MEASUREMENT & DISTRIBUTION

The Benchmark Isn't the Asset

What history shows about when proprietary measurement creates distribution, when it becomes a business of its own, and when monetizing it destroys the thing that made it valuable.

AUGUST 2026 10 MINUTE READ

A company finds itself holding unusual data about its own market. It knows things about the institutions it sells to that those institutions do not know about each other. So it publishes an index. Every company, city, hospital, supplier, or agency gets a number and a position.

The ones at the bottom will want to know why. The ones at the top will want to tell people. Either way there is now a reason to talk to a potential customer that did not exist the week before, and it arrives without a cold email.

It sounds close to unfair. Sometimes it is.

But the historical record on this strategy is much stranger than the pitch. Some rankings became powerful ways into commercial relationships. Some became excellent standalone media properties that never sold a single unit of anything else. Some committed their owners to forty years of permanent conflict management. And one of the most influential regulatory rankings ever published became so consequential, commercially and politically, that the institution behind it shut the whole thing down.

The interesting question is not whether measurement can create distribution. It can. The question is what separates the versions that worked from the versions that ate their owners.

ONE

How powerful this can get

Start with the case that removes any doubt about the mechanism.

From 2003, the World Bank published an annual report called Doing Business, which ranked national economies on how difficult it was to start and run a company there. Registration times, credit access, tax filing, contract enforcement. Country rankings arrived in 2005 and eventually covered up to 190 economies.

Governments did not treat it as a research publication. They treated it as a scoreboard. More than seventy countries established dedicated policy committees or working groups aimed specifically at improving their indicator scores. Narendra Modi's administration set a public target of reaching at least 50th place, an improvement of more than a hundred positions. In 2012 Vladimir Putin set a target of the top twenty by 2018. Indonesia went for the top forty. Morocco for the top fifty.

Then those governments went looking for help. The World Bank sold advisory services to countries, paid for by those countries, on the regulatory environments the Bank itself was scoring.

> The mechanism worked almost too well. It produced attention, behavioral change, and paying demand for improvement, all pointed back at the institution holding the pen.

What followed was not a simple consulting scandal, and it is worth being precise about it. In June 2020, World Bank staff raised internal concerns about the numbers for Azerbaijan, China, Saudi Arabia, and the United Arab Emirates. An external investigation by the law firm WilmerHale found that senior leadership had pressured staff to improve China's ranking in the 2018 edition at a time when the Bank was seeking Chinese support for a capital increase. An internal audit reported that management had pressured nine of fifteen staff on the 2018 and 2020 editions. On 16 September 2021, the World Bank discontinued the report.

The manipulations were not crude. Contemporaneous analysis noted that in none of the documented cases did anyone simply edit a country's final score. The changes were methodological. A weighting here, an indicator definition there, each individually defensible.

That detail is the whole lesson. Once a score carries real consequence and the publisher both controls the methodology and earns money from the measured, every methodological decision becomes commercially and politically loaded. The expert panel the Bank convened said so plainly, recommending that the Bank stop selling consulting aimed at improving country scores, on the grounds that it should not simultaneously score a business environment and accept payment to change it.

TWO

Where the pressure comes from

Now a version that has survived, and where the pressure originates somewhere else entirely.

Companies like BitSight and SecurityScorecard rate the cybersecurity posture of organizations from the outside. Passive scanning, exposed services, certificate hygiene, leaked credentials. They do this without asking, which means most companies have a rating whether they know it or not.

Consider how this actually arrives for the rated company. A supplier gets a call from a customer that represents a meaningful share of its revenue. The customer's third party risk team has been looking at the supplier's score and does not like it. Contract renewal is in eleven weeks. An obscure number that the supplier had never heard of on Monday is, by Thursday, a problem with a date on it.

The supplier goes to the rating company. Not to buy anything, initially. To find out what the number is built from, to see which findings are attributed to domains it does not own, to dispute the wrong ones, and then to watch the score while it fixes the right ones. Some of that is free. The monitoring is not.

The critical point is what created the urgency. The rating company did not. It published a number. The customer's procurement and risk process is what turned that number into a deadline.

Worth noting how the free part came about, since it is often described as an ingenious funnel design. In June 2017 the US Chamber of Commerce, working with more than two dozen large member companies that were themselves being rated, published a set of Principles for Fair and Accurate Security Ratings. It was modeled explicitly on the Fair Credit Reporting Act. Rated organizations were to get access to their own rating and the underlying data, the right to challenge and correct it, and disputed ratings were to be marked as disputed until resolved.

In other words, the rated companies demanded transparency as a condition of the category's legitimacy, and the raters conceded it. That it also happens to be a useful commercial doorway is a downstream accident rather than a design.

THREE

The company that gave away the standard

The most instructive case runs opposite to instinct, and it is worth sitting with.

Press Ganey was founded in 1985 by two University of Notre Dame professors, a medical anthropologist and a statistician, who had noticed that hospital managers kept asking them how to measure whether patients were satisfied. For roughly two decades the company owned that question. Its own instrument, its own comparative database, its own definition of good. By the mid 2010s its outpatient survey was the most widely used in the United States.

Then the federal government built a competing standard and made it mandatory.

Medicare and the Agency for Healthcare Research and Quality developed HCAHPS, a standardized patient experience survey. From 2007, hospitals in the main Medicare payment system had to collect and publicly report it or take a payment reduction. The results went onto a public government website where any patient could compare hospitals. Then in October 2012, under the Affordable Care Act, Medicare began withholding one percent of base operating payments to create an incentive pool of roughly $850 million, redistributed on quality performance, with patient experience as a substantial component. Current descriptions put HCAHPS at around a quarter of a hospital's value-based purchasing score.

Press Ganey became one of many government-approved vendors permitted to administer someone else's survey.

On the face of it, that is a demotion. A company that owned a category standard was reduced to running a federal form. What happened instead is that it got considerably more valuable. Revenue was around $281 million in 2014. It listed on the New York Stock Exchange in May 2015, went private, and was sold in June 2019 for $4.2 billion. In 2025 Qualtrics announced it would acquire the business for $6.75 billion.

Those numbers reflect four decades of building, several acquisitions, and an expansion well beyond patient surveys, so nobody should read them as a clean measurement of one regulatory change. But the structural shift is worth understanding, because it explains why the company was safe in a way its predecessors were not.

Medicare owns the questions. Medicare owns the weights. Medicare receives the data, computes the consequential number, publishes it, and attaches the money. Press Ganey sells everything arranged around that obligation: survey administration, sampling design, submission, compliance management, comment analysis, its own supplementary instruments, workforce surveys, benchmarking against a database no single hospital system could assemble, and improvement consulting.

> There is no version of the sentence "we paid Press Ganey and our score went up" that makes sense. The only path from payment to a better number runs through patients having a better time in the hospital.

Look at what the criticism attaches to as a result. Press Ganey has been argued about intensely for over a decade, and almost none of it concerns the company's honesty. A 2012 study in the Archives of Internal Medicine found higher patient satisfaction associated with greater spending and higher mortality. A 2015 Hastings Center paper warned that chasing satisfaction scores could push clinicians toward honoring requests for treatments patients should not have. A large literature argues about whether satisfaction scoring drove inappropriate opioid prescribing, and the findings are genuinely mixed.

All of that is an attack on the metric. When a government owns the standard, the fight over whether the standard is any good happens between the government and its critics. The vendor administering it is not in the room.

FOUR

What it costs to own both sides

Some organizations do insist on owning the standard and selling improvement against it. It is possible. The price is that the question never closes.

The Joint Commission accredits American hospitals, and its accreditation carries what is called deemed status, meaning it can determine whether a hospital participates in Medicare at all. There is not much larger a consequence available in the sector. Its affiliate, Joint Commission Resources, established in 1986, sells education, publications, and consulting to hospitals, including help preparing for the survey its parent conducts.

This has been contested for a quarter century. Congressional calls for investigation, two Government Accountability Office reviews, a bill that would have moved oversight to Medicare, a Senate subcommittee inquiry, and eventually a formal federal consultation on whether accreditors with consulting arms create a perception of conflict. None of it broke the model.

The reason is a firewall built in 1987 and taken unusually seriously. Separate offices, telephone numbers, and computer systems. A contractual disclaimer telling client hospitals that the accreditor is not informed they used the consultants. Consultants barred from communicating with surveyors about specific facility decisions or from discussing which surveyors get assigned where. Legal review of all consulting promotional material. A compliance officer with authority to dismiss.

There is a quieter control too. Responding to the earliest criticism, the organization noted that consulting had gone to fewer than two percent of accredited institutions, and that most of the affiliate's net revenue came from education, publications, and international work. Keeping the improvement business small relative to the measurement business is a defense an outsider can check.

The GAO's eventual finding was carefully worded. The two organizations had taken steps designed to protect facility-specific information. It reached no conclusion about whether a conflict actually existed, and both sides read it as vindication.

That is the realistic best case for anyone who owns the standard and sells the remedy. Not resolution. Procedural adequacy, re-examined by outsiders every decade or so, funded permanently.

FIVE

A different way to break

One more, briefly, because it fails somewhere the others do not.

EcoVadis rates the sustainability performance of suppliers, and large corporate buyers use those ratings in procurement, commonly requiring a minimum score to stay on an approved supplier list. A low number can remove a company from bidding. The consequence again belongs to somebody else, since the buyer created it, and EcoVadis sells the assessment along with improvement tooling and training.

The exposure here is not that paying improves your score, because every rated supplier pays. It is that the assessment rests on evidence the supplier selects and submits about itself. That is a provenance problem rather than a conflict problem, and it is the same one that damaged American college rankings, where decades of unaudited self-reported data eventually produced a manipulation scandal at a top-ranked university and a boycott by leading professional schools.

Measurement businesses fail from several directions. Commercial conflict is only one of them.

What actually separated them

Set the individual histories aside. Five things distinguish the versions that produced commercial relationships from the versions that produced press coverage.

The number has to matter to somebody other than you

In every case that generated real commercial behavior, there was a third party who could withhold something concrete. Medicare withholding payment. A customer withholding a contract renewal. A buyer withholding RFP eligibility. An insurer setting terms. Investors voting shares.

Compare that to prestige rankings. For twenty-five years, a govtech media company has run an annual survey ranking American cities on their use of technology, with awards presented at a national municipal conference. Cities compete for it enthusiastically and it costs them nothing to enter. It is a well-run business, funded by technology vendors who underwrite the survey in order to be near the cities. What it is not is a way for those vendors to acquire customers, because no city buys anything as a result of where it placed. The audience is peers and suppliers, and neither can withhold anything from a city.

> Reputation creates attention. Consequence creates action. These are not points on the same scale.

There has to be a date

Severity alone is not what moves institutions. Timing is. Every conversion in this history ran through a decision that was about to be made: a reporting deadline, an accreditation window, a vendor review, a renewal, a procurement threshold.

This has a practical consequence for anyone building the motion. An interesting score can sit in an inbox indefinitely, and usually does. An obligation with a date attached cannot, because somebody inside the measured organization is now accountable for it on a specific Tuesday.

A named score is a different object from a report

Industry research builds authority. A named, attributable score builds a subject.

"Permitting in mid-sized American cities takes an average of nine weeks" is content. Somebody might quote it. "Your city ranks 84th, here is the dataset, here is the calculation, and here are the three jurisdictions nearest you in population that score higher" is an event with an owner.

Do not overread this. The named score produces attention and, sometimes, engagement. Whether that engagement becomes a purchase is decided by the consequence, not by the personalization. The personalization is what gets the number read by a person rather than filed.

The safest improvement product changes reality, not the number

This is the sharpest thing in the historical record, and it can be reduced to one question.

THE TEST

Could a customer's payment improve their score through any path that does not run through improved reality?

If no such path exists, you can sell improvement openly, market it in plain language, and defend it indefinitely. Press Ganey advertises helping hospitals improve their scores, in those words, and has done so for years without a credibility problem, because the number is computed by the federal government from what patients report.

If such a path does exist, careful wording will not save you. The accusation attaches to the structure. It will be made eventually, by a competitor or a journalist or a legislator, and the honest answer will be that the path exists and you chose not to use it.

Doing Business is the warning on the other side. The path existed. It ran through methodology, held by an institution that was also being paid by the countries it scored, and it was used.

Owning everything is powerful and fragile at the same time

The instinct is to own the data, the methodology, the score, the publication, and the improvement product. It looks like a moat. It is also five liabilities stacked in one entity, and the more consequential the score becomes, the heavier the stack gets.

What the durable businesses have in common is that each gave something up. Press Ganey surrendered the standard and became the best operator of the government's. The security ratings industry accepted governance written by the companies it rates and now competes partly on whose methodology has been validated by outside parties. The Joint Commission kept both halves and paid for the separation with forty years of firewall maintenance and periodic federal scrutiny.

When the European Union finally legislated on this question for sustainability ratings, the Commission's opening position was an outright ban on rating providers selling consulting to the entities they rate. The regulation adopted in late 2024 did not do that. It permitted the combination on condition of structural separation, with the supervising authority explicitly warning that the separation must not be superficial or a matter of legal formality. Three bodies in three sectors across four decades looked at this and reached the same conclusion, which is that the conflict is manageable and never finished.

Three businesses that get mistaken for each other

Most of the confusion in this area comes from a category error. There are at least three distinct businesses here and they have different economics, different staffing, and different failure modes.

Research as marketing

You publish, people read, your name carries more weight in a room. It produces credibility and inbound attention. It does not produce a pipeline you can forecast.

Measurement as the product

Someone pays for the score, the certification, the data, or the comparison itself. This can be a very good business. Notice that when the measured party is the one paying, the ranking has become the sales mechanism for the measurement, and the higher-value product on the other side often never gets built because it stops being necessary.

Measurement as distribution

The score creates a consequential reason for the measured organization to enter a commercial relationship about something else. This is the rarest of the three and the one with the most conditions attached.

None of these is better than the others. The damage comes from believing you have the third when you actually have the first, because then an entire commercial organization gets built against pipeline that does not arrive, and the index gets blamed for it.

The questions worth asking first

If you are holding proprietary data and considering publishing a score, the first question is not whether you can rank your prospects. You almost certainly can. These are closer to the ones that decide the outcome.

01 Who cares about this number besides the organization receiving it?

02 What can that outside party actually withhold, in money, access, eligibility, or permission?

03 When is the decision made? Is there a date, and does it recur?

04 Can the measured organization opt out of being scored, and what happens to the whole model if the important ones do?

05 Who inside the measured organization becomes responsible for responding, and is that an operator or a communications function?

06 What can you sell that genuinely improves the underlying thing being measured?

07 Could paying you improve the score through any other route?

Question five is worth dwelling on. The difference between a score that lands on an operator and a score that lands on a press office is roughly the difference between a commercial conversation and a statement to the media. The organizations that made this work tended, deliberately or not, to send the number to the person who owned the process rather than the person who owned the narrative.

One honest limitation, since it bears on how much weight any of this can carry. No organization in this history publishes the rate at which measured entities become customers. Not one. The structural conditions are well documented across decades and sectors. The conversion economics are not documented anywhere, by anybody. What history supports is the mechanism and the conditions under which it operates, not a number you can put in a model.

• • •

The reversal

The instinct is that the valuable thing is the benchmark. You built it, you own it, nobody else has the data, and every year it gets harder to replicate. All of that can be true and still miss where the value sits.

A benchmark matters only at the moment somebody with power acts on it. Everything else about it, the methodology, the coverage, the brand, the press cycle, is infrastructure for producing that moment. The companies that built something durable around measurement either found a consequence that already existed and stood next to it, positioned themselves beside one that a government or a customer or a regulator had created, or kept both halves and accepted that conflict management would be a permanent line item forever.

> The benchmark is not the asset. The consequence is.

Which suggests a different way to think about the whole exercise. The strongest measurement businesses in this history do not need anyone to fear the publisher. They need their customers to care intensely about a decision that somebody else is about to make.

SOURCES

World Bank Independent Evaluation Group, evaluation of the Doing Business project, on country reform committees and government targets, 2021.

WilmerHale, Investigation of Data Irregularities in Doing Business 2018 and Doing Business 2020, September 2021; World Bank announcement of discontinuation, 16 September 2021; external expert panel review, September 2021.

US Chamber of Commerce, Principles for Fair and Accurate Security Ratings, June 2017.

Centers for Medicare & Medicaid Services and Agency for Healthcare Research and Quality, HCAHPS survey documentation and approved vendor materials; Hospital Value-Based Purchasing provisions of the Affordable Care Act, effective October 2012.

Press Ganey company history and product documentation; transaction reporting on the May 2015 listing, the June 2019 sale, and the 2025 acquisition announcement.

Fenton, Jerant, Bertakis and Franks, The Cost of Satisfaction, Archives of Internal Medicine, 2012; Hastings Center Report, 2015; subsequent literature on patient satisfaction and opioid prescribing.

US Government Accountability Office, GAO-07-79, Hospital Accreditation: Joint Commission on Accreditation of Healthcare Organizations' Relationship with Its Affiliate, December 2006; congressional correspondence, 2001 and 2004; CMS request for information on accrediting organization conflicts of interest, December 2018.

EcoVadis methodology documentation and scorecard materials; buyer threshold practices as described by procurement advisers.

Regulation (EU) 2024/3005 on the transparency and integrity of ESG rating activities, adopted November 2024; ESMA consultation on draft regulatory technical standards, 2025.

Contemporaneous reporting on college ranking data integrity, 2022 and 2023, including the settlement of student litigation and the withdrawal of leading professional schools.

Figures are drawn from company disclosures, regulatory documents, government evaluations, and contemporaneous reporting. Where a source is a company's own material, it is used for what the organization says it sells rather than for how well it works. No organization examined publishes conversion data from measurement to commercial relationship.