Raymond J Biederman
Founding Partner, MBCB Attorneys
CEO, Proteus Discovery Group
Ray.Biederman@proteusdiscovery.com

Sean P Burke
Founding Partner, MBCB Attorneys
Director, Proteus Discovery Group
Sean.Burke@mbcblaw.com

March 31, 2022

Via Electronic Mail
James F. Humphreys Complex Litigation Center
George Washington Law School
humphreyscenter@law.gwu.edu

Re: GW Law School "A New Framework" Request for Public Comment

Dear New Framework Authors:

Thank you for the opportunity to comment on "The Discovery Proportionality Model: A New Framework." While this comment may be critical of the New Framework in certain respects, we sincerely thank the contributors to the New Framework who gave their expertise, time, and resources to this significant attempt to move the discussion of proportionality forward.

This comment offers certain general critiques and analyzes how those shortcomings undercut specific portions of the New Framework's proposal.

THE AMOUNT IN CONTROVERSY IS CONTROVERSIAL.

The New Framework focuses "on cumulative costs and burdens compared with the amount in controversy and importance of issues at stake…" (P. 10).

In a surprise to no one, litigants don't often agree. If the starting point for the New Framework is forging some understanding between opposing parties and the court about the amount in controversy, therein lies the New Framework's first problem.

The New Framework hopes to add certainty to the proportionality question by focusing on the amount in controversy – and what is proportional to it – but this calculation cannot be satisfied by looking at a simple historical average of *similar cases* (a difficult concept in its own right) adjudicated prior to 2008 as presented by "Litigation Costs in Civil Cases: Multivariate Analysis (Lee III and Willging, 2010), and which forms the basis of the most central assumption of the New Framework.

In many cases, the amount in controversy is itself a hotly contested and controversial issue. While damages may be ascertainable with mathematical certainty in certain cases, in most cases, we submit, that the amount in controversy is a question for resolution by the fact finder after discovery is well under way or complete.

Because the amount in controversy cannot be established with any degree of certainty at the outset of litigation, the wisdom of relying upon that assumption seems to be a fatal flaw of the model.

THE TOTAL COST OF DISCOVERY IS A MOVING TARGET

The New Framework purports to cap discovery costs as a fraction of the amount in controversy. Even if the amount in controversy is ascertainable, the resulting cap on discovery costs is not a panacea for managing discovery.

Every litigant knows that attorney review makes up the vast majority of eDiscovery cost: 77% as demonstrated by "GW Discovery Proportionality Model – Judicial Detail View". Under a total cost model, an enterprising litigant may say "just produce it all *without* review and then clawback what you need to." Such an approach would lower the total cost of discovery, presumably allowing for more discovery without any consideration of the value of that discovery. This approach, however, completely misses the point of the proportionality inquiry and is inconsistent with the reasonable and universal belief that discovery should not be a court-sanctioned fishing expedition.

Further, litigant choices often skew the total cost of discovery and can be manipulated. For example, as shown by the Lee III and Willging research relied upon by the New Framework, choice of law firm and that firm's may be a cause of the non-proportional cost and that choice is completely up to the producing party. *See* p. 8 of Lee III and Willging (noting that "a defendant represented by an attorney in a firm of more than 500 attorneys had costs more than double (156% larger) those of a sole practitioner, all else equal.")

MARGINAL UTILITY REMAINS IMPORTANT AS RECOGNIZED BY THE NEW FRAMEWORK.

The New Framework hopes to "break [] away from traditional practices, which assess whether discovery is proportional based primarily on the marginal utility".

While the New Framework purports to depart from a marginal cost analysis, it doesn't. Instead, the New Framework specifically relies on concepts like custodian prioritization and high/low burden assessments, which are exercises in marginal utility. For example, the purpose of the high/low designation appears to be based on the relative or incremental value of the proposed discovery in meeting case requirements versus the cost of acquiring ESI from specific data stores and/or custodians.

Similarly, an argument consisting only of "It costs too much" without proving that there is limited value to some particular aspect of the case should not be sufficient to demonstrate lack of proportionality A claim of non-proportionality should be supported by an analysis that demonstrates the lack of value in acceding to the request, the cost of complying with the request, while keeping in mind the likely value of the case..

If "The beginning of wisdom is the definition of terms", the New Framework Falls Short

"The New Framework modelling is based on *standard inputs*, including unit costs and degrees of *burden* for discovery under *typical* circumstances projected by the team of experts" (emphasis added). And yet, the authors admit that "the inputs are highly dependent on individual circumstances…" The obvious tension here underscores the inherent difficulty present when the New Framework relies upon and uses terms like "standard," "burden," and "typical" throughout.

For example, "burden" and "importance" are subjective terms, not wholly objective concepts. Burden and importance are not subject to ordinal rankings within matters. Rather, burden and importance are best viewed as continuous values, the ranking of which can be in constant flux as

a matter evolves, and of course, have no relationship to each other across matters – a necessary function of a "framework." A "burdensome" data store may have de minimis value (i.e., undue burden) in one matter and be central (high importance necessitating undertaking that "burden") in another matter. Similarly, data stores believed to be of no value at the commencement of a matter may take on primary importance as a case evolves.

There is little agreement on the concepts of "standard," "burden," "typical" or "atypical". The use of these terms throughout The New Framework understates the complexity of what is described. Because experts and lawyers disagree on the meaning of these terms, their pervasive presence in the New Framework is problematic.

The aforementioned challenges manifest themselves in several ways in sections 01-05 of the New Framework.

SECTION 01: CUSTODIANS

Section 01, which "sets out objective criteria used for guidance to prioritize custodians. In the end, all the decisions represent judgment calls *by those making them*…" intertwines both objective and subjective criteria. (P 12). Absent objective criteria that can be universally applied, there is no true framework, but rather, a set of tools that may or may not be useful in each circumstance, and entirely dependent upon the individual matter at hand.

To this end, Section B sets out four excellent criteria for prioritizing (or demonstrating the marginal utility of) discoverable information: materiality, strength, uniqueness, and likelihood of finding significant discoverable information on a particular data source. (P. 14). Tellingly though, only "uniqueness" and "likelihood" are objective as they are quantifiable.[1] "Materiality" and "strength" are not qualifiable, they are subjective judgements.

The New Framework recognizes the tension between establishing a "framework" versus a set of tools that may be useful given particular circumstances in its discussion of a "Standardized Report Format." The Framework states "no single format can effectively handle all cases", yet goes on to say, "this can then be used as a template for the report" (P. 15).

The New Framework itself calls into question the efficacy of its core goal when it excepts "cases involving a multitude of custodians".[2] (P. 15). The New Framework presents as **"designed to be *universal and to apply to all cases large and small*, including complex commercial litigation (B2B), single plaintiff, class action, and investigations/trade secrets matters"** (emphasis added). (P. 10). If it is truly universal, then the templates and standards that make up the Framework must be universally applicable. Again, this is the power and benefit of a "framework." If it is not universal, and rather, applies only to a subset of litigation, then The New Framework should clearly define the bounds of its applicability.

Section 01 closes with a description of what the Framework considers "atypical use cases" that "will require different handling". (P. 15). The examples of "atypical" are highly debatable. While they may not fit into the custodian assessment pro forma provided in Appendix A, they cannot be called atypical:

[1] "Uniqueness" and "likelihood" can be considered objective, as it is quantifiable and demonstrable using, for example, statistically valid data sampling techniques. The defensible argument here is one of marginal utility: That is to say, there is no (or too little) value to producing the ESI from a particular data source or custodian, as either none can be found (assuming search criteria have been agreed) or it is duplicative in content (based on agreed search criteria) of other sources produced. All other factors are highly rebuttable subjective judgments.

[2] What exactly is a "multitude" of custodians? More than 5, 10, 30, 100?

- **Departed Employees:** Given (1) that the U.S. workforce is highly mobile, and (2) litigation is based in the past, to designate departed employees as atypical is not reasonable.
- **Non-custodian Data:** In many matters non-custodial data is highly relevant and anything but atypical.[3] HR policy and procedure documents, training manuals, product manuals and descriptions, employee policies, marketing material, data handling policies and procedures, and myriad others (along with centralized system data) are typical to any number of matter types from contractual disputes to employee discrimination, to trademark and patent infringements, and so on.
- **International Custodians:** If the Framework presented is designed to apply to complex commercial litigation, it is difficult to conceive of international custodians being "atypical." Cross-border disputes have been increasingly common for some time, not to mention employee transfers, secondments, and other factors. The complexities of data privacy regulations do not render international custodians *atypical* or irrelevant.

If these are to be considered "atypical" then The New Framework is only applicable to small, easily defined cases, with few custodians that are all currently employed by the responding party, where non-custodial data is immaterial, and no custodian or data is located outside the U.S.

SECTION 02: DEFINING DATA-SOURCE BURDEN AND EFFORT

"The New Framework's standard approach requires the designation of data sources by degree of burden." (P. 16) This reorients the interpretation to 26(b)(1) away from the importance of the issues at stake and the importance (and relevance) of the discovery in resolving the issues. Rather, the focus is now exclusively on "how much data can I get based on the amount in controversy?"

Section 02 lists "specialized data sources not considered", but these sources must be addressed in an increasing number of cases, especially given the movement of data to the cloud and the ubiquitous nature of self-developed databases.

By way of example, backup media may indeed be unduly burdensome and easily shown to be entirely unnecessary, but if the central actor in the matter is separated from the responding company prior to any preservation, then the backup taken in the ordinary course of business may be crucial, regardless of increased burden. In other words, what The New Framework presentation hardens into a low priority custodian and low priority data, is now a high priority custodian and high priority data.

The insistence on transforming common ESI sources in particular matter types into the atypical, or outliers to be addressed by some other means hobbles The New Framework. Consequently, the Burden Assessment Tool (Appendix E) and its ability to "automatically assign" low/medium/high/highest burden classifications is of limited utility. (P. 39). The burden of collecting, processing, and reviewing, for example, text messages or chat and the emojis contained within them does not change based on matter type and issues. Its importance does.

In the final analysis, we believe that the "burden assessment" has it backwards. Under a marginal utility regime, the designation of data sources must first be a function of importance of the source to understanding the matter and/or to the merits of the claims and defenses. Only after this analysis is done, can the degree or magnitude of burden be assessed.

SECTION 03: DISCOVERY COST PROJECTIONS

[3] Further, the rise and prevalence of collaborative documents wherein multiple custodians work in the same document at the same time will require analysis of what we consider "non-custodial" moving forward.

The New Framework is right in noting the "wide ranges in per-gigabyte [or] per-document discovery costs." And that "many variables complicate the analysis…" It attributes "65%-75% to attorney review." (P. 20). However, that the percentage distributions offered for attorney review, collection, and processing are only very broadly illustrative.

Thus, the statement that "These proportions offer a useful crosscheck in verifying estimates" is debatable as it implies that budgeting estimates that don't conform to the breakdown are problematic. There are any number of variables in a matter that can materially alter these percentages (for example, significantly increased collection costs, but not impact processing, hosting, or review costs.)

Also, to say that "the hourly or per-device rates for data collection have remained relatively flat over the period 2004-2020", is counterfactual. (P. 21). If anything, the exponential increase in competition among service providers during that period has generally lowered collection rates. Similarly suspect, the Framework notes that volume of information being collected has been reduced during this period. If anything, the sheer explosion of data has dramatically increased the volume of data collected, even when accounting for an increased focus on proactive information governance.

Importantly, this begs the question: as eDiscovery pricing continues its downward trajectory (or the amount of data per custodian or device decreases) does this translate into ever broader discovery requests being acceptable? The Framework implies this to be the case if total cost as a percentage of amount in controversy has supremacy.

SECTION 04: PROJECTING A REBUTTABLE, PRESUMPTIVELY REASONABLE DISCOVERY-RATIO RANGE BASED ON "AMOUNT IN CONTROVERSY" AND "IMPORTANCE OF ISSUES AT STAKE" PROPORTIONALITY FACTORS

This section opens with an about-face by rediscovering the importance of "importance." And, in arguing against a marginal utility approach states that "Although this [the marginal utility] argument can resolve many, perhaps most, discovery disputes, Rule 26(b)(1) has a broader scope". (P. 24).

Even accepting that 26(b)(1) has a "broader scope", if the marginal utility approach resolves most discovery disputes, it serves its purpose.[4] And, if marginal utility works in many if not most cases to solve the problem, in which circumstances should The New Framework be given precedence?

The Framework then promptly jettisons the importance factor and focuses solely on the proportionality component of 26(b)(1): "A comparison of discovery costs and the amount in controversy results in a ratio, which can inform the proportionality analysis" which it notes is a "subjective judgment, which is fraught with difficulty and imprecision", (p. 24), and worse, "vulnerable to gamesmanship and a fair degree of skepticism…subject to self-policing." (Pg. 25).

And the goal of all this subjectivity, imprecision, and gamesmanship is not actually to argue a demonstrable and quantifiable undue burden (which marginal utility can and does), but rather simply to arrive at "a workable 'discovery budget'…" (P. 24).

[4] Besides, marginal utility ignores neither amount in controversy nor importance. It is premised on it. Marginal utility by definition is the ability or inability to satisfy a need by consuming more of something (in this case, discovery). It is a calculation of proportion concerning both importance and amount in controversy (e.g., worth).

Having jettisoned all 26(b)(1) factors other than a budget v. amount in controversy The New Framework rests its argument on the *2010 Litigation Costs in Civil Cases: Multivariate Analysis* which found that a "1 percent increase in the amount in controversy was associated with a 0.25 percent increase in total litigation costs". (P. 25).

The New Framework then works through a series of percentage calculations, estimates, and calibrations and then arbitrarily assigns an importance to them. While the New Framework cautions that "If different criteria are employed, the party must be prepared to defend them", the New Framework fails to justify or defend its own criteria.

The calculations discussed in Section 4 rest heavily, if not entirely, on the FJS Multivariate Analysis, so accordingly, its findings are discussed below.

LITIGATION COSTS IN CIVIL CASES: MULTIVARIATE ANALYSIS REPORT TO THE FJC (LEE III AND WILLGING, 2010)

We are not statisticians and do not purport to question the methodology used or analysis provided in this research. Instead, because The New Framework relies on this work for its most foundational premise – calculating an appropriate amount in controversy/eDiscovery ratio to arrive at a proportionality framework – we consider if this study is appropriate for the task. Several factors indicate that it may not be.

The research is done using "cases terminated in the fourth quarter of 2008". (P. 1).[5] This means that the cases under analysis date to when electronic data discovery was just emerging as a critical factor in litigation. 2008 was a radically different eDiscovery environment than today, and this alone calls into question its usefulness when projecting findings against the current environment.

Of the 16 hypotheses H_1 (high stakes), H_5 (ESI disputes), H_{10} (factual complexity), H_{11} (contentiousness among parties), and H_{14} (nonmonetary stakes) are the most pertinent to the New Framework. (P. 3).

The study apparently isolates the tested hypotheses by removing other factors to best attribute cost impact solely to the single hypothesis being considered. Unsurprisingly, they report that each of the above noted factors will increase costs, all else being equal. As we know, however, when factual complexity increases, or stakes increase, or contentiousness increases, all else is rarely equal.

It is difficult to infer from the study how these aspects combine and to what degree they raise the cost of discovery in a contemporary scenario beyond the 0.25:1 factor relied upon. Remember too, that the study's cost comparisons are cases without ESI discovery and include a broad range of cases that today would have significantly different eDiscovery profiles.

Finally, the discovery "level" was measured as an ordinal-level variable as were case complexity and contentiousness on a seven-point scale (pg. 9). It is difficult to imagine useful data regarding eDiscovery costs and the impact of the variables in today's environment based on such a ranking and the timeframe in which it was made. Perhaps alternative sources be found to support the Framework's proposition.

SECTION 05: EFFECTIVE USE OF HEAT MAP AND DISCOVERY ROADMAP

In this section, The New Framework returns to a marginal utility argument and the primacy of burden over importance: "A party can decide to add or eliminate custodians or reduce or add more

[5] All page references in this section refer to the study unless otherwise noted.

gigabyte volume per custodian or data source of discovery, based on the discovery roadmap". (P. 28).

The Framework lists "prioritization of custodians and their data sources by importance of information resolving the issues" but suggests several things that conflict with its premise. (P. 28). Several aspects of the guidance call the Framework's usefulness into question:

1. It restricts the analysis to the "five common data sources" which would be helpful only in a subset of matters.
2. It introduces the concept of adding or reducing GB volume/custodian. The Framework goes on to explain that, for example, this could be "shortening the period of time under review". (P. 29). The time period under review is a rather rudimentary concept that should be well-defined prior to collection. A respondent cannot simply shorten it for budgetary purposes.
3. The application of averages (whether by cost/GB or volume) speaks to the "budget forecasting" focus of the Framework as opposed to understanding likely actual costs. It is hard to imagine the requestor accepting budget forecasts based on industry averages as opposed to actual costs for discussion purposes, let alone resolving discovery scope disputes.
4. "Cumulative discovery costs" for selected custodians has little meaning beyond an illustrative budget unless one can simply trade "expensive" custodians or restrict data discovery concerning them for less expensive ones. This also demonstrates why averages are not useful. High priority custodians – and their data volumes – will likely always incur more expense than low priority custodians for any number of reasons. In short, calculating average costs on "standard" data based only on volume estimates (even if based on the party's own historical averages only provides an initial budget illustration. One fails to see how this then elevates to an argument of proportionality, let alone importance, which surely demands proof of actual cost against the value of that data to importance of issues and resolution. (P. 47)

We agree that the New Framework does "better inform Rule 26", "help parties better craft an ESI protocol", and "develop a phased-discovery plan." It will also "better inform cost estimates when negotiating production format decisions". (P. 31). All very important.

However, the New Framework does not provide "key data points to meet its 'burden of making specific objection…'" (P. 31). Historical averages and estimates lack the specificity required for this. Neither averages nor estimates are specific. Because of this, The New Framework contention that the "reasonable average of discovery costs typically incurred…frames proportionality assessments" is easily rejected. (P. 32).

CONCLUSION

Again, we sincerely thank the contributors who drafted The New Framework and hope the criticisms offered are taken in the spirit intended: to advance the conversation about discovery proportionality to help both sides reorient to litigate cases on their merits rather than on discovery battles. If the New Framework is seen as another useful tool and a reference, it is a welcome one. But, no framework, not new, not old can be a one size fits all solution.