



WORKING PAPER · SEPTEMBER 2026

The Bracket

How economics set aside value in use, how the seller picked it up, and why the debt is now due

Michael Brey

Value Lattice

CONTENTS

Executive Summary	3
What This Means for a Buyer	3
1 The question	4
2 The bracket: how economics set value in use aside	5
3 The bracket goes institutional in private business	13
4 The marketing turn: value in use is named and handed to the supplier	16
5 The seller inherits the question	18
6 The seller's case, and where it breaks	21
7 The governance vacancy, and what stands around it	25
8 The size of the debt	27
9 Outcome pricing makes the instrument load-bearing	28
10 Resolving the debt	29
About the Author	31

Executive Summary

Every purchase carries two numbers. One is what you paid. The other is what it did for you. The first is on the invoice. The second is not on anything.

This paper calls that asymmetry the bracket, and traces where it came from. In 1776 Adam Smith distinguished value in use from value in exchange, illustrated the difference with water and diamonds, and then built the rest of his book on exchange. The economists who followed him developed a rich theory of how utility explains price, and later, in welfare economics, real instruments for measuring the value of public works. The appraisal half of those instruments migrated into private business as the discounted business case. The evaluation half, the mandated check after the fact on whether the value arrived, never migrated into private enterprise technology purchasing. Business measurement grew up on the exchange side of Smith's line, because that is the side prices make auditable.

When the question of value in use finally came back into commercial practice, it came back through marketing, as something the supplier proposes. And from roughly 2003, the enterprise software industry built the only working apparatus for quantifying value in use that most private buyers have ever encountered, delivered it free, and filed the resulting software category under Sales. I helped build it. It was built to answer the seller's question, which is whether a purchase will be made. It was never scored against the buyer's question, which is whether the value arrived.

Since 2024, enterprise software vendors have begun charging for outcomes they define and meter themselves. In that market, that makes the buyer's missing instrument load-bearing for the first time. The paper's conclusion is that the party whose money is at risk needs its own meter, and that sellers who adapt to that early will do better than those who defend the old arrangement.

Where the paper interprets rather than cites, it says so. The historians support a narrower claim than the one it is tempting to make, and section 2.9 states the limit.

What This Means for a Buyer

If you buy enterprise technology, this paper argues four things you can act on.

The number that matters is not on any instrument you hold. Your organization can tell you what it paid for a system to the cent. It cannot tell you, on any document it authored, what the system was supposed to deliver in business terms, whether that arrived, or by how much it missed. That is not a gap in your organization. It is a gap in the whole profession, and it has a history.

The instrument you have been using belongs to the seller. The business case, the benchmark, the return model, and the value realization review that accompanied your last major purchase were, in most cases, built by the vendor's value engineering function, using the vendor's method, in the vendor's tool. They were built to help the purchase decision get made. Two decades of evidence show that realized value runs at roughly half to two-thirds of what those instruments projected, and no vendor value function has published a scored comparison of its projections with outcomes.

Outcome pricing has turned this from a knowledge gap into a billing gap. When a vendor charges per "resolution," the vendor defines the resolution and runs the meter. Your own data, the tickets, the transcripts, the records of when a human was asked for, can check that meter. Most buyers do not use it that way, and most contracts do not give them the right to.

The fix is a buyer-authored statement of value, made before signing, verified after. Not a service-level agreement, which measures the seller's activity. A statement of the business outcome, in your terms, with the measure and the consequence written down, that the seller responds to rather than writes. Section 10 says how that differs from what procurement already does, and why attribution is the hard part.

The rest of the paper is the argument. It is a history, because the problem has one, and a hostile reader will test the history first.

1. The question

The asymmetry between what you paid and what it did is so familiar that it reads as a fact of nature rather than as a choice somebody made. It is a choice, and it has a history. The purpose of this paper is to make the choice visible, to follow its consequences forward, and to show where it has left private buyers in a year when vendors have begun to charge for outcomes they define and meter themselves.

Three things need saying before the history begins.

First, none of the people in this story were foolish. Smith set value in use aside because he could not see how to reason from it to price, and he was right that he could not, on the definition of utility he was using. The marginalists solved that problem and kept working on price. The welfare economists built instruments for measuring value in use, for public works, and only the front half of them, the appraisal, was carried across into private technology purchasing. The accountants built on exchange value because exchange value can be audited. The marketers named value in use because they wanted to sell more effectively, which is their job. The value engineers built projection tools because that is what closes deals, which is theirs. Each step was locally sensible. The result is a hole that nobody dug on purpose.

Second, this is not a paper against sellers. I spent fifteen years on the seller's side of enterprise software, building the value cases and the benchmarking apparatus this paper describes. What I describe, I helped make. The argument is not that the apparatus was built in bad faith. It is that it was built to answer a question, it answers that question, and it has been asked to answer a different question it was never scored against.

Third, the conclusion is that the party whose money is at risk needs its own meter. That is not a preference. It follows from the record, and the record is the rest of this paper.

2. The bracket: how economics set value in use aside

2.1 The distinction was old before Smith used it

Aristotle made it. In the *Politics* he wrote that “of everything which we possess there are two uses,” the proper use and the use in exchange, and illustrated it with a shoe, which is “used for wear, and is used for exchange.” Emil Kauder states the consensus: “It is generally accepted that Aristotle was the first who created the concept of the value-in-use.”¹

The medieval scholastics developed the distinction into what Schumpeter calls “a fragmentary but genuine subjective or utility theory of exchange value.” Schumpeter reads Pufendorf as distinguishing a good's aptitude for use from its exchange price and

¹ Aristotle, *Politics*, Book I, part 9, 1257a6-13, trans. Jowett. Emil Kauder, *A History of Marginal Utility Theory* (Princeton University Press, 1965), p. 15.

letting the latter be governed by scarcity; the terms are Schumpeter's characterization rather than Pufendorf's own.² John Law stated the water and diamond contrast in 1705, seventy-one years before Smith, and resolved it in the same paragraph: "Water is of great use, yet of little Value; Because the Quantity of Water is much greater than the Demand for it. Diamonds are of little use, yet of great Value, because the Demand for Diamonds is much greater, than the Quantity of them."³ Galiani, in 1751, made value depend on utility and scarcity together, which Schumpeter treats as the high point of the pre-Smith utility tradition.⁴ Condillac published a utility-and-scarcity theory of value in the same year as the *Wealth of Nations*.⁵

So when Smith wrote his paragraph, the distinction was ancient, the paradox was known, and the resolution was in print. What Smith did was different, and it is the thing that matters.

2.2 Smith's decision

Book I, chapter IV of the *Wealth of Nations*:

"The word *VALUE*, it is to be observed, has two different meanings, and sometimes expresses the utility of some particular object, and sometimes the power of purchasing other goods which the possession of that object conveys. The one may be called 'value in use;' the other, 'value in exchange.' The things which have the greatest value in use have frequently little or no value in exchange; and, on the contrary, those which have the greatest value in exchange have frequently little or no value in use. Nothing is more useful than water: but it will purchase scarce any thing; scarce any thing can be had in exchange for it. A diamond, on the contrary, has scarce any value in use; but a very great quantity of other goods may frequently be had in exchange for it."⁶

And the next paragraph opens: "In order to investigate the principles which regulate the exchangeable value of commodities, I shall endeavour to shew, First, what is the real measure of this exchangeable value..."

² Joseph A. Schumpeter, *History of Economic Analysis* (Oxford University Press, 1954), Part II, ch. 2. Pufendorf's own terms are aptitude, *pretium vulgare*, and *pretium eminens*; the use/exchange framing is Schumpeter's.

³ John Law, *Money and Trade Considered* (Edinburgh, 1705), ch. 1.

⁴ Ferdinando Galiani, *Della Moneta* (1751). Characterized in Schumpeter, *History*, Part II, ch. 6, §3(a), and Kauder, p. 24. Neither Galiani nor Pufendorf is quoted from the original; the characterizations are the historians'.

⁵ Étienne Bonnot de Condillac, *Le Commerce et le Gouvernement* (1776).

⁶ Adam Smith, *An Inquiry into the Nature and Causes of the Wealth of Nations* (1776), Book I, ch. IV, para. 13 (Glasgow ed. I.iv.13; Cannan ed. 1904). The following sentence opens para. 14. Smith does not use the phrase "paradox of value"; the label is later. For effectual demand see Book I, ch. VII; for precious stones, Book I, ch. XI.

That is the bracket. Value in use is named, illustrated, and set down. The investigation proceeds on exchange value. Smith does not say value in use is unimportant, and he does not abandon utility entirely: effectual demand governs market price in chapter VII, and in chapter XI he explains the price of precious stones by a demand that “arises altogether from their beauty,” enhanced by scarcity. What he does not do is build on value in use. The measure he goes looking for is a measure of exchange.

The historians of economic thought who take the Austrian side of the story are direct about this, and it is fair to say they are the Austrian side. Schumpeter, in the *History of Economic Analysis*, says that in chapter IV Smith “severs ‘value in exchange’ completely from ‘value in use,’” and that in doing so he was “falling far below the level reached by many older authors and particularly by Galiani.” Elsewhere in the same work:

“Having distinguished value in use and value in exchange, he dismisses the former by pointing to what has been called above the ‘paradox of value’, which he evidently did believe to be a bar to progress on this line, thereby barring, for the next two or three generations, the door so auspiciously opened by his French and Italian predecessors. No talk about his ‘recognizing the role of demand’ can alter this fact.”

Kauder:

*“Adam Smith had an unfortunate influence on the further development of the value explanation. After reading the *Wealth of Nations* many economists reached the conclusion that a further discussion of the value-in-use was meaningless... in his *Wealth of Nations* he misled several generations of readers.”*⁸

Kauder adds that in his Glasgow lectures Smith taught that scarcity and utility determine market price. He knew the resolution and left it out of the book.⁹

There is another side, and a fair paper names it. Samuel Hollander, in *The Economics of Adam Smith*, and Robertson and Taylor before him, argue that Smith had a working theory of demand and that the “dismissal” reading overstates the case.¹⁰ Schumpeter’s sentence about “no talk about his recognizing the role of demand” is Schumpeter waving that side away, not refuting it. Mark Blaug, who is more cautious than either Schumpeter

⁷ Schumpeter, *History of Economic Analysis*, Part II, ch. 3, Reader’s Guide to the *Wealth of Nations* (both the “severs” passage and the “barring the door” passage). Page numbers vary by edition; the Part and chapter references are stable.

⁸ Kauder, p. 28.

⁹ Kauder, p. 28, referring to Smith’s *Lectures on Jurisprudence*.

¹⁰ Samuel Hollander, *The Economics of Adam Smith* (University of Toronto Press, 1973); H. M. Robertson and W. L. Taylor, “Adam Smith’s Approach to the Theory of Value,” *Economic Journal* 67 (1957): 181-198.

or Kauder, adds a correction this paper accepts: the classical economists did not set utility aside because they thought it unmeasurable, but because, on the definition they used, they “saw no relationship between utility, in their sense of the term, and demand.”¹¹

What all sides agree on is enough for this paper. Smith named value in use, and the discipline he founded built its apparatus on the other side of the line.

2.3 Ricardo narrows the field, and the classical exceptions

Ricardo, in the *Principles* of 1817, opens on value by quoting Smith’s paragraph and then makes two moves. First: “Utility then is not the measure of exchangeable value, although it is absolutely essential to it.” Second, he sets aside goods whose value comes from scarcity alone, “some rare statues and pictures, scarce books and coins, wines of a peculiar quality,” because they “form a very small part of the mass of commodities daily exchanged in the market,” and restricts the analysis to “such commodities only as can be increased in quantity by the exertion of human industry.”¹²

So by 1817 the field has been narrowed twice. Value in use is out as a measure. Scarcity goods are out as a class. What remains is the exchange value of reproducible goods, and for those, cost of production is the natural standard. Böhm-Bawerk later described the result: “The stress was thus thrown upon the value of freely reproducible goods. In this way it came about that ‘cost’ was held to be the ‘ultimate standard of value.’”¹³

The bracket was never total, and a paper that pretended otherwise would be caught. Say grounded value in utility in 1803. Ricardo himself, in chapter XX, argues at length with Say and Lauderdale about the difference between value and riches, which is the use and exchange distinction under other names. Senior, Lloyd, Dupuit, and Gossen all worked on utility before 1871, and Mill in 1848 made value in use the upper limit of exchange

¹¹ Mark Blaug, *Economic Theory in Retrospect* (Irwin, 1962; 5th ed. Cambridge University Press, 1997), ch. 2, “The Measure and Cause of Value.”

¹² David Ricardo, *On the Principles of Political Economy and Taxation* (1817; 3rd ed. 1821), ch. I, sec. I. The third-edition text is quoted.

¹³ Eugen von Böhm-Bawerk, “The Ultimate Standard of Value,” trans. C. W. Macfarlane, *Annals of the American Academy of Political and Social Science* 5, no. 2 (September 1894): 149-208, at 152. German original: “Der letzte Maßstab des Güterwertes,” *Zeitschrift für Volkswirtschaft, Socialpolitik und Verwaltung* 3 (1894).

value.¹⁴ The classical exceptions existed. What they did not do was change what the discipline measured.

2.4 The marginalists bring utility back, to explain price

In 1871 two books appeared independently, Jevons's *Theory of Political Economy* in England and Menger's *Grundsätze* in Vienna, followed in 1874 by Walras in Lausanne. Hayek's judgment is the standard one: "the complete independence of the work of the three founders is quite certain."¹⁵ They are usually treated as one revolution, and Jaffé's caution that they should be "de-homogenized" applies with force here: Menger rejected mathematics and measurement outright, while Jevons proposed a calculus of pleasure and pain.¹⁶

Jevons opened with the claim that "value depends entirely upon utility."¹⁷ Menger: "The determining factor in the value of a good, then, is neither the quantity of labor or other goods necessary for its production nor the quantity necessary for its reproduction, but rather the magnitude of importance of those satisfactions with respect to which we are conscious of being dependent on command of the good."¹⁸

This matters to the present argument in a specific way. The marginalists reopened value in use as the thing that explains exchange value. Schumpeter: they "established what A. Smith, Ricardo, and Marx had believed to be impossible, namely, that exchange value can be explained in terms of use value."¹⁹ Some of them, Jevons and later Edgeworth, also proposed to measure utility directly, and in the twentieth century Fisher and Frisch tried.²⁰ What none of them proposed, because it was not their question, was to check a delivered outcome against a promise made before purchase. Utility entered economics as a determinant of price. It did not enter as a quantity a buyer verifies afterward.

¹⁴ Jean-Baptiste Say, *Traité d'économie politique* (1803); Ricardo, *Principles*, ch. XX, "Value and Riches, Their Distinctive Properties"; Nassau Senior, *An Outline of the Science of Political Economy* (1836); W. F. Lloyd, *A Lecture on the Notion of Value* (1833); Jules Dupuit (1844), cited below; H. H. Gossen, *Entwicklung der Gesetze des menschlichen Verkehrs* (1854); J. S. Mill, *Principles of Political Economy* (1848), Book III, ch. I.

¹⁵ F. A. Hayek, introduction to Carl Menger, *Principles of Economics*, trans. Dingwall and Hoselitz (1934 introduction; Mises Institute ed.).

¹⁶ William Jaffé, "Menger, Jevons and Walras De-Homogenized," *Economic Inquiry* 14, no. 4 (1976): 511-524.

¹⁷ W. S. Jevons, *The Theory of Political Economy* (Macmillan, 1871), ch. I.

¹⁸ Carl Menger, *Grundsätze der Volkswirtschaftslehre* (Braumüller, 1871), ch. III; Dingwall and Hoselitz trans., pp. 146-147.

¹⁹ Schumpeter, *History*, Part IV, on the marginal utility revolution.

²⁰ F. Y. Edgeworth, *Mathematical Psychics* (1881); Irving Fisher, *Mathematical Investigations in the Theory of Value and Prices* (1892); Ragnar Frisch, *New Methods of Measuring Marginal Utility* (1932).

2.5 Böhm-Bawerk, 1894, a partisan witness

In 1894 Eugen von Böhm-Bawerk published “The Ultimate Standard of Value” in the *Annals of the American Academy of Political and Social Science*, translated by C. W. Macfarlane.²¹ The paper is an intervention in the dispute between utility and cost as the ultimate determinant of value, and its conclusion is that cost is a proximate leveling mechanism while marginal utility governs: “The cost therefore exercises, as it were, only a vice-regency.”²²

On page 151 he describes what the classical theory had done:

*“The classical theory of value, as is well known, divided the domain of the phenomena of value. A distinction was drawn between ‘value in use’ and ‘value in exchange.’ The ‘value in use’ of goods was thought to rest entirely upon utility, but beyond this passing reference to the domain of utility the classical theory did not trouble itself about value in use.”*²³

Böhm-Bawerk was fighting a dispute when he wrote that, and he was fighting it on the Austrian side. His “did not trouble itself” is an overstatement, as section 2.3 shows. Edgeworth answered him that September, conceding the main point under Ricardian assumptions while insisting that disutility remained a co-ordinate standard, and Böhm-Bawerk replied in December.²⁴ Nobody in that exchange was thinking about buyers. Böhm-Bawerk is not the source of this paper’s argument, and his sentence is quoted here as a contemporary’s description of the classical emphasis, not as a neutral historian’s verdict. It is accurate about the emphasis. It is not the last word about the exceptions.

2.6 What Marshall settled, and what he set aside

A careful reader will object that Marshall resolved the utility-versus-cost dispute in 1890 with the scissors: “We might as reasonably dispute whether it is the upper or the under blade of a pair of scissors that cuts a piece of paper, as whether value is governed by utility or cost of production.”²⁵

²¹ Böhm-Bawerk, “The Ultimate Standard of Value,” cited above. SAGE DOI 10.1177/000271629400500201. The *Annals* text and its offprint are in the public domain; cite the 1894 original rather than any print-on-demand reproduction.

²² Böhm-Bawerk, p. 153.

²³ Böhm-Bawerk, p. 151.

²⁴ F. Y. Edgeworth, “Professor Böhm-Bawerk on the Ultimate Standard of Value,” *Economic Journal* 4, no. 15 (September 1894): 518-521; Böhm-Bawerk, “One Word More on the Ultimate Standard of Value,” *Economic Journal* 4, no. 16 (December 1894): 719-724. Modern treatment: Motohiro Okada, “Revisiting the Böhm-Bawerk-Edgeworth Controversy,” *Review of Political Economy* 29, no. 4 (2017): 636-651.

²⁵ Alfred Marshall, *Principles of Economics* (Macmillan, 1890; 8th ed. 1920), Book V, ch. III, §7.

Marshall settled the question of what determines price: utility and cost jointly, with the balance shifting by time period. Value in use survives in Marshall as consumers' surplus, which he develops in Book III and calls "the economic measure of this surplus satisfaction."²⁶ That is a genuine measure of value in use, and it is the ancestor of every instrument discussed in the next section. What Marshall did not do, and did not claim to do, was apply it to a firm checking whether a supplier's promise was kept. The price question was closed. The audit question was not opened.

2.7 The instrument economics did build, and where it went

Here the paper owes a concession, because a hostile economist will make it if the paper does not. Economics did build an instrument for measuring value in use. It built several, and they are in daily use.

Dupuit, in 1844, wrote a paper titled "On the Measurement of the Utility of Public Works," which is a buyer-side measurement of value in use by the state.²⁷ Marshall's consumers' surplus gave the concept its standard form. Hotelling in 1938 and Hicks in 1941 gave it rigour. The U.S. Flood Control Act of 1936 required that the benefits of federal works exceed their costs, which made benefit measurement a legal obligation.²⁸ Cost-benefit analysis grew from there into a profession with its own journals and standards. The UK Treasury's Green Book requires it for public spending and its Magenta Book requires evaluation after the fact.²⁹ Contingent valuation, hedonic pricing, and travel-cost methods were built to measure the value in use of things with no market price. NICE, in the UK, decides what the health service will buy by measuring value in use in quality-adjusted life years, and the seller does not hold the meter.³⁰ Flyvbjerg's work on megaprojects is a published systematic comparison of pre-decision projections with realized outcomes, which is the very thing this paper will say is missing from enterprise technology.³¹

²⁶ Marshall, *Principles*, Book III, ch. VI, "Value and Utility."

²⁷ Jules Dupuit, "De la mesure de l'utilité des travaux publics," *Annales des ponts et chaussées* (1844).

²⁸ Harold Hotelling, "The General Welfare in Relation to Problems of Taxation and of Railway and Utility Rates," *Econometrica* 6 (1938); J. R. Hicks, "The Rehabilitation of Consumers' Surplus," *Review of Economic Studies* 8 (1941); Flood Control Act of 1936, 33 U.S.C. §701a.

²⁹ HM Treasury, *The Green Book: Central Government Guidance on Appraisal and Evaluation*; HM Treasury, *The Magenta Book: Central Government Guidance on Evaluation*.

³⁰ National Institute for Health and Care Excellence, *NICE health technology evaluations: the manual*.

³¹ Bent Flyvbjerg, Nils Bruzelius, and Werner Rothengatter, *Megaprojects and Risk* (Cambridge University Press, 2003), and subsequent work.

So the claim that no instrument exists is false, and this paper does not make it. The claim it makes is narrower, and it survives. Consumers' surplus and its Hicksian refinements were pure theory. The applied instruments built on them, from Dupuit onward, were built for public goods, public works, and public spending, because the state is a buyer that must justify its purchases to people who did not choose them. Hedonic pricing and real-options analysis have private uses, and the balanced scorecard measures non-price value inside a firm; none of them verifies a supplier's promise.

What migrated into private technology purchasing is the front half of the public method. The discounted business case that accompanies every large system purchase is cost-benefit appraisal by another name. What did not migrate is the back half: the mandated evaluation, after the fact, of whether the appraised benefit arrived. The Green Book requires appraisal; the Magenta Book requires evaluation; a private company buying a system is bound by neither, runs no Gateway review, and is required by no one to look back. The welfare economists built the buyer's instrument in two halves. Private business took the half that helps a purchase get approved and left the half that checks it at the door of the public sector.

The one private-sector exception worth naming, because a reviewer will raise it, is the outcome-based pharmaceutical contract, in which a private insurer pays a drug maker according to measured patient outcomes. Those exist, they are metered by the payer, and they show that a buyer-held meter is possible in private commerce. They also show how rare it is: a handful of contracts, in one industry, where the outcome is clinical and the payer already holds the clinical record.

There is also a later moment, better documented than Smith's, when economics formally stopped trying to measure utility as a magnitude at all. Pareto in 1906, Robbins in 1932, and Hicks and Allen in 1934 established the ordinal view: preferences can be ranked, not measured, and comparisons between persons are outside the science.³² That turn is the one a modern economist will recognize, and it does not undermine the paper's story. It confirms it. The discipline that had set value in use aside for price theory in 1776 declared it unmeasurable for pure theory in 1934, and left the measuring to the applied branches that served the state.

³² Vilfredo Pareto, *Manuale di economia politica* (1906); Lionel Robbins, *An Essay on the Nature and Significance of Economic Science* (1932); J. R. Hicks and R. G. D. Allen, "A Reconsideration of the Theory of Value," *Economica* 1 (1934).

2.8 Marx, placed

Marx opens *Capital* with the same pair. “The utility of a thing makes it a use value... Use values become a reality only by use or consumption.” And: “Exchange value, at first sight, presents itself as a quantitative relation.”³³ For Marx, use-value is an objective property of the commodity, realized in consumption and abstracted from in exchange, which is why his analysis moves immediately to labor as the common substance. For Menger, value is a subjective magnitude imputed by the valuing person. Marx brackets use-value on purpose, as a method. Either way, the quantity that matters to a buyer, what the thing did once bought, is not on his ledger either.

2.9 What the record supports and what it does not

What the record supports: the classical economists named value in use, set it aside as a measure, and built the discipline on exchange value. The marginalists brought utility back to explain price. Welfare economics built real instruments for measuring value in use and applied them to public spending. Marshall closed the price question. Historians on the Austrian side say the bracket was Smith’s error; historians on the other side say it was a reasonable emphasis; both agree on what the discipline went on to measure.

What the record does not say: that Smith made his choice “because prices are observable” or “for analytical convenience.” Those are this paper’s interpretations. The closest textual support is Smith’s own pivot, from utility to “the power of purchasing other goods,” and Ricardo’s restriction to goods “increased in quantity by the exertion of human industry.” I hold the observability reading as a reasonable one, and I hold it as mine.

And the claim that private business measurement inherited the bracket while public measurement did not is the paper’s own argument. The next sections make it.

3. The bracket goes institutional in private business

3.1 Measurement follows the numbers

Once economics had settled on exchange value as its object, the professions that measured private business followed. Accounting records transactions, which are exchanges. Capital budgeting discounts cash flows, which are exchanges anticipated. Financial reporting

³³ Karl Marx, *Capital*, vol. 1 (1867), ch. 1, §1, trans. Moore and Aveling (1887).

reconciles to money moved. None of this is a criticism. Audit requires a record, and exchange leaves one.

Corporate finance does have a buyer-side instrument on paper. The post-completion audit of capital projects has been in the textbooks since the 1960s, and surveys in the 1990s found that most large firms claimed to conduct them.³⁴ Two things are true about it. It is weakly practiced, which every study of it says. And it is applied to the firm's own projects, not to a vendor's claims about what a purchased capability would deliver. The instrument exists in finance. It has never been pointed at the supplier.

Three twentieth-century developments are worth placing, because each is sometimes offered as the missing instrument and none of them is.

3.2 Miles, 1947: the buyer asks what function he is buying

Lawrence D. Miles was a purchasing agent at General Electric. In 1947 he was assigned by Harry Erlicher, GE's vice president of purchases, "to develop a system which would significantly improve the results of the company's cost reduction efforts."³⁵ What he built was value analysis, and its founding move was to change the question: not "what material am I buying?" but "what function am I buying?"³⁶

Miles defined value in the 1961 first edition of *Techniques of Value Analysis and Engineering*: "Value is stated as the minimum dollars which must be expended in purchasing or manufacturing a product to create the appropriate use and esteem factors." And: "Use value: The properties and qualities which accomplish a use, work, or service."³⁷ The U.S. Navy Bureau of Ships adopted the method in 1954 and, because its function was engineering, renamed it value engineering.³⁸ The Society of American Value Engineers was founded in 1959 with Miles as its first president.³⁹

³⁴ Richard A. Brealey and Stewart C. Myers, *Principles of Corporate Finance* (McGraw-Hill, various editions), on post-audit; Lawrence A. Gordon and Mary D. Myers, "Postauditing Capital Projects: Are You in Step with the Competition?," *Management Accounting* (January 1991).

³⁵ SAVE International, "The History of the First Five Years of the Society," written by the Society's historians through 1964.

³⁶ As reported in Robert B. Stewart, *Value Optimization for Project and Performance Management* (Wiley, 2010), ch. 1.

³⁷ Lawrence D. Miles, *Techniques of Value Analysis and Engineering* (McGraw-Hill, 1961), pp. 1, 3. The often-repeated one-line definition, "value is the lowest price you must pay to provide a reliable function or service," does not appear in the 1961 text.

³⁸ SAVE International history, cited above; U.S. Department of Defense, SD-24 Value Engineering Guidebook (February 2025), also gives 1954.

³⁹ SAVE International history. Organizing meeting 17 March 1959; incorporation 22 October 1959; Miles elected first president 5 December 1960.

Miles was a buyer. His method was about function, about what the purchased thing was for. And the term “value engineering,” which the enterprise software industry adopted for a sales function half a century later, was coined by a Navy procurement bureau for a buyer-side discipline. When SAP’s Bill McDermott described his company’s “value engineering” program to Computerworld in 2004 as something that “started a year ago,” he was using, for a vendor’s sales apparatus, a name the buyer’s side had held since 1954.⁴⁰

But value analysis was a cost-reduction discipline for physical products. It asked the minimum cost to deliver a function. It did not ask, after purchase, whether a promised business outcome had arrived. It was the closest thing the private buyer’s side built, and it stopped at the loading dock.

3.3 TCO, 1987: the cost side, fully instrumented

Gartner popularized total cost of ownership in 1987, with Bill Kirwin credited as the originating analyst.⁴¹ Lisa Ellram gave it academic form in the purchasing literature in 1993 and 1995.⁴²

TCO is a real instrument and a good one. It answers what an asset will cost across its life, which buyers had been answering badly. But it is a cost instrument. It sits on the exchange side of Smith’s line, and its post-purchase use is supplier cost performance, not benefit verification. UK government purchasing guidance says so directly: TCO “does not assess how well a solution offers value for money, or how well it fits with an organisation’s strategic objectives.”⁴³ It instrumented the half of the purchase that was already the better-instrumented half.

3.4 EVA, 1983: the shareholder’s measure

Stern Stewart developed Economic Value Added in 1983, put it into commercial use in 1989, popularized it in Bennett Stewart’s *The Quest for Value* in 1991, and registered it as a trademark in 1994.⁴⁴ It measures residual return on capital employed. It is a measure of

⁴⁰ Bill McDermott interview, *Computerworld*, 22 April 2004.

⁴¹ Total cost of ownership popularized by Gartner in 1987; Bill Kirwin is credited as the originating analyst in multiple secondary sources. A Gartner-authored page naming Kirwin was not located; the attribution is well supported but secondary.

⁴² Lisa M. Ellram, “Total Cost of Ownership: Elements and Implementation,” *International Journal of Purchasing and Materials Management* 29, no. 4 (1993): 2-11; Ellram, “Total cost of ownership: an analysis approach for purchasing,” *International Journal of Physical Distribution & Logistics Management* 25, no. 8 (1995): 4-23.

⁴³ UK Cabinet Office and Home Office, “Total Cost of Ownership: things to consider,” v1, published on gov.uk.

⁴⁴ Stern Value Management company history (developed 1983); USPTO registration no. 1,856,961 (first use in commerce 24 November 1989; registered 4 October 1994); G. Bennett Stewart III, *The Quest for Value* (HarperBusiness, 1991).

the value a firm creates for the providers of its capital, not of the value a firm received from a thing it bought. It is mentioned only because “value” in its name leads people to expect otherwise.

By the end of the 1980s, a private buyer had instruments for cost across the life of an asset, for the function-per-dollar of a physical product, for the return the firm delivered to its investors, and, on paper, for auditing its own capital projects. It had no instrument, in use, for whether a purchased capability delivered the outcome it was bought for. In the private sector, the bracket was fully institutional.

4. The marketing turn: value in use is named and handed to the supplier

4.1 Economic value to the customer, 1978

The first commercial instruments for value in use in business markets came from marketing, and they came a decade before the phrase “value proposition.” Shapiro and Jackson in 1978 set out pricing to the customer’s economic gain, and Forbis and Mehta in 1981 gave it the name that stuck, economic value to the customer: a supplier’s calculation of what its product was worth to a buyer relative to the alternative, used to set price.⁴⁵ The supplier computes the customer’s value in use in order to capture a share of it. That is the pattern, and it did not change.

4.2 The value proposition, 1988

In 1988, two McKinsey consultants, Michael Lanning and Edward Michaels, wrote a staff paper called “A Business Is a Value Delivery System.” It introduced a phrase: “A business unit promises customers some value, a combination of benefit and price. It thus offers a value proposition.”⁴⁶ The business proposes. The customer receives a proposition.

⁴⁵ Benson P. Shapiro and Barbara B. Jackson, “Industrial Pricing to Meet Customer Needs,” *Harvard Business Review* (November-December 1978); John L. Forbis and Nitin T. Mehta, “Value-based strategies for industrial products,” *Business Horizons* 24, no. 3 (1981): 32-42.

⁴⁶ Michael J. Lanning and Edward G. Michaels, “A Business Is a Value Delivery System,” McKinsey Staff Paper no. 41 (June 1988).

4.3 Customer value models, 1998

Anderson and Narus, in the *Harvard Business Review*: “How do you define value? Can you measure it? What are your products and services actually worth to customers? Remarkably few suppliers in business markets are able to answer those questions.” Their method is the customer value model, and its purpose is stated plainly: “By creating and using what the authors call customer value models, suppliers are able to figure out exactly what their offerings are worth to customers... they can use value models to provide evidence to customers of their accomplishments.”⁴⁷ The 2007 book that followed carries the purpose in its subtitle: *Value Merchants: Demonstrating and Documenting Superior Value in Business Markets*.⁴⁸

This is the point at which value in use comes back into private commercial practice as a quantity someone proposes to measure, and it comes back as a supplier’s instrument. The supplier builds the model, documents the value, and provides evidence of its accomplishments. The customer is the subject of the measurement, not its author.

4.4 Service-dominant logic, 2004: named, located, and left without an instrument

In January 2004, Vargo and Lusch published “Evolving to a New Dominant Logic for Marketing” in the *Journal of Marketing*.⁴⁹ On page 6 it contains marketing’s own acknowledgment of the bracket:

“At that time, Smith’s focus on exchange value represented a departure from the more accepted focus on value in use, and it had critical implications for how economists, and later marketers, would view exchange.”⁵⁰

Foundational premise seven states that “the enterprise can only make value propositions”; “the consumer must determine value.”⁵¹ The 2008 revision added premise ten: “Value is always uniquely and phenomenologically determined by the beneficiary.”⁵²

⁴⁷ James C. Anderson and James A. Narus, “Business Marketing: Understand What Customers Value,” *Harvard Business Review* 76 (November-December 1998): 53-65.

⁴⁸ James C. Anderson, Nirmalya Kumar, and James A. Narus, *Value Merchants: Demonstrating and Documenting Superior Value in Business Markets* (Harvard Business School Press, 2007).

⁴⁹ Stephen L. Vargo and Robert F. Lusch, “Evolving to a New Dominant Logic for Marketing,” *Journal of Marketing* 68, no. 1 (January 2004): 1-17.

⁵⁰ Vargo and Lusch (2004), p. 6.

⁵¹ Vargo and Lusch (2004), pp. 10-11, foundational premise 7.

⁵² Stephen L. Vargo and Robert F. Lusch, “Service-dominant logic: continuing the evolution,” *Journal of the Academy of Marketing Science* 36, no. 1 (2008): 1-10, foundational premise 10; restated as Axiom 4 in Vargo and Lusch, *Journal of the Academy of Marketing Science* 44, no. 1 (2016).

By 2008, marketing had named value in use as the real thing, located its determination in the buyer, and made the enterprise's role the offering of propositions. What it did not do was propose any instrument by which the beneficiary verifies, after purchase, that the proposition was kept. Service-dominant logic is a theory of exchange written for marketing scholars. The operational work that followed measures the supplier's orientation toward service-dominant thinking, not the buyer's realized outcome.⁵³

The vocabulary was complete. The instrument gap it opened was about to be filled, on the seller's side, by people whose job was to close deals.

5. The seller inherits the question

5.1 The function is named

By the vendor's own account, enterprise software value engineering was named and staffed as a sales function around 2003. Chakib Bouhdary, who became SAP's Chief Value Officer, described it:

*"When we came up with the value engineering program, the value engineering program was one of the most dramatic changes we made in the way we go to markets... It's basically at our costs, we'll do the analysis... We have a benchmarking database... about 50 surveys that we developed with 10,000 customers... We've been doing this since 2003."*⁵⁴

He called it a "reverse selling model." The vendor benchmarks the prospect against its customer database, identifies what it calls "leakage of value," and produces the return case and the key performance indicators. Oracle's Insight program is described on Oracle's own site as a "complimentary service" that brings in "experienced teams from top-tier strategy firms" to "develop a compelling business case."⁵⁵ A recruiter in the field summarizes the origin as the industry remembers it: the first business value teams were created "nearly two decades ago" at companies like SAP and Oracle, staffed by people from the major strategy firms.⁵⁶

⁵³ Ingo O. Karpen, Liliana L. Bove, and Bryan A. Lukas, "Linking Service-Dominant Logic and Strategic Business Practice: A Conceptual Model of a Service-Dominant Orientation," *Journal of Service Research* 15, no. 1 (2012): 21-38.

⁵⁴ Chakib Bouhdary, interview, Apps Run The World, published 10 September 2015; content references the SAP co-CEO era, so the interview dates to approximately 2013.

⁵⁵ Oracle Insight, oracle.com/corporate/insight, accessed September 2026. Secondary sources report a 2003 launch; that date is not verified from Oracle and is not asserted in the text.

⁵⁶ The Value Search Team, "5 Key Skills Your Value Engineers Need," September 2024. A practitioner's recollection, cited as such.

I was one of those people. The account that follows is mine, and I give it as evidence rather than as complaint.

5.2 What we built, and what it was for

The apparatus had four parts, and the same four appear in every vendor's version.

The benchmark: a database of customer key performance indicators, assembled from surveys, used to place a prospect against a peer set and identify the gap. That gap became the “value leakage.”

The business case: a model that converted the leakage into a projected financial return, discounted over a period, compared against the purchase and implementation cost, and presented as a return on investment.

The value hypothesis: a mapping from product capabilities to the drivers of the projected benefit. If you deploy this module, this process gets faster; if this process gets faster, inventory turns more often; if inventory turns more often, this much working capital is released. Every number in the business case sat at the end of a chain like that.

And the value realization review: a schedule for re-surveying the customer's key performance indicators after go-live and comparing them to the baseline and to the forecast.

I want to be precise about what these were for, because the seller's case in the next section depends on it, and because a hostile reader will look for a contradiction between this section and that one. The apparatus was built to get a purchase decision made, and at that it was effective. The benchmark created urgency. The business case gave the champion a number to carry to the finance committee. The value hypothesis gave the number a story. The realization review gave the customer confidence that someone would be accountable afterward. I never saw a business case built with numbers anyone believed to be false. That is not the problem.

Here is the load-bearing observation. The inputs to the business case came from the customer. The structure of the model came from us. The benchmark came from us. The value hypothesis came from us. The tool that ran the realization review came from us, and the review compared the customer's outcome to our forecast, using our method, in our

system. At no point did the buyer hold an instrument of their own. They held ours, because they had none.

And here is what happened to the realization reviews, because it answers a question the next section raises. We ran them. Deal by deal, they were real: a re-survey, a comparison, a report. They were run with the customer and signed off by the customer's sponsor. What never happened, in my experience and in anything I have been able to find published since, is aggregation. Nobody added them up. Nobody asked, across a hundred deals, what fraction of the projected value arrived. The reviews existed. The scorecard did not.

5.3 What the apparatus became

Over two decades, the vendor value engineering function became three things.

A discipline with a job title. “Value engineer,” “business value consultant,” “value advisor.” A recognized specialty with its own career path and recruiters.

A software category. Gartner maintains a peer-review market called “Value Selling,” filed under Sales.⁵⁷ Gartner's guidance to technology providers is to move “beyond just ROI calculators” to platforms spanning marketing, selling, customer success, and expansion.⁵⁸ The category claims post-sale operation. It does not claim to operate on the buyer's behalf, because it is sold to the seller, and its post-sale function is described in terms of retention.

And the owner of the post-sale question. The Technology Services Industry Association, the trade body for vendor service organizations, states it in one line: customer success “is responsible for ensuring your company fulfills on the promise to deliver the expected solution outcomes to the customer. They are the value realization engine for your organization.”⁵⁹ The vendor named the post-sale question, kept it, and put it in a department whose performance metric is net revenue retention, which is the fraction of last year's subscription revenue that renews and expands.

The buyer, whose money is at risk, is still where Smith left them. They can see what they paid. They cannot see, on any instrument of their own, what it did.

⁵⁷ Gartner Peer Insights, “Value Selling” market, filed under Sales, accessed September 2026.

⁵⁸ Gartner, “Use Value Management Platforms to Assess and Convey Value to Buyers and Customers,” research note 4009169, as summarized publicly by a vendor in that market. The note was not accessed directly.

⁵⁹ Jared Raftery, TSIA blog, 20 May 2021; TSIA customer growth research page, 2026.

6. The seller's case, and where it breaks

6.1 The case, made at full strength

If this paper is read by the people who run vendor value organizations, and I hope it is, they will make an argument, and it is a good one. I will make it for them at full strength, because a weak version would be recognized instantly by anyone who makes the argument for a living.

Before roughly 2003, enterprise technology was bought on feature checklists, reference visits, and executive relationships. Vendor value engineering introduced quantification where there was none. It built benchmark databases from thousands of customer surveys, which no single buyer could build, because no single buyer sees hundreds of implementations. It created a vocabulary of key performance indicators that let a chief financial officer and a chief information officer argue about the same number. It imposed a discipline on the buyer: name the outcome before you sign. It did this at the vendor's expense, with people from the best strategy firms, and it turned "trust me" into "here is the model, here are your inputs, challenge them."

Buyers needed it. Research cited by the vendors themselves finds that only 17 percent of companies have a formal value management practice of their own.⁶⁰ The peer-reviewed literature confirms the behavior works: value-based selling improves salesperson performance beyond conventional approaches, in a study of 944 salespeople and managers, and the same study finds that organizational value assessment tools help.⁶¹ Value quantification capability improves firm performance.⁶² Vendor-sponsored survey data associates value management adoption with better on-time, on-budget, and on-value delivery.⁶³

The realization reviews were real, they were run with the customer, and they were signed by the customer's sponsor. Vendors do not publish aggregate outcome comparisons

⁶⁰ Corporate Executive Board, Applications Executive Council research, undated, as cited in a vendor value advisory presentation to the Americas' SAP Users' Group, October 2023.

⁶¹ Harri Terho, Andreas Eggert, Wolfgang Ulaga, Alexander Haas, and Eva Böhm, "Selling Value in Business Markets: Individual and Organizational Factors for Turning the Idea into Action," *Industrial Marketing Management* 66 (2017): 42-55; 944 salespeople and managers in 43 Finnish sales organizations. See also Terho, Eggert, Haas and Ulaga, *Industrial Marketing Management* 45 (2015), n = 816.

⁶² Andreas Hinterhuber, "Value quantification capabilities in industrial markets," *Journal of Business Research* 76 (2017): 163-178. The effect is on firm performance, not individual sales manager performance, and weakens with market dynamism.

⁶³ SAP/AMR Value Management Survey 2010, 400+ companies, as cited in the October 2023 presentation. A self-reported practice-adoption survey; the multipliers vary across retellings.

because customers forbid the disclosure; non-publication is not non-measurement. And a buyer-held instrument is not neutral either: under outcome pricing, a buyer who meters “resolutions” has every incentive to under-report, and no seller will sign against outcomes it cannot control, like adoption and data quality. The honest arrangement is a jointly agreed definition with audit rights, not a buyer’s unilateral standard.

Twenty years of voluntary adoption, a professional job title, a market category, and a body of supporting research is what earned authority looks like. If buyers had wanted a different apparatus, they had two decades and the money to build one. They did not. That is the case.

6.2 Where it breaks

It breaks in six places, and every one is in the seller’s own evidence or my own experience.

First, the studies cited in the seller’s case measure the seller’s outcome. They take sales performance, or practice adoption, as the thing explained. None of them measures whether the projected number arrived. When that literature says value-based selling “works,” it means it helps deals close. That is the seller’s question, and the answer is yes. It is not evidence about the buyer’s question and was never designed to be.

Second, the one published field test of a value calculator at the deal level is not encouraging, and it should be read carefully. Pöyry, Parvinen and Martens studied 1,193 sales meetings and 301 offers at one international business-to-business service firm. They found that calculator use in sales calls was associated with smaller deals and did not affect conversion or cycle time.⁶⁴ This is one firm, and the artifact studied is a salesperson’s calculator, not the multi-week business case a value engineering team builds; selection is a plausible explanation, if calculators were deployed on smaller opportunities. The finding is suggestive, not general. What it does establish is that the seller’s own literature has not shown the instrument to do even the job it was built for, at the level where it would need to be shown.

Third, the evidence on what actually reduces a buyer’s doubt points at evidence, not models. Anderson and Wynstra, studying what they call “ambiguity about superior value,” the buyer’s doubt about whether claimed savings will materialize, found that

⁶⁴ Essi Pöyry, Petri Parvinen, and Robin-Vincent Martens, “Effectiveness of value calculators in B2B sales work: Challenges at the sales-call level,” *Journal of Business Research* 126 (2021): 350-360.

reference customers and pilot programs were equally effective at reducing it.⁶⁵ What reduced doubt was something the buyer could observe. The model was not among the things tested, and I do not claim the study found against it. I claim that in the purchasing study closest to the question, the things that worked were things that showed the buyer an outcome.

Fourth, the figures the vendors cite contradict the authority claim. A large vendor's value advisory organization, presenting to its user group in 2023, cited research from the Corporate Executive Board showing realized value at 50 to 65 percent of expected, with the largest single tranche of the shortfall, 20 to 25 points, attributed to "poor absorption and value tracking."⁶⁶ The research is CEB's and its date is not given, so it should be read as the vendor's chosen framing rather than a 2023 measurement. But read what the framing says. The vendor's value function presents a 35 to 50 percent gap between projection and outcome, locates the biggest cause in the phase the vendor apparatus does not own, and offers this as a reason to buy more value services. It is an accurate statement of the debt, made by the party that ran the experiment.

Fifth, the reviews were run and never scored. The seller's case is right that realization reviews exist and that customers sign them. I ran them. What has never been published, by any vendor value function I could find after targeted search, is an aggregation: across the customer base, what fraction of projected value arrived.⁶⁷ Confidentiality explains why individual results are not published. It does not explain why no vendor has published the aggregate, which discloses nothing about any customer. A discipline that has run its instrument for twenty years and never published its hit rate has not earned authority over projections. It has earned authority over closing.

Sixth, voluntary adoption is the wrong test. Buyers adopted the vendor's apparatus because it was free and theirs was absent. The 17 percent figure is not evidence that the vendor's apparatus was validated. It is evidence of a vacuum. A person with no scale will use the one the grocer offers, and that tells you about the absence of scales, not the

⁶⁵ James C. Anderson and Finn Wynstra, "Purchasing Higher-Value, Higher-Price Offerings in Business Markets," *Journal of Business-to-Business Marketing* 17, no. 1 (2010): 29-61. The finding quoted is that reference customers and pilot programs "appear to be equally effective as value evidence" in reducing ambiguity about superior value.

⁶⁶ Vendor value advisory presentation to the Americas' SAP Users' Group, October 2023, citing Corporate Executive Board research, undated.

⁶⁷ Targeted searches run 21 September 2026 across vendor value organizations, analyst coverage, and academic databases for any published aggregate comparison of pre-sale projections to realized outcomes across a customer base. None found. Stated as "could not find," not "does not exist." Search terms available on request.

honesty of grocers. Grocers' scales are checked by a weights-and-measures inspector, which is to say, by a standard the grocer did not write. There is no inspector for enterprise software value. Nobody certifies the vendor's benchmark or audits the vendor's realization review. Until there is, the only party in the transaction who can stand in for the inspector is the one whose money is on the scale.

6.3 On the seller's strongest point

The seller's case makes one argument that survives the six breaks, and it deserves a direct answer. A buyer-held meter is not neutral either. Under outcome pricing, a buyer who defines and counts "resolutions" alone has an incentive to under-count, and a seller cannot be held to outcomes that depend on the buyer's own adoption and data.

That is correct, and it does not lead where the seller wants it to. The answer is not that the buyer's meter is neutral. It is that the buyer is the party whose money is at risk, and the standard a purchase is verified against should be written by that party, with the seller responding to it, and with both sides holding audit rights over the count. A buyer-written statement that ignored the buyer's own obligations would be one no seller could sign, and the seller is right to say so. So the statement names them: the adoption the buyer commits to, the data the buyer will provide, the baseline both sides accept, and how the consequence is allocated if the shortfall traces to the buyer's side rather than the seller's. That is what "the seller responds to it" means. The seller does not write the standard; the seller negotiates its terms on the buyer's paper. Both parties act. One of them governs. For two decades, the one that governed was the seller, and the record of that arrangement is in section 8.

6.4 The verdict, in the buyer's terms

I call the vendor-built era a failed experiment, and a fair reviewer will ask whether the word is earned, since nothing was hypothesized, controlled, or scored. Here is the sense in which it is. For twenty years, private buyers ran a live test of a proposition: that an apparatus built for the seller's decision could stand in for the buyer's own. They ran it with real money, at scale, across every major purchase, and they let the seller keep the score. The proposition was testable. The data existed, deal by deal, in the realization reviews. The party holding the data never added it up, and the party at risk never asked. An experiment is not failed by a bad result. It is failed when it runs for two decades and

nobody who could score it does. That is what happened, and the figures in section 8 are what the unscored result looks like from outside.

It left a debt. The instrument for governing value in use in private purchasing, the thing Smith set aside, welfare economics built for the state, and every private profession since declined to build, still does not exist on the buyer's side. The vendor's version exists, is filed under Sales, and is metered by the party being measured. That is the debt, and the next sections show how large it is and why it has just become urgent.

7. The governance vacancy, and what stands around it

7.1 What procurement already does, and where it stops

A procurement leader reading this will say, correctly, that procurement already writes statements of work, acceptance criteria, service levels, and key performance indicators into every master agreement, and that supplier performance management is a mature discipline. All true. The distinction this paper turns on is between activity and outcome.

A service-level agreement measures the seller's activity: availability, response time, throughput. It can be met in full while the business outcome the purchase was made for does not arrive. Performance-based contracting, which has a genuine purchasing literature, moves the measure closer to the outcome, and Selviaridis and Wynstra's review of 241 studies is the reference on it.⁶⁸ But performance-based contracts in the literature are overwhelmingly about physical assets and industrial services, and in most of them the performance measured is the supplier's delivery of the service. Gainshare arrangements go further and measure a shared result, and they are rare, confined to a few contract types, and almost unknown in enterprise software. The nearest thing procurement has built usually stops at the service boundary. It seldom reaches the number the business case promised.

7.2 Public procurement has the instrument

The reason the paper is confined to private enterprise purchasing is that the public sector is different, and the difference proves the point. The UK's Gateway Review 5 is titled "Operations review and benefits realization" and is a formal check, after go-live, on

⁶⁸ Kostas Selviaridis and Finn Wynstra, "Performance-based contracting: a literature review and future research directions," *International Journal of Production Research* 53, no. 12 (2015): 3505-3540.

whether the benefits in the business case arrived.⁶⁹ The Magenta Book requires evaluation of what was delivered. In the United States, the Clinger-Cohen Act of 1996 and OMB Circular A-11 require post-implementation reviews of federal technology investments.⁷⁰ NICE decides what the health service buys by a measured standard of value in use.

Public buyers have the instrument because they are answerable for their purchases to people who did not choose them. Private buyers are answerable to no one in that way, and the instrument was never built for them. The public sector shows that a buyer-held meter is possible, practical, and routine. The private sector shows what happens without one.

7.3 The standard that was absorbed

For a few years, private technology buyers had something close. In 2006 the IT Governance Institute published Val IT, a governance standard for the value of IT-enabled investments, with version 2.0 in 2008.⁷¹ Its purpose was to answer two questions: “Are we doing the right things?” and “Are we getting the benefits?” It had processes for value tracking and for closing out an investment against achieved value, and its own text conceded that those were “rarely carried out today.”⁷² In 2012, ISACA folded it into COBIT 5, where its content survives as the benefits-delivery objectives, and no standalone edition has appeared since.⁷³ Related standards exist: ISO/IEC 38500 on IT governance, the European standard EN 12973 on value management, and PMI’s benefits realization practice guide.⁷⁴ Each holds a piece of the ground. None is a buyer’s instrument for verifying a vendor’s promised value, and no purchasing organization I have encountered uses any of them for that purpose.

Benefits realization management exists as a discipline, built at Cranfield by Ward, Daniel and Peppard.⁷⁵ It is an information systems and project management literature. It governs

⁶⁹ UK Office of Government Commerce, Gateway Review 5: Operations review and benefits realization; now administered by the Infrastructure and Projects Authority.

⁷⁰ Clinger-Cohen Act of 1996 (Information Technology Management Reform Act), Pub. L. 104-106; OMB Circular A-11, Part 7, Capital Programming Guide, on post-implementation review.

⁷¹ IT Governance Institute, *Enterprise Value: Governance of IT Investments, The Val IT Framework 2.0* (July 2008); version 1.0 released March 2006.

⁷² Val IT 2.0, on processes IM8 and IM10.

⁷³ ISACA, *COBIT 5: A Business Framework for the Governance and Management of Enterprise IT* (2012), which “consolidates and integrates the COBIT 4.1, Val IT 2.0 and Risk IT Framework”; the benefits-delivery content survives as objectives EDM02 and APO05 in COBIT 5 and COBIT 2019.

⁷⁴ ISO/IEC 38500:2008, *Corporate governance of information technology*; EN 12973:2000, *Value Management*; Project Management Institute, *Benefits Realization Management: A Practice Guide* (2019).

whether your own program delivered. It does not sit in purchasing, and the profession that owns the contract that promised the benefits has not adopted it.

FinOps governs cloud spend. Technology business management classifies cost. Both have moved their language toward “technology value” in the last two years, and both remain cost disciplines in their instruments. Each of these stands around the hole. What is in the hole is the question Smith set aside: after the exchange, in a private company, did the use arrive? The seller governs that, in a department measured on renewal.

8. The size of the debt

The gap between projected and realized value in enterprise technology is not a secret. It is measured constantly, by the analysts and consultants who advise both sides. What is missing is not the measurement of the gap but any buyer-side instrument for closing it.

- **Bain, June 2026.** Of 951 executives surveyed, 37 percent had targeted AI cost reductions of 11 to 20 percent. Nearly 40 percent of those who measured landed at 0 to 10 percent. Ninety percent of companies whose AI underdelivered plan to increase AI budgets next year. Forty-four percent are funding the next AI wave from prior automation savings that Bain says “have consistently come in below target.”⁷⁶
- **KPMG, July 2026.** Among more than 2,000 leaders at companies above \$50 million in revenue, 7 percent report an established, measurable return on AI investment. Forty-nine percent had delayed or reduced AI initiatives after costs exceeded expected value.⁷⁷
- **Gartner, October 2024.** Among 3,186 CIOs and 1,126 other executives across 88 countries, 48 percent of digital initiatives meet or exceed their business outcome targets.⁷⁸

⁷⁵ John Ward, Elizabeth Daniel, and Joe Peppard, “Building Better Business Cases for IT Investments,” *MIS Quarterly Executive* 7, no. 1 (March 2008); Ward and Daniel, *Benefits Management: Delivering Value from IS & IT Investments* (Wiley, 2006), 2nd ed. retitled *Benefits Management: How to Increase the Business Value of Your IT Projects* (Wiley, 2012).

⁷⁶ Bain & Company, “Your AI Budget Is Growing. Your Returns Aren’t. Here’s Why,” Automation and AI Pathfinder Survey, 1 June 2026, n = 951.

⁷⁷ KPMG Global AI Pulse, Q2 2026, as reported July 2026; 2,000+ leaders, 20 countries, companies above \$50 million revenue.

⁷⁸ Gartner CIO and Technology Executive Survey, press release 22 October 2024.

- **McKinsey and Oxford, 2012.** Across more than 5,400 IT projects, large projects ran 45 percent over budget and delivered 56 percent less value than predicted.⁷⁹
- **Panorama Consulting, 2014.** Sixty-six percent of ERP implementers realized less than half of anticipated benefits.⁸⁰
- **Corporate Executive Board, undated, as cited by a vendor value organization in 2023.** Realized value at 50 to 65 percent of expected.⁸¹

On the most conservative reading, roughly half of enterprise technology initiatives miss their outcome targets, and the realized value of those that proceed runs at somewhere between half and two-thirds of what was projected. That has been true of ERP in 2014, digital initiatives in 2024, and AI in 2026. It has been true for the whole of the vendor value engineering era. Bain's finding that 44 percent of large companies are funding the next wave on the unmaterialized savings of the last describes a system compounding an unverified projection into the next, with no instrument anywhere in the chain that checks whether the first one arrived.

That is the debt. Not that projections were wrong. Projections are always wrong. It is that nobody on the buying side has held the instrument that would say by how much, in time to act.

9. Outcome pricing makes the instrument load-bearing

Until recently the debt could be carried. A projection that missed by a third was a disappointment absorbed into the next budget cycle. The price had been fixed at signing.

That has changed, and the change has dates.

On 28 August 2024, Zendesk announced outcome-based pricing for its AI agents: "Our pricing is now directly tied to the outcomes delivered by AI agents, meaning customers will only incur costs for issues that are resolved autonomously by AI."⁸² On 12 September

⁷⁹ Michael Bloch, Sven Blumberg, and Jürgen Laartz, "Delivering large-scale IT projects on time, on budget, and on value," McKinsey and University of Oxford BT Centre, October 2012.

⁸⁰ Panorama Consulting, 2014 ERP Report, n = 192.

⁸¹ As note 66.

⁸² Zendesk, "Zendesk First in CX Industry to offer Outcome-Based Pricing for AI Agents," 28 August 2024.

2024, Salesforce announced Agentforce at two dollars per conversation.⁸³ On 25 June 2026, Salesforce went further: “With pay-per-resolution pricing, you are charged when Help Agent resolves an issue autonomously, start to finish. If the customer asks for a human or walks away unhappy, there is no charge. Your cost is tied to outcomes, not activity.”⁸⁴

Pricing on outcomes is not new in commerce; cost-per-click advertising has done it since the early 2000s and gainshare outsourcing before that. What is new is its arrival in enterprise software, where the outcome being priced is a business result inside the buyer’s own operation. And the mechanics are the same as they have always been. In every one of these arrangements, the definition of the billable outcome is written by the vendor. The measurement of whether it occurred is performed by the vendor. Zendesk notes that customers “can monitor automated resolution usage” through “an in-product dashboard,” which is the seller’s instrument.⁸⁵

For two hundred and fifty years, the private buyer’s inability to measure value in use was a gap in knowledge. As of 2024, it is a gap in the invoice. When the price is the outcome, the party that defines and meters the outcome sets the price.

Three things a buyer can do about that now, before any larger instrument exists. Negotiate the contractual definition of the billable outcome, because “resolution” is a word and words are negotiable. Secure audit rights over the meter. And use the data the buyer already holds: the tickets, the transcripts, and the record of every time a customer asked for a human are in the buyer’s own systems, and they can check the vendor’s count. Most buyers do not, and most contracts do not say they may.

10. Resolving the debt

10.1 The buyer writes the standard

The resolution follows from the record. A verification instrument designed, operated, and metered by the party being verified is an instrument of sales, however honest the people

⁸³ Salesforce press release, 12 September 2024.

⁸⁴ Salesforce, “Salesforce Launches Agentforce Help Agent That Deploys in Minutes and Only Charges for Resolutions,” 25 June 2026.

⁸⁵ Zendesk, 28 August 2024, cited above.

running it, and the evidence of two decades is that it was never scored against the buyer's question because it was never built to be.

So the buyer writes the standard. The buyer, not the seller, states the business outcome the purchase is for, the measure by which it will be verified, and the consequence if it is not. The statement is made before the purchase, in the buyer's terms, on the buyer's document. The seller responds to it. Verification runs afterward against that document, with audit rights on both sides over the count.

A procurement leader will say that outcome measures already appear in contracts: gainshare clauses on cost savings, availability targets in performance-based logistics, first-contact-resolution rates in outsourcing. They do, and the value statement differs from them in three ways, not one. First, the measured quantity sits in the buyer's own ledger, in the working capital released or the revenue retained, rather than at the service boundary where the seller's activity is counted. Second, the baseline and the attribution rule are fixed before signing, so that the later comparison has something to compare against. Third, the buyer authors the document. A service level can be met in full while the business outcome the purchase was made for does not arrive, and in the figures in section 8 it routinely is. A gainshare clause is closer, and it is the exception that shows the rule: it exists in a few contract types, in a few industries, and almost never in enterprise software.

10.2 The hard part is attribution

A procurement leader will ask the right question here: how does a buyer isolate the vendor's contribution from everything else that changed? The market moved, the headcount changed, three other projects went live. This is the reason buyers accepted vendor models in the first place. Attribution is expensive, and the vendor's benchmark was a cheap substitute for it.

The honest answer is that attribution is never perfect and does not need to be. It needs to be stated in advance. A value statement that names its measure and its baseline before signing has already done most of the attribution work, because it has fixed what will be compared to what. What destroys attribution is not noise; it is the absence of a prior statement to compare against. The vendor's business case was such a statement, and its defect was authorship, not existence.

I have written elsewhere, at book length, about what such a statement contains and how a buying organization operates it. This paper is not the place for that. Its job was to show why the buyer's instrument has never existed in private purchasing, who has been holding its substitute, and why the substitute has just become a billing mechanism.

10.3 Why sellers who adapt early will be better off

A word to the people I used to work alongside, because this paper will be read as an attack and it is not one.

The vendor value engineering era ends the way every free substitute ends: when the party paying for it builds its own. A seller who has been building value cases for twenty years is better placed than anyone to respond to a buyer's standard, because they already know what a credible claim looks like from the inside. The sellers who suffer in the transition will be the ones who defend the old arrangement, in which they held the instrument, as if it were a right. The sellers who benefit will be the ones who recognize that responding well to a buyer-written standard is a competitive advantage, because a buyer who can verify value can also tell which vendors delivered it.

Value in use was set aside by economics in 1776, measured by economics for the state from 1844, computed by suppliers for pricing from 1978, staffed as a vendor sales function by 2003, and metered by the vendor for billing since 2024. At no point in that sequence has the private buyer held the meter. The party whose money is at risk needs one.

About the Author

Michael Brey spent fifteen years in business value engineering for enterprise software vendors before writing *The Value Standard*, which sets out the buyer-held instrument this paper argues for. He publishes at valuelattice.com. This is a working paper; corrections to any sourced claim are welcome and will be recorded.