

The Future of Commerce Using Physical AI

Embodied AI arrived in retail and a large share of it has already been taken out again. This review grades what stayed, what was withdrawn, and what both imply for the half-million people a year who enter the frontline commerce workforce.

Kat Moore, EdD, MBA, Industrial Research Fellow · The Glass Lab

18 graded deployment & withdrawal records

56 numbered sources · grade on every table entry

3 BLS series pulled from the public API

Market & economic review · not investment advice

Physical AI reached the American selling floor over a decade, and a substantial part of it has since been removed. Amazon closed its last 15 Amazon Go stores in February 2026, having already removed Just Walk Out from every US Amazon Fresh grocery store in April 2024, while licensing the same checkout technology into more than 360 third-party venues. Walmart cancelled its shelf-scanning robots in 2020 and is now installing electronic shelf labels in all 4,597 US stores. Kroger took a \$2.585 billion charge on automated fulfilment ten months after Walmart committed, conditionally, to 400 automated store backrooms. This review assembles eighteen deployment and withdrawal records into one graded ledger and reads the pattern by how much change the customer has to absorb. The second half turns to the workforce, where the aggregate is not collapse: US retail employment in June 2026 was 15,459,700, near its February 2020 level, while retail's share of national employment fell from 10.19% to 9.72%. The Bureau of Labor Statistics models 313,600 fewer cashier jobs by 2034 alongside 542,600 cashier openings a year. The training problem in commerce is throughput through

a changing entry-level role, and the report specifies a vendor-independent curriculum, a measurement standard for labour-hours claims, and the workforce cost of automation that gets withdrawn.

1. Headline numbers

Physical AI arrived on the American selling floor over roughly a decade, and a substantial share of it has since been taken out again. The withdrawals are better documented than the installations, because a closure is a filing and an installation is a press release. That asymmetry shapes this report: the strongest evidence in the selling-floor half concerns systems that were removed, and the strongest evidence in the workforce half comes from government series and randomised experiments rather than from vendors.

The table below is the short version. Each figure is graded at first substantive use, and the grade travels with the number through the tables. Where a number is a company's own claim, the sentence containing it says so.

Key number	Value (as of August 2026)	Grade	Ref
Amazon-operated checkout-free stores in the US	0 retail: 15 Go closed Feb 2026; Just Walk Out had already been removed from all US Fresh stores in April 2024, so the 57 Fresh closures were not checkout-free at close. 40+ Amazon-operated breakrooms still running	reported	1–4
Just Walk Out third-party locations	360+ across five countries; ~150 added during 2025	self-published	1, 5, 6
Largest disclosed retail automation impairment identified in this review	\$2.585bn impairment and related charges (Kroger, FQ3 2025), including an accrued cash termination payment of about \$350m to Ocado	reported	19–22
Walmart electronic shelf label coverage	2,300 stores mid-2026; all 4,597 US stores targeted by end-2026	reported	28, 29

Self-checkout partially withdrawn	12,000+ of Dollar General's 20,000+ stores changed; ~300 removed entirely	reported	33, 34
US retail trade employment, June 2026	15,459,700, below February 2020's 15,519,200	verified	42
Retail share of US nonfarm employment	10.19% (Feb 2020) to 9.72% (Jun 2026)	verified levels, author's arithmetic	42
Projected cashier employment change, 2024–34	–313,600 (–9.9%), the largest projected numeric decline of any US occupation in the 2024–34 projections (BLS Employment Projections table 1.5)	modelled	44
Projected annual cashier openings, 2024–34	542,600 per year, about 17x the average annual net decline	modelled	44
Effect of an AI assistant on frontline support throughput	+14% issues resolved per hour; +34% for novices, negligible for experts	verified	51
Largest randomised frontline AI experiment	5,940 agents; gains in performance quintiles 1–3, quality decline in quintile 5	reported	52
Ceiling on aggregate AI productivity, ten years	a cumulative increase in total factor productivity of no more than 0.66%	modelled	48

Table 1. Key figures, with verification grade. Citation numbers refer to the reference list.

2. Scope, definitions, method and the grading scheme

This report is an Industrial Research Fellowship review by Kat Moore, EdD, MBA, hosted by the Glass Lab, the Institute's airspace integration, aviation training and STEM workforce group. The fellow's background is in sales and in education, and the report is organised around that pairing: what embodied AI does to the selling

floor, and what it does to the people standing on it. It is a market and economic review. It is not investment advice, and no security is discussed.

Physical AI here means artificial intelligence embodied in equipment that senses and acts in a store or a fulfilment building: ceiling and shelf cameras that infer what a shopper picked up, mobile robots that photograph shelves, robotic arms in a backroom, electronic labels that change a price on a shelf edge, and the software assistants that frontline staff now carry in a handheld. Online recommendation systems and demand forecasting sit outside the scope except where an experiment on them bears directly on what a human associate does.

Four grades are used, and they are the Institute's standard. Verified means independently replicated or peer-reviewed, or pulled directly from a primary government series. Reported means a named analyst, operator, agency or news organisation stated it without independent confirmation. Self-published means the number originates with the company selling the thing. Modelled means it was computed rather than measured, including every projection. A modelled figure is never written as a measurement in this report, and where the author has performed arithmetic on other people's figures, the arithmetic is labelled as the author's.

Two limits on the evidence base should be stated before the argument begins. First, no retailer or vendor has published a matched-store controlled comparison of any shelf-scanning robot, checkout-free installation, or computer-vision loss prevention system. The category's return-on-investment case rests on supplier arithmetic, and this report treats every vendor effect claim accordingly. Second, the industry's own shrink benchmark was discontinued: the National Retail Federation ended the National Retail Security Survey after 32 editions, the last covering fiscal 2022, so loss-prevention business cases are now written against a baseline nobody measures consistently ^{36,37}.

3. Part I. What is actually installed: a deployment ledger

The selling-floor half of this report starts with a ledger rather than a thesis. Eighteen programmes are listed below, each with the scale it reached and its status in August 2026. The ledger is not a random sample. It over-represents programmes large enough that their ending was newsworthy, and it under-represents quiet installations that are still running, so it cannot be read as a survival rate for the category.

Counting the ledger anyway is instructive, with that caveat attached. Of the eighteen records, ten were closed, cancelled, converted or materially reduced, one entered bankruptcy, one is a commitment rather than an installed base, and six are operating. The classification rule is the status cell in Table 2 as written: a programme partially withdrawn, de-installed from part of its estate, or reported shut down is counted in the first group, not among the operating. That count is the author's, over a non-random ledger, and is graded modelled. It is offered for one purpose: any plan to put embodied AI on a selling floor should carry an explicit assumption about how long the system will stay, and the assumption is not decades.

The pattern that runs through the ledger is easier to see once the systems are sorted by what they ask of the customer. Checkout-free shopping did not hold in the weekly-shop grocery format under any operator in this ledger, while the same technology is operating and expanding in venue concessions where the basket is small and the alternative is a long queue. Systems that changed only what staff do behind the scenes are, without exception here, still running or expanding.

System	Operator	Scale reached	Status	Grade	Ref
Amazon Go	Amazon	15 US stores at close	Closed, early Feb 2026	reported; severance terms self-published	1–3
Amazon Fresh (US)	Amazon	57 US stores at close	Closed; investment redirected to Whole Foods and delivery	reported	1, 4
Amazon Fresh (UK)	Amazon	19 stores	14 closed, 5 converted to Whole Foods, Sep 2025; ~250 jobs at risk	reported	8, 9
Just Walk Out, licensed	Amazon and venue operators	360+ locations,	Operating and expanding,	self-published	1, 5, 6, 7

			five countries	concentrated in stadiums and airports		
	Just Walk Out, breakrooms	Amazon	40+ North American fulfilment centres	Operating	self-published	1
	Shop&Go	Aldi UK	Greenwich trial store	Trial ended	reported	11
	GetGo	Tesco UK	Multiple convenience sites	Converted to hybrid: staffed and self-service tills offered	reported	11
	Just Walk Out at Holborn	Sainsbury's	One store	Removed, Feb 2025	reported	10
	Bossa Nova shelf robot	Walmart US	~500 stores piloted; 1,000 announced	Contract ended Nov 2020; vendor cut ~half its staff and exited retail	reported	12, 13
	Tally shelf robot	Simbe, with ~60 chains in 10 countries	4.7m autonomous hours, 1.8m km claimed	Operating; all scale figures are the vendor's	self-published	14
	Tally Spot fixed cameras	Simbe	Launched Jan 2025	Operating; installed base not disclosed	self-published	17
	Digital Teammate robot	Badger Technologies (Jabil)	1,000+ robots claimed	Operating; vendor figure	self-published	18
	Ocado automated	Kroger	3 CFCs closed, 1 cancelled, 3 spokes	Withdrawn; \$2.6bn charge,	reported	19-22

fulfilment centres		closed earlier	\$350m to Ocado		
Grocery micro-fulfilment	Takeoff Technologies	Multiple grocer sites	Chapter 11, May 2024; sold via 363 process to Woolworths	verified filing	24, 25
Symbotic accelerated pickup and delivery backrooms	Walmart	400 units committed over several years	Commitment, not an installed base	self-published	26, 27
Electronic shelf labels	Walmart with VusionGroup	2,300 stores mid-2026; 4,597 targeted	In rollout, the only chain-wide saturation in the ledger	self-published (company target)	28, 29
Self-checkout	Dollar General	12,000+ of 20,000+ stores changed	Partially withdrawn on shrink grounds; ~300 stores fully de-installed	reported	33, 34
Blue Jay multi-arm station	Amazon	Launched Oct 2025	Reportedly shut down within ~6 months; single-outlet, unconfirmed	reported, weak	39, 40

Table 2. Deployment ledger: physical AI on the US and UK selling floor, 2016-2026. Status as of August 2026.

4. Checkout-free stores: closures and third-party licensing, 2024-2026

On 27 January 2026 Amazon said it would close every Amazon Fresh and Amazon Go store in the United States, 72 locations in total, with most shutting that first weekend of February. Amazon's statement frames the decision as a refocus on

Whole Foods, where it plans more than 100 new stores, and on delivery; staff were offered redeployment or severance of 90 days at full pay and benefits ¹. Store counts of 57 Fresh and 15 Go come from trade and general press on 27 and 28 January ^{2,3,4}, and the grade is reported. Amazon has never published per-store economics for the format and never disclosed how many employees were affected.

What that closure does and does not establish needs care. It is not evidence that sensor-fusion checkout does not work, and Amazon says in the same statement that it continues to expand the technology. The confound is that Amazon Fresh had assortment, siting and price-position problems that have nothing to do with cameras, and the grocery banner was weak overall. The defensible reading is narrower: the most-photographed embodied-AI retail concept of the last decade did not scale as a store business under its own originator, which is a hard fact for anyone arguing that checkout-free technology transforms a store.

Amazon reports Just Walk Out running in more than 360 third-party locations across five countries, roughly 150 of them added during 2025, concentrated in stadiums and airports, plus more than 40 North American fulfilment-centre breakrooms ^{1,5,6}. Nissan Stadium concessions are one documented example ⁷. All of these counts originate with Amazon and have not been independently audited, so the grade is self-published. Location is also not equivalent to store, since a concession kiosk counts the same as a shop.

British operators produced the cleanest convergence in the record. Amazon announced in September 2025 that it would close all 19 UK Amazon Fresh stores, 14 closing and five converting, with around 250 jobs at risk ^{8,9}. Sainsbury's removed Just Walk Out from its Holborn store in February 2025 ¹⁰. Aldi ended the Greenwich Shop&Go trial, and Tesco moved GetGo to a hybrid format that offers both staffed and self-service tills ¹¹. Four independent operators reached the same design answer inside eighteen months, which is a stronger signal than any single retreat, although UK grocery is unusually price-competitive and none of the four published the transaction economics behind the decision.

5. Shelf-scanning robots: deployments, withdrawals and business-model change

Walmart ended its Bossa Nova shelf-scanning contract in November 2020 after piloting in about 500 stores and announcing an expansion to 1,000; Bossa Nova then cut roughly half its staff and wound down retail robotics ^{12,13}. Walmart never

gave an official reason. Reporting at the time attributed the decision to three separate things: online order picking already had associates in every aisle producing the same shelf data as a by-product, simpler human methods worked and cost less, and Walmart US chief executive John Furner was concerned about how shoppers reacted to the machines. The robot lost to a workflow change rather than to a better robot, which is a different lesson from the one usually drawn.

The category did not die. Simbe, on its tenth anniversary, claims deployment with nearly five dozen retail chains across ten countries, 4.7 million autonomous operating hours, 1.8 million kilometres travelled, 44.8 billion shelf photographs and 18 billion price tags scanned ¹⁴. Every one of those is a marketing figure with no third-party audit and no denominator, and nearly five dozen retail leaders includes trials. Badger Technologies, a Jabil division, reports more than 1,000 robots deployed and markets them as a Digital Teammate ¹⁸. Both grades are self-published and the numbers should be read as scale indicators only.

Survivors in this category changed the sale rather than the hardware. Simbe sells a data subscription in which the robot is the sensor platform, and Badger's naming does the same work in the other direction, positioning the machine beside the associate rather than in place of them. Simbe's published client outcomes are 98% on-shelf availability, 90% improved price and promotion accuracy, more than 50 hours a week returned to store teams, 42% of detected shelf gaps fixed instantly, and in an earlier statement a 60% drop in out-of-stocks with online fulfilment time halved ^{14,15,16}. No control group, no baseline, no named retailer, no disclosed method.

One of those claims deserves separate attention because it is the hinge of the workforce argument later in this report. More than 50 hours per week returned to store teams is a labour-hours claim that does not say what happened to the hours. Whether they came off the payroll or were redeployed into service, replenishment or fulfilment is the single most important fact for a worker, and it is unstated. Simbe's January 2025 launch of Tally Spot, fixed ceiling and shelf cameras watching high-turn and high-shrink zones continuously, suggests where the economics are heading ¹⁷: a camera above the produce table observes the same shelf a thousand times a day, while a mobile robot pays a mobility cost for every observation it takes. Anyone planning a deployment should price observations rather than robots, while noting that Simbe sells both and its complement-not-replace framing is not neutral.

6. Electronic shelf labels: rollout scale and the dynamic-pricing evidence

Walmart expects electronic shelf labels in all 4,597 US stores by the end of 2026, up from 2,300 stores in mid-2026, against an original 2024 plan that targeted only the 2,300; the supplier is VusionGroup^{28,29}. Walmart says a price change that used to take two days now takes minutes, which is the company's own figure and describes task duration rather than labour hours or headcount. The 4,597 number is a target for end-2026, not a completed rollout, and Walmart has published no shrink or labour-hours effect.

This is the only embodied-AI-adjacent store system in the ledger approaching chain-wide saturation in the United States, and the reason is visible in what it touches. Changing a paper price tag is repetitive, has no customer-facing judgement in it, and its removal is not something a shopper experiences. Compare that with checkout-free exit or a robot in an aisle, where the labour case is contested and the customer is party to the change. The saturation pattern in Table 2 sorts along that line rather than along technical difficulty.

Electronic labels also generated the sharpest public-interest controversy in the category. In August 2024 Senators Elizabeth Warren and Bob Casey wrote to Kroger warning that the labels would enable surge pricing and price-setting by consumer profile; Kroger replied that it does not and has never engaged in surge pricing³¹. That exchange framed the technology for a general audience and it remains the reference point in most coverage.

The first large-scale evidence supports the retailer. Robert Sanders of UC San Diego Rady, Ioannis Stamatopoulos of UT Austin McCombs and Robert Bray of Northwestern Kellogg studied 114 US grocery stores across four states with more than 180 million product-level observations from October 2022 to mid-2024 and found no evidence that electronic labels enabled surge pricing³⁰. Unexplained short-term price spikes affected 0.0050% of products before installation and 0.0006% after, a decrease. Price changes stayed chain-wide rather than store-local and tracked wholesale costs, competitor prices and manufacturer promotions. The grade is reported, since it is an SSRN working paper that is not yet peer-reviewed, covering one unnamed chain in a window that closes before most of the current rollout, and it measures what retailers did during a specific inflationary period rather than what the hardware makes possible. The senators' separate concern

about camera-based demographic inference at shelf displays is untouched by this data.

7. Automated fulfilment: Kroger's withdrawal and Walmart's expansion

Kroger closed three Ocado-powered automated customer fulfilment centres at Pleasant Prairie, Frederick and Groveland, cancelled a planned Charlotte facility, took \$2.585 billion of impairment and related charges in fiscal Q3 2025, a figure that includes an accrued cash termination payment of about \$350 million to Ocado ^{19,20,21,22,23}. It expects the retreat to improve e-commerce profitability by about \$400 million in 2026 by shifting to in-store fulfilment. Interim chief executive Ron Sargent ordered a site-by-site review in September 2025 before deciding, and Kroger had already paused Ocado development in September 2023 and closed three spoke facilities in March 2025 that did not meet its benchmarks.

The impairment is an accounting entry rather than cash lost, and it bundles other e-commerce changes, so the headline figure should not be read as money that left the building. Kroger's own explanation is order density and suburban delivery distance, which is a problem of American geography rather than a verdict on Ocado's engineering; the same technology continues to run in denser markets. What the episode does establish is that centralised robotic fulfilment lost to the store as a fulfilment node in this geography, at this order density.

Kroger's experience was not idiosyncratic. Takeoff Technologies, the leading grocery micro-fulfilment specialist, filed Chapter 11 in Delaware on 30 May 2024 after operating losses in 2022 and 2023, owing roughly \$12.9 million in trade and unsecured claims, and was sold through a customer-funded 363 process to Woolworths ^{24,25}. The filing states liabilities rather than causes, and the diagnosis that volume per site never reached the level the capital required is trade-press analysis. The pure-play vendor serving mid-size grocers failed before the largest deployment did.

Walmart moved the opposite way. Company guidance puts about 65% of stores serviced by automation and about 55% of fulfilment-centre volume through automated facilities by the end of fiscal 2026, with an expected 20% improvement in average unit cost ²⁶. Symbotic completed its acquisition of Walmart's Advanced Systems and Robotics business on 28 January 2025 for \$200 million in cash and up to \$350 million in contingent consideration, and signed a commercial agreement

under which Walmart pays \$520 million for the development programme and, if performance criteria are achieved, is committed to purchasing and deploying systems for 400 accelerated pickup and delivery centres over a multi-year period ²⁷. All of those figures are company guidance rather than audited outcomes, serviced by automation is Walmart's own definition, and secondary reports claiming the 400 units were already deployed by March 2026 conflict with Symbotic's multi-year language. The two largest US grocers reached opposite conclusions about the same technology within about ten months, Walmart committing in January 2025 and Kroger withdrawing in its quarter ending November 2025, and the variable that separates them is fixed cost per order at achievable local volume, where Walmart's store density gives it a denominator Kroger did not have.

8. Self-checkout and the transfer of cost

Self-checkout is the largest installed base in this ledger by a wide margin, and it is the system now being partially unwound. Dollar General pulled self-checkout from more than 12,000 of its 20,000-plus stores during 2024 and removed it entirely from roughly 300 of its highest-theft locations; Target limited self-checkout to ten items or fewer in March 2024, and Five Below announced in 2024 that it was removing it from its highest-risk stores ^{33,34}. Whether all three policies still stand in 2026 is not established here. A LendingTree survey of more than 2,000 US consumers in December 2025 found 27% of self-checkout users admitting they deliberately did not scan at least one item ³⁵. A 2023 survey by the same publisher reported 15%, but the two are separate online panels and the wording and screening are not established as identical, so they are not treated here as a trend. That figure is self-reported admission in an online panel and is not a measurement of loss; self-report on theft is unreliable in both directions.

The retailers' own causal attribution is theft, and they have not published the shrink arithmetic behind the decisions. What can be said with more confidence is that the system moved a cost rather than removing one, from payroll into shrink and into the customer's own time at the terminal. A labour-hours saving became a profit-and-loss saving only if the two transferred costs stayed smaller than the payroll it replaced, and several large operators concluded they did not.

Verifying any of that is now harder than it was. The National Retail Federation discontinued the National Retail Security Survey after 32 editions, the last covering fiscal 2022 ³⁶. That final report attributed 36% of inventory loss to external theft,

29% to employee theft, 27% to process and control failures and errors, 6% to unknown causes and 1% to other ³⁶. NRF had previously retracted a headline organised-retail-crime statistic after a Retail Dive analysis found an error ³⁷. The widely circulated shrink figure of 1.6% and around \$112 billion trace to fiscal 2022 data and should not be presented as current.

Two things follow for anyone building a loss-prevention business case. The baseline is unmeasured, and process and control failures accounted for 27% of shrink in the last measurement, more than a quarter of the loss, which a camera at the exit is not positioned to see. Everseen, the leading computer-vision loss-prevention vendor, reports analysing 140,000 self-checkout lanes daily and commissioned a Forrester Total Economic Impact study in September 2024, but the underlying per-retailer shrink deltas are not public ³⁸. No retailer or vendor has published a controlled measurement of shrink reduction from computer vision at checkout.

9. What the selling-floor record supports

Seven conditions can be drawn from Table 2 with varying confidence, and they are set out in Table 3 with the evidence attached to each. They are engineering and siting guidance rather than laws, and each one rests on a small number of operators in two national markets.

A fifth observation is about measurement rather than siting. Every effect claim in the selling-floor literature is a vendor's, and the industry benchmark that would let a retailer check one has been discontinued. Before an operator buys any of this, the useful discipline is to define the measurement first: the store-weeks, the matched comparison stores, the baseline, and the disposition of any hours the system frees. A supplier unwilling to be measured that way is selling a forecast.

Condition	What the record supports	Evidence
Hard throughput spike, dwell time worthless to the operator	Checkout-free sensing works commercially: stadiums, airports, breakrooms	360+ licensed locations, self-published ^{1,5,6,7}
Weekly shop, dwell time valuable, price competition intense	Checkout-free did not hold: US and UK store formats closed or converted	72 US closures; four UK operators retreated [1-4, 8-11]

Task is repetitive with no customer-facing judgement	Automation saturates chain-wide	ESLs in 2,300 stores heading to 4,597 ^{28,29}
Task requires the customer to change behaviour at the exit	Optionality wins; enforced single paths are removed	Tesco GetGo hybrid; Dollar General de-installs ^{11,33,34}
Observation is needed continuously in a fixed zone	Fixed cameras beat mobility on cost per observation	Tally Spot launch, vendor framing ¹⁷
Data the robot collects is a by-product of existing work	The robot loses to the workflow change	Walmart-Bossa Nova, pickers already in aisle ^{12,13}
Order density below the level the capital requires	Centralised automation fails; the store wins as the node	Kroger \$2.6bn; Takeoff Chapter 11 [19-25]

Table 3. Siting rules supported by the 2016-2026 deployment record.

10. Part II. The aggregate labour series, before any argument

US retail trade employment was 15,459,700 in June 2026, seasonally adjusted, which is below the February 2020 level of 15,519,200 and below January 2019's 15,617,400. Over the same period total nonfarm employment rose from 152,293,000 to 158,984,000. These come directly from the Bureau of Labor Statistics public API, series CES4200000001 and CES0000000001, and are graded verified ⁴². Retail's share of national employment therefore fell from 10.19% to 9.72%, which is the author's arithmetic on two verified levels rather than a BLS-published series.

The headline is flat rather than falling. Retail is shrinking as a share of American work while the rest of the economy grows, without shedding a visible mass of jobs. Any argument that automation has already emptied the selling floor is not supported by the aggregate series, and this report says so before it says anything else.

Two measurement caveats matter more than usual here. The CES series counts jobs at retail-trade establishments under NAICS 44-45 rather than retail occupations, so

a cashier at a hospital gift shop is classified elsewhere. E-commerce fulfilment work has migrated into transportation and warehousing under NAICS 49, which makes retail look worse than the work actually is. Anyone using the retail series as a proxy for selling-floor labour is measuring the wrong denominator by a growing margin.

Pay moved in the direction that complicates the substitution story. Average weekly earnings of all retail trade employees rose from \$593.12 in January 2019 to \$787.87 in June 2026, a 32.8% nominal increase, again pulled directly from the BLS API, series CES4200000011, and graded verified⁴³. CPI-U (BLS series CUUR0000SA0) rose about 27% over the same span, so the real gain is on the order of five points rather than thirty-three, and that deflated figure is the author's arithmetic. Weekly earnings also move with hours, so a shift toward longer scheduled shifts would inflate the figure without any hourly raise. Employers raised the price of the labour they were said to be replacing, which is both the strongest incentive to automate and evidence that the remaining roles were revalued.

11. Composition: which retail occupations the projections move

The aggregate hides the shape of the change, and the occupational projections show it precisely. The Bureau of Labor Statistics models cashiers falling from 3,157,200 in 2024 to 2,843,600 in 2034, a decline of 313,600 or 9.9%, which is the largest projected numeric decline of any US occupation in the 2024-34 projections (BLS Employment Projections table 1.5), and the agency explicitly attributes it to self-service checkout and online sales^{32,44}. The 2024 base-year levels are survey estimates and are solid. The 2034 endpoints are model output built on assumed technology-adoption and consumer-spending paths, and the agency's own prior decade projections for cashiers were badly wrong in the other direction.

Decline is concentrated in transaction handling rather than across the store. Supervision falls only 5%. Stockers and order fillers, which is physical and spatially variable, is projected to grow by 235,000 or 8.5%. That occupation is SOC 53-7065, a transportation and material-moving code whose employment sits in warehousing as well as in stores, and BLS does not publish the split, so it cannot be read as selling-floor growth without that qualification. That boundary is the same one the deployment ledger drew: embodied AI has taken the repetitive, fixed-position, judgement-free work and has not crossed into the variable physical work or the persuasive work.

Retail salespersons are projected essentially flat at -0.5%, which the Occupational Outlook Handbook describes as little or no change, with about 586,000 annual openings for retail sales workers, while retail trade as a sector is projected to lose more jobs than any other at -1.2% ^{32,44}. Retail salespersons is a broad occupational code spanning car sales and apparel floors with very different exposure, and flat national employment can conceal large compositional and geographic churn. Taken at face value, the projections say the transactional half of the floor job is being automated and the diagnostic, persuasive half is not.

Occupation (SOC)	2024	2034	Change	Median annual wage	Grade
Cashiers (41-2011)	3,157,200	2,843,600	-313,600 (-9.9%)	\$31,190	modelled
Retail salespersons (41-2031)	3,936,700	3,917,100	-19,600 (-0.5%)	\$34,580	modelled
First-line supervisors of retail sales workers (41-1011)	1,432,600	1,360,300	-72,300 (-5.0%)	\$47,320	modelled
Customer service representatives (43-4051)	2,814,000	2,660,300	-153,700 (-5.5%)	not listed here	modelled
Stockers and order fillers (53-7065)	2,764,800	2,999,800	+235,000 (+8.5%)	not listed here	modelled

Table 4. BLS occupational projections, 2024-34. Base-year levels are survey estimates; 2034 endpoints are modelled. Source: BLS Employment Projections ⁴⁴.

12. Scoring thirteen years of automation forecasts

Frey and Osborne assigned cashiers a 0.97 probability of computerisation, ranked 657th of 702 occupations, with counter and rental clerks also at 0.97, retail salespersons at 0.92, stock clerks and order fillers at 0.64, and first-line supervisors of retail sales workers at 0.28 ⁴⁵. The prediction can be scored against outcomes only for the period that has elapsed. Cashier employment stood at 3,157,200 in

2024, more than a decade after a 0.97 computerisation probability was assigned to the occupation; the further 9.9% decline to 2034 is BLS model output and cannot serve as an ex-post test. The direction was right and the magnitude was not. Frey and Osborne were explicit that their probabilities express technical feasibility rather than adoption, cost, regulation or consumer preference, so quoting 0.97 as a prediction of job loss misreads their own claim. A good deal of retail strategy was built on that misreading.

Changing the unit of analysis changes the answer by a factor of five. Arntz, Gregory and Zierahn applied a task-based measure to the same question and found 9% of US jobs at high automation risk against Frey and Osborne's 47% ⁴⁶. Their estimate relies on PIAAC self-reported task data, and workers over-report task variety, which biases it downward, so the honest reading is that the range is wide. The methodological point survives the uncertainty. The unit of analysis matters: a retail job is a bundle of tasks with very different exposure.

The best available ex-post test found slower growth rather than destruction. Georgieff and Milanez examined 21 countries over 2012 to 2019 and found employment grew everywhere, while employment in occupations at high automation risk grew 6% against 18% in low-risk occupations, with low-educated workers already concentrated in the high-risk group and becoming more so ⁴⁷. The window closes before generative AI, and the risk classification inherits the limitations of the task-based measure it builds on. Six per cent against eighteen was measured over seven years; sustained, a gap of that size produces a very different labour market without a single layoff, though the extrapolation is the author's and is not something the study measured. Acemoglu supplies the macroeconomic ceiling: a cumulative increase in total factor productivity of no more than 0.66% over ten years, and likely under 0.53%, implying US GDP roughly 0.93 to 1.16% higher in 2034 ⁴⁸. That is a Hulten's-theorem aggregation assuming no new tasks and no reorganisation of production, exactly the channels his own earlier work says matter most, and it is modelled rather than measured.

Two historical analogues are usually offered as reassurance and both need qualification. Basker found that barcode scanning raised retail labour productivity by low single-digit percentages in the years after adoption, yet cashier employment grew after scanners spread in the 1980s ⁵⁵. Larger checkout-time figures circulate in secondary coverage and are not used here. Scanners automated a step inside the cashier's task, while self-checkout removes the cashier from the transaction and hands the work to the shopper, so the mechanism differs and the analogy supports

scepticism about forecasts rather than a prediction about this one. Bessen's ATM case is cited more often still: tellers per branch fell from about 20 in 1988 to about 13 in 2004 while urban branches rose roughly 43%, so total teller employment grew ⁵⁶. The channel was cost-to-scale, since cheaper branches meant more branches. US teller employment has fallen substantially since the smartphone and US store counts are not rising, so anyone citing the ATM story as reassurance is citing the first half of a two-act history.

13. The entry-level signal and the displacement that leaves no record

Brynjolfsson, Chandar and Chen, using ADP payroll microdata, find workers aged 22 to 25 in the most AI-exposed occupations showing a 16% relative employment decline after conditioning on firm-time effects, while employment for experienced workers in the same occupations held stable ⁴⁹. In raw terms, 22 to 25 year-olds in the highest-exposure quintiles saw employment fall about 6% from late 2022 to September 2025, against 5 to 13% growth in the lowest three quintiles. Customer service representatives are one of the two named exemplar occupations, which puts the signal adjacent to retail. It is a working paper on a single proprietary payroll panel that over-represents large firms, exposure is an occupational index rather than observed tool use, the authors say they cannot establish causation, and the headline moved from 13% to 16% between drafts.

Set against that, the Budget Lab at Yale finds the occupational mix of US employment has changed about 1% since the release of ChatGPT, within historical range and comparable to the early personal computer and early internet periods, with no discernible economy-wide disruption 33 months in, measured to the study's own endpoint in late 2025 rather than to this report's as-of date ⁵⁰. Occupational dissimilarity is built to catch reallocation between occupations and is close to blind to task change within one, which is precisely what is happening to cashiers. A null from that measure is not evidence of no effect. Both findings are carried here.

The mechanism that matters most for commerce may generate no statistics at all. Internal Amazon documents obtained by the New York Times in October 2025 show the automation team expecting to avoid hiring more than 160,000 US workers by 2027 and more than 600,000 by 2033, saving roughly 30 cents per item packed and shipped, with a stated ambition of automating 75% of operations;

executives told the board that robotics could hold US headcount flat while sales doubled by 2033⁴¹. These are leaked planning documents describing an ambition, Amazon disputed the framing, and avoided hires is a counterfactual against Amazon's own growth assumptions that cannot be audited.

Take the mechanism seriously even if the numbers are not auditable. Hiring avoidance produces no layoffs, no WARN notices and no separations. Nearly every publicly funded workforce programme in the United States is triggered by a separation event: WIOA Title I dislocated worker funding, state rapid response teams, and most employer severance obligations. A displacement mechanism that never separates anyone is invisible to all of them. The people affected are those who would have been hired, who are disproportionately young and disproportionately entering through the roles the Stanford paper flags.

14. What the assistant does to the worker, and to which worker

The strongest evidence in this report concerns software assistants rather than robots. Brynjolfsson, Li and Raymond studied 5,179 customer support agents given a generative AI assistant and found issues resolved per hour rose 14% on average, with a 34% improvement for novice and low-skilled workers and minimal effect on experienced high performers⁵¹. The mechanism they identify is a training claim rather than a tooling claim: the assistant propagated the tacit practice of high performers to newcomers. Identification rests on a staggered rollout at one firm rather than random assignment, the setting is scripted text support, and the outcome measured is throughput rather than sales or margin.

Alibaba's field experiment replicates the levelling result at ten times the scale and then breaks the comfortable version of it. Ni and colleagues randomised 5,940 after-sales agents, 2,895 treated and 3,045 control, across roughly 2.56 million chats and 0.39 million customer ratings, covering 15.6% of the official customer service workforce⁵². Intention-to-treat effects were modest: issue-identification time down 8.2%, chat duration down 1.1%, dissatisfaction down 3.4% against baseline, ratings up 0.042 points. At full usage they reach a 32.3% cut in issue-identification time, 15.4% lower dissatisfaction and a 0.184-point rating gain. The distributional result is the important one. Usage fell monotonically from 39.3% among bottom-quintile agents to 18.3% among top-quintile agents, gains concentrated in quintiles one through three, and top-quintile agents got worse on both subjective and objective

quality, with higher customer retrials and shift-away time up 23.8%. Almost a fifth of treated agents never used the tool at all.

A universal rollout with universal training therefore applies an intervention that helps the bottom three-fifths of a workforce and degrades the top fifth. The paper is a July 2026 preprint run with the platform whose product is being evaluated, with a co-author employed there, in Chinese after-sales chat rather than on an American selling floor, and the top-performer decline is inferred from shift-away time and retrials rather than observed directly. Discounted for all of that, it remains the largest randomised evidence available and it points the same way as the smaller studies.

Dell'Acqua and colleagues add the failure mode a worker cannot see. Among 758 BCG consultants working with GPT-4 in 2023, access to the model on tasks inside its capability frontier raised task completion 12.2%, speed 25.1% and human-rated quality 40%, with below-median performers gaining 43% against 17% for above-median performers; on tasks outside the frontier, assisted consultants performed 19 percentage points worse than unassisted ones ⁵³. The researchers constructed the boundary, which is cleaner than any real job has, and the model is a 2023 vintage, so the 19-point penalty is not a current measurement. The transferable finding is that the boundary is invisible from the inside, which turns AI training from tool instruction into judgement instruction.

How much of this reaches sales is open, and the one direct answer is dispersion. Fang and colleagues ran six randomised generative-AI experiments on an online retail platform and found effects on sales ranging from no detectable impact to plus 16.3%, the largest in the pre-sale service chatbot workflow ⁵⁴. The authors include platform consultants and an employee of the platform, the setting is online, and none of the six workflows involves an associate on a selling floor. The same technology, in the same firm, produced anything from zero to sixteen per cent depending on which task it touched.

Study	Population	Design	Headline effect	Low performers	High performers	Overall
Ni et al. 2026 ⁵²	5,940 Alibaba after-sales agents	Randomised, agent level	-8.2% issue-identification time (ITT)	Gains concentrated in quintiles 1-3	Quality declined; shift-away time +23.8%	Quality declined

Brynjolfsson, Li and Raymond 2025 ⁵¹	5,179 support agents, one US firm	Staggered rollout	+14% issues resolved per hour	+34%	minimal	v
Dell'Acqua et al. 2025 ⁵³	758 BCG consultants	Randomised, task level	+12.2% completion inside frontier	+43%	+17%; -19pp outside frontier	
Fang et al. ⁵⁴	Online retail platform workflows	Six randomised experiments	0 to +16.3% on sales	not reported	not reported	n (P

Table 5. Frontline AI experiments, ranked by relevance to a selling floor.

15. De-automation as a workforce event

Every reversal in Table 2 was a workforce event, and none was managed as one in public. When Amazon closed 72 stores, staff were offered redeployment or 90 days of severance at full pay and benefits, a decent package by sector standards, and Amazon never disclosed how many people were affected¹. When Walmart cancelled Bossa Nova, associates in 500 stores went back to a workflow without a robot in it. When Dollar General pulled self-checkout from more than 12,000 stores, someone had to staff lanes again. The research literature on automation and labour is almost entirely about installation. The commerce record says removal happens often enough to need its own plan.

Three costs fall on workers when a system is withdrawn. The first is stranded skill. Competence built around a specific system, knowing which exceptions the gate throws on a Saturday afternoon, how to reseal a robot that has parked itself in dairy, which shelf tags the label system silently fails to update, is real expertise, and almost none of it is documented or credentialed. When the system goes, the expertise is worth nothing and the worker holds no record that it existed.

The second is role reversion. Systems on the floor do not simply add themselves to an existing job; the job is redesigned around them. A checkout-free store has hosts and exception handlers where it once had cashiers. A self-checkout bank has one attendant covering six lanes and an intervention queue. Reverting that design requires the old role back, and the people who could perform it at speed have moved on or were never hired. Restaffing 12,000 Dollar General stores requires

cashiers who can run a lane quickly under pressure, and the pipeline that produced them had been thinned for several years. That last claim is the author's inference from the sequence of events, not a measurement, and no retailer has published re-staffing data that would test it.

A third cost compounds. In a conventional store the register is where a new hire learns the assortment, the regulars, the price points and the rhythm of the floor, and where supervisors watch performance closely enough to decide who to promote. Automating the register removes the observation post along with the task. The projections show first-line supervisors of retail sales workers declining only 5% ⁴⁴, so demand for people who can run a floor persists while the route by which they used to be identified narrows. That is a structural problem in the internal labour market and it appears in no employment count.

An operator can address most of this at installation time and at almost no cost. Record what the role becomes, in writing, at the level of tasks and exceptions. Issue the worker a portable record of the competencies acquired. Keep the pre-installation workflow documented well enough that someone who never performed it could restore it, which is a business continuity requirement as much as a workforce one. None of it can be reconstructed afterwards. Whether reversals themselves raise turnover in the affected stores is unmeasured, and it is listed among the open questions in section 20.

16. Training a frontline workforce for equipment that changes yearly

The training problem in commerce is a throughput problem. BLS models 542,600 cashier openings a year against an average annual net decline of about 31,400, a ratio of roughly seventeen to one, and about 586,000 annual openings for retail sales workers ⁴⁴. Even in the fastest-shrinking occupation in America, replacement demand exceeds automation by more than an order of magnitude. Openings are dominated by separations and the separations model is calibrated on historical turnover, so if automation produces fewer, better-paid and more skilled floor roles, openings will fall faster than the model implies. Taken as it stands, the number moves the task from retraining the displaced to continuously equipping the churning.

That has an unforgiving consequence for how instruction is designed. At 542,600 cashier openings a year against a stock of about 3.16 million, the entry-level role

turns over fast enough that a curriculum sitting outside the work will not be delivered, whatever the training budget says. Instruction has to happen inside the shift, in short units, assessed by doing rather than by seat time. That follows from the openings-to-stock ratio; no median-tenure figure is asserted here, because none was located for this occupation.

Equipment turnover is the second constraint. Table 2 contains programmes that launched and ended inside two years, and vendor interfaces change faster than that. A curriculum organised around a specific system has a shorter useful life than the time needed to write it, accredit it and train the trainers. The alternative is to teach the invariants and treat the interface as a short, repeatable swap. Six invariants recur across the systems in this report, and together they are the specification in Table 6.

Exception handling comes first. Every deployed system generates exceptions and the human role converges on resolving them: the gate that flags a basket, the terminal that will not accept an item, the robot parked in an aisle, the shelf label that disagrees with the register. This is diagnostic work under time pressure in front of a customer, and it is the most consistently underestimated part of every role redesign in the record. Ground truth maintenance is second. Sensing systems are only as good as the human-maintained planogram, price file and label, and Simbe's claim that 42% of detected gaps are fixed instantly presupposes a person who can act on a detection ¹⁴. The competency is understanding what the system believes about the store and how one's own actions change that belief.

Boundary detection is third, and the evidence for it is the strongest in this report. Dell'Acqua's jagged frontier shows assisted workers performing 19 percentage points worse than unassisted ones outside the model's competence, with the boundary invisible from the inside ⁵³. A frontline worker cannot tell whether a customer's question falls inside or outside the assistant's frontier. Calibration is trainable through productive failure: route trainees to questions the assistant answers confidently and wrongly, then have them articulate the tell. That transfers across vendors, because the failure mode belongs to the model class rather than to the interface.

Disclosure to the customer is fourth. Staff are asked to explain what a camera does, why a price changed on a shelf edge and what happens to a receipt, usually with no script and no accurate understanding of their own. The Warren and Casey correspondence and the pricing study around it show this is a live public concern even where the evidence favours the retailer ^{30,31}. Fifth is the diagnostic and

persuasive half of selling, which the projections say survives: retail salespersons flat at -0.5% while cashiers fall 9.9% ⁴⁴. It has the strongest existing pedagogy, the weakest existing measurement, and the clearest link to revenue.

Two design decisions follow from the experimental evidence, and they are the least conventional recommendations here. First, deployment and training have to be differentiated by current performance. The Alibaba result, with usage at 39.3% in the bottom quintile against 18.3% in the top, gains in quintiles one to three, and quality declining in quintile five ⁵², means a universal rollout applies a treatment that harms the best performers. Differentiation requires a performance baseline, so the first deliverable of a frontline AI programme is measurement of the existing workforce rather than a course. That is an instructional design decision and it belongs to whoever owns training.

Second, the credential has to be portable. Turnover means an employer's training investment leaks to competitors within a year, which is the standard reason retail underinvests in it. A vendor-neutral, employer-recognised record of the competencies in Table 6 turns a leaking private investment into a sector asset and makes the churn cheaper for everyone in it. The Institute maintains verifiable credential infrastructure on open standards for exactly this pattern; the barrier is agreement on what the competencies are rather than the technology to record them.

Competency	Why it survives equipment turnover	Evidence anchor	Assessment
Exception handling under customer observation	Every system generates exceptions; only the exception types change	Role redesign in all checkout-free and self-checkout deployments ^{1,11,33}	Timed live resolution with a customer present, scored on resolution and on manner
Ground truth maintenance	Sensing depends on a human-maintained planogram, price file and label	42% of gaps fixed instantly presupposes an actor ¹⁴	Audit of a seeded set of induced discrepancies
	The jagged frontier is a	-19pp outside the frontier ⁵³	Calibration exercise:

Boundary detection on an AI assistant	property of the model class, not the interface		confident wrong answers, trainee must flag before acting
Disclosure and consent explanation	Public concern persists regardless of which sensing system is installed	Warren and Casey letter; ESL pricing study ^{30,31}	Scripted and unscripted explanation, scored for accuracy not reassurance
Diagnostic and persuasive selling	The projections say this half of the role persists	Retail salespersons -0.5% vs cashiers -9.9% ⁴⁴	Observed sale with a rubric; attach to margin where measurable
Differentiated assignment of AI tools	Tool effects run opposite ways by performance quintile	Alibaba quintile reversal ⁵²	Manager competency: baseline first, then assign, then re-measure

Table 6. Competency specification for frontline commerce roles under physical AI. Every competency is stated to be independent of any specific vendor system.

17. What an operator should do, in order

The sequence below is what the evidence in this report supports for a business putting physical AI on a selling floor. It is ordered because the later steps are worthless without the earlier ones. Nothing in it requires a large capital budget, and the first three steps cost staff time rather than money.

Start by defining the measurement, before the purchase. Name the comparison stores, the store-weeks, the baseline metric and who computes it. Require any supplier claiming a labour-hours saving to state the disposition of the hours, meaning whether they leave the payroll or move to another task, and put that in the contract. The absence of this discipline is why the entire category's return-on-investment case currently rests on supplier arithmetic. Then baseline the workforce, since the Alibaba result makes differentiated assignment impossible without knowing who the top and bottom performers are.

Site the system according to the rules in Table 3, install it with the role redesign written down, and hold a documented rollback path. Assume a mean time to withdrawal measured in years rather than decades, given that ten of the eighteen programmes in the ledger were closed, cancelled, converted or materially reduced. No survival rate is claimed from a ledger this report has already described as non-random and closure-biased; this is a planning assumption with no numeric basis asserted. Issue the workers a portable record of what they learned. Re-measure at a fixed interval and be willing to publish, because the first retailer to release a matched-store comparison of any of these systems will set the standard the rest of the sector is measured against.

18. Risk register

The register below covers risks to an operator deploying physical AI in commerce and risks to the workforce affected by it. Likelihood and impact are the author's judgement on the evidence assembled here, and are graded modelled. Leading indicators are chosen to be observable without privileged access.

#	Risk	Likelihood	Impact	Leading indicator	Mitigation
R1	System withdrawn within three years, stranding process and training investment	High	Medium	Vendor funding round without a named reference customer; pilot not expanded after 12 months	Documented rollback path; portable credentials; contract exit terms
R2	Vendor effect claims do not replicate in the operator's stores	High	High	No matched-store study offered; effect quoted without a baseline	Measurement defined before purchase; staged rollout with control stores
R3	Labour hours freed are not	Medium	High	Hours claim stated without disposition;	Contractual disclosure of hours

		converted into service, sales or shrink reduction			no post-install schedule change	disposition; schedule audit at 90 days	
	R4	Universal AI assistant rollout degrades top-quintile performance	Medium	Medium	Usage inversely correlated with performance rank; retrials up among best agents	Baseline first, differentiated assignment, opt-out for top quintile ⁵²	
	R5	Cost transfer rather than cost removal: shrink and customer time absorb the saving	Medium	High	Shrink rising while labour hours fall; queue complaints	Measure shrink and abandonment alongside labour; ten-item limits or attendant ratios	
	R6	Loss-prevention business case written against an unmeasured baseline	High	Medium	Reliance on pre-2023 industry shrink figures	Internal baseline; note that the last industry survey put 63% of loss internal or procedural ³⁶	
	R7	Entry-level hiring quietly contracts, narrowing the supervisor pipeline	Medium	High	Fewer postings at entry grade while supervisor vacancies persist	Explicit apprentice posts; make the observation post part of the role redesign ^{41,44,49}	
	R8	Regulatory action on	Medium	Medium	State legislation on	Disclosure training;	

	in-store sensing and dynamic pricing			shelf-edge pricing or biometric inference; further congressional correspondence	avoid demographic inference at displays ³¹
R9	Public backlash to visible machines on the floor	Medium	Medium	Negative local coverage; store-level complaint volume	Customer-facing explanation as a trained competency; night-hours operation ¹²
R10	Density assumption fails and centralised automation is stranded	Medium	High	Orders per site below underwriting case for two consecutive quarters	Store-node fulfilment as default; stage-gate capital on density [19–25]

Table 7. Risk register. Likelihood and impact are the author's judgement, not measurements.

19. Scenarios to 2035

Four scenarios are set out below. They differ in one variable, which is where the machine sits relative to the customer, and they resolve differently for employment and for training. Likelihoods are the author's judgement on the evidence in this report and are graded modelled. All employment figures are anchored on the BLS 2024 base of 3,157,200 cashiers and 3,936,700 retail salespersons ⁴⁴.

The scenario this report judges most likely is the second, in which sensing and labelling saturate the back of the store and the customer-facing exit stays optional. It is the direction every operator in Table 2 moved once they had run the numbers, and it is consistent with both the projections and the productivity ceiling.

Two signposts would move the judgement quickly. If a large grocer publishes a matched-store comparison showing a shrink or availability effect from computer vision that survives scrutiny, the capital case for full sensing coverage strengthens and scenario three becomes reachable. If Amazon's US headcount flattens while its sales grow, in the pattern its internal documents describe ⁴¹, the hiring-avoidance

mechanism is confirmed and the entry-level contraction in scenario four becomes the central case for logistics-adjacent commerce work.

Scenario	Mechanism	Selling floor in 2035	Workforce consequence	Likelihood
1. Reversion	Sensing capital fails to earn out; further withdrawals follow Kroger and Dollar General	Machines confined to backrooms and price labels	Cashier decline shallower than the BLS model; training reverts to conventional retail practice	Low
2. Hybrid saturation	Back-of-store sensing and labelling saturate; exit stays optional; assistants go to every associate	Every store instrumented, every customer path choosable	Composition shifts as projected; the binding constraint is onboarding 500,000+ people a year into a changed role	High
3. Sensed store	A published controlled result unlocks capital for full-coverage vision	Continuous shelf and basket sensing; checkout as one option among several	Cashier decline exceeds the model; exception handling and ground truth become the floor job	Medium-low
4. Quiet contraction	Hiring avoidance rather than layoffs; entry-level postings shrink across commerce and logistics	Similar to 2026 in appearance, thinner in staffing	No layoff statistics move; the supervisor pipeline narrows; workforce programmes triggered by	Medium

Table 8. Scenarios to 2035. Likelihood is the author's judgement, graded modelled.

20. Open questions and what is not known

Several things this report would need in order to be more useful do not exist in public. No retailer or vendor has published a matched-store controlled comparison of a shelf-scanning robot, a checkout-free installation, or computer-vision loss prevention. No operator has published the disposition of labour hours freed by any of these systems, which is the single fact that determines whether the technology raises what people can do or removes their hours. Amazon has never published per-store economics for Go or Fresh, so the closure of 72 stores cannot be attributed cleanly to the technology.

On the workforce side, four questions are open and answerable with data someone already holds. Does a system withdrawal raise turnover in the affected stores, and by how much? What happens to promotion rates into first-line supervision after the register is automated, given that supervisor demand falls only 5% in the projections ⁴⁴? Do the quintile effects found in Alibaba's after-sales operation appear on a physical selling floor, where the work is spoken, mobile and unscripted? And does Amazon's actual US headcount trajectory since October 2025 match the hiring-avoidance plan described in the leaked documents ⁴¹?

Three items in this report are weaker than the rest and are marked here so they are not carried forward as settled. The Blue Jay shutdown rests on single-outlet reporting accessed second-hand with no Amazon confirmation and no public definition of what shut down means ³⁹. The Alibaba top-performer quality decline is inferred from shift-away time and customer retrials rather than directly observed ⁵². The claim in section 15 that Dollar General's re-staffing was constrained by a thinned cashier pipeline is the author's inference from the sequence of events and has no supporting measurement.

21. Conclusions

The selling-floor record between 2016 and 2026 sorts cleanly by one variable, which is how much of the change the customer has to absorb. Electronic shelf labels are heading for all 4,597 Walmart stores because nobody outside the building

experiences them ²⁸. Checkout-free stores closed as a format and survived as a licensed product in stadiums and airports, where a queue at halftime is the operator's whole problem ^{1,5}. Shelf robots survived by being sold as a data subscription instead of a labour substitute ^{14,18}. Centralised robotic fulfilment lost \$2.6 billion of book value at Kroger while Walmart put the same category of machine inside stores it already owns ^{19,26,27}. None of that is a statement about what the technology can do. It is a statement about where it earns.

On the workforce side the picture is neither collapse nor continuity. Retail employment in June 2026 was 15,459,700, below its February 2020 level while national employment grew by 6.7 million ⁴². The BLS models the largest numeric occupational decline in the country in cashiers, and models 542,600 cashier openings a year alongside it ⁴⁴. Both are true at once, and the second one is the operative fact for anyone designing training. Half a million people a year enter a role whose content is being rewritten by equipment with a shorter service life than the training cycle.

Experimental evidence says the technology can raise what a frontline worker does, and says it does so unevenly enough that a uniform rollout will harm the best performers ^{51,52,53}. That makes deployment an instructional design problem before it is an IT problem, and it puts the decision in the hands of whoever owns training. The competencies that survive equipment turnover are specified in Table 6, the measurement discipline an operator needs is in section 17, and the things nobody has measured are listed in section 20. This report will be revised when someone publishes a matched-store comparison, and the Institute will publish that revision whichever way the numbers fall.

References

1. ¹ Amazon, corporate statement on refocusing grocery investment on Whole Foods Market and delivery (aboutamazon.com), announced 27 January 2026, page updated 12 May 2026.
SELF-PUBLISHED
2. ² Newsweek, reporting on the closure of all Amazon Fresh and Amazon Go stores, 27-28 January 2026.
REPORTED
3. ³ Axios, coverage of the Amazon Fresh and Amazon Go closures, 27-28 January 2026.
REPORTED

4. ⁴ Supermarket News, store-count reporting on the Amazon Fresh and Amazon Go closures, 28 January 2026.
REPORTED
5. ⁵ SiliconANGLE, Just Walk Out third-party deployment count, 18 December 2025.
REPORTED
6. ⁶ Retail TouchPoints, Just Walk Out third-party location reporting, 2025-2026.
REPORTED
7. ⁷ Axios Nashville, Just Walk Out concessions at Nissan Stadium, 8 January 2026.
REPORTED
8. ⁸ CNBC, Amazon to close all 19 Amazon Fresh stores in the United Kingdom, 23 September 2025.
REPORTED
9. ⁹ CNN, Amazon Fresh UK closure coverage, 23 September 2025.
REPORTED
10. ¹⁰ Grocery Gazette, Sainsbury's removes Just Walk Out from its Holborn store, 4 February 2025.
REPORTED
11. ¹¹ The Grocer, coverage of the end of Aldi's Greenwich Shop&Go trial and Tesco's move of GetGo to a hybrid till format, 2024-2025.
REPORTED
12. ¹² Wall Street Journal, Walmart ends its Bossa Nova shelf-scanning contract, November 2020; relayed by CNBC, TechCrunch and Retail Dive, 2 November 2020.
REPORTED
13. ¹³ The Robot Report, Bossa Nova layoffs and wind-down of retail robotics operations, 2020.
REPORTED
14. ¹⁴ Simbe Robotics, tenth-anniversary press release (PR Newswire), 12 November 2025.
SELF-PUBLISHED
15. ¹⁵ Fast Company, Simbe client-outcome claims relayed from the company, 2025.
SELF-PUBLISHED
16. ¹⁶ Supply Chain Dive, Simbe out-of-stock and fulfilment-time claims relayed from the company.
SELF-PUBLISHED
17. ¹⁷ Simbe Robotics, Tally Spot announcement, January 2025; reported by Grocery Dive and The Robot Report.
SELF-PUBLISHED
18. ¹⁸ Badger Technologies (a Jabil division), Digital Teammate positioning and deployment count.
SELF-PUBLISHED

19. ¹⁹ Kroger, announcement of automated fulfilment network review and facility closures, 18-19 November 2025.
SELF-PUBLISHED
20. ²⁰ Grocery Dive, Kroger Ocado facility closures and impairment, November-December 2025.
REPORTED
21. ²¹ Supply Chain Dive, Kroger-Ocado compensation and 2026 e-commerce profitability guidance, December 2025.
REPORTED
22. ²² Forbes, analysis of Kroger's automated fulfilment retreat, 8 December 2025.
REPORTED
23. ²³ Chain Store Age, Kroger fulfilment restructuring coverage, November 2025.
REPORTED
24. ²⁴ Takeoff Technologies, Chapter 11 petition, US Bankruptcy Court for the District of Delaware, 30 May 2024.
VERIFIED
25. ²⁵ Grocery Dive, Takeoff Technologies bankruptcy and section 363 sale to Woolworths, 2024.
REPORTED
26. ²⁶ Walmart investor meeting guidance on automation coverage and unit-cost improvement (Tampa), relayed by Fox Business, Chain Store Age and Retail TouchPoints, 2025-2026.
SELF-PUBLISHED
27. ²⁷ Symbotic, investor relations release on the acquisition of Walmart's Advanced Systems and Robotics business and the 400 accelerated pickup and delivery centre commitment, October 2025.
SELF-PUBLISHED
28. ²⁸ CNBC, Walmart to install electronic shelf labels in all US stores by the end of 2026, 21 March 2026.
REPORTED
29. ²⁹ Grocery Dive and Retail Dive, Walmart electronic shelf label rollout with VusionGroup, 2024.
REPORTED
30. ³⁰ Robert Sanders (UC San Diego Rady), Ioannis Stamatopoulos (UT Austin McCombs) and Robert Bray (Northwestern Kellogg), study of electronic shelf labels and retail price adjustment across 114 stores and 180+ million product-level observations; SSRN working paper, publicised by UC San Diego, 22 July 2025.
REPORTED
31. ³¹ Senators Elizabeth Warren and Bob Casey, letter to Kroger on electronic shelf labels and dynamic pricing, August 2024, and Kroger's response.
VERIFIED

32. ³² US Bureau of Labor Statistics, Occupational Outlook Handbook entries for cashiers and retail sales workers, Employment Projections 2024-2034, released 28 August 2025.
MODELLED
33. ³³ NPR, retailers pulling back on self-checkout, 18 March 2024.
REPORTED
34. ³⁴ NBC News, Dollar General, Target and Five Below self-checkout reductions, 2024.
REPORTED
35. ³⁵ LendingTree, self-checkout survey of more than 2,000 US consumers, December 2025.
REPORTED
36. ³⁶ National Retail Federation, National Retail Security Survey, final edition covering fiscal 2022, and NRF's statement discontinuing the series after 32 editions.
REPORTED
37. ³⁷ Retail Dive, analysis of the National Retail Federation's retracted organised-retail-crime statistic and of the survey's discontinuation.
REPORTED
38. ³⁸ Everseen, product materials on self-checkout lane coverage, and Forrester Total Economic Impact study commissioned by Everseen, September 2024.
SELF-PUBLISHED
39. ³⁹ Business Insider, reporting that Amazon's Blue Jay station was shut down, relayed by BGR, 5 March 2026. Single outlet, unconfirmed by Amazon.
REPORTED
40. ⁴⁰ Amazon, Blue Jay multi-arm robotic station launch announcement, 22 October 2025.
SELF-PUBLISHED
41. ⁴¹ The New York Times, reporting on internal Amazon documents describing automation and hiring-avoidance plans, 21 October 2025.
REPORTED
42. ⁴² US Bureau of Labor Statistics, Current Employment Statistics, series CES4200000001 (retail trade, all employees) and CES0000000001 (total nonfarm, all employees); retrieved from the BLS public API, August 2026.
VERIFIED
43. ⁴³ US Bureau of Labor Statistics, Current Employment Statistics, series CES4200000011 (average weekly earnings, retail trade); retrieved from the BLS public API, August 2026.
VERIFIED
44. ⁴⁴ US Bureau of Labor Statistics, Employment Projections program, 2024-34 occupational projections and occupational separations and openings tables; retrieved from data.bls.gov, August 2026.
MODELLED

45. ⁴⁵ Carl Benedikt Frey and Michael A. Osborne, 'The Future of Employment: How Susceptible Are Jobs to Computerisation?', Oxford Martin School (2013); Technological Forecasting and Social Change 114 (2017), 254-280. Probabilities read from the appendix ranking table.

VERIFIED

46. ⁴⁶ Melanie Arntz, Terry Gregory and Ulrich Zierahn, 'The Risk of Automation for Jobs in OECD Countries: A Comparative Analysis', OECD Social, Employment and Migration Working Paper No. 189 (2016).

VERIFIED

47. ⁴⁷ Alexandre Georgieff and Anna Milanez, 'What happened to jobs at high risk of automation?', OECD Social, Employment and Migration Working Paper No. 255, 21 May 2021. Research funded by the Google.org foundation.

VERIFIED

48. ⁴⁸ Daron Acemoglu, 'The Simple Macroeconomics of AI', Economic Policy 40(121), 2025, 13-58; NBER Working Paper 32487.

MODELLED

49. ⁴⁹ Erik Brynjolfsson, Bharat Chandar and Ruyu Chen, 'Canaries in the Coal Mine? Six Facts about the Recent Employment Effects of Artificial Intelligence', Stanford Digital Economy Lab, November 2025 version, using ADP payroll microdata.

REPORTED

50. ⁵⁰ Martha Gimbel, Molly Kinder, Joshua Kendall and Maddie Lee, 'Evaluating the Impact of AI on the Labor Market: Current State of Affairs', The Budget Lab at Yale, October 2025, with subsequent CPS updates.

REPORTED

51. ⁵¹ Erik Brynjolfsson, Danielle Li and Lindsey Raymond, 'Generative AI at Work', Quarterly Journal of Economics 140(2), 2025, 889-942; NBER Working Paper 31161.

VERIFIED

52. ⁵² Xiao Ni, Yiwei Wang, Tianjun Feng, Lauren Xiaoyuan Lu, Yitong Wang and Congyi Zhou, 'Generative AI in Action: Field Experimental Evidence from Alibaba's Customer Service Operations', preprint, July 2026. One co-author is employed by Alibaba Group.

REPORTED

53. ⁵³ Fabrizio Dell'Acqua, Edward McFowland III, Ethan Mollick, Hila Lifshitz-Assaf, Katherine Kellogg, Saran Rajendran, Lisa Kraye, Francois Candelon and Karim Lakhani, 'Navigating the Jagged Technological Frontier', Organization Science (2025); first circulated as Harvard Business School Working Paper 24-013 (2023).

VERIFIED

54. ⁵⁴ Lu Fang, Zhe Yuan, Kaifu Zhang, Dante Donati and Miklos Sarvary, 'Generative AI and Sales Productivity: Field Experiments in Online Retail', working paper, arXiv:2510.12049v5.

REPORTED

55. ⁵⁵ Emek Basker, 'Raising the Barcode Scanner: Technology and Productivity in the Retail Sector', *American Economic Journal: Applied Economics* 4(3), 2012.

VERIFIED

56. ⁵⁶ James Bessen, 'Learning by Doing: The Real Connection between Innovation, Wages, and Wealth' (2015) and 'How Computer Automation Affects Occupations', NBER Summer Institute (2016).

REPORTED

Market and economic review, not investment advice. Figures carry the verification grade under each reference: **verified** (peer-reviewed or independently replicated), **reported** (a named source stated it, not independently confirmed), **self-published** (the organisation's own figure) and **modelled** (computed here, not measured). A modelled figure is never presented as a measurement. Corrections are welcome and will be recorded.