SUPPLY CHAIN RESEARCH · EDITORIAL FEATURE

The Automation Payback Period

The Five Assumptions Inside a Warehouse Automation Business Case

Automation business cases resolve to one number. That number rests on five separate forecasts about the next decade: wages, labour scarcity, volume, throughput, and utilization. The payback is only as good as the weakest of them, and nobody goes back to check.

Independent, vendor-neutral research for supply chain and technology leaders

Supply Chain Research | supplychainresearch.com | 2026

Contents

The Five Assumptions Inside a Warehouse Automation Business Case

01 Executive summary

A warehouse automation proposal arrives at the investment committee as a single figure. The system costs a stated amount, it removes a stated quantity of labour, and it therefore pays for itself in a stated number of years. The number is usually somewhere between two and four, which is comfortably inside the horizon over which capital committees are willing to think, and the decision follows. What the single figure conceals is that it is not a measurement but the product of five separate forecasts about the next decade: that labour costs will keep rising at the modelled rate, that labour will remain scarce, that volume will reach the projection, that the system will achieve its designed throughput, and that realized utilization will approximate the design point. Each of those is a forecast, each can be wrong, and the payback figure is only as reliable as the weakest of them.

This article examines those five assumptions against the public record and finds that each has failed somewhere in the past several years, in cases documented through securities filings, bankruptcy dockets, and company disclosures rather than through vendor case studies. It also examines what the payback ratio omits, which is substantial and cuts in both directions: it understates cost by excluding maintenance, spares, software, integration, and the operational flexibility surrendered, and it understates benefit by excluding peak survivability, safety, and service consistency. We say honestly that automation frequently earns its capital and that a business built on manual labour at scale has its own accumulating problems. The argument here is not against automating but against underwriting a decade-long capital commitment on a point estimate that nobody revisits. The remedy is to name the assumptions, model the range, fix the process before mechanising it, prefer reversible steps where demand is uncertain, and go back afterward to compare what the system achieved with what the case promised.

5
separate multi-year forecasts compressed into a single payback figure

86.9%
of one public automation supplier's annual revenue attributable to a single customer

21%
the five-year growth rate one market house cut its mobile-robot forecast to, from 26 percent

The argument in brief

- **The payback is five forecasts, not one number.** Wage growth, labour scarcity, volume, throughput, and utilization are each assumptions about the next decade. Compressing them into a point estimate hides where the case is fragile.
- **The system is sized for the peak and runs at the average.** Capacity is set by a handful of weeks, then the capital charge continues at full rate for the rest of the year while the system runs well below its design point.
- **Volume risk is the one that actually breaks cases.** The public record contains capital-intensive operations whose demand never arrived. A payback computed on forecast volume is a payback computed on a forecast.

- **The ratio is wrong in both directions.** It omits maintenance, spares, software, integration, and lost flexibility on the cost side, and peak survivability, safety, and consistency on the benefit side. Fixing only one side produces the opposite error.
- **Prefer the reversible step, then go back and check.** Where demand is uncertain, option value is worth a worse payback figure. And almost nobody re-tests the business case against realized performance, which is the only way the discipline improves.

02 Five forecasts wearing one number

The structural problem with a payback period is not that it is calculated incorrectly. In most business cases the arithmetic is sound and the inputs are defensible on the day they are assembled. The problem is that the output format destroys the information a decision-maker most needs, which is where the case is fragile. Figure 1 unpacks what the single figure contains.

Figure 1. A payback period is five forecasts wearing one number. Labour cost, labour scarcity, volume, throughput, and utilization are each multi-year assumptions, and the payback fails if any one of them does.

Consider what has to be true for a three-year payback to be realised. Labour cost must rise over the period at approximately the modelled rate, because the saving is calculated against the labour the system displaces and a slower rise means a smaller saving. Labour must remain difficult enough to obtain that the displaced roles would otherwise have been filled at that cost, because a loosening market changes both the price and the urgency. Volume must reach the projection, because the saving per unit only converts into total saving when the units arrive. The system must achieve its designed throughput in the buyer's actual operation, with the buyer's actual product mix and error rates, rather than in the vendor's reference configuration. And the realized utilization must approximate the design point, because a system running at seventy percent of its capacity delivers seventy percent of the modelled saving while costing one hundred percent of the modelled capital.

Each of these is a forecast about a decade, and forecasts about a decade are unreliable in ways everyone acknowledges in the abstract and few reflect in the arithmetic. What makes the payback format particularly unhelpful is that it presents the conclusion with a precision the inputs do not support. A figure of three point two years implies a confidence that the underlying model cannot possibly justify, and it invites a decision process that compares three point two against a hurdle rather than examining which of the five assumptions the organization is least sure about. The number is a summary of a model, and summarising a model well requires showing its sensitivity, not concealing it.

There is a second and subtler problem, which is that the five assumptions are correlated rather than independent. A softening labour market usually accompanies weaker demand, so the scenario in which wage growth disappoints is frequently the same scenario in which volume disappoints, and the two errors compound rather than offsetting. A product mix shift that reduces achievable throughput often also changes the volume profile. Treating the assumptions as independent, which a point estimate implicitly does, understates the probability of the joint downside because it assumes the bad outcomes will not arrive together. In practice they arrive together more often than not, because they share underlying causes in the demand environment.

None of this argues against building the model, which is necessary, or against automating, which is frequently correct. It argues for a different output. A business case that reports a range, identifies which assumption the range is most sensitive to, and states what the organization would need to believe for the investment to fail, gives a committee something to reason about. A business case that reports three point two years gives the committee a number to compare against a threshold, which is a decision procedure that cannot distinguish a robust case from a fragile one.

There is a governance reason the point estimate persists that is worth naming, because it explains why the format survives despite being known to be inadequate. Capital committees operate under time pressure across many proposals and need a comparable basis on which to rank them, and a single payback figure supplies exactly that: it is one number, it is comparable across dissimilar investments, and it maps directly onto a threshold. A range with exposed sensitivities is more informative and considerably harder to rank against other proposals presented as points. The format persists because it serves the process, not because anyone believes it captures the risk.

The way out of that bind is not to abandon comparability but to standardise on a richer summary. A committee that required every capital proposal to state its central estimate, its downside scenario, and the single assumption to which the result is most sensitive would retain comparability while restoring the information the point estimate destroys. Proposals could still be ranked, and the ranking would additionally reveal which of them are robust and which are leveraged bets on a forecast. This is a small change in reporting convention with a large effect on the quality of the decision, and it requires no additional analysis, since the underlying model already contains everything needed to produce it.

03 The labour assumption

The labour assumption deserves first examination because it supplies most of the benefit in almost every automation case, and because the way it is usually modelled has a systematic bias. Figure 2 shows the shape of the series that matters.

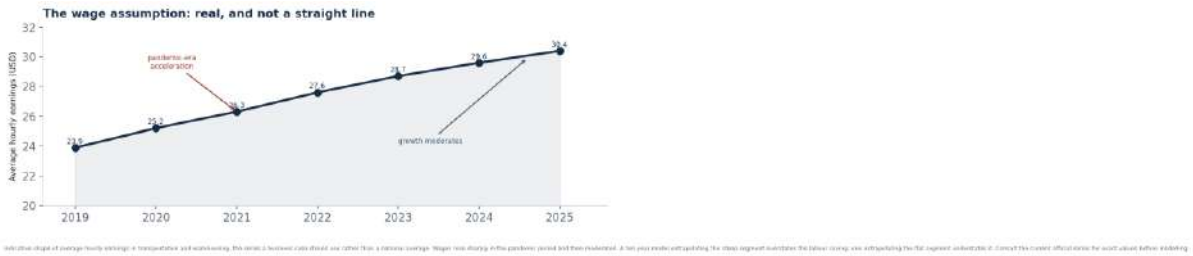

Figure 2. Indicative shape of average hourly earnings in transportation and warehousing. Wages rose sharply in the pandemic period and then moderated. A model extrapolating the steep segment overstates the saving.

The first discipline is to use the right series. National average wage growth is not the relevant input; the relevant input is compensation in transportation and warehousing, in the buyer's own labour market, for the specific roles the system displaces, including the fully-loaded cost of benefits, overtime, agency premiums, and turnover. Business cases frequently use a headline national figure because it is easy to obtain, and the resulting estimate can miss substantially in either direction depending on the local market. An organization operating in a tight regional labour market with heavy agency use may have a fully-loaded cost well above the published average, which strengthens the case; one in a looser market may be well below it.

The second discipline concerns the projection. The sector experienced sharp wage acceleration in the pandemic period and then a moderation, and a model built during or shortly after the acceleration that extrapolates it linearly for ten years will overstate the saving substantially. This is a general hazard of building capital models at the top of a cycle, and it is acute here because the automation investment wave and the wage acceleration coincided, meaning that a large number of business cases were constructed using the steepest available segment of the series as the basis for a decade-long extrapolation. The correction is straightforward: model a range that includes a scenario in which real wage growth in the sector is modest, and observe what that does to the payback.

The third element is labour availability, which is analytically distinct from labour cost and is frequently conflated with it. Scarcity has real value independent of price: an operation that cannot reliably staff a shift has a service problem that no wage rate solves, and automation that removes the dependency is buying certainty rather than saving money. This is a legitimate and frequently decisive benefit, and it should be valued explicitly rather than smuggled into the wage assumption. The distinction matters because the two can move in opposite directions, and an operation that automates purely for scarcity reasons will accept a much longer payback than one automating purely for cost reasons, provided it has said so.

Turnover deserves separate attention as a component of the labour case, because it is where the fully-loaded cost of manual operation is most frequently understated. High turnover imposes recruitment

cost, training cost, a productivity ramp for each replacement, elevated error rates during the ramp, and supervisory load, none of which appears in an hourly rate. Where turnover is high, the true cost of the manual alternative exceeds the wage bill by a meaningful margin, and a business case using the wage rate alone understates the benefit of automation. This is one of the places where a rigorous analysis strengthens rather than weakens the automation case, and it is worth doing properly for that reason.

A further refinement concerns which labour the system actually displaces, because business cases routinely overstate this by assuming that removing a task removes a headcount. In practice an automated system removes a proportion of the work associated with several roles rather than the entirety of any one, and the residual work has to be reorganised before the headcount reduction can be realised. An operation that automates picking may find that the pickers become exception handlers, replenishers, and quality checkers, roles that the automation created or intensified, and that the net reduction is a fraction of the tasks removed. This is not a criticism of the technology; it is a description of how work reorganises around it.

The disciplined method is to model the future-state operating structure explicitly, role by role and shift by shift, rather than converting displaced hours into headcount at an average rate. That exercise forces the organization to specify who does the exception handling, who maintains the equipment, who manages the interface between the automated and manual portions, and what supervisory structure the new operation requires. The resulting figure is usually lower than the arithmetic conversion and considerably more defensible, and producing it before approval avoids the common outcome in which the projected saving is never realised because the operating model was never designed.

04 Sized for the week that pays for the year

The throughput and utilization assumptions are best examined together, because they are joined by a sizing decision that almost every operation makes for sound reasons and that has consequences the payback arithmetic rarely reflects. Figure 3 shows the pattern.

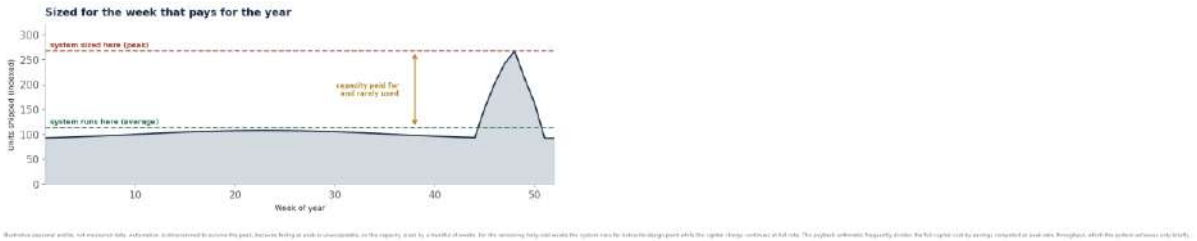

Figure 3. Illustrative seasonal profile, not measured data. Automation is dimensioned to survive the peak, so capacity is set by a handful of weeks while the capital charge continues at full rate through the rest of the year.

An automated system must be sized to handle the operation's peak, because failing at peak is the one outcome nobody will accept. For most consumer-facing supply chains the peak is several times the average, concentrated into a few weeks, and the automation is therefore dimensioned around a demand level that occurs for a small fraction of the year. This is the correct engineering decision. Its financial consequence is that the operation pays a capital charge sized to the peak across all fifty-two weeks, while realising peak-rate savings for perhaps six of them.

The arithmetic error that follows is common and consequential. A business case that computes savings as displaced labour hours multiplied by the system's rated throughput is computing savings at the peak rate and applying them across the year, which materially overstates the return. The correct calculation applies the throughput actually achieved in each period against the volume actually present in that period, which produces a lower and more honest figure. Committees reviewing automation proposals should ask directly whether the savings were computed at rated capacity or at expected realized volume, and the answer is diagnostic: cases built at rated capacity are systematically optimistic.

There is a further gap between rated and achieved throughput that has nothing to do with volume. A system's rated throughput is established under favourable conditions, with a representative product mix, clean data, and everything functioning. Real operations have exception handling, damaged items, mis-picks, mixed unit types, replenishment interruptions, and maintenance windows, all of which reduce effective throughput below the rating. The gap is frequently substantial and it is knowable in advance from reference sites operating similar mixes, provided the buyer asks about achieved rather than rated performance and asks it of a site with comparable characteristics rather than a showcase.

The manual alternative deserves a fair comparison on the same basis, because a labour-based operation flexes in ways an automated one does not. Manual capacity can be increased by adding people and hours, at a price and within limits, so a manual operation sized for the average can surge for the peak by absorbing cost temporarily. An automated system sized for the average cannot surge at all, which is why it is sized for the peak in the first place. The genuine comparison is therefore between a manual operation that carries surge cost during peaks and an automated operation that carries excess capacity year-round, and which is cheaper depends on the shape of the demand profile. Operations with sharp,

narrow peaks and low baselines are the least favourable case for capital-intensive automation, and they are frequently the ones where the peak pain makes automation most attractive emotionally.

The utilization assumption has a further dimension that sizing discussions frequently miss, which is availability. A system rated at a given throughput achieves that rate only while it is running, and scheduled maintenance, unscheduled downtime, software updates, and recovery from jams or faults all subtract from available hours. Mature installations report availability figures well below one hundred percent, and the gap compounds with the throughput gap discussed above: a system achieving eighty-five percent of rated throughput during ninety percent of scheduled hours delivers roughly three quarters of the modelled output. A business case that models neither factor overstates the saving by a quarter before any other assumption is questioned.

Availability is also the dimension on which the manual and automated alternatives differ most sharply in character. A manual operation degrades gradually: if people are absent, throughput falls proportionally and the work continues more slowly. An automated operation tends to fail discretely: it runs at rate or it stops, and when it stops the operation has no partial capacity at all unless a manual fallback has been designed and staffed. That difference should influence both the availability assumption and the contingency planning, and it is a further reason to keep a manual capability alive rather than dismantling it entirely once the automation is commissioned.

05 The volume forecast that did not arrive

Of the five assumptions, volume is the one that has produced the largest recent failures, because it is the assumption on which the entire saving depends and the one over which the buyer has least control. If the units do not arrive, the saving per unit is irrelevant.

The most instructive recent case is not a warehouse but a manufacturing operation, and it is worth examining precisely because the public record is unusually complete. A European battery manufacturer built capital-intensive capacity against projected demand from automotive customers, and filed for Chapter 11 protection in November 2024 with roughly five and eight tenths billion dollars of debt against approximately thirty million dollars of cash. The proximate causes included the cancellation of a major order by one automotive customer earlier that year and reported output far below the plant's nameplate capacity. The company subsequently filed for bankruptcy in its home jurisdiction in March 2025. The volume that justified the capacity did not arrive, and no amount of operational efficiency could compensate, because the fixed cost of the capacity was incurred regardless.

The lesson transfers directly to warehouse automation even though the setting differs, because the financial structure is the same: a large capital commitment made in advance against a demand projection, with a cost base that does not scale down if the projection proves optimistic. The manual alternative has the opposite structure, in that its cost scales roughly with volume because labour can be reduced when demand falls. That difference is the essential trade in the automation decision, and it is systematically underweighted because business cases model the expected volume rather than the distribution of possible volumes. An operation automating on the assumption of growth is taking a leveraged position on that growth.

The corollary is that automation is most defensible where volume is most predictable, and least defensible where it depends on a growth projection that the organization does not control. An operation with stable, mature volumes and a clear labour problem is a strong candidate; an operation automating in anticipation of a demand increase that has not yet materialised is taking a position on that increase and should say so explicitly in the business case. The question a committee should ask is what happens to the return if volume comes in fifteen or twenty percent below plan, which is well within normal forecasting error, and whether the organization could live with the answer.

Volume risk has a structural feature that distinguishes it from the other four assumptions and makes it deserve disproportionate attention. The other assumptions concern the relationship between the system and the work: whether the throughput is achievable, whether the utilization matches the design, whether the labour saved costs what was assumed. Those are properties the organization can investigate, test, and to some degree control. Volume is a property of the market, arriving from customers who owe the organization nothing, and no amount of operational excellence produces units that customers did not order.

This asymmetry should shape how the case is stressed. Operational assumptions can be validated through reference sites, pilots, and careful specification, and the residual uncertainty after that work is modest. Volume uncertainty cannot be reduced by any amount of diligence about the equipment, and it

remains after everything else has been settled. A business case that has been stressed thoroughly on throughput and utilization while accepting the volume projection at face value has examined the assumptions it could control and left untouched the one it could not, which is precisely the wrong allocation of scrutiny. The first question about any automation proposal should concern the confidence interval around the volume forecast, and the answer should determine how much irreversibility the organization is willing to accept.

06 What the market forecasters did next

An organization justifying automation against a growth projection is usually relying, directly or indirectly, on published market forecasts for the sector. Those forecasts have themselves been revised substantially, and the revisions are worth examining. Figure 4 shows them.

Figure 4. Warehouse automation project orders declined in 2024, and a leading market-intelligence house cut its mobile-robot projections after a vendor-by-vendor recount reduced its estimates of earlier years.

The sequence is instructive. Warehouse automation project orders declined by roughly three percent in 2024, against a period in which the prevailing expectation had been continued rapid growth. A leading market-intelligence house then revised its mobile-robot forecasts downward in 2025, lowering its projection for the end of the decade, trimming the five-year growth rate from around twenty-six percent to around twenty-one, and reducing its estimates of the market size in earlier years by roughly eight percent following a vendor-by-vendor recount. The last of these is the most telling, because it is a restatement of history rather than a revision of the future: the market had been smaller than the same house had previously reported.

These figures come from a market-research firm and are an interested source in a specific sense, in that such firms sell forecasts and analysis to both vendors and investors in the sector they cover, and they are flagged accordingly. The direction of the revision is nonetheless the relevant signal, and it is the direction that a buyer building a case on projected sector growth needs to notice. A business case constructed in the enthusiasm phase, using growth assumptions drawn from forecasts that were subsequently cut, inherits the error, and organizations that committed capital on that basis are living with the consequences irrespective of how the forecasts have since been revised.

The broader methodological point is that sector growth forecasts are a poor foundation for an individual operation's business case in any event. What matters to a particular warehouse is its own volume, which depends on its own customers and markets, and the correlation between sector-wide automation adoption forecasts and one operation's throughput is weak. A business case that leans on market growth is frequently doing so because the operation's own volume projection is less encouraging, which is a signal worth attending to. The disciplined approach is to model the operation's own demand under a range of scenarios and to treat sector forecasts as context rather than as input.

There is a general lesson in the recount that goes beyond this sector, and it concerns how market-size estimates are constructed. A market figure is typically built by summing estimated revenues across identified vendors, using a mixture of disclosed figures, company guidance, channel checks, and modelled estimates for the private companies that will not disclose. Each of those inputs carries error, the errors are not independent because the same optimistic environment shapes many of them, and the

resulting total inherits a bias in the direction of the prevailing sentiment. When the sentiment turns and the house revisits its estimates vendor by vendor, the correction arrives all at once and reveals how much of the earlier figure was modelled rather than measured.

For a buyer this has a practical consequence. Market-size and growth figures should be treated as directional indicators of sentiment rather than as measurements, and they should never appear as inputs to a business case that determines a capital commitment. Where a proposal cites sector growth as support, the appropriate question is what the figure would have to be for the case to fail, and whether the organization would notice in time. A case whose viability depends on the accuracy of a third-party market projection has borrowed someone else's uncertainty and added it to its own.

07 Five records worth reading first

The evidence base for this article is deliberately drawn from public filings and formal disclosures rather than from vendor case studies, because case studies are selected for their outcomes and filings are not. Figure 5 assembles five records that an investment committee would find more useful than any reference visit.

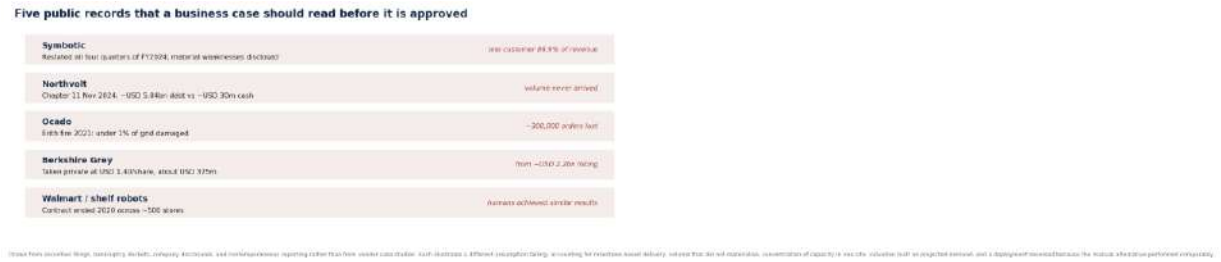

Figure 5. Five public records drawn from securities filings, bankruptcy dockets, company disclosures, and contemporaneous reporting. Each illustrates a different assumption failing.

The first concerns revenue recognition and customer concentration at a major automation supplier. In November 2024 the company restated all four quarters of its prior fiscal year after determining that it had recognised costs, and consequently revenue under percentage-of-completion accounting, before the associated milestones were achieved. It disclosed material weaknesses in internal control and delayed its annual report. Its filings also disclose that a single customer accounted for approximately eighty-seven percent of revenue in that year. Neither fact says anything about whether the systems work, and both are directly relevant to a buyer assessing supplier durability over a ten-year asset life, because a supplier dependent on one customer has a risk profile the buyer inherits.

The second is the battery manufacturer discussed in the previous section, which illustrates volume risk in its purest form. The third is a grocery automation operator whose highly automated fulfilment centre suffered a fire in 2021 caused by a collision between robots, in which damage to well under one percent of the storage grid nonetheless resulted in roughly three hundred thousand lost orders and a material revenue impact, following an earlier and more severe fire at another site. The lesson is about concentration rather than fire safety: automating a large share of capacity into a single densely-engineered facility concentrates operational risk in a way that a distributed manual network does not, and the business case rarely prices that concentration.

The fourth is a robotics supplier that listed publicly at a valuation of roughly two point two billion dollars and was subsequently taken private for approximately three hundred and seventy-five million, at one dollar and forty cents per share, according to its own disclosure. The relevance to a buyer is supplier viability: a ten-year asset requires a supplier capable of supporting it for ten years, and the sector's valuation trajectory is a reasonable proxy for how much consolidation and distress the buyer should expect. The fifth is a large retailer that ended its contract for shelf-scanning inventory robots across approximately five hundred stores in 2020, having planned an expansion months earlier, on the reported basis that human workers achieved comparable results. That case is the cleanest available reminder that the manual alternative is sometimes simply competitive, and that a deployment can be reversed after the capital is spent.

A common objection to using these cases is that each involves circumstances specific to its company, and that a well-run buyer would not repeat them. That objection has some force and it misses what the cases are for. They are not offered as predictions of what will happen to any particular investment, and none of them establishes a base rate, since the public record contains no systematic sample of automation outcomes. They are offered as existence proofs: demonstrations that each of the five assumptions has failed in a documented, verifiable instance involving competent organizations with substantial resources and professional advisers.

That is the appropriate standard for a business case exercise, because the purpose is not to estimate the probability of failure but to establish which failure modes are real enough to warrant modelling. An assumption that has never failed anywhere can reasonably be treated as robust. An assumption that has failed publicly, at scale, within the last several years, at organizations that had every reason to get it right, needs a scenario in the model. The cases in this section move each of the five assumptions from the first category into the second, and a committee that has read them will ask better questions than one working from vendor references alone.

08 How sensitive the answer actually is

Having identified the assumptions, the useful next step is to establish how much the conclusion depends on each, which is a routine analysis that automation business cases frequently omit. Figure 6 demonstrates the effect.

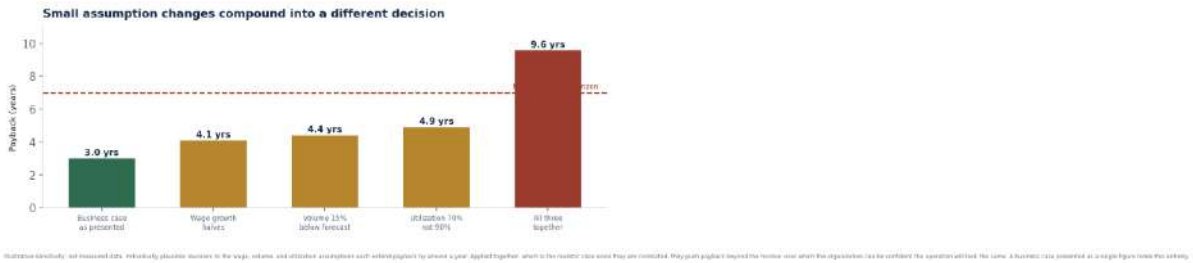

Figure 6. Illustrative sensitivity, not measured data. Individually plausible revisions to wage, volume, and utilization assumptions each extend payback by around a year; applied together, which is the realistic case, they push it beyond the planning horizon.

The illustration takes a business case presenting a three-year payback and applies three revisions, each of which is well within the range of ordinary forecasting error. Halving the assumed rate of wage growth extends the payback by roughly a year. Volume arriving fifteen percent below forecast extends it by a comparable amount. Realized utilization at seventy percent rather than ninety extends it further. Any one of these leaves the investment defensible on most hurdle rates. Applied together, the payback extends beyond the horizon over which the organization can be confident the operation will retain its current shape, at which point the investment is no longer a payback calculation but a strategic bet.

The reason the combined case matters more than the individual ones is the correlation noted earlier. These three revisions are not independent draws; they share a common driver in the demand environment, so the world in which volume disappoints is disproportionately the world in which wage pressure eases. A sensitivity analysis that varies one assumption at a time, which is the standard format, systematically understates the joint downside for exactly this reason. Presenting a small number of coherent scenarios, in which the assumptions move together as they plausibly would, is more informative than a table of one-at-a-time variations even though it looks less rigorous.

There is a straightforward reframing that makes this tractable for a committee, borrowed from investment analysis. Rather than asking what the payback is, ask what would have to be true for the investment to fail, and then ask how likely the committee thinks that is. If the answer is that the investment fails only if volume falls by a third and wages stop rising entirely, the case is robust and the committee can approve it with confidence. If the answer is that it fails if volume comes in modestly below plan, the case is fragile and the committee is being asked to take a position on the demand forecast rather than on the automation. Both may be worth doing; they are different decisions and should be discussed differently.

09 What the payback ratio leaves out

A critique of automation business cases that focused only on optimistic assumptions would be incomplete and would mislead in a predictable direction, because the payback ratio omits material items on both sides of the calculation. Figure 7 sets out the omissions.

Figure 7. The payback ratio is wrong in both directions at once. It understates cost by excluding maintenance, spares, software, integration, and lost flexibility, and understates benefit by excluding peak survivability, safety, and consistency.

On the cost side, the omissions are familiar to anyone who has operated automation and absent from most cases that justify buying it. Maintenance is a continuing cost that rises as the asset ages, and spares for specialised mechanical systems are neither cheap nor always quickly available. The software layer that orchestrates the automation requires licensing, upgrades, and integration with the warehouse management and enterprise systems, and that integration is a project in its own right that frequently costs a meaningful fraction of the mechanical spend. Staff must be retrained, and the skill profile of the operation shifts toward technical roles that command higher wages and are harder to recruit, which partly offsets the labour saving the case is built on.

The largest omitted cost is flexibility, and it is the hardest to quantify. A manual operation can be reconfigured for a new product profile, a new customer requirement, or a new fulfilment model within weeks, at modest cost. A heavily automated operation is configured for the product dimensions, order profile, and process flow it was designed around, and material changes to any of those may require re-engineering that ranges from expensive to impractical. In a stable operation this costs nothing. In an operation whose product mix or channel structure is evolving, it is a real and recurring constraint that will shape commercial decisions for the life of the asset, and it belongs in the case as a stated risk even if it cannot be given a number.

On the benefit side, the omissions are equally substantial and cut the other way. Peak survivability has genuine value: an operation that can handle its peak without a hiring surge avoids recruitment cost, training cost, quality degradation, and the risk of being unable to staff at all, and in a service-sensitive business the avoided failure is worth more than the labour saved. Safety improvements from removing manual handling reduce injury rates and the associated costs, which are large and frequently underestimated. Consistency and accuracy improve, reducing errors, returns, and the service failures that damage customer relationships. And space utilization frequently improves substantially, which either releases property cost or defers a facility expansion, an item that rarely appears in a labour-focused payback and can be worth more than the labour.

The implication is that a buyer who corrects only the cost side of the calculation will systematically reject automation that it should undertake, which is the opposite error to the one this article has spent

most of its length examining and is equally damaging. The disciplined position is to correct both sides: to build the full cost of ownership including maintenance, software, integration, retraining, and flexibility, and to build the full benefit including peak survivability, safety, consistency, and space, and to compare those rigorously. Cases built this way are more work and they are considerably more likely to survive contact with the operation, in both directions.

The flexibility cost deserves a more concrete treatment than a general caution, because it can be estimated even though it cannot be measured precisely. The method is to identify the specific changes the operation considers plausible over the asset life, such as a shift in average order size, the addition of a new product format, a change in packaging dimensions, or a move into a different fulfilment model, and to establish for each what accommodating it would require of the automated configuration. Vendors will answer these questions during selection, and the answers range from routine reconfiguration to substantial re-engineering to outright impossibility.

Once that mapping exists, the flexibility cost can be expressed as a set of conditional statements rather than a single number: if the operation moves to smaller average order sizes, the system requires modification at an estimated cost; if it adds a product format outside the handled dimensional range, that volume must be processed manually. These conditionals are considerably more useful to a committee than an unquantified warning about reduced flexibility, because they identify which strategic options the investment forecloses and at what price. An organization that knows the automation would make a plausible future channel shift expensive has information it can weigh, and it may still proceed, having decided that the shift is unlikely or that the cost is acceptable.

10 The vendor quote, and where it comes from

Buyers encountering this category for the first time will hear payback periods quoted with apparent confidence, frequently in the range of eighteen months to three years, and it is worth understanding how such figures are produced and what they are worth.

A vendor or integrator quoting a typical payback is generalising from its own installed base, which is a legitimate thing to do and carries three specific biases. The first is selection: the installations the vendor knows best and cites most readily are the successful ones, because unsuccessful deployments generate less reference material and less enthusiasm from the customer. The second is that the payback in a vendor's model is usually computed on the mechanical scope the vendor supplied, excluding the integration, software, and process work the customer performed around it, which understates the total investment. The third is that the labour saving is typically computed at rated throughput, for the reasons set out earlier.

None of this makes vendor figures useless, and treating them as fabrications would be both unfair and unhelpful. They are best read as an indication of what the technology has achieved in favourable circumstances, which is genuine information. What they cannot do is substitute for the buyer's own model built on the buyer's own volume, mix, wage structure, and operating profile, and a buyer that adopts a vendor payback figure has adopted a set of assumptions from a party with an interest in the decision. The appropriate use of a vendor quote is as a challenge to the buyer's own model: if the buyer's figure is far worse than the vendor's, the difference should be explicable in terms of specific differences in volume, mix, or wage structure, and if it is not explicable, one of the two models is wrong.

The same caution extends to industry association research and market-house benchmarks, which are widely cited in this category. Industry association surveys are conducted among members who are, by definition, participants in the sector, and market houses sell to both vendors and investors. This does not make their work unreliable, and much of it is careful. It does mean that essentially every published figure about automation adoption, payback, and benefit originates somewhere in the ecosystem that benefits from adoption, and that a buyer looking for a disinterested benchmark will not find one. The most reliable external evidence available to a buyer is the public filings of the suppliers and operators themselves, which is why this article has leaned on them, and the buyer's own measurements once a system is running.

One further practice is worth recommending because it converts the vendor's information advantage into something the buyer can use. Rather than asking a vendor what payback its customers typically achieve, which invites a selected answer, ask for the achieved throughput, availability, and maintenance cost at three named reference sites with a product mix and volume profile comparable to the buyer's own, and ask to speak to the operations manager rather than the executive sponsor. Vendors that have such sites will supply them, and the resulting figures are specific, checkable, and directly usable as inputs to the buyer's own model rather than as a substitute for it.

11 The fairness case: automation earns its capital

This article has spent considerable length on how automation business cases fail, and a reader who concluded that warehouse automation is generally a poor investment would have taken away something this article does not intend and the evidence does not support.

The clearest point in fairness is that automation demonstrably works at scale, and the same public record that supplies the cautionary cases supplies the successes. Highly automated grocery fulfilment operates at densities and accuracy levels that manual picking cannot approach. Large-scale automated distribution has been deployed and expanded by operators who have every ability to measure whether it pays and who continue to invest, which is a meaningful revealed preference. Within the same market data that showed order declines and forecast revisions, particular segments continued to grow substantially, driven by operators expanding programmes they had already tested. Automation that is well matched to a stable, high-volume, predictable operation is frequently a strong investment, and the argument of this article is about underwriting discipline rather than about the technology.

A second point is that labour scarcity and wage pressure are real, and the direction of the labour argument is sound even where its precise extrapolation is not. Operations that depend on being able to recruit large numbers of people for physically demanding work at short notice have a structural fragility that has been demonstrated repeatedly, and the value of removing that dependency is not fully captured by any wage-based calculation. An organization that has repeatedly failed to staff its peak, or that has paid substantial agency premiums to do so, has evidence in its own records that the manual alternative carries costs the wage rate does not show.

A third point is that the failures examined here are predominantly failures of business-case discipline and sizing rather than of technology. The battery manufacturer did not fail because its capital equipment did not work; it failed because the demand did not arrive. The restatement at the automation supplier concerned revenue recognition, not system performance. The fire concerned concentration of capacity in a single site. In each case the equipment did approximately what it was designed to do, and what failed was a forecast, a control environment, or a design decision about concentration. That is an argument for better underwriting, which this article recommends, not for avoiding the category.

A fourth point deserves emphasis because it cuts against this article's own framing. Payback is a poor primary criterion for a strategic capability investment in any domain, and applying it rigorously would cause an organization to reject investments in resilience, safety, and service quality that it should make on other grounds. An operation may reasonably automate because it wants to be able to handle peak without a hiring surge, because it wants to reduce injuries, or because it wants service consistency that its customers value, and each of those is a legitimate basis for a decision that a payback calculation would not support. The recommendation is not to hold automation to a stricter payback test but to be explicit about which basis the decision actually rests on, and to stop presenting strategic capability decisions in the language of a return calculation that was never doing the work.

A fifth point in fairness concerns the counterfactual that critics of automation rarely examine, which is what happens to an operation that does not automate while its competitors do. Where automation delivers structurally lower cost per unit at scale, the operations that forgo it may find their cost position deteriorating relative to the market over the same decade the payback would have covered, and the loss from that erosion does not appear in any business case because it is the cost of the decision not taken. An organization comparing automation against its current state is comparing against a baseline that will not hold if the sector moves.

This cuts against the caution recommended elsewhere in this article and it belongs in the analysis for that reason. The disciplined response is not to assume competitive erosion, which would restore the optimism the article has been arguing against, but to model it explicitly where there is evidence for it: whether competitors are in fact automating, whether the resulting cost differential is material at the buyer's scale, and how quickly it would show up in price competition. Where that evidence is strong, the case for acting is stronger than the payback alone suggests. Where it rests on the assumption that everyone else must be doing something, it is worth nothing, and the distinction between the two is exactly the sort of thing a properly built case makes visible.

12 Automate a process you have already fixed

The single most consequential piece of practical guidance in this article is one that predates automation by decades and is routinely ignored: mechanising a process does not improve it, so the process should be fixed before it is mechanised rather than afterward.

The failure mode is easy to describe and common to observe. An operation with a workflow that has accumulated exceptions, workarounds, and redundant steps decides to automate it, and the automation is specified to replicate the existing workflow because that is what the operation does and because redesigning it would delay the project. The result is a mechanised version of an inefficient process, and the inefficiency is now embedded in capital equipment where it is far more expensive to remove than it was when it existed only in procedure. The organization has spent substantial money to perform its existing process faster, when a proportion of the process should not have been performed at all.

The discipline that avoids this is sequential and unglamorous. Map the current process and measure it. Identify the steps that exist because of historical decisions, system limitations, or accumulated exceptions rather than because the work requires them, and remove them. Redesign what remains for the outcome the operation needs rather than for the way it has been done. Run the redesigned process manually for long enough to establish that it works and to measure its actual performance. Only then specify automation against the redesigned process. This sequence takes longer and it produces both a better automation specification and, frequently, a discovery that a portion of the anticipated benefit is available without capital expenditure at all.

That last point is worth stating directly because it changes the decision. Process improvement, slotting optimisation, better scheduling, and workflow redesign frequently deliver a material share of the benefit that the automation business case attributes to the automation, at a fraction of the cost and with none of the irreversibility. An organization that has not exhausted these before committing capital does not know how much of the projected benefit requires the capital, and it may be paying for improvements it could have obtained by other means. Establishing the improved manual baseline first is therefore not merely good project sequencing; it is the only way to know what the automation is actually buying, because the correct comparison for the automated case is the best achievable manual operation rather than the current one.

There is an organizational reason the sequence gets inverted that deserves acknowledgement, because it explains why the advice is so widely known and so rarely followed. Process redesign has no vendor, no proposal, and no natural sponsor, whereas automation arrives with a supplier who will help build the business case, a project structure, and a visible outcome that a career benefits from. Improving a process quietly over six months produces a smaller and less legible achievement than commissioning a system, even where the former delivers more value per unit of cost and risk. The incentive gradient runs toward capital, and the organizations that resist it do so deliberately.

The corrective is structural rather than exhortatory. An organization that requires every automation proposal to document what was attempted without capital, what it achieved, and why the residual gap

requires equipment will find that a proportion of proposals do not survive the requirement, and that the ones that do are considerably stronger. This is the same discipline that public procurement applies when it requires justification for a sole-source award: not a prohibition, but an obligation to state the reasoning where the default would otherwise apply unexamined. It costs little, it is easy to administer, and it redirects attention to the cheapest available improvements before the expensive ones are considered.

13 An underwriting protocol, and a scoring rubric

The principles above combine into a protocol for underwriting an automation investment, and a rubric a committee can apply to judge whether a case in front of it has been built properly. Figure 8 sets out the discipline.

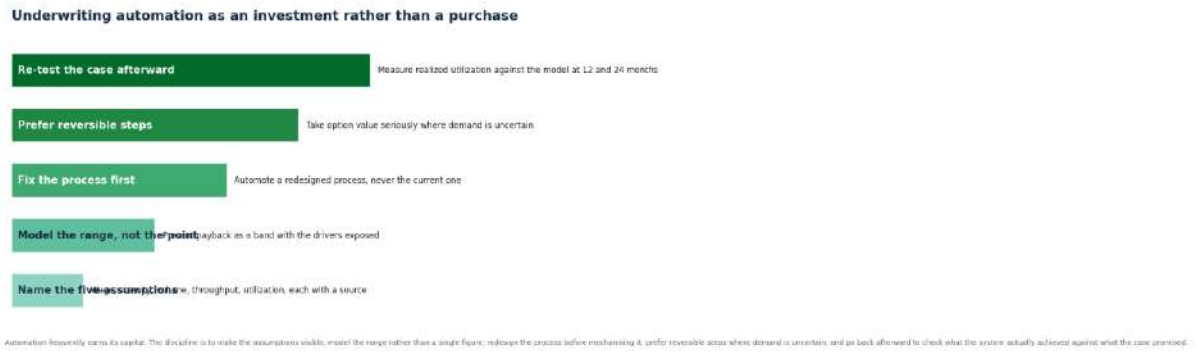

Figure 8. Underwriting automation as an investment rather than a purchase: name the assumptions, model the range, fix the process first, prefer reversible steps, and re-test the case against realized performance.

The protocol runs as follows. State the five assumptions explicitly, each with its source and its basis, using sector-specific and locally-relevant labour data rather than national averages. Present the payback as a range with the sensitivity exposed, including at least one coherent downside scenario in which the correlated assumptions move together. Build the full cost of ownership including maintenance, spares, software, integration, and retraining, and the full benefit including peak survivability, safety, consistency, and space. Redesign and stabilise the process manually before specifying automation against it. Prefer reversible steps where demand is uncertain. And commit in advance to re-testing the case against realized performance at defined intervals after commissioning.

A scoring rubric

The dimensions below distinguish an underwritten investment from a purchased proposal.

Dimension	Underwritten	Purchased
Assumptions	Named, sourced, individually challengeable	Embedded in a single payback figure
Labour data	Sector and local, fully loaded, with turnover	National average hourly rate
Throughput basis	Expected realized volume and mix	Rated capacity applied year-round
Downside	Correlated scenario, stated failure condition	One-at-a-time variation, or none
Cost scope	Maintenance, software, integration, retraining	Mechanical scope as quoted
Benefit scope	Peak, safety, consistency, space, as well as labour	Labour hours removed

Dimension	Underwritten	Purchased
Process	Redesigned and stabilised before specification	Existing workflow mechanised as-is
Follow-up	Re-tested against realized performance	Never revisited after approval

A case scoring in the left column tells a committee what it is taking a position on and what would have to happen for the position to fail. A case scoring in the right column tells the committee a number. The rubric does not make automation less attractive; applied properly it strengthens good cases, because the omitted benefits are frequently larger than the omitted costs. It ensures the decision is made on the actual basis rather than on a figure whose fragility nobody examined.

14 Reversibility, and the value of the smaller step

A dimension that scarcely appears in automation business cases, and that deserves to be weighed alongside the return, is reversibility: how easily the organization could undo the decision if the assumptions prove wrong. Figure 9 positions the options.

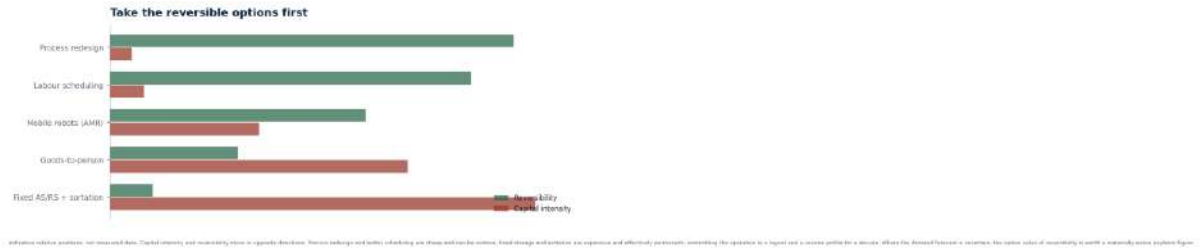

Figure 9. Indicative relative positions, not measured data. Capital intensity and reversibility move in opposite directions. Where demand is uncertain, the option value of reversibility is worth a materially worse payback figure.

The relationship is close to monotonic. Process redesign is cheap and entirely reversible. Improved scheduling and slotting are cheap and reversible. Mobile robots are moderately expensive, redeployable within a facility and to some extent between facilities, and can be scaled up or down in units rather than in a single commitment. Goods-to-person systems are expensive and substantially fixed. Fixed storage and sortation infrastructure is the most expensive and effectively permanent, committing the operation to a layout, a throughput envelope, and a product profile for the life of the building.

Standard capital appraisal handles this badly, because a payback calculation values only the expected return and is blind to the option structure. Where the future is uncertain, the ability to change course has value in itself, and a reversible option with a modestly worse expected return may be worth more than an irreversible one with a better expected return, precisely because it preserves the ability to respond to information that has not arrived yet. This is standard reasoning in investment analysis and it is rarely applied in warehouse capital decisions, where the comparison is usually conducted on payback alone.

The practical implication is a staged approach that many operations would benefit from and that vendors have limited incentive to propose. Take the reversible steps first, in order of increasing commitment: fix the process, improve the scheduling and slotting, deploy modular equipment that can be redeployed, and only then consider fixed infrastructure, by which point the operation has real data about its own volumes, its own achievable throughput, and the actual benefit of each step. Each stage produces measurement that improves the next decision, and the organization retains the ability to stop if the assumptions are not holding.

The objection to staging is that it may be less efficient in aggregate than a single integrated design, and this is sometimes true: a purpose-built integrated system can achieve densities and flows that an incrementally assembled one cannot. That is a real trade and it should be evaluated rather than assumed. The relevant question is how confident the operation is in its ten-year volume and mix projections. Where that confidence is high, integrated design is defensible and the efficiency gain is worth the irreversibility. Where it is low, which describes most operations candidly assessed, staging is

worth its inefficiency because it preserves the ability to be wrong without having committed the full capital to the error.

15 Conclusion: go back and check

Warehouse automation is one of the largest capital commitments a supply chain organization makes, it shapes the operation for a decade, and it is routinely approved on the basis of a single number that compresses five multi-year forecasts into a figure with a decimal point. That is the argument of this article stated as compactly as it can be. The five forecasts concern labour cost, labour scarcity, volume, throughput, and utilization; each is uncertain, they are correlated rather than independent, and the format in which the case is presented conceals all of that from the committee that has to decide.

The evidence assembled here comes from filings and dockets rather than from case studies, because those sources are not selected for their outcomes. They show a major supplier restating a full year of results and disclosing that a single customer accounted for most of its revenue; a capital-intensive manufacturer entering bankruptcy with billions in debt because the volume did not arrive; a highly automated operator losing hundreds of thousands of orders from damage to a fraction of one site; a robotics supplier taken private at a small fraction of its listing valuation; and a large retailer ending a robot programme because people were achieving comparable results. None of these is an indictment of automation. Each is a record of an assumption that did not hold, which is exactly what a business case needs to consider and rarely does.

What a disciplined organization does follows from that. It names the five assumptions and sources them properly, using sector and local labour data rather than national averages, and including the turnover cost that makes the manual alternative more expensive than its wage rate suggests. It presents a range rather than a point, with a coherent downside in which the correlated assumptions move together, and it states plainly what would have to be true for the investment to fail. It builds the full cost including maintenance, software, integration, retraining, and lost flexibility, and the full benefit including peak survivability, safety, consistency, and space, because correcting only one side produces a systematic error in a known direction. It fixes and stabilises the process before mechanising it, which frequently reveals that part of the projected benefit needs no capital at all. And where the demand outlook is uncertain, it prefers the reversible step and accepts a worse headline figure for the option to change course.

The final discipline is the one most consistently absent and the cheapest to adopt. Almost no organization goes back, twelve or twenty-four months after commissioning, to compare what the system actually achieved against what the business case promised: realized throughput against rated, realized utilization against design, actual labour reduction against modelled, actual maintenance cost against budget. Without that comparison the organization learns nothing transferable from a decision worth tens of millions, and the next business case is built with the same optimism as the last because nothing contradicted it. An organization that institutes this single review will improve its underwriting faster than any amount of analytical sophistication applied before approval, because it will finally have evidence about its own forecasting rather than about the vendor's. The number in the business case was never the point. Whether the operation is better off, measured afterward against what was promised, is the only test that matters, and it is the one almost nobody runs.

16 Methodology, caveats, and sources

Methodology

- This article draws on securities filings, bankruptcy dockets, official labour statistics, company disclosures, and contemporaneous reporting, current to mid-2026. Supply Chain Research is independent and accepts no payment from the automation suppliers, integrators, or market-research firms discussed.
- Evidence is drawn from public filings rather than vendor case studies wherever possible, because case studies are selected for their outcomes and filings are not. Figures originating with market-research firms or industry associations are identified as interested sources.

Caveats

- Figures 1, 3, 6, 7, and 9 are conceptual illustrations of structure or arithmetic rather than measured data, and are labelled as such. Figure 6 in particular uses invented inputs to demonstrate sensitivity behaviour.
- Figure 2 shows the indicative shape of a labour-cost series rather than exact published values. Any business case should use the current official series for the relevant sector and geography rather than the illustration here.
- Market-size and forecast-revision figures come from a market-intelligence firm that sells research to participants in the sector it covers. They are interested estimates; the direction of the revision, rather than the precise magnitude, is the reliable signal.
- The company cases are drawn from public disclosures and reporting and are used to illustrate assumption failure, not to make any claim about the quality of any supplier's equipment or the conduct of any management team.
- Vendor-quoted payback ranges are described in general terms because they vary widely by system type, scope, and operation, and because published examples are selected by the parties quoting them.
- Accounting matters described, including a restatement and disclosed material weaknesses, are summarised from company announcements and filings and should be read against the full original disclosures rather than this summary.

Sources

[1] Symbotic Inc. Annual report on Form 10-K, fiscal year 2024, including customer concentration.

[2] Symbotic Inc. Company update on the restatement of fiscal year 2024 results.

[3] Northvolt AB. Chapter 11 case docket and filings.

[4] Interact Analysis. Warehouse automation project orders declined in 2024 (interested source: market-research firm).

[5] Interact Analysis. Mobile robot market forecast reduced (interested source: market-research firm).

[6] InternetRetailing. Estimated revenue impact of the 2021 automated fulfilment centre fire.

[7] US Bureau of Labor Statistics. Warehousing and storage industry data (NAICS 493).

[8] Federal Reserve Bank of St Louis. Average hourly earnings, transportation and warehousing (series CES4300000003).

Additional context drawn from company announcements concerning the 2023 take-private of a robotics supplier, from contemporaneous reporting on the 2020 discontinuation of a retail shelf-scanning robot programme, and from published labour turnover statistics. Figures originating with market-research firms or industry associations are identified as interested sources. This article is analysis, not investment, engineering, or operational advice, and its conclusions should be validated against your own circumstances before any decision.