*Draft of October 8, 2019 – Please do not circulate*

# *Incentivizing Lawyers as Teams*

Edward Cheng, Paul H. Edelman, & Brian T. Fitzpatrick

## I. Introduction

The study of law has long been interested in incentives, usually in the form of how judges or legislators can best incentivize the rest of us to take desired behaviors. This, of course, has been the contribution of law and economics to our field. The payment of lawyers has been no different. Although judges and legislators need not often decide how lawyers should be paid because lawyers and their clients can work such things out for themselves in the marketplace, sometimes they do, such as when judges have to appoint lawyers for litigants who are unable to select lawyers for themselves. This happens when litigants are minors, incapacitated, absent class members, or have their cases consolidated in multidistrict litigation. If judges cannot closely monitor the work of these lawyers and instruct them as the absent litigants would instruct them, then the judges have to rely on something else to ensure that the lawyers do what the absent litigants would want them to do. For many decades now, the answer encouraged by scholars has been to rely on economic theories of incentives.

With rare exception, however, the models scholars encourage judges to use have always assumed that the judges are hiring only one lawyer or one pre-existing law firm. That is, scholars have focused on single-principal-single-agent economic models. Thus, for example, the leading theory of how judges should pay lawyers recommends paying a percentage of what the lawyer recovers *plus* an hourly rate for the time expended to get the recovery. This formula takes advantage of the best of the contingency system—incentivizing the lawyer to recover as much as possible—while mitigating the worst of the system—premature settlement.

But sometimes—most frequently in multidistrict litigation—judges do not hire only a single lawyer or a single law firm. They assemble a *team* of lawyers or law firms. The lawyers are selected despite having no preexisting relationship among them and asked to work together for the common good of the relevant litigants. The question we ask in this paper is: how should judges pay *teams* of lawyers?

The standard one-principal-one-agent model does not answer the question. That is, per the above, perhaps the team qua team should be paid a percentage of what the team recovers plus for the number of hours the team worked, but then how should that be allocated among each member of the team? Should each member of the team share equally? Then what incentive would any member have to do anything? Should each member be paid in proportion to his hours? Then what incentive would any member have to work efficiently? The one-principal-one-agent model breaks down when there are multiple agents.

This is where we step in. There is an extensive literature in economics and industrial organization that attempts to answer the question of how a principal can best incentivize a team of agents. In this article, we bring this literature to bear on how judges should compensate the teams of lawyers they hire, with a special focus on the context where they do it most often: multidistrict litigation.

In the course of doing so, we find even this literature lacking in one important respect: most results from one-principal-multiple-agent models suggest that it would be best to compensate the agents in part based on the relative contribution each agent made to the team's output. But, with rare exception, the literature does not tell principals how they might go about figuring out what those relative contributions are. In this article, we instruct principals how to do this. We offer two formulas whereby judges can ask team members to report on *each other's* relative contributions but not their own contribution. Our formulas prevent team members from manipulating their own allocations and colluding on the allocations of others. We believe these formulas not only fill an important gap in the preexisting literature but can give judges in multidistrict litigation a practical way to properly incentivize the members of the teams they assemble.

This is not a small matter. Over half of all civil litigation pending in federal courts right now has been rolled into multidistrict litigations. Without sound theory behind them, judges have devised ad hoc compensation schemes and these schemes are often deeply controversial and lead to prolonged satellite litigation over fees. It is well past time to bring order to this massive piece of our system of justice. It is our hope and belief that bringing the unexplored economic theory to bear on this question and devising practical ways to implement this theory can settle these controversies and improve outcomes for multidistrict litigants and lawyers alike.

## II. When Judges Assemble Teams

Typically in our system of justice, clients hire their own lawyers and decide for themselves how to pay them. But not always. For example, it is well known that judges must appoint lawyers for minors, incapacitated persons, and absent class members. When judges make these appointments (at least outside multidistrict litigation), judges usually appoint only one lawyer or one pre-existing law firm.

But sometimes judges decide to appoint *teams* of lawyers. These teams are literally assembled by the judge from scratch; the lawyers are often from different firms and have no preexisting contractual relationships with one another. Sometimes judges do this outside multidistrict litigation, but it is very common within multidistrict litigation. Under the federal Multidistrict Litigation Act, any time more than one case presenting the same factual question is filed in a federal court, a special panel of federal judges can decide to consolidate and transfer all such cases to a single federal judge for pretrial purposes. This can be hundreds, thousands, or even tens of thousands of cases. But each of these cases will already have its own lawyer, which can mean hundreds, thousands, or even tens of thousands of lawyers to deal with. What does the judge do? The judge assembles a small *team* of lawyers to handle all of the pretrial work that is "common" to the consolidated cases: they draft and oppose dispositive and non-dispositive motions on common issues, take discovery on common issues, litigate class certification where applicable, and negotiate settlements. Many of the decisions this small team makes will bind all of the litigants in the multidistrict litigation either as a technical matter or as a practical matter despite the fact the litigants may not have consented to (or even opposed) representation by the judge's team.

No one on these teams appointed to do work on behalf of the common good expects to do it out of the goodness of his heart—i.e., for free. Indeed, they probably wouldn't take the appointment, or, if they did, they probably wouldn't do a very good job, if they were expected to work for free. Thus, as in other contexts where judges appoint lawyers for others, the judge must find some way for the litigants in multidistrict litigation to compensate the judge's team. As we explain later, however, judges have largely been left to do so on their own without the benefit of the economic theory they have come to rely upon in other contexts.

What have judges come up with in the meantime? The current state of the art is for judges to set aside a percentage—usually in the 6-12% range—of any settlements that arise from the multidistrict litigation and use those monies to pay the team.[1] How do they allocate these funds? They rely on a variety of methods. Sometimes the judges defer to one lawyer on the team or a small subcommittee of the team to do it and if none of the other team members objects, then that allocation sticks.[2] Sometimes—and if someone does object under the other methods—the judges do it themselves by asking the lawyers to describe what they did to contribute to the litigation.[3] But no matter which of these methods is used, the factor most important in everyone's mind is the same: each team member's relative "lodestar"—the number hours the team member worked multiplied by his normal hourly rate.

Many people are unhappy with the state of the art. As we noted above, it is well known that paying lawyers based on their relative lodestars gives them an incentive to be inefficient with their time. In addition, deferring to one member or a small subcommittee of the team to divide the pie—including their own pieces—empowers the few to enrich themselves at the expense of the many. It is no exaggeration to say that the allocation of lawyers' fees in multidistrict litigation has led to fights that are even more ferocious than the fights the same lawyers wage against the defendants they sue.[4]

---

[1] See William Rubenstein, 5 Newberg on Class Actions § 15:117 (5th ed.).

[2] *See* Victor v. Argent Classic Convertible Arbitrage Fund L.P., 623 F.3d 82, 85 (2d Cir. 2010) (providing lead counsel with discretion to allocate the attorneys' fees); In re Vitamins Antitrust Litig., 398 F. Supp. 2d 209, 213–14 (D.C. Cir. 2005) (giving the three lead co-counsel authority to split the attorneys' fees into three pools and allocate to plaintiffs' lawyers accordingly); In re Initial Public Offering Sec. Litig., No. 21 MC 92(SAS), 2011 WL 2732563, at *1 (S.D.N.Y. July 8, 2011) (describing the method where one firm acted on behalf of a committee of lead counsel in order to propose a fee allocation); *see also See, e.g.*, Order and Reasons Allocating Common Benefit Fees in the Knauf Aspect of This Litigation, In re Chinese Drywall Prods. Liab. Litig., No. 2:09-md-02047-EEF-JCW, at 7-8 (E.D. La. Feb. 4, 2019) (describing the method of appointing a Fee Allocation Committee to recommend the appropriate allocation for the team of attorneys).

[3] *See* In re TFT-LCD (Flat Panel) Antitrust Litig., No. M 07-1827 SI.MDL. No. 1827, 2013 WL 1365900, at *7–16 (N.D. Cal. Apr. 3, 2013) (allocating the attorneys' fees with assistance by a Special Master). For examples where the court allocated the funds after objections are made by some of the lawyers, see *Victor*, 623 F.3d at 84, *In re Vitamins Antitrust Litig.*, 2011 WL 2732563, at *214, and Order and Reasons, *supra* note 2, at 8–9.

[4] *See* Charles L. Becker, Shanin Specter & Thomas R. Kline, *How Not to Manage a Common Benefit Fund: Allocating Attorneys' Fees in* Vioxx *Litigation*, 9 DREXEL L. REV. 1 (2016); Amanda Bronstad, *Critical Mass: Class Needs Replacement as Embattled Michael Avenatti Exits. Plus, Fee Battle Roils Mesh Litigation*, LAW.COM (Apr. 3, 2019, 12:15 PM),

Our hope and belief are that bringing the economic and industrial organization literature on incentivizing teams to multidistrict litigation—and improving upon this literature with our own contribution—can change all this.

## III. The Existing Literature on Incentivizing Lawyers

There is an extensive literature in law and economics on how clients should compensate their lawyers when they cannot closely monitor them. One example of this situation is when judges hire lawyers for others. Obviously, the absent litigants cannot monitor the lawyers hired by the judge; if they could, there would be no reason for the judge to do the hiring. But even the judges are thought to be poor monitors: judges have a lot to do and they are largely reactive and dependent on the adversarial process to bring matters to their attention and to structure their decision making. When judges hire lawyers for others, this process is largely absent. Thus, scholars often recommend that, where possible, judges rely on the invisible hand of incentives to make sure these lawyers do a good job for their clients.

The central concern in this literature is agency costs: how does the principal stop the agent from doing what is good for the agent rather than for the principal when the principal cannot monitor the agent? The literature suggests the following:

- It is better to pay the lawyer a percentage of what the lawyer recovers rather than by the hour, otherwise the lawyer will work longer than the client would want and have no reason to recover as much as possible for the client.[5]

---

https://www.law.com/2019/04/03/critical-mass-class-needs-replacement-as-embattled-michael-avenatti-exits-plus-fee-battle-roils-mesh-litigation/; Amanda Bronstad, *'Structural and Procedural Flaws' Foul Up Fees in Syngenta Settlement*, LAW.COM (Aug. 28, 2019 7:55 PM), https://www.law.com/nationallawjournal/2019/08/28/structural-and-procedural-flaws-foul-up-fees-in-syngenta-settlement/.

[5] *See* In re Synthroid Mktg. Litig., 325 F.3d 974, 979 (7th Cir. 2003) ("Contingent-fee arrangements are used when it is difficult to monitor counsel closely; otherwise some different arrangement, such as hourly rates, is superior."); Richard A. Epstein, *Class Actions: The Need for a Hard Second Look* 9, CIVIL JUSTICE REPORT (Mar. 2002) ("One possible fee arrangement is for the informed client who is capable of monitoring his lawyer's work to pay an hourly fee. . . . But whatever its good sense in ordinary litigation, it is stillborn in class action litigation for the simple reason that diffuse class members have no real way to monitor the behavior of their lawyers. The great danger is that lawyers will pad their bills and leave the shareholders . . . empty-handed.").

- But paying the lawyer a percentage presents a problem: if the percentage is less than 100%, then the lawyer will want to work less than the client would want the lawyer to. This is because the lawyer bears 100% of the opportunity cost of working but benefits less than 100% from that work. This leads to premature settlement.[6]

- Various solutions to premature settlement have been proposed. One solution is to increase the lawyer's percentage as the recovery is improved. But this does not completely solve the problem.[7] A better solution is to pay the lawyer both a percentage *and* by the hour, but even this does not entirely solve the problem unless a complicated third-party insurance scheme is employed.[8]

- Nonetheless, the leading theory on how to pay a lawyer when the client cannot monitor closely is with a percentage of the lawyer's recovery *plus* the lawyer's usual hourly rate.

- There is an easy way to figure out the lawyer's usual hourly rate. But what about the percentage? The leading theory here is to hold an auction to determine which lawyer is willing to take the representation for the lowest percentage.[9]

All of this is well and good. But all of it assumes that the judge is hiring only one lawyer. What is the judge to do if the judge wanted to assemble a

---

[6] *See* STEVEN SHAVELL, FOUNDATIONS OF ECONOMIC ANALYSIS OF LAW 435 (2004) ("Under contingency fee arrangements, the lawyer's incentives are different, leading the lawyer to press for settlement more often than when the settlement offer exceeds the expected judgment net of litigation costs, because the lawyer bears all the litigations costs but obtains only a percentage of the settlement."); Epstein, *supra* note 6, at 9 ("It is well known that in some cases the ordinary contingent fee lawyer will settle a case *sooner* than might be in the best interest of his client, and do so in order to reduce the costs of additional litigation (which fall largely or entirely on him).").

[7] *See* John C. Coffee, Jr., *Understanding the Plaintiff's Attorney: The Implications of Economic Theory for Private Enforcement of Law Through Class and Derivative Actions*, 86 COLUM. L. REV. 669, 697 (1986).

[8] *See* Kevin M. Clermont & John D. Currivan, *Improving on the Contingent Fee*, 63 Cornell Law Review 529, 546-50 (1978). For the variation that perfects the formula, *see* A. Mitchell Polinsky & Daniel Rubinfeld, *Aligning the Interests of Lawyers and Clients*, 5 American Law and Economics Review 165, 166-69 (2003).

[9] *See* Jonathan R. Macey & Geoffrey P. Miller, *The Plaintiffs' Attorney's Role in Class Action and Derivative Litigation: Economic Analysis and Recommendations for Reform*, 58 U. CHI. L. REV. 1, 106–08 (1991); *see also* Epstein, *supra* note 6, at 7 (endorsing the auction approach).

team of six unrelated lawyers? The results above tell the judge how much the team qua team should get: they should get some percentage plus their normal hourly rates. But how much should each lawyer on the team get? Maybe the lawyers can work this out for themselves, but maybe they can't. All of the analysis discussed above is based on one-principal-one-agent models; none of it considers what to do when the principal is hiring multiple agents.

This is a massive oversight for the reason we stated above: more than half of all federal civil litigation is wrapped up in multidistrict litigation where judges frequently assemble teams to handle the pretrial work for all these cases.

To our knowledge, there is only one paper that tackles the question of how best to pay a team of lawyers. The paper is a very good one by Charlie Silver at the University of Texas and Geoff Miller at NYU.[10] They even address the multidistrict litigation context directly. They recommend what we will describe below as the creation of a monitor who holds a residual on the team's profits. In particular, they recommend that whichever lawyer has the greatest number of clients in the multidistrict litigation be appointed to assemble, oversee, and decide how to pay the lawyers who do the pretrial work. The catch is that the lawyer so appointed cannot assign the pretrial work to himself. But because he has the biggest stake in the litigation, he will both enjoy the biggest upside when the team he assembled does well as well as pay the biggest share of the fees for the work those lawyers do. This gives him the incentive to pick, compensate, and monitor wisely. It is an ingenuous solution. But no judge has been attracted to it because the lawyer with the biggest portfolio of clients is almost always a lawyer who generated those clients by advertising on television or the like—not necessarily by being a crackerjack litigator—and no judge thus far has wanted to give a lawyer like that control of an important case.

But, as we show below, creating a monitor with a residual on the team's profits is only one of several ways that economists and industrial organization scholars have shown that a team can be optimally incentivized. We believe judges may find more appealing some of the other options. In the next section, we bring this literature to the law for the first time and explain how judges can use it to compensate lawyers in multidistrict litigation. As we will see, even

---

[10] Charles Silver & Geoffrey P. Miller, *The Quasi-Class Action Method of Managing Multi-District Litigations: Problems and a Proposal*, 63 VAND. L. REV. 107 (2010).

this literature is missing an important piece and we will finish this article by filling in that piece.

## IV. The Literature on Incentivizing Teams

Largely unbeknownst to legal scholars, there is an extensive literature in economics and industrial organization that seeks to devise compensation schemes to overcome the problem of agency costs in one-principal-multiple-agent contexts. This literature is quite nuanced and the results that it generates depend on various assumptions that are inputted into the models. As such, it is impossible to summarize it all here. Rather, we set forth below the findings that we believe are most relevant our focus: the judges who assemble lawyer teams in multidistrict litigation:

- If the work product of each member of the team is dependent on how well other members of the team do their work—e.g., if the lawyer who is handling the summary judgment papers will do a better job if the lawyer handling discovery does a better job, something we think is true—then each member of the team should be compensated based on how much the *other* members of the team contribute.[11] One way to do this is to compensate each member based on how much the team qua team produces. That is, to compensate each team member more the more the team produces.

- But if team members are compensated only by how well others do, then they may shirk on what they contribute.[12] This is a classic moral hazard problem. Thus, team members should be compensated

[11] See Stanley Baiman & Joel S. Demski, Economically Optimal Performance and Control Systems, 18 J. Acct. Res. 184, 198-99 (1980); Bengt Holmstrom, Moral Hazard in Teams, Bell J. Econ. 324, 335; Hal R. Varian, Monitoring Agents with Other Agents, 146 J. Inst. & Theoretical Econ. 153, 153 (1990).

[12] See Armen A. Alchian & Harold Demsetz, Production, Information Costs, and Economic Organization, 62 Am. Econ. R. 777, 780 (1972); Edward P. Lazear & Sherwin Rosen, Rank-Order Tournaments as Optimum Labor Contracts, 89 J. Pol. Econ. 841, 842 (1981); Dilip Mookherjee, Optimal Incentive Schemes with Many Agents, 51 R. Econ. Stud. 433, 443-44 (1984) (showing that agents may collude to shirk if they are paid only based on ordinal rankings); Claudia M. Landeo & Kathryn E. Spier, 119 J. Econ. Behavior & Org. 496, 497 (2015).

based *both* on how much the team does *and* how much they contribute to the team's output relative to others.[13]

- It is often easy enough to determine—especially in multidistrict litigation—how much the team qua team produces: these litigations usually result in the recovery of a pot of money for the plaintiffs. The difficulty is figuring out the relative contribution of each team member. Ideally, team members will be compensated in proportion to the *distance* between their relative contributions—e.g., a lawyer that is *twice* as responsible as another for the team's production should be awarded *twice* the compensation on this component of the formula—but, if that is not possible, even imprecise measures of relative contribution like an ordinal ranking of who contributed the most may be just as good, or, at the very least, are usually better than nothing.[14]

- How can principals figure out, even imprecisely, what each team member's relative contribution was? Most of the papers in this literature are unhelpful; they assume some "technology" will exist to do it. But the most helpful papers suggest one of three methods: 1) to permit team members to decide amongst themselves what their relative contributions were;[15] 2) to give one team member a residual (i.e., what is left over after the others are paid) and ask him to

[13] See Hideshi Itoh, Cooperation in Hierarchical Organizations: An Incentive Perspective, 8 J. L. Econ. & Org. 321, 322-23 (1992); Ines Macho-Stadler & J. David Perez-Castrillo, 11 Intl. J. Indus. Org. 73, 90-91 (1993); Yeon-Koo Che & Seung-Weon Yoo, Optimal Incentives for Teams, 91 Am. Econ. R. 525, 538 (2001); Enno Siemsen et al., Incentives that Induce Task-Related Effort, Helping, and Knowledge Sharing in Workgroups, 53 Mangement Sci. 1533, 1533, 1535, 1539, 1546 (2007); Alex Gershkov et al., How to Share it Out: The Value of Information in Teams, 162 J. Econ. Theory 261, 262-63 (2016); Tim Baldenius et al., Relational Contracts with and between Agents, 61 J. Acct. & Econ. 369, 370, 381 (2016).

[14] See Edward P. Lazear & Sherwin Rosen, Rank-Order Tournaments as Optimum Labor Contracts, 89 J. Pol. Econ. 841, 863 (1981); Jerry R. Green & Nancy L. Stokey, A Comparison of Tournaments and Contracts, 91 J. Pol. Econ. 349, 349 (1983); Bengt Holmstrom, Moral Hazard in Teams, Bell J. Econ. 324, 335-337; Alex Gershkov et al., How to Share it Out: The Value of Information in Teams, 162 J. Econ. Theory 261, 271-282 (2016).

[15] See Hideshi Itoh, Coalitions, Incentives, and Risk Sharing, 60 J. Econ. Theory 410, 425 (1993). Some papers find that delegation undermines the ability of the principal to induce optimal performance because the agents can collude against the compensation scheme. See Bengt Holmstrom & Paul Milgrom, Regulating Trade Among Agents, 146 J. Inst. & Theoretical Econ. 85 (1990). It is not clear to us if this is possible if the pool of compensation starts from a percentage of the team qua team's output.

monitor and pay the others;[16] and 3) rather complicated, multi-stage techniques to induce team members to make honest reports of each other's contributions.[17]

In summary, the results from the literature suggest that team members should be paid based in part on how much the team produces and in part based on some measure of how much they contributed relative to the others. Other than the complicated, multi-stage techniques, however, the literature offers only vague aspirations as to how those relative contributions can be determined. But all the papers—even those with only vague aspirations—suggest one thing in common: relying in some way on team members themselves to determine what their relative contributions were.[18]

Can any of this be made useful to multidistrict litigation? We think so.

To begin with, it is encouraging to note that much of what judges do now is consistent with this literature. For example, judges usually start with a pool of attorney's fees equal to a *percentage of what the team recovers* for the plaintiffs that they then allocate in way that is designed to be *related to relative contribution*. This captures both a team-based component and a relative-contribution component.

Moreover, the methods judges use to calculate the relative-contribution component are consistent with the literature, too. As we noted, judges will often defer to the team itself if it can work out its own allocation. When judges do not or cannot—as we noted above, these allocations are often controversial and not all team members will agree with an allocation proposed by other team members—judges base the allocation on each team member's relative lodestar (i.e., hours worked multiplied by their normal hourly rate). Relative lodestar

[16] See Armen A. Alchian & Harold Demsetz, Production, Information Costs, and Economic Organization, 62 Am. Econ. R. 777, 782 (1972).

[17] See Ching-to Ma, et al., Stopping Agents from "Cheating," 46 J. Econ. Theory 355, 363 (1988); Hal R. Varian, Monitoring Agents with Other Agents, 146 J. Inst. & Theoretical Econ. 153, 165 (1990).

[18] See Eugene Kandel & Edwad P. Lazear, Peer Pressure and Partnerships, 100 J. Pol. Econ. 801, 811 (1992); See Hideshi Itoh, Cooperation in Hierarchical Organizations: An Incentive Perspective, 8 J. L. Econ. & Org. 321, 325 (1992).

is a proxy for relative contribution, and, as we noted, even imprecise measures of relative contribution can be used to incentivize multiple agents.

On the other hand, relative lodestar method is a pretty poor proxy for relative contribution—and, the worse the proxy, the further the compensation scheme may be from optimal. As we know, the lodestar method gives lawyers the incentive to inflate the number of hours they say they worked. Moreover, every hour does not have equal impact on a litigation. This is why the lodestar method has been rejected by the law-and-economics literature in the one-principal-one-agent models. As we explain in the next part, we think it should be rejected in one-principal-multiple-agent models as well.[19]

In fairness to the judges, they have tried to improve upon the lodestar method. Judges often attempt to adjust each team member's lodestar to reflect the fact that some work was more important to the case than others. For example, the lawyers who review documents during discovery often see their lodestars discounted relative to lawyers who do other things. But how do judges figure out the relative value of the tasks that make up a litigation? They either rely upon their own knowledge of the case—which we are assuming here to be imperfect because the entire point of this inquiry is to figure out what to do when the principal is unable to adequately monitor the team—or, as we noted, they defer to one member or a small subcommittee of the team to do it. As we explain next, however, deferring to one member or a subcommittee of the team comes with obvious shortcomings we think we can overcome.

## V. How Can Lawyers' Relative Contributions Be Determined?

As we noted above, optimally incentivizing a team of lawyers will require the judge in multidistrict litigation to figure out what each of their relative contributions were to the team's output. But how can the judge figure this out if we take as the opening assumption that the judge is not a good position to monitor his team? The literature we canvassed above largely converges on one idea: peer assessments. One idea is to let the peers sort it

[19] Many papers in the economics and industrial organization literature assume that the ideal would be to base compensation on each agent's "effort." See, e.g., Ines Macho-Stadler & J. David Perez-Castrillo, 11 Intl. J. Indus. Org. 73, 79 (1993). But isn't the number of hours worked equivalent to effort? And, if so, would the lodestar method be a good method according to this literature? The answer is no. The models in this literature assume that a given level of effort necessarily produces a certain outcome: e.g., low effort produces a weak outcome and high effort produces a good outcome. But in the real world there is no necessary correlation in between how many hours someone works and what kind of outcome they achieve.

out on its own. Another idea is to give one peer the residual on the team's output and ask that peer to decide. A final idea is a complicated multi-stage game where peers are induced to honestly report on each other. Below, we assess each of these ideas here and then show how we think we can improve upon them.

The first idea—let the peers sort it out on their own—is fine as far as it goes; as we noted, many judges try to use it already. But it requires the consent of all the team members, and, for the reasons we have stated above, that is not often forthcoming. Rather, fee disputes in multidistrict litigation are some of the ugliest, most ruthless disputes we have ever seen in the law.

The second idea—to give one peer a residual and ask him to decide—is fine as far as it goes, too: we can pay someone to monitor the team and figure out how to pay them to produce the biggest residual. The problem with this idea is the residual: where does this extra money come from? In the existing literature it comes from the principal, but, in multidistrict litigation, the ultimate principals are the plaintiffs. If we take more money from them, we are undermining one of the purposes of multidistrict litigation, which is to try to make the plaintiffs' whole for their injuries; every extra dollar we siphon off to others further defeats that purpose. We suppose it is possible to take the residual from the other lawyers on the team, but, with less money for the monitor to throw around, won't it be harder to induce the team to perform well? We think so. Professors Silver and Miller solve this problem by taking the residual from the lawyer who has the most clients and giving it right back to him again. It is an ingenious solution, but judges have not warmed to it because the lawyers with the most clients are not necessarily the best litigators. It is possible that the judge could assign the residual to a different lawyer, one with fewer clients and a smaller stake but one that the judge might deem more competent at litigation. In theory, we believe this could work just as well as the Silver-Miller solution and that it is worth trying. But any monitor who does not already have a stake in the litigation will cost too much money.

The third idea will not be attractive for multidistrict litigation. Thus far, the multi-stage techniques have only been devised for two- or three-person teams.[20] But multidistrict litigation teams are much bigger. In addition, all the techniques require each team member's work to be split into multiple yet simultaneous stages. This is simply not how litigation unfolds. In short, as clever as these techniques are, they are not yet ready for prime time, and, even

[20] *See* Ma, *supra* note 17, at 363; Varian, *supra* note 17, at 165.

if they continue to improve, we suspect judges will always deem them too cute and game-like for the serious business of life and death litigation against the largest corporations in the world.

In short, we do not think any ideas presently in the literature give judges a realistic way to use peer assessment in their multidistrict litigation compensation formulas.

As we noted, judges have not been idle in this vacuum. They have developed their own peer assessment method: they usually take each lawyer's relative lodestar and adjust it based on how important one member or a small subcommittee of the team thinks each lawyer's hours were to the case. What is wrong with this method? What is wrong with it is that the lawyers who are assessing the others are obviously self interested. The more they discount other lawyer's lodestars, the more money is left over for them. Unlike the self-interested monitor with a residual, this self interest is not good. The monitor-with-a-residual has to decide what to pay each of the other lawyers *before* they produce the residual for him: he therefore has to be careful about not being too greedy for himself.[21] By contrast, judges rely on one lawyer or a small subcommittee to decide what their peers get *after* the case is over. It is not hard to see why these sorts of assessments will not be very accurate ones unless the same team will be handling future multidistrict litigation together,[22] which they rarely do. No economic model of which we are aware recommends that some team members decide whatever everyone on the team—including themselves!—get after the team has already finished its work. In short, if we are going to rely on peer assessment, it either needs to take place ex ante, as with the monitor-with-a-residual, or, if it takes place ex post, it needs to be free of self interest.

This is where we think we can make our contribution. We have devised two formulas that allow peers to assess each other ex post but in a way that prevents them from profiting from their own assessments. Before we describe our formulas, however, we should acknowledge that we are aware of two other efforts similar to ours. These papers do not engage any of the above literature because they are not concerned with schemes that will incentivize optimal efforts ex ante but with ex post divisions of windfalls. Nonetheless, we believe these papers could be used in the one-principal-multiple-agent compensation

---

[21] See Armen A. Alchian & Harold Demsetz, Production, Information Costs, and Economic Organization, 62 Am. Econ. R. 777, 782 (1972).

[22] See Claudia M. Landeo & Kathryn E. Spier, 119 J. Econ. Behavior & Org. 496, 497, 509 (2015).

schemes we described above and we described their ideas below. For the reasons we explain, however, we think their ideas are not as helpful to judges in multidistrict litigation as our own formulas.

The first prior effort is from Steven Brams and Alan Taylor.[23] Their work is from what is known as the "divide the dollar" literature: how best to divide a windfall so that everyone goes away happy. The most famous application of this literature is the I-cut-you-choose method for dividing a cake between two people. In one of their books, they address the question of how a team of workers should divide an unexpected bonus. They recommend the following: 1) ask each team member to propose how he would allocate the team's compensation; 2) take the average of these allocations to derive a baseline allocation for each member; and 3) pay each member exactly what he allocated for himself, but do so in the order from the least "greedy" to the most "greedy" (i.e., start with the team member with the smallest difference between his self allocation and what the average allocation from everyone else was). What happens if there is not enough money to pay everyone? The last team member or members are left with only a portion their self allocations or perhaps even nothing at all. Their logic is that this method will cause each team member not allocate too much to himself, driving everyone toward self allocations that match what they think others will want to give them—i.e., honest reports.

Frankly, we think the Brams-Taylor method is brilliant, but we believe it suffers from two features that will limit its usefulness to judges in multidistrict litigation. First, in theory, the method could leave one or more lawyers on the judge's team with nothing—after years upon years of hard work. Maybe this will never happen because team members will be so fearful of being left with nothing that they always request too little rather than too much for themselves. But what if it does happen? What will the judge do? Stick to the formula or make a modification? If the judges will make a modification—and we think they would—then might that make everyone less fearful to begin with, undermining the very discipline the method was supposed to impose on everyone in the first place? We think so. Second, the method assumes that all of the members of the team know enough about each other to rate one another. But what if some members don't know what some other members did? Some multidistrict litigation teams can consist of two or three dozen lawyers and not everyone will know what everyone else has done.

[23] STEVEN J. BRAMS & ALAN D. TAYLOR, FAIR DIVISION: FROM CAKE-CUTTING TO DISPUTE RESOLUTION (1996).

How can you make the baseline from which the payment order is derived if there is missing information? Professors Brams and Taylor do not say.

The second prior effort is from Geoffroy de Clippel and two co-authors.[24] It, too, comes from the divide-the-dollar literature. They recommend asking each team member to rate the relative contributions of each pair of *other* team members. For example, for a team of three members (A, B, and C), A would be asked what the contribution of B was relative to C; B would be asked what the contribution of A was relative to C; and C would be asked what the contribution of A was relative to B. They derive a family of formulae that take these pairwise ratings and produce an overall allocation that makes it impossible for any firm to affect their own allocation in their reports of others.[25] Thus, there is little reason for team members not to give honest reports.

Frankly, we think the de Clippel method is brilliant, too. But, like the Brams-Taylor method, it suffers from a feature that will make it unattractive in multidistrict litigation. Like the Brams-Taylor method, it causes problems for de Clippel if some team members do not know enough to rate all other team members. For the method to work at all, every possible pair of team members must have received at least one relative-contribution report from someone, and, for the method not to leave unallocated money, two reports are needed for every possible pair. We wonder how many multidistrict litigations will be unable to deliver on these demands.

Fortunately, none of the problems with Brams-Taylor and de Clippel are present in our own formulas. Neither of our formulas will leave anyone with nothing unless everyone says that person should get nothing. Both of them require less information to work. Although it is theoretically possible for a team member to affect his own allocation with our formulas, we believe that as a practical matter it will be impossible for anyone to predict how our

---

[24] Geoffroy de Clippel, Herve Moulin & Nicolaus Tideman, *Impartial Division of a Dollar*, 139 J. ECON. THEORY 176 (2008).

[25] For three firms, there is a unique formula: if we let $r_{BC}$ be the report from A about the relative contribution of B to C, and similarly $r_{AC}$ is B's report of the relative contribution of A to C, then de Clippel's allocation to C is $\frac{1}{1+r_{BC}+r_{AC}}$. Note that this allocation is not dependent on any report from C. Similar formulae give the allocations to A and B. For more than three firms, de Clippel derives a family of solutions that are too complex to set forth here.

formulas will affect his own allocation, and, as such, we think our formulas are just as impervious to manipulation as de Clippel's.

Not unlike the prior approaches, both of our formulas use reports from team members about how they would allocate fees among their peers. In particular, we start by asking each member of the team how they would allocate fees among all the other team members for whom they have enough knowledge. Thus, for example, we would ask each member on a team of four to tell us how they would allocate a hypothetical $100 among the other three members—or, if they don't know enough for all three, among as many as they do know. The team members in this example might respond with the following allocation matrix:

Figure 1

| Rating / Rater | Firm 1 | Firm 2 | Firm 3 | Firm 4 |
|---|---|---|---|---|
| Firm 1 | NA | 0.50 | 0.40 | 0.10 |
| Firm 2 | 0.70 | NA | 0.30 | NA |
| Firm 3 | 0.85 | NA | NA | 0.15 |
| Firm 4 | NA | 0.50 | 0.50 | NA |

Our first formula simply uses least-squares optimization to find the overall allocation that mostly closely fits the team member reports. We offer two ways to do this. One way is to minimize the error between the optimized distribution and each of the reports for each pair of firms.[26] The other way is to minimize the error between the optimized distribution and each firm's slate of reports.[27] The advantage of the former is that firms' pairwise assessments may be more reliable. The advantage of the latter is it requires less data to compute an overall distribution. In many cases, the techniques produce very similar results, but a court could use whichever optimization technique it thought was best in the circumstances or it could examine both techniques and

[26] See Section 2.1 of the Technical Appendix for details.
[27] See Section 2.2 of the Technical Appendix for details.

pick an allocation that falls in between them. For the allocation matrix in Figure 1, our first formula produces the following results:

Figure 2

| | Firm 1 | Firm 2 | Firm 3 | Firm 4 |
|---|---|---|---|---|
| Optimum Pairwise | 0.48 | 0.24 | 0.21 | 0.07 |
| Optimum by Firm | 0.48 | 0.24 | 0.21 | 0.07 |

In other words, the court could award 48% of the overall fee pot to Firm 1, 24% to Firm 2, 21% to Firm 3, and 7% to Firm 4.

Our second formula uses Bayesian modeling to estimate the overall allocation. Bayesian modeling views the firms' reports as noisy data and looks for the "true" values within that noise. We again offer two ways to do this. One way is a linear model where we assume the error between the individual reports of a firm's contribution and the "true" contribution are uncorrelated and then run simulations to estimate the allocation.[28] The other way is a compositional model where we allow the errors to be correlated with one another and then run the simulations.[29] Given that allocations are a zero-sum game, the error technically *is* correlated and should not be modelled as independent. Nonetheless, again, the models will often agree with one another. For the fee allocation in Figure 1, our second formula generates the following results:

[28] See Section 3.1 of the Technical Appendix.
[29] See Section 3.2 of the Technical Appendix.

Figure 3

| | Firm 1 | Firm 2 | Firm 3 | Firm 4 |
|---|---|---|---|---|
| Bayes Linear | 0.49 | 0.21 | 0.21 | 0.09 |
| Bayes Composite | 0.41 | 0.28 | 0.24 | 0.06 |

In other words, the court could award between 41% and 48% of the overall fee pot to Firm 1, 21-24% to Firm 2, 21-24% to Firm 3, and 6-9% to Firm 4, depending on which model it selected or whether it wanted to use some sort of average of both models.

Each of the two approaches, optimization and the Bayesian method, have certain advantages. Optimization has the benefit of being clear in its motivation—to find the allocation that minimizes certain measures of error—and so it is (relatively) easy to explain to all of the parties involved. The Bayesian approach, viewing the reports as noisy data and looking for the true values within that noise, has a certain intuitive appeal but it also allows for error estimates in the allocation that may be informative for the parties as way to assess the recommended division.

We think our formulas are just as simple as the methods set forth by Brams-Taylor and de Clippel yet we believe they overcome the limitations of those methods.

First, our methods are easy to motivate and explain. The optimization methods choose an allocation that best fits the reports of the firms. The Bayesian models are the best estimates for the fair allocation given the noisy reports from the firms. We believe that these are standard concepts that firms can understand and appreciate. Moreover, as indicated by our examples, the resulting allocations based on these methods are remarkably consistent with each other.

Second, our methods will rarely, if ever, allocate nothing to a firm unless everyone else says that person should get nothing. This should make our formulas more palatable to courts that the Brams-Taylor technique, which could leave the firm that contributed the most to the litigation without a penny in fees.

Third, each of our formulas requires less information to work than either the Brams-Taylor or de Clippel method. By way of example, neither of their methods are able to solve even the simple four-firm allocation matrix in Figure 1!

There is one area, however, in which our formulas cannot offer any improvement over the Brams-Taylor and de Clippel methods: the latter proved mathematically that it is impossible for one firm to affect its allocation in its reports of the other firms. Like Brams-Taylor, we cannot insure that. But like Brams-Taylor, we think it would be *extraordinary* for a firm to be able to predict how its ratings of other firms will end up affecting its own allocation. The firm would have to be able to predict what other firms were going to say about still other firms. The only way for a firm to know that would be to collude with other firms on their ratings a problem that no current method can avoid.

We are not sure how often collusion will manifest itself because, if everyone's reports are kept confidential, colluders will not be able to verify whether their counterparties followed through, which usually undermines such schemes. But if we wanted to do something to affirmatively prevent collusion, there are many options.[30] One option is simply to prohibit firms from discussing their reports with one another and to punish firms who violate this prohibition. Another option is for the court to randomly drop some reports from the data inputted into the formulas in order to mitigate the effects of collusion—and thereby the incentive to engage in it—by potentially excluding one side of it from the analysis.

More ambitiously, we can modify our Bayesian model to be resistant to collusion by "learning" which firms' reports depart from the general consensus and then giving those reports less weight in the final analysis.[31] For example, suppose we have ten firms, and Firms 4 and 5 decide to collude with each other. Figure 4 shows what ratings Firms 4 and 5 would have given honestly, versus what ratings Firms 4 and 5 give under a collusive agreement to inflate each other's scores. (Note that firms are not allowed to rate themselves.)

---

[30] As we noted earlier, when there are more than four firms, de Clippel's method turns into a family of methods that allocated the whole pot. As they note, some of the members of that family are less prone to the effects of collusion than others. See supra note 24. Nevertheless, such choices are buried enough that it would be hard to make transparent which firms are contributing to the problem.

[31] See Section 3.3 of the Technical Appendix.

Figure 4

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Baseline 4 | 0.42 | NA | 0.28 | NA | **0.19** | NA | 0.11 | NA | NA | NA |
| Baseline 5 | 0.36 | 0.26 | 0.21 | **NA** | NA | 0.17 | NA | NA | .0.79 | 0.083 |
| | | | | | | | | | | |
| Collusion 4 | 0.32 | NA | 0.21 | NA | **0.39** | NA | 0.08 | NA | NA | NA |
| Collusion 5 | 0.21 | 0.16 | 0.13 | **0.39** | NA | 0.10 | NA | NA | NA | NA |

Figure 5 shows the performance of our various methods under collusive conditions. For the baseline (honest) set, we see that all of the methods arrive at similar outcomes. For the collusion set (where Firm 4 and 5 collude), under the optimization and basic Bayesian methods, Firm 4 is able to mildly increase its allotted portion. The resistant Bayesian model, however, all but eliminates any effect from the collusive conduct, and achieves outcomes similar to honest reporting.

Figure 5

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Baseline Set | | | | | | | | | | |
| Optimum Pairwise | 0.24 | 0.16 | 0.14 | 0.11 | 0.07 | 0.12 | 0.0 | 0.06 | 0.032 | 0.026 |
| Optimum by Firm | 0.24 | 0.16 | 0.13 | 0.12 | 0.08 | 0.12 | 0.05 | 0.06 | 0.029 | 0.024 |
| Bayes Linear | 0.25 | 0.15 | 0.13 | 0.11 | 0.09 | 0.11 | 0.05 | 0.06 | 0.028 | 0.028 |
| Bayes Composite | 0.23 | 0.16 | 0.13 | 0.11 | 0.09 | 0.10 | 0.05 | 0.06 | 0.032 | 0.029 |
| Collusion Set | | | | | | | | | | |
| Optimum Pairwise | 0.20 | 0.14 | 0.11 | **0.14** | **0.09** | 0.11 | 0.05 | 0.08 | 0.058 | 0.043 |
| Optimum by Firm | 0.20 | 0.15 | 0.11 | **0.16** | **0.09** | 0.10 | 0.04 | 0.07 | 0.045 | 0.036 |
| Bayes No Resist | 0.24 | 0.15 | 0.13 | **0.13** | **0.08** | 0.10 | 0.05 | 0.06 | 0.028 | 0.028 |
| Bayes Resist | 0.25 | 0.14 | 0.13 | **0.11** | **0.07** | 0.10 | 0.05 | 0.06 | 0.027 | 0.027 |

Of course, it goes without saying that even if our formulas on improvements on the preexisting methods, our formulas are not perfect. For one thing, even though our formulas require less information to be able to generate results, they do require at least some overlapping reports to succeed. Indeed, it is possible that certain information network structures will produce quirky allocations, such as if there is only one firm that knows enough about another firm did rate it. In addition, even though the literature recommends using peer reports to make relative assessments and courts currently largely do so as well, there may be ways in which reports can be biased by irrelevant considerations or cognitive limitations. For example, it is possible that firms will have trouble making distinctions between firms that contributing little to a litigation because it is more difficult to make distinctions between small numbers (4% versus 5%) than big numbers (40% versus 50%). In addition, firms may waste resources preening and trying to exaggerate their contributions to affect their ratings. Even more perniciously, ratings may be affected by racial or gender bias.

We obviously cannot solve all of the problems with peer reports in this paper. But we hope we solved one. We hope that our formulas will give courts ways to turn peer reports into reliable measures of how much each member of a team of lawyers contributed to the litigation, and thereby improve their ability to award attorneys' fees to those teams.

## VI. Conclusions

# Technical Appendix

October 7, 2019

## 1 Problem Specification

Suppose that $N$ firms, labeled 1, 2, ..., $N$, have worked together on a litigation matter. A court has issued a collective award of attorneys fees, and the objective is to distribute the sum according to "desert" as defined by the firms. To begin, we ask each firm to rate the relative contribution made by each of the other firms. We label these observed ratings as $S_{ij}$, $0 < S_{ij} < 1$ where $i$ is the rater, and $j$ is the rated firm. To prevent self-dealing, a firm may not rate itself. In addition, because some firms may not have sufficient contact with some of the other firms to make an educated rating, some of the $S_{ij}$ may be missing. So for example, we might have a score matrix that looks like Table 1.

| | Rating for | | | |
|---|---|---|---|---|
| Rater | Firm 1 | Firm 2 | Firm 3 | Firm 4 |
| Firm 1 | NA | 0.50 | 0.40 | 0.10 |
| Firm 2 | 0.70 | NA | 0.30 | NA |
| Firm 3 | 0.85 | NA | NA | 0.15 |
| Firm 4 | NA | 0.50 | 0.50 | NA |

Table 1: Example Score Matrix

Note that under this construction, the row sums ($\sum_j S_{ij}$) necessarily equal 1, since each rating firm makes *relative* assessments among the firms for which it has information. Missing values are NA, as distinct from a zero contribution, which are not permitted in most of the models that follow. Given this dataset, the goal of the approaches below is to arrive at a justified estimate for the contribution made by each firm to the litigation as a whole. We will denote by $\alpha_j$ the final recommended allocation for firm $j$.[1]

# 2 Optimization

Our first approach to the fee allocation problem is optimization. We will choose allocations $\{\alpha_j\}$ so as to minimize the error between our allocations and the reports provided by the firms themselves. We will analyze two different measures. The first is based on pair-wise error and the second is based on the individual error of the firms.

## 2.1 Pairwise Error

Suppose that firm $i$ reports on the relative contributions $S_{ij}$ and $S_{ik}$ of the two firms $j$ and $k$. Then we would want our final allocations $\alpha_j$ and $\alpha_k$ to satisfy

$$\frac{S_{ij}}{S_{ik}} \approx \frac{\alpha_j}{\alpha_k}.$$

A natural measure of the error produced by our allocation relative to firm $i$'s assessment of firms $j$ and $k$ is

$$\left(\frac{S_{ij}}{S_{ik}} - \frac{\alpha_j}{\alpha_k}\right)^2$$

but this has a rather severe drawback — it is not symmetric in the firm labels $j$ and $k$. Alternatively we could choose the symmetric measure

$$\left(\alpha_k\, S_{ij} - \alpha_j\, S_{ik}\right)^2 .$$

This formulation has two advantages. It is symmetric in the labels and it leads to a well-behaved quadratic optimization function. Algebraically we have

$$\left(\alpha_k\, S_{ij} - \alpha_j\, S_{ik}\right)^2 = \left(\alpha_k\, S_{ik}\right)^2 \left(\frac{S_{ij}}{S_{ik}} - \frac{\alpha_j}{\alpha_k}\right)^2 .$$

So this latter measure amounts to a weighted version of our non-symmetric measure. This weighting will magnify the error when both the report $S_{ik}$ and the allocated amount $\alpha_k$ is large. The former may happen because the reporting firm only has

[1] Further note that even a simple example like Table 1 already violates the de Clippel informational requirements. For example, Firms 1 and 2 are never rated together.

information on a few other firms, thus the relative values will tend to be larger, or because this particular firm merited a large allocation. In either case, we think large reports should be held to stricter requirements forcing the error they produce to be less.

Summing over all firms and their reports we are left the following quadratic optimization problem:

$$\text{Minimize} \sum_{i=1}^{N} \sum_{\substack{j,k \neq i \\ S_{ij}, S_{ik} \neq 0}} (\alpha_k \, S_{ij} - \alpha_j \, S_{ik})^2$$
$$\text{subject to } \sum_i \alpha_i = 1$$
$$\alpha_i \geq 0 \ \forall i \in \{1, 2, \ldots, N\}$$

## 2.2 Individual Error

Another approach would be to focus on the individual assessments of the firms. Suppose firm $i$ reports a relative share $S_{ij}$ for firm $j$. Then we would want

$$S_{ij} \approx \frac{\alpha_j}{\sum_{\{k | S_{ik} \neq 0\}} \alpha_k}.$$

Reasoning as before we might wish to quantify the error in this allocation as

$$\left( S_{ij} - \frac{\alpha_j}{\sum_{\{k | S_{ik} \neq 0\}} \alpha_k} \right)^2.$$

Unlike the earlier measure, this one exhibits no particular asymmetry. It is, however, rather ill-behaved for optimization purposes. It also has the property of weighting all errors equally, even if the actual allocations involved are rather small. We can create both an easier optimization problem as well as a more meaningful weighting by using

$$\left( S_{ij} \sum_{\{k | S_{ik} \neq 0\}} \alpha_k - \alpha_j \right)^2 = \left( \sum_{\{k | S_{ik} \neq 0\}} \alpha_k \right)^2 \left( S_{ij} - \frac{\alpha_j}{\sum_{\{k | S_{ik} \neq 0\}} \alpha_k} \right)^2$$

as our quantification of the error in the allocation for firm $i$'s estimate for the relative share for firm $j$. Note that now we are weighting the error by the amount of the allocation that firm $i$ observes. This leads us to the well-behaved quadratic optimization problem

$$\text{Minimize} \sum_{i=1}^{N} \sum_{\substack{j \neq i \\ S_{ij} \neq 0}} \left( S_{ij} \sum_{\{k | S_{ik} \neq 0\}} \alpha_k - \alpha_j \right)^2$$

$$\text{subject to } \sum_i \alpha_i = 1$$

$$\alpha_i \geq 0 \ \forall i \in \{1, 2, \ldots, N\}$$

# 3 Bayesian Model

An alternative approach to the problem is to use a statistical model in which each firm's "true" contribution to the litigation is modelled as a latent vairable, $\alpha_j$. Contributions are measured on a relative basis, so the vector $\boldsymbol{\alpha}$ is a unit simplex, that is $\alpha_j > 0$, and $\sum_j \alpha_j = 1$. The $\alpha_j$'s are latent and therefore not directly observed. Instead, we observe the relative ratings given by other firms, which provide information about the $\alpha_j$'s indirectly and with error. In what follows, we develop Bayesian models of increasing sophistication to estimate $\alpha_j$.

## 3.1 Linear Approach

One straightforward approach is to view the observed scores ($S_{ij}$) as consisting of the "true" contribution of firm $j$ relative to all of the other firms rated by rater $i$ plus some Gaussian error. So, for example, if we let $i$ be the rating firm, $j$ be the rated firm, and $X_{ik}$ be an indicator variable for when firm $i$ has sufficient contact with firm $k$ to evaluate its contribution, then the observed score can be modelled as:

$$S_{ij} = \mu_{ij} + \epsilon_{ij} \tag{1}$$

where $\mu_{ij} = \frac{\alpha_j}{\sum_{k \neq i} \alpha_k X_{ik}}$, and $\epsilon_{ij} \sim N(0, \sigma_j^2)$. Here we give the $\alpha_j$'s a flat Dirichlet prior, since they constitute a simplex, and $\sigma_j$'s receive uninformative inverse-gamma priors. (**?**).

The model in Equation 1 can be extended to account for idiosyncratic raters. For example, to the extent that certain raters produce "noisier" signals than others, we can take a random effect approach, where the variance of the error term depends on the rater, i.e., $\sigma_i$. Using a random effect may also have the benefit of making the model resistant to strategic behavior, since the model will characterize a rater whose ratings significantly deviate from the norm as "noisy," and thus an aberrant rater's scores will be effectively downweighted. We will explore these ideas further in the compositional model below.

## 3.2 Compositional Approach

A significant limitation to the linear approach is that it fails to account for correlation among the observation errors. Because $\boldsymbol{\alpha}$ is a simplex, error in measuring one firm's contribution should negatively correlate with the error in measuring other firms' contributions. For example, suppose Firm A rates the relative contributions of Firms B & C. If A underestimates B's contribution, then it must necessarily overestimate C's contribution, since the total relative contributions of B & C must sum to 1.

The log-ratio approach to compositional data proposed in **?** helps model this correlated error. Let's assume that we have $N$ firms involved in the division, and that the true (latent) contribution for each firm is represented by $\alpha_j$, where $j = 1, \ldots N$. Because the vector $\boldsymbol{\alpha}$ is a simplex, it is completely defined by its first $n = N - 1$ elements.[2] So instead of focusing on the individual $\alpha_j$'s, we can focus on their log-ratios, namely:

$$\mu_j = \log\left(\frac{\alpha_j}{\alpha_N}\right), \qquad j = 1, \ldots, n.$$

We can then view the observed log-ratios to be the true log-ratios plus some error term. Let $s_{ij}$ be the observed contribution of firm $j$ as judged by firm $i$, and the vector $\boldsymbol{s_i}$ contain all of the contributions observed by firm $i$ except the last term $(s_{iN})$, namely $(s_{ij}, j = 1, \ldots, n)$. Then, we can model those contributions as:

$$\log\left(\frac{\boldsymbol{s_i}}{s_{iN}}\right) = \boldsymbol{\mu} + \boldsymbol{\epsilon_i}, \tag{2}$$

where $\epsilon_i \sim \mathcal{N}^n(0, \Sigma)$,and $i$ indexes the firm doing the judging. Here, $\alpha_j, j = 1, \ldots, n$, can have uniform priors on [0,1], or perhaps a weakly informative normal prior (with appropriate constraints) since we know that the relative contributions will tend to be in the lower part of that interval.

For ease of notation, we can borrow the notation from **?** and express the model as $\boldsymbol{s_i} \sim \mathcal{L}^n(\mu, \Sigma)$. The covariance matrix $(\Sigma)$ captures interdependencies among the firm contributions. For example, if Firm A and Firm B worked on the same aspect

[2] $\alpha_N = 1 - \alpha_1 - \alpha_2 \ldots - \alpha_n$. As **?** shows, the results are invariant to which $\alpha_j$ is chosen to be the $\alpha_N$– in other words, the results are invariant to permutations of the firms.

of a case, then we should expect negative covariance between the error associated with A's contribution and B's contribution.

One final complication to the log-ratio approach is the issue of missing data. Recall that to prevent self-dealing, Section 1 specified that firms may not rate themselves. In addition, some firms may not have sufficient information to rate all of the other firms involved in the litigation. The estimated contributions provided by the raters are therefore only the *relative* contributions of the firms for which the rater has enough information. The approach presented thus far, however, requires complete information.

Fortunately, because the logratio approach is based on a multivariate normal model, it can handle relative contributions (known as subcompositions) with ease through a linear transformation. (**?**). If $\boldsymbol{s_i} \sim \mathcal{L}^n(\mu, \Sigma)$ is a complete composition as seen in Equation 2, then subcomposition $\tilde{\boldsymbol{s}}_{\boldsymbol{i}}$ has the following distribution:

$$\begin{aligned}\tilde{\boldsymbol{s}}_{\boldsymbol{i}} &\sim \mathcal{L}^n(\tilde{\mu}, \tilde{\Sigma}) \\ \tilde{\mu} &= Q\mu \\ \tilde{\Sigma} &= Q\Sigma Q^T,\end{aligned}$$

where we construct matrix $Q$ based on which indicies are chosen for the subcomposition. In particular, let $M$ be the number of firms in the subcomposition with $m = M - 1$, and $N$ be the number of firms in the complete composition with $n = N - 1$. Then

$$Q = F_M Z F_N^{\intercal} H^{-1},$$

where $F_k$ is the identity matrix $I_k$ with an appended last column of -1's (i.e., $[I_{k-1} : -j_k]$), Z is the $M \times N$ selection matrix that creates the subcomposition, and $H$ is defined as $H = I + J$, where $I$ is the identity matrix and $J$ is a matrix of ones. (**?**).

We apply this property frequently throughout the remainder of the chapter to do estimates using the compositional model. For brevity, we will assume the use of this transformation to handle relative contributions whenever necessary without explicitly mentioning it.

## 3.3 Collusion Resistance

A limitation of the basic compositional model is that it treats the rater firms as interchangeable scientific instruments, so that each set of observed ratings is like any other. However, certain judging firms may be more knowledgable or competent, and thus they may have lower error in estimating contributions. Of even greater concern, two firms may attempt to collude with each other — for example, they may agree to rate each other at levels far in excess of their desert.

One way to address these concerns is to introduce a random effect into the compositional model to account for rater variability. To do this, we can use a decomposition of the covariance matrix:

$$\Sigma = \Omega D \Omega^T,$$

where D is a diagonal matrix of eigenvalues, $\sigma_j$, corresponding to the "scale" of the covariance. We then assume that D has takes the form:

$$D = \gamma_i \tilde{D},$$

where $\tilde{D}$ is a diagonal matrix with elements $\tilde{\sigma_j}$, $0 \leq \tilde{\sigma_j} \leq 1$, and $\gamma_i$ is the random effect measuring the variability (or reliability) of the rating firm $i$. We assume that the $\gamma_i$ arise from a common normal distribution with zero mean and common variance.

The random-effects model directly addresses accuracy differences among different firms. It also helps resist collusion. Consider this example: suppose that there are 10 firms in the pool, and Firm 4 and Firm 5 collude to give higher ratings to each other. Since Firm 4's estimate of Firm 5's contribution will deviate from the assessments made by the other eight firms, the model will estimate $\gamma_4$ to be quite high, effectively downweighting Firm 4's observations. A similar downweighting will occur due to Firm 5's estimate of Firm 4's contributions. Such collusion protection is less effective when the number of firms is small, because detecting "outlier" ratings will be difficult if not impossible. But the random effect model provides some incentive and assurance against collusive behavior when there is a substantial number of firms.

# 4 Examples of Implementation

**Simulation 1.** The first simulation uses the example data introduced in Section 1 and reproduced in Table 2. The authors constructed the dataset by taking the ground truth (obviously not observed by the model) and then playing the role of each of the rating firms and doing rough estimates of the other firms. So, for example, Firm 2 only rates Firm 1 and Firm 3, whose contributions in truth should be in a 5:2 ratio, but which we made roughly 70/30. As seen in Table 2, the models do a reasonable job estimating the "true" contribution values.[3] The compositional model is notably less accurate, and we suspect this is because of the limited data available

[3]The Bayesian models were estimated using MCMC methods using the Stan statistical modeling platform. Visual checks of the trace plots suggest that the linear model has difficulty mixing, likely because of the (unmodeled) correlated errors. The compositional model exhibits no such issues.

(and relatively large number of parameters) in a four firm problem. The posterior distributions of the estimates for the compositional model exhibit a lot of variance, probably due to overfitting. For example, as seen in Table 3, the credibility intervals for the compositional model's estimates are much wider that those for the linear model. These issues (and the relative success of the other methods) suggest that future work might consider some kind of regularization for the compositional model when dealing with small numbers of firms.

| | Rating for | | | |
|---|---|---|---|---|
| Rater | Firm 1 | Firm 2 | Firm 3 | Firm 4 |
| Firm 1 | NA | 0.50 | 0.40 | 0.10 |
| Firm 2 | 0.70 | NA | 0.30 | NA |
| Firm 3 | 0.85 | NA | NA | 0.15 |
| Firm 4 | NA | 0.5 | 0.5 | NA |
| Truth | 0.50 | 0.25 | 0.20 | 0.05 |
| Opt Pair | 0.48 | 0.24 | 0.21 | 0.07 |
| Opt Indiv | 0.48 | 0.24 | 0.21 | 0.07 |
| Bayes Linear | 0.49 | 0.21 | 0.21 | 0.09 |
| Bayes Comp | 0.41 | 0.28 | 0.24 | 0.06 |

Table 2: Data and Results for Simulation 1

| | Credibility Interval for | | | |
|---|---|---|---|---|
| | Firm 1 | Firm 2 | Firm 3 | Firm 4 |
| Linear (50% CI) | (0.49, 0.49) | (0.21, 0.21) | (0.21, 0.21) | (0.09, 0.09) |
| Linear (95% CI) | (0.41, 0.50) | (0.20, 0.26) | (0.20, 0.26) | (0.07, 0.09) |
| Comp (50% CI) | (0.22, 0.52) | (0.15, 0.42) | (0.14, 0.31) | (0.04, 0.08) |
| Comp (95% CI) | (0.01, 0,89) | (0.01, 0.85) | (0.01, 0.66) | (0.01, 0.17) |

Table 3: Credibility Intervals for Bayesian Estimates in Simulation 1

**Simulation 2.** The second simulation uses the data in Table 4. The dataset was constructed similarly to Simulation 1, except that it involves a larger group of firms and therefore more ratings. We also informally increased the amount of measurement error and introduced more missing data values. With six firms, all of the models perform well at recovering the "truth." Once again, the credibility intervals for the

Bayesian compositional model's estimates are much wider that those for the Bayesian linear model, as seen in Table 5,[4] and visual checks of the trace plots yielded results similar to Simulation 1.

| | Rating for | | | | | |
|---|---|---|---|---|---|---|
| Rater | Firm 1 | Firm 2 | Firm 3 | Firm 4 | Firm 5 | Firm 6 |
| Firm 1 | NA | 0.60 | 0.30 | NA | 0.10 | NA |
| Firm 2 | 0.50 | NA | 0.20 | 0.30 | NA | NA |
| Firm 3 | NA | NA | NA | 0.50 | 0.25 | 0.25 |
| Firm 4 | NA | 0.45 | 0.30 | NA | 0.15 | 0.10 |
| Firm 5 | 0.50 | 0.50 | NA | NA | NA | NA |
| Firm 6 | 0.60 | NA | NA | 0.35 | 0.05 | NA |
| Truth | 0.30 | 0.30 | 0.15 | 0.15 | 0.05 | 0.05 |
| Opt Pair | 0.29 | 0.27 | 0.14 | 0.17 | 0.06 | 0.07 |
| Opt Indiv | 0.29 | 0.27 | 0.14 | 0.17 | 0.06 | 0.07 |
| Bayes Linear | 0.28 | 0.28 | 0.11 | 0.17 | 0.08 | 0.08 |
| Bayes Comp | 0.32 | 0.28 | 0.14 | 0.15 | 0.05 | 0.05 |

Table 4: Data and Results for Simulation 2

| | Credibility Interval for | | | | | |
|---|---|---|---|---|---|---|
| | Firm 1 | Firm 2 | Firm 3 | Firm 4 | Firm 5 | Firm 6 |
| Linear (50% CI) | (0.28, 0.30) | (0.28, 0.30) | (0.11, 0.15) | (0.15, 0.17) | (0.05, 0.08) | (0.06, 0.08) |
| Linear (95% CI) | (0.24, 0.32) | (0.25, 0.35) | (0.11, 0.18) | (0.10, 0.19) | (0.02, 0.08) | (0.05, 0.14) |
| Comp (50% CI) | (0.22, 0.41) | (0.18, 0.32) | (0.10, 0.18) | (0.10, 0.19) | (0.04, 0.07) | (0.04, 0.06) |
| Comp (95% CI) | (0.04, 0.71) | (0.05, 0.58) | (0.10, 0.40) | (0.10, 0.40) | (0.01, 0.12) | (0.01, 0.10) |

Table 5: Credibility Intervals for Bayesian Estimates in Simulation 2

**Simulation 3.** The problem with manual encoding is that measurement errors are introduced in a haphazard and uncontrolled way. Thus, the "truth" is no longer really the truth, since our manual coding may inadvertently bias things in some direction. For the third simulation, we adopted a more systematic approach: We began with the "true" distribution among 10 firms. Then for each rater, we added independent normal error to each component, where the standard deviation of the

[4]Hereafter, we will forego reporting credibility intervals in the interests of brevity.

error was set at 20% of the true value (to model the fact that people are perhaps less precise when observing larger quantities). We then randomly dropped some of these components, although to make things more realistic, more "involved" firms (those entitled to a greater share) were more likely to observe a greater number of their peers. Finally, we renormalized the observations so that the relative shares all added to one. The resulting set of observations was labelled the "Baseline" set for the simulation.

To simulate collusive behavior, we then assumed that Firms 4 and 5 agreed to inflate each other's scores. Table 6 shows an example of the baseline (no collusion) set of observations for Firms 4 and 5, and then the affected (collusion) set for the same two firms. Note that as a result of the collusion, Firm 4's rating for Firm 5 is about twice as large, whereas Firm 5, who would have not rated Firm 4 at all, instead rates Firm 4 greater than Firm 1, the greatest contributor in the set. Those inflated scores then have concomitant downstream effects on the other firms rated.

| | Rating for Firm | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| Baseline 4 | 0.42 | NA | 0.28 | NA | **0.19** | NA | 0.11 | NA | NA | NA |
| Baseline 5 | 0.36 | 0.26 | 0.21 | **NA** | NA | 0.17 | NA | NA | 0.079 | 0.083 |
| | | | | | | | | | | |
| Collusion 4 | 0.32 | NA | 0.21 | NA | **0.39** | NA | 0.08 | NA | NA | NA |
| Collusion 5 | 0.21 | 0.16 | 0.13 | **0.39** | NA | 0.10 | NA | NA | NA | NA |

Table 6: Baseline and Collusion Data for Simulation 3

Table 7 displays the true contribution for each of the ten firms in Simulation 3, and the results of the various models on the Baseline and Collusion datasets. Three results are especially worthy of note. First, the outputs of the four models are again quite similar. Second, the models seem to have some natural resistance to collusion, probably because they rely on the observations of so many other firms. Despite Firm 4 and Firm 5's significant attempts to inflate their scores, they are at best only able to modestly affect outcomes by a few percentage points, which is often within the general noise we see in the table. Third, the Bayesian collusion resistant model seems able (at least in this example) to negate the distortion created by Firm 5's collusive efforts.

One should bear in mind that all of these results are subject to random variation. We construct the baseline dataset through random processes, and the process of constructing the collusion set (which attempts to distort the baseline set) also has

| | Allocation | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| Truth | 0.25 | 0.15 | 0.15 | 0.10 | 0.10 | 0.10 | 0.05 | 0.05 | 0.025 | 0.025 |
| **Baseline Set** | | | | | | | | | | |
| Opt Pair | 0.24 | 0.16 | 0.14 | 0.11 | 0.07 | 0.12 | 0.05 | 0.06 | 0.032 | 0.026 |
| Opt Indiv | 0.24 | 0.16 | 0.13 | 0.12 | 0.08 | 0.12 | 0.05 | 0.06 | 0.029 | 0.024 |
| Bayes No Resist | 0.25 | 0.15 | 0.13 | 0.11 | 0.09 | 0.11 | 0.05 | 0.06 | 0.028 | 0.028 |
| Bayes Resist | 0.23 | 0.16 | 0.13 | 0.11 | 0.09 | 0.10 | 0.05 | 0.06 | 0.032 | 0.029 |
| **Collusion Set** | | | | | | | | | | |
| Opt Pair | 0.20 | 0.14 | 0.11 | **0.14** | **0.09** | 0.11 | 0.05 | 0.08 | 0.058 | 0.043 |
| Opt Individual | 0.20 | 0.15 | 0.11 | **0.16** | **0.09** | 0.10 | 0.04 | 0.07 | 0.045 | 0.036 |
| Bayes No Resist | 0.24 | 0.15 | 0.13 | **0.13** | **0.08** | 0.10 | 0.05 | 0.06 | 0.028 | 0.028 |
| Bayes Resist | 0.25 | 0.14 | 0.13 | **0.11** | **0.07** | 0.10 | 0.05 | 0.06 | 0.027 | 0.027 |

Table 7: Results for Simulation 3

random aspects. Further, the estimation procedure for the Bayesian model, which uses Markov Chain Monte Carlo (MCMC) methods, also has random aspects. To get a better sense of average model performance for the Bayesian models, we ran a procedure similar to Simulation 3 (dataset generation as well as model estimation) twenty times, measuring the performance of the models by calculating the sum of square errors between the estimated and true parameters, namely:

$$\text{Performance Metric} = \sum_j (\hat{\alpha}_j - \alpha_j)^2$$

The results of this exercise are seen in Table 8. It shows that the non-collusion-resistant model predictably performs less well on the collusion dataset than on the baseline dataset. This result is of course expected, since the colluders are distorting the observations. The collusion-resistant model, however, seems able to counteract some of the collusion, getting on average closer to the "truth" than the non-resistant model.

| | Dataset / Bayesian Model | | |
|---|---|---|---|
| | Baseline / No Resist | Collusion / No Resist | Collusion / Resist |
| Mean Performance | 1.49E-3 | 2.65E-3 | 1.98E-3 |

Table 8: Average Mean Squared Error from 20 Runs of Simulation 3

# 5 Discussion

Both the optimization and statistical methods offer more viable and flexible methods for solving the fee division problem. They are more easily understood and interpreted: The optimization method seeks a compromise that minimizes the error between it and the firm reports. The statistical method determines the most likely underlying "true" allocation assuming that the firm reports are noisy estimates of this truth. The proposed methods are also capable to handling incomplete data and are robust to collusion, with the collusion resistant Bayesian model being particularly so. In addition, in simulations, the methods consistently yield allocations that accord with each other and that are close to the ground truth. The close agreement between the optimization and statistical models may seem remarkable at first, but further consideration may suggest that it should be unsurprising. As we know from the Gauss-Markov Theorem in the linear regression context, the least squares estimator (optimization approach) is the best linear unbiased estimator (statistical approach). Perhaps what we see here is a compositional data variant of the Gauss-Markov Theorem, although further research would be needed to ascertain which precise methods are linked and under what conditions.

## 5.1 Self-Reporting

One potentially significant limitation of our proposal is its prohibition on self-reports — firms rating themselves. This limitation raises two concerns. First, firms may object to these procedures because they cannot provide any direct input on the share they receive. Part of this objection may rest on some kind of participation value, but it may be also related to information costs. Arguably, the firm with the best (although not unbiased) information on what Firm A did during a litigation is Firm A itself. By not allowing self-reporting, we may be throwing away valuable information. Second, if firms are unable to self-report, what incentive do they have to make careful assessments about other firms? When a firm is self-reporting, it is highly motivated to get other firm contributions right, because its own contribution will be assessed relative to those other contributions. When a firm cannot self-report, the relationship between its assessment of other firms and what the rater firm ultimately receives may appear much more tenuous.

The good news is that our methods are actually capable of handling self-reporting.[5] Just as the methods exhibit natural resistance to collusion, they also resist distortions caused by overly generous self-reports. Much of this resistance likely comes

[5] Our thanks to Ben Alarie for pointing out this possibility.

from the dense, interlocking web of relative ratings that effectively prevents any one set of ratings from distorting the outcomes. In the case of the collusion resistance model, there are further helpful incentives at play. Due to the presence of the random effect term, a self-report that significantly departs from the underlying consensus effectively causes the model to "discount" that rater firm's report. Thus, if a firm wishes its report to be taken seriously, it would be wise to rate itself and the other firms carefully and accurately.

To test how our methods handled self-reports, we extended the simulation in Table 7 to have self-reports. We estimated the allocations for four sets of data, all based on the original baseline set from Simulation 3: i) the baseline set as before; ii) the baseline set plus unbiased self-reports; iii) the baseline set with self-reports where Firm 3 doubled its own rating; iv) the baseline set with self-reports where Firm 9 quadrupled its own rating. As seen in Table 9, the models appear unfazed by the addition of self-reports, whether unbiased or distortive.

## 5.2 Data Concerns

Our proposal has significantly more relaxed data requirements than the de Clippel rule, since it merely requires that every firm be rated by some other firm, not that every pair of firms be rated by some other firm. However, for the methods to be effective, not only do the firms have to be rated, but there must be some degree of interconnection between the various ratings. It is unlikely that any firm will have no ratings, but small-share firms may have only one point of connection to the broader group. For example, A, B, and C all work together, but D only interacts with C. Such small-share firms are somewhat vulnerable to the whims of their connection point. Another structure that could be troubling is a dumbbell-shaped social network where one firm is the link between two groups of firms (ie., a low-degree vertex with high betweenness). In this situation, the relative weights within the groups may be well estimated, but the weight between the two groups will depend entirely on the link firm. Whether this dumbbell shape creates problems in practice though is unclear. After all, one suspects that how satisfied a firm is with its allocation depends on how its portion compares to firms with which it is familiar, not portions received in the foreign half of the dumbbell.

Another potential concern is that ratings will be correlated with share size. This bias can occur in two ways. One possibility is that prominence can psychologically affect the perception of a firm's contribution. The other is that people may not assess small differences in proportions very well – for example, the difference between 33% and 50% is better understood than the difference between 3.3% and 5%. The

| | Rating for Firm | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| Truth | 0.25 | 0.15 | 0.15 | 0.10 | 0.10 | 0.10 | 0.05 | 0.05 | 0.025 | 0.025 |
| **Opt Pair** | | | | | | | | | | |
| No Self-Report | 0.24 | 0.16 | 0.14 | 0.11 | 0.07 | 0.12 | 0.05 | 0.06 | 0.032 | 0.026 |
| Unbiased Self-Report | 0.23 | 0.17 | 0.14 | 0.11 | 0.08 | 0.11 | 0.05 | 0.06 | 0.024 | 0.025 |
| Firm 3 Inflated Self | 0.23 | 0.16 | 0.16 | 0.11 | 0.08 | 0.11 | 0.05 | 0.06 | 0.026 | 0.027 |
| Firm 9 Inflated Self | 0.22 | 0.16 | 0.13 | 0.11 | 0.09 | 0.11 | 0.05 | 0.06 | 0.060 | 0.024 |
| **Opt Indiv** | | | | | | | | | | |
| No Self-Report | 0.24 | 0.16 | 0.13 | 0.12 | 0.08 | 0.12 | 0.05 | 0.06 | 0.029 | 0.024 |
| Unbiased Self-Report | 0.23 | 0.16 | 0.14 | 0.12 | 0.08 | 0.11 | 0.05 | 0.06 | 0.025 | 0.025 |
| Firm 3 Inflated Self | 0.22 | 0.16 | 0.16 | 0.11 | 0.08 | 0.11 | 0.05 | 0.06 | 0.028 | 0.029 |
| Firm 9 Inflated Self | 0.22 | 0.16 | 0.13 | 0.11 | 0.09 | 0.11 | 0.05 | 0.06 | 0.053 | 0.023 |
| **Bayes No Resist** | | | | | | | | | | |
| No Self-Report | 0.25 | 0.15 | 0.13 | 0.11 | 0.09 | 0.11 | 0.05 | 0.06 | 0.028 | 0.028 |
| Unbiased Self-Report | 0.24 | 0.16 | 0.13 | 0.12 | 0.08 | 0.12 | 0.05 | 0.06 | 0.027 | 0.027 |
| Firm 3 Inflated Self | 0.23 | 0.15 | 0.14 | 0.10 | 0.07 | 0.12 | 0.05 | 0.06 | 0.027 | 0.027 |
| Firm 9 Inflated Self | 0.22 | 0.16 | 0.13 | 0.11 | 0.07 | 0.12 | 0.05 | 0.06 | 0.030 | 0.028 |
| **Bayes Resist** | | | | | | | | | | |
| No Self-Report | 0.23 | 0.16 | 0.13 | 0.11 | 0.09 | 0.10 | 0.05 | 0.06 | 0.032 | 0.029 |
| Unbiased Self-Report | 0.22 | 0.16 | 0.13 | 0.10 | 0.09 | 0.12 | 0.05 | 0.06 | 0.025 | 0.025 |
| Firm 3 Inflated Self | 0.26 | 0.17 | 0.13 | 0.10 | 0.06 | 0.12 | 0.05 | 0.06 | 0.025 | 0.025 |
| Firm 9 Inflated Self | 0.22 | 0.16 | 0.13 | 0.10 | 0.09 | 0.12 | 0.05 | 0.06 | 0.032 | 0.028 |

Table 9: Results for Self-Report Simulation

extent that these biases are true is an empirical question and will require additional investigation. For now, we only argue that the direction of the bias is completely unclear ex ante. A rater may perceive a large-share firm as being overly important because of its prominence, or a rater may perceive a small-share firm as being overly important because of a bias toward awarding equal shares to each entity. Similarly, any round-off error associated with small values presumably cuts in both directions. And round-off error can be reduced in litigations featuring a large number of firms by simply reducing the maximum number of firms than any rater can assess (which will have the effect of increasing the relative proportions involved).

A system based on peer ratings may also be susceptible to preening as well as racial or gender bias concerns. Since allocations are based on peer perception, firms may waste resources trying to look good. Similarly, peer ratings may incorporate unstated biases against women or minority attorneys. We concede that these are potential problems with any system reliant on peer ratings, but nothing in our proposal exacerbates them. The analog to preening in the lodestar context is running up billable hours, a far more effective and predictable way of increasing one's share of the fees than preening. As for discrimination, the weights used in the lodestar method is also a way in which bias can creep into its "objective" calculus. And at least for our statistical method, in cases with a large number of firms, it may be possible to use covariates to check for possible race effects both for raters and ratees.

Finally, we note that nothing in our proposal requires that the raters in our scheme be identical to the rated firms. So a court could easily add to the dataset its own assessment of the firms (or the assessments of consultants or special masters) to address any of the above concerns. For reasons stated in the Introduction, we doubt that courts or special masters will often have such superior information so as to justify this strategy. Also, a single set of ratings is unlikely to influence the outcome, for the same reasons that make our methods collusion resistant. Nonetheless, courts can use this option as a safety valve, especially if they weight the "neutral" ratings more heavily in the model.