Noisy Factors in Law

Adriana Z. Robertson,† Pat Akey†† & Mikhail Simutin‡

For years, academic experts have championed the widespread adoption of the "Fama-French" factors in legal settings. Factor models are commonly used to perform valuations, performance evaluations, and event studies across a wide variety of contexts, many of which rely on data provided by Professor Kenneth French. Yet these data are beset by a problem that the experts themselves did not understand: In a companion article, we documented widespread retroactive changes to French's factor data. These changes are the result of discretionary changes to the construction of the factors, and they materially affect a broad range of estimates.

In this Article, we show how these retroactive changes can have enormous impacts in precisely the settings in which experts have pressed for their use. We provide examples of valuations, performance analyses, and event studies in which the retroactive changes have a large—and even dispositive—effect on an expert's conclusions. Our analysis has several implications. First, it demonstrates that these data are not sufficiently reliable to be used by experts. Second, it demonstrates a phenomenon we call the law of conservation of judgment: Methodologies that appear objective still rely on judgment of one kind or another. Rather than eliminating judgment, they simply move it around. Finally, our analysis points to the problems that arise from the commingling of academic and commercial interests.

† Donald N. Pritzker Professor of Business Law, The University of Chicago Law School.

†† Associate Professor of Finance, University of Toronto; Visiting Professor, INSEAD.

‡ Professor of Finance, University of Toronto. We thank Lucian Bebchuk, Alon Brav, Ryan Bubb, Ed Cheng, Quinn Curtis, Elisabeth de Fontenay, Jared Ellias, Jill Fisch, Joe Grundfest, Cam Harvey, Scott Hirst, Colleen Honigsberg, Marcel Kahan, Louis Kaplow, Jonathan Klick, Brian Leiter, Saul Levmore, Dorothy Lund, John Morley, Mariana Pargendler, Elizabeth Pollman, Roberta Romano, Paolo Saguato, Holger Spamann, George Vojta, and Michael Weber for valuable suggestions and discussions. This Article has benefited from comments by workshop participants at Columbia Law School, George Mason University Antonin Scalia Law School, Georgetown University Law Center, Harvard Law School, Stanford Law School, UC Berkeley School of Law, the University of Chicago Law School, the University of Oxford Faculty of Law, the University of Pennsylvania Carey Law School, the University of Toronto Faculty of Law, the University of Virginia School of Law, and the Washington University School of Law, as well as at the American Law and Economics Association Annual Meeting, the Corporate & Securities Litigation Workshop, the Labex ReFi-NYU-SAFE/LawFin Law & Banking/Finance Conference, and the Utah Winter Deals Conference. Robertson gratefully acknowledges the support of the Douglas Clark and Ruth Ann McNeese Faculty Research Fund. Katy Beeson and Levi Haas provided exceptional research assistance. All errors are our own.

INTRODUCTION

In 2013, Chief Executive Officer Michael Dell and a consortium of investors completed a management buyout of the eponymous technology company Dell Inc.[1] The transaction ultimately led to one of the most important appraisal cases of the last decade: *Dell, Inc. v. Magnetar Global Event Driven Master Fund Ltd.*[2] Management began discussing the idea of a buyout in June 2012.[3] As part of those discussions, investment bankers would have probably started to analyze the intrinsic value of the company. To

---

1 *See In re* Appraisal of Dell Inc., 2016 WL 3186538, at *2 (Del. Ch. May 31, 2016), *aff'd in part, rev'd in part sub nom.* Dell, Inc. v. Magnetar Glob. Event Driven Master Fund Ltd., 177 A.3d 1 (Del. 2017).

2 177 A.3d 1 (Del. 2017).

3 *See id.* at 6.

assess that value, the bankers would have done a discounted cash flow (DCF) analysis, which requires an estimate of the risk premium associated with the firm.[4]

Unhappy with the deal price, a group of shareholders brought an appraisal action in the Delaware Court of Chancery in 2016.[5] Like the bankers had in 2012, the parties would have hired experts to perform valuations of the firm in the lead-up to this appraisal action. And just like the bankers, these experts—perhaps from a litigation consulting firm—would have needed to come up with an estimate of the risk premium to use in the DCF analysis that they presented to the court.

The fact that different experts came to different valuations should come as no surprise. Like the use of expert witnesses elsewhere, judicial valuations are well-known "battles of the experts," where selective presentation and interpretation of available evidence are commonplace.[6] But what is much more surprising is that even if the consultants had performed *exactly the same analysis* as the bankers in 2012 and had diligently followed the best practices of valuation taught in MBA programs around the world, they may well have come to different valuations.

Unlike the typical gap between valuations presented by dueling experts, this difference would have left the bankers and the consultants scratching their heads. Surely one of them must have made a mistake: How can it be that the exact same analysis—looking at data from the same sources, inputting the exact same time period, and using the exact same computer program—generated materially different risk premia?[7] The answer is that the data changed. Buried deep within the DCF is a regression analysis used to estimate the firm's cost of capital. That analysis, in turn, often relies on a financial dataset known as the Fama-French factors, which is provided to the public free of charge on

---

4 *See, e.g.*, JONATHAN BERK & PETER DEMARZO, CORPORATE FINANCE 331 (5th ed. 2020).

5 *See infra* Part II.A.

6 *See, e.g.*, Cede & Co. v. Technicolor, Inc., 2003 WL 23700218, at *14 (Del. Ch. Dec. 31, 2003) ("[S]elective quotations . . . are certainly not unexpected in an adversarial process—especially in a 'battle of the experts' appraisal trial."); *see also* Keith Sharfman, *Valuation Averaging: A New Procedure for Resolving Valuation Disputes*, 88 MINN. L. REV. 357, 359 (2003) (describing "the phenomenon of dueling experts" as "a concern for the law of evidence generally," but particularly important in valuations); Christopher Tarver Robertson, *Blind Expertise*, 85 N.Y.U. L. REV. 174, 177 (2010) ("[I]n almost every case, the factfinder sees a 'battle of the experts.'").

7 *See infra* Part II.A.

Professor Kenneth French's website.[8] In doing so, he is providing a tremendously valuable service. These data, to which we refer as the Fama-French data, are quite literally *the* standard, and are used across a huge variety of empirical applications in finance.[9] And a little-known fact, even among financial economists, is that material retroactive changes are made to the data quite regularly.[10] Because of the frequency and magnitude of these retroactive changes, we refer to the factors as "noisy."

Obviously, nothing in this description is specific to *Dell, Inc.* After all, any analysis that conformed to best practices would have been similarly affected. It is not even really a story about discounting: the retroactive changes that we recently documented, and the implications of these changes, extend across a huge swath of empirical finance.[11] Instead, it is a story about an empirical approach[12] so widely accepted among academic experts that it became the standard operating procedure. From there, it was vigorously, and successfully, promoted by experts in a variety of legal contexts. All of this was done entirely in good faith, and it is hard to blame experts for advocating for the use of a standard academic approach. Unfortunately, however, it turned out that the experts did not really understand what was going on under the hood. As a result, they had no idea that retroactive changes

---

8 *See Current Research Returns*, KENNETH R. FRENCH (last updated July 2024), https://perma.cc/B25L-LUKD.

9 *See infra* Part I.A.3; *see also, e.g.*, Shuba Srinivasan & Dominique M. Hanssens, *Marketing and Firm Value: Metrics, Methods, Findings, and Future Directions*, 46 J. MKTG. RSCH. 293, 294 (2009) (describing the Fama-French factors as the "starting point for tackling the marketing valuation question"); Lasse Heje Pedersen, Shaun Fitzgibbons & Lukasz Pomorski, *Responsible Investing: The ESG-Efficient Frontier*, 142 J. FIN. ECON. 572, 592 (2021) (using the Fama-French factors as the primary controls in an analysis of the effect of environmental, social, and governance scores on returns); Stefano Ramelli & Alexander F. Wagner, *Feverish Stock Price Reactions to COVID-19*, 9 REV. CORP. FIN. STUD. 622, 632 (2020) (using the Fama-French factors as the primary controls in an analysis of the effect of the COVID-19 pandemic shock on equity values).

10 *See* Pat Akey, Adriana Z. Robertson & Mikhail Simutin, Noisy Factors? The Retroactive Impact of Methodological Changes on the Fama-French Factors 1–3 (July 25, 2024) (Rotman Sch. Mgmt. working paper) (available on SSRN) [hereinafter Akey et al., Impact of Methodological Changes].

11 *See infra* Part I.A.3.

12 Formally, the Fama-French model is distinct from French's data, and one could use other versions of the factors in an analysis. In practice, applying the Fama-French model almost always means relying on French's data. We discuss some institutional reasons for the reliance on the standard data source in Parts I.B and V.D.

to the Fama-French data were large and frequent enough to materially affect their analyses.[13]

These retroactive changes result from both revisions to the underlying raw data used to construct the factors and changes in the methodology used to construct them.[14] While revisions to the underlying data can explain a large fraction of the retroactive changes in the early part of the time series (up to about the mid-1960s), they account for almost none of the changes since then.[15] Rather, those changes—which affect the data most likely to be used in a wide variety of legal contexts—are caused by updates to the computer code used to construct the variables from the underlying raw data.[16] While the changes may be perfectly sensible—after all, it is widely acknowledged that the construction involves a multitude of arbitrary choices[17]—both the decision to implement them and the timing of implementation is entirely discretionary.[18] Importantly, we find no evidence that these discretionary changes improved the overall performance of the model.[19]

Because the Fama-French model in general, and the Fama-French data in particular, are so ubiquitous, the consequences of the noisy factors have bled into law in a wide variety of contexts. For the purposes of this Article, we divide these contexts into

---

13 For example, Professor Robert Dittmar characterized conversations within the scholarly community about the noisy factors to a journalist as "feel[ing] a little like group therapy," and observed that "[a]lmost all of us who work in this field have tried to re-create the Fama-French data that Ken posts on his website, and you get really close, but you're never quite there." Mary Childs & Justina Lee, *Upstarts Challenge a Foundation of Modern Investing*, BLOOMBERG (Mar. 11, 2024), https://www.bloomberg.com/news/features/2024-03-11/a-fight-over-factor-investing-tests-a-pillar-of-modern-finance. In addition to French's website, the data can be accessed through the Wharton Research Data Services (WRDS). In describing French's data, WRDS notes that they "incorporate any revisions in the historical underlying data, and thus computations that use the most recent vintage of this set may differ from computations that use an earlier vintage. The revisions are typically very small and this set is most commonly used in academic studies . . . ." *Fama French Research Portfolios and 3 Factors*, WHARTON RSCH. DATA SERVS., https://wrds-www.wharton.upenn.edu/pages/support/manuals-and-overviews/fama-french/fama-french-research-portfolios-and-factors. As discussed in more detail below, we find that a substantial amount of the changes are due not to the underlying historical data but to the construction of the factors. *See infra* notes 61–65 and accompanying text. For more information about WRDS, see *infra* note 54 and accompanying text.

14 *See* Akey et al., Impact of Methodological Changes, *supra* note 10, at 2–3.

15 *See id.* at 12–13.

16 *See id.*

17 *See, e.g.*, Mathias Hasler, *Is the Value Premium Smaller than We Thought?*, CRIT. FIN. REV. (forthcoming) (manuscript at 1) (available on SSRN).

18 *See* Akey et al., Impact of Methodological Changes, *supra* note 10, at 13.

19 *See id.* at 4.

three groups: valuation, performance evaluation, and event studies. Together, the three contexts demonstrate both how deeply enmeshed the factors are in law and the seriousness of the problems that the noisy factors create. They also lay bare a basic truth that we call *the law of conservation of judgment*: Methodologies that appear objective often smuggle in judgment of one kind or another. Rather than eliminating judgment, they simply move it around, often to some place we would not think to look.

In the valuation context, we focus on the role of experts and expert techniques in judicial proceedings. Relying on the academic finance literature, experts—particularly academic experts—have pressed for the widespread adoption of the Fama-French factors in judicial valuation.[20] In doing so, however, they inadvertently introduced a source of noise that they did not anticipate. This noise turns out to be substantial: the *Dell, Inc.* example demonstrates that the retroactive changes to the Fama-French data *alone* can generate gaps that are as large as those created by dueling experts. But unlike the classic dueling-experts setting, here, the experts themselves would have nothing to offer by way of explanation for the difference in valuations.

Turning to performance evaluation, we focus on the impact of the noisy factors on advice from and decisions by fiduciaries. While careless and disloyal fiduciaries exist, our focus is on a diligent fiduciary acting in good faith. We discuss several settings in which she might—relying on the standard prescription from the academic finance literature—employ the noisy factors in her analysis. Here, our illustrative example begins with the five largest actively managed domestic stock mutual funds. We show that depending on when an analyst downloaded the Fama-French data, her conclusion about the performance of these funds would vary substantially: whether a particular fund over- or underperformed the market, and even the relative ranking of the funds, can depend on when the data were downloaded. This leaves our fiduciary in a difficult situation: Surely the answer to whether a mutual fund over- or underperformed the market should not depend on which version of the data she used, and yet its *measured* performance very much does. And, of course, she has no way of knowing whether the next version of the factors might cause the

[20] *See infra* Part II.B.

estimated performance to change yet again. This makes it difficult, to say the least, for our fiduciary to know how to proceed.[21] This concern is not hypothetical: less than four months after our companion article on this subject, *Noisy Factors? The Retroactive Impact of Methodological Changes on the Fama-French Factors*, was first made public, a report to the managers of the world's largest sovereign wealth fund cited our finding and described it as a problem for evaluating the performance of the fund.[22]

Finally, we turn to event studies, which are used by both courts and scholars to determine the impact of an action, intervention, or other event on the financial performance of a traded security. For example, they are used to answer questions like "Did a stock price drop after a misstatement was corrected?" or "Does hostile activism by an activist hedge fund disproportionately benefit investors?" Obviously, the answer to these questions should not depend on when the data were downloaded. And yet, as with the other two contexts, it often does. We demonstrate this with an illustrative example drawn from the hedge-fund-activism literature. Using data generously shared by the leading scholars in this field, we ask whether the market reaction to hostile and nonhostile hedge fund activism is the same, on average. We find that the answer to this question depends on which version of the Fama-French data we use. The noisy factors, in other words, are noisy enough to change the results of a large-scale event study analysis. Given that courts rely on event studies extensively in securities litigation settings, this finding should be a major cause for concern.

The Fama-French factors originated in the academy. Two finance professors, Professor Eugene Fama (who went on to win a Nobel Prize[23]) and Professor Kenneth French, developed the methodology and published it in a foundational academic article.[24] The Fama-French data are provided by an academic (Professor

---

21 Naturally, the same is true of an investor making a decision on her own behalf. While we acknowledge this as an important issue, it is not the focus of this Article.

22 *See* Rob Bauer, Charlotte Christiansen & Trond Døskeland, A Review of the Active Management of Norway's Government Pension Fund Global 33 (Jan. 3, 2022) (available on SSRN) (noting that "[f]actor models have many potential difficulties," one of which is displayed in "a recent study by Akey, Robertson[,] and Simutin (2021) show[ing] that using Fama and French (2015) data in factor models is not without measurement issues").

23 *See Eugene F. Fama—Facts*, THE NOBEL PRIZE (Sept. 3, 2024), https://perma.cc/PH6C-LJVA.

24 *See generally* Eugene F. Fama & Kenneth R. French, *Common Risk Factors in the Returns on Stocks and Bonds*, 33 J. FIN. ECON. 3 (1993) [hereinafter Fama & French, *Common Risk Factors*].

French), on a website hosted by his academic institution.[25] They are ubiquitous in scholarly research,[26] and academics were big proponents of their use in law.[27] And yet, as it turns out, the experts did not really understand what it was that they were advocating for.

Our analysis of the noisy factors—and their impact on such a broad array of legal contexts—is a stark illustration of how expert analysis can go wrong when imported into a legal context. We do not take this to mean that experts have nothing useful to contribute, and we believe it would be a mistake to eschew expert analysis entirely. But it does illustrate just how precarious expert analysis can turn out to be. While all empiricists, including financial economists, know that empirical results are sensitive to model inputs and assumptions, until very recently, no one would have picked the noisy factors as an area for concern. Given this, it is hard to be confident that we can reliably identify ex ante where the next problem might arise.

Our analysis also points to an even more concerning phenomenon: the commingling of academic work and financial interests. After our previous article had been circulating in the academic community for two years, Fama and French published a research note that appeared to be a response to our findings.[28] In it, they acknowledged publicly—to our knowledge, for the first time—that the factors that are posted on French's Dartmouth University Tuck School of Business webpage are produced by staff at Dimensional Fund Advisers (DFA),[29] one of the world's largest asset managers.[30] While Fama and French's long-standing affiliations with DFA are well-known and properly disclosed, to our knowledge, this was the

---

25 *See Current Research Returns*, *supra* note 8.

26 This includes the authors of this Article: each of us have relied upon the Fama-French data in several prior academic articles. We mention this in part to make clear that our discussion in this Article is not meant as a criticism of scholars or other experts who have relied upon the Fama-French data.

27 *See infra* Parts II–IV.

28 *See generally* Eugene F. Fama & Kenneth R. French, *Production of U.S. Rm-Rf, SMB, and HML in the Fama-French Data Library* (Univ. of Chi. Booth Sch. of Bus., Working Paper No. 23-22, Dec. 2023) [hereinafter Fama & French, *Production*].

29 *See id.* at 5 ("Under our guidance, Dimensional employees produce the monthly updates, post them on a Dartmouth server, maintain the computer code, and until 2021 updated our CRSP-Compustat links.").

30 For example, Pensions & Investments ranked Dimensional Fund Advisers at number twenty-one on its list of asset managers ranked by total worldwide institutional assets under management as of December 31, 2023. *Managers Ranked by Total Worldwide Institutional Assets Under Management*, PENSIONS & INVS., https://www.pionline.com/largest-money-managers/2024-full-list.

first public disclosure that the *data* are produced by a large, for-profit asset manager.[31]

The remainder of this Article proceeds as follows. In Part I, we introduce factor models in finance in general, and the Fama-French factors in particular, and explain their ubiquity. We also explain what our finding about noisy factors means for their estimation. In Parts II through IV, we show how the noisy factors matter in three common settings: valuations (Part II), performance analyses (Part III), and event studies (Part IV). The fact that the impact of the noisy factors went undetected for so long represents a failure of the expert community. We take a step back in Part V and discuss some of the broader implications of our analysis. We then briefly conclude.

I. THE (NOISY) FAMA-FRENCH FACTORS

We begin this Part by introducing factor models in finance. While they might seem like an arcane economic concept, the intuition behind them is quite simple. They also happen to be extremely useful in a wide variety of empirical applications. Next, we discuss one particular factor model—the Fama-French model—which has risen to the top of the heap. We then discuss how factor models like the Fama-French model are estimated, including the data that are required. As we will see, the Fama-French model is very simple to implement, which has no doubt contributed to its ubiquity. Finally, we explain our finding about noise in the Fama-French data—the standard dataset that is used to estimate the model—and what it means for estimates that rely on it.

A. What Are Factor Models and Why Are They Used?

In a nutshell, factor models provide a way for a researcher to estimate what an asset's return "should" be. The basic insight is that, in a competitive market with many buyers and many sellers, the return on an investable asset should be proportional to the risk associated with that asset.[32] But since some risks can be mitigated through diversification, not all risks affect returns the same way. An investor who owns a stake in both an ice cream shop and an umbrella stand makes money, rain or shine. And if

---

[31] *See* Fama & French, *Production*, *supra* note 28, at 5.

[32] *See generally* Franciso Barillas & Jay Shanken, *Comparing Asset Pricing Models*, 73 J. FIN. 715 (2018) (describing and evaluating competing models to measure this risk).

she splits her investments up even further—say by investing in both a grocery store and a tech company—she can further insulate herself from the vagaries of chance.

Of course, there is a limit to how much risk she can diversify away: after all, every business is still participating in the overall economy. But the key is that some of the risk—the idiosyncratic component—can be diversified away. Moreover, just because a particular investor does not diversify risk away does not mean that she *could not*, and in a competitive capital market, another investor who is fully diversified could always come along. That diversified investor would not be worried about the idiosyncratic risk of a new investment, so the investment opportunity would look more attractive to her. She would, accordingly, be willing to pay slightly more for the asset, thereby bidding up the price and pushing down the return. In a competitive market, we would expect this to keep happening until the price of the asset—and every other asset—simply reflects the nondiversifiable risks associated with it. The extent to which an asset's return moves with a nondiversifiable (or "priced") risk factor is known as its "exposure" to that factor, which is sometimes referred to as the asset's "beta."[33] Other things equal, if asset *A* has a lower exposure to a priced risk factor than asset *B*, investors will be content to earn a lower return from asset *A*—and thus will be willing to pay a higher price for it—than asset *B*.

1. Single-factor (CAPM) model.

The most intuitive factor model is the Capital Asset Pricing Model (CAPM), which is a single-factor model. Under the CAPM, each asset's expected return is determined solely by the asset's sensitivity to the return of the market as a whole.[34] In other words, the expected return (in excess of a risk-free investment) is:

$$R^e_{i,t} = \alpha_i + \beta_{i,1} \times R^e_{m,t} \tag{1}$$

The CAPM predicts that alpha (α), which captures the extent to which an asset over- or underperforms, should be zero ex ante.[35]

---

[33] *See* BERK & DEMARZO, *supra* note 4, at 379–83, 456–61.

[34] *See id.* at 421–28.

[35] *See id.* The model relates the expected return of the asset to the expected return of the market. Since expected returns are not observable, the model is usually estimated with *actual* returns. *See id.* at 475 (describing a typical time horizon as "two years of weekly return data or five years of monthly return data").

As a result, the asset's ex ante expected return is simply its beta (β) multiplied by the return on the market (again, in excess of a risk-free investment). In other words, the asset's expected, or "fair," return should be proportional to its exposure to the market. The difference between the return on the market and the return on a risk-free investment is often called the market risk premium, or simply the market return.[36]

2. Fama-French three-factor model.

In the early 1990s, Professors Fama and French found that supplementing the CAPM with two additional factors improved the model's success at explaining returns. This finding was first articulated in their foundational 1993 article *Common Risk Factors in the Returns on Stocks and Bonds*,[37] which remains one of the most cited articles in financial economics.[38] This has become known as the Fama-French three-factor model, or simply the three-factor model.

The intuition behind the CAPM extends to the three-factor model: other things equal, an asset with a higher exposure to one of the three priced factors (again, captured by a beta) will command a higher return than an asset with a lower exposure. Mathematically, this is summarized as:

$$R^e_{i,t} = \alpha_i + \beta_{i,1} \times R^e_{m,t} + \beta_{i,2} \times HML_t + \beta_{i,3} \times SMB_t \qquad (2)$$

The only difference between Equation 1 and Equation 2 is the addition of two factors—HML and SMB—along with their associated betas. HML, also known as the value factor, represents the return on a portfolio of *high* (H) book-to-market stocks *minus* (M) the return on a portfolio of *low* (L) book-to-market stocks (hence, "high minus low," or HML). High book-to-market stocks—stocks of companies with relatively more assets compared to the value ascribed to the company by the stock market—are colloquially known as value stocks,[39] since they trade at a low price and can therefore be thought to represent good value to investors. Conversely, low book-to-market stocks—stocks of companies with

---

[36] In other words, $R^e_{i,t}$ refers to the return (in excess of the risk-free rate) of asset i in period t. α_i refers to the alpha of asset i. $\beta_{i,1}$ refers to the beta of asset i. $R^e_{m,t}$ refers to the return (in excess of the risk-free rate) of the market in period t.

[37] Fama & French, *Common Risk Factors*, *supra* note 24.

[38] As of July 2024, the article had over 36,000 citations on Google Scholar.

[39] *See* BERK & DEMARZO, *supra* note 4, at 507.

relatively few assets compared to the value ascribed to the company by the stock market—are colloquially known as growth stocks[40] on the theory that the market must be anticipating that the company will grow quickly to justify the high valuation. On average, value stocks have earned higher returns than growth stocks over the past several decades, and the difference between the two is known as the value premium.[41]

SMB, also known as the size factor, has a similar structure. It represents the return on a portfolio of stocks of *small* (S) companies *minus* (M) the return on a portfolio of *big* (B) companies (hence, "small minus big," or SMB). Like value stocks, small stocks have tended to earn higher returns than big stocks, although the difference (known as the size premium[42]) is often thought to be smaller than the value premium.[43]

Putting all of this together, the three-factor model says that an asset's expected return can be estimated by calculating its exposure to each of the three priced factors—the market, value, and size—and then multiplying each of these exposures with its respective premium. To the extent that an asset (whether it be a stock, a mutual fund, or anything else) has a higher return than that, it has outperformed. To the extent that it has a lower return, it has underperformed. Estimating the expected return of an asset and computing its ex post performance relative to that expected return made its way into law through its adoption by modern finance.

3. Applications of factor models.

It is not an exaggeration to say that factor models are ubiquitous in modern finance. In this Section, we sketch out the three empirical applications with the most direct implications for legal settings.

a) Valuation. The textbook approach to valuing an asset that generates income is the discounted cash flow (DCF) technique.[44] The income could be associated with a company, a

---

40 *See id.*

41 *See generally* Lu Zhang, *The Value Premium*, 60. J. FIN. 67 (2005). The value premium has deteriorated substantially in recent years. *See* Eugene F. Fama & Kenneth R. French, *The Value Premium*, 11 REV. ASSET PRICING STUD. 105, 108 (2020) (showing that over the 1963–2019 sample period, the value premium was much larger in the first half than in the second half).

42 *See* BERK & DEMARZO, *supra* note 4, at 507.

43 *See generally* Clifford Asness, Andrea Frazzini, Ronen Israel, Tobias J. Moskowitz & Lasse H. Pedersen, *Size Matters, If You Control Your Junk*, 129 J. FIN. ECON. 479 (2018).

44 We mean this quite literally. *See infra* note 102 and accompanying text.

project, an asset (like a factory), a contractual right, or any other claim. The technique works by projecting the net income stream associated with the asset (i.e., its cash flows) into the future, and then figuring out what a claim on those cash flows is worth today. Because cash today is worth more than cash next year, and a sure bet is worth more than a risky payoff, future cash flows are adjusted—or discounted—for both time and risk.

The standard way to estimate the right discount rate is to use a factor model. After all, factor models are designed to estimate what an asset's return should be given that asset's exposure to priced risk factors.[45] Just as importantly, they are easy to estimate empirically,[46] which makes them practical to implement. And finally, they are relatively intuitive, making them more appealing in contexts where the valuation will have to be explained to non-specialists.

To go from factor exposures (i.e., betas) to a discount rate, all a researcher, practitioner, or other analyst needs to do is multiply each beta by its respective factor premium (in the case of the three-factor model, this is the market risk premium, the value premium, and the size premium), and add in the risk-free rate. In many valuation contexts, the standard practice is to use a single-factor CAPM,[47] which can also be estimated using the Fama-French data.[48]

b) Performance evaluation. A second application is performance evaluation. The basic intuition of this application is perhaps even simpler than valuation. Since alpha captures the extent to which an asset over- or underperforms the factor model, it is often used as a measure of ex post performance. For this reason, a positive alpha is interpreted as "beating the market."[49]

There are several legal and financial contexts in which we might want to measure performance. For example, we might

---

45 *See supra* notes 33–43 and accompanying text.

46 *See infra* Part I.B.

47 *See* Kenneth Ayotte & Edward R. Morrison, *Valuation Disputes in Corporate Bankruptcy*, 166 U. PA. L. REV. 1819, 1827 (2018) ("Among academic finance scholars, two approaches to calculating the required return on equity [in a DCF context] are most common and widely advocated: the capital asset pricing model (CAPM) and the Fama-French three-factor model."). The authors went on to note that the standard CAPM is the dominant method used in a variety of contexts, including by investment banking advisers and chief financial officers. *Id.* at 1829.

48 *See infra* Part I.B.

49 *See infra* notes 126–27 and accompanying text; *see also* Larry R. Gorman & Robert A. Weigand, Measuring Alpha Based Performance: Implications for Alpha Focused, Structured Products 14 (Nov. 28, 2007) (available on SSRN).

want to know whether a particular asset manager is doing a good job. Just looking at a portfolio's return is not an adequate measure of performance because it does not take into account the fact that investments with greater exposure to priced factors tend to have higher returns. Focusing on alpha instead isolates the component of returns not attributable to priced risk. The same logic applies to a particular firm: we can use a firm's alpha to evaluate whether and to what extent that firm beat the market.[50]

c) Event studies. A third category of applications for factor models is to study whether an asset's return was unusually high or low around a particular event. The asset could be a company's stock, an investment fund, or anything else. This is known as an event study. To perform an event study, one typically estimates a factor model in the period leading up to (but not including) the event in question. The estimated betas from this analysis are then used to calculate the asset's predicted return during the event window. Subtracting that predicted return from the asset's actual return yields its abnormal return during the time period of interest.

Event studies have a myriad of uses in law and finance. They are used extensively in securities law both to determine whether a company's share price "really" declined during a particular time period and to measure how much it declined. They are also used extensively in scholarly contexts when researchers wish to estimate the impact of some policy, intervention, or other event. In these contexts, a researcher will typically perform an event study on a large number of firms and will study the average abnormal performance of the affected firms.[51]

B. How Are Factor Models Estimated?

Estimating a factor model is simple, which is part of its appeal. All one needs is software capable of running a linear regression (Microsoft Excel will do just fine, as will any number of widely used statistical or general-purpose programming languages), the historical returns on the asset of interest, and the historical returns of the factors. A linear regression can be estimated where the dependent (or outcome) variable is the return on the asset of interest, and the independent variables are the returns of each of the factors. The estimated coefficients on each of the factors represent the betas—or the asset's exposure to each of

---

50 *See* Gorman & Weigand, *supra* note 49, at 14.

51 We provide an example of such an analysis in Part IV.C.

the factors. The intercept represents the asset's alpha. A positive or negative alpha indicates whether the asset over- or underperformed relative to its factor exposure.

Since an analyst will typically have easy access to returns of the asset in question, the only other data she will need is the return on the factors. If she is estimating a three-factor model, she is in luck: those are freely available to anyone with an internet connection through Professor French's online data library, hosted by his home institution.[52] This library contains a wealth of data, including daily, weekly, and monthly returns of the three factors for both domestic and foreign markets.[53] Because the data are updated regularly, the files typically cover the entire period from the 1920s to a few months prior to the present day. For convenience, we refer to these files as the Fama-French data. These data are also distributed through the Wharton Research Data Services, a widely used source of academic data in finance.[54]

While one could, in theory, construct one's own factors for use in an analysis, the overwhelming majority of researchers prefer to rely on the Fama-French data for a variety of reasons.[55] Constructing her own factors raises the concerns that the researcher might be manipulating the data for her own purposes; using data provided by an arm's-length third party eliminates this concern. Moreover, because the Fama-French data are so widely used and are provided by a highly regarded third party, she does not need to explain what they are or answer pointed questions about why she chose to use them. Additionally, the fact that the data are free means that cost is not a barrier.[56] Finally, she can also estimate a single-factor CAPM using the Fama-French data without having to download any additional data. In short, there is no discernable upside to constructing her own factors—only substantial downsides.

When she estimates the regression, the researcher will choose what time period she wants to use. For example, she might be interested in the performance of a group of mutual funds during the period from 2005 to 2010. This is known as the sample

---

52 *See Current Research Returns*, *supra* note 8.

53 *See id.*

54 *Home*, WHARTON RSCH. DATA SERVS., https://perma.cc/2GG3-9YSD. More than 530 institutions, representing 75,000 individual users in 38 countries, subscribe to this service. *Id.*

55 For a discussion of the reasons why researchers use the standard data, see *infra* Part V.D. While there are no doubt some academic papers in which the authors construct their own factors, we are unable to come up with any examples.

56 We return to these incentives in Part V.D.

period. For the discussion that follows, it is crucial to keep in mind that the sample period—the period being analyzed—is distinct from the date that the analyst downloaded the data. After all, she could be interested in studying 2005–2010 performance in June 2011, August 2016, or January 2022. In all three cases, the sample period is the same: 2005–2010. One would hope that an asset's 2005–2010 performance wouldn't depend on whether she looks at the data in 2016 or 2022.

C. The Noisy Factors

It turns out that it does. Specifically, in a companion article, we showed that there are substantial retroactive changes to the Fama-French data.[57] When we then compare the Fama-French data from each available year to every other available year, we find changes throughout the series. For example, using data beginning in 1964 and comparing the 2005 vintage of the Fama-French data to the 2006 vintage, we find that the monthly factor returns differ more than half the time.[58] To be clear, we hold the sample period constant when we make these comparisons, so these differences reflect retroactive changes in the data for the *same sample period*. So, for example, if one downloads the data for a given period—say, the year 2002—the data will be different based on when the download took place. If one downloaded the data for the year 2002 at the beginning of 2005, one would get a different set of numbers than if one downloaded data for the same time period at the beginning of 2006. These differences can be substantial. Again, just comparing the 2005 and 2006 vintages, 66% of monthly returns for the value factor differ by more than

---

[57] *See* Akey et al., Impact of Methodological Changes, *supra* note 10, at 1. To do so, we initially used the Wayback Machine to download data from Professor French's website. In November 2022, after our previous article had been circulating for over a year, French's website was updated to provide vintages of the factors from the start of our initial sample period, which enabled us to obtain data for all years. We therefore updated the analysis to reflect this newly available information. We continue to rely on data from the Wayback Machine for other data required in our analysis.

[58] The extent to which the returns differ across vintages varies by factor. For example, the vast majority of monthly HML returns (98%) and SMB returns (96%) differ between the adjacent 2005 and 2006 vintages. The market risk premium is the most consistent between these two adjacent vintages, but even here, 49% of monthly returns exhibit retroactive changes. *Id.* at tbl.1.

1% annually,[59] a very substantial change. The differences between vintages tend to get even larger when we extend the time period between them.[60]

There are only three possible explanations for the retroactive changes to the Fama-French data: (1) the underlying raw stock return and accounting data are changing, (2) the computer code that is used to construct the factors using that data is changing, or (3) both are changing. We investigate this using archived versions of the raw data needed to construct the factors. We then use these data to construct our own versions of the factors by running the same code on the archived versions of the raw data.[61] By comparing changes in these fixed-code factors (constructed using archived data) to the changes in the archived versions of the Fama-French data, we can assess the extent to which the changes in the latter are driven by changes in the underlying data. This reveals a striking pattern. Using the earliest available and most recent data, we find that changes due solely to data updates—as measured by changes in the fixed-code factors—explain almost half (42%–49%, depending on the factor) of the changes in the distant past (from 1926 to the mid-1960s).[62] Changes in the data relating to the more recent past (the mid-1960s to the most recent data available), in contrast, explain essentially none of the changes in the Fama-French data.[63] Instead, they are driven by discretionary changes to the computer code used to construct the factors.[64] Interestingly, while the Fama-French data change (and these changes materially affect results), the changes do not seem to represent improvements. Rather, using standard statistical techniques designed to compare the performance of different asset pricing models, we find no evidence that later vintages perform

---

59 *Id.*

60 *Id.*

61 Because the code used to construct the Fama-French data is not public, we created our own version based on publicly available descriptions of the methodology. We did so by relying on descriptions in the literature, as well as on information provided by Professor French on his website. *See generally* Fama & French, *Common Risk Factors*, *supra* note 24; James L. Davis, Eugene F. Fama & Kenneth R. French, *Characteristics, Covariances, and Average Returns: 1929 to 1997*, 55 J. FIN. 389 (2000); *Description of Fama/French Factors*, KENNETH R. FRENCH, https:perma.cc/CMZ7-D2BJ; *Variable Definitions*, KENNETH R. FRENCH, https://perma.cc/FM2F-MGMK.

62 *See* Akey et al., Impact of Methodological Changes, *supra* note 10, at fig.2.

63 *Id.*

64 *See infra* Part V.C.

better than earlier ones.[65] This is important for two reasons. First, it suggests that whatever is causing the Fama-French data to change does not seem to be leading to an overall improvement—or, for that matter, a deterioration—in the model's ability to price assets. And second, it means that there is no particular reason to think that using the most recent vintage will lead to more accurate estimates.

These retroactive changes have substantial effects on estimated alphas and betas. For example, switching between the 2010 and 2022 factor vintages causes 28% of single stock alphas—estimated using textbook techniques—to change by more than 100 basis points (i.e., one percentage point) per year.[66] The estimated betas also change substantially: switching between 2010 and 2022 causes 13% of market betas to change by more than 0.1.[67] Assuming a market risk premium of about 5% per year, this represents a difference in the estimates of the cost of equity of about 50 basis points per year. The differences are also large for the HML and SMB factors: switching vintages causes 26% of HML betas and 9% of SMB betas to change by more than 0.1.[68] In contrast, when we use our fixed code factors, which have no discretionary changes, the effects virtually disappear: instead of 28%, only 1.3% of alphas change by more than 1%, a more than twenty-fold reduction,[69] with similarly reduced effects on estimated betas.[70] This confirms that the effects are driven by the

---

65 Specifically, we implement two types of model comparison tests from the financial economics literature. The first type is Gibbons, Ross, and Shanken (GRS) tests, which use the model to price a series of test assets. *See generally* Michael R. Gibbons, Stephen A. Ross & Jay Shanken, *A Test of the Efficiency of a Given Portfolio*, 57 ECONOMETRICA 1121 (1989). When comparing two models using this metric, a "better" model is one that achieves alphas that are closer to zero. In implementing the GRS tests, we use different vintages of the Fama-French factor data as the models. We use the standard test assets from the asset pricing literature: twenty-five portfolios (sorted into size quintiles and book-to-market quintiles) and seventeen industry portfolios, both from French's website. To avoid the problem of having to select test assets (and concerns that French's test asset data has its own changes), we also implement squared Sharpe ratio tests. When comparing two models using this metric, the one with the higher squared Sharpe ratio is "better." Here again, we use different vintages of the Fama-French factor data as the models in implementing these tests. In both cases, we find no consistent evidence that the changes to the factors are causing the model's performance to either improve or deteriorate. Akey et al., Impact of Methodological Changes, *supra* note 10, at 27–28.

66 Akey et al., Impact of Methodological Changes, *supra* note 10, at fig.7.

67 *Id.*

68 *Id.*

69 *Id.*

70 *Id.*

discretionary changes to the Fama-French factors and not by updates to the raw data.

The effect is not limited to individual stocks: we find that switching vintages has a similar effect on mutual fund alphas and betas. When studying mutual funds, we follow the standard approach and use one year of data in our baseline analyses. We find that switching between the 2010 and 2022 vintages causes 53% of estimated mutual fund alphas to change by more than 1% per year, and 37% of statistically significant alpha estimates to lose significance.[71] In other words, the extent to which a fund is considered to have under- or outperformed the market can change dramatically solely based on when the analyst downloaded the data. And, perhaps more importantly, her conclusions about whether a fund manager over- or underperformed—in a way that is statistically significant—are highly sensitive to when the analyst downloaded the data. It goes without saying that this should have no bearing on her evaluation. And indeed, if we used the fixed-code factors—which eliminate the discretionary changes to the construction of the factor—this is what we would see. When using fixed-code factors, only 0.24% of estimated alphas change by more than 1%, and less than 3% lose statistical significance.[72]

* * *

In sum, we find that there are substantial changes to the Fama-French factors, which in turn have enormous effects on estimates that rely on that data. These changes are driven by changes to the methodology used to construct the factors, and not by updates to the underlying raw data. They are, in other words, the product of judgment, exercised far away from the end user of the data. These discretionary choices become even more concerning in light of the recent admission by Fama and French—which came years after *Noisy Factors? The Retroactive Impact of Methodological Changes on the Fama-French Factors* began circulating in the academic community—that the Fama-French factors are actually produced by employees at DFA,[73] a $750 billion asset manager with a vested interest in the performance of the value factor.[74] Even if, as they represent, Fama and French "continue to determine the rules,

---

71 Akey et al., Impact of Methodological Changes, *supra* note 10, at fig.5.

72 *Id.*

73 *See* Fama & French, *Production*, *supra* note 28, at 5.

74 *See infra* Part V.C.

definitions, and process used to form [the factors],"[75] this commingling raises thorny questions, particularly given that, as we showed in our previous work, the discretionary decisions consistently improve the performance of the value factor without improving the performance of the overall model.[76] We return to this issue in Part V.C. Because of the magnitude and extent of these retroactive changes, we sometimes refer to the Fama-French data as "the noisy factors."

II. VALUATION WITH NOISY FACTORS

Having explained the Fama-French data and how they are used, we now show how much the retroactive changes in the data matter for legal applications. We begin with valuation. Since we already know that they affect beta estimates,[77] it is easy to see why discount rates calculated using those betas would also be affected. What may be more surprising is the magnitude of this effect.

To show this, we begin with an illustrative example: the high-profile appraisal of Dell Inc.[78] In that setting, it turns out that the noisy factors generate a gap in beta estimates that is as large as the gap between the estimates put forward by the dueling experts.[79] To be clear, we do not believe that this gap was caused by the noisy factors—indeed, we have no reason to think that they had anything to do with it. Nonetheless, we think that this comparison is instructive. To the extent that the discrepancies caused by dueling experts are large and concerning, this implies that the discrepancies caused by the noisy factors can be equally large and equally concerning. Of course, *Dell, Inc.* is just an illustration, and § 262 appraisal actions are just one example of valuation. In Part II.B, we discuss other legal contexts where the same issue arises.

A. The Dell Appraisal

In June 2012, a well-known investor approached Chief Executive Officer Michael Dell about whether he would consider

---

75 *See* Fama & French, *Production*, *supra* note 28, at 5.

76 *See id.* at 4–5.

77 *See supra* Part I.C.

78 *See Dell, Inc.*, 177 A.3d 1.

79 Because valuations rely on beta estimates, this gap in betas will in turn affect the ultimate valuation.

leading a management buyout of Dell Inc.[80] After a lengthy process, including discussions with several financial advisers and prospective acquisition partners, as well as several rounds of offers and counteroffers, the final offer put forward by Mr. Dell and his private equity backer was approved at a special meeting of Dell's stockholders on September 12, 2013.[81] A group of dissenting former stockholders exercised their appraisal rights under § 262 of the Delaware General Corporate Law.[82] Vice Chancellor J. Travis Laster held a four-day trial in the Delaware Chancery Court in October 2015[83] to determine the fair value of the shares.[84]

As is typical in an appraisal action, both sides engaged experts, each of whom performed a DCF analysis to value the company.[85] As is also typical, the two experts came to markedly different valuations.[86] Somewhat less typically, Vice Chancellor Laster chose to perform his own DCF analysis and used "DCF methodology exclusively to derive a fair value of the Company."[87]

After evaluating the inputs and assumptions of the valuations performed by the parties' experts, Vice Chancellor Laster selected from each the parts that he found most convincing.[88] Because of this, *In re Appraisal of Dell Inc.*[89] is a convenient setting to illustrate how the noisy factors could affect results in an appraisal action. To do so, we now switch from the record of what actually happened to what might have happened.

---

80 *See In re* Appraisal of Dell Inc., 2016 WL 3186538, at *2 (Del. Ch. May 31, 2016), *aff'd in part, rev'd in part sub nom. Dell, Inc.*, 177 A.3d 1.

81 *See id.* at *19.

82 *See* DEL. CODE ANN. tit. 8 § 262 (2013).

83 *See In re Appraisal of Dell Inc.*, 2016 WL 3186538, at *1.

84 As the Delaware Supreme Court explained in *Dell, Inc.*, § 262 "allows stockholders who perfect their appraisal rights to receive 'fair value' for their shares as of the merger date instead of the merger consideration. The appraisal statute requires the Court of Chancery to assess the 'fair value' of such shares and . . . 'take into account all relevant factors.'" 177 A.3d at 5.

85 *See In re Appraisal of Dell Inc.*, 2016 WL 3186538, at *45.

86 The expert for the dissenting stockholders performed a DCF and concluded that the fair value of the company on the closing date was $28.61 per share. *Id.* The expert for Dell performed his own DCF and concluded that its fair value on the closing date was $12.68 per share. *Id.* The deal price was $13.75 per share. *See Dell, Inc.*, 177 A.3d at 5.

87 *In re Appraisal of Dell Inc.*, 2016 WL 3186538, at *51. The Delaware Supreme Court held that the Chancery Court's decision to put no weight on the deal price was erroneous. *See Dell, Inc.*, 177 A.3d at 5.

88 *See Dell, Inc.*, 177 A.3d at 5 (describing the Chancery Court as having relied "exclusively on its own discounted cash flow ('DCF') analysis").

89 2016 WL 3186538 (Del. Ch. May 31, 2016), *aff'd in part, rev'd in part sub nom. Dell, Inc.*, 177A.3d 1.

To begin, let us suppose—as was in fact the case—that Dell management sought the advice of an investment bank in this process. Drawing on what she learned in her MBA classes, a diligent banker tasked with this project might have started by downloading the Fama-French data.[90] She would have used those data, along with Dell's stock returns, to estimate the beta of the company's stock by running a regression. If she had followed standard best practices, she would have estimated a single-factor CAPM beta using five years of monthly data.[91]

For argument's sake, suppose that she had downloaded the data in May 2012.[92] This file contained data through the end of the first quarter of 2012. Supposing that she wanted to retain the most up-to-date data available, she would have included data from April 2007 through the end of March 2012 in her analysis. Had she done this, she would have come up with a CAPM beta of 1.306. She would then have plugged this beta into a formula to estimate the cost of capital, which in turn would have served as the discount rate in her DCF model.

Fast forward a couple of years to late 2015, shortly before the Chancery Court's decision in the appraisal action. Suppose that in November 2015, an associate at a litigation consulting firm performed the exact same analysis as our banker had in 2012. The only difference is that the consultant would have downloaded the data three and a half years later. Like the banker, suppose that he included data from April 2007 through the end of March 2012 and estimated a CAPM beta using five years of monthly data. Had he done that, using the version of the Fama-French data that he downloaded in November 2015, he would have obtained a CAPM beta of 1.352. Like the banker, he would have used this to calculate a discount rate.

We can see right away that the consultant's discount rate—estimated in late 2015—would be higher than the banker's, leading to a lower valuation. Crucially, this would be true holding constant all modeling decisions. Of course, things generally are not

---

90 *See Current Research Returns*, *supra* note 8.

91 *See, e.g.*, TIM KOLLER, MARC GOEDHART & DAVID WESSELS, VALUATION 283–84 (6th ed. 2015) (providing an example of how to estimate a company's beta and using five years of monthly data).

92 We chose this date for convenience. The precise date upon which a banker might have begun the analysis is not clear from *Dell, Inc.* or *In re Appraisal of Dell Inc.* However, it stands to reason that a well-known investor would have performed—or asked someone else to perform—a valuation of the company before approaching Mr. Dell in June 2012. *See supra* note 80 and accompanying text.

held constant due to the incentives faced by expert witnesses. But the noisy factors add an extra dimension to the gap between valuations. And importantly, this dimension is entirely hidden, not just from judges but from the experts themselves.

The Dell illustration is especially instructive on this point. When he got to the portion of his own DCF where he had to come up with a beta estimate, Vice Chancellor Laster had this to say:

> The experts disagreed about beta. [The dissenting shareholders' expert] derived a beta of 1.35 by analyzing the Company's peers. [Dell's expert] derived a beta of 1.31 by analyzing weekly observations over a two-year period. A beta specific to the Company is more targeted than a blended beta calculated from peer companies, particularly when both experts opined that the Company had few peers. This decision uses [Dell's expert]'s beta.[93]

Vice Chancellor Laster's discussion is almost eerie in light of what we have already seen about the noisy factors. By coincidence, the experts retained by the parties ended up proposing beta estimates that are, to three significant digits, identical to the estimates obtained by our hypothetical banker and consultant. This is particularly consequential given that, in cases such as this, we would expect the data to be downloaded at different times by different players in the process.

Far from implying that the noisy factors do not matter, this highlights just how important they are. The gap created solely by the noisy factors, holding everything else equal, is as large as the gap created by experts that *did* diverge dramatically on methodology. And that gap is actually quite conservative; in untabulated results, we find that many plausible deviations from best practices, combined with the noisy factors, substantially increase the extent to which the beta estimates diverge.

Perhaps more importantly, the gap created by the noisy factors would have gone entirely unexplained. Before our previous work was made public, no research or commentary of which we are aware indicated that downloading the Fama-French data at a different time could affect results at all, let alone that it was likely to have a material effect. And even now that we are aware of the phenomenon, we have no basis for saying that any one factor vintage produces estimates that are more accurate, reliable,

---

[93] *In re Appraisal of Dell Inc.*, 2016 WL 3186538, at *49 (citations omitted).

or otherwise better than any other.[94] In other words, while experts are expected to provide evidence, which the court can then evaluate, the only explanation here is "because the data changed." This explanation is not particularly satisfying, nor is it a normatively desirable basis for driving the results in an appraisal action.

Of course, the Delaware Supreme Court reversed and remanded the Chancery Court's appraisal award.[95] It did so not because of any errors in Vice Chancellor Laster's DCF analysis, but rather because the Chancery Court chose to give no weight to the deal price, something that the Delaware Supreme Court held to be an abuse of discretion.[96] While the Delaware Supreme Court opined that "the record as distilled by the trial court suggests that the deal price deserved heavy, if not dispositive, weight,"[97] it declined to instruct the Chancery Court to enter a judgment at the deal price, leaving the door open for the Chancery Court to "weigh a variety of factors in arriving at fair value."[98]

Even though the Delaware Supreme Court declined to create a presumption in favor of market prices in *Dell, Inc.*, the decision was widely viewed as supporting the proposition that deal prices are at least probative, especially in appraisals involving arm's-length transactions.[99] The doctrine has continued to evolve in the years since *Dell, Inc.* Courts have continued to note the problems associated with DCF valuation,[100] but the Delaware Supreme Court has made it clear that its recent appraisal decisions have not "ruled out using any recognized valuation methods to support

---

94 *See supra* note 65 and accompanying text.

95 *See Dell, Inc.*, 177 A.3d at 19.

96 *See id.* at 23–24 ("[T]here is a dissonance between the key underpinnings of the decision to disregard the deal price and the facts as found, and this dissonance distorted the trial court's analysis of fair value.").

97 *Id.* at 23.

98 *Id.* at 44.

99 *See, e.g.*, Victor Lewkow, Meredith E. Kotler & Mark E. McDonald, *Analysis of Delaware Supreme Court's Dell Appraisal Decision*, HARV. L. SCH. F. ON CORP. GOVERNANCE (Dec. 19, 2017), https://perma.cc/W69T-JJDD ("*Dell* thus indicates that only compelling evidence of market failure will justify departing from deal price in cases involving arm's-length mergers."); Norbert B. Knapke II & Daniel E. Wolf, *Negotiated Deal Price Is Best Evidence of Fair Value—Delaware Dispels the Dell Appraisal Overhang*, KIRKLANDPEN (Kirkland & Ellis LLP), Dec. 19, 2017, at 1, 2 ("While it declined to create a presumption in favor of the deal price, the Supreme Court's opinion was unequivocal in its view that in *Dell* the deal price was the best indicator of value.").

100 *See e.g.*, Verition Partners Master Fund Ltd. v. Aruba Networks, Inc., 210 A.3d 128, 141 (Del. 2019) (noting the "imprecision" of DCF valuation, including the need to estimate "(i) future free cash flows; (ii) the weighted average cost of capital (*including the stock's beta*); and (iii) the perpetuity growth rate" (emphasis added)).

fair value."[101] And of course, § 262 appraisal actions are not the only context in which courts must perform a valuation. We discuss several other examples in the next Section.

B. Experts and Judicial Valuation

Experts, especially academic experts, have long been strong proponents of the use of factor models for valuation. This is true both for factor models in general and for the Fama-French model in particular. Their support is entirely understandable. After all, these models are the textbook approach in financial economics, and it is precisely the role of the expert to apply standard expert approaches.

If you open a standard corporate finance textbook and flip to the valuation section, it is all but guaranteed that the DCF approach will feature prominently.[102] The single-factor CAPM remains commonly used in that setting (which is why we used it in the Dell illustration in Part II.A), but scholars have suggested that the three-factor model would be an improvement.[103] And moreover, as we have seen, because the Fama-French data can be—and sometimes are—used to estimate a single-factor model, the noisy factors can still have an effect in a single-factor setting.

Another context beyond § 262 appraisals in which judicial valuation plays an outsized role is in restructurings under Chapter 11 of the Bankruptcy Code.[104] Most fundamentally, in a traditional reorganization, the court distributes claims (typically in the form of debt and equity) in the debtor company to its creditors. In order

---

[101] Fir Tree Value Master Fund, LP v. Jarden Corp., 236 A.3d 313, 323–24 (Del. 2020).

[102] *See, e.g.*, KOLLER ET AL., *supra* note 91, at 135 (describing enterprise-level DCF as "a favorite [valuation approach] of practitioners and academics"); *see also* STEPHEN J. LUBBEN, CORPORATE FINANCE 194 (3rd ed. 2021) (describing DCF as "the most common valuation method used in many legal settings"); JEFFREY J. HAAS, CORPORATE FINANCE 83 (2nd ed. 2021) ("The [DCF] method, or a variant thereof, is the most common valuation method employed by the financial community today."); STEPHEN A. ROSS, RANDOLPH W. WESTERFIELD, JEFFREY JAFFE & BRADFORD D. JORDAN, CORPORATE FINANCE 182 (12th ed. 2019) (introducing the "diverse applications of . . . [DCF] valuation").

[103] *See, e.g.*, Ayotte & Morrison, *supra* note 47, at 1837 (noting that "[a] substantial number of scholars believe that the model is superior to the CAPM"). Nonetheless, Professors Kenneth Ayotte and Edward Morrison's research suggests that courts may be hesitant to accept the Fama-French model—at least in the corporate bankruptcy context—because of its minimal adoption in the valuation industry. *See id.*

[104] *See generally* 11 U.S.C. § 101 et seq. (defining insolvency, among other things, in § 101(32)(A)–(B)).

for it to do so, the court must first assign a value to the restructured company.[105] It should therefore come as little surprise that many important articles on judicial valuation are focused on the bankruptcy context.[106]

DCF valuations are the norm in this setting. In a recent article, Professors Kenneth Ayotte and Edward Morrison examined almost twenty years of Chapter 11 valuation disputes. Of the 141 cases they identified and analyzed, 122 used a DCF valuation technique.[107] The discount rate was often a contentious feature, and they found that experts fought over the discount rate 46% of the time.[108] They also found that the experts in these disputes often estimate discount rates in ways that diverge from both the CAPM and Fama-French model, and they were highly critical of these departures.[109] Instead, they "recommend[ed] that courts consistently apply the CAPM."[110] Of course, as we saw in the Dell illustration, a CAPM beta calculated using the Fama-French data is vulnerable to the effects of noisy factors. In other words, the noisy factors undermine the best valuation method, done correctly.

DCF valuations also appear in many other areas of law. Delaware courts may have cooled on DCF valuations in the appraisal context,[111] but they continue to rely on DCF valuations in establishing entire fairness. While decisions of corporate officers

---

[105] *See* Ayotte & Morrison, *supra* note 47, at 1824 ("Key moments in a Chapter 11 reorganization hinge on valuation.").

[106] *See, e.g.*, Douglas G. Baird & Donald S. Bernstein, *Absolute Priority, Valuation Uncertainty, and the Reorganization Bargain*, 115 YALE L.J. 1930, 1952–63 (2006) (arguing that uncertainty about judicial valuations can explain many observed departures from absolute priority in corporate reorganizations); *see also* Anthony J. Casey & Julia Simon-Kerr, *A Simple Theory of Complex Valuation*, 113 MICH. L. REV. 1175, 1182 (2015) (arguing that valuations are no different from other forms of fact-finding and therefore should be governed by traditional evidentiary rules); Ayotte & Morrison, *supra* note 47, at 1821–23 (analyzing twenty years of Chapter 11 valuation disputes).

[107] *See* Ayotte & Morrison, *supra* note 47, at 1832.

[108] *See id.* at 1833. Another feature of the DCF approach—the projected cash flows—was even more contentious. In their sample, Professors Ayotte and Morrison found that the experts disputed these in 74% of cases. *See id.*

[109] *See id.* at 1841–42.

[110] *Id.* at 1842. While DCF is the most theoretically rigorous approach, Professors Ayotte and Morrison have argued that bankruptcy judges struggle to police deviations from best practices by experts. As a result, they may be better off using the simpler—but less rigorous—multiples-based approach. For the same reason, Professors Ayotte and Morrison advocated for the use of market-based measures wherever possible. *See* Ayotte & Morrison, *supra* note 47, at 1846.

[111] *See supra* notes 99–101 and accompanying text.

and directors are, in the ordinary course, subjected to the deferential business judgment rule,[112] this is just a presumption. Under certain circumstances, it can be rebutted and replaced with the much more stringent entire fairness standard.[113] When that happens, "the defendants must establish to the *court's* satisfaction that the transaction was the product of both fair dealing *and* fair price."[114] The fair price prong[115] of the analysis is "largely equivalent to the fair value determination in an appraisal proceeding,"[116] and courts do indeed rely on DCF valuations.[117] The primary difference between the two is that the purpose of an appraisal is to pick a single number; for entire fairness, "the court's task is . . . to

---

[112] *See* HOLGER SPAMANN, SCOTT HIRST & GABRIEL RAUTERBERG, CORPORATIONS IN 100 PAGES 36–37 (2nd ed. 2021).

[113] As recently explained by the Delaware Chancery Court, entire fairness

> applies to board action where there exists actual conflicts of interest . . . including (1) when a plaintiff pleads facts that call into question the disinterestedness and independence of a sufficient number of directors; (2) when the transaction was effectuated by a controlling or dominating shareholder, and (3) when a plaintiff pleads a fraud-on-the-board theory and the attendant illicit manipulation of a board's deliberative processes by self-interested corporate fiduciaries.

In re Pattern Energy Grp. Inc. S'holders Litig., 2021 WL 1812674, at *31 (Del. Ch. May 6, 2021) (quotation marks omitted).

[114] Reis v. Hazelett Strip-Casting Corp., 28 A.3d 442, 459 (Del. Ch. 2011) (emphasis in original) (quotation marks omitted) (quoting Cinerama, Inc. v. Technicolor, Inc., 663 A.2d 1156, 1163 (Del. 1995)).

[115] While we refer to this as the fair price "prong," the entire fairness analysis is a unitary, rather than a bifurcated, test. *See* Weinberger v. UOP, Inc., 457 A.2d 701, 711 (Del. 1983) ("The concept of fairness has two basic aspects: fair dealing and fair price. . . . However, the test for fairness is not a bifurcated one as between fair dealing and price. All aspects of the issue must be examined as a whole since the question is one of entire fairness.").

[116] Owen v. Cannon, 2015 WL 3819204, at *31 (Del. Ch. June 17, 2015) (determining fair price under the entire fairness standard by reference to the determination of fair value in an appraisal proceeding); *see also* ACP Master, Ltd. v. Sprint Corp., 2017 WL 3421142, at *18 (Del. Ch. July 21, 2017), *aff'd*, 184 A.3d 1291 (Del. 2018) ("The economic inquiry called for by the fair price aspect is the same as the fair value standard under the appraisal statute.").

[117] *See, e.g.*, *In re* S. Peru Copper Corp. S'holder Derivative Litig., 52 A.3d 761, 816–17 (Del. Ch. 2011), *aff'd sub nom*. Ams. Mining Corp. v. Theriault, 51 A.3d 1213 (Del. 2012) (calculating a fair price by balancing three values, the first of which was a DCF); *In re* Dole Food Co., S'holder Litig., 2015 WL 5052214, at *35–37 (Del. Ch. Aug. 27, 2015) (discussing and modifying the DCF valuation replied upon in its entire fairness analysis); *Owen*, 2015 WL 3819204, at *31 (relying on a DCF valuation in an entire fairness analysis); *ACP Master, Ltd.*, 2017 WL 3421142, at *28 (relying, inter alia, on a DCF analysis to determine fair price for the purposes of entire fairness).

determine whether the transaction price falls within a range of fairness."[118]

Courts also rely upon DCF valuations in a wide variety of settings outside of traditional corporate and bankruptcy law. DCF is used to measure damages in contract,[119] international arbitration,[120] tort,[121] and tort-like[122] claims. It is also used to estimate asset values in as disparate of areas as tax[123] and family law.[124] To our knowledge, the Fama-French data have not been widely adopted in any of these contexts, so the noisy factors have not (yet) had an effect on these areas.

Defenders of the Fama-French factors—as well as those who are skeptical of hyperbolic claims in law review articles—might respond that while the noisy factors will benefit one side or the

---

[118] *In re Dole Food Co.*, 2015 WL 5052214, at *33. The focus on a range of fair values attenuates the problem of the noisy factors somewhat. Nevertheless, at least in cases where the value is close to the line, they may still be enough to tip the scales one way or another.

[119] *See, e.g.*, Indeck Energy Servs., Inc. v. Merced Cap., L.P., 2021 WL 5815740, at *2 (N.Y. App. Div. Dec. 7, 2021) (affirming the trial court's decision to accept a DCF approach to valuing damages in a breach of contract claim, and noting that "[m]any authorities recognize that the most reliable method for determining the value of a business is the discounted cash flow method" (alterations and quotation marks omitted) (quoting Lippe v. Bairnco Corp., 288 B.R. 678, 689 (S.D.N.Y. 2003))); Energy Cap. Corp. v. United States, 302 F.3d 1314, 1333 (Fed. Cir. 2002) (endorsing a DCF approach to calculating damages for breach of contract against the United States and noting that the discount rate accounts for both time and risk).

[120] *See, e.g.*, Enron Corp. Ponderosa Assets, L.P v. Argentine Republic, ICSID Case No. ARB/01/3, Award, ⁋ 386 (May 22, 2007) (noting that "the Tribunal is persuaded that the DCF method offers a reliable approach" in awarding damages in the context of international investment dispute).

[121] *See, e.g.*, Proctor Tr. Co. v. Upper Valley Press, Inc., 405 A.2d 1221, 1225 (Vt. 1979) (holding in a fraud action that a DCF valuation was "one of the approved methods" of arriving at the damages resulting from the alleged fraud); N. Am. Title Co. v. Liberty Title Co., 2008 WL 2227244, at *8 (Cal. Super. Ct. Apr. 9, 2008) (holding that a "discounted cash flow method of valuation . . . does appear to represent an accepted method of appraisal for valuation of the business 'pre-tort'").

[122] *See, e.g.*, Elk v. United States, 87 Fed. Cl. 70, 92–93 (2009) (employing a DCF approach to determining tort-like damages).

[123] *See, e.g.*, Gross v. Comm'r, 78 T.C.M. (CCH) 201 (T.C. 1999), *aff'd*, 272 F.3d 333 (6th Cir. 2001) ("We have for many years relied on a discounted cash-flow analysis to determine the present value of one or more future cash-flows."); Est. of Jones v. Comm'r, 118 T.C.M. (CCH) 143 (T.C. 2019) (concluding, for the purpose of valuating limited partnership interests in the gift-tax context, that a DCF method "is more appropriate" than an alternative method).

[124] *See, e.g.*, Adams v. Adams, 945 N.E.2d 844, 869 (Mass. 2011) (holding that "[t]he special master should have elected to employ some variant of the discounted cash flow method" in valuing a husband's partnership interest in a divorce action); Sharp v. Sharp, 449 S.E.2d 39, 44 (N.C. Ct. App. 1994) (affirming the trial court's adoption of the DCF valuation of a residential subdivision in the context of a divorce).

other in any specific valuation, there is no reason to think that they systematically affect the results in any particular way. The effects are, in other words, unbiased. As a result, a defender might argue that they are still valid estimators.

We have three responses to this argument. The first is that, even assuming that the result is unbiased, this alone is not a reason to adopt a method. If it were, we could save a lot of time and expense by just flipping a fair coin: heads the defendant wins; tails the plaintiff wins. This is totally unbiased (after all, that is what it means for a coin to be fair), but it is not a remotely credible means of assessing value. Something more than unbiasedness is surely required. The valuation method must be credible.

A second reason not to use the Fama-French factors, even if they lead to unbiased valuations, is that this is only true ex ante. Ex post, one vintage will always end up yielding results more favorable to one party in any particular instance. Once a change has been made to the factors, it is child's play for an analyst (or expert witness) to try out a valuation using all the available factor vintages. Consequently, with every vintage update comes another opportunity for experts to pick the one that yields results that are most favorable to her client. And since no vintage is any better—in any objective sense—than any other, there is no reason for her not to do so. The fact that litigation typically occurs years after the event in question makes this problem even worse, since she can expect to have many equally valid vintages to choose from. In other words, while the valuations may not be biased ex ante, it would be naïve to expect them to be anything but biased ex post.

Finally, while we know with a fair degree of certainty what the effects of past discretionary changes to the Fama-French factors have been, we have no way of predicting how they might change in the future. The nature of discretionary changes is just that: they are discretionary and not the result of some algorithm or rule. This is compounded by the fact that Fama and French have shown no interest in publicly releasing the code that generates the factors or in making any firm commitments about future changes. This is fair enough—after all, it is their data, and they can do what they wish with it. This is especially true since the users of the data are not the ones paying for it.[125] But without any

[125] Notwithstanding the fact that the data is, of course, theirs to do with as they please, the fact that they have continued to keep their code private is contrary to the emerging scholarly consensus in favor of data and code sharing, which is thought to facilitate replication

assurances about what future discretionary changes will or will not be made, any use of the Fama-French factors in the future is implicitly relying on *them* and whatever judgment *they* exercise (and, depending on the details of the production process, perhaps also that of the DFA employees who produce the factors). Having access to the code would at least allow for an easy way to audit any discretionary changes that were made. Without that, it is hard to see how a judge could assess the credibility of any analysis that relied on the data without assessing the credibility of the people creating it.

* * *

It goes without saying that valuations—and the potentially multimillion-dollar judgments that go along with them—should not depend on something as arbitrary as the date on which an expert downloaded some data. Nor should it depend on discretionary changes to an algorithm implemented deep in the bowels of a large asset manager that has nothing to do with the case at issue. Obviously then, experts should stop using these data, and judges should not accept as credible any analysis that relies on them. Perversely, the more sophisticated and rigorous—at least from the perspective of the finance literature—the valuation technique, the more likely it is to be affected by the retroactive changes to the Fama-French data. This is a manifestation of the law of conservation of judgment: what appears on the surface to be a more objective, scientific technique is also one where it is harder to pinpoint the locus of judgment and discretionary decision-making.

III. PERFORMANCE ANALYSIS WITH NOISY FACTORS

A second context in which noisy factors have legal consequences is performance evaluation. Here, our focus is primarily on what the noisy factors mean for a variety of different fiduciaries, each of which, we presume, is seeking to discharge her duties with loyalty and diligence. While it is possible that the noisy factors could expose them to potential liability, a bigger problem is

and improve the credibility of academic research. Scholarly journals in a variety of fields have data and code sharing policies, including some of the most prominent publications in finance and economics. *See, e.g.*, *Journal of Finance Data and Code Sharing Policy*, AM. FIN. ASS'N, https://perma.cc/Y5CU-Y45B; *Data and Code Sharing Policy*, J. FIN. ECON., https://perma.cc/FT2L-YHSX; *Data and Code Availability Policy*, AM. ECON. ASS'N, https://perma.cc/9RZX-SQ3V.

that a standard tool relied upon by these fiduciaries—who, collectively, are responsible for safeguarding trillions of dollars in assets—yields conflicting and contradictory results.

Just as DCF is the textbook approach to valuation, the standard way to evaluate performance is to use a factor model, and in particular a model using the Fama-French factors.[126] This is especially true with respect to mutual funds and other investment funds.[127] Whereas it was changes in beta estimates that affected valuation in Part II, here the effects are driven by changes in alphas. Since we know that the noisy factors affect alpha estimates for both mutual funds and individual stocks,[128] it is obvious that they will also affect analyses that rely on those alphas.

As with our discussion of valuation, we begin with an illustrative example. Here, we consider a hypothetical investment adviser analyzing mutual funds in order to advise her client. As this illustration makes clear, the noisy factors can dramatically affect the results of a standard performance analysis. This is a very real concern. Within a few months of when we first made our research public, an expert report to the Norwegian Ministry of Finance on the Norwegian Government Pension Fund Global referenced the noisy factors and explained that they create difficulties for evaluating the performance of the fund.[129] If the noisy factors are a problem for the world's largest sovereign wealth fund,[130] they might also be a problem for other fiduciaries with fewer resources at their disposal.

---

126 *See, e.g.*, Wayne E. Ferson, *Investment Performance Evaluation*, 2 ANN. REV. FIN. ECON. 207, 209 (2010) (noting that "[t]he most famous performance measure is alpha"). Professor Wayne Ferson went on to note that "[h]undreds of papers provide evidence about alphas," *id.* at 212, and that the approach using and building on the Fama-French model "is reflected prominently in academic studies," *id.* at 211. We note, of course, that this does not mean that the Fama-French model is the only approach that is used.

127 *See, e.g.*, RICHARD A. BREALEY, STEWART C. MYERS & FRANKLIN ALLEN, PRINCIPLES OF CORPORATE FINANCE 210 (12th ed. 2017) ("The Fama-French model finds its widest use as a way of measuring the performance of mutual funds, pension funds[,] and other professionally managed portfolios."); KOLLER ET AL., *supra* note 91, at 281 ("Given the strength of Fama and French's empirical results, the academic community now measures risk with a model commonly known as the Fama-French three-factor model."); ROSS ET AL., *supra* note 102, at 391–92 (providing two exercises in which the reader is instructed to download the Fama-French data from Professor French's website and use it for mutual fund performance analysis).

128 *See supra* notes 66–72 and accompanying text.

129 *See* BAUER ET AL., *supra* note 22, at 33.

130 *The World's Biggest Sovereign Wealth Funds—In One Chart*, WORLD ECON. F. (Feb. 12, 2021), https://www.weforum.org/stories/2021/02/biggest-sovereign-wealth-funds-world-norway-china-money/. In 2021, the fund managed over $1.3 trillion. *The Fund*, NORGES BANK INVESTMENT MANAGEMENT, https://perma.cc/76D9-7CXZ.

A. Mutual Fund Performance Analysis

Millions of people in the United States rely on investment advisers to help them make financial decisions. Under the Investment Advisers Act of 1940[131] (Advisers Act), these advisers are fiduciaries.[132] To keep things simple, suppose a client comes in asking for help selecting between actively managed mutual funds.[133] Because we want to see how the results of an adviser's analysis might differ using a later factor vintage, let us suppose that the conversation occurred in July 2012. To avoid cherry-picking, let us further suppose that the funds under consideration are the five largest (in terms of total assets under management) actively managed domestic equity mutual funds as of the end of 2011.

Predicting mutual fund performance is notoriously difficult. While there was considerable evidence of persistence in mutual fund performance in the past,[134] more recent evidence suggests that good performance in the past does not, on average, predict good performance in the future.[135] At the same time, there are reasons to

---

[131] 15 U.S.C. § 80b-1 et seq.

[132] The statutory basis for these fiduciary duties is rooted in § 206 of the Advisers Act. *See* Transamerica Mortg. Advisors, Inc. v. Lewis, 444 U.S. 11, 17 (1979) ("[Section] 206 establishes federal fiduciary standards to govern the conduct of investment advisers." (quotation marks omitted) (quoting Santa Fe Indus., Inc. v. Green, 430 U.S. 462, 471 n.11 (1963))). Courts have interpreted the statute as reflecting common law equitable principles. *See* Sec. & Exch. Comm'n v. Cap. Gains Rsch. Bureau, Inc., 375 U.S. 180, 191–94 (1963) (discussing the history of the Advisers Act and observing that Congress recognized investment advisers to be fiduciaries).

[133] It is well established that the typical actively managed mutual fund underperforms the market. *See, e.g.*, Diane Del Guercio & Jonathan Reuter, *Mutual Fund Performance and the Incentive to Generate Alpha*, 64 J. FIN. 1673, 1673 (2014) (noting that "[t]he typical actively managed U.S. equity fund earns a negative after-fee alpha" and describing this underperformance as "well-documented"). Notwithstanding this, they remain an important part of the market. As of year-end 2020, the proportion of U.S. equities held by actively managed domestic equity mutual funds and ETFs (14%) was about the same as the proportion held by index domestic equity mutual funds and ETFs (also 14%). INV. CO. INST., 2021 INVESTMENT COMPANY FACT BOOK fig. 2.9 (2021).

[134] *See, e.g.*, Mark M. Carhart, *On Persistence in Mutual Fund Performance*, 52 J. FIN. 57, 57 (1997) (describing mutual fund persistence as "well documented in the finance literature").

[135] *See* James J. Choi & Kevin Zhao, *Carhart (1997) Mutual Fund Performance Persistence Disappears Out of Sample*, 10 CRITICAL FIN. REV. 263, 264–66 (2021) (showing that the persistence documented by Carhart is absent in a more recent time period); *see also* BERLINDA LIU & GAURAV SINHA, U.S. PERSISTENCE SCORECARD, MID-YEAR 2021, at 1 (2021) (demonstrating that only 4.8% of the actively managed domestic mutual funds in the top quartile of performance in June 2019 remained there two years later).

stay away from poorly performing funds.[136] These options are likely to be dominated by an ultra-low-cost broad-based index fund.[137]

The most standard approach to evaluating mutual fund performance is to use a factor model to estimate the fund's alpha.[138] Given this, our hypothetical adviser might start by downloading the Fama-French data and using them to estimate a three-factor model for each of the funds under consideration. Each fund's alpha captures its risk-adjusted performance, so a higher alpha is better, and a fund with a positive alpha beat the market. To keep things simple, let us suppose that the adviser performs her analysis on each fund in each of the last five years (2007–2011) and then calculates the average alpha for each of the funds.[139] Having done so, she would have found that Fund *A* had the highest (a modestly positive) alpha, followed by Fund *B* (at about zero) and *C* (modestly negative). Fund *D* was considerably behind, with an estimated alpha of about −2%, and Fund *E* was the real laggard, at less than −3%.

Let us suppose that, on the basis of this advice, and even with the caveat that past performance is not a guarantee of future returns, the client chooses to invest in Fund *A*. Suppose further that five years later, in 2017, the client decides to reevaluate his portfolio and seeks out a second adviser. After he explains to his second adviser why he chose Fund *A*, she decides to repeat the analysis that the first adviser performed back in 2012. Like the first adviser, she starts by downloading the Fama-French data. She then

---

[136] For example, a fund that is underperforming because of high fees is likely to continue to charge high fees and, therefore, continue to provide poor net-of-fee returns. Similarly, a fund with high turnover—which tends to depress returns—is likely, absent a sharp change in management style, to continue to have high turnover. *See e.g.*, Claudia Champagne, Aymen Karoui & Saurin Patel, *Portfolio Turnover Activity and Mutual Fund Performance*, 44 MANAGERIAL FIN. 326, 331 (2018) (arguing that turnover is negatively related to returns). The same goes for a highly concentrated fund: since underdiversification tends to reduce risk-adjusted returns, absent a sharp change in style, such a fund is likely to continue to perform poorly on a risk-adjusted basis.

[137] *But see* Pat Akey, Adriana Z. Robertson & Mikhail Simutin, *Closet Active Management of Passive Funds* 11–21 (Rotman Sch. Mgmt., Working Paper No. 3874582, 2021) (showing that about a third of U.S. index funds and ETFs are more active than the median actively managed fund and that more active index funds and ETFs have lower performance).

[138] Betas are also useful to evaluate whether a fund's exposure matches the strategy it presents to investors. *See supra* notes 33, 35–36, and accompanying text.

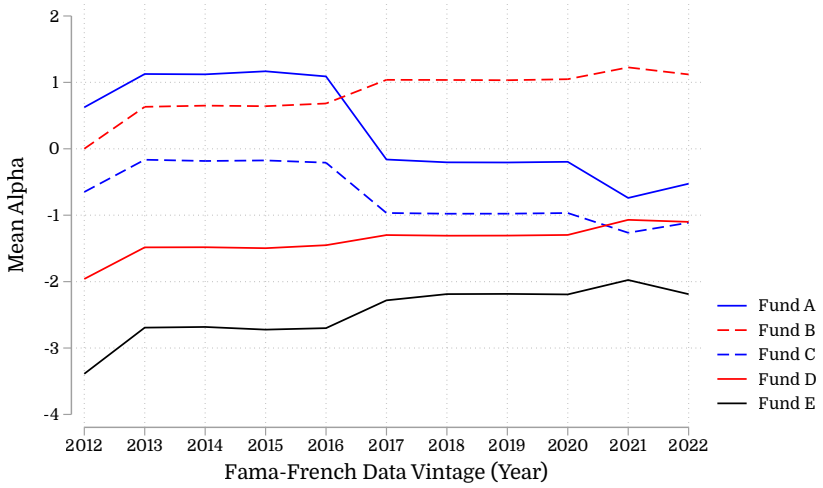
[139] As discussed above, it is common to compute one-year alphas in the mutual fund space. *See, e.g.*, Mikhail Simutin, *Cash Holdings and Mutual Fund Performance*, 18 REV. FIN. 1425, 1431 n.4 (2014) (using twelve-month returns). In untabulated results, we find that if she had instead computed five-year alphas, the adviser would have gotten a similar, albeit more attenuated, pattern.

performs *exactly the same analysis* as the first adviser, estimating the performance of each of the five funds in the 2007–2011 period. After doing so, she scratches her head: far from being the best performer, Fund *A* now seems to have considerably *underperformed* Fund *B*. Fund *A*'s alpha, she explains to the increasingly agitated client, was slightly *negative* during the five-year period leading up to his initial investment, and it was Fund *B* that had a modestly positive alpha. The second adviser shrugs and suggests that perhaps the first adviser made a mistake. The client is understandably upset.

Perhaps the investor also reevaluated his portfolio at the nine-year mark, in 2021. At that point, after performing independent analysis (still looking at the performance of the same five funds between 2007 and 2011), a third adviser's news would have been even more upsetting to her client. Far from being around zero, her estimate of Fund *B*'s alpha would be substantially greater than 1%. Fund *A*, meanwhile, would seem to have substantially underperformed, with an alpha of about −0.75%. And there would also be changes further down the list: Fund *C* did not substantially outperform Fund *D* at all; it was the other way around, although at least she agrees that they both had negative alphas. Understandably, the client concludes that his first adviser was utterly incompetent.

Of course, all three advisers were equally competent, since all performed exactly the same analysis. The only difference between the inputs in 2012, 2017, and 2021 was the date on which the Fama-French data were downloaded. Figure 1 summarizes the results of this analysis using each of the available factor vintages. A few features stand out from this figure. First, it demonstrates that retroactive changes can have a large effect on estimated alphas. These effects, moreover, differ substantially across funds. While some (Funds *B*, *D*, and *E*) seem to improve fairly consistently across vintages, the estimated performance of Funds *A* and *C* improves between the 2012 and 2014 vintages before deteriorating. Again, we stress that the only thing changing in Figure 1 is the date when the Fama-French data were downloaded; everything else, including the sample period, is identical.

FIGURE 1: FUND PERFORMANCE, 2007–2011
VARYING ONLY FAMA-FRENCH DATA VINTAGE



B. Fiduciaries and Performance Analysis

The effect of the noisy factors on performance analysis puts fiduciaries in a difficult position. Sticking with the example in Part III.A, the adviser genuinely cannot tell whether Fund *A* or Fund *B* performed better, or whether Fund *A* outperformed or underperformed the market during the sample period. The conclusions she would reach using different data vintages are contradictory, and there is no way for her to judge which vintage provides more accurate results. Nor, of course, would it have helped much if she had simply performed her analysis at a single point in time, using a single vintage. While this would perhaps avoid the confusion of *seeing* the contradictory results, it would not change the fact that her conclusion would, in fact, be the result of which vintage she happened to use.

It is unlikely, but not impossible, that the events described in the preceding Section could give rise to liability. The private right of action under the Advisers Act is limited,[140] and there is very little case law on adviser recommendations that implicate only the duty of care.[141] More to the point, like any standard of care,

---

[140] *See Transamerica*, 444 U.S. at 24.

[141] At least one federal court has suggested that egregious failures to competently investigate before providing investment advice can be grounds for liability under § 206(2) of the Advisers Act. *See* Sec. & Exch. Comm'n v. Duncan, 2021 WL 4197386, at *15

the standard must be applied prospectively, not retrospectively. The first adviser would have had no reason to suspect that the factors might change, making it nonsensical to say that she had failed to act with due care. And since we find no evidence that the factors are getting better, there is no reason to think that the analysis by the second adviser is any more accurate anyway, making this notion even more preposterous.

That preposterousness is precisely the point. All three advisers in the illustration performed exactly the same analysis, and they exhibited exactly the same level of care. And yet, by the time 2017—or 2021—rolled around, the first adviser might not be able to explain why her analysis differed from the contemporaneous one. As a result, in the unlikely event that the irate customer managed to persuade the Securities Exchange Commission (SEC) to take action against her, the adviser might find it hard to defend herself, particularly if she no longer had a copy of the original code and data she used to perform the analysis.[142] The reason for this, of course, is the noisy factors.

Quite apart from the (low but not zero) risk of liability is the fact that our adviser simply *does not know* how well the five funds performed, either in isolation or relative to each other. And to the extent that she thinks she does (perhaps because she only performed the analysis at a single point in time and is not aware of the noisy factors), her confidence—however genuine and (until now) reasonable—would be misplaced. The unfortunate reality is that the noisy factors create the most problems for the very fiduciaries that were the most diligent, relying on the best practices supported by academic experts.

There are a variety of other contexts in which some person or group, acting in a fiduciary capacity, is required to evaluate the performance of a portfolio or company. While the law typically would not require them to do so using a factor model, a prudent

(D. Mass. Sept. 15, 2021) (discussing a series of failures and holding that "[t]he SEC sustained its burden to prove that Defendant was negligent by failing to employ reasonable care to avoid misleading his clients," which constituted "negligence under Section 206(2)" (quotation marks omitted) (quoting Sec. & Exch. Comm'n v. Cap. Gains Rsch. Bureau, Inc., 375 U.S. 180, 194 (1963))).

142 While not technically a fiduciary, a broker would be in a similar position. Under the SEC's Regulation Best Interest, brokers must act in the best interest of their clients when making a recommendation. *See* 17 C.F.R. § 240.15*l*-1(a)(1) (2019). This includes "exercis[ing] reasonable diligence, care, and skill to . . . [u]nderstand the potential risks, rewards, and costs associated with the recommendation," and to "[h]ave a reasonable basis to believe that the recommendation is in the best interest of a particular retail customer." *Id.* § 240.15*l*-1(a)(2)(ii)(A)–(B).

fiduciary might well seek to employ a textbook technique in discharging her duties.

For example, a wide variety of retirement plans, including both defined contribution (such as 401(k) and 403(b)) and defined benefit plans, are governed by the Employee Retirement Income Security Act of 1974[143] (ERISA). Under ERISA, retirement fund trustees, including 401(k) plan managers, owe fiduciary duties to plan participants and beneficiaries.[144] These duties are "derived from the common law of trusts," and "[i]n determining the contours of an ERISA fiduciary's duty, courts often must look to the law of trusts."[145] These duties include a duty of prudence and, in the context of 401(k) plans, an ongoing duty to monitor the investment options in the menu.[146]

It stands to reason that standard performance-analysis techniques might be helpful for a fiduciary seeking to prudently oversee the portfolio of a defined benefit plan, or to monitor the performance of the options available in a defined contribution plan menu. After all, it is hard to know that you are prudently managing a portfolio or to evaluate the options on a menu if you do not know how they are doing. A prudent fiduciary might therefore want to employ a factor model to measure risk-adjusted performance.[147]

This notion is consistent with two recent academic articles. In the first, Professors Ian Ayres and Quinn Curtis addressed the issue of dominated funds in 401(k) plans—funds that are almost certainly worse investment opportunities than at least one other option on the menu.[148] As the authors pointed out, the current

---

143 Pub. L. No. 93-406, 88 Stat. 829 (codified as amended in scattered sections of 29 U.S.C.).

144 *See* 29 U.S.C. § 1002(21)(A)(i); *see also id.* § 1104(a)(1).

145 Tibble v. Edison Int'l, 575 U.S. 523, 528–29 (2015).

146 *See id.* at 530 ("[A] fiduciary normally has a continuing duty of some kind to monitor investments and remove imprudent ones.").

147 *See* Max M. Schanzenbach & Robert H. Sitkoff, *Reconciling Fiduciary Duty and Social Conscience: The Law and Economics of ESG Investing by a Trustee*, 72 STAN. L. REV. 381, 427 (2020) (explaining that the duty of prudence permits trustees to select an investment opportunity "provided that the investment fits within a diversified overall investment strategy with portfolio-level risk-return objectives reasonably suited to the trust").

148 *See* Ian Ayres & Quinn Curtis, *Beyond Diversification: The Pervasive Problem of Excessive Fees and "Dominated Funds" in 401(k) Plans*, 124 YALE L.J. 1476, 1481 (2015) ("On average, 401(k) menus in our sample provide investors sufficient options to diversify, but investors in many plans bear costs well in excess of retail index funds—and these costs are unlikely to be fully mitigated by returns.").

legal and regulatory regime is ill-equipped to handle this problem.[149] While Professors Ayres and Curtis were primarily focused on high-fee funds, the concern about fees is founded on the impact that they have on net-of-fee investment performance.[150] Accordingly, risk-adjusted net-of-fee performance is a reasonable metric for a fiduciary to consider (perhaps in addition to fees alone) in evaluating funds. One could easily imagine a plan sponsor, persuaded that dominated funds should be removed from a plan, relying on a factor model as part of her analysis.

The second article explicitly relies on a factor model. In it, Professor Ayres, this time with Professor Edward Fox, argued that fiduciaries should explicitly consider an investment option's alpha—i.e., its return in excess of a factor model—before recommending or investing in that option, rather than sticking with a low-cost, broadly diversified mutual fund or ETF.[151] They argued that these "alpha duties," as they term them, are consistent with current fiduciary law.[152] Naturally, in order to evaluate an investment's alpha, the fiduciary would first have to calculate it (or, more likely, ask an adviser or consultant to calculate it). This requires a factor model. And while Professors Ayres and Fox did not use the Fama-French factors for their primary analysis, they noted that their proposed approach can "easily be generalized" to other factor models, including the Fama-French model.[153]

The directors of mutual funds and other investment companies regulated under the Investment Company Act of 1940[154] ('40 Act funds) are in a somewhat similar position. They too have fiduciary duties, including a statutory obligation to review the fund's advisory contract annually.[155] While not obligated to, a diligent trustee seeking to use the best available means to evaluate the current adviser's performance might well use the Fama-French model to do so. The fact that this textbook approach leads

---

149 *See id.* at 1507–08 (observing that the themes underpinning case law interpreting ERISA's § 404(c) safe harbor provision, 29 U.S.C. § 1104(c), are a "poor fit for the realities of investor choice").

150 *See id.* at 1481 ("We show that the primary problem for investors in 401(k) plans is not loss due to lack of diversification, but loss due to excessive fees.").

151 *See* Ian Ayres & Edward Fox, *Alpha Duties: The Search for Excess Returns and Appropriate Fiduciary Duties*, 97 TEX. L. REV. 445, 450 (2019).

152 *See id.* at 496–97 (aligning "alpha duties" with the Restatement (Third) of Trusts's approach to active investment).

153 *Id.* at 464.

154 15 U.S.C. § 80a-1 et seq.

155 *Id.* § 80a-15(c).

to highly inconsistent results should be troubling not just to trustees seeking to discharge their duties, but also to the beneficiaries of those duties—namely, mutual fund investors.

The investment advisers to '40 Act funds are themselves subject to a statutory fiduciary duty,[156] which in their case prohibits them from charging excessive fees.[157] The standard for liability, developed in *Gartenberg v. Merrill Lynch Asset Management, Inc.*[158] and now known as the *Gartenberg* standard, is a multifactor analysis that includes consideration of "the nature and quality of the service" provided by the adviser.[159] In addition to forming the basis for potential liability for advisers,[160] this obligation feeds back to the fund directors. In an express nod to the *Gartenberg* standard,[161] the SEC requires fund directors to disclose "factors relating to both the board's selection of the investment adviser, and its approval of the advisory fee and any other amounts to be paid under the advisory contract," including a discussion of "the investment performance of the fund and the investment adviser."[162] That performance needs to be evaluated somehow.

Even more generally, factor models are used to describe the performance of mutual funds. Since the noisy factors mean that the results change, they pose a problem for evaluating the truthfulness of a mutual fund's disclosure. Perhaps even more problematic is the fact that it is hard to interpret a performance metric that changes dramatically for reasons that the experts cannot explain.[163]

---

156 *Id.* § 80a-35(b).

157 *See id.* § 80a-36(b). For a more detailed discussion of fee liability under § 36(b), see generally Quinn Curtis & John Morley, *An Empirical Study of Mutual Fund Excessive Fee Litigation: Do the Merits Matter?*, 130 J.L. & ECON. 275 (2014).

158 694 F.2d 923 (2d Cir. 1982).

159 *Id.* at 930. The Supreme Court has since endorsed the *Gartenberg* standard. Jones v. Harris Assocs. L.P., 559 U.S. 335, 336 (2010) ("The *Gartenberg* standard . . . accurately reflects the compromise that is embodied in § 36(b), and it has provided a workable standard for nearly three decades.").

160 *But see generally* John Morley & Quinn Curtis, *Taking Exit Rights Seriously: Why Governance and Fee Litigation Don't Work in Mutual Funds*, 120 YALE L.J. 84 (2010) (discussing the structural and institutional problems with mutual fund fee litigation).

161 Disclosure Regarding Approval of Investment Advisory Contracts by Directors of Investment Companies, 69 Fed. Reg. 39,798, 39,801 n.31 (June 30, 2004) (noting that "[c]ourts have used similar factors in determining whether investment advisers have met their fiduciary obligations under Section 36(b) of the Investment Company Act" and citing *Gartenberg*).

162 *Id.* at 39,801.

163 Stepping briefly outside the fiduciary context, third-party analysis also relies on performance analysis to evaluate mutual funds.

Performance analysis also arises outside of the mutual fund context. Consider, for example, the case of executive compensation, where a member of the board's compensation committee might want to evaluate the CEO's performance. While imperfect, a company's stock performance is routinely used to assess how well the company is doing. There are many ways to evaluate a stock's performance, but as we know, one textbook approach is to look at its alpha during the relevant sample period.[164] To the extent that the firm's performance is attributable to the firm's managers, this also gives a measure of the manager's performance.

* * *

To be clear, our claim is not that fiduciaries are currently required to use the Fama-French data in all, or even any, of the examples discussed in this Part. Rather, our point is simply that at least some of them almost certainly *do*, and the current expert consensus is that they probably *should*. Far from acting wrongfully, those that have done so in the past were simply following the best expert advice available to them. Yet the fact that doing so can leave them with wildly varying results—for reasons that the experts themselves cannot explain—is deeply troubling. Given what we know now, it is hard to see how a fiduciary could rely on data that yield substantially different results because of something as arbitrary as when she accessed the data.

IV. EVENT STUDIES WITH NOISY FACTORS

A final context in which we consider the legal consequences of the noisy factors is event studies. In addition to their extensive use in litigation, event studies are a mainstay of legal scholarship. A unique feature of the event study context is that, in addition to being used to measure quantities, event studies are commonly used for binary categorizations. For example, an expert might present an event study to answer the question, "Did a stock price fall after a misstatement was corrected?" Or a scholar might rely on the technique to answer the question, "Do mandatory disclosure rules benefit investors?" While the litigation context generally focuses on the price reaction of a single firm on a single day, academic studies

---

[164] *See supra* notes 126–27 and accompanying text.

often look across a large number of firms to establish a more general relationship.

We begin this Part with an illustrative example of the latter type of analysis, drawn from a well-established area of research in corporate governance: the impact of activist hedge funds. We then discuss a wide variety of contexts in which event studies are relied upon by courts and parties, and by legal scholars, respectively.

A. Hedge Fund Activism

A classic topic in corporate governance is the impact of activist investors on shareholder value.[165] One branch of this literature has studied the impact of hedge funds on target firms. While this remains an active area of research, one robust result is that, around the time an activist hedge fund announces that it is planning to target a particular firm, that target firm's share price tends to jump.[166] The first thing we note is that this result is extremely robust to changing factor vintages, and nothing in our analysis casts doubt on this finding. Rather, we use other features of the data to illustrate the impact the noisy factors can have on event study analyses.

Not all activism events are hostile. For example, an activist fund might simply make an investment or submit a shareholder proposal. Other events, such as launching a proxy contest or a takeover bid, are clearly hostile.[167] Suppose that we are

---

[165] For a sampling of articles on this topic published within the last fifteen years, see generally Marcel Kahan & Edward B. Rock, *Hedge Funds in Corporate Governance and Corporate Control*, 155 U. PA. L. REV. 1021 (2007); Alon Brav, Wei Jiang, Frank Partnoy & Randall Thomas, *Hedge Fund Activism, Corporate Governance, and Firm Performance*, 63 J. FIN. 1729 (2008) [hereinafter Brav et al., *Activism, Governance, and Performance*]; Robin Greenwood & Michael Schor, *Investor Activism and Takeovers*, 92 J. FIN. ECON. 362 (2009); Alon Brav, Wei Jiang & Hyunseob Kim, *The Real Effects of Hedge Fund Activism: Productivity, Asset Allocation, and Labor Outcomes*, 28 REV. FIN. STUD. 2723 (2010); Lucian A. Bebchuk, Alon Brav & Wei Jiang, *The Long-Term Effects of Hedge Fund Activism*, 115 COLUM. L. REV. 1085 (2015); John C. Coffee Jr. & Darius Palia, *The Wolf at the Door: The Impact of Hedge Fund Activism on Corporate Governance*, 41 J. CORP. L. 545 (2016); Alon Brav, Wei Jiang, Song Ma & Xuan Tian, *How Does Hedge Fund Activism Reshape Corporate Innovation?*, 130 J. FIN. ECON. 237 (2018).

[166] *See, e.g.*, Brav et al., *Activism, Governance, and Performance*, *supra* note 165, at 1730 (finding a 7%–8% abnormal return around the announcement of activism).

[167] The hedge fund activism data we use, produced by Professor Alon Brav, Professor Wei Jiang, and economist Hyunseob Kim, and collaborators, codes an event as hostile if it involved a (1) proxy contest, (2) lawsuit, (3) takeover bid, (4) threat of a lawsuit or proxy

interested in knowing whether, on average, the market reacts differently to hostile events compared to events that are not. To answer this, we obtain the updated version of the activist hedge fund database maintained by Professors Alon Brav, Wei Jiang, and John Barry, and their collaborators.[168] Using standard techniques, we estimate the abnormal returns of the target firms around the time that the activism event was announced.[169] We then ask whether the abnormal returns around hostile events differ from those around events that are not.[170]

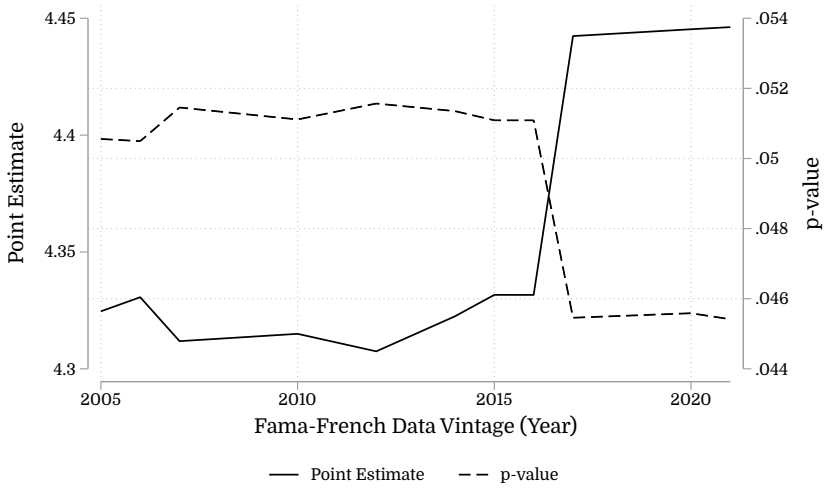
fight, or (5) proposal or public letter indicating hostile intentions or language, such as asking the management to resign. *See generally* Alon Brav, Wei Jiang & Hyunseob Kim, *Hedge Fund Activism: A Review*, 4 FOUNDS. & TRENDS FIN. 185 (2010) [hereinafter Brav et al., *Hedge Fund Activism*].

168 The data are an updated sample using the same data collection procedure as in Brav et al., *Activism, Governance, and Performance*, *supra* note 165, at 1737–38, and Brav et al., *Hedge Fund Activism*, *supra* note 167. More information is available on Professor Jiang's website. *Hedge Fund Activism*, WEI JIANG, https://perma.cc/UZW4-8J79.

169 We consider all events from the beginning of the sample period (1994) through the end of 2004. This allows us to perform an identical analysis using each of the factor vintages available to us. This leaves a total of 1,281 events affecting 1,081 distinct target firms during that eleven-year period. We estimate the betas for each event with a three-factor model using daily stock return and factor data. We include data from within 365 days before the activism was first announced, omitting the 30 days immediately before the event. To remain in the sample, the firm must have at least 100 days of data. We use these betas to compute abnormal returns in the two weeks leading up to the announcement through the end of the week following it (i.e., beginning 13 days before the announcement through the 7 days that follow it). If this yields 15 trading days (three weeks), we use this period to compute the cumulative abnormal return. If this yields only 14 trading days (because the market is not open every day), but abnormal returns are available for the eighth day after the announcement, we add this to the 14 other available days. Otherwise, the event is excluded from the sample.

170 Because we are interested in abnormal returns around the date of the first announcement, we use the variable that indicates that the event started as hostile.

FIGURE 2: DO HOSTILE EVENTS HAVE HIGHER ABNORMAL RETURNS?
VARYING ONLY FAMA-FRENCH DATA VINTAGE



It turns out that the answer to that question, at least according to conventional techniques, depends on when we downloaded the Fama-French data. Figure 2 summarizes the results of this analysis. As in the other illustrations, the only thing we vary is the data vintage. In all cases, we keep the sample period, the computer code, and all the other data identical. We plot the point estimates—which capture the average difference in the share price reaction between hostile and nonhostile events—and the associated *p*-values in the solid and dashed lines, respectively.[171] As with Figure 1, the x-axis shows the factor vintage in question.

Using the pre-2017 vintages, the *p*-values are above 0.05, meaning that each of these point estimates would be described as insignificant at conventional statistical levels. Beginning with the 2017 vintages, however, the *p*-values fall comfortably below 0.05, so the point estimates would be considered statistically significant. Applying the standard rule of thumb in empirical work, a researcher relying on a pre-2017 vintage would conclude that whether an activism event is hostile makes no difference to the share price reaction. In contrast, a researcher relying on a

[171] In other words, we estimate a regression where the dependent variable is the cumulative abnormal return of the target firm, and the independent variable is an indicator variable equal to one if the event was hostile. We estimate the regression using heteroskedasticity-robust standard errors.

post-2017 vintage would conclude that there *is* a difference: the abnormal returns around hostile activist events are larger than the returns around nonhostile events.

One obvious takeaway from this analysis is the well-known fact that using a *p*-value cutoff of 0.05—or 5%—is arbitrary and problematic. Many scholars in various fields have been making this point for years,[172] and we certainly agree. We note, however, that despite these arguments, *p*-value cutoffs continue to be widely used, including by courts in the securities litigation context.[173]

It is clear that the lines in Figure 2 bounce around throughout the figure, although the size of the jumps varies substantially. It is also worth pointing out that while the largest jump in Figure 2 is between the 2016 and 2017 vintages, the Dell analysis in Part II precedes either of these vintages. Accordingly, that result cannot be attributed to there being something special about those two vintages.

B. Securities Litigation and Event Studies

Securities fraud is the most obvious real-world legal context in which the effect of the noisy factors on event studies matters. Scholars were a driving force behind the adoption of event studies by the courts. Not only did they develop the underlying theories and techniques; they also advocated for the adoption of modern financial theory and techniques.[174] Some also serve as expert witnesses.

Scholars have been calling attention to the importance of event studies in securities fraud for decades, and there is broad acceptance of the fact that event studies are critical for establishing three of the six elements of a securities fraud claim under SEC

---

[172] For an accessible discussion of this problem, see generally Andrew Gelman & Eric Loken, *The Statistical Crisis in Science*, 102 AM. SCIENTIST 460 (2014).

[173] *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."); *see also* Jill E. Fisch & Jonah B. Gelbach, *Power and Statistical Significance in Securities Fraud Litigation*, 11 HARV. BUS. L. REV. 55, 62–63 (2021) (pointing out that courts use event studies both in the application of the *Daubert* standard and in assessing legal sufficiency).

[174] *See, e.g.*, Daniel R. Fischel, *Use of Modern Finance Theory in Securities Fraud Cases Involving Actively Traded Securities*, 38 BUS. LAW. 1, 17–19 (1982) (arguing that the market model of stock returns should be used as a basis for determining liability and damages in securities fraud cases under the SEC's Rule 10b-5).

Rule 10b-5,[175] as well as in assessing damages.[176] Because of this, scholars have described event studies as "critical,"[177] an "essential element of a securities fraud claim,"[178] "so entrenched in securities

---

[175] *See* 17 C.F.R. § 240.10b-5 (2025). The six elements of a Rule 10b-5 claim with respect to a publicly traded security are: (1) a material misrepresentation or omission, (2) scienter, (3) a connection with the purchase or sale of a security, (4) reliance, (5) economic loss, and (6) a causal connection between the material misrepresentation and the loss. *See* Dura Pharm., Inc. v. Broudo, 544 U.S. 336, 341–42 (2005). Event studies are used to establish the first, fourth, and sixth elements. *See infra* note 176.

[176] *See, e.g.*, Sanjai Bhagat & Roberta Romano, *Event Studies and the Law: Part II: Empirical Studies of Corporate Law*, 4 AM. L. & ECON. REV. 380, 398 (2002) [hereinafter Bhagat & Romano, *Empirical Studies*] ("The doctrine makes plain that event studies have a dual role in securities litigation. They can be critical for determining both liability and damages."); Michael J. Kaufman & John M. Wunderlich, *Regressing: The Troubling Dispositive Role of Event Studies in Securities Fraud Litigation*, 15 STAN. J.L. BUS. & FIN. 185, 187 (2009) (arguing that "a properly conducted event study is not just a helpful way to present evidence of essential elements of a securities fraud action, it has become a substantive and essential element of a securities fraud claim itself"); Jonah B. Gelbach, Eric Helland & Jonathan Klick, *Valid Inference in Single-Firm, Single-Event Studies*, 15 AM. L. & ECON. REV. 495, 496 (2013):

> [S]ingle-firm event studies . . . are especially important in the context of securities litigation. A plaintiff alleging securities fraud under SEC Rule 10b-5 must establish six basic elements . . . . Event studies can be used to address directly the materiality and loss causation elements. Additionally, financial economics is highly relevant to establishing reliance, which can be tightly linked to the appropriateness of using event studies to address materiality and loss causation.

See also Alon Brav & J.B. Heaton, *Event Studies in Securities Litigation: Low Power, Confounding Effects, and Bias*, 93 WASH. U. L. REV. 583, 585 (2015) ("After the Supreme Court endorsed the fraud-on-the-market doctrine in *Basic Inc. v. Levinson* in 1988, event studies became so entrenched in securities litigation that they are viewed as necessary in every case."); Jill E. Fisch, Jonah B. Gelbach & Jonathan Klick, *The Logic and Limits of Event Studies in Securities Fraud Litigation*, 96 TEX. L. REV. 553, 556 (2018) ("Use of the event study methodology has become ubiquitous in securities fraud litigation. 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."); Fisch & Gelbach, *supra* note 173, at 56–57:

> In securities fraud cases, event studies are used in several ways, including analyzing the efficiency of the market in which the securities trade, measuring the price impact of the fraudulent disclosures, determining whether there is a causal relationship between the fraud and the plaintiffs' economic losses, and computing the amount of damages. Although courts vary in the extent to which they require the use of an event study and the degree to which they accept other evidence with respect to these issues, a properly conducted event study is often a critical factor.

[177] Bhagat & Romano, *Empirical Studies*, *supra* note 176, at 398; Fisch & Gelbach, *supra* note 173, at 57.

[178] Kaufman & Wunderlich, *supra* note 176, at 187.

litigation that they are viewed as necessary in every case,"[179] and "preferred or even required."[180]

Courts also rely on evidence from event studies at two crucial procedural stages in securities fraud cases: motions for summary judgment and motions for class certification.[181] And, as Professors Jill Fisch and Jonah Gelbach pointed out, since parties make decisions about whether to litigate at all in the shadow of prevailing judicial standards, plaintiffs "are unlikely even to file a complaint unless they can support their claims with an event study likely to pass muster."[182]

The noisy factors mean that, at least some of the time, an abnormal return that looks "real" (based on the standard statistical threshold of 95% confidence)—and is therefore interpreted as having established, for example, loss causation—might no longer be "real" using a different vintage of the Fama-French data. And the converse will also be true: an event study conducted using one vintage could indicate a lack of statistical significance—thereby, for example, persuading a court to grant summary judgment for the defendant—when an identical analysis performed using a different vintage could have come back as statistically significant.

It is troubling, to say the least, that liability, settlement amounts, and (in the unlikely event that a securities fraud action makes it that far) damages awards could depend on something as arbitrary as when an analyst downloaded the data used in the analysis. This is particularly true in light of the fact that there is no particular reason to think that any one vintage is better, in the statistical sense, than any other.[183]

Unfortunately, it is very clear that the Fama-French data are used in securities litigation. This is not to say that they are always

---

179 Brav & Heaton, *supra* note 176, at 585.

180 Fisch et al., *supra* note 176, at 556.

181 *See* Fisch & Gelbach, *supra* note 173, at 61:

> Although securities fraud cases rarely go to trial and, as a result, judicial efforts to calculate damages are virtually non-existent, litigants also proffer event studies with respect to damages on motions for summary judgment as well as at the motion for class certification in response to Rule 23's requirement that damages can be calculated on a class-wide basis.

Since their article was published, the Supreme Court has only increased the importance of the class certification stage. *See* Goldman Sachs Grp., Inc. v. Ark. Tchr. Ret. Sys., 141 S. Ct. 1951, 1960 (2021) (holding that courts "should be open to *all* probative evidence" of price impact at the class certification stage (emphasis in original) (quotation marks omitted) (quoting *In re* Allstate Corp. Sec. Litig., 966 F.3d 595, 613 n.6 (7th Cir. 2020))).

182 Fisch & Gelbach, *supra* note 173, at 61.

183 *See supra* note 65 and accompanying text.

used perfectly. Indeed, many commentators have pointed out that the implementation of event studies by "experts" often deviates from the methodologies used in peer-reviewed scholarship in financial economics.[184] For example, experts in securities fraud actions sometimes include an index of firms in the same industry as a factor in the regression,[185] something that is inconsistent with both finance theory and best practice in empirical finance.

Notwithstanding this, it is not difficult to find examples of cases where at least one of the experts performed an event study using the Fama-French factors. One recent example of this is *In re Allstate Corp. Securities Litigation*,[186] where plaintiffs brought two putative class action complaints against Allstate under § 10(b) of the Securities Exchange Act of 1934[187] and Rule 10b-5.[188] After reviewing reports from both the plaintiff's and defendant's experts, the court granted the motion for class certification.[189] The plaintiff's expert, as it turns out, "calculated the expected returns on shares of Allstate's common stock by applying the widely accepted Fama-French Three-Factor Model."[190] While expert reports are not always easy to find—particularly for cases that settle—we found several other instances in recent years where experts relied on the Fama-French model (or a modified version thereof) in an analysis that was accepted by the court.[191]

---

184 *See, e.g.*, Brav & Heaton, *supra* note 176, at 583 ("[T]he [event study] methodology litigants use in court differs from the methodology that economists apply in their research."); Fisch et al., *supra* note 176, at 557 (explaining that "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").

185 *See, e.g.*, *infra* notes 190–91 and accompanying text.

186 2020 WL 7490280 (N.D. Ill. Dec. 21, 2020).

187 15 U.S.C. § 78 et seq.

188 *See In re Allstate Corp. Sec. Litig.*, 2020 WL 7490280, at *1.

189 *See id.* at *2.

190 Expert Report of John D. Finnerty, Ph.D. in Support of Lead Plaintiff's Motion for Class Certification ¶ 46, *In re Allstate Corp. Sec. Litig.*, 2020 WL 7490280 (No. 1:16-CV-10510). Dr. John Finnerty did not disclose the source of the factor data he relied upon, but based on his description, we are almost certain that he used the standard Fama-French data. Dr. Finnerty supplemented the three-factor model with a fourth "factor": "the returns on an industry index of common stocks that are comparable to Allstate." *Id.* ¶ 54. While this fourth factor is not supported by finance theory, there is no reason to expect its inclusion to change the effect of noisy factors on estimates.

191 Many of these reports were written by Dr. Finnerty. *See, e.g.*, *In re* Vale S.A. Sec. Litig., 2019 WL 11032303, at *14 (S.D.N.Y. Sept. 27, 2019) ("[B]ased on Dr. Finnerty's model, Lead Plaintiffs have carried their burden of showing that damages can be calculated on a classwide basis. This computation of damages is also consistent with Lead Plaintiffs' theory of fraud."). For example, Dr. Finnerty had "calculated the expected returns on the

Fortunately, in many relevant settings, the noisy factors will not be noisy enough to be dispositive, at least when it comes to establishing liability. After all, cases are usually brought when there has been an unusually large drop in the company's share price. When this happens, even a substantial change in the predicted return (the part of the event study analysis that is affected by the noisy factors) might not be enough to change the conclusion. It could still be dispositive in close cases where statistical significance is borderline. While it is hard to judge the frequency of such cases, they are arguably the ones where the analysis matters the most, since they are precisely the cases where uncertainty is highest. And, it goes without saying, even in cases where noisy factors are not dispositive for liability, they will always have an impact on the valuation of damages.

We are far from the first scholars to point out methodological problems with relying on event studies in securities litigation. Professor Gelbach, along with Professors Eric Helland and Jonathan Klick, pointed out an important statistical problem with the single-firm event studies commonly used in securities litigation.[192] A few years later, Professor Brav and scholar J.B. Heaton extended this critique by pointing out a series of other statistical and methodological problems with single-firm event

Vale [American Depository Receipts] by applying the widely accepted Fama-French Three-Factor Model." Expert Report of John D. Finnerty ¶ 34, *In re Vale S.A. Sec. Litig.*, 2019 WL 11032303, at *14 (No. 1:15-CV-09539). The expert supplemented the three-factor model with three additional factors. *See id.* ¶ 43. Dr. Finnerty had done this before. *See, e.g.*, City of Ann Arbor Emps.' Ret. Sys. v. Sonoco Prod. Co., 827 F. Supp. 2d 559, 584 (D.S.C. 2011) (denying the defendant's motions to exclude Dr. Finnerty's expert testimony and for summary judgment because "Dr. Finnerty's opinions are sufficient to create a genuine issue of material fact on the elements of loss causation and damages"). Here again, Dr. Finnerty had performed an event study using the Fama-French three-factor model, supplemented by an industry factor. *See* Declaration of John D. Finnerty, Ph.D. in Support of Loss Causation and Market Efficiency at 5, *City of Ann Arbor Emps.' Ret. Sys.*, 827 F. Supp. 2d at 584 (No. 4:08-CV-02348). Going as far back as 2009, Dr. Finnerty has conducted analysis in this fashion. *See* Silverman v. Motorola, Inc., 259 F.R.D. 163, 174 (N.D. Ill. 2009) (holding that the plaintiffs had met their burden and granting plaintiffs' motion for class certification). As above, Dr. Finnerty had performed an event study using the Fama-French three-factor model, supplemented with an industry factor. *See* Expert Report of John D. Finnerty, Ph.D. at 16, *Silverman*, 259 F.R.D. at 174 (No. 1:07-CV-04507). This time, Dr. Finnerty *did* disclose the source of his factor data. It was, indeed, the Fama-French data from Professor French's website. *Id.* at 68.

[192] *See* Gelbach et al., *supra* note 176, at 496–98. Their primary statistical concern arises from the non-normality of stock returns. *See id.* at 495–96.

studies.[193] Professors Gelbach and Klick returned to this problem a few years later, this time with Professor Fisch, and used the Supreme Court's decision in *Halliburton Co. v. Erica P. John Fund, Inc.*[194] to point out further problems with the ways that courts use event studies in securities litigation.[195] Most recently, Professors Fisch and Gelbach argued that the default "95% confidence" threshold is often unwarranted in securities litigation.[196]

We take no issue with any of these arguments. Rather, our point about the impact of the noisy factors represents an additional reason to proceed with caution when using event studies in securities law. Unfortunately, the problem of noisy factors would remain even if all the problems pointed out by other scholars were solved, since it is rooted in the underlying data that financial economists also routinely rely on in high-quality, peer-reviewed work.[197]

C. Scholarly Applications of Event Studies

Event studies are also ubiquitous in scholarly legal settings. As Professors Sanjai Bhagat and Roberta Romano explained in their seminal two-part article, *Event Studies and the Law*, "[e]vent studies are among the most successful uses of econometrics in policy analysis."[198] Since the time of that writing, their use and influence has become even more widespread. Unfortunately, for the reasons illustrated in Parts II and III, the noisy factors pose a problem for their use as an analytical tool. Precisely because event studies have been used so successfully—and so extensively—we limit ourselves to a brief sampling of relevant articles

---

193 These problems include the low statistical power, the inability to average away confounding effects, and an upward bias in detected price impacts. Brav & Heaton, *supra* note 176, at 586.

194 573 U.S. 258 (2014).

195 *See generally* Fisch et al., *supra* note 176.

196 *See* Fisch & Gelbach, *supra* note 173, at 613–14.

197 For example, in an empirical exercise meant to illustrate some of the problems with single-firm event studies, Professor Brav and Dr. Heaton relied on Professor French's factor data and used a four-factor model (consisting of the three Fama-French factors supplemented with the well-established Carhart momentum factor). Brav & Heaton, *supra* note 176, at 595 n.28. We note this not as a criticism, but rather to reinforce the point that the more rigorous the empirical analysis, the more likely it is to be affected by the noisy factors. Going forward, one easy near-term solution is for experts to use arm's-length, transparent versions of the factors. As discussed in more detail below, we have made the code used to construct our fixed-code factors freely available and invite experts to use them. *See infra* note 230 and accompanying text.

198 Sanjai Bhagat & Roberta Romano, *Event Studies and the Law: Part I: Technique and Corporate Litigation*, 4 AM. L. & ECON. REV. 141, 142 (2002) [hereinafter Bhagat & Romano, *Technique and Corporate Litigation*].

across a variety of contexts related to corporate and securities law. The same is true for their analytical cousins, regressions where the dependent variable is an alpha derived from a factor model. While these are distinct types of analysis, both are affected in similar ways by the noisy factors, so we consider them together for the sake of parsimony.

A first context is the market for corporate control. One strand of this literature, introduced in Part IV.A, is shareholder activism. Relying heavily on event studies, this literature has established a robust, positive relationship between hedge fund activism and the performance of target firms in both the short[199] and long run.[200] A related literature has studied "negative activism," where activists take a short position in target firms,[201] and may even try to drive down their prices.[202] Event studies have also been used to study other questions related to mergers and acquisitions, including impacts on the share price of acquirers[203] and the impact of mergers on innovation.[204]

A second broad category is corporate governance. As Professors Bhagat and Romano pointed out, "[v]irtually all of the important mechanisms of corporate governance have been subjected to event study analysis."[205] This has continued in the twenty years since

---

[199] *See, e.g.*, Brav et al., *Activism, Governance, and Performance*, *supra* note 165, at 1730 (finding that hedge fund activism has a positive short-term impact on target firms, with no reversal over the next year).

[200] *See, e.g.*, Bebchuk et al., *supra* note 165, at 1123–30 (finding no evidence that the short-term gains following activist interventions are followed by reversal over the subsequent five years).

[201] *E.g.*, Barbara A. Bliss, Peter Molk & Frank Partnoy, *Negative Activism*, 97 WASH. U. L. REV. 1333, 1338 (2020).

[202] *See generally, e.g.*, Joshua Mitts, *Short and Distort*, 49 J. LEGAL STUD. 287 (2020) (studying the impact of pseudonymous and nonpseudonymous attacks on target firms).

[203] *See, e.g.*, Bhagat & Romano, *Empirical Studies*, *supra* note 176, at 394–95 (summarizing the literature and noting that "[d]epending on the sample period and sample considered, studies document average bidder returns that cover the range from positive, economically small, and statistically insignificant, to negative, economically small, and statistically insignificant"); *see also* Laurence Capron & Nathalie Pistre, *When Do Acquirers Earn Abnormal Returns?*, 23 STRATEGIC MGMT. J. 781, 781 (2002) (investigating the conditions under which acquirers earn abnormal returns); Sara B. Moeller, Frederik P. Schlingemann & René M. Stulz, *Wealth Destruction on a Massive Scale? A Study of Acquiring-Firm Returns in the Recent Merger Wave*, 60 J. FIN. 757, 763 (2005) (investigating the impact of acquisitions on acquiring firm shareholders).

[204] *See, e.g.*, Darren Filson, Saman Olfati & Fatos Radoniqi, *Evaluating Mergers in the Presence of Dynamic Competition Using Impacts on Rivals*, 58 J.L. & ECON. 915, 922–23 (2015) (showing that the abnormal returns of rival pharmaceutical companies around merger announcements predict postmerger changes in the combined firm's R&D intensity).

[205] Bhagat & Romano, *Empirical Studies*, *supra* note 176, at 401. Professors Bhagat and Romano went on to discuss several relevant studies. *See id.* at 401–09.

their article. For example, event studies have been used to study the impact of shareholder rights,[206] managerial entrenchment,[207] and lax corporate governance[208] on shareholder value. They have also been used to study specific corporate governance arrangements, such as staggered boards,[209] dual class shares,[210] majority voting for directors,[211] limited liability,[212] Delaware's corporate opportunity waiver,[213] and "golden leashes" for activist-nominated

---

[206] *See, e.g.*, Paul Gompers, Joy Ishii & Andrew Metrick, *Corporate Governance and Equity Prices*, 118 Q.J. ECON. 107, 108–09 (2003) (developing an index measuring shareholder rights and finding that firms with stronger shareholder rights had higher returns). *But see* Jens Frankenreiter, Cathy Hwang, Yaron Nili & Eric Talley, *Cleaning Corporate Governance*, 170 U. PA. L. REV. 1, 42–45 (2021) (correcting the index in Gompers et al., *supra*, and repeating their analysis).

[207] *See, e.g.*, Lucian Bebchuk, Alma Cohen & Allen Ferrell, *What Matters in Corporate Governance?*, 22 REV. FIN. STUD. 783, 812–13 (2009) (developing an index measuring manager entrenchment and finding that higher levels of entrenchment are associated with negative abnormal returns); Jay B. Kesten, *Managerial Entrenchment and Shareholder Wealth Revisited: Theory and Evidence from a Recessionary Financial Market*, 2010 BYU L. REV. 1609, 1642–43 (finding that the relationship between high entrenchment and negative stock returns identified by Professors Lucian Bebchuk, Alma Cohen, and Allen Ferrell disappeared during the 2008 financial crisis).

[208] *See, e.g.*, Ofer Eldar, *Can Lax Corporate Law Increase Shareholder Value? Evidence from Nevada*, 61 J.L. & ECON. 555, 580–82 (2018) (finding that Nevada reincorporation does not harm shareholder value).

[209] *See, e.g.*, Olubunmi Faleye, *Classified Boards, Firm Value, and Managerial Entrenchment*, 83 J. FIN. ECON. 501, 514–15 (2007) (finding a strong positive relationship between a firm's decision to destagger its board and its performance); Mira Ganor, *Why Do Managers Dismantle Staggered Boards?*, 33 DEL. J. CORP. L. 149, 185–87 (2008) (finding a null relationship between a firm's decision to destagger its board and firm performance); Lucian A. Bebchuk, Alma Cohen & Charles C.Y. Wang, *Reexamining Staggered Boards and Shareholder Value*, 125 J. FIN. ECON. 637, 640–46 tbls.1–4 (2017) (finding that staggered boards reduce shareholder value under a variety of specifications).

[210] *See, e.g.*, Valentin Dimitrov & Prem C. Jain, *Recapitalization of One Class of Common Stock into Dual-Class: Growth and Long-Run Stock Returns*, 12 J. CORP. FIN. 342, 352–53 (2006) (finding that dual-class recapitalizations increase shareholder value); Scott B. Smart, Ramabhadran S. Thirumalai & Chad J. Zutter, *What's in a Vote? The Short- and Long-Run Impact of Dual Class Equity on IPO Firm Values*, 45 J. ACCT. & ECON. 94, 112–13 (2008) (finding that unifying share classes increases shareholder value).

[211] *See, e.g.*, Jay Caia, Jacqueline L. Garner & Ralph A. Walkling, *A Paper Tiger? An Empirical Analysis of Majority Voting*, 21 J. CORP. FIN. 119, 127–33 (2013) (concluding that majority-voting proposals appear to be a matter of form over substance).

[212] *See, e.g.*, Mark I. Weinstein, *Share Price Changes and the Arrival of Limited Liability in California*, 32 J. LEGAL STUD. 1, 10 (2003) (finding that the introduction of limited liability in California had no effect on share prices); Mark I. Weinstein, *Don't Buy Shares Without It: Limited Liability Comes to American Express*, 37 J. LEGAL STUD. 189, 221 (2008) (finding that the adoption of limited liability had little effect on the value of American Express).

[213] *See, e.g.*, Gabriel Rauterberg & Eric Talley, *Contracting Out of the Fiduciary Duty of Loyalty: An Empirical Analysis of Corporate Opportunity Waivers*, 117 COLUM. L. REV. 1075, 1133–36 (2017) (finding a weak share price increase after the adoption of corporate opportunity waivers).

director candidates.[214] Yet others have focused on specific features of shareholder voting, such as shareholder proposals[215] and the use of proxy advisers.[216]

Finally, event studies are used to study the impact of legal and regulatory changes driven by legislatures,[217] courts,[218] regulators,[219] or a combination thereof[220] across a variety of areas of law.[221] As Professors Bhagat and Romano pointed out, event studies have also been used to study the wealth effects of corporate litigation more broadly,[222] and more recent work has extended

---

214 *See, e.g.*, Matthew D. Cain, Jill E. Fisch, Sean J. Griffith & Steven Davidoff Solomon, *How Corporate Governance Is Made: The Case of the Golden Leash*, 164 U. PA. L. REV. 649, 685–94 (2016) (finding that golden leashes are positively related to stock returns of firms facing activist attention).

215 *See, e.g.*, John G. Matsusaka, Oguzhan Ozbas & Irene Yi, *Can Shareholder Proposals Hurt Shareholders? Evidence from Securities and Exchange Commission No-Action-Letter Decisions*, 64 J.L. & ECON. 107, 117–19, 125 (2021) (finding that shareholder proposals reduce firm value).

216 *See, e.g.*, David F. Larcker, Allan L. McCall & Gaizka Ormazabal, *Outsourcing Shareholder Voting to Proxy Advisory Firms*, 58 J.L. & ECON. 173, 203 (2015) (finding that outsourcing voting to proxy advisory firms may induce boards to decrease shareholder value).

217 *See, e.g.*, Haidan Li, Morton Pincus & Sonja Olhoft Rego, *Market Reaction to Events Surrounding the Sarbanes-Oxley Act of 2002 and Earnings Management*, 51 J.L. & ECON. 111, 122–25 (2008) (finding a positive relationship between the Sarbanes-Oxley Act of 2002 and stock returns, particularly for firms that engaged in more earnings management). *See generally* Joel F. Houston, Chen Lin & Wensi Xie, *Shareholder Protection and the Cost of Capital*, 61 J.L. & ECON. 677 (2018) (finding evidence that weakening litigation rights increases the cost of capital).

218 *See, e.g.*, Sean J. Griffith & Natalia Reisel, *Dead Hand Proxy Puts and Shareholder Value*, 84 U. CHI. L. REV. 1027, 1061 (2017) (finding no negative reaction to the Delaware Chancery Court's three "dead hand proxy put" rulings). *See generally* Peter Molk, *Delaware's Dominance and the Future of Organizational Law*, 55 GA. L. REV. 1111 (2021) (finding a negative reaction after the Delaware Supreme Court undermined the state's commitment to responsive limited liability company law).

219 *See, e.g.*, Randolph Beatty & Padma Kadiyala, *Impact of the Penny Stock Reform Act of 1990 on the Initial Public Offering Market*, 46 J.L. & ECON. 517, 532–38 (2003) (finding evidence that speculative issuers migrated into the nonpenny range after the Penny Stock Reform Act of 1990); Allen Ferrell, *Mandatory Disclosure and Stock Returns: Evidence from the Over-the-Counter Market*, 36 J. LEGAL STUD. 213, 245–47 (2007) (finding that the 1964 imposition of mandatory disclosure requirements on the over-the-counter market had a positive effect on stock returns).

220 *See, e.g.*, Bo Becker, Daniel Bergstresser & Guhan Subramanian, *Does Shareholder Proxy Access Improve Firm Value? Evidence from the Business Roundtable's Challenge*, 56 J.L. & ECON 127, 154–57 (2013) (finding evidence that proxy access increased shareholder value).

221 *See* Bhagat & Romano, *Empirical Studies*, *supra* note 176, at 390–414 (discussing a variety of examples).

222 *See* Bhagat & Romano, *Technique and Corporate Litigation*, *supra* note 198.

this application of event studies to analyze cross-sectional differences in the impact of corporate litigation.[223]

Like the use of event studies in practice, the scholarly use of event studies in corporate law scholarship has also been criticized. In a recent working paper, Professors Emiliano Catan and Marcel Kahan identified several conceptual and methodological problems with the use of event studies to answer questions in corporate governance. Instead, they argued that scholars should focus on unlevered firm returns.[224]

The noisy factors constitute a separate cause for concern from those raised by Professors Catan and Kahan. In particular, they are likely to be a problem whenever scholars rely on the Fama-French data and find results where the statistical significance is not overwhelming. Based on our work to date, we think that robust results, and results that have been replicated in many contexts, are less likely to be sensitive to the noisy factors than those that are not. But there are many published articles that contain results that fall into the latter group. At least some of these results are likely to be affected, through no fault of their authors. This point bears repeating: the scholarly community was not aware of the noisy factors until very recently. It would therefore be unreasonable and unfair to hold the scholars responsible in the event that previously published results turn out to be sensitive to the retroactive changes in the Fama-French data.

V. BROADER IMPLICATIONS

The most obvious implication of the noisy factors is that experts should stop using the Fama-French data, especially in legal settings. But there are two other, broader lessons that we can draw from this episode. First is the law of conservation of judgment: judgment can never be removed from an analysis; it can only be moved around. Second, there are the risks that arise when the interests of experts diverge from those of the legal system. In this instance, these diverging interests take two very different

---

[223] *See, e.g.*, Utpal Bhattacharya, Neal Galpin & Bruce Haslem, *The Home Court Advantage in International Corporate Litigation*, 50 J.L. & ECON. 625, 633–38 (2007) (finding that U.S.-firm defendants experience smaller drops in share price than foreign firms upon the announcement of a lawsuit in U.S. federal court).

[224] *See generally* Emiliano Catan & Marcel Kahan, *Corporate Governance and Firm Value* (Feb. 2025) (Eur. Corp. Governance Inst., L. Working Paper No. 824/2024) (available on SSRN).

forms: the interests that arise when academic work becomes entangled with commercial or other economic interests and the incentives of academics to adopt "standard" data and approaches, rather than asking too many questions.

A. The Fama-French Factor Data Are Not Appropriate for Legal Settings

As the analysis in both our companion article and this Article has shown, financial analyses that rely on the Fama-French data, as well as the conclusions that follow from those analyses, are not reliable as evidence in court.

Indeed, our analysis casts doubt on whether expert evidence that relies on the noisy factors should even be admissible. In federal court, this question is governed by the standard developed by the Supreme Court in *Daubert v. Merrell Dow Pharmaceuticals, Inc.*[225] and incorporated into Federal Rule of Evidence (FRE) 702 in 2000. FRE 702 requires that

> (a) the expert's scientific, technical, or other specialized knowledge will help the trier of fact to understand the evidence or to determine a fact in issue;
> (b) the testimony is based on sufficient facts or data;
> (c) the testimony is the product of *reliable* principles and *methods*; and
> (d) the expert's opinion reflects a reliable application of the principles and methods to the facts of the case.[226]

The *Daubert* standard allows judges to consider, among other factors, "whether the theory or technique can be (and has been) tested, whether it has been subjected to peer review and publication, [and] its known or potential error rate and the existence and maintenance of standards controlling its operation."[227] Judges were admonished to focus "solely on principles and methodology, not on the conclusions that they generate."[228]

Given the analysis in this Article and its companion piece, it is hard to see how the Fama-French data can be used to support expert testimony that satisfies the *Daubert* standard and FRE 702 requirements. The fact that the Fama-French factor data change regularly, and that these changes materially affect

---

225 509 U.S. 579 (1993).
226 FED. R. EVID. 702 (emphasis added).
227 *Daubert*, 509 U.S. at 580.
228 *Id.*

analyses that rely on that data, undermines any argument about reliability. Furthermore, the fact that the changes are entirely discretionary and up to the creators to implement with no meaningful discussion underscores the lack of reliable standard controlling the data's construction. This is only exacerbated by the fact that code used to generate the data has not been made public. While *caveat emptor* might be good enough for some settings, it is not good enough to sustain an expert analysis.

It is also not good enough when it comes to a fiduciary discharging her duties. To be sure, fiduciary duties do not require perfection. But it is hard to imagine that a prudent person would rely on a compromised data source in the management of her own affairs.[229] To see why, we need only imagine that the fiduciary was relying on a tool that randomly gives different numbers. Even if that random number generator was unbiased, so that on average it gave the right result, it is hard to justify relying on that tool when another is easily available. And another tool—without this quality—is available: we have made the code that we used to construct the fixed-code factors freely available online so that anyone can use it.[230]

The same goes for scholars: it is hard to see how one can justify continuing to use these data in light of our findings. This is particularly true when an easy and free substitute is available. In addition to the obvious benefits of not relying on compromised data, declining to use French's data helps to build norms of code sharing in the academy, which has long-term and broad-based benefits. And if lofty appeals to the scholarly enterprise are not enough, we can appeal to self-interest: After all, what scholar wants to write a paper that she already knows is likely to fail to replicate in a few years? Even for purely self-interested reasons, our results demonstrate the downside of relying on French's data.

To be clear, nothing in our analysis should be taken to mean that any expert that relied on the Fama-French data in the past did anything wrong. After all, all three of us have used it in our scholarly work, as have thousands of other scholars over the past

---

229 Under ERISA, a fiduciary is required, inter alia, to "discharge his duties . . . with the care, skill, prudence, and diligence under the circumstances then prevailing that a prudent man acting in a like capacity and familiar with such matters would use in the conduct of an enterprise of a like character and with like aims." 29 U.S.C. § 1104(a)(1)(B).

230 The data are available at *Noisy Factors Replication Code*, BOX, http://www.law.uchicago.edu/law-finance/code/NoisyFactors.

thirty years.[231] Before our research was made public, if one had asked twenty financial economists to list the ten most likely loci of problems with empirical analysis in finance, the Fama-French data almost certainly would not have made any of their lists. But now that we do know about these problems, there is no excuse to continue using them.

B. Expert Analysis and the Law of Conservation of Judgment

A much broader implication of our analysis is the pervasiveness of the law of conservation of judgment: even what appears to be an entirely technical, technocratic exercise inevitably involves an enormous amount of discretionary decision-making. To be clear, this is not a law of science like gravity, nor is it a law the way the Securities Exchange Act of 1934 is. Instead, it is a description of a general and widespread phenomenon: that expertise and fancy methods do not eliminate judgment and discretion; they can only move it around. And often, the fancier the technique, the less able even the expert is to identify the loci and effect of that discretion.

To illustrate, we can return to the example of an appraisal action before the Delaware Court of Chancery (i.e., the type of action that *In re Appraisal of Dell Inc.* was an example of). We can imagine a continuum: at one extreme, the judge (or, more precisely, the chancellor or vice chancellor) could simply decide, in her judgment, what she thinks the value of Dell was on a particular date. At the opposite extreme, she might rely on a black-box machine learning model that was trained on reams and reams of data, and which ultimately spits out a number. The type of DCF model routinely used in valuation lies somewhere in between.

All of these involve judgment. When the judge is the one deciding, the judgment is right there in the open. Presumably, she will provide reasons in her opinion. At the other extreme, the judgment is hidden in a myriad of places, including the type of algorithm used, the trading data provided to that algorithm, and the internal workings of the software package used to implement the analysis. At each of these steps, the expert who built the model exercised judgment: There were multiple options available to him, and he had to pick one. Each one unquestionably affected the final valuation. It is just that the *way* that they affected the

---

[231] *See supra* note 26.

outcome is opaque, including, in all likelihood, to the expert himself. Even if he wanted to transparently explain them (and their implications on the bottom-line number) to the judge, he probably could not. The DCF model lies somewhere in between. Some of the discretionary choices that go into implementing a DCF valuation can be explained in a sensible way, while others (like which factor vintage is used to estimate the betas) are likely to be beyond even the expert's understanding. And so, the judgment and discretionary decisions that went into the changes to the Fama-French factors can and will affect valuations that rely on those factors. But, until quite recently, this would not even have been on the most careful expert's radar.

In short, even the most technical and technocratic analyses still involve substantial amounts of judgment. The question is not whether judgment is removed—leaving some objective truth—but rather who is exercising the judgment, to what extent it is and can be explained, and what effect it has on the valuation.

Event studies are similarly affected. Take, for example, the use of event studies to assess materiality in securities litigation.[232] Under the federal securities laws, a fact is material if there is a "substantial likelihood that the . . . fact would have been viewed by the reasonable investor as having significantly altered the 'total mix' of information made available."[233] A materiality determination is notoriously slippery, and there are a variety of different methods available to parties to establish or rebut materiality. We can think of these approaches as lying on a continuum, just like our approaches to valuation. At one extreme, there is the type of "I know it when I see it" approach that Professor Daniel Fischel once called "the traditional model."[234] This is unabashedly an exercise of applying judgment.

Other approaches that courts have adopted are further down the continuum. For example, take the "rule of thumb" approach articulated in SEC's Staff Accounting Bulletin (SAB) No. 99, which discusses the use of quantitative thresholds like 5% to assess the materiality of items or statements in accounting statements.[235] Rather than endorsing or rejecting such an approach, the SEC staff's opinion was that while a numerical threshold can

---

[232] *See supra* notes 175–76 and accompanying text.

[233] TSC Indus. v. Northway, Inc., 426 U.S. 438, 449 (1976).

[234] *See* Fischel, *supra* note 174, at 1, 2.

[235] *See* Staff Accounting Bulletin No. 99, 64 Fed. Reg. 45,150 (Aug. 12, 1999).

be "an initial step in assessing materiality," it "is only the beginning" and "cannot appropriately be used as a substitute for a full analysis of all relevant considerations."[236] Here, too, there is a substantial amount of discretion: What should the denominator for the 5% be—profits, revenues, assets, or something else? Which other factors must be considered, and how should they be weighed against each other? While this will not produce certainty, it is still fairly straightforward for an appellate court to review these exercises of discretion.[237]

At the other extreme is a full-blown event study with all the bells and whistles. Here, the outcome will be affected by all sorts of discretionary choices, including—but by no means limited to—the choices that go into the construction of the factors. In short, all of these involve discretion; the questions are just who is exercising that discretion and how it is presented and explained.

Finally, there are fiduciaries. In some sense, the exercise of judgment is inherent in a fiduciary's role.[238] She cannot escape this by relying on what appears to be a technical or technocratic analysis. As illustrated in the example in Part III.A, the noisy factors markedly change the results of the "textbook" analysis and would similarly affect any advice that followed from that analysis. There is, in short, no way to get around discretion. The only question is whether the judgment is exercised by the fiduciary or by someone else who is not focused on the best interest of the principal.

One way to address the issues that arise from the law of conservation of judgment is by starting with a simple analysis. The downside of simple analyses—which is why more sophisticated methods exist—is that they tend to have well-known flaws. That can be turned into a virtue: after conducting a simple analysis, if the expert goes on to find different results using a more sophisticated method, she should try to articulate with particularity, based on the known flaws in the simple approach, both why the sophisticated approach yields different results and why those

---

[236] *Id.* at 45,151.

[237] For a clear illustration of this, see *Litwin v. Blackstone Group, L.P.*, 634 F.3d 706 (2d Cir. 2011), where the Second Circuit vacated the district court's dismissal of securities fraud claims under the Securities Act of 1933. *Id.* at 723. The district court had held that the statements at issue were not material as a matter of law, using the quantitative and qualitative factors in SAB 99. *See id.* at 713–15. The Second Circuit applied the same test and came to the opposite conclusion. *See id.* at 716–23.

[238] After all, we arguably would not need fiduciary duties at all if the principal could specify in an enforceable way exactly what she wanted the agent to do.

results are more reliable. If she is unable to do so, she should proceed with caution. In particular, she may be better off relying on an approach that she knows is imperfect—but the imperfections of which she understands—than on a sophisticated approach that she thinks is right, but does not know why.

Judges, for their part, should be skeptical of elaborate technical approaches that lead to meaningfully different answers from simple, easy-to-understand alternatives. While there may be good reasons why these simple alternatives are wrong or inappropriate, it should be up to the expert offering the more complex analysis to explain exactly why that is, and to provide a compelling justification for why her approach is better. And, of course, the judge should never be fooled by arguments that a complex technical analysis is free of judgment or discretion. If an expert claims that this is the case, she is either lying or does not understand her own method well enough to know where that judgment is.

C. The Entanglement of Academic and Commercial Interests

Because judgment is an unavoidable feature of all analyses, it is essential to think carefully about the interests and incentives of those that exercise judgment. This is particularly important when they do so in ways that are not entirely transparent.

Until quite recently, the documentation surrounding the provenance of the Fama-French data was quite parsimonious. It was posted on Professor French's academic website, which carried a Tuck School of Business at Dartmouth URL.[239] While it had been well-known, and clearly disclosed, that Professors Fama and French both had financial relationships with DFA,[240] there was nothing to suggest that the Fama-French *data* had anything to do with that. Even if French was not personally running the code and updating the website to keep the data current, most scholars probably assumed that he had delegated this task to a graduate student or research assistant at Dartmouth.

But it turns out that this is not what was happening. After we began circulating the preliminary version of *Noisy Factors? The Retroactive Impact of Methodological Changes on the Fama-*

---

[239] *See Description of Fama/French Factors*, KENNETH R. FRENCH https://perma.cc/CMZ7-D2BJ.

[240] *See, e.g.*, *Consulting Relationships*, KENNETH R. FRENCH, https://perma.cc/DJY6-25XC; *see also* Eugene F. Fama, Vita (Aug. 2020) (available at https://perma.cc/HJH5-JFWV); *About Us*, DIMENSIONAL, https://www.dimensional.com/us-en/who-we-are/about-us#board-of-directors (listing both Fama and French as Directors of FDA).

French Factors in the fall of 2021, a few well-connected finance scholars suggested to us that DFA was behind the data. But these people were not able to point to any evidence to support this. As we dug deeper into the research, we stumbled onto the source code of French's webpage. While the webpage has a "mba.tuck.dartmouth.edu" URL, its source code indicated that it was "[d]eveloped by Dimensional Fund Advisors Web Team."[241] Of course, the fact that the website was developed by a team at DFA does not necessarily tell us who is producing the data, particularly since the source code also says that "[a]ll images and code are property of Ken French."[242] It was not until two years later, in November 2023, that we received definitive proof of the relationship between the factor data and DFA. On November 10, 2023, Professor Fama sent an email to us, as well as to several scholars that were thanked in the acknowledgements section of the working paper versions of this Article as well as our companion article.[243] Attached to that email was a draft of their piece, *Production of U.S. Rm-Rf, SMB, and HML in the Fama-French Data Library*, which was posted on SSRN three days later. And on page five was (to our knowledge, for the first time) the disclosure that "Dimensional Fund Advisors' research group" has been "helping with the updates" to the factors since 2003.[244] The paragraph went on to explain that Fama and French "continue to determine the rules, definitions, and process used to form factor portfolios. Under [their] guidance, Dimensional employees produce the monthly updates, post them on a Dartmouth server, maintain the computer code, and until 2021 updated [their] CRSP-Compustat links."[245]

This matters because the discretion in the construction of the factors has consistently led to improvements in the performance of the value factor. Specifically, Fama and French acknowledged the implementation of three discretionary changes that affect the value factor in their report: First, in August 2016, they made a change to their calculation of book equity in light of Financial Accounting Standard Board's Financial Accounting Standard

---

241 *Home*, KENNETH R. FRENCH, https://perma.cc/28V4-BFH2 (displaying the source code for French's website).

242 Presumably this refers to the HTML code that makes up the website, and not the computer code used to produce the factors. *See id.*

243 Email from Eugene Fama, Professor of Fin., Univ. of Chi. Booth Sch. of Bus., to Pat Akey, Author, et al. (Nov. 10, 2023) (on file with authors).

244 Fama & French, *Production*, *supra* note 28, at 5.

245 *Id.*

(FAS) 109, which was issued in 1993.[246] Then, in August 2020, they ended an earlier (presumably pre-1993) response to FAS 106, an accounting rule change that had been issued in 1990.[247] According to the report, they did so because they concluded that it "had little impact on the cross-section of book-to-market equity,"[248] which is the characteristic that is used to construct the value factor.[249] Finally, in September 2021, they changed the process that they used to link corporate balance sheet data to stock return data.[250] Specifically, they moved from the files that they had developed and updated using their own internal processes from 1992 to 2021 to a third-party, publicly available linking file.[251] We find that the first two led to an improvement in the performance of the value factor, as calculated in the period since 1993, the year that the original paper proposing the factor model was published.[252] The third—which represents a move *away* from discretion and toward an objective, third-party rule—*decreased* the performance of the value factor.[253] We are pleased that nothing in the Fama-French report is inconsistent with our findings.[254] To the contrary, we view it as a successful replication of our work.

To understand why this matters, we need to pause to say a few words about DFA. DFA is one of the largest asset managers in the world, with about $750 billion in assets under management

---

246 *See id.* at 4 (opting not to add Deferred Taxes and Investment Tax Credits to book equity for fiscal years 1993 and later in accordance with FAS 109's improvements in accounting for deferred income taxes).

247 *See id.* at 3–4.

248 *Id.* at 4.

249 BERK & DEMARZO, *supra* note 4, at 507.

250 *See* Fama & French, *Production*, *supra* note 28, at 4–5.

251 *See id.* at 5.

252 We assess the performance of the value factor using its Sharpe ratio, which is a measure of risk-adjusted performance. The Sharpe ratio of French's value factor increased—corresponding to improved performance—after each of these changes. The fixed-code version of the factor did not, indicating that the changes were not due to changes in the underlying raw data. *See* Akey et al., Impact of Methodological Changes, *supra* note 10, at fig 3 panel A; *see also* Fama & French, *Production*, *supra* note 28, at tbl.1 panel B (showing that each of these changes led to a higher return on the HML factor in the affected period).

253 The Sharpe ratio of French's value factor fell modestly after this change. The Sharpe ratio of the fixed-code version of the factor also fell, but very slightly. *See* Akey et al., Impact of Methodological Changes, *supra* note 10, at fig.3 panel A; *see also* Fama & French, *Production*, *supra* note 28, at tbl.1 panel B (showing that this change led to a lower return on the HML factor in the affected period).

254 *See generally* Fama & French, *Production*, *supra* note 28.

as of early 2024.[255] It is known for its scientific investing techniques, which draw on the expertise of Fama and French.[256] Over time, it has come to be associated specifically with the value factor,[257] and focuses on "financial science"[258] and "harvesting beta"[259] rather than chasing alpha. As a consequence, the desirability of its investment strategy is associated with the performance of the value factor.[260] Unfortunately, the value factor had, quite famously, underperformed over the past several decades leading up to the initial release of our companion piece.[261] This, in turn, led to growing skepticism from both the financial press and investment advisers about the wisdom of such a strategy.[262] DFA used

---

[255] *See* Dimensional Fund Advisors LP Uniform Application for Investment Adviser Registration and Report by Exempt Reporting Advisers (Form ADV), at item 5 (Mar. 28, 2024).

[256] For example, the first line of DFA's website for individual investors in the United States says: "The scientific pursuit of a better way to invest." *Dimensional Investing*, DIMENSIONAL, https://perma.cc/2EWW-5LDB. Scrolling down slightly, the text reads: "Rely on science, not speculation" and "Dimensional is driven by an evidence-based approach, Nobel Prize–winning insights, and decades of expertise applying financial science to real-world portfolios." *Id.*

[257] More recently, it has also come to be associated with the profitability factor. Profitability was one of the two additional factors that Fama and French added to their initial three factor model when they developed their five-factor model. *See generally* Eugene F. Fama & Kenneth R. French, *A Five-Factor Asset Pricing Model*, 116 J. FIN. ECON. 1 (2015).

[258] *See Dimensional Investing*, *supra* note 256.

[259] *See e.g.*, Wes Cril, *Expectations vs. Reality in Value Funds*, DIMENSIONAL (Mar. 3, 2023), https://www.dimensional.com/us-en/insights/expectations-versus-reality-in-value-funds ("A process that stays the course in its pursuit of value can therefore boost the odds of harvesting the premium when value stocks outperform.").

[260] Historically, DFA was also associated with a size strategy. Today, DFA offers funds targeting both large- and small-cap stocks, as well as all-cap funds. Its flagship large-cap value fund is currently substantially larger—in terms of assets under management—than its small-cap value fund.

[261] *See, e.g.*, Larry Swedore, *It's Too Soon to Say the Value Premium Is Dead*, MORNINGSTAR (Sept. 27, 2023), https://perma.cc/HK62-7NCN ("The underperformance of U.S. value stocks since the Great Recession has received much attention from the financial media, and prompted at least some investors to conclude that value investing is dead."); *An Exceptional Value Premium,* DIMENSIONAL (Oct. 5, 2020), https://www.dimensional.com/us-en/insights/an-exceptional-value-premium ("It's probably not news to most value investors that the value premium has struggled over the past decade."); Jeremy Wang, *The Underperformance of Value: Is This Time Different?*, VISTA CAP. PARTNERS (Apr. 26, 2024), https://perma.cc/94RR-RZPA (emphasis in original):

> Over the past decade, however, value stocks have returned "just" 9.9% per year, while growth stocks have returned 14.4% per year. Quite simply, the value premium has turned negative—value stocks have *underperformed* growth stocks by nearly 5% per year. This has led many to declare that value investing is dead.

[262] *See, e.g.*, Justina Lee, *Is Value Dead? Debate Rages Among Quant Greats from Fama to AQR*, BLOOMBERG (June 3, 2020), https://www.bloomberg.com/news/articles/2020-06-03/quant-fight-2020-inside-wall-street-s-big-argument-on-value.

its marketing material to push back against these concerns, including, notably, by citing French's data—the data that they have since acknowledged is produced by DFA—as evidence of the performance of the value factor.[263] This performance, we now know, was retroactively improved by the discretionary decisions made over the course of the past decade.[264] And those changes, in turn, leaked into the plethora of legal settings discussed in Parts II through IV.

To be clear, we have no way of ascertaining the motivations for the changes that were implemented to the construction of the factors. Nor do we have any idea what input, if any, DFA employees had in that process. But this is the problem with the commingling of academic and pecuniary interests: once it occurs, it becomes difficult to know where the first ends and the second begins. As Professor Luigi Zingales has argued, economists (or, for that matter, any other group of experts) are vulnerable to capture.[265] Full and fair disclosure of conflicting interests is one option to address this risk. While it is impossible to know for sure, we suspect that academics might have viewed the factors differently if they had been named "the DFA factors." While disclosure represents a bare minimum, there is a great deal of evidence that disclosure alone is ineffective.[266] Therefore, the best way to guard against that capture is to avoid the conflict.

D. The Incentive to Take the Path of Least Resistance

The scholars who have been relying on a data source that they did not fully understand for decades are responding to their own set of incentives. Typically, when a scholar downloads and relies upon the Fama-French data in an academic study, she is using them as control variables of one kind or another. When she does so, it is generally with an eye toward satisfying a highly skeptical referee, journal editor, or other reader. And as we discussed in Part I.B, there are sensible reasons why a skeptical

---

263 *See id.*

264 *See* Akey et al., Impact of Methodological Changes, *supra* note 10, at fig. 1 panel B (showing that the sum of discretionary changes between 2005 and 2022 resulted in a consistently and substantially higher calculated return on the HML factor).

265 *See generally* Luigi Zingales, *Preventing Economists' Capture*, *in* PREVENTING REGULATORY CAPTURE: SPECIAL INTEREST INFLUENCE AND HOW TO LIMIT IT 124 (Daniel Carpenter & David Moss eds., 2013).

266 *See generally* OMRI BEN-SHAHAR & CARL E. SCHNEIDER, MORE THAN YOU WANTED TO KNOW: THE FAILURE OF MANDATED DISCLOSURE (2014).

reader might want her to use a standard dataset: after all, whatever problems the data might have, at least our skeptic is pretty confident that it is at arm's length from the author in question.[267]

Of course, this does not mean that the noisy factors are not a very serious problem for scholarly work that relies upon them.[268] But it may help to understand why so many scholars overlooked them for so long. We can easily see how other incentives also contributed: it is probably a better strategy to adopt the same approach that everyone else is using than to spend time developing a deep understanding of everything that approach entails. After all, most well-established results are well-established for a good reason. It might be nice for a scholar to fully understand every component of her analysis, but even if it were feasible, doing so would inevitably take time and energy away from her own original research. This is a risky strategy in a competitive academic environment, where she is rewarded for developing increasingly sophisticated techniques and uncovering novel findings.

CONCLUSION

Academics were the ones who brought the Fama-French factor data into legal settings. As it turned out, they did not fully understand their provenance, their construction, or the effects of discretionary decisions on the analyses that rely upon them. These effects, it turns out, were very large and pervasive and now cast doubt on the reliability of any analysis that uses them. As a result, experts should stop using these data, at least in legal settings; they simply do not rise to the required level of reliability. This episode is a demonstration of the law of conservation of judgment: judgment can never be removed from an analysis, it can only be moved around, often to a place one would not think to look. Finally, our analysis highlights problems that arise from the commingling of academic and commercial interests. While disclosure is one solution to this problem, it is probably safer to avoid the conflict entirely.

---

[267] *See supra* Part I.B. Having dueling experts construct all their own intermediate data raises exactly the same concerns. As bad as the risk of something like the noisy factors is, it is not clear that it is better to let expert witnesses construct their own factors. After all, those experts will have incentives to, at the very least, break ties in favor of their clients' positions. This concern is somewhat attenuated in the fiduciary context, but expecting fiduciaries (or their advisers) to construct their own intermediate data is costly (particularly if they do so by hiring outside experts) and error prone (particularly if they do not).

[268] Naturally, the same reasoning applies for the applications of event studies in legal scholarship.