), since \( u_t \) is iid and has mean 0 and variance 1.\(^{168}\) The approach taken until Part III.B.4 of this Article can be regarded as being equivalent to assuming that the variance does not actually change across dates, so that \( \sigma_t = \sigma \) for all dates \( t \). In that case we would have \( \varepsilon_t = \sigma u_t \), with \( u_t \) simply being the result of defining the standardized daily excess return, i.e., \( u_t \equiv \sigma^{-1} \varepsilon_t \).

We can allow for the possibility that \( \sigma_t \) varies over time by allowing it to depend on variables that might be expected to drive such variation. First, we allow for the possibility that the date-\( t \) conditional variance of the date-\( t \) excess return, \( \sigma_t^2 \), varies with the value of the squared returns for our three industry index variables.\(^{169}\) Second, we allow for the possibility that \( \sigma_t^2 \) varies with the value of the VIX volatility index.\(^{170}\)

---

\(^{167}\) The variance of \( \varepsilon_t \) equals the expected value of its square since \( \varepsilon_t \) has mean zero.

\(^{168}\) This follows because \( \text{Variance}[\varepsilon_t] = \text{Variance}[\sigma_t u_t] = E[(\sigma_t u_t)^2] - (E[\sigma_t u_t])^2 \). Since \( u_t \) is independent of everything and has mean zero, we have \( E[\sigma_t u_t] = 0 \). Further, \( E[(\sigma_t u_t)^2] = E_{u_t}(E[\sigma_t^2 | u_t] u_t^2) = E[\sigma_t^2] E[u_t^2] = E[\sigma_t^2] \). The first equality follows by the law of iterated expectations; the second follows because \( \sigma_t \) and \( u_t \) are independent by assumption; and the third follows because \( u_t \) has mean zero and variance 1.

\(^{169}\) Contemporaneous values of the squared returns of the indexes—values for date \( t \)—are not yet known during date \( t \). Thus if we were trying to estimate a model of daily excess returns from the perspective of investors, it would be inappropriate to include the contemporaneous squared index returns in our model explaining \( \sigma_t \). However, in conducting an event study used for securities fraud litigation, it is appropriate to include these variables because we are attempting to determine whether event-date excess returns are unusual enough to be considered statistically significant, given not only the information available to investors when they traded on event dates, but also available information about what else happened that date for related securities.

\(^{170}\) See footnote 165, supra concerning the VIX index.
And third, we allow for the possibility that \( \sigma_t^2 \) varies with the GARCH terms described above, namely \( \sigma_{t-1}^2 \) and \( \varepsilon_{t-1}^2 \). Our model for \( \sigma_t^2 \) is:\(^{171}\)

\[
\sigma_t^2 = \sigma_0^2 + V_t^2 \theta_1 + \sigma_{t-1}^2 \theta_2 + \varepsilon_{t-1}^2 \theta_3,
\]

where \( V_t^2 \) is a column vector of the squared returns for the industry peer index, our version of the S&P 500 energy sector index, and our version of the energy and construction index; and the various \( \theta \) parameters connect each explanatory variable on the right hand side to the daily conditional variance. Notice that if all explanatory variables were 0, we would have constant conditional variance with \( \sigma_t^2 = \sigma_0^2 \), which is why we give the intercept in the conditional variance equation the name \( \sigma_0^2 \).

With this model in hand, the next question is how to estimate it. The model is obviously interconnected with our market model, since both involve the daily excess return \( \varepsilon_t \).\(^{172}\) If we were willing to assume not just that the \( u_t \) component of the daily excess return is iid, but also that it follows a normal distribution, then it would be straightforward to write down the model’s log likelihood function and maximize it with respect to the parameters \( \alpha, \beta, \sigma_0^2, \theta_1, \theta_2, \) and \( \theta_3 \); many statistical software packages such as Stata do this using canned routines that are easy to invoke.\(^{173}\)

**B. Estimation Results**

Here we report the results from estimating the model described in section A of this appendix. The first two columns report the coefficient estimates for the parameters in the market model, i.e., the betas that explain Halliburton’s daily return. The coefficient estimates and estimated standard errors are very similar to those in Table 1, as should be expected given the robustness result described in footnote 178, *infra*.

---

171 We also considered the possibility that
172 Recall from Part II.B.1 that the market model is

\[ H_t = \alpha + M_t \beta + \varepsilon_t. \]

173 In Stata, the routine in question is the “arch” command, using the options “arch(1)”, “garch(1)”, and “het(…)”, where the ellipsis in the last option is replaced with a list of names of the variables in \( V_t^2 \) and \( V_{t-1}^2 \) from our equations in the text.
The coefficient estimates for the conditional variance equation—the estimates of the \( \theta \) parameters, using our notation from above—appear in Table 10. The coefficient estimates are statistically insignificant for the squared daily returns of the S&P 500 Energy index and the Engineering & Construction Index (further testing indicated that they are also jointly insignificant). However, both the VIX index and the industry peer index squared daily returns coefficients are highly statistically significant, which tells us that there is variation in the conditional variance of Halliburton’s excess return. The coefficient on the squared returns of the
industry peer index (0.330) is several times the magnitude of the VIX coefficient (0.080), indicating that volatility in the excess returns of Halliburton’s direct competitors plays a comparatively more important role than volatility in the stock market overall (which is what VIX measures). The coefficients on the GARCH parameters show that there is a statistically significant coefficient on the lagged variance term; the coefficient on the square of lagged excess returns is twice the size of the coefficient on the lagged variance, though it is statistically significant only at the 10% level.

Table 10: Estimates from Market Model for Conditional Variance

<table>
<thead>
<tr>
<th>Square of daily return for:</th>
<th>Coefficient</th>
<th>Standard error*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Industry Peers**</td>
<td>0.330</td>
<td>0.083</td>
</tr>
<tr>
<td>S&amp;P 500 Energy**</td>
<td>0.068</td>
<td>0.080</td>
</tr>
<tr>
<td>Engineering &amp; Construction**</td>
<td>0.080</td>
<td>0.061</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Level of:</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>VIX**</td>
<td>0.044</td>
<td>0.017</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>GARCH lag parameters for:</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Variance of excess return</td>
<td>0.128</td>
<td>0.062</td>
</tr>
<tr>
<td>Square of realized excess return</td>
<td>0.244</td>
<td>0.147</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Intercept:</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma_0^2$</td>
<td>-10.024</td>
<td>0.520</td>
</tr>
</tbody>
</table>

* Estimated standard errors computed using robust covariance estimator to account for possibility of non-normality in standardized excess returns, $u_t$.

** Variables expressed in units of standard deviation, so that all these coefficients represent the effect of an increase of one-unit of standard deviation.

174 Because the VIX index and all squared returns variables have been divided by their respective sample standard deviations, the coefficient magnitudes may meaningfully be compared.
We used the coefficient estimates from Table 9 to estimate daily excess returns for both those dates in the estimation period and those dates on which alleged corrective disclosures occurred. That is, we used these coefficients to estimate the set of $\varepsilon_t$ values:

$$\hat{\varepsilon}_t = H_t - \hat{\alpha} - M_t \hat{\beta},$$

where hats denote estimates and we recall that $H_t$ and $M_t$ are, respectively, the observed date-$t$ returns for Halliburton and the observed vector of date-$t$ returns for the market indexes.

To estimate the conditional variance $\sigma_t^2$, we used

$$\hat{\sigma}_t^2 = \hat{\sigma}_0^2 + \hat{V}_{t-1}^{2} \hat{\delta}_1 + \hat{V}_{t-2}^{2} \hat{\delta}_2 + \hat{\delta}_3 \hat{\delta}_3 + \hat{\varepsilon}_{t-1}^2 \hat{\delta}_4,$$

where hats again denote estimates.

With our estimates of the conditional variance and the excess return for each date $t$, we calculated estimates of the standardized excess returns $u_t$ for each date:

$$\hat{u}_t \equiv \frac{\hat{\varepsilon}_t}{\sqrt{\hat{\sigma}_t^2}}$$

C. Statistical Significance Testing Methodology

In this section we discuss how to use the estimated standardized estimated excess returns, computed in section B just above, to test for statistical significance of the event date estimated excess return. It is helpful to first consider what one might do if the standardized excess

\footnote{The only trick here is that we need estimates of $\hat{\sigma}_t^2$ and $\hat{\varepsilon}_t^2$ for the date before the initial date in our estimation sample. We use Stata’s default approach of using the estimated unconditional variance; see Stata Corporation, \textit{Autoregressive conditional heteroskedasticity (ARCH) family of estimators}, \url{http://www.stata.com/manuals13/tsarch.pdf} at 12-13 (“Priming” tab, “arch0” option defaults). This is the method suggested in Timothy Bollerslev, \textit{Generalized Autoregressive Conditional Heteroskedasticity}, 31 J. Econometrics 307 (1986); see Matteo M. Pelagatti and Francesco Lisis, \textit{Variance Initialisation in GARCH Estimation}, paper presented at Complex data modeling and computationally intensive statistical methods for estimation and prediction, Milan September 14-16, 2009, available at \url{http://www2.mate.polimi.it/ocs/viewpaper.php?id=127&cf=7}.}
return $u_t$ were. It is possible that this is true, even though the overall excess return $\varepsilon_t = \sigma_t u_t$ is not normal, because it is possible that all the non-normality in $\varepsilon_t$ is the result of the heterogeneity in the conditional variance $\sigma_t$.\footnote{See, e.g., discussion in Campbell, Lo, and MacKinlay, supra note ___, at 480-481 (explaining that even when the variable we have called $u_t$ is normal, heterogeneity in $\sigma_t$ will necessarily cause excess kurtosis, and thus non-normality, in $\varepsilon_t$).} Perhaps for this reason, Andrew Baker in his student note uses a test that is justified by assuming that $u_t$ is normal. Leaving aside the fact that Baker does not allow for the possibility of autoregressive conditional heteroskedasticity, and some basically unimportant differences in the ways we and Baker estimate the model parameters,\footnote{He uses FGLS, whereas we use QMLE, but both are consistent, which is sufficient for present purposes.} there is one key difference between his approach and our preferred method. Baker uses estimated model parameters to calculate the standardized estimated excess return $\hat{u}_t$ for event dates (essentially as we do below), which he compares to critical values based on the standard normal distribution; we do not use critical values based on the normal distribution.

To be clear, the approach Baker takes is appropriate if $u_t$ really does have a normal distribution. However, the normality assumption concerning $u_t$ could be wrong, and as in Part II.B.1 for the homogeneous variance case, a method of testing for statistical significance that ignores actually existing non-normality is unreliable. One approach to dealing with this issue is to allow alternative forms for the distribution of $u_t$ (perhaps broad enough to include the normal distribution as one possibility). For example, the Stata software package allows users to specify a number of alternative distributions. However, such an approach works only if the chosen distribution is itself correct, and there is no reason to make that assumption, either.

Fortunately, there is a simple and robust alternative, due to a famous result in the econometric theory literature. Even when $u_t$ is not actually distributed normally, under very broad conditions it turns out that the model parameters $\alpha, \beta, \sigma_0^2, \theta_1, \theta_2, \theta_3,$ and $\theta_4$ will be appropriately estimated if we use the maximum likelihood estimator that would follow from assuming normality of $u_t$.\footnote{See Tim Bollerslev and Jeffrey M. Wooldridge, Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances, 11 Econometric Reviews 143 (1992) (proving that the so-called Quasi-Maximum}
For any date $t$, we may think of the estimate $\hat{u}_t$ as the estimated standardized excess return, because it is computed by dividing the estimated excess return by an appropriate estimate of its (date-specific) conditional standard deviation. And since $\hat{\epsilon}_t$ and $\hat{\sigma}_t^2$ are appropriate estimates of $\epsilon_t$ and $\sigma_t^2$, respectively, it follows that $\hat{u}_t$ is an appropriate estimate of the true standardized excess return $u_t$ defined above. By the same arguments used in GHK and referred to above, this means that we may use the sample quantiles of the collection of standardized estimated excess returns $\hat{u}_t$ to serve as critical values for event date standardized estimated excess returns.

In other words, once we calculate the $\hat{u}_t$ estimates, we simply follow the exact SQ approach to testing for statistical significance that we described in Part II.B.1 with respect to non-standardized estimated excess returns. This approach works because, once we standardize estimated excess returns by dividing by the conditional standard deviation, we have eliminated any variance heterogeneity. What remains is a version of the estimated excess return whose distribution can be reliably estimated using the SQ test approach.

---

179 Formally, the Slutsky theorem tells us that the probability limit passes through a continuous function, so that a continuous function of consistent estimators is consistent for the continuous function of the probability limits of those estimators. Here that means that $\hat{u}_t$ is consistent for $u_t$. See ___.