

# Predicting the “Unpredictable”: The Power of Litigation Analytics

by Charlie Price

## I. Introduction to Litigation Analytics

Litigation is chaos. It’s crammed with uncertainties, hazards, and surprises. As lawyers, we’re constantly trying to predict the future. What are the strongest theories of liability and the chances of the plaintiff prevailing on each one? Whose witnesses will be more believable? What damages will be awarded? How long will the case last? And how much will it cost? How will the judge rule on key evidentiary and dispositive motions? What strategies will most efficiently and effectively advance our client’s goals? When, and on what terms or conditions, will we end or resolve the lawsuit?

Like it or not, we are prognosticators, forecasters (and statisticians and mathematicians). It may not be *why* we went to law school, but it’s what is called for by our jobs. And our predictions drive our decision making. Better predictions mean better decisions, which lead to better results—i.e., better decisions, settlements, outcomes, and advice. Lawyers who grasp this principle have an unmistakable advantage over those who do not. And those who go about gathering the data to inform their decisions have yet a greater edge.

**Example #1: Motion Practice.** Imagine, for instance, being able to tell your client something like this when discussing whether to file a potential motion: *“Our judge has ruled on 386 motions for summary judgment and granted 18% of them. That percentage is 10% higher for cases filed by [XYZ Law Firm]. When she grants them, she most often relies on these four cases. When she denies them, she often relies on these two cases. It takes her about 91 days to rule. Finally, if we staff it this way, it costs us about \$XXX to prepare this type of motion. Given these facts, our risk analysis and settlement history, and the average verdict range, we recommend filing this motion.”* While that level of insight is not easy to achieve, it’s a great way to show how you’re spending money (and your time and energy)—and how it’s making a difference.

**Example #2: Budgeting.** Another example involves budgeting. We know that accurate budgeting (which is just a form of prediction) is hugely important for clients. That said, most outside counsel dislike the process. Why? Because it’s hard, or at least it seems so. There’s this myth, an often costly one, that pervades many law firms and law departments—that some things, such as expected legal costs, are just too difficult to estimate. They happily rely on the familiar disclaimers: “lawsuits are crapshoots,” “I can’t control what the other side does,” “litigation is unpredictable,” or similar mantras. But then again, there are lawyers (a small subset of the profession) who can reliably predict future litigation costs. And while their proficiency might seem like dumb luck, they can do it every time. Why? Because they have a handle on data and analytics and apply principles of science (statistics and math).

These are just a couple examples. There are dozens of other opportunities for inhouse and outside lawyers to collect and use data to detect patterns and trends that will improve the delivery of legal services—data about law firms, parties, judges, briefs, orders, motions, opinions, and other filings that we can then use to make better predictions about how opposing counsel, a judge, a jury, or a party will behave. Yet as a whole, the legal industry has lagged behind other businesses and professions in establishing metrics to measure cost and assess performance. For instance, the same company that requires mountains of data when considering a \$100,000 investment in a piece of equipment might be perfectly content turning over a \$50 million lawsuit to a lawyer who will not employ a quantitative analysis to assess case outcome probabilities and whose own record of trial losses, case management, cost control, and decision errors is unknown. This makes no sense. So, what can or should lawyers be thinking about when it comes to decision making? How can they implement statistics-based decision-making to their practices?

## **II. Using Data to Improve Litigation Judgment**

Here are some practical strategies for deploying data to enhance and improve litigation judgment and decision making.

### **A. Know which litigation metrics you should be tracking.**

Right now, law firms mostly track and use data that benefits them, i.e., revenue per attorney, total expense by practice group, and billable hours, to name a few. But all have access to other data, including information necessary for predicting litigation costs. Past budgets (and spend-versus-budget data) can provide good insight into what a matter has (or should) cost—and offer important learning opportunities. Further, historical discovery costs, expert costs, attorney-billing and -efficiency tendencies, average case cycle times, and judicial/litigant/attorney characteristics can also be beneficial.

Whether you are an outside law firm or corporate legal department, here are some metrics you'd love to know about your litigation and settlement history:

- Total cost per matter broken out by stage—early, middle, late, and trial;
- Average litigation costs per matter-type;
- Average, minimum, and maximum settlements by case type and venue;
- Average cycle-time of matters (i.e., how long the matter is open) by matter-type, judge, and jurisdiction; and
- Motion success ratios (and outcome-to-cost ratios) by case-type.

As you complete this analysis in multiple cases, you'll gain opportunities for data mining, predictive analysis, and greater accuracy of projections over time. And you will be able to

build in new metrics to improve your predictions. As new cases come in, you can query against prior cases with a similar profile to gain even deeper case insights and settlement strategies. Finally, for corporate legal departments in particular, consider tracking these additional litigation metrics:

- **Overall legal spend.** How much you spent this year (or quarter) compared to last year (or quarter).
- **Spend versus budget.** Tracking budgets is one of the best ways to predict and control costs for legal departments. Budgets also provide benchmarks for what matters *should* cost, so you can set your own internal budget based on past spending.
- **Billing data.** Analyze costs and productivity of law firms and individual lawyers. Initially, focus on hours and costs of legal services. Later, expand your analysis to discovery costs, expert costs, efficiency of time spent, fee spikes, and when (at what stage of litigation) is the best time to settle. Pay particular attention to time spent managing discovery disputes.
- **Outside and inside spending as a percentage of company revenue.** This can help show department's efficiency over time.
- **Invoice savings and savings-versus-budget.** Track expense guideline reductions, timekeeper rate reductions, and hours-billed reductions to show you are reviewing and enforcing your own billing and expense guidelines.
- **Litigation exposure over time.** Tracking maximum and minimum exposure, along with likely outcome, allows you to keep stakeholders apprised of potential risks. This metric also allows you to categorize cases, predict outcomes (to some degree), and provide information concerning return on investment. After you make these forecasts, determine the accuracy of them. This must be done with sufficient regularity and rigor so that conclusions can be drawn about exposure.
- **Win rates and total cost per matter** (settlement plus outside legal fees);
- **Billing rates** by firm, attorney, and region, including rate increases;
- Average, minimum, and maximum **hourly rates by matter-type and region**; and
- **Average spend** per matter by law firm.

The last four categories can help guide decisions regarding counsel-selection. Law departments are always trying to lower legal costs without compromising quality or outcome. This means determining which law firms deliver the most value (and best results) in each geographical region. Assuming you have this information, you can make rough judgments about the relative value your outside lawyers provide. The goal is threefold. You want to put yourself in a position to: (1) track costs at reasonable intervals throughout the life of each matter; (2) forecast future costs of your litigation matters; and (3) apply statistical methods of analyzing litigation cost, such as decision trees and probabilistic analysis, to make informed, intelligent decisions about resolution.

**B. Create after-action reports at the conclusion of cases.**

Examine your own settlement history—nationally, geographically, and across similar types of cases. Identify the factors that went into your settlement decisions. How long did they take to happen? Who or what was responsible for making them happen? When is the earliest they could have happened? How much did it cost you to get to those settlements?

Where appropriate, conduct unflinching postmortems at the conclusion of your lawsuits. Analyze what went right or wrong and what factors contributed to the end result. You can use these reports to more effectively analyze settlements and the timing of them. For instance, if outside counsel recommends settling a case for \$100,000 on the eve of trial, have them identify the facts or factors that support that recommendation. Then examine that information to determine (1) when it became known or apparent to you or outside counsel and (2) how it could have been discovered, obtained, or delivered sooner. This exercise can help decrease case cycle time and improve the early evaluation process, both of which can lower litigation costs.

### C. Consider the “expected value” of your litigation strategies.

Expected value is a term used to describe the value of an event averaged over all possible outcomes. Here is a simple example involving a basketball game. Over his 21-year NBA career, LeBron James has attempted 11,000 free throws and made around 8,100 of them (a 74% success rate). Based on this data, a fair estimation of his chance of making his next free throw is 74%. And the *expected value* or EV of his next free throw attempt is .74 points. He either makes the free throw (and scores one point) or misses it (and scores zero points); but on average, with one free throw attempt, he’s expected to score .74 points. That is the expected value of one LeBron James free-throw attempt averaged.

The expected value of any of our actions is the product of two variables—the odds the action will allow us to gain something of value, and the value of the gain to us. Whether we are aware of it or not, it’s a concept we all use in some way, shape, or form when we make litigation decisions. We consider our options and determine: (1) the likely “cost” (in time, money, energy, or some other resource) of each option; (2) the potential value to us if we achieve the desired outcome; and (3) the probability that the option we pick will actually produce the desired outcome. It’s basically a math equation, the only tricky part being that we need to fill in three key numbers—the *cost* of the action being considered, the *value* of the desired outcome, and the *probability* the strategy will work.<sup>1</sup>

---

<sup>1</sup> While it’s not always easy to do, we must learn to express probabilities numerically (e.g., “there’s a **65% chance** of our motion being granted” or “we believe that **7 out of 10 times** we will get a defense verdict”) versus subjectively (e.g., “there’s a **good chance** we’ll win our motion” or “it’s **very likely** that the judge will toss their claim”). These

#### **D. Understand base rates and how they can improve decision making.**

As lawyers, we're often expected to predict things. It's part of the job—but a part we resist for fear of looking foolish. One way to avoid that fate is to pay attention to base rates, which are real-world background condition that exists for whatever you're trying to predict—i.e., the relevant baseline data about similar situations that can help inform your decision, prediction, or forecast. Base rates are important because they help us avoid over-relying on misleading pieces of information that can send us veering off in wrong directions. Here is the classic illustration used to explain the problem:

The Renzettis live in a small house at 84 Chestnut Avenue. Frank Renzetti is forty-four and works as a bookkeeper for a moving company. Mary Renzetti is thirty-five and works part-time at a day care. They have one child, Tommy, who is five. Frank's widowed mother, Camila, also lives with the family. How likely is it that the Renzettis have a pet?

That seems like a hard question, doesn't it? Some of you might have just shrugged your shoulders and thought "no idea." More likely, though, your brain shifted into a storytelling mode to figure out the answer: *Hmmmm ... Frank Renzetti. Sounds Italian ... Frank probably grew up with lots of siblings. But he only has one kid. He probably prefers a bigger family ... Plus, Frank's mom lives with them. She could help watch the pet during the day while the parents are at work.* Accordingly, based on the narrative you've created—which feels really compelling!—you conclude there is a high likelihood the Renzettis have a pet. Or you might have created a different story as to why they don't have one. Either way, it's a wild guess that feels "right" in your brain. But you would be much better off starting with a reliable base rate. According to Google, for instance, about 66% of U.S. households (86.9 million homes) own a pet. That's a relevant base rate. That should be our starting point. Once we have that base rate (and assuming we deem it reliable), we can *incrementally* adjust it up or down based on the specifics of the particular case (The Renzetti Story).

This is how good predictions are made—by starting with the outside view (the base rate) before turning to the inside view (the situational details). Because once you know the base rate for whatever you're trying to predict (and you are comfortable speaking in numbers), you can calculate the "expected value" of your decisions.

#### **E. Make data-driven predictions about how opposing counsel, a judge, a jury, or a party will behave.**

---

latter subjective expressions can lead to bad decisions, because human beings ascribe different meanings to each term.

While every matter is unique, there are clear patterns and trends when it comes to litigation costs. Legal analytics allow you to compare your billing rates against data from across the country. They can give you a significant advantage in rate negotiations, reveal new savings opportunities, and help increase value received from outside counsel.

There are lots of third-party analytics products that can help with this process. Some popular analytics products are Serengeti Tracker, TyMetrix Legal Analytics, Sky Analytics, Westlaw Edge, LexisNexis Context, and Lex Machina. Two worth noting are LexisNexis Context and Westlaw Edge. They extract massive amounts of information about law firms, parties, judges, briefs, orders, motions, opinions, and other filings that you can then use to make data-driven predictions about how opposing counsel, a judge, a jury, or a party will behave. Among other things, they allow you to:

- Analyze judge tendencies, including how frequently specific judges grant certain types of motions, win/loss data, time to disposition;
- Determine which attorneys handle which kinds of cases and whether they went to trial or tend to settle;
- Compare granted motions with denied motions, as well as dockets, case strategies, and tendencies;
- Determine how likely a given judge is to grant or deny evidentiary and dispositive motions, as well as how frequently the judge cites any given case in deciding such motions;
- Analyze judge tendencies, including how frequently specific judges grant certain types of motions, win/loss data, time to disposition.
- Determine which attorneys handle which kinds of cases and whether they went to trial or tend to settle, and
- Compare granted motions with denied motions, as well as dockets, case strategies, and tendencies. While most of the data is scraped from PACER, and is therefore limited to federal court activity, it is still helpful information.

### **III. Conclusion**

Malcolm Gladwell, in his book *Outliers: The Story of Success*, explained that “[s]uccess is a function of persistence and doggedness and the willingness to work hard for twenty-two minutes to make sense of something that most people would give up on after thirty seconds.” Remember this quote when you encounter what seems like insurmountable uncertainty. Avoid the knee-jerk reaction to surrender after thirty seconds. Instead, take a breath, and focus on how you can reduce—not necessarily eliminate—that uncertainty through your data and analytics. Don’t obsess about what you don’t know; focus on what you do know. Understand the value of base rates; learn basic statistics; measure things; collect data; and make, track, and learn from your decisions and predictions. Because

where decisions have potentially significant consequences, such as in litigation, these habits will give you a competitive advantage in this highly competitive field.

Finally, there are entire library shelves devoted to forecasting, decision analysis, and predictive analytics—as well as how to deal with all the cognitive biases, misperceptions, and misinformation that interferes with optimal decision making. If you are interested in learning more, here are a few good books to consider:

- *Winning at Litigation through Decision Analysis*, by John Celona
- *Thinking in Bets* and *How to Decide*, both by poker legend Annie Duke
- *Superforecasting* by Philip Tetlock and Dan Gardner
- *Beyond Right and Wrong* by Randall Kiser
- *Data-Driven Law* by Ed Walters
- *Predictably Irrational* by Dan Ariely
- *Moneyball* by Michael Lewis. This is one of my favorites. It's about how the Oakland A's used analytics to improve their forecasting and decision making. And while practicing law is nowhere near as amenable to data analysis as baseball, there are some great lessons in there.