



# CFS Working Paper Series

No. 552

Jill E. Fisch, Jonah B. Gelbach and Jonathan Klick

## After Halliburton: Event Studies and Their Role in Federal Securities Fraud Litigation

## The CFS Working Paper Series

presents ongoing research on selected topics in the fields of money, banking and finance. The papers are circulated to encourage discussion and comment. Any opinions expressed in CFS Working Papers are those of the author(s) and not of the CFS.

The Center for Financial Studies, located in Goethe University Frankfurt's House of Finance, conducts independent and internationally oriented research in important areas of Finance. It serves as a forum for dialogue between academia, policy-making institutions and the financial industry. It offers a platform for top-level fundamental research as well as applied research relevant for the financial sector in Europe. CFS is funded by the non-profit-organization Gesellschaft für Kapitalmarktforschung e.V. (GfK). Established in 1967 and closely affiliated with the University of Frankfurt, it provides a strong link between the financial community and academia. GfK members comprise major players in Germany's financial industry. The funding institutions do not give prior review to CFS publications, nor do they necessarily share the views expressed therein.

# AFTER HALLIBURTON: EVENT STUDIES AND THEIR ROLE IN FEDERAL SECURITIES FRAUD LITIGATION

**August 1, 2016**

*Jill E. Fisch*<sup>\*</sup>  
*Jonah B. Gelbach*<sup>\*\*</sup>  
*Jonathan Klick*<sup>\*\*\*</sup>

## *Abstract*

Event studies have become increasingly important in securities fraud litigation after the Supreme Court's decision in *Halliburton II*. Litigants have used event study methodology, which empirically analyzes the relationship between the disclosure of corporate information and the issuer's stock price, to provide evidence in the evaluation of key elements of federal securities fraud, including materiality, reliance, causation, and damages. As the use of event studies grows and they increasingly serve a gatekeeping function in determining whether litigation will proceed beyond a preliminary stage, it will be critical for courts to use them correctly.

This Article explores an array of considerations related to the use of event studies in securities fraud litigation. It starts by describing the basic function of the event study: to determine whether a highly unusual price movement has occurred and the traditional statistical approach to making that determination. The Article goes on to identify special features of securities fraud litigation that distinguish litigation from the scholarly context in which event studies were developed. The Article highlights the fact that the standard approach can lead to the wrong conclusion and describes the adjustments necessary to address the litigation context.

---

\* Perry Golkin Professor of Law, University of Pennsylvania Law School.

\*\* Professor of Law, University of Pennsylvania Law School.

\*\*\* Professor of Law, University of Pennsylvania Law School.

We thank Bernard Black, Ryan Bubb, Merritt Fox, Jerold Warner and participants at Rutgers-Camden, the February 2016 Penn/NYU Law and Finance Symposium, and USC-Gould for their comments, questions, and suggestions.

We use the example of six dates in the Halliburton litigation to illustrate these points.

Finally, the Article highlights the limitations of event studies – what they can and cannot prove – and explains how those limitations relate to the legal issues for which they are introduced. These limitations bear upon important normative questions about the role event studies should play in securities fraud litigation.

## Table of Contents

Introduction.....	3
I. The Role of Event Studies in Securities Litigation .....	7
II. The Theory of Financial Economics and the Practice of Event Studies: An Overview .....	19
A. Steps (1)-(4): Estimating a Security’s Excess Return .....	20
B. Step (5): Statistical Significance Testing in an Event Study ...	22
III. The Event Study as Applied to the <i>Halliburton</i> Litigation.....	28
A. Dates and Events at Issue in the Halliburton Litigation.....	28
B. An Illustrative Event Study of the Six Dates at Issue in the Halliburton Litigation .....	31
IV. Special Features of Securities Fraud Litigation and their Implications for the Use of Event Studies .....	37
A. The Inappropriateness of Two-Sided Tests .....	38
B. Non-normality in Excess Returns .....	42
C. Multiple event dates of interest.....	47
1. When the question of interest is whether any disclosure had an unusual effect .....	48
2. How should events be grouped together to adjust for multiple testing?.....	51
3. When the question of interest is whether both of two event dates had an effect of known sign.....	53
D. Dynamic Evolution of the Excess Return’s Standard Deviation .....	55
V. Evidentiary Challenges to the Use of Event Studies in Securities Litigation.....	61
A. The Significance of Insignificance .....	62
B. Dealing with Multiple Pieces of News on an Event Date.....	63

C.    Power and Type II Error Rates in Event Studies Used in Securities Fraud Litigation .....	65
Conclusion .....	68
Appendix A:    Implementing the SQ Test .....	70
Appendix B: Correcting for Multiple Testing When There are Multiple Pairs of Alleged Misrepresentation and Alleged Corrective Disclosure Dates .....	72
Appendix C: Allowing for Heterogeneity in the Distribution of Excess Returns .....	75
A.    A Model that Allows for Heterogeneity in the Variance of Excess Returns .....	75
B.    Estimation Results .....	77
C.    Statistical Significance Testing Methodology .....	80

## 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.<sup>1</sup> Although Halliburton failed to persuade the Supreme Court to overrule *Basic, Inc. v. Levinson*,<sup>2</sup> which had approved the fraud-on-the-market (FOTM) presumption of reliance in private securities fraud litigation, it did persuade the Court to allow defendants to introduce evidence of lack of price impact at class certification.<sup>3</sup> 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.”<sup>4</sup>

---

<sup>1</sup> *Halliburton Co. v. Erica P. John Fund, Inc.*, 134 S. Ct. 2398 (2014) (“*Halliburton I*”).

<sup>2</sup> 485 U.S. 224 (1988).

<sup>3</sup> *Halliburton II*, 134 S. Ct. at 2417.

<sup>4</sup> *Halliburton II*, 134 S. Ct. at 2416.

The concept of price impact or price distortion<sup>5</sup> is a critical component of securities fraud litigation. The importance of price impact can be traced back to the Supreme Court's decision in *Basic*, in which the Court adopted a theory of open market securities fraud based on the premise that, in an efficient market, materially fraudulent statements result in plaintiffs purchasing or selling securities at a different price than they would have in a market undistorted by fraud.<sup>6</sup> Although the *Basic* decision considered price impact in the context of determining plaintiffs' reliance on fraudulent statements, price impact is critical to other elements of securities fraud, including loss causation, materiality and damages.

The challenge, post-*Basic*, is to determine whether, in a given case, the allegedly fraudulent statements affected stock price. This task is not trivial -- stock prices fluctuate continuously in response to a variety of issuer and market developments as well as "noise" trading. To address the question, litigators began introducing event studies.<sup>7</sup> Use of the event study methodology has become ubiquitous in securities fraud litigation, and courts have accepted event studies as evidence of price impact at various stages of litigation -- at the pleadings stage, on a motion for class certification, and at the merits.<sup>8</sup> Indeed, many courts have concluded that the use of an event study is preferred or even required to establish one or more of the necessary elements of the plaintiffs' case.<sup>9</sup>

Event studies have their origins in the academic literature.<sup>10</sup> Financial economists use event studies to measure the relationship

---

<sup>5</sup> Fraudulent information has price impact if, in the counterfactual world in which the disclosures were accurate, the price of the security would have been different. One of us has used the term price distortion to encompass both fraudulent information that moves the market price and information that distorts the market by concealing the truth. See Jill E. Fisch, *The Trouble with Basic: Price Distortion After Halliburton*, 90 Wash. U. L. Rev. 895, 897 n. 8 (2013).

<sup>6</sup> *Basic Inc. v. Levinson*, 485 U.S. 224 (1988).

<sup>7</sup> See, e.g., *In re Oracle Sec. Litig.*, 829 F. Supp. 1176, 1181 (N.D. Ca. 1993) ("Use of an event study or similar analysis is necessary to isolate the influences of [the allegedly fraudulent] information").

<sup>8</sup> See, e.g., Alon Brav & J.B. Heaton, *Event Studies in Securities Litigation: Low Power, Confounding Effects, and Bias*, 93 Wash. U. L. Rev. 583, 585 (2016) ("event studies soon became so entrenched in securities litigation that they are viewed as necessary in every case").

<sup>9</sup> See, e.g., *Bricklayers and Trowel Trades Int'l Pension Fund v. Credit Suisse Sec. (USA) LLC*, 752 F.3d 82, 86 (1st Cir. 2014) (The usual—it is fair to say 'preferred'—method of proving loss causation in a securities fraud case is through an event study . . . .").

<sup>10</sup> See, e.g., *United States v. Schiff*, 602 F.3d 152, 173 n.29 (3d Cir. 2010)

between stock prices and various types of events.<sup>11</sup> The core contribution of the event study is its ability to differentiate between price fluctuations that reflect the range of typical variation for a security, and a highly unusual price movement. Although an event study cannot evaluate causation directly, price impact can be inferred from a highly unusual price movement that occurs immediately after an event and has no other potential causes.

The use of event studies in securities fraud litigation entails a number of challenges, however. For one thing, it is unclear that courts fully understand event study methodology and, in particular, its limits. 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?”<sup>12</sup> Counsel responded to Justice Alito’s question by stating that: “Event studies are very effective at making that sort of determination.”<sup>13</sup> In reality, however, 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. Event studies do not speak to the rationality of those price changes.

Second, event studies only measure the movement of a stock price in response to the release of unanticipated material information. In circumstances in which fraudulent statements falsely confirm prior statements, however, the stock price would not be expected to move.<sup>14</sup> Event studies are not capable of measuring the effect of these so-called confirmatory disclosures on stock price.<sup>15</sup> Similarly, in cases involving

---

("An event study . . . is a statistical regression analysis that examines the effect of an event[, such as the release of information,] on a dependent variable, such as a corporation's stock price.").

<sup>11</sup> See generally S. P. Khotari & Jerold B. Warner, *Econometrics of Event Studies*, *Handbook of Corporate Finance: Empirical Corporate Finance*, Volume A (2006) (describing the event study literature and conducting census of event studies published in five journals for the years 1974 through 2000).

<sup>12</sup> Transcript of Oral Argument at 24, *Halliburton II*, 134 S. Ct 2398 (No. 13-317).

<sup>13</sup> *Id.*

<sup>14</sup> See, e.g., *Greenberg v. Crossroads Sys., Inc.*, 364 F.3d 657, 665-66 (5th Cir. 2004) (“[C]onfirmatory information has already been digested by the market and will not cause a change in stock price.”).

<sup>15</sup> As we discuss below, courts have responded to this limitation by allowing plaintiffs to show price impact indirectly through event studies that show a price drop on the date of an alleged corrective disclosure.

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.<sup>16</sup>

Third, there are important differences between the scholarly contexts for which event studies were originally designed and the use of event studies in securities fraud litigation. For example, academics originally designed the event study methodology to measure the significance of a single event across multiple firms, the significance of multiple events at a single firm, or the significance of multiple events at multiple firms.<sup>17</sup> By contrast, the event studies used in securities fraud litigation attempt to measure the impact of a single event at a single firm.<sup>18</sup> This application has important methodological implications for the proper use of event studies. In addition, determining whether to characterize a price movement as highly unusual is the product of methodological choices, including choices about statistical power. In the litigation context, those choices have normative implications that courts have not considered.

In this article, we examine the event study methodology and its use in securities fraud litigation. Part I identifies the variety of legal issues in securities fraud litigation to which event studies are relevant and demonstrates why the concept of a highly unusual price movement is central to all these legal issues. Part II explains how event studies work. Part III conducts a stylized event study using data related to the ongoing *Halliburton* litigation. Part IV identifies the special features of securities fraud litigation that require adjustments to the standard event study approach and demonstrates how a failure to incorporate these features can

---

<sup>16</sup> This sort of problem, which we discuss below, has arisen in cases; *see, e.g.*, *The Archdiocese of Milwaukee Supporting Fund, Inc. v. Halliburton Co.*, Civ. Act. No. 3:02-CV-1152-M (N.D. Tex. Nov. 4, 2008), at 20 (explaining that Halliburton’s Dec. 7, 2001 disclosure contained “two distinct components,” a corrective disclosure of prior misstatements and new negative information and denying class certification because plaintiff was unable to demonstrate that it was more probable than not that the stock price decline was caused by the former); *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.”).

<sup>17</sup> *See, e.g.*, Brav & Heaton, *Event Studies*, note 8, *supra*, at 586 (“almost all academic research event studies are multi-firm event studies (“MFESs”) that examine large samples of securities from multiple firms”).

<sup>18</sup> *See* Jonah B. Gelbach, Eric Helland, & Jonathan Klick, *Valid inference in single-firm, single-event studies*, 10 *Am. L. Econ. Rev.* 1093 (2013).



lead to the wrong conclusion. Part V highlights the limitations of event studies – what they can and cannot prove – and explains how those limitations relate to the legal issues for which they are introduced. We also note the normative/policy implications of these limitations.

A review of judicial use of event studies raises troubling questions about the capacity of the legal system to incorporate social science methodology. Our analysis demonstrates that the proper use of event studies in securities fraud litigation requires care, both in a better understanding of the event study methodology and in an appreciation of its limits.

## I. THE ROLE OF EVENT STUDIES IN SECURITIES LITIGATION

In this Part we take a systematic look at the different questions that event studies might answer in a securities fraud case.<sup>19</sup> As noted above, the use of event studies in securities fraud litigation is widespread. As litigants and courts have become familiar with the methodology, they have expanded the range of purposes for which an event study may be introduced.

The starting point for these developments was the Supreme Court's decision in *Basic, Inc. v. Levinson*.<sup>20</sup> In *Basic*, the Court accepted the FOTM presumption which holds that "the market price of shares traded on well-developed markets reflects all publicly available information, and, hence, any material misrepresentations."<sup>21</sup> The Court observed that the typical investor, in buying or selling stock at the price set by the market, "does so in reliance on the integrity of that price"<sup>22</sup> As a result, the Court concluded that an investor's reliance could be presumed, for purposes of a 10b-5 claim, if the following requirements were met: (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.<sup>23</sup>

---

<sup>19</sup> To succeed on a federal securities fraud claim, the plaintiff must establish the following elements: "(1) a material misrepresentation (or omission); (2) scienter, i.e., a wrongful state of mind; (3) a connection with the purchase or sale of a security; (4) reliance; (5) economic loss; and (6) 'loss causation,' i.e., a causal connection between the material misrepresentation and the loss." *Dura Pharms., Inc. v. Broudo*, 544 U.S. 336, 341-42 (2005) (internal citations omitted).

<sup>20</sup> 485 U.S. 224 (1988).

<sup>21</sup> *Id.*, at 246

<sup>22</sup> *Id.* at 247.

<sup>23</sup> *Id.* at 248 n. 27.

The Court's decision in *Basic* was influenced by a law review article by Professor Daniel Fischel of the University of Chicago Law School.<sup>24</sup> 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.<sup>25</sup> Although the *Basic* decision focused on the reliance requirement, Fischel argued in his article that the only relevant inquiry in a securities fraud case 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.<sup>26</sup> Moreover, Fischel stated that the effect of fraudulent conduct on market price could be determined through a blend of financial economics and applied statistics. Although Fischel did not use the term “event study” in this article, he described the event study methodology.<sup>27</sup>

The lower courts initially responded to the *Basic* decision by focusing extensively on the efficiency of the market in which the securities traded.<sup>28</sup> The leading case on market efficiency, *Cammer v. Bloom* involved a five factor test: “(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.”<sup>29</sup>

Economists serving as expert witnesses generally use event studies to address the fifth *Cammer* factor.<sup>30</sup> 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 generally look at multiple information or news events—some relevant to the litigation in question

---

<sup>24</sup> Daniel R. Fischel, Use of Modern Finance Theory in Securities Fraud Cases Involving Actively Traded Securities, 38 BUS. LAW. 1 (1982).

<sup>25</sup> *Id.* at 9-10.

<sup>26</sup> *Id.* at 13.

<sup>27</sup> *Id.* at 17-18.

<sup>28</sup> See Fisch, *Trouble with Basic*, note 5, *supra*, at 911.

<sup>29</sup> 711 F. Supp. 1264, 1286-87 (D.N.J. 1989). Some courts have added additional factors. See, e.g., *Krogman v. Sterritt*, 202 F.R.D. 467, 477-78 (N.D. Tex. 2001) (identifying three additional factors).

<sup>30</sup> 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).

and some not—for a stock and evaluate the extent to which these events are associated with price changes in the expected directions.<sup>31</sup>

A number of commentators have questioned the centrality of market efficiency to the *Basic* presumption, disputing either the extent to which the market is efficient to the extent presumed by the *Basic* court<sup>32</sup> or the relevance of market efficiency altogether.<sup>33</sup> Financial economists do not consider the *Cammer* factors to be reliable for purposes of establishing market efficiency in academic research.<sup>34</sup> Nonetheless, it has become common practice for both plaintiffs and defendants to submit event studies that address the extent to which the market price of the securities in question respond to publicly reported events for the purpose of addressing *Basic*'s requirement that the securities trade in an efficient market.<sup>35</sup>

The *Basic* decision signaled a broader potential role for event studies, however. By framing its analysis in terms of the harm resulting from a misrepresentation's effect on stock price rather than on the autonomy of investors' trading decisions, *Basic* distanced federal securities litigation from the individualized tort of common law fraud.<sup>36</sup> In this sense, *Basic* was transformative – it introduced a market-based approach to federal securities fraud litigation.<sup>37</sup> Price impact is a critical component of this

---

<sup>31</sup> See, e.g., *Halliburton II* at 2415 (explaining that “EPJ Fund submitted an event study of various episodes that might have been expected to affect the price of Halliburton’s stock, in order to demonstrate that the market for that stock takes account of material, public information about the company.”).

<sup>32</sup> See, e.g., Jonathan R. Macey, et al., *Lessons from Financial Economics: Materiality, Reliance, and Extending the Reach of Basic v. Levinson*, 77 VA. L. REV. 1017, 1018 (1991) (citing “substantial disagreement . . . about to what degree markets are efficient, how to test for efficiency, and even the definition of efficiency.”); Baruch Lev & Meiring de Villiers, *Stock Price Crashes and 10b-5 Damages: A Legal, Economic, and Policy Analysis*, 47 Stan. L. Rev. 7, 20 (1994) (stating that “overwhelming empirical evidence suggests that capital markets are not fundamentally efficient”). Notably, Lev & de Villiers concede that markets are likely information-efficient, which is the predicate requirement for FOTM. See *Id.* at 21 (“While capital markets are in all likelihood not fundamentally efficient, widely held and heavily traded securities are probably ‘informationally efficient.’”).

<sup>33</sup> See, e.g., Brief of Law Professors as Amici Curiae at 4, *Halliburton II*, 134 S. Ct. 2398 (2014) (No. 13-317) (arguing that inquiry into market efficiency to show reliance was a “mistake and counterproductive”); Fisch, *The Trouble with Basic*, note 5, *supra*, at 898 (stating that “market efficiency is neither a necessary nor a sufficient condition to establish that misinformation has distorted prices”).

<sup>34</sup> See Brav & Heaton, *Event Studies*, note 8, *supra*, at 601.

<sup>35</sup> See, e.g., *Halliburton II*, slip op. at 19 (describing event study submitted by the plaintiff to establish market efficiency).

<sup>36</sup> See Fisch, *Trouble with Basic*, note 5, *supra*, at 914.

<sup>37</sup> Fisch, *Trouble with Basic*, note 5, *supra*, at 916.

approach, because, in the absence of an impact on stock price, plaintiffs who trade in reliance on the market price are not defrauded. As the Supreme Court subsequently noted in *Halliburton II*, “[i]n the absence of price impact, Basic’s fraud-on-the-market theory and presumption of reliance collapse.”<sup>38</sup>

The importance of price impact extends beyond the reliance requirement. In *Dura Pharmaceuticals*, the plaintiffs, relying on Basic, filed a complaint in which they alleged that, at the time they purchased Dura stock, its price had been artificially inflated due to Dura’s alleged misstatements.<sup>39</sup> The Supreme Court held that, while artificial price inflation at the time of the plaintiffs’ purchase might address the reliance requirement, plaintiffs were also required to plead and prove the separate element of loss causation.<sup>40</sup> Key to the Court’s reasoning was that purchasing at an artificially inflated price did not automatically cause economic harm in that an investor might purchase at an artificially inflated price and subsequently sell while the price was still inflated.<sup>41</sup>

Following *Dura*, courts have allowed plaintiffs to establish loss causation in various ways, but the standard approach involves the use of an event study “to demonstrate both that the economic loss occurred and that this loss was proximately caused by the defendant’s misrepresentation.”<sup>42</sup> Practically speaking, plaintiffs in the post-*Dura* era need to plead price impact both at the time of the misrepresentation<sup>43</sup> and price impact on the alleged corrective disclosure date. However, in *Halliburton I*, the Supreme Court explained that plaintiffs do not need to prove loss causation in order to avail themselves of the *Basic* presumption, since this presumption has to do with “transaction

---

<sup>38</sup> *Halliburton II* at 2414.

<sup>39</sup> *Dura Pharmaceuticals, Inc. v. Broudo*, 544 U.S. 336 (2005).

<sup>40</sup> *Id.* at 346. The PSLRA had codified the loss causation requirement, which had previously been developed by the lower courts. Pub. L. No. 104-67, § 101(b), § 21D(b)(4), 109 Stat. 737, 747 (codified at 15 U.S.C. § 78u-4(b)(4) (2006)). See Jill E. Fisch, *Cause for Concern: Causation and Federal Securities Fraud*, 94 IOWA L. REV. 811, 820–22 (2009) (describing judicial development of the loss causation requirement).

<sup>41</sup> *Id.* at 343.

<sup>42</sup> Michael J. Kaufman & John M. Wunderlich, *Regressing: The Troubling Dispositive Role of Event Studies in Securities Fraud Litigation*, 15 Stan. J.L. Bus. & Fin. 183, 198 (2009)

<sup>43</sup> The former requirement is not necessary in cases involving confirmatory disclosures. See *infra* at notes 62-65 and accompanying text (discussing confirmatory disclosures)..

causation”—the decision to buy the stock in the first place, which occurs before any evidence of loss causation could exist.<sup>44</sup>

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....”<sup>45</sup> Indeed, many courts have refused to admit expert testimony with respect to loss causation that did not include an event study.<sup>46</sup>

Proof of price impact for purposes of analyzing reliance and causation also overlaps with the materiality requirement.<sup>47</sup> The Court has defined material information as information that has a substantial likelihood to be “viewed by the reasonable investor as having significantly altered the 'total mix' of information made available.”<sup>48</sup> Because market prices are a reflection of investors' trading decisions, information that is relevant to those trading decisions has the capacity to impact stock prices and, similarly, information that does not affect stock prices is arguably immaterial.<sup>49</sup> As the Third Circuit explained in

---

<sup>44</sup> *Erica P. John Fund, Inc. v. Halliburton Co.*, 563 U.S. 804, 812 (2011) (“*Halliburton I*”). As to the merits, though, plaintiffs must also demonstrate a causal link between the two events – the initial misstatement and the corrective disclosure. See, e.g., *Aranaz v. Catalyst Pharmaceutical Partners Inc.*, 302 F.R.D. 657, 671-72 (S.D.Fla.2014) (describing and rejecting defendants' argument that other information on the date of the alleged corrective disclosure was responsible for the fall in stock price). *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 a price impact from the misstatements).

<sup>45</sup> *Bricklayers & Trowel Trades Int'l Pension Fund v. Credit Suisse Secs.*, 752 F.3d 82, (1st Cir. 2014).

<sup>46</sup> 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.”)

<sup>47</sup> See, e.g., *Erica P. John Fund, Inc. v. Halliburton Co.*, 718 F.3d 423, 434 n. 10 (5th Cir. Tex. 2013) (observing that “there is a fuzzy line between price impact evidence directed at materiality and price impact evidence broadly directed at reliance”).

<sup>48</sup> *Basic*, 485 U. S. at 232.

<sup>49</sup> See Frederick C. Dunbar & Dana Heller, *Fraud on the Market Meets Behavioral Finance*, 31 DEL. J. CORP. L. 455, 509 (2006) (“The definition of immaterial information . . . is that it is already known or . . . does not have a statistically significant effect on stock price in an efficient market.”). Cf. Donald C. Langevoort, *Basic at Twenty: Rethinking Fraud on the Market*, 2009 WIS. L. REV. 151, 173-77 (arguing that, in some cases, material information may not affect stock prices).

*Burlington Coat Factory*: “In the context of an “efficient” market, the concept of materiality translates into information that alters the price of the firm’s stock.”<sup>50</sup> Event studies can be used to demonstrate the impact of fraudulent statements on stock price, providing evidence that the statements are material.<sup>51</sup> The lower courts have, on occasion, accepted the argument that the absence of price impact demonstrates the immateriality of alleged misrepresentations.<sup>52</sup>

A statement can be immaterial because it is unimportant or because it conveys information that is already known to the market.<sup>53</sup> The latter argument is known as the “truth on the market” defense, since the argument is that the market already knew the truth. According to the truth on the market defense, an alleged misrepresentation that occurs after the market already knows the truth cannot change market perceptions of firm value because any effect of the truth will already have been incorporated into the market price.<sup>54</sup>

In *Amgen*,<sup>55</sup> the parties agreed that the market for Amgen’s stock was efficient and that the statements in question were public, but disputed the reasons why *Amgen*’s stock price had dropped on alleged corrective disclosure dates.<sup>56</sup> Specifically, the defendants argued that, because the truth regarding the alleged misrepresentations was publicly known before plaintiffs purchased their shares, plaintiffs did not trade at a price

---

<sup>50</sup> *In re Burlington Coat Factory Sec. Litig.*, 114 F.3d 1410, 1425 (3d Cir. N.J. 1997).

<sup>51</sup> See, e.g., *In re Sadia*, 269 F.R.D. 298, 302 (S.D.N.Y. 2010); *In re Gaming Lottery Sec. Litig.*, 2000 U.S. Dist. LEXIS 1558, 2000 WL 193125, \*1 (S.D.N.Y. 2000) (describing the event study as “an accepted method for the evaluation of materiality damages to a class of stockholders in a defendant corporation.”).

<sup>52</sup> See *In re Merck & Co. Sec. Litig.*, 432 F.3d 261 (3d Cir. 2005); *Oran v. Stafford*, 226 F.3d 275, 282 (3d Cir. 2000) (Alito, J.) (“[I]n an efficient market the concept of materiality translates into information that alters the price of the firm’s stock . . . .” (internal quotation marks omitted)); *In re Burlington Coat Factory Sec. Litig.*, 114 F.3d 1410, 1425 (3d Cir. N.J. 1997).

<sup>53</sup> See *Conn. Ret. Plans & Trust Funds v. Amgen Inc.*, 660 F.3d 1170, 1177 (9th Cir. Cal., 2011) 660 F.3d, at 1177 (A “truth-on-the-market defense . . . “is a method of refuting an alleged misrepresentation’s materiality”).

<sup>54</sup> See, e.g., *Aranaz v. Catalyst Pharm. Partners*, 302 F.R.D. 657, 670–71 (S.D. Fla. 2014) (seeking to introduce evidence that price was not distorted by alleged misrepresentations because the market knew the truth).

<sup>55</sup> *Amgen Inc. v. Conn. Ret. Plans & Trust Funds*, 133 S. Ct. 1184 (2013).

<sup>56</sup> See, e.g., Defendants’ Brief in Opposition to Class Certification, *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).

that was impacted by the fraud.<sup>57</sup> Although the majority in Amgen concluded that proof of materiality was not required at the class certification stage, it acknowledged that the defendant's proffered "truth-on-the-market" evidence could potentially refute materiality."<sup>58</sup>

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.<sup>59</sup> The 1934 Act provides that plaintiffs may recover actual damages, which must be proved.<sup>60</sup> A plaintiff who can prove damages has obviously proved she sustained an economic loss. At the same time, a plaintiff who cannot prove damages cannot prove she suffered an economic loss. Thus the economic loss and damages determination elements merge into one. A number of courts have rejected testimony or reports by damages experts that failed to include an event study.<sup>61</sup>

Notably, while the price impact at the time of the fraud (required in order to obtain the Basic presumption of reliance) is not the same as price impact at the time of the corrective disclosures (loss causation under *Dura*),<sup>62</sup> in many cases, the parties may seek to address both elements with a single event study. This is most common in cases that involve alleged fraudulent confirmatory statements. Misrepresentations that falsely confirm market expectations will not lead to an *observable change in price*.<sup>63</sup> But this does not mean they have no *price impact*.

---

<sup>57</sup> *Amgen Inc. v. Conn. Ret. Plans & Trust Funds*, 133 S. Ct. 1184, 1193-94 (2013). As a lower court had put it, "FDA announcements and analyst reports about Amgen's business [had previously] publicized the truth about the safety issues looming over Amgen's drugs.", citing *Connecticut Retirement Plans and Trust Funds v. Amgen Inc.*, 660 F.3d 1170, 1177 (9<sup>th</sup> Cir. 2011).

<sup>58</sup> *Amgen*, 133 S.Ct. at 1201.

<sup>59</sup> 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*.

<sup>60</sup> 15 U.S.C. § 78bb(a)(1).

<sup>61</sup> See, e.g., *Fecht v. Northern Telecom Ltd. (In re Northern Telecom Ltd. Sec. Litig.)*, 116 F. Supp. 2d 446, 460 (S.D.N.Y. 2000) (terming expert's testimony "fatally deficient in that he did not perform an event study or similar analysis"); *In re Executive Telecard Sec. Litig.*, 979 F. Supp. 1021, 1025 (S.D.N.Y. 1997) (concluding that the "reliability of the Expert Witness' proposed testimony is called into question by his failure to indicate. . . whether he conducted an 'event study'").

<sup>62</sup> See *Halliburton I* (distinguishing between reliance and loss causation); see also Fisch, *Trouble with Basic*, note 5, *supra* (highlighting the distinction and terming the former ex ante price distortion and the latter ex post price distortion).

<sup>63</sup> See, e.g., *FindWhat Investor Group v. FindWhat.com*, 658 F.3d 1282, 1310 (11<sup>th</sup> Cir. 2011) ("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

The relevant price impact is simply counterfactual: the price would have fallen had there not been fraud.

In cases where plaintiffs allege confirmatory misrepresentations, event study evidence has no probative value related to alleged misrepresentation dates, since the plaintiffs' own allegations predict no change in price. Thus there will be no *observed* price impact on alleged misrepresentation dates. However, a change in observed price will ultimately 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. A showing that the stock price responded to a subsequent corrective disclosure can provide indirect evidence of counterfactual price impact of the alleged misrepresentation.<sup>64</sup> 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.

*Halliburton II* presented this scenario.<sup>65</sup> Plaintiffs alleged that Halliburton made a variety of fraudulent confirmatory disclosures that artificially maintained the company's stock price. Initially, defendants had argued that the plaintiff could not establish loss causation because Halliburton's subsequent corrective disclosures did not impact the stock price.<sup>66</sup> When the Supreme Court held, in *Halliburton I*, that the plaintiffs were not required to prove loss causation on a motion for class certification,<sup>67</sup> Halliburton argued, on remand, "that the evidence it had presented to disprove loss causation also demonstrated that none of the alleged misrepresentations actually impacted Halliburton's stock price, i.e., there was a lack of 'price impact,' and, therefore, Halliburton had rebutted the Basic presumption."<sup>68</sup> Halliburton attempted to present

---

in the stock price. This is so because the market has already digested that information and incorporated it into the price.").

<sup>64</sup> 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. Secs.*, 281 F.R.D. 134, 143 (S.D.N.Y. 2012) (finding that stock price's negative reaction to corrective disclosure served to defeat defendant's argument of lack of price impact).

<sup>65</sup> *Erica P. John Fund, Inc. v. Halliburton Co.*, 134 S. Ct. 2398 (2014).

<sup>66</sup> *Archdiocese of Milwaukee Supporting Fund, Inc. v. Halliburton Co.*, 2008 WL 4791492 (N.D. Tex. 2008).

<sup>67</sup> *Erica P. John Fund, Inc. v. Halliburton Co.*, 131 S.Ct. 2179, 2187 (2011).

<sup>68</sup> See *Erica P. John Fund, Inc. v. Halliburton Co.*, 309 F.R.D. 251, 255-56 (N.D. Tex. 2015)



“extensive evidence of no price impact,” evidence that the lower courts ruled was “not appropriately considered at class certification.”<sup>69</sup>

The Supreme Court disagreed. In *Halliburton II*, Chief Justice Roberts explained 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; 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.<sup>70</sup> Under these circumstances, the Chief Justice concluded that prohibiting a court from relying on this same evidence to evaluate whether the fraud affected stock price “makes no sense.”<sup>71</sup>

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 general—“or allowing consideration of direct evidence as well.”<sup>72</sup> 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.

On remand, the trial court considered the event study submitted by Halliburton’s expert, which purported to find that neither the alleged misrepresentations nor the corrective disclosures<sup>73</sup> identified by the

---

<sup>69</sup> *Erica P. John Fund, Inc. v. Halliburton Co.*, 718 F.3d 423, 435 n.11 (5th Cir. 2013) vacated and remanded, 134 S. Ct. 2398 (2014).

<sup>70</sup> *Halliburton Co. v. Erica P. John Fund, Inc.*, 134 S. Ct. 2398, 2417 (2014).

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).

<sup>71</sup> *Halliburton Co. v. Erica P. John Fund, Inc.*, 134 S. Ct. 2398, 2415 (2014).

<sup>72</sup> *Id.* at 2417.

<sup>73</sup> As the court noted, “Measuring price change at the time of the corrective disclosure, rather than at the time of the corresponding misrepresentation, allows for the fact that many alleged misrepresentations conceal a truth.” *Erica P. John Fund, Inc. v. Halliburton Co.*, 309 F.R.D. 251, 262 (N.D. Tex. 2015)

plaintiff impacted Halliburton's stock price.<sup>74</sup> After carefully considering the event studies submitted by both parties, which addressed six corrective disclosures, the court found that Halliburton had successfully demonstrated a lack of price impact as to five of the dates, granting class certification with respect to the Dec. 7<sup>th</sup> alleged corrective disclosure.

Following *Halliburton II*, several other lower courts have considered defendants' use of event studies to demonstrate the absence of price impact. In *Local 703 v. Regions Fin. Corp.*,<sup>75</sup> 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 *Halliburton II* "did not fully consider this evidence."<sup>76</sup> Accordingly, the court vacated and remanded "for fuller consideration... of all the price-impact evidence submitted below."<sup>77</sup> 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.<sup>78</sup> 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 absence of price impact.<sup>79</sup> This decision places the burden of persuasion concerning price impact squarely on defendants.<sup>80</sup>

In *Aranaz v. Catalyst Pharmaceutical Partners Inc.*,<sup>81</sup> 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

---

<sup>74</sup> *Id.* at 262. The court noted that the expert attributed the one date on which the stock experienced a highly unusual price movement as a reaction to factors other than Halliburton's disclosure. *Id.*

<sup>75</sup> *Local 703, I.B. of T. Grocery & Food Employees Welfare Fund v. Regions Fin. Corp.*, 762 F.3d 1248 (11th Cir. Ala. 2014)

<sup>76</sup> *Id.* at 1258.

<sup>77</sup> *Id.*

<sup>78</sup> *Local 703, I.B. of T. Grocery & Food Emples. Welfare Fund v. Regions Fin. Corp.*, 2014 U.S. Dist. LEXIS 162403 (S.D. Ala. 2014).

<sup>79</sup> Defendants argued that their expert's event study "conclusively finds no price impact on January 20, 2009 [the date of the corrective disclosure]." *Id.* at \*25.

<sup>80</sup> See Merritt B. Fox, *Halliburton II: It All Depends on What Defendants Need to Show to Establish No Impact on Price*, 70 *Bus. Law.* 437 (2015) (draft at 20-23) (describing the resulting statistical burden this approach would impose on defendants to rebut the presumption).

<sup>81</sup> *Aranaz v. Catalyst Pharmaceutical Partners Inc.*, 302 F.R.D. 657 (S.D.Fla.2014).

impact on stock price.<sup>82</sup> 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 disclosure<sup>83</sup> but argued that other statements in the two publications caused the “drastic changes in stock price.”<sup>84</sup> The court concluded that, because the defendant had the burden of proving that “price impact is inconsistent with the results of their analysis,”<sup>85</sup> 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.<sup>86</sup>

The lower court decisions following *Halliburton II* demonstrate the growing importance of event studies. The most recent trial court decision as to class certification in the *Halliburton* litigation itself<sup>87</sup> demonstrates as well the challenges for the court in evaluating the event study methodology, an issue we will consider in more detail in Part III below.

Significantly, as reflected in the preceding discussion, proof of price impact is relevant to multiple elements of securities fraud. A single event study may provide evidence relating to materiality, reliance, loss causation, economic loss, and damages. Although 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 implies that plaintiff will be unable to establish entitlement to damages. These

---

<sup>82</sup> Under *Halliburton I* and *Amgen*, this limit is appropriate. The district court in *Halliburton* took the same approach on remand following *Halliburton II*; see *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 [actually] 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.”).

<sup>83</sup> *Aranaz* at \*22.

<sup>84</sup> *Id.* at \*29.

<sup>85</sup> *Id.* at 32.

<sup>86</sup> *Id.* at \*27.

<sup>87</sup> *Erica P. John Fund v. Halliburton Co.*, 309 F.R.D. 251 (N.D. Tex. 2015).

observations explain why event studies play such a central role in securities fraud litigation.

Loss causation and price impact have taken center stage at the pleading and class certification stages. If the failure to establish price impact is fatal to plaintiffs' case, defendants benefit by making that challenge at the pleading stage, before plaintiffs can obtain discovery<sup>88</sup> or by preventing plaintiffs from obtaining the leverage of class certification.<sup>89</sup> Accordingly much of the Supreme Court's jurisprudence on loss causation and price impact has been decided in the context of pre-trial motions.

*Basic* was decided on a motion for class certification. A key factor in the Court's analysis was the critical role that a presumption of reliance would play in enabling plaintiff to address Rule 23's commonality requirement.<sup>90</sup> As the Court explained "Requiring proof of individualized reliance from each member of the proposed plaintiff class effectively would have prevented respondents from proceeding with a class action, since individual issues then would have overwhelmed the common ones."<sup>91</sup> By facilitating class certification, *Basic* has been described as transforming private securities fraud litigation.<sup>92</sup>

Defendants have responded by attempting to increase the burden imposed on the plaintiff to obtain class certification. In *Halliburton I*, the lower courts accepted defendant's argument that plaintiffs should be required to establish loss causation at class certification.<sup>93</sup> In *Amgen*, the defendants argued that the plaintiff should be required to establish materiality in order to obtain class certification.<sup>94</sup> Notably, in both cases, defendants' objective was to require the plaintiff to prove price impact, through an event study, at a preliminary stage in the litigation rather than at the merits stage.

---

<sup>88</sup> 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).

<sup>89</sup> See, e.g., *Halliburton I* Transcript, note 12, *supra*, at 23 (Justice Scalia: "Once you get the class certified, the case is over, right?").

<sup>90</sup> *Basic Inc. v. Levinson*, 485 U.S. 224, 242 (U.S. 1988).

<sup>91</sup> *Basic Inc. v. Levinson*, 485 U.S. 224, 242 (U.S. 1988).

<sup>92</sup> See, e.g., Langevoort, note 49, *supra*, at 152 (stating that "[t]ens of billions of dollars have changed hands in settlements of 10b-5 lawsuits in the last twenty years as a result of *Basic*").

<sup>93</sup> *Archdiocese of Milwaukee Supporting Fund, Inc. v. Halliburton Co.*, 597 F.3d 330, 2010 U.S. App. LEXIS 3226 (5th Cir. Tex., 2010); *Archdiocese of Milwaukee Supporting Fund, Inc. v. Halliburton Co.*, 2008 U.S. Dist. LEXIS 89598 (N.D. Tex., Nov. 4, 2008).

<sup>94</sup> *Amgen Inc. v. Connecticut Retirement Plans and Trust Funds*, 133 S.Ct. 1184 (2013).

Similarly, the Court's decision in *Dura Pharmaceuticals* was issued in the context of a motion to dismiss for failure to state a claim.<sup>95</sup> The complaint ran afoul of even the pre-*Twombly*<sup>96</sup> pleading standard by failing to allege any corrective disclosure associated with a loss.<sup>97</sup> The *Dura* court held that the plaintiffs' failure to plead loss causation meant that the complaint failed to show entitlement to relief as required under Rule 8(a)(2),<sup>98</sup> 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. Given the importance of the event study in securities litigation, it important to understand both the methodology involved and the limitations of its results.

## II. THE THEORY OF FINANCIAL ECONOMICS AND THE PRACTICE OF EVENT STUDIES: AN OVERVIEW

The theory of financial economics adopted by courts for purposes of securities litigation is based on the premise that publicly released information concerning a security's price will be incorporated into market price quickly. This premise is known in financial economics as the semi-strong form of the efficient market hypothesis,<sup>99</sup> 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 quickly inflate the firm's stock price over the level that true conditions warrant. Conversely, statements that correct such inflationary misrepresentations will quickly lead the stock price to fall.

---

<sup>95</sup> *Dura Pharmaceuticals, Inc. v. Broudo*, 544 U.S. 336 (U.S., 2005).

<sup>96</sup> *Bell Atlantic Corp. v. Twombly*, 550 U.S. 544 (2007).

<sup>97</sup> *Dura Pharmaceuticals, Inc. v. Broudo*, 544 U.S. 336, 347 (2005) ("the complaint nowhere ... provides the defendants with notice of what the relevant economic loss might be or of what the causal connection might be between that loss and the misrepresentation concerning Dura's" product).

<sup>98</sup> *Id.* at 346. Fed. R. of Civ. P. 8.

<sup>99</sup> 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).

Financial economists began using event studies to measure how much stock prices respond to various types of news.<sup>100</sup> Typically event studies focus not on the level of a stock's price, but on the percentage change in stock price, which is known as the stock's observed "return." In its simplest form, an event study compares a stock's return on a day when news of interest hits the market to the range of returns typically observed for that stock, taking account of what would have been expected given general changes in the overall market on that day. For example, if a stock typically moves up or down by no more than one percent in either direction, but the stock rises by 2% on a date of interest, after controlling for relevant market conditions, then the stock return moved an unusual amount on that date. What range is "typical," and thus how large a return must be to be considered unusual, is a question that event study authors answer using statistical significance testing.

A typical event study has five basic steps: (1) identify one or more appropriate event dates, (2) calculate the security's return on each event date, (3) determine the security's expected return for each event date, (4) subtract the actual return from the expected return to compute the excess return for each event date and (5) evaluate whether the resulting excess return is statistically significant.<sup>101</sup> We treat these five steps in two sections.

#### A. Steps (1)-(4): Estimating a Security's Excess Return

Experts typically address the first step, selecting the event date, by using the date on which the representation or disclosure is publicly made.<sup>102</sup> For purposes of public market securities fraud, the information must be communicated widely enough that the market price can be

---

<sup>100</sup> For a history of the use of event studies in academic scholarship see A. Craig MacKinlay, *Event Studies in Economics and Finance*, 35 J. Econ. Lit. 13 (1997).

<sup>101</sup> See Jonathan Klick & Robert H. Sitkoff, *Agency Costs, Charitable Trusts, and Corporate Control: Evidence from Hershey's Kiss-Off*, 108 Colum. L. Rev. 749, 798 (2008).

<sup>102</sup> The event study literature contains an extensive treatment of the appropriate choice of event window, a topic that we do not consider in detail here. See Alex Rinaudo & Atanu Saha, *An Intraday Event Study Methodology for Determining Loss Causation*, 2 J. Fin. Persp. 161, 163 (2014) (observing that the typical event window for an event study is a single day but advocating instead "intraday event study methodology relying on minute-by-minute stock price data."); Allen Ferrell & Atanu Saha, *The Loss Causation Requirement for Rule 10b-5 Causes of Action: The Implications of Dura Pharmaceuticals, Inc. v. Broudo*, 63 BUS. LAW. 163, 167-68 (2007) (discussing factors affecting choice of event window).

expected to react to the information.<sup>103</sup> The second step, calculating a security's actual return, requires only public information about daily security prices.<sup>104</sup>

The third step is to determine the security's expected return on the event date, given market conditions that might be expected to affect the firm's price even in the absence of the news at issue. Event study authors do this by using statistical methods to separate out components of a security's return that are based on overall market conditions from the component due to firm-specific information. Market conditions typically are measured using a broad index of other stocks' returns on each date considered in the event study, and/or an index of returns of others firms engaged in similar business (since firms engaged in common business activities are likely to be affected by similar types of information). To determine the expected return for the security in question, an expert will estimate a regression model that controls for the returns to market and/or industry stock indexes.<sup>105</sup> The estimated coefficients from this model can then be used to measure the expected return for the firm in question, given the performance of the index variables included in the model.

The fourth step is to calculate the "excess return"<sup>106</sup>, which one does by subtracting the expected return from the actual return on the date in question. Thus the excess return is the component of the actual return that cannot be explained by market movements on the event date, given the regression estimates described above. So, the excess return measures the stock's reaction to whatever news occurred on the event date.

A positive excess return indicates that the firm's stock increased more than would be expected based on the statistical model; a negative excess return indicates that the stock fell more than the model predicts it should have. Figure 1 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. The upwardly sloped straight line represents the

---

<sup>103</sup> In some cases, litigants may dispute whether information is sufficiently public to generate a market reaction; in other situations, leakage of information before public announcement may generate an earlier market reaction. These specialized situations can be addressed by tailoring the choice of event date.

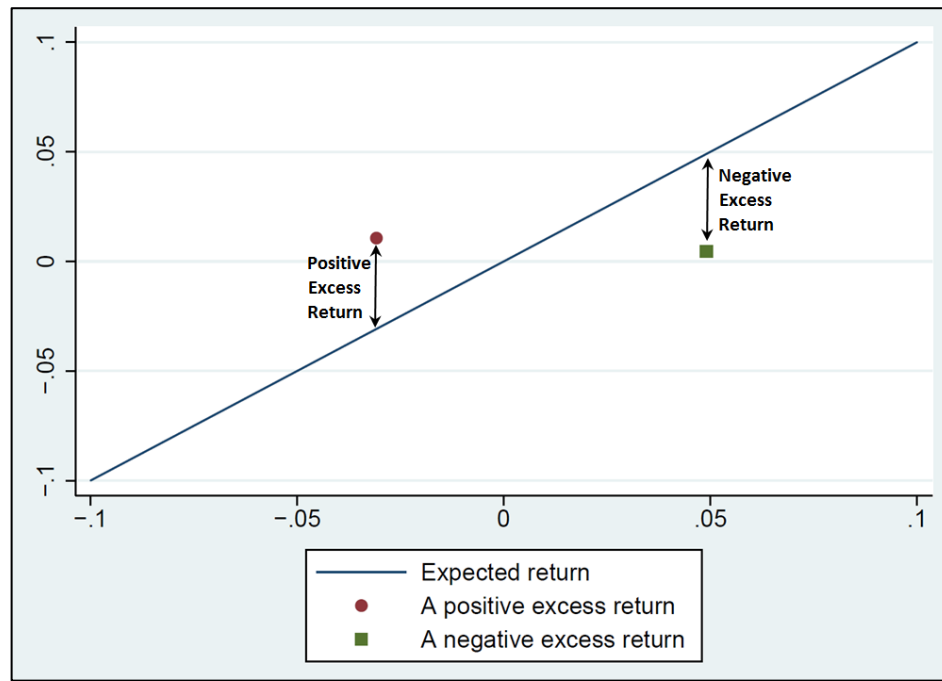
<sup>104</sup> Recall that a security's daily return on a particular date is the percentage change in the security over the preceding date.

<sup>105</sup> As one pair of recent commentators have noted, the "failure to make adjustments for the effect of market and industry moves nearly always dooms an analysis of securities prices in litigation." Brav & Heaton, *Event Studies*, note 8, *supra*, at 590.

<sup>106</sup> The term "abnormal return" is interchangeable with excess return. We use only "excess return" in this Article in order to avoid confusing "abnormal returns" with non-normality in the distribution of these returns.

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 lies above the line where actual and expected returns are equal, so this point indicates a positive excess return. By contrast, at the point plotted with a square, the actual return is below the line where the actual and expected returns are equal, so the excess return is negative.

**Figure 1: Illustrating the Calculation of Excess Returns from Actual and Expected Returns**



*B. Step (5): Statistical Significance Testing in an Event Study*

Our fifth and final step is to determine whether the estimated excess return is statistically significant. The use of statistical significance testing is designed to distinguish stock price changes that are just the result of typical volatility from those that are sufficiently unusual that they are likely a response to the alleged corrective disclosure.

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%. The hypothesis that the coin is fair, so that the chance of a heads is 50%, is an example of what statisticians call a *null hypothesis*: a maintained assumption about the object of statistical study, which will be dropped only if the statistical evidence is sufficiently inconsistent with the assumption.

Since one can expect random variation to affect the share of heads in 100 coin flips, 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%. Even though these results do not equal exactly the baseline level, they are close enough that the evidence is too weak to reject the null hypothesis that the coin is fair. On the other hand, common sense and statistical methodology suggest that if 89 of 100 tosses yielded heads, it would be strong evidence that the coin were biased toward heads. A finding of 89 heads would then allow us to reject the null hypothesis that the coin is fair.

Event study tests of whether a stock price moved in response to information are similar to the coin-toss example. They seek to determine whether the stock's excess return was highly unusual on the event date. The null hypothesis in an event study is that the news at issue did not have any price impact. Under this null hypothesis, the stock's return should reflect only the usual relationship between the stock and market conditions on the event date; in other words, the stock's return should be the expected return, together with normal variation. Our baseline expectation for the stock's excess return is that it should be zero. Normal variation, however, will cause the stock's actual return to differ somewhat from the expected return. Statistical significance testing focuses on whether this deviation—the actual excess return on the event date—is far enough from zero so as to be highly unusual.

What counts as highly unusual? Typically courts and experts have assumed that an event-date effect is statistically significant 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.<sup>107</sup> In this

---

<sup>107</sup> See, e.g., *Erica P. John Fund, Inc. v. Halliburton Co.*, 309 F.R.D. 251, 262 (N.D. Tex. 2015) (“To show that a corrective disclosure had a negative impact on a company's share price, courts generally require a party's expert to testify based on an event study that meets the 95% confidence standard, which means ‘one can reject with 95% confidence the null hypothesis that the corrective disclosure had no impact on price.’”) (citing Fox, *It All Depends*, note 80, *supra*, n. 17 at 442); Cf. Brav & Heaton, *Event Studies*, note 8, *supra*, 596-99 (questioning whether requiring statistical significance at the 95% confidence level for securities fraud event studies is appropriate). The genesis of the 5% significance level is most probably its use by R. A.

situation, experts might say that there is statistically significant evidence at the 5% level or, equivalently, “at level 0.05”, or “with 95% confidence.”<sup>108</sup>

Experts typically assume that, in the absence of any fraud-related event, a stock’s excess returns—that is, the typical variability not driven by the news at issue in litigation—will follow a normal distribution,<sup>109</sup> an issue we discuss in more detail in Part IV. 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.<sup>110</sup> Experts assuming normality of excess returns and using the 95% confidence level often determine that the excess return is highly unusual if it is greater than 1.96 standard deviations. For example, if the standard deviation of a stock’s excess returns is 1.5%, an expert might declare an event date’s excess return statistically significant if and only if it is more than 2.94 percentage points from zero.<sup>111</sup> In this example, the expert has determined that the “critical value” is 2.94: any value of the event date excess return greater in magnitude than this value will lead the expert to determine that the excess return is statistically significant. A lower value for the excess return would lead to a finding of statistical insignificance.

When an event date excess return is statistically significant, courts will treat the size of the excess return as a measure of the price effect associated with the news at issue. One consequence is that the excess return may then be used as a basis for damages determination. On the other hand, if the excess return is statistically insignificant, then the statistical evidence is too weak to meet the plaintiff’s burden of persuasion that the information affected the stock price.

Note that a statistically insignificant finding may occur even when the excess return is directionally consistent with the plaintiff’s

---

Fisher in his influential textbook; see R. A. Fisher, *STATISTICAL METHODS FOR RESEARCH WORKERS*, 47 (1925) (cited in Wikipedia, *1.96*, at [https://en.wikipedia.org/wiki/1.96#cite\\_note-11](https://en.wikipedia.org/wiki/1.96#cite_note-11), last accessed on July 19, 2016).

<sup>108</sup> That is not to say that the event study can determine whether this price effect is “rational” in the substantive sense that Justice Alito seems to have had in mind; see TAN 12, *supra*. The measured price impact represented by the excess return is simply the effect that is empirically evident from investor behavior in the relevant financial market.

<sup>109</sup> See, e.g., Brav & Heaton, *Event Studies*, note 8, *supra*, 591, n. 17 (“standard practice still rests heavily on the normality assumption”).

<sup>110</sup> The standard deviation is a measure of how spread out a large random sample of the variable is likely to be. 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.

<sup>111</sup> This figure arises because 1.96 times 1.5 is 2.94. As we discuss in Part IV, *infra*, there are a number of potential problems with this typical approach.

allegations. In such a case, the evidence is consistent with the plaintiff's theory of the case but the size of the effect is too small to be statistically significant. Such an outcome may sometimes occur even when the null hypothesis was really false, i.e., there really was a price impact due to the news on the event date.

This last point hints at an inherent tradeoff reflected in statistical significance testing. When one conducts a statistical significance test, there are four possible outcomes. These four categories of statistical inference are summarized in Table 1. Two of these are correct inferences: the test may fail to reject a null hypothesis that is really true, or the test may reject a null hypothesis that is really false. The first of these cases correctly determines that there was no price impact (the upper left box in Table 1). The second case correctly determines that there was a price impact (the lower right box in Table 1). Given that there really was a price impact, the probability of correctly making this determination is known as the test's power.<sup>112</sup>

The other two outcomes are incorrect inferences. The first mistaken inference involves rejecting a null hypothesis that is actually true. This is known as a Type I error (top right box in Table 1). The probability of this result, given that the null hypothesis is true, is known as a test's size.<sup>113</sup> The second incorrect inference is failing to reject a null hypothesis that is actually false (lower left box in Table 1); this is known as a Type II error.<sup>114</sup>

---

<sup>112</sup> Thus power is the probability of winding up in the lower right box in Table 1, given that we must wind up in one of the two lower boxes; it is the ability of the test to identify a price impact when it actually exists.

<sup>113</sup> For this reason, a test with significance level of 5% is sometimes said to have size 0.05.

<sup>114</sup> Given that the null hypothesis is false, so that we must wind up in one of the two lower boxes in Table 1, the probability of a Type II error equals one minus the test's power. See Brav & Heaton, *Event Studies*, note 8, *supra*, 614, n. 26 (citing Paul D. Ellis, *The Essential Guide to Effect Sizes: Statistical Power, Meta-Analysis, and the Interpretation of Research Results* 52 (2010) ("Statistical power describes the probability that a test will correctly identify a genuine effect.")).

**Table 1: Four Categories of Statistical Inference**

	<u>Don't Reject Null</u> Test does not find statistically significant price effect	<u>Reject Null</u> Test finds statistically significant price effect
<u>Null is true</u> No highly unusual price effect	Accurate finding of no price effect	Type I error (Size)
<u>Null is false</u> Highly unusual price effect	Type II error	Accurate finding of price effect (Power)

The tradeoff that arises in statistical significance testing is simple: reducing a test's Type I error rate means increasing its Type II error rate, and vice-versa.<sup>115</sup> As noted above, event study authors usually use a confidence level of 95%, which is the same as a Type I error rate of 5%. The Type II error rate associated with this Type I error rate will depend on the typical range of variability of excess returns, but it has recently been pointed out that insisting on a Type I error rate of 5% when using event studies in securities fraud litigation can be expected to cause very high Type II error rates.<sup>116</sup> Another way to put this is that event studies used in securities litigation are likely to have very low power—very low probability of rejecting an actually false null hypothesis—when we insist on keeping the Type I error rate as low as 5%.<sup>117</sup> We discuss this very important issue further in Part V.

A final issue related to statistical significance concerns who bears the burden of persuasion if the defendant seeks to use event study evidence to show that there was no price impact related to an alleged misrepresentation. *Halliburton II* states that “defendants must be afforded an opportunity before class certification to defeat the presumption through evidence that an alleged misrepresentation did not

---

<sup>115</sup> To be sure, it is sometimes true that two tests have the same Type I error rate but different Type II error rates (or vice-versa). However, the Type II error rate for a given test—such as the significance testing approach typically used in event studies—can be reduced only by increasing the Type I error rate (and vice-versa).

<sup>116</sup> See Brav & Heaton, *Event Studies*, note 8, *supra*, 593-597 (demonstrating that, as a result, the standard event study will fail to detect smaller price impacts reliably).

<sup>117</sup> See *id.*, for an excellent in-depth discussion.

actually affect the market price of the stock.”<sup>118</sup> But it does not tell us what statistical standard will apply to defendants’ evidence. As Merritt Fox discusses, one view is that the defendant must present statistically significant evidence that the price changed in the direction *opposite* to the plaintiff’s allegations.<sup>119</sup> Alternatively, the defendant might have to present evidence that is sufficient only to persuade the court not that its own evidence of the absence of price impact is more persuasive than plaintiff’s affirmative evidence of price impact.<sup>120</sup>

As Fox has noted in other work, the choice of applicable legal standard will have considerable impact on the volume of cases that are able to survive beyond a preliminary stage.<sup>121</sup> Further, Fox points out, a variety of factors affect the choice of approach, including social policy considerations about the appropriate volume of securities fraud litigation.<sup>122</sup> Perhaps for this reason, as of the present writing, the question of Rule 301’s applicability is being litigated at the Fifth Circuit by the *Halliburton* parties.<sup>123</sup> A full discussion of these issues is beyond the scope of the present Article. For concreteness, we will simply follow the approach taken by the District Court in the ongoing *Halliburton* litigation. While that court has found that “both the burden of production and the burden of persuasion are properly placed on Halliburton,”<sup>124</sup> the court does not understand this conclusion to mean that the defendant must affirmatively disprove the plaintiff’s allegations statistically. Rather, Halliburton need only “persuade the Court that its expert’s event studies are more probative of price impact than the Fund’s expert’s event studies.”<sup>125</sup> The rest of the court’s opinion makes clear that this means treating both sides’ event studies as if they are testing whether the statistical evidence is sufficient to establish that there is statistically significant evidence of a price impact as discussed above. We will therefore continue to concentrate on that approach throughout this Article.

---

<sup>118</sup> *Halliburton Co. v. Erica P. John Fund, Inc.*, 134 S.Ct. 2398, 2417 (U.S.,2014).

<sup>119</sup> See Fox, *It All Depends*, note 80, *supra*, 447-454.

<sup>120</sup> *Id.*, 454-455. As Fox discusses, Federal Rule of Evidence 301 provides some support for this second approach. *Id.*, 457. However, Fox also points out a number of complicating issues as to the applicability of Rule 301 to 10b-5 actions.

<sup>121</sup> Merritt B. Fox, *Halliburton II: What It’s All About*, 1 J. Fin. Reg. 135, 139-41 (2015).

<sup>122</sup> *Id.* at 141.

<sup>123</sup> See, e.g., Brief of Appellants Halliburton Co. and David J. Lesar, *Erica P. John Fund, Inc., v. Halliburton Co.*, No. 15-11096, at 52-60 (C.A.5) (February 8, 2016) and Brief of the Lead Plaintiff - Appellee and the Certified Class, *Erica P. John Fund, Inc., v. Halliburton Co.*, No. 15-11096, at 49-58 (C.A.5) (March 28, 2016).

<sup>124</sup> *Erica P. John Fund, Inc., v. Halliburton Co.*, 309 F.R.D. 251, 260 (2015).

<sup>125</sup> *Id.*

The foregoing discussion summarizes the basic methodology of event studies used in securities litigation. In the next Part, we present our own stylized event study of dates involved in the ongoing *Halliburton* litigation both to illustrate the principles described above and to identify some important refinements that experts and courts should make.

### III. THE EVENT STUDY AS APPLIED TO THE *HALLIBURTON* LITIGATION

This Part uses data and methods from the opinions and expert reports in the *Halliburton* case to illustrate and critically analyze the use of an event study to measure price impact. Our objective is, initially, to provide a basic application of the theory described in the preceding Part for those readers having limited familiarity with the operational details. We then introduce, in Part IV, problems with the application of the basic approach and demonstrate the implications of making the necessary adjustments to respond to these problems.

#### A. *Dates and Events at Issue in the Halliburton Litigation*

Plaintiffs in the *Halliburton* litigation alleged that between the middle of 1999 and the latter part of 2001,<sup>126</sup> Halliburton and several of the company's officers—collectively referred to here as simply “Halliburton”—made false and misleading statements about various aspects of the company's business.<sup>127</sup> 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.<sup>128</sup> For purposes of illustration, consider two of the allegedly fraudulent statements:

---

<sup>126</sup> We focus on the class period at issue at the time of the most recent district court order, which ran from July 22, 1999 to December 7, 2001. The class period referred to in the operative complaint began slightly earlier, on June 3, 1999. FCAC at 1, ¶ 1. The difference is immaterial for our purposes.

<sup>127</sup> Fourth Amended Consolidated Complaint, *Archdiocese of Milwaukee Supporting Fund, Inc., v. Halliburton Co.*, 3:02-CV-1152-M, No. 246-11, at 1-2, ¶ 2 (N.D.Tex., April 4, 2006) [henceforth “FCAC”].

<sup>128</sup> A defense expert report lists 25 distinct dates on which plaintiffs or their expert alleged misrepresentations; see Expert Report of Lucy P. Allen, *Archdiocese of Milwaukee Supporting Fund, Inc., v. Halliburton Co.*, 3:02-CV-1152-M, No. 572-2, at 5-7, ¶ 10 (September 10, 2014) [henceforth, “Allen Report”].

- (1) Plaintiffs alleged that in its 1998 10-K report filed on March 23, 1999, 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.<sup>129</sup>
- (2) On November 8, 2001, 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.”<sup>130</sup> Plaintiffs also alleged that this representation was false and misleading.<sup>131</sup>

Both the alleged misrepresentations described above were confirmatory, in the sense that the plaintiffs alleged that Halliburton, rather than accurately informing the market of negative news, falsely confirmed prior good news that was no longer accurate.<sup>132</sup> The alleged result was that Halliburton’s stock price was inflated because it remained at a higher level than it would have had Halliburton disclosed accurately. Since false confirmatory misrepresentations do not constitute “new” information, even under the plaintiff’s theory, neither of the two statements above would have been expected to cause an increase in Halliburton’s market price. As a result, in considering the price impact of the alleged misrepresentations, the district court allowed plaintiffs to focus on whether subsequent alleged corrective disclosures were associated with reductions in Halliburton’s stock price.<sup>133</sup>

---

<sup>129</sup> FCAC, note 127, *supra*, at 58, ¶ 74.

<sup>130</sup> FCAC, note 127, *supra*, at 126, ¶ 189.

<sup>131</sup> FCAC, note 127, *supra*, at 126-28, ¶ 190.

<sup>132</sup> See *Archdiocese of Milwaukee Supporting Fund, Inc. v. Halliburton Co.*, 2008 U.S. Dist. LEXIS 89598 (N.D. Tex. 2008).

<sup>133</sup> *Erica P. John Fund v. Halliburton Co.*, 309 F.R.D. 251, 262 (N.D. Tex. 2015) (“Measuring price change at the time of the corrective disclosure, rather than at the time of the corresponding misrepresentation, allows for the fact that many alleged misrepresentations conceal a truth.”). As discussed in Part I, 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

On July 25, 2015, the district court issued its most recent order and memorandum opinion concerning class certification.<sup>134</sup> By this point of a litigation that had been ongoing for more than 13 years, the event studies submitted by the parties' experts<sup>135</sup> focused on six dates on which plaintiffs Halliburton had issued alleged corrective disclosures: December 21, 2000;<sup>136</sup> June 28, 2001;<sup>137</sup> August 9, 2001;<sup>138</sup> October 30, 2001;<sup>139</sup> December 4, 2001;<sup>140</sup> and December 7, 2001.<sup>141</sup>

The trial court concluded in its July 2015 decision, after weighing two competing expert reports, that five of these alleged corrective disclosures did not have a statistically significant price impact. For that reason, the district court denied class certification with respect to these five dates.<sup>142</sup> 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

---

though the report is false in that the company has actually lost money (presumably when that loss is disclosed the share price will fall).

*Nathenson v. Zonagen Inc.*, 267 F.3d 400, 419 (C.A.5 (Tex.), 2001). In contrast, 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, leading us to analyze the corrective disclosure date. *But see* Part IV.C.3, *infra*, which considers the situation when plaintiffs must establish price impact on both an alleged misrepresentation date and an alleged corrective disclosure date.

<sup>134</sup> 309 F.R.D. 251.

<sup>135</sup> See Expert Report of Chad Coffman (NDTX, 3:02-cv-01152-M CM/ECF No. 589, App. 004-138) (plaintiffs' expert) [henceforth, "Coffman Report"] and Expert Report of Lucy P. Allen (NDTX, 3:02-cv-01152-M CM/ECF No. 573-1) (09/10/14) (defendants' expert).

<sup>136</sup> On this date, "Halliburton announced a \$120 million charge which included \$95 million in project costs, some of which allegedly should not have been previously booked." Coffman Report at 3, ¶ 8 (citing FCAC at ¶ 150).

<sup>137</sup> On this date, "Halliburton disclosed that" third-party "Harbison-Walker asked for asbestos claims related financial assistance from Halliburton." Coffman Report at 3, ¶ 8 (citing FCAC at ¶ 170).

<sup>138</sup> On this date, Halliburton's "2Q01 10-Q included additional details regarding asbestos claims." Coffman Report at 3, ¶ 8 (citing FCAC at ¶ 178).

<sup>139</sup> On this date, "Halliburton issued a press release announcing the Mississippi Verdict." Coffman Report at 3, ¶ 8 (citing Halliburton 8-K).

<sup>140</sup> On this date, "Halliburton announced Texas judgment and three other Judgments." Coffman Report at 3, ¶ 8 (citing FCAC at ¶ 191).

<sup>141</sup> On this date, "Halliburton announced Maryland verdict." Coffman Report at 3, ¶ 8 (citing FCAC at ¶ 191).

<sup>142</sup> 309 F.R.D. at 280.



from the *Basic* presumption. The court therefore certified a class action with respect to the alleged misrepresentations associated with December 7, 2001.<sup>143</sup>

*B. An Illustrative Event Study of the Six Dates at Issue in the Halliburton Litigation*

Following the approach outlined in Part II, we apply the event study to the six dates listed in section A, *supra*. For our first step, selection of an appropriate event, we follow the parties and analyze the dates of alleged corrective disclosures.<sup>144</sup>

Next, we use the market model to construct Halliburton's estimated return.<sup>145</sup> 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

---

<sup>143</sup> *Id.* Halliburton subsequently requested and received permission to pursue an interlocutory appeal of the class certification order, pursuant to Rule 23(f); see *Erica P. John Fund, Inc. v. Halliburton Co.*, 2015 WL 10714013 (C.A.5, 2015). The issues on appeal concern not the statistical aspects of event study evidence, but rather are related to the District Court's determination that Halliburton may not, at the class certification stage, provide non-statistical evidence challenging the status of news as a corrective disclosure. *Erica P. John Fund, Inc. v. Halliburton Co.*, 309 F.R.D. 251, 260-262 (N.D.Tex. 2015). As of this writing, oral argument in the Fifth Circuit is scheduled for the afternoon of August 31, 2016. *Erica P. John Fund, Inc. v. Halliburton Company, et al.*, No. 15-11096, June 15, 2016 (5th Cir.) (Docket entry on July 12, 2016). 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.].

<sup>144</sup> We do not independently address the legal question as to whether the disclosures made on the designated event dates are appropriately classified as corrective disclosures, as the trial court determined that whether a disclosure was correctly classified as corrective was not properly before the court at the class certification stage. 309 F.R.D. at 261-62 ("the issue of whether disclosures are corrective is not a proper inquiry at the certification stage.").

<sup>145</sup> Since the possibility of unusual stock return behavior is the object of an event study in the case, these dates should be removed from the set used in estimating the market model, and we do exclude them. 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).

performance of the S&P 500 Energy Index during the class period.<sup>146</sup> The plaintiffs' expert pointed out that this index is dominated by "petroleum refining companies, not energy services companies like Halliburton."<sup>147</sup> In his own market model, he therefore added a second index intended to more closely reflect the performance of Halliburton's industry peers,<sup>148</sup> which we also included.<sup>149</sup> 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.<sup>150</sup> 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<sup>151</sup> are set forth in Table 2.<sup>152</sup> These estimates indicate that Halliburton's daily stock return moves 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. This makes the industry peer index a good tool for estimated Halliburton's expected return in the absence of fraud. 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.

---

<sup>146</sup> 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.

<sup>147</sup> Coffman Report at 16, ¶ 28.

<sup>148</sup> 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.

<sup>149</sup> We calculated the return on this index in the same way as the return on the energy index described in note 146, *supra*; we took the list of included companies from Exhibit 3b of the Coffman Report.

<sup>150</sup> We took the list of companies for this index from the Allen Report, note 128, *supra*, at 12, n. 20.

<sup>151</sup> These estimates are calculated using the ordinary least squares estimator.

<sup>152</sup> 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 logarithm of one plus the daily return, as experts sometimes do. For simplicity we therefore decided to stick with the raw daily return.

**Table 2: Market Model Regression Estimates**

Variable	Coefficient Estimate	Estimated standard error
Industry Peer Index	0.903	0.031
Energy Index	0.210	0.048
E&C Index	0.033	0.036
Intercept	-0.001	0.001
Root mean squared error	1.745%	
Number of dates	593	

We then use these market model coefficient estimates to calculate daily estimated excess returns for the six event dates excluded from estimation of the model. We calculated the contribution of each index to each date's expected return by multiplying the index's Table 2 coefficient estimate by the observed value of the index on the 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, i.e., the variable plotted on the horizontal axis of Figure 1 and Figure 3. The excess return for each event date is then found by subtracting each date's estimated expected return from its actual return.

Table 3 reports the actual, estimated expected, and estimated excess returns for each of the six alleged corrective disclosure dates in the *Halliburton* litigation, sorted from 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. On three of the dates, the estimated expected return was also negative, indicating that typical market factors would be expected to cause Halliburton's stock price to fall, even in the absence of any unusual event. For the other three dates, market developments would have been expected to cause an increase in Halliburton's stock price. This means the estimated excess returns on those dates will imply larger price drops than are reflected in the actual returns. Finally, the estimated excess return column in Table 3 shows that the 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.

**Table 3: Actual, Expected, and Excess Returns for Event Dates**

Event Date	Actual Return	Estimated Expected Return	Estimated Excess Return
December 7, 2001	-42.4%	0.3%	-42.7%
August 9, 2001	-4.5%	0.6%	-5.1%
December 4, 2001	-0.7%	2.9%	-3.6%
December 21, 2000	-2.0%	-0.8%	-1.2%
October 30, 2001	-5.2%	-4.3%	-0.9%
June 28, 2001	-3.8%	-3.1%	-0.8%

The next step is to test these estimated excess returns for statistical significance in order to determine whether they are highly unusual, that is, whether they are large enough such that we should reject the possibility that they might simply reflect essentially random movements in price.

For the moment, we adopt the standard assumption that Halliburton stock's excess returns follow a normal distribution. Our Table 2 above reports that the root mean squared error for our Halliburton market model—which is an estimate of the standard deviation of excess returns—was 1.745%. If we multiply 1.96 times 1.745, we obtain a critical value of 3.42%. In other words, in the absence of unusual events affecting Halliburton's stock price, we can expect that 95% of Halliburton's excess returns will take on values between -3.42% and 3.42%.<sup>153</sup> For an alleged corrective disclosure date, excess returns must be negative to support the plaintiff's theory, so a typical expert would determine that an event date excess return drop of 3.42% or more is statistically significant.

In the first column of Table 4, we again present the estimated excess returns from Table 3. 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 December 4, 2001; they are statistically insignificant for the other three dates.

---

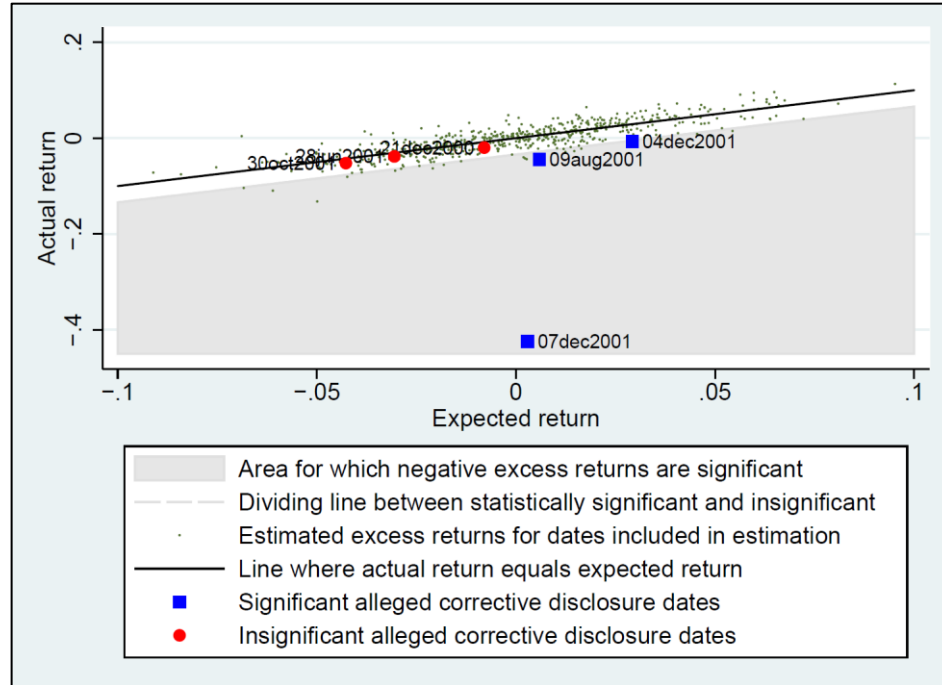
<sup>153</sup> This follows because 1.96 times 1.745 equals 3.4202.

**Table 4: Standard Significance Testing for Event Dates  
(sorted by magnitude of estimated excess return)**

Event Date	Estimated Excess Return	Critical Value	Statistically Significant Using Standard Approach?
December 7, 2001	-42.7%	-3.42%	Yes
August 9, 2001	-5.1%	-3.42%	Yes
December 4, 2001	-3.6%	-3.42%	Yes
December 21, 2000	-1.2%	-3.42%	No
October 30, 2001	-0.9%	-3.42%	No
June 28, 2001	-0.8%	-3.42%	No

We can illustrate the standard approach by again using a graph that relates actual and expected returns. As in earlier figures, Figure 2 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.

**Figure 2: Scatter Plot of Actual and Expected Returns for Alleged Corrective Disclosure Dates and for Observations in Estimation Period**



In addition, the figure includes three larger circles and three larger squares. The 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 4 tells us estimated excess returns were negative (below the upwardly sloped line), but were not statistically significant according to the standard approach. The 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 4—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 3, *supra*, points in this region have statistically significant price drops according to the standard approach. In sum, our event study shows price impact for three dates, and it fails to show price impact for the other three.

#### IV. SPECIAL FEATURES OF SECURITIES FRAUD LITIGATION AND THEIR IMPLICATIONS FOR THE USE OF EVENT STUDIES

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, each of these assumptions is false in the context of the *Halliburton* litigation. The court did take appropriate account of the falsity of the third assumption, involving multiple comparisons,<sup>154</sup> but it failed even to address the other three.

Yet violations of any of these assumptions will render the standard approach to testing for statistically significant unreliable. 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.

In the remaining sections of this Part, we explain these four assumptions in more detail, and we show that they are unsustainable in the context of the *Halliburton* event study conducted in Part III.

---

<sup>154</sup> 309 F.R.D. 251, 265-267 (2015).

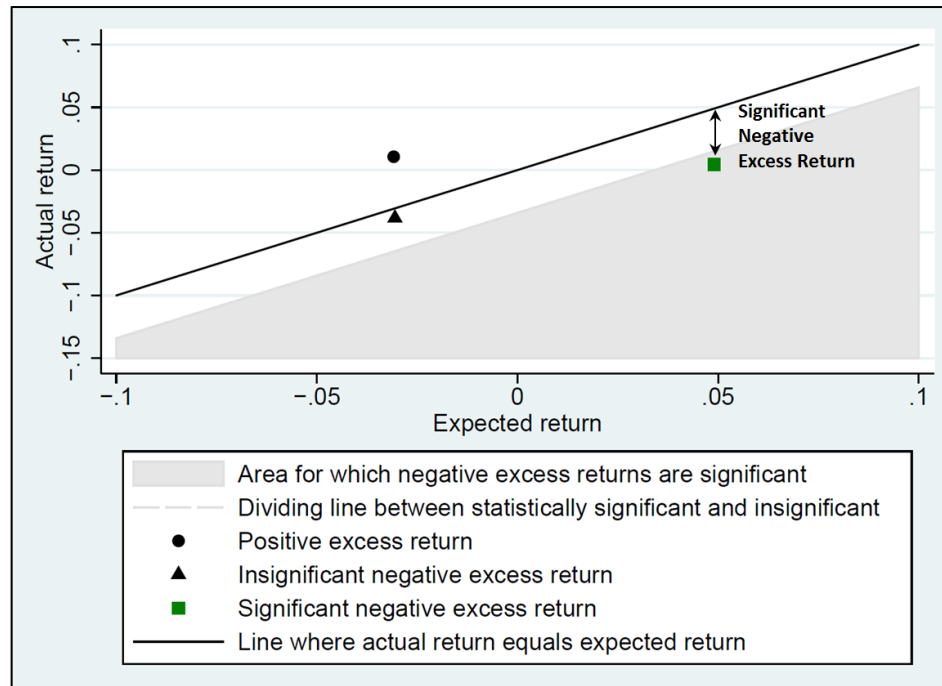
*A. The Inappropriateness of Two-Sided Tests*

In a purely academic study, economic theory may not predict whether an event date excess return can be expected to be positive or negative. For example, an announced merger might be either good or bad for a firm's market valuation. In such cases, statistical significance is appropriately tested by checking whether the estimated excess return is large in magnitude regardless of its sign. In other words, either a very large drop or a very large increase in the firm's stock price constitutes evidence against the null hypothesis that the news had no impact on stock price. Such tests are known as "two-sided" tests of statistical significance, since a large value of the excess return on either side of zero provides evidence against the null hypothesis.

In event studies used in securities fraud litigation, by contrast, price must move in a specific direction to support plaintiff's case. For example, an alleged corrective disclosure is expected to cause the stock price to fall. Thus, tests of statistical significance based on event study results should be conducted in a "one-sided" way, so that an estimated excess return is considered statistically significant only if it moves in the direction consistent with the allegations of the party using the study. The one-sided/two-sided distinction is one that courts and expert witnesses regularly miss, and it is an important one.

Figure 3 illustrates this point. As in Figure 1, the upwardly sloped line indicates the set of points where the actual and excess returns are equal. The shaded area in Figure 3 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.



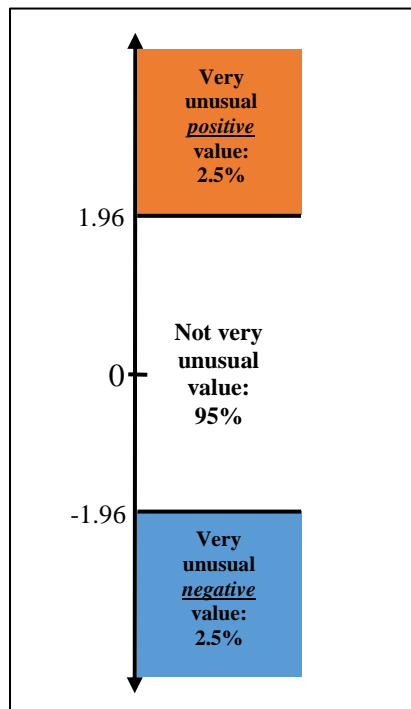
**Figure 3: Illustrating Statistical Significance of Excess Returns**

Consider the points indicated by a circle and a square in Figure 3, which are equally far from the actual-equals-expected line, but in opposite directions. The circle depicts a point that has a positive excess return. Even though the circle is sufficiently far away from the line, the point has the wrong sign for an alleged corrective disclosure date, and no court would consider such evidence a basis on which to find for the plaintiff. The square, in contrast, 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. Finally, consider the point indicated by a triangle. This point is in the direction consistent with the plaintiff's allegations—a negative excess return, for an alleged corrective disclosure—but at this point the actual and expected returns are too close for the excess return to be statistically significant. For an alleged corrective disclosure date, only the square would provide the plaintiff's required proof.

If no litigant would present evidence of a statistically significant price movement in the wrong direction, why does the two-sided approach matter? The reason is that the practical effect of this approach is to reduce the Type I error rate for the tests used in event studies from

the stated level of 5% to half that size, i.e., to 2.5%. To see why, consider Figure 4. Higher points in the figure correspond to larger and more positive estimated excess returns. The shaded regions correspond to the sets of excess returns that are further from zero than the critical value of 1.96 standard deviations used by experts who deploy the two-sided approach. For each shaded region, the probability that a randomly chosen excess return will wind up in that region is 2.5%. Thus the probability an excess return will be in either region, and thus that the null hypothesis would be rejected if event study experts followed usual two-sided practice—is 5% in total, which is the desired the Type I error rate.

**Figure 4: The Standard Approach to Testing on an Alleged Corrective Disclosure Date with a Type I Error Rate of 5% (Measured in standard deviation units)**



However, on an alleged corrective disclosure 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. In other words, only estimated

excess returns that are in the bottom shaded region in Figure 4 would meet the plaintiff's burden. As we have seen, this region contains 2.5% of the probability when there is no actual effect of the news in question.<sup>155</sup> This means that a finding of statistical significance would occur only 2.5% of the time when the null hypothesis is true—or half as frequently as the 5% that courts and experts say they are attempting to apply.<sup>156</sup>

Although a reduction in Type I errors is desirable with all else held equal, as we discussed in Part II.B, *infra*, there is a tradeoff between Type I and Type II error rates. As a result of this tradeoff, the Type II error rate of a test rises—possibly dramatically—as the Type I error rate is reduced. This means that using a Type I error rate of 2.5 percent in an event study induces many more false negatives than using a Type I error rate of 5 percent.<sup>157</sup>

This mistake is easily corrected. Rather than base the critical value on the two-sided testing approach, one merely uses a one-sided critical value. In terms of Figure 4, that means choosing the critical value so that a randomly chosen excess return would turn up in the bottom shaded region 5% of the time, given that the news of interest actually had no impact. Still maintaining the assumption that excess returns are normally distributed, the relevant critical value is -1.645 times the standard deviation of the stock's excess returns.<sup>158</sup> In our application, this yields a critical value for an event date excess return of -2.87%; any excess return more negative than this value will yield a finding of statistical

---

<sup>155</sup> The fact that two-tailed tests are erroneous has been noted in recent literature. See Brav & Heaton, *Event Studies*, note 8, *supra*, at 614 (2015); see also Edward G. Fox, Merritt B. Fox, and Ronald J. Gilson, *Economic Crisis and the Integration of Law and Finance: The Impact of Volatility Spikes*, 116 Colum. L. Rev. 325, 407 (2016) (acknowledging that the usual two-tailed test delivers a Type I error rate of only 2.5%); Fox, *It All Depends*, note 80, *supra*, 464 (same). Those authors seem to accept 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.

<sup>156</sup> A method that delivers many more false negatives than claimed surely raises important *Daubert* and Fed. R. EvId. 702 concerns. See *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579, 594 (U.S. Cal., 1993).

<sup>157</sup> We discuss power implications of this issue in Part V.C, *infra*, at 59.

<sup>158</sup> This is so because a normally distributed random variable will take on a value less than -1.645 times its standard deviation 5% of the time. If one were testing for statistical significance on the date of a non-confirmatory alleged misrepresentation, one would use a critical value of 1.645 times the standard deviation of the excess return, since a normally distributed random variable will take on a value *greater* than 1.645 times its standard deviation 5% of the time.

significance.<sup>159</sup> This is a considerably less demanding critical value than the -3.42% based on the two-sided approach. Consequently, an appropriate switch to the one-sided test will correct an erroneous finding of no statistical significance whenever the estimated excess return is between -3.42% and -2.87%.

As it happens, none of the estimated excess returns in Table 4 has a value in this range, so correcting this error does not affect any of the statistical significance determinations we made in Part III for Halliburton. But that is just happenstance; had any of the estimated excess returns fallen in this range, our statistical significance conclusion would have changed. Further, Halliburton's median daily market value was \$17.6 billion over the estimation period, so the range of estimated excess returns that would have led to a switch—i.e., -3.42% to -2.87%—corresponds to a range of Halliburton market value of nearly \$100 million. In other words, using the erroneous approach would, in the case of Halliburton, require a market value drop of almost \$100 million more than should be required to characterize the drop as highly unusual.

#### B. *Non-normality in Excess Returns*

Recall that, as discussed above, we characterize an excess return as highly unusual by looking at the distribution of excess returns on days where there is no news. The standard event study assumes that this distribution is normal.<sup>160</sup> There is no good reason, however, to assume that excess stock returns are actually normally distributed, and there is considerable evidence against normality.<sup>161</sup> 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.<sup>162</sup>

In the case of Halliburton, we found strong evidence that the excess returns distribution was non-normal over the class period. Summary

---

<sup>159</sup> This critical value is the product of -1.645 and the estimated standard deviation of 1.745%:  $-1.645 \times 1.745\% = -2.87\%$ .

<sup>160</sup> See generally Jonah B. Gelbach, Eric Helland, and Jonathan Klick, *Valid Inference in Single-Firm, Single-Event Studies*, 15 AM. L. & ECON. REV. 495 (2013) [henceforth, “GHK”] (discussing normal distribution).

<sup>161</sup> For early evidence on non-normality, see Stephen J. Brown and Jerold B. Warner, *Using Daily Stock Returns: The Case of Event Studies*, 14 J. Fin. Econ. 3, 4-5 (1985). For more recent evidence in the single-firm, single-event context, see GHK, note 160, *supra*, at 511, 534-538.

<sup>162</sup> 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.

statistics indicate that Halliburton's excess returns exhibit negative skew: they are more likely to have positive values than negative ones. Further, the distribution has fat tails, with values far from the distribution's center than would be the case if excess returns were normally distributed. Formal statistical tests reinforce this story: Halliburton's estimated excess returns systematically fail to follow a normal distribution over the estimation period.<sup>163</sup>

We illustrate the role of the normality assumption in Figure 5, which plots various probability density functions for excess returns. Roughly speaking, a probability density function tells us the frequency with which a given value of the excess return is observed. The probability of observing an excess return value less than, say,  $X$  is the area between the horizontal axis and the probability density function, for all values less than  $x$ . The curve plotted with a solid line in the top part of Figure 5 is the familiar density function for a normal distribution (also known colloquially as a bell curve) with standard deviation equal to 1. The shaded area to the left of the point where the excess return equals  $-1.645$  equals  $0.05$ , reflecting the fact a normal random variable will take on a value less than  $-1.645$  units of standard deviation  $5\%$  of the time. To put it differently, the  $5^{\text{th}}$  percentile of standard normal distribution is  $-1.645$ , which is why we use this figure for the critical value to test for a price drop at significance level  $5\%$  when excess returns are normally distributed.

The curve plotted with a dashed line in the top part of Figure 5 is the probability density function for a different distribution. Compared to the standard normal distribution, the left-tail percentiles of this second distribution are compressed toward its center. That means fewer than  $5\%$  of this distribution's excess returns will take on a value less than  $-1.645$ ; the  $5^{\text{th}}$  percentile of this distribution is closer to zero, equal to roughly  $-1.36$ . Thus, when the distribution of excess returns is compressed toward zero relative to the normal distribution, we must use a more forgiving critical value—one closer to zero—to test for a significant price drop.

The bottom graph in Figure 5 again plots the standard normal distribution's probability density function with a solid line. In contrast to the top graph, the curve plotted with a dashed line now depicts a distribution of excess returns whose left-tail percentiles are splayed out compared to the normal distribution. The  $5^{\text{th}}$  percentile is now  $-2.35$ , so

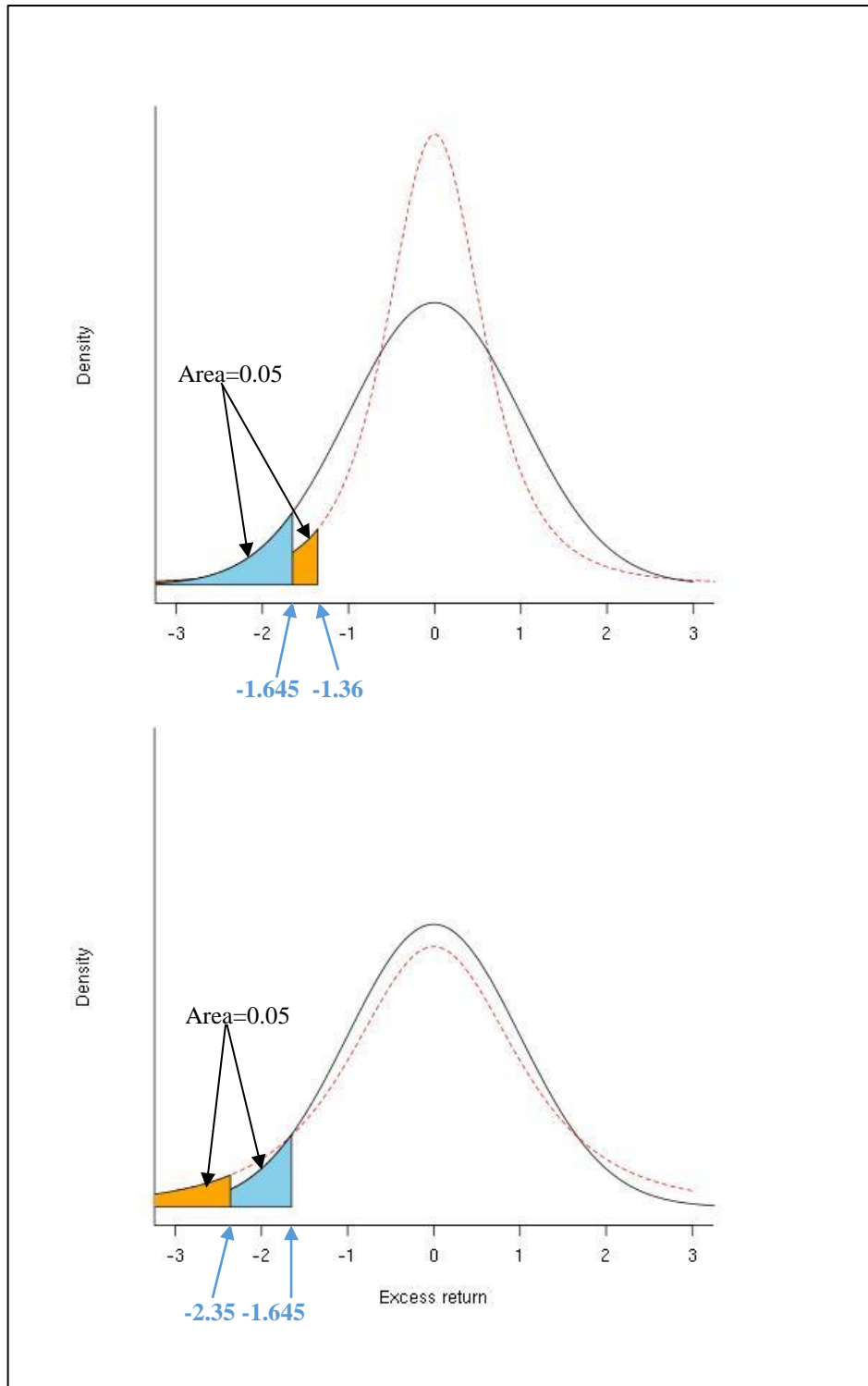
---

<sup>163</sup> 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.

that we must use a more demanding critical value—one further from zero—to test for significance.

As this discussion illustrates, the assumption that excess returns are normally distributed is not innocuous: if the assumption is wrong, an event study analyst might use a very different critical value from the correct one.

Figure 5: Illustrating Non-Normality



It might seem a daunting task to determine the true distribution of the excess return. However, 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.<sup>164</sup> This result provides a simple correction to the normality assumption: instead of using the features of the normal distribution to determine the critical value for statistical significance testing, the critical value we use is the 5<sup>th</sup> percentile of the distribution of excess returns estimated using our market model.<sup>165</sup> GHK describe this percentile approach as the “SQ test”, since the approach relies for its theoretical justification on the branch of theoretical statistics that concerns the behavior of sample quantiles; for our purposes, though, sample quantiles are simply percentiles.

For a statistical significance test with Type I error rate of 5%, the SQ test entails using a critical value equal to the 5<sup>th</sup> percentile of the estimated excess returns distribution (among non-event dates). Among the 593 non-event dates in our class period estimation sample, the 5<sup>th</sup> percentile is -3.08%.<sup>166</sup> According to GHK’s SQ test, then, this is the

---

<sup>164</sup> GHK, *Valid Inference*, note 160, *supra*. GHK actually use somewhat different notation; the estimated excess return described in the present Article is the same as GHK’s  $\hat{\gamma}$  regression parameter. With this difference noted, our point about statistical properties is demonstrated in GHK’s Appendix B. *Id.*, at 538-539. This result is practically useful 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.

<sup>165</sup> The SQ test 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, e.g.*, William H. Greene, *ECONOMETRIC ANALYSIS* (7<sup>th</sup> Ed. 2011). Unpublished tabulations from the data GHK used show that the SQ test performs extremely well even when using estimation period sample sizes considerably lower than the 250 days used here. The underlying 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 5<sup>th</sup> 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 percentiles of the estimated excess returns distribution. For details, *see* section 5.1 of GHK, *Valid Inference*, note 160, *supra*, at 517-520.

<sup>166</sup> We find the 5<sup>th</sup> percentile of a sample by multiplying the number of dates in the sample by 0.05, which yields 29.65. Conventionally, this means that the 5<sup>th</sup> percentile



value we should use as the critical value for testing whether event date excess returns are statistically significant. 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 conclude that an alleged corrective disclosure date's estimated excess return is statistically significant if it is less than -3.08%.

Note that this critical value is greater than the value of -2.87% found in section IV.A, *supra*, where we maintained the assumption of normality. Thus relaxing the normality assumption has the effect of making the standard for a finding of statistical significance about 0.21 percentage points more demanding.<sup>167</sup> Although this correction does not affect our determination as to any of the six event dates in our Halliburton event study, it is nonetheless potentially quite important because 0.21 percentage points corresponds to a range of Halliburton's market value of nearly \$40 million.

As we discuss in Appendix A, the SQ test has both statistical and operational characteristics that make it very desirable. First, it 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. Thus the operational demands of using the SQ test are thus minor, and we think experts and courts should adopt it. And second, the SQ test not only is appropriate in many instances where the normality assumption fails, but also is always appropriate when the normality assumption is valid. Thus there is no cost to using the SQ test, by comparison to the standard approach of assuming normality.

### C. Multiple event dates of interest

The approaches to statistical significance testing discussed above were all designed for situations involving the analysis of a single event date. As we have seen, however, there are six alleged corrective disclosure dates at issue in the *Halliburton* litigation. The distinction is important.

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

---

lies between the 29<sup>th</sup> and 30<sup>th</sup> most negative estimated excess returns; in our sample, these are -3.089066% and -3.074954% (the shares of estimated excess returns less than or equal to these values are 4.89% and 5.06%. Their midpoint is -3.08201%, which is our estimate of the 5<sup>th</sup> percentile.

<sup>167</sup> That is, an estimated excess return must now be more negative than -3.08%, rather than -2.87%, to be found statistically significant.

more bites at the same apple, and the odds the apple will be eaten up increase with the number of bites. At the same time, however, securities litigation differs from the example in that multiple events do not always relate to the same fraud. Corrective disclosures relating to different misstatements are different pieces of fruit. We discuss the multiple comparison adjustment first, in section 1, and then, in section 2, we explain an approach for determining when such an adjustment is warranted. In section 3, we address the very different statistical problem raised by a situation in which a plaintiff must prove *both* the existence of price inflation on the date of an alleged misrepresentation *and* the existence of a price drop on the date of an alleged corrective disclosure.<sup>168</sup>

*1. When the question of interest is whether any disclosure had an unusual effect*

In our event study analysis so far, we have tested for statistical significance as if each of the six event dates' estimated excess returns constituted the only one being tested. As mentioned above, this means the probability of finding *at least* one event date's estimated excess return significant will be considerably greater than the desired Type I error rate of 5%. The defendants raised the multiple comparison issue in the *Halliburton* litigation, and it played a substantial role in the court's analysis.<sup>169</sup>

Various statistical approaches exist to correct for this problem.<sup>170</sup> One approach is called the Holm-Bonferroni *p*-value correction. The district court used this approach in *Halliburton*.<sup>171</sup> To understand this correction, it is first necessary to explain the term *p*-value. The *p*-value can be viewed as another way of describing statistical significance. In terms of our prior analysis, if the estimated excess return for a single date

---

<sup>168</sup> Cases that present a combination of the questions addressed in sections 1 and 2 are more complicated notationally and mathematically; we address such cases in Appendix A to this Article.

<sup>169</sup> See, e.g., 309 F.R.D. 251, 265-57 (2015) (finding that "a multiple comparison adjustment is proper in this case").

<sup>170</sup> . 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 multiple testing methodology involves some fairly technical mathematical details, we will not discuss it in detail. A brief but exceedingly clear discussion appears in Hervé Abdi, *Holm's Sequential Bonferroni Procedure*, in Neil Salkind (Ed.), *ENCYCLOPEDIA OF RESEARCH DESIGN*, Thousand Oaks, CA: Sage (2010).

<sup>171</sup> See, e.g., discussion at 309 F.R.D. 251, 265-67 (2015).

was found statistically significant at the 5% level, then the  $p$ -value for that date must be less than or equal to 0.05. If, on the other hand, the estimated excess return was found not statistically significant, then the  $p$ -value must be above 0.05. We will refer to  $p$ -values that are computed as if only a single date were being tested as “usual”  $p$ -values; this allows us to distinguish between usual and multiple comparison-adjusted  $p$ -values.

Calculating the usual  $p$ -value for an alleged corrective disclosure date when using the one-sided SQ test requires counting up the number of estimated excess returns from the market model estimation period that are more negative than the estimated excess return on the event date, and then dividing 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 5; the third column reports whether price impact was found statistically significant using the one-sided SQ test. Note that 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 5: Controlling for Multiple Testing  
Using the Holm-Šidák Approach**

Event Date	Excess Return	One-Sided SQ Approach, Ignoring Multiple Testing Issue		One-Sided SQ Approach, With Šidák Correction for Multiple Testing	
		$p$ -value	Statistically Significant?	$p$ -value	Statistically Significant?
December 7, 2001	-42.7%	0	Yes	0	Yes
August 9, 2001	-5.1%	0.0017	Yes	0.0034	Yes
December 4, 2001	-3.6%	0.0269	Yes	0.0787	<b><u>NO</u></b>
December 21, 2000	-1.2%	0.2222	No	0.6340	No
October 30, 2001	-0.9%	0.2609	No	0.7795	No
June 28, 2001	-0.8%	0.3013	No	0.8837	No

The fourth column of the table reports  $p$ -values that are corrected for multiple testing.<sup>172</sup> The final column reports whether the Holm-Šídák  $p$ -value is less than 0.05, in which case there is statistically significant price impact even after adjusting for the presence of multiple tests.<sup>173</sup> Table 5 shows that after correcting for multiple testing, we find 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.

---

<sup>172</sup> There are different flavors of  $p$ -values that correct for multiple comparisons. The type we have reported in the Table is known as Šídák. To calculate the Šídák  $p$ -value for the event date with the lowest usual  $p$ -value is just that usual  $p$ -value; thus the  $p$ -value for the excess return on December 7, 2001, is unaffected by the correction for multiple comparisons. Let the second lowest usual  $p$ -value be called  $p_2$  (December 4, 2001, in our event study). The formula for the Šídák  $p$ -value for this date is  $p_{S2} = 1 - (1 - p_2)^2$ . The logic of this formula is that the probability of independently drawing two excess returns that are more negative than the usual  $p$ -value actually observed on this date, i.e.,  $p_{S2}$ , is  $(1 - p_2)^2$ ; thus the probability of *not* drawing a more negative excess return is  $p_{S2}$ . The value  $p_{S2}$  is thus the probability of taking two draws from the excess returns distribution and observing at least one with a more negative excess return than  $p_2$ . It can be shown that when this probability is less than 0.05, the underlying statistic is statistically significant at the 5% level.

For the event date with the third lowest usual  $p$ -value, which we will call  $p_3$ , the formula for the Šídák  $p$ -value for this date is  $p_{S3} = [1 - (1 - p_3)^3]$ ; again the logic is that this is the probability of drawing repeatedly (now, three times) from the excess returns distribution and obtaining an excess return that is more negative than the date in question. In general, let the usual  $p$ -value for the date with the  $m^{\text{th}}$ -lowest usual  $p$ -value be  $p_m$ ; then the Šídák  $p$ -value for this date is  $p_{Sm} = 1 - (1 - p_m)^m$ . We note also that for small values of  $p_m$  and small values of the exponent  $m$ , Šídák  $p$ -values are well approximated by  $m \times p_m$ , which is known as the Bonferroni  $p$ -value. In our application it turns out not to matter which of the two approaches we use, though in general the Šídák  $p$ -value is more accurate than the Bonferroni  $p$ -value. The District Court in the *Halliburton* litigation addressed the choice between Bonferroni and Šídák  $p$ -values, because experts in the case debated which was more appropriate, though in this case the choice makes no difference to the actual statistical significance determinations. *Erica P. John Fund, Inc. v. Halliburton Co.*, 309 F.R.D. 251, 265-67 (N.D. Tex. 2015).

<sup>173</sup> That is, we consider the price impact on the date with the second-lowest  $p$ -value to be significant if and only if its Šídák  $p$ -value is less than 0.05 *and* the . If date 6'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 Šídák  $p$ -value is less than 0.05; if not, we stop, but if so, we turn to date 3's price impact, and so on.

2. *How should events be grouped together to adjust for multiple testing?*

A critical threshold question before applying a multiple comparison adjustment is to determine which, if any, of a plaintiff's multiple alleged corrective disclosure dates should be grouped together. In the preceding section we grouped all dates together, because that is the approach the District Court took in the *Halliburton* litigation. However, it is not clear that this is the best—or even a good—approach. As noted, using multiple event dates gives the plaintiff an advantage by increasing the chance of achieving statistical significance with respect each transaction.

How do we identify which disclosure dates to group together? A full analysis of this mixed question of law and advanced statistical methodology is beyond the scope of this Article, but one simple solution is to draw an analogy to general principles of claim preclusion. Rule 18(a)'s generous claim joinder rule allows, but does not require, a plaintiff to bring all possible claims in a single lawsuit.<sup>174</sup> Thus, a plaintiff might choose to bring separate actions, with only a subset of alleged corrective disclosure dates at issue in each action. The rules of claim preclusion impose a limit on plaintiffs' power to litigate multiple claims independently, however, by looking to whether two claims are sufficiently closely related.<sup>175</sup> If so, a judgment on one such claim will preclude a separate cause of action on the second.

We suggest that if a losing judgment in Claim 1 would preclude a plaintiff from prevailing on Claim 2, then it is reasonable for the district court to consider all alleged corrective disclosure dates for the two claims together for purposes of multiple comparisons. Contrariwise, if losing on Claim 1 would not preclude as to Claim 2, then, we suggest, the alleged corrective disclosure dates related to the two claims should be treated separately. This rule would ensure that in addressing multiple alleged corrective disclosure dates, courts require a consistent quantum of statistical evidence to obtain class certification across collections of dates concerning the same or related misstatements—i.e., claims which plaintiffs would naturally be expected to litigate together. Basing this test

---

<sup>174</sup> Fed. R. of Civ. P. Rule 18(a) (“A party asserting a claim . . . may join, as independent or alternative claims, as many claims as it has against an opposing party.”).

<sup>175</sup> Whether the claims are closely enough related is likely to be governed by Restatement (Second) of Judgments § 24's “transaction” test. The Restatement is of course not *per se* binding on federal courts, but the Supreme Court has endorsed the Restatement's approach. *See, e.g., U.S. v. Tohono O'Odham Nation*, 563 U.S. 307, 316 (U.S., 2011) (“The now-accepted test in preclusion law for determining whether two suits involve the same claim or cause of action depends on factual overlap, barring ‘claims arising from the same transaction.’ *Kremer v. Chemical Constr. Corp.*, 456 U.S. 461, 482, n. 22, 102 S.Ct. 1883, 72 L.Ed.2d 262 (1982); see also Restatement (Second) of Judgments § 24 (1980).”).

on the law of claim preclusion prevents future plaintiffs from gaming the system by attempting to bring multiple lawsuits in order to avoid the multiple comparison adjustment. At the same time, our rule would not penalize a plaintiff for bringing two unrelated claims in the same action—thereby respecting and reinforcing the baseline set by Rule 18(a).

To illustrate with respect to Halliburton, five of the six alleged corrective disclosures analyzed here involved allegations related to Halliburton's asbestos liabilities. The sixth alleged corrective disclosure date (December 21, 2000) involved Halliburton's statements regarding merger-related and other issues.<sup>176</sup> Assuming that the asbestos-related fraud allegations are sufficiently separate from the merger and other allegations that judgment in one set of claims would not preclude the other, the District Court should have treated the December 21, 2000, date separately from the other five alleged corrective disclosure dates. This means that there would be no necessary correction for multiple comparisons for December 21, 2000; statistical significance testing for that date would follow the usual practice. For the other five dates, the relevant number of tests would be 5, rather than 6 as used by the District Court.<sup>177</sup>

It can be shown that this change would not affect any of the statistical significance conclusions in our *Halliburton* event study. However, the change would have made a difference in other circumstances. For example, had the usual  $p$ -value for December 21, 2000, been below 0.05, it would again be considered statistically significant using our approach to grouping alleged corrective disclosure dates.<sup>178</sup> This example shows the importance of a court's approach to determining the number of relevant dates for purposes of adjusting for multiple testing.<sup>179</sup>

---

<sup>176</sup> Allen Report, note 128, *supra*, at 7-8, ¶ 11.

<sup>177</sup> It is true that this rule would require the District Court to engage in a (provisional) claim preclusion analysis that would otherwise be unnecessary. However, such analysis will usually not be all that cumbersome, and it provides a principled basis for determining when a multiple comparisons adjustment is appropriate.

<sup>178</sup> Recall from Table 5 (at page 49, *supra*), that the usual  $p$ -value for this date is 0.2222, whereas the  $p$ -value after correcting for multiple testing in the way the District Court endorsed was 0.6340. Suppose the usual  $p$ -value had been 0.04. Then the District Court-endorsed approach—treating December 21, 2000, as part of the same group as the other five dates for multiple testing purposes—would have yielded a Holm-Šidák  $p$ -value of 0.0784. Thus the District Court's approach would not find statistical significance, whereas our preclusion-based approach would.

<sup>179</sup> Still another issue that arises here involves the problem that would arise if a plaintiff's expert tested some dates but then excluded consideration of them from her

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

There is another side to the multiple comparison adjustment. 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. At class certification, the plaintiff need not establish loss causation, so only price impact on Date 1 would be at issue. However, both dates are relevant for merits purposes, since the plaintiff will have to prove both that the alleged misrepresentation caused the stock price to rise and that the alleged corrective disclosure caused the price to drop.

When the plaintiff is required to show price impact for *both* Date 1 *and* Date 2, the situation differs from the one considered above, where it was sufficient for the plaintiff to show price impact as to *any* of multiple dates. This case is the polar opposite of that presented in Halliburton and requires a different statistical adjustment. In the case in which two events must both be shown to have statistical significance, the statistical threshold for finding price impact must be adjusted to be less demanding than if only a single date is being analyzed.

To see why, consider what would happen if we used a traditional one-sided test for each date separately, separately demanding a 5% Type I error rate for each. For each day, we have seen that the probability of finding statistical significance when there was no actual price impact is 1 in 20. Because these significance tests are roughly independent,<sup>180</sup> the

---

expert report, in order to hold down the magnitude of the multiple testing correction. Halliburton suggested that the plaintiffs had done just that, *see* 309 F.R.D. at 266, and argued that *all* dates on which news similar to the plaintiff's alleged corrective disclosures should be considered for purposes of determining the magnitude of the multiple testing correction. The judge rejected these allegations of unscrupulous behavior. *Id.*

<sup>180</sup> There are two potential reasons to question independence of the estimated excess returns. First, suppose Date 1 involves an alleged misrepresentation and Date 2 an alleged corrective disclosure. If the alleged fraud is a real one, then the magnitudes of the excess returns on Dates 1 and 2 will be correlated. However, this fact is irrelevant to Type I error rate considerations in statistical significance testing. Such testing imposes the null hypothesis that there was actually no material fraud, in which case there is no reason to think the excess returns will be correlated. Second, though, the estimated excess returns will have a bit of dependence because they are calculated from the same estimated market model, whose estimated coefficients will be common to the two event date excess returns. However, this dependence can be shown to vanish as the number of

probability that *both* tests will reject when each null hypothesis is true is only 1 in 400, i.e., one-quarter of one percent.<sup>181</sup> To put it differently, requiring each date separately to have a 5% Type I error rate for a finding of statistical significance is equivalent to requiring a Type I error rate of just 0.25% in determining whether the plaintiff has met its merits burden as to the alleged misrepresentation in question. This is obviously a *much* more demanding standard than the 5% Type I error rate that courts and experts say they are using.<sup>182</sup>

To make the adjustment, 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 say they are applying—we should determine that price impact is significant on both days if each date has a usual *p*-value of less than 0.2236.<sup>183</sup> Using the one-sided SQ approach, this means that

- the estimated price impact for the alleged corrective disclosure date is statistically significant if the estimated excess return for that date is less than estimated excess returns for fewer than 22.4% of the dates in the estimation period; and
- the estimated price impact for the alleged misrepresentation date is statistically significant if the estimated excess return for the alleged corrective disclosure date is greater than estimated excess returns for fewer than 22.4% of the dates in the estimation period.

The resulting test has a 5% Type I error rate, i.e., a 5% chance of erroneously making a finding of statistical significance as to both dates considered together.

To illustrate using our *Halliburton* example, think of December 21, 2000, as Date 2, and imagine that the alleged corrective disclosure on that date had been associated not with a confirmatory disclosure but a non-confirmatory alleged misrepresentation on Date 1. In that case, the plaintiff would have to prove both that the stock price rose an unusual

---

dates in the estimation period grows, and with 593 we would expect very little to persist.

<sup>181</sup> This is the case because  $1/20$  times itself is  $1/400$ , which is one-fourth of  $1/100$ —or, equivalently, a quarter of a percent.

<sup>182</sup> In terms of confidence level, the actual standard amounts to 99.75% confidence rather than the claimed 95%.

<sup>183</sup> This is true because the probability of finding that two independent 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 jointly significant if each has a usual *p*-value less than this level.



amount on Date 1 *and* that it fell by an unusual amount following the alleged corrective disclosure on December 21, 2000. Recall that the usual  $p$ -value for the December 21, 2000, estimated excess return was 0.2222.<sup>184</sup> This value just makes the 0.2236 cutoff. If the hypothetical Date 1 estimated excess return had a usual  $p$ -value of 0.2236 or lower, then both arms of our test would be met. In such a case, a court should find that the plaintiff would have carried its burden to show both a material change in price for the alleged misrepresentation and loss causation as to the 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.

We know of no case where our argument has even been made, but the argument is grounded in the same statistical analysis applied by the court in *Halliburton*. Concededly a court could take the view that for any single piece of statistical evidence to be credited, that single piece must meet the 5% Type I error rate—even if that means that a party who must show 2 pieces of evidence is actually held to the radically more demanding standard of a 0.25% Type I error rate.<sup>185</sup> We believe that such a view is indefensible on probability grounds.<sup>186</sup>

#### D. Dynamic Evolution of the Excess Return's Standard Deviation

For a traditional 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 fraud-related 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,

---

<sup>184</sup> See the second column of Table 5 at page 49, *supra*.

<sup>185</sup> We note that our point is especially important for those situations in which there are more than just 2 dates in question. For example, if there were 5 dates, then the true Type I error rate required when a court requires the plaintiff to meet the 5% Type I error rate separately for each date would be less than 0.00003% (which is approximately 1 in 3.2 million—or 1/20 raised to the fifth power).

<sup>186</sup> We recognize that courts might, as a matter of evidence law, adopt an approach that makes no sense in probability terms. However, we think there is no good reason to do so in the present context, and we think that once courts are aware of the issue they should agree.

differed substantially in the two periods, then 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 A, *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. But 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.<sup>187</sup>

Consider an extreme example to illustrate. Suppose that the standard deviation of a stock's excess return is usually 1%, and for simplicity assume that the excess returns always have a normal distribution. An expert who assumes the standard deviation is 1% on an alleged corrective disclosure date therefore will determine that the excess return for that date is statistically significant if it is less than -1.645%.<sup>188</sup> But suppose that on the date of the alleged corrective disclosure, the standard deviation happens to be much greater than usual—e.g., 2%. Then the actual Type I error rate for the expert's test of statistical significance is about 21%—more than four times the desired level.<sup>189</sup> What has happened here is that the increase in the standard deviation on the alleged corrective disclosure date means that the excess return is more

---

<sup>187</sup>See Fox, Fox, and Gilson, *Volatility Spikes*, note 155, *supra*, at 357; Andrew C. Baker, Note, *Single-Firm Event Studies, Securities Fraud, and Financial Crisis: Problems of Inference*, 68 *Stan. L. Rev.* 1207, 1246-55 (2016); Allen Report, note 128, *supra*, at 97-106.

<sup>188</sup> Recall that for a normally distributed random variable, which has mean 0 and standard deviation 1, the probability of taking on a value less than -1.645 is 0.05, i.e., 5%.

<sup>189</sup> It is a fact of probability theory that the probability that a normally distributed random variable with standard deviation  $\sigma$  takes on a value less than -1.645 is the same as the probability that a normally distributed random variable with standard deviation 1 takes on a value less than  $-1.645/\sigma$ . Setting  $\sigma$  equal to 2, the resultant probability is 0.2054, or roughly 21%.

likely to take on values further from the average of zero. Consequently, the excess return on this date is more likely than usual to correspond to a price drop of more than 1.645%. The opposite result would occur if the standard deviation were *lower* on the alleged corrective disclosure date. With a standard deviation of only  $\frac{1}{2}$  on that date, the Type I error rate would fall to 0.05%, which is one one-hundredth of the desired level.<sup>190</sup> Ignoring the alleged corrective disclosure date's difference in standard deviation in this situation would make false negatives (Type II errors) much more common than would a test that uses a correct critical value for the alleged corrective disclosure date excess return.

Changes in volatility are a potentially serious concern in at least some cases. Fox, Fox & Gilson show that the stock market has experienced volatility spikes in connection with every major economic downturn in the past 85 years, including the 2008-2009 financial crisis.<sup>191</sup> As they point out, the effect of a volatility spike is to raise the necessary threshold for demonstrating materiality or price impact with an event study, thereby increasing the Type II error rate of standard event study tests.<sup>192</sup>

Event studies can be adjusted to deal with the problem of dynamic changes in standard deviation. To do so, one must use a model that is capable of estimating the standard deviation of the event-date excess return both for dates used in the estimation period—our “usual” dates from above—and for those dates that are the object of the price impact inquiry. The details of doing so are fairly involved, requiring both a substantial amount of mathematical notation and a discussion of some technical econometric issues. Accordingly, we relegate these details to Appendix C, which appears at the end of this Article, providing only a brief conceptual summary here. We use a statistical model that allows the standard deviation of excess returns to vary on a day-to-day basis—whether due to the evolution of market- or industry-level return volatility or to the evolution of Halliburton's own return volatility. To compute the *p*-value for each event date, we use the model's estimates to re-scale the excess returns for non-event dates so that all these dates have the same standard deviation as the event date in question. We then use the re-

---

<sup>190</sup> Setting  $\sigma$  equal to  $\frac{1}{2}$ , the probability in question is the probability that a normally distributed random variable with standard deviation 1 takes on a value less than -3.29, which is 0.0005, or roughly 0.05%.

<sup>191</sup> Fox, Fox, & Gilson, note 187, *supra*, at 335-336.

<sup>192</sup> *Id.*, at 357 (stating that a volatility spike “can result in a several-fold increase in Type II error—that is, securities fraud claims will fail when they should have succeeded”).

scaled excess returns to conduct one-sided SQ tests with correction for multiple testing, as discussed in the sections above.

We analyze here only the question of firm-specific variance.<sup>193</sup> Using the approach detailed in the appendix, we find that the standard deviation in Halliburton's excess returns does not remain stable but rather evolves over our time period in at least three important ways. First, Halliburton's excess returns have greater standard deviation on days when the industry peer index returns have greater standard deviation. Second, Halliburton's excess returns are more variable on days when a measure of overall stock market volatility suggests this volatility is greater.<sup>194</sup> Third, the standard deviation 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).

Using the model estimates described in Appendix A, we tested for normality of the re-scaled excess returns.<sup>195</sup> We found that the data resoundingly reject the null hypothesis that the white noise term  $u_t$  is

---

<sup>193</sup> That is, we do not focus on market-level changes; *see also* Fox, Fox, and Gilson, *Volatility Spikes*, note 155, *supra*, at 346-61 for an extensive analysis of variations in market volatility.

<sup>194</sup> 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. Its use as a variance forecasting tool has recently been advocated in Baker, *Single-Firm Event Studies*, note 187, *supra*, following such use by in an event study used in securities fraud litigation, *see* Expert Report of Mukesh Bajaj at 45, *In re Fed. Home Loan Mortg. Corp. (Freddie Mac) Sec. Litig.*, 281 F.R.D. 174 (S.D.N.Y. 2012) (No. 1:09-MD-2072 (MGC)), 2011 WL 7473682 (cited in Baker, *Single-Firm Event Studies*, note 187, *supra*, at n. 217 at 1245). 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 the VIX index is much less important in explaining the variance of Halliburton's excess returns than is volatility in the industry peer index.

<sup>195</sup> We used the same method as in section IV.B, *supra*; *see* TAN 163, *supra*.

distributed normally.<sup>196</sup> Accordingly, it is unreliable to base a test for statistical significance on the assumption that  $u_t$  follows a normal distribution.<sup>197</sup> We therefore use the SQ test approach described in the section IV.B, *supra*.

Table 6 reports  $p$ -values from our earlier and new results. The first three columns involve what we have called “usual”  $p$ -values, which are computed as if one date’s statistical significance is being tested at a time. The first column of these three reports the usual  $p$ -values from Table 5, which were computed from statistical significance tests that impose the assumption that the standard deviation of Halliburton’s excess returns is the same on all dates. The second column reports usual  $p$ -values computed from our model that allows the standard deviation to evolve over time. Our third column shows that when we ignore the issue of multiple tests, our conclusions from statistical significance testing are the same whether we account for dynamics in the daily standard deviation or not (three of the dates are found significant using both approaches, and the other three are not).

The last three columns of Table 6 provide  $p$ -value and significance testing results when we take into account the fact that there are 6 alleged corrective disclosure dates.<sup>198</sup> For five of the six dates, the significance conclusion is unaffected by allowing Halliburton’s excess return standard deviation to vary over time. However, for December 4, 2001, the  $p$ -value drops substantially once we account for the possibility of evolving standard deviation: it falls from 0.0787, which is noticeably above the significance threshold of 0.05, to 0.03, which is noticeably below the threshold. Allowing for the evolution of standard deviation thus matters critically in *Halliburton*, provided that we also account for the multiple dates on which alleged corrective disclosures must be assessed statistically.

---

<sup>196</sup> 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.

<sup>197</sup> Baker appears to have done exactly this in his simulation study; see Baker, *Single-Firm Event Studies*, note 187, *supra*, at 1246 (referring to the use of  $t$ -statistics to determine rejection rates).

<sup>198</sup> For details see the discussion of the Holm-Šidák approach described in notes 172 and 173, *supra*.

**Table 6: Controlling for Evolution in the Volatility of Halliburton's Excess Returns**

Event Date	Usual $p$ -Value (no accounting for multiple tests)			Holm-Šidák $p$ -Value (Accounting for Six Tests)		
	Assuming Constant Standard Deviation	Allowing Dynamic Standard Deviation	Statistical Significance	Assuming Constant Standard Deviation	Allowing Dynamic Standard Deviation	Statistical Significance
	December 7, 2001	0	0	Both	0	0
August 9, 2001	0.0017	0.002	Both	0.0034	.003	Both
December 4, 2001	0.0269	0.010	Both	0.0787	.030	<b>Dynamic Only</b>
December 21, 2000	0.2222	0.256	Neither	0.6340	.694	Neither
October 30, 2001	0.2609	0.317	Neither	0.7795	.881	Neither
June 28, 2001	0.3013	0.298	Neither	0.8837	.851	Neither

What drives this important reversal for the December 4, 2001, alleged corrective disclosure? For that date, our model yields an estimated standard deviation of 1.5%. This is lower than the value of 1.745% in the constant-variance model underlying Table 5, and that is part of the story. But there is more to it. When we assumed constant variance across dates, there were 16 estimation period dates that had a more negative estimated excess return than the one for December 4, 2001. Once we allowed for the standard deviation to evolve over time, *all but one of these 16 dates* had an estimated standard deviation greater than 1.5%. 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.<sup>199</sup>

<sup>199</sup> 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  $u_t = \frac{\varepsilon_t}{\sigma_t}$ , we have

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

In sum, the standard deviation on December 4, 2001, was a bit on the low side, while dates in the left tail of the excess returns distribution had very high standard deviations. When we multiply by the scale factor to make all other dates comparable to December 4, 2001, the re-scaled excess returns for left-tail dates move toward the middle of the distribution. This result indicates that the December 4, 2001, excess return, is actually more unusual than it appears when we fail to account for dynamic evolution in the standard deviation. Once we correct that failure, we find that the alleged corrective disclosure date of December 4, 2001, is sufficiently unusual as to be statistically significant.

## V. EVIDENTIARY CHALLENGES TO THE USE OF EVENT STUDIES IN SECURITIES LITIGATION

The foregoing Parts have explained the role and methodology of event studies as well as identifying several adjustments required to make the event study methodology reliable for addressing issues of price impact, materiality, loss causation and damages in securities fraud litigation. We turn, in this Part, to the limitations of event studies – what they can and cannot prove. Although event studies have become popular because of the apparent scientific rigor that they bring to analysis of the relationship between disclosures and stock price movements, the question that they answer is not identical to the underlying legal questions for which they are offered as evidence. In addition, characteristics of real world disclosures may limit the ability of an event study to determine the relationship between a specific disclosure and stock price. Finally, the use of the event study methodology incorporates a judgment about the relative importance of Type I and Type II errors that masks an underlying policy judgment about the social value of securities fraud litigation. These concerns have not received sufficient attention by the courts that are using event studies to decide securities cases.

---

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.

### A. *The Significance of Insignificance*

An event study answers a very specific question: was the stock price movement on the event date highly unusual? More precisely, the study asks whether it would have been extremely unlikely to observe the excess return on the event date in the absence of some unusual firm-specific event. In the case of a securities fraud event study, the firm-specific event is a fraudulent statement or a corrective disclosure.

Importantly, event study evidence of a highly unusual excess return rebuts the null hypothesis of no price effect. But failure to rebut the null hypothesis does not necessarily mean there was no price impact of an alleged misrepresentation. An event date's excess returns might be in the direction consistent with the plaintiff's allegations but be too small to be statistically significant. Thus, a conclusion of statistical insignificance does not prove the null hypothesis; it simply fails to reject it. That is, the standard event study does not show that the information did not affect stock price; it simply shows that the information did not have a statistically significant effect.<sup>200</sup>

This limitation raises two concerns. One is the burden of proof. As Merritt Fox has explained, an open issue following the Supreme Court's decision in *Halliburton II* concerns the appropriate burden of proof for a defendant seeking to rebut a plaintiff's showing of price impact at the class certification stage.<sup>201</sup> On the other hand, this distinction may be largely cosmetic. To the extent that the plaintiff will have the burden of proof at summary judgment or at trial to establish materiality, reliance and causation, the plaintiff will presumably need to offer an event study that demonstrates a highly unusual price effect at that time. Although there may be practical or policy reasons for holding the plaintiff to a lesser burden at class certification, any such relief for the plaintiff will be short-lived.<sup>202</sup>

This in turn introduces the second concern. To what extent should courts consider additional evidence of price impact in a case in which even a well-constructed event study is unlikely, or unable, to reject the null

---

<sup>200</sup> See Brav & Heaton, *Event Studies*, note 8, *supra*, at 587 (“Courts err because of their mistaken premise that statistical insignificance indicates the probable absence of a price impact.”).

<sup>201</sup> See Fox, *It All Depends*, note 80, *supra*.

<sup>202</sup> We note that deferring the dismissal of a case to, say, summary judgment would create some settlement value, since both the prospect of summary judgment and the battle over class certification involve litigation costs. We leave for another day a full discussion of the importance of these costs in the long-running debate over the empirical importance of procedure in generating the filing of low-merit cases.



hypothesis? We consider this question in more detail in sections B and C below.

*B. Dealing with Multiple Pieces of News on an Event Date*

There are at least two ways in which the question answered by an event study differs from the legally relevant question. First, event studies cannot determine whether the event in question *caused* the highly unusual excess return.<sup>203</sup> 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 the customer was terminating activity in one of the regions where it uses Halliburton's services. The customer's statement, rather than Halliburton's corrective disclosure, might be the cause of a drop in stock price.

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 customer's announcement would fully or partially offset the effect of the corrective disclosure, at least within the limits of the power that appropriate statistical tests can provide. In that case, there will be no highly unusual change in Halliburton'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 event occurs at the same time as the legally relevant alleged event. We might term this additional event a confounding event. If multiple unusual events—events that would affect the stock price even aside from any industry-wide or idiosyncratic developments—occurred on the event date, then even an event study that controls for market- or industry-level factors will 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 measure is the net market response to these two

---

<sup>203</sup> Even if the event study were capable of identifying causality, it would not be able to specifically determine the reasons for a stock price reaction. Thus, as noted above, the correct response to Justice Alito's question at oral argument in *Halliburton II* is that, by themselves, event studies are incapable of distinguishing between a rational and irrational response to information.

developments. Without further refinement, it would not distinguish the sources of this response.

The event study methodology might be refined to deal with some possible confounding events. 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 events.<sup>204</sup> 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.<sup>205</sup> As a result, a study that looks at price movements during the day may be able to separate out the effect of disclosures that took place at different times.

When multiple sources of news are released at exactly the same time, however, no event study can, by itself, separate out the effects of the different news. The event study can only tell us whether the net effect of all the news was associated with an unusually large price drop or rise.

The results of the event study could still 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

---

<sup>204</sup> See discussion in Brav & Heaton, *Event Studies*, note 8, *supra*, at 607 (discussing intraday event studies and citing a case, *In re Novatel Wireless Sec. Litig.*, 910 F. Supp. 2d at 1218-21, in which the court held that an expert's testimony as to such a study was admissible).

<sup>205</sup> 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, note 128, *supra*, n. 93 at 43. We note that if two pieces of news are released very close in time to each other, that might raise special challenges related to the limited amount of trading typically seen in a short enough window; this issue is beyond the scope of the present Article.

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. Alternatively, experts might use quantitative content analysis, e.g., measuring the relative frequencies of two types of news in headlines of articles published following the news.<sup>206</sup> While the release of multiple pieces of news on the same date complicates the use of event studies to measure price impact, event studies might be useful in at least some of those cases. On the other hand, as this discussion suggests, an event study is likely to be incapable of definitively resolving the question of price impact, and a court considering a case involving confounding disclosures will have to determine the role of other evidence in addressing the question.

Lurking in the shadows of this discussion is the question of *why* information events might occur at the same time in a way that would complicate the use of an event study. Although the presence of confounding events could result from random chance, it could also be that an executive shrewdly decides to release multiple pieces of information simultaneously. Specifically, the overwhelming judicial reliance on event studies might create an incentive for issuers and corporate officials to bundle corrective disclosures with other information in a single press release or filing. If the presence of overlapping news makes it difficult or impossible for plaintiffs to marshal admissible and useful event study evidence, defendants may strategically structure their disclosures to impede plaintiffs' ability to establish price effect. The possibility of such strategic behavior raises important questions about the admissibility of non-event study evidence.

### C. Power and Type II Error Rates in Event Studies Used in Securities Fraud Litigation

The focus of courts and experts in evaluating event studies has been on whether an event study establishes a statistically significant price impact. As we discussed briefly in regard to Table 1 in Part II.B, *supra*, the 95% standard for statistical significance is a requirement that the Type I error rate be less than 5%. But Type I errors are only one of two ways an event study can lead to erroneous inference. An event study

---

<sup>206</sup> See, e.g., David Tabak, *Making Assessments about Materiality Less Subjective through the Use of Content Analysis*, at 13 (Mar. 13, 2007) (unpublished manuscript), available at [http://www.nera.com/extImage/PUB\\_Tabak\\_Content\\_Analysis\\_SEC1646-FINAL.pdf](http://www.nera.com/extImage/PUB_Tabak_Content_Analysis_SEC1646-FINAL.pdf).

leads to a Type II (false negative) error when it fails to reject a null hypothesis that really is false—i.e., when it fails to detect something unusual that really did happen on a date of interest.

As we discussed in section II.B, *supra*, for a given statistical test there is a tradeoff between Type I and Type II error rates: choosing to tolerate fewer false positives necessarily creates more false negatives. Thus, by insisting on a 5% Type I error rate, courts are implicitly insisting on both a 5% rate of false positives and some particular rate of false negatives. Recent work has pointed out that in single firm event studies used in securities litigation, requiring a Type I error rate of only 5% yields an extremely high Type II error rate.<sup>207</sup>

To illustrate, suppose that a corrective disclosure by an issuer actually causes a price drop of 2%. We assume for simplicity that the issuer's excess returns are normally distributed with standard deviation 2%.<sup>208</sup> A properly executed event study will reject the null hypothesis of no effect on that date if and only if the estimated excess return is less than -1.645%. The probability that this will occur—also known as the power of the test—is 57%.<sup>209</sup> This means that the Type II error rate is 43%.<sup>210</sup> In other words, 43% of the time, the event study will fail to find a statistically significant price impact. Notably, this error rate is many times greater than the 5% Type I error rate.

As this example illustrates, the Type II error rate that results from insisting on a Type I error rate of 5% can be quite high. We see no reason to believe that this trade-off is appropriate from a normative perspective.<sup>211</sup> The use of the 5% Type I error rate is what is

---

<sup>207</sup> Brav & Heaton, *Event Studies*, note 8, *supra*, at 597 (discussing fact that Type II error rate is 73.4% for a stock with normally distributed excess returns having standard deviation 1.5%, when true event-related price impact is a drop of 2%).

<sup>208</sup> This magnitude for the standard deviation was not atypical in 2014, for example; see Brav & Heaton, *Event Studies*, note 8, *supra*, at 13 (Table I) (showing that the average value of the standard deviation of excess returns was 2.0% among firms whose standard deviations put them in the sixth decile of 4,298 firms studied for 2014).

<sup>209</sup> Because the excess return is assumed normally distributed with standard deviation 2%, the scaled random variable that equals one-half the excess return will have a normal distribution with mean 0 and standard deviation 1%. Since the corrective disclosure causes a 2% drop, the event study described in the text will yield a finding of statistical significance whenever -1 plus this scaled random variable is less than the ratio (-1.645/2). The probability of that event—the test's power in this case—can be shown to equal 0.5704.

<sup>210</sup> Since the probability of a Type II error is 1 minus the power of the test, the probability of a Type II error is 0.4296, which implies a Type II error rate of 43%.

<sup>211</sup> Fox, Fox, and Gilson, *Volatility Spikes*, note 155, *supra*, at 368-372, do take this view.

traditionally used in the academic literature on financial economics,<sup>212</sup> but there are numerous differences between those academic event studies and the ones used in securities litigation. As we have already seen, the one-sided/two-sided distinction is one such difference, as is the frequent role of multiple event dates.

In addition, most academic event studies involve the averaging of event date excess returns over multiple firms. This averaging often will both (i) greatly reduce the standard deviation of the statistic that is used to test for statistical significance,<sup>213</sup> and (ii) greatly reduce the importance of non-normality.<sup>214</sup> Thus the event studies typically of interest to scholars in their academic work are atypical of event studies that are used in securities litigation. Whatever the merits of the convention of insisting on a Type I error rate of 5% in academic event studies, we think the use of that rate in securities litigation is the result of happenstance and inertia rather than careful weighing of the costs and benefits of the tradeoff in Type I and Type II errors.

This observation suggests that the current approach to using event studies in securities litigation warrants scrutiny. As long as courts continue to insist on a Type I error rate of 5%, Type II error rates in securities litigation will be very high. This means that event study evidence of a significant price impact is much more convincing than event study evidence that fails to find a significant price impact. To put it in evidence law terms, at the current 5% Type I error rate, a finding of significant price impact may be considerably more probative than a failure to find significant price impact.

That raises two questions. First, what Type I error rate *should* courts insist on, and how should they determine that rate? Second, if event study evidence against a significant price impact has limited probative value, does that change the way courts should approach *other* evidence that is usually thought to have limited probative value? For example, one approach might be to allow financial industry professionals to be qualified as experts for purposes of testifying that an alleged corrective disclosure could be expected to cause price impact, both for the class certification purposes on which we have focused and as to other merits questions. The logic of this idea is simple: when event study evidence fails to find a significant price impact, that evidence has limited probative value, so the value of general, non-statistical expert opinion

---

<sup>212</sup> See Brav & Heaton, *Event Studies*, note 8, *supra*, at 614 (citing *United States v. Hatfield*, 795 F.Supp.2d 219, 234 (E.D.N.Y. 2011) (questioning this practice)).

<sup>213</sup> See Brav & Heaton, *Event Studies*, note 8, *supra*, at 605.

<sup>214</sup> See GHK, note 160, *supra*, at 510.

will be comparatively greater in such cases than in those cases in which event study evidence does find a significant price impact.<sup>215</sup> These are complex questions that go to the core of the appropriate role of event studies in securities fraud litigation.<sup>216</sup>

## CONCLUSION

Event studies play an important role in securities fraud litigation. In the wake of *Halliburton II*, that role will 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 explained the event study methodology and explored a variety of considerations related to the use of event studies in securities fraud litigation, highlighting the ways in which the litigation context differs from the empirical context of many academic event studies.

A key lesson from this Article is that courts and experts should pay more attention to methodological issues. We identify four methodological considerations and demonstrate how they can be addressed. 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. Fourth, event studies must be modified appropriately to account for the possibility that stock price volatility varies across time.

Even with these adjustments, event studies have their limits. We discuss some evidentiary challenges that confront the use of event studies in securities litigation. First, failing to reject the null hypothesis is not the same as proving that information did not have a price effect. As

---

<sup>215</sup> Further, such an approach would reduce the incentive for managers to strategically release bad news in ways that would defeat the usefulness of event studies (*see* section 0, *supra*), since doing so opens the door to more subjective expert testimony that is likely to be easy for plaintiffs to obtain.

<sup>216</sup> A full discussion of the normative implications of the 5% Type I error rate is beyond the scope of this Article. Two of us consider this question further in Jill Fisch & Jonah Gelbach, *The Policy Tradeoff in Statistical Significance: What it is, and How to Deal with It* (tba).

a result, the legal impact of an event study may depend critically on which party bears the burden of proof and the extent to which courts permit the introduction of non-event study evidence on price impact. Second, both accidental and strategic bundling of news may make event study evidence more difficult to muster. Third, event studies used in securities litigation are likely to be plagued by very high ratios of false negatives to false positives—that is, they are much more likely to yield a lack of significant evidence of an actual price impact than they are to yield significant evidence of price impact when there really was none. We suggest that this imbalance of Type II and Type I error rates warrants further analysis.

**APPENDIX A: IMPLEMENTING THE SQ TEST**

For convenience and comparison's sake, Table 7 summarizes the steps necessary to implement the standard approach assuming normality 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 uses the data to guide the expert regarding what critical value to use.

Finally, the SQ test adds no assumptions that are not already relied on by the standard approach.<sup>217</sup> 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. Thus we emphasize that there are circumstances in which the one-sided SQ test will perform systematically better than the standard approach, and there are no circumstances under which it will systematically perform worse. 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, and (iii) the SQ test adds virtually no additional work for an expert or court to understand. There is no good reason to use the standard approach, but not to use the SQ test. And, leaving aside the other issues we are about to discuss, there is good reason *to* use the SQ test.

---

<sup>217</sup> See note 165, *supra*.



**Table 7: Comparing the Standard Approach and the SQ Test Approach for Testing the Null Hypothesis of No Event Date Effect on an Alleged Corrective Disclosure Date**

<b>Step</b>	<b>Standard Approach</b>	<b><i>SQ Test Approach</i></b>
<b>1. Estimate market model</b>	Use ordinary 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.)	<i>Same as standard approach.</i>
<b>2. Calculate estimated excess return for event date</b>	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.	<i>Same as standard approach.</i>
<b>3. Calculate critical value</b>	Critical value is product of -1.96 (if two-sided critical value used) or -1.645 (if one-sided critical value used) and the standard error of regression from market model.	<i>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<sup>th</sup> percentile of these estimated excess returns. This number is the critical value.</i>
<b>4. Determine whether to reject null hypothesis of no event date effect</b>	Reject null hypothesis if event-date estimated excess return from step 2 is more negative than critical value from step 3.  Otherwise do not reject null hypothesis.	<i>Same as standard approach (except that critical value is found as described in cell just above).</i>

**APPENDIX B: CORRECTING FOR MULTIPLE TESTING WHEN THERE ARE MULTIPLE PAIRS OF ALLEGED MISREPRESENTATION AND ALLEGED CORRECTIVE DISCLOSURE DATES**

Here we briefly consider a situation like the following, in which we are concerned with proper statistical significance testing as to the merits. Plaintiff alleges that:

- on Date 1A, defendant made a non-confirmatory fraudulent statement, and then there was a corrective disclosure on Date 1B;
- on Date 2A, defendant made another non-confirmatory fraudulent statement, and then there were corrective disclosures on Dates 2B, 2C, 2D, and 2E;
- on Date 3A, defendant made a confirmatory fraudulent statement, and there was a corrective disclosure on Date 3B.

Suppose for the moment that all three alleged misrepresentations are within a common scope of claim preclusion, so that under our analysis in section IV.C.2 above, statistical significance methodology should account for the presence of all dates related to the three alleged misrepresentations. The plaintiff will win as to alleged misrepresentation 1 if she carries her burden as to Date Set 1 (Date 1A and Date 1B); or as to Date Set 2 (Date 2A and any one of Dates 2B, 2C, 2D, or 2E); or as to Date set 3 (Date 3B). We assume that the court wants to ensure that there is a 5% chance of erroneously finding for the plaintiff as to any of the three alleged misrepresentations.

Following the logic of section IV.C.3, the court should require a Type I error rate of 1.7% with respect to each Date Set considered separately, because if the court does that, there will be a 5% chance that at least one of the Date Sets is erroneously found to be statistically significant.<sup>218</sup> Date Set 1 involves two dates, both of which must meet the relevant statistical threshold to for Date Set 1 to be found significant. With a Type I error rate of 1.7% for Date Set 1 overall, the appropriate Type I error rate to apply separately to Dates 1A and 1B is 13%.<sup>219</sup>

---

<sup>218</sup> To be precise, the Type I error rate is 1.69524275%. To see this, observe that the probability that at least one of the 3 Date Sets will erroneously be found statistically significant is 1 minus the probability that none of them will. Using the above Type I error rate, the probability equals  $1 - (1 - 0.0169524275)^3$ , which equals 0.05, or 5%.

<sup>219</sup> Since the probability that both Date 1A and Date 1B will have an estimated excess return with a  $p$ -value of  $q$  is  $q^2$ , and since we want  $q$  to equal 0.017, we have  $q = \sqrt{0.017}$ , which is approximately 0.13, or 13%.

Date Set 2 involves five dates. For the plaintiff to win as to this date set, she must show that the estimated excess return meets the court's threshold as to Date 2A and any one of Dates 2B, 2C, 2D, or 2E. The court can achieve its desired overall Type I error rate for Date Set 2 by requiring a 13% Type I error rate for Date 2A considered separately, together with a 13% Type I error rate for any one of Dates 2B, 2C, 2D, or 2E. As with Date Set 1, this approach will achieve an overall Type I error rate of 1.7% for Date Set 2.<sup>220</sup> The 13% Type I error rate for any one of Dates 2B, 2C, 2D, or 2E can be achieved by requiring a Type I error rate of approximately 3.4% for each of these four dates considered separately.<sup>221</sup>

Finally, Date Set 3 involves only one date, Date 3B (since the alleged misrepresentation on Date 3A is confirmatory). Thus the appropriately  $p$ -value for Date 3B considered separately is 1.7%.

We summarize these findings for this example in Table 8. The key take-away point from this discussion is that taking appropriate account for multiple testing in securities litigation is fact-bound, depending importantly on how the various alleged events interact with each other according to the plaintiff's allegations.<sup>222</sup>

---

<sup>220</sup> The math is the same as in 219, *supra*; the only difference is that the 13% Type I error rate for Date 1B is replaced by a Type I error rate of 13% for finding any of Dates 2B, 2C, 2D, or 2E significant.

<sup>221</sup> This is so because the probability that at least one of the 4 Dates Date 2A, 2B, 2C, or 2D will erroneously be found statistically significant is 1 minus the probability that none of them will. Using the above Type I error rate, the probability equals  $1 - (1 - 0.03421642936)^4$ , which equals roughly 0.034, or 3.4%.

<sup>222</sup> Careful readers will have noted that the discussion in this section has not made use of Holm's sequential approach to determining which Date Sets, and which of the Dates within them, should be treated as statistically significant. The discussion in this section is thus the simple, non-sequential Šidák approach. Using Holm's sequential approach would involve greater power and should be done in practice; we have not used it in this section since it takes more notation to explain and illustrate.

**Table 8: Separately Considered  $p$ -Values that Yield a 5% Type I Error Rate in Finding for the Plaintiff as to Any Alleged Misrepresentation Date of the Three Date Sets Considered in Our Example**

Type I Error Rate for Each Date Considered Separately	
<i>Date Set 1</i>	
<b>Date 1A</b>	<b>13%</b>
<b>Date 1B</b>	<b>13%</b>
(Both dates' $p$ -values must be less than 0.13 for plaintiff to win as to alleged misrepresentation on Date 1A.)	
<i>Date Set 2</i>	
<b>Date 2A</b>	<b>13%</b>
<b>Date 2B</b>	<b>3.4%</b>
<b>Date 2C</b>	<b>3.4%</b>
<b>Date 2D</b>	<b>3.4%</b>
<b>Date 2E</b>	<b>3.4%</b>
(Date 2A $p$ -value must be less than 0.13 and $p$ -value must be less than 0.034 for at least one of Dates 2B-2E for plaintiff to win as to alleged misrepresentation on Date 2A.)	
<i>Date Set 3</i>	
Date 3A	Confirmatory alleged misrepresentation, so no relevant data
<b>Date 3B</b>	<b>1.7%</b>
(Date 3B $p$ -value must be less than 0.017 for plaintiff to win as to alleged misrepresentation on Date 3A.)	

### APPENDIX C: ALLOWING FOR HETEROGENEITY IN THE DISTRIBUTION OF EXCESS RETURNS

In this appendix we drop 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.<sup>223</sup> 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”).<sup>224</sup> The GARCH model class allows the variance of a security’s return on a given date 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

In the absence of any special news on date  $t$ , our basic model may be written as

$$H_t = \alpha + M_t' \beta + \varepsilon_t,$$

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

---

<sup>223</sup> This index is the VIX index published by the Chicago Board Options Exchange. This index, sometimes known as the “fear index”, may be associated with the volatility of firms’ daily returns, which can be accounted for by modeling individual stock variance to take account of movements in the VIX. See Baker, note 187, *supra*, at 1245. We include this index here to account for such concerns.

<sup>224</sup> The GARCH model was first developed in detail in Timothy Bollerslev, *Generalized Autoregressive Conditional Heteroskedasticity*, 31 J. Econometrics 307 (1986), building on seminal work by Robert Engle, *Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of UK Inflation*, 50 *Econometrica* 987 (1982). For an excellent and generally accessible discussion, see Chapter 12 of John Y. Campbell, Andrew W. Lo, and A. Craig MacKinlay, *THE ECONOMETRICS OF FINANCIAL MARKETS* (1997). For a use of GARCH in the law and finance literature, see Mark I. Weinstein, *Don’t Buy Shares without It: Limited Liability Comes to American Express*, 37 J. Leg. Stud. 189 (2008).

Halliburton's return; and  $\varepsilon_t$  is the excess return for date  $t$  given that no special news occurred.

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

$$\varepsilon_t = \sigma_t u_t,$$

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  $E[\varepsilon_t^2 | I_t] = E[(\sigma_t u_t)^2 | I_t]$ ,<sup>225</sup> which can be shown to equal  $E[\sigma_t^2 | I_t]$ .<sup>226</sup> The approach taken until Part IV.D 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.,  $u_t \equiv \varepsilon_t / \sigma_t$ .

We allow for the possibility that  $\sigma_t$  varies over time through three channels. 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.<sup>227</sup> Second, we allow for the possibility that  $\sigma_t^2$  varies with the value of the VIX

<sup>225</sup> The variance of  $\varepsilon_t$  equals the expected value of its square since  $\varepsilon_t$  has mean zero.

<sup>226</sup> 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 variance 1.

<sup>227</sup> 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 take an ex post view by including these variables: 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.

volatility index.<sup>228</sup> And third, we allow for the possibility that  $\sigma_t^2$  varies with the GARCH terms described above, namely the one-period lags  $\sigma_{t-1}^2$  and  $\varepsilon_{t-1}^2$ . Our model for  $\sigma_t^2$  is:

$$\sigma_t^2 = \sigma_0^2 + V_t' \theta_1 + \sigma_{t-1}^2 \theta_2 + \varepsilon_{t-1}^2 \theta_3,$$

where  $V_t$  is a column vector of the squared returns for the industry peer index, our version of the S&P 500 energy sector index, our version of the energy and construction index, and the VIX 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  $\sigma_0^2$ . Finally, we note that since the standard deviation is the square of the variance, the above model for the daily variance  $\sigma_t^2$  implicitly is also a model for the daily standard deviation  $\sigma$ .

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  $\varepsilon_t$ .<sup>229</sup> 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.<sup>230</sup> We estimate the model under this assumption and then return to the question of non-normality below.

### B. Estimation Results

Here we report the results from estimating the model described in section A of this appendix. Table 9 reports 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

---

<sup>228</sup> See footnote 223, *supra* concerning the VIX index.

<sup>229</sup> We can write the market model as

$$H_t = \alpha + M_t \beta + \varepsilon_t,$$

where  $M_t$  is a vector of the industry index variables and  $\alpha$  and  $\beta$  are parameters.

<sup>230</sup> 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  $V_{2t}$  from our equations in the text.

standard errors are very similar to those in Table 2, as should be expected given the robustness result described in footnote 235, *infra*.

**Table 9: Estimates from Market Model for Excess Return, from Model that Allows Heterogeneity in Conditional Variance**

	Market Model for Excess Return	
	Coefficient	Standard error*
<i><math>\beta</math> parameters</i>		
Industry Peers	0.941	0.034
S&P 500 Energy	0.162	0.055
Engineering & Construction	0.026	0.035
<i><math>\alpha</math> parameters</i>		
Intercept	-0.001	0.001
* Estimated standard errors computed using robust covariance estimator to account for possibility of non-normality in standardized excess returns, $u_t$ .		

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, coefficients are highly statistically significant for both the VIX index and the industry peer index squared daily returns, which tells us that there is market-associated 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 the volatility in the stock market overall that VIX measures.<sup>231</sup> The coefficients on the GARCH parameters show that there is a statistically significant coefficient on the autoregressive term (i.e., the lagged variance term). The conditional heteroscedasticity coefficient (i.e., the coefficient on the square of lagged excess returns) is twice the

<sup>231</sup> 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. Thus, the effect of a one-unit movement in the VIX variable included in our model is comparable to the effect of a one-unit movement in the square of each of our industry indexes.



size of this coefficient, though it is statistically significant only at the 10% level.

**Table 10: Estimates from Market Model for Conditional Variance**

	Coefficient	Standard error*
<i>Square of daily return for:</i>		
Industry Peers**	0.330	0.083
S&P 500 Energy**	0.068	0.080
Engineering & Construction**	0.080	0.061
<i>Level of:</i>		
VIX**	0.044	0.017
<i>GARCH lag parameters for:</i>		
Autoregressive term (lagged variance)	0.128	0.062
Conditional heteroskedasticity term (square of lag in realized excess return)	0.244	0.147
<i>Intercept:</i>		
$\sigma_0^2$	-10.024	0.520
* Estimated standard errors computed using robust covariance estimator to account for possibility of non-normality in standardized excess returns, $u_t$ .		
** Variables are scaled so that all these coefficients represent the effect of an increase of one-unit of standard deviation of each regressor.		

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 daily variance  $\sigma_t^2$ , we used

$$\hat{\sigma}_t^2 = \hat{\sigma}_0^2 + V_t' \hat{\theta}_1 + \hat{\sigma}_{t-1}^2 \hat{\theta}_2 + \hat{\varepsilon}_{t-1}^2 \hat{\theta}_3,$$

where hats again denote estimates.<sup>232</sup> (The estimated conditional daily standard deviation  $\hat{\sigma}_t$  may be calculated by taking the square root of the estimated conditional daily variance.)

### 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. We first construct the estimated scale factor discussed in section IV.D of the main text. To do so for some date  $t$ , we first calculate the conditional daily variance for the event date,  $\hat{\sigma}_*^2$  and the conditional daily variance for date  $t$ ,  $\hat{\sigma}_t^2$ . The estimated scale factor is the square-root of the ratio of these estimates:

$$\widehat{SF}_t \equiv \sqrt{\frac{\hat{\sigma}_*^2}{\hat{\sigma}_t^2}}.$$

We then create the re-scaled estimated excess return as

$$\hat{r}_t \equiv \widehat{SF}_t \times \hat{\varepsilon}_t,$$

where  $\hat{\varepsilon}_t$  is the estimated excess return for date  $t$ . Under the null hypothesis that there was no special price impact on the event date, the re-scaled estimated excess return  $\hat{r}_t$  has the same limiting distribution as the event date estimated excess return. Consequently, we may use the collection of  $\hat{r}_t$  estimates for testing statistical significance of the event

---

<sup>232</sup> 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, *Autoregressive conditional heteroskedasticity (ARCH) family of estimators*, <http://www.stata.com/manuals13/tsarch.pdf> at 12-13 ("Priming" tab, "arch0" option defaults). This is the method suggested in Timothy Bollerslev, *Generalized Autoregressive Conditional Heteroskedasticity*, 31 J. Econometrics 307 (1986); see also Matteo M. Pelagatti and Francesco Lisis, *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 <http://www2.mate.polimi.it/ocs/viewpaper.php?id=127&cf=7>.

date estimated excess return just as we would use the unscaled estimated excess returns in the absence of variance heterogeneity. Which tests are appropriate depend on the distribution of the underlying white noise term  $u_t$ .

It is helpful to first consider what one might do if the white noise term  $u_t$  were normal. In this case, the re-scaled excess returns—i.e., the  $r_t$ —are themselves normal; all non-normality in the unconditional distribution of  $\varepsilon_t$  is due to the heterogeneity in the daily standard deviation  $\sigma_t$ .<sup>233</sup> The usual critical values based on the normal distribution are justified when the normality assumption holds.<sup>234</sup> However, this assumption could be wrong, and as in Part IV.B for the homogeneous variance case, a method of testing for statistical significance that ignores actually existing non-normality is unreliable.

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  $\alpha, \beta, \sigma_0^2, \theta_1, \theta_2$ , and  $\theta_3$  will be appropriately estimated if we use the maximum likelihood estimator that would follow from incorrectly assuming normality of  $u_t$ .<sup>235</sup>

For any date  $t$ , consider the estimated white noise term  $\hat{u}_t \equiv \frac{\hat{\varepsilon}_t}{\hat{\sigma}_t}$ . Since  $\hat{\varepsilon}_t$  and  $\hat{\sigma}_t^2$  are appropriate estimates of  $\varepsilon_t$  and  $\sigma_t^2$ , respectively, it follows that  $\hat{u}_t$  is an appropriate estimate of the true white noise terms  $u_t$  defined above.<sup>236</sup> This means that we may use the empirical distribution of estimated white noise terms to estimate features of the distribution of the true white noise term. We then used Stata's **sktest** command to test for

---

<sup>233</sup> See, e.g., discussion in Campbell, Lo, and MacKinlay, note 224, *supra*, 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  $\varepsilon_t$ ).

<sup>234</sup> This is the approach taken by Baker, note 187, *supra*.

<sup>235</sup> 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 Likelihood Estimator that results “when a normal log-likelihood is maximized but the assumption of normality is violated” yields a consistent and asymptotically normal estimator of parameter estimates under broad conditions and specifically analyzing GARCH models).

<sup>236</sup> 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  $u_t$ .

normality of the white noise term. The test rejected normality resoundingly, with a  $p$ -value of 0.025, due to substantial kurtosis.<sup>237</sup>

The “usual”  $p$ -values computed in section IV.B were then computed using the SQ test applied to the event date estimated excess return, using the sample quantiles of the re-scaled estimated excess returns. That is, each event date’s “usual”  $p$ -value equals the share of estimation-period dates that had re-scaled estimated excess returns greater than the event date estimated excess return.

---

<sup>237</sup> We note in addition that the 5<sup>th</sup> sample quantile of the estimated white noise term (-1.742) was somewhat more negative than the 5<sup>th</sup> quantile of the standard normal distribution (-1.645), which is relevant to the comparative performance of the SQ test and the usual test assuming normality of the white noise term. As discussed in section IV.B, the empirical magnitude of such non-normality will vary with context.

## Recent Issues

All CFS Working Papers are available at [www.ifk-cfs.de](http://www.ifk-cfs.de).

No.	Authors	Title
551	Matthias M. M. Buehlmaier and Josef Zechner	<i>Financial Media, Price Discovery, and Merger Arbitrage</i>
550	Paul Schneider, Christian Wagner and Josef Zechner	<i>Low Risk Anomalies?</i>
549	Daniel A. Rettl, Alex Stomper and Josef Zechner	<i>The Stability of Dividends and Wages: Effects of Competitor Inflexibility</i>
548	Youchang Wu, Russ Wermers, Josef Zechner	<i>Managerial Rents vs. Shareholder Value in Delegated Portfolio Management: The Case of Closed-End Funds</i>
547	Thomas Dangl and Josef Zechner	<i>Debt Maturity and the Dynamics of Leverage</i>
546	Thomas Dangl, Otto Randl, and Josef Zechner	<i>Risk Control in Asset Management: Motives and Concepts</i>
545	Otto Randl and Josef Zechner	<i>Sovereign Reputation and Yield Spreads: A Case Study on Retroactive Legislation</i>
544	Alexander Muermann and Thomas Rauter	<i>Prestige and Loan Pricing</i>
543	Daniela Kremslehner and Alexander Muermann	<i>Asymmetric Information in Automobile Insurance: Evidence from Driving Behavior</i>
542	Ester Faia	<i>Sovereign Risk, Bank Funding and Investors' Pessimism</i>