

The Focus On Event Studies In Class Certification

Law360, New York (December 22, 2015, 5:10 PM ET) --



Torben Voetmann



Paul Hinton



Matthew Aharonian

Over the nearly 30 years since *Basic v. Levinson*, it is perhaps surprising how few decisions have focused on the implementation of the event study — an essential element of fraud-on-the-market cases in shareholder class actions. U.S. District Judge Barbara M. Lynn's recent class certification decision in *Halliburton* is an exception.[1] As a consequence of last year's U.S. Supreme Court ruling in *Halliburton* (*Halliburton II*), Judge Lynn's decision explicitly focuses on the analyses of the expert witnesses.[2]

Halliburton II upheld the fraud-on-the-market doctrine, but opened the door for defendants to present price impact evidence at the class certification stage to challenge the presumption of reliance.[3] The Supreme Court's 2014 *Halliburton* ruling has been interpreted in various ways across the legal industry, especially as it relates to the evidentiary demands that the decision raised for demonstrating price impact at the class certification stage. This evidence can also serve to knock out alleged corrective disclosures and potentially reduce the estimated price inflation and the length of the class period. Interestingly, *Halliburton II* did not provide a rigorous economic definition of price impact,[4] although it did identify event studies as a source of evidence of price impact.[5]

While Judge Lynn's decision presents a thoughtful evaluation of the event study evidence presented by experts — and may signal a new focus for future securities class certification hearings — it is important to recognize that many arguments made by plaintiffs in this case were idiosyncratic. Thus, the implications of the class decision for class certification in future cases will depend in large part on the particular circumstances. This article examines four particularly interesting methodological findings in Judge Lynn's class decision from an economist's perspective and describes their potential future influence.

Methodology Rulings in Judge Lynn's Class Decision

Judge Lynn noted that “[t]he Supreme Court did not state expressly in Halliburton II whether plaintiffs or defendants must carry the burden of persuasion to show price impact or lack thereof,” but found that in this case “the burdens of production and persuasion to show lack of price impact are properly placed on Halliburton.”[6] As a consequence, Halliburton’s expert’s event study evidence and critiques of the opposing expert’s analysis became a focus of the class decision, which ruled on issues of econometric method and statistical inference, and provided a singular methodological profile for the case.[7] Of particular note, Judge Lynn ruled on four methodological practices and their application in the context of the specific facts of this case:[8]

- Judge Lynn ruled it was appropriate in this case for the event study model to be estimated using data from the class period — a departure from the convention of excluding class period data when calibrating the model.
- The notion of price impact occurring over a period of more than one day was inconsistent with market efficiency in this case, absent a compelling reason for a delayed price response.
- Adding a specific peer index to the defendant’s expert’s event study model (which benchmarked returns using broader industry indices) was preferred because the resulting model provided a higher level of in-sample explanatory power (i.e., gave the event study model a higher R-squared over the estimation window).
- The initial plaintiff’s expert’s apparent decision to identify relevant disclosure event dates from the data — based on whether a day’s standardized abnormal return was greater than a typical 5 percent critical threshold — required statistical adjustment because this approach posed a multiple comparisons issue.[9]

These practices, if applied generally, would not be inconsequential to the outcome of statistical testing and measurement of price impact. Therefore, it is important for plaintiffs and defendants to determine whether the rationale behind these methodological rulings applies to their specific cases. It remains to be seen how influential the court’s particular methodological choices and the rationale for them in the class decision will be in future cases.

Model Estimation Period

The choice of estimation period for the event study model does not affect the observed change in price on the day of the alleged corrective disclosure, but it does affect the measured abnormal price reaction and statistical tests of the specific price impact.[10] Selection of the model estimation period is a matter of expert judgment. Using a “clean” period (i.e., a period prior to and unaffected by the alleged fraud) to calibrate an event study model has been the more conventional approach in academic and litigation work. However, the class decision is noteworthy because it found that in this case, the class period itself — excluding alleged corrective disclosure days — represented a “clean” period appropriate for estimating the event study model.[11]

Experts consider many factors when choosing a period that represents “normal price movements” for a company. However, if the company’s returns undergo fundamental changes near alleged disclosure

dates — changes that are unrelated to the alleged corrective disclosures — then the choice of an appropriate estimation period needs more consideration. Examples of such changes include general changes in capital markets (e.g., a financial crisis), the company's business strategy, its operations, or its competitive environment. Using the class period data to calibrate the event study model can always be criticized because the days used to represent normal circumstances may be affected by the alleged fraud. Judge Lynn's class decision suggests that using this approach in litigation may be less vulnerable to criticism, as long as there are valid economic reasons for doing so.

Speed of Price Adjustment

Extending an event window from one day to two or three days — as used in many academic studies — has the potential to pick up an extended price impact that could either reinforce or reverse a more immediate initial price impact of an alleged corrective disclosure. However, a longer event window is also more likely to pick up responses to unrelated events. While longer event windows may, in principle, neither favor or disfavor either side, in practice, they may be selected opportunistically. Economic studies cited in the ruling showed, "that when firms publish their latest earnings or announce dividend changes, the major part of the price adjustment occurs within five to 10 minutes of the announcement."^[12] The court interpreted that the disclosure in question was of a similar class of news. Indeed, for disclosures of this type, Judge Lynn's class decision recognizes that a time period of less than a day can be appropriate for certain types of news. This raises the prospect of event study analysis using intradaily returns in future cases, and poses separate methodological issues not discussed here.^{[13],[14],[15]}

Judge Lynn's class decision found that "a compelling explanation" was needed to justify using a two-day window to measure the price impact of the alleged disclosures. In the absence of such an explanation, the court relied on conventional one-day price impact tests.^[16] The court adopted, the one-day event window convention that was used by both experts, but it also looked at intraday price movements to provide more insight into the timing of the market's reaction to news. The intraday price movements did not contradict the finding that there was no price reaction measured using closing prices and a one-day event window. Although an intraday event study analysis was not relied on by the court or the experts, the discussions of intraday analysis in this case suggests that more detailed analysis using high-frequency data, including intraday event studies, may be relied upon in future class certification proceedings.

Comparing Price Response to an Index of Peers

Experts will often include an index of stock price returns for a smaller group of peer companies in their event study model. This is to better control for price impacts that are not specific to the company being analyzed or to the disclosure events being alleged. The plaintiff's expert in this case argued that an index of Halliburton's competitors — labeled the "Analyst Index" — was appropriate to include in the event study model. Judge Lynn's class decision argued that the Analyst Index would permit the price impact studies to better control for Halliburton's specific industry (as discussed above) and "[b]ecause the addition of the Analyst Index constructed by [the plaintiff expert] increases the explanatory power of [the defendant's expert] model, the court is persuaded that it should be utilized in measuring the statistical significance of the price" impact.^[17] The defendant's expert apparently did not voice any material objection to this modification, suggesting agreement as to its appropriateness in this case.

The explicit preference for a model on the basis of increased in-sample explanatory power is particularly noteworthy, though it should not be viewed as an absolute guiding principle for future cases. In the

event study context, while an increase in explanatory power may be viewed as desirable (all else equal) there are important instances where it may not be the right thing to pursue. Consider the case where adding a peer-company index does improve the explanatory power of an event study model prior to the alleged disclosures. However, if the alleged disclosures are such that they have external ramifications for the companies in the peer-company index as well, then adding the index may potentially bias estimates of the price impact attributable to the alleged disclosure. That is, if all peer companies experience a price drop on that day due to the company-specific disclosure, adding the peer-company index to the model will treat a portion of the company drop as if it is an unrelated industry event, when it is not.

The class decision demonstrates an increasing willingness of the courts to make adjustments to experts' models and may ultimately lead to judges explicitly developing their own event studies for guidance. The onus will, therefore, increasingly fall upon counsel and experts to provide appropriate guidance and intuitive reasoning when exceptions to apparent rules-of-thumb do occur.

Testing Price Impact Across Multiple Corrective Disclosures

Judge Lynn's class decision embraced a statistical multiple-comparison adjustment named the Holm-Bonferroni method (the "Holm-Bonferroni adjustment") to collectively examine multiple corrective disclosures. The court's reasoning was that "the substantial number of comparisons" being tested for statistical significance required an adjustment because, without adjustment, testing for more corrective disclosures has a natural tendency of finding significance in any one purely by chance, i.e., an increased rate of false positives (type I error). The Holm-Bonferroni adjustment is one method for controlling the risk that an alleged disclosure date (from a family of many) is found to have caused a price impact, when it did not.

The rationale of the adjustment was explained by defendants' expert as follows:

[I]magine rolling a 20-sided die with 19 white sides and one red side. If the die is rolled once, it would be surprising if the die landed red-side up since the likelihood of this occurring is only 5 percent(1 out of 20). However, if the die is rolled 100 times, it would be much less surprising that the die landed red-side up one time since, on average, the die should land red-side up five times for every 100 rolls.[18]

However, while the Holm-Bonferroni adjustment controls for the risk of a false positive when considering multiple disclosure dates, it does so at the expense of reducing what is called the power of the test (type II error). The court explicitly recognized this by noting that the adjustment increases the risk that "we observe an abnormal return that is not statistically significant and conclude that there was no price impact, when in fact there was one." [19]

As a simple example, consider the case where there are two corrective disclosure dates being tested, and each is significant at a 3 percent level of confidence — stronger than the usual 5 percent cutoff used to test significance. The Holm-Bonferroni adjustment would recalibrate the 5 percent statistical cutoff to 2.5 percent, in recognition of the two dates being tested. Now both price impacts in our example would fail to be statistically significant at a combined 5 percent level. If there was broad consensus in our example that each of the two disclosures were curative, the Holm-Bonferroni adjustment would lead to a failure to recognize these two events as having had price impact. Similarly, consider an example with five statistically significant and uncorrelated disclosures at the 1 percent level: the recalibrated level of significance using the Holm-Bonferroni adjustment leaves none of the five disclosures as material.

The court's decision to adopt the Holm-Bonferroni adjustment is important because it can make it harder to find valid price impacts more generally. While other adjustments for multiple comparisons exist that are less vulnerable to the problem of loss of statistical power, the general use of such adjustments raises deep statistical and economic questions of applicability in the event study context. It is not always the case that event studies should combine multiple disclosure events within a multiple comparisons framework. This is especially likely to be true when there are few event dates identified prior to any statistical modeling, and which are based on sound economic rationale and flow of new news.

The class decision was unusual in that the court was compelled to consider an atypical situation — one where plaintiffs alleged a large number of disclosure dates apparently based on an after-the-fact examination of calculations that (mechanically) identified abnormal returns. This is a clear example of data-mining, and some form of multiple comparisons adjustment would be recommended by most statistical experts in this setting. However, it will be up to the experts to determine whether such adjustment is genuinely necessary (and appropriately balances the different types of statistical risks appropriately) for the standard class certification case where there are a small number of predetermined corrective disclosures, having strong, data-independent economic reasons for being tested for significance individually.

The adoption of a multiple-comparisons adjustment in the class decision provides a warning to plaintiffs to use objective criteria to identify distinct corrective disclosures. If other courts follow this precedent, plaintiffs that cherry pick alleged corrective disclosure dates from a large number of possible candidates will face the prospect of price impact that will be tested collectively and found to lack statistical significance. But as the court alluded, it may be difficult to determine whether purported objective criteria for selecting “key” corrective disclosures were reverse-engineered to give the best results. As a result of the class decision, whether the price impact of corrective disclosures should be tested individually as economically distinct events, or collectively — as partial disclosures of the same information or to account for data mining — may become another important battleground for class certification experts.

Summary

Judge Lynn's class decision provides an interesting benchmark for how courts may compare the methodological choices of experts in securities class certification price impact analysis. The class decision qualified its assessments of the expert evidence as relevant to the particular case. Those qualifications may mean it will not have a strong precedent influence on future class certifications. However, an effect of the decision may be to encourage courts to be more active in adjudicating event study methodology differences between experts going forward.

—By Torben Voetmann, Paul Hinton and Matthew Aharonian, The Brattle Group Inc.

Torben Voetmann is a principal of The Brattle Group in San Francisco. He focuses on cases that involve complex economic and financial issues and has worked with clients and experts on securities cases and valuation disputes related to capital markets and financial institutions.

Paul Hinton is a principal of The Brattle Group in New York. He has 20 years of experience directing projects and providing expert testimony in securities litigation and enforcement.

Matthew Aharonian is a senior associate of The Brattle Group in Cambridge, Massachusetts, and Toronto. He has consulting and research experience with financial issues involving securities pricing, business valuation, risk management, capital markets and corporate finance.

The opinions expressed are those of the author(s) and do not necessarily reflect the views of the firm, its clients, or Portfolio Media Inc., or any of its or their respective affiliates. This article is for general information purposes and is not intended to be and should not be taken as legal advice.

[1] No. 3:02-cv-1152-M (N.D. Tex., July 25, 2015)

[2] *Halliburton Co. v. Erica P. Jon Fund Inc.* 134 S.Ct. 2398 (2014) held that defendants may rebut the *Basic v. Levinson* presumption of reliance at the class certification stage by demonstrating that the security's price was not impacted by the alleged misrepresentations or omissions.

[3] Whether the *Halliburton II* decision allows defendants to prove a lack of price impact by showing that a disclosure preceding a stock price decline did not correct any alleged misrepresentation is the subject of a pending appeal of the class decision. Motion for Leave to Appeal under Fed. R. Civ. P. 23(f), Nov. 4, 2015.

[4] The use of an event study as the de facto test of price impact in a litigation context can be unsatisfying to economists because it may have low power to detect economically meaningful price impacts and also leaves open the possibility that price impact measured in this way includes the effects of news other than the corrective disclosure(s) at issue. Separating out the effects of other news would involve a causation inquiry that *Halliburton II* deferred to the merits phase (although the pending appeal of the class decision seeks to clarify a related question of how to determine whether a disclosure corrected an alleged misrepresentation). The resulting disconnect between the economic definition of price impact and the implied legal definition used to test price impact at the class certification stage will likely be a subject of dissent in future cases.

[5] See e.g., *Halliburton II*, p. 19, "After all, plaintiffs themselves can and do introduce evidence of the existence of price impact in connection with 'event studies' — regression analyses that seek to show that the market price of the defendant's stock tends to respond to pertinent publicly reported events."

[6] *The Erica P. John Fund Inc., et al. v. Halliburton Co., et al.*, No. 3:02-cv-1152-M (N.D. Tex., July 25, 2015), p. 7. Judge Lynn also noted that the "[f]und must demonstrate that the *Basic* presumption [of reliance] applies," by showing "that *Halliburton's* stock traded in an efficient market[.]" However, she stated that *Halliburton II* "clarifies that securities fraud defendants may rebut the *Basic* presumption at the class certification stage by presenting evidence of lack of price impact." *Id.*, p. 6.

[7] The burden to provide evidence of lack of price impact is itself interesting. From an economist's point of view, it is not obvious as to what constitutes evidence of a lack of price impact to a specific event. Hypothesis tests in the event study framework are designed the other way around — they test for statistically significant price impacts. They do not provide evidence of a lack of impact, i.e., failure to reject the null hypothesis is not evidence that the null hypothesis is true. The test is not structured in this way, which is why economists say "we fail to reject the null hypothesis." It is never asserted that "we accept the null hypothesis."

[8] *The Erica P. John Fund Inc., et al. v. Halliburton Co., et al.*, No. 3:02-cv-1152-M (N.D. Tex., July 25, 2015), pp. 21-29.

[9] The court considered the Bonferroni versus the Holm-Bonferroni adjustment. Ultimately, the court narrowed its selection of the latter to this particular case, but did not offer a roadmap for determining which method would be appropriate on future cases. The court's mention of the "prospect of unacceptably high levels of Type II errors" as a reason for applying the Holm-Bonferroni adjustment, which may be linked to the discussion that it is the defendant who bears the "burdens of production and persuasion to show lack of price impact" where the rebuttable presumption under Basic applies. *Id.*, p. 7.

[10] The model estimation period is the collection of days used as a control group in testing whether the price impact of a news event is statistically significant.

[11] *The Erica P. John Fund Inc., et al. v. Halliburton Co., et al.*, No. 3:02-cv-1152-M (N.D. Tex., July 25, 2015), p. 21. The opposing experts did not differ on whether the 633-day class period itself should be used. Instead, they differed on which days during the class period that should be excluded to test "whether Halliburton's stock price movement was significantly different from what would have been predicted under normal circumstances."

[12] *The Erica P. John Fund Inc., et al. v. Halliburton Co., et al.*, No. 3:02-cv-1152-M (N.D. Tex., July 25, 2015), p. 30.

[13] See e.g., Starks, Laura T., "Discussion of "Market Microstructure: An Examination of the Effects on Intraday Event Studies," *Contemporary Accounting Research*, Vol. 10, No. 2, 1994, pp. 383-386.

[14] See e.g., Andersen, Torben G., Tim Bollerslev, Francis X. Diebold and Clara Vega, "Real-time price discovery in global stock, bond and foreign exchange markets," *Journal of International Economics*, Vol. 73, 2007, pp. 251-277.

[15] See e.g., Hotchkiss, Edith S. and Tavy Ronen, "The Informational Efficiency of the Corporate Bond Market: An Intraday Analysis," *The Review of Financial Studies*, Vol. 15, No. 5, 2002, pp. 1325-1354.

[16] *The Erica P. John Fund Inc., et al. v. Halliburton Co., et al.*, No. 3:02-cv-1152-M (N.D. Tex., July 25, 2015), p. 29-30 ("The court finds that, in this case, the use of a two-day window is inappropriate to measure price impact in an efficient market. An efficient market is said to digest or impound news into the stock price in a matter of minutes; therefore, an alleged corrective disclosure released to the market at the start of Day 1, coupled with an absence of price impact throughout Day 1, followed by a price impact on Day 2, will not show price impact as to the alleged corrective disclosure.").

[17] *The Erica P. John Fund Inc., et al. v. Halliburton Co., et al.*, No. 3:02-cv-1152-M (N.D. Tex., July 25, 2015), p. 29.

[18] *The Erica P. John Fund Inc., et al. v. Halliburton Co., et al.*, No. 3:02-cv-1152-M (N.D. Tex., July 25, 2015), pp. 22-23.

[19] Brav, Alon and J.B. Heaton, "Event Studies In Securities Litigation: Low Power, Confounding Effects, And Bias," *Forthcoming Washington University Law Review*, March 2015, p. 11.
