Six Signs Your Planning System is Modeling the Stop, Not the Item

Why Furniture Distribution Breaks Every Assumption a Traditional TMS is Built On

AUGUST 2026 7 MIN READ

A two-person crew is forty minutes into a delivery, for which the actual plan allowed only 20. The sectional came up the stairwell without trouble. It will not make the turn at the top. One of the crew is on the phone to the branch asking whether to take it back or leave it in the hallway. Three deliveries are booked behind this one, and the second is a wardrobe that has to be assembled in the room it is destined for.

The route was built the night before. It assumed every stop would take the same amount of time, because the address was the only thing the planner could model.

That assumption is borrowed, and it is borrowed from a business that earns differently. In a parcel network the stop genuinely is the planning unit. Service time per stop is roughly predictable, the vehicle is interchangeable, the driver's task is uniform, and mode is decided before the order reaches operations, so a planner can model the network as a set of addresses with a fixed cost of visiting each one. Furniture distribution violates every one of those assumptions, because

almost every operational variable derives from what is being delivered rather than from where it is going.

The reflex is to call this a bad day, or a difficult building, or a crew that should have measured first. It is none of those. It is what happens when a planning system treats the stop as the unit of work in a category where almost nothing about the work is a property of the stop.

There is now a commercial reason this matters more than it did three years ago. The US third-party big-and-bulky last-mile market grew at 10.6% a year from 2017 through 2025, and is projected to grow at 5.1% through 2027. Growth has roughly halved. At 10.6% an operator can absorb structural inefficiency and still look like it is winning. At 5.1% the margin has to come out of the operation.

Six signs it is not coming out yet.

THE 30-SECOND READ

- Assembly time, crew size, mode eligibility and access all vary by SKU rather than by address. A route built on an averaged service time is not slightly late. It is arithmetically incapable of finishing, and the error is discovered at the third stop.
- A furniture trailer fills its cube long before it reaches its weight limit. Standard utilization measures are tonnage-based, so the loss is invisible in the number the operation reports and the region gets flagged for consolidation it has already done.
- Relay runs and unassigned vehicles break trip-level attribution. When two drivers move one load, cost and pay have to be calculated per driver and per segment. A system reading a trip total produces numbers that are structurally wrong rather than slightly imprecise.
- Most furniture operators run wholesale truckload, consumer white glove, purchased capacity and direct e-commerce at once, on separate systems that cannot see each other. Loads that should

combine do not, and under-filled capacity is deferred instead of consolidated.

- Furniture is refused at the door for reasons that are not operational faults, and almost nobody captures why in a form that can change anything upstream.
- A delivery slot is inventory with a price. Pricing it is the one operational change in this category that generates revenue rather than reducing cost, and very few operators do it.

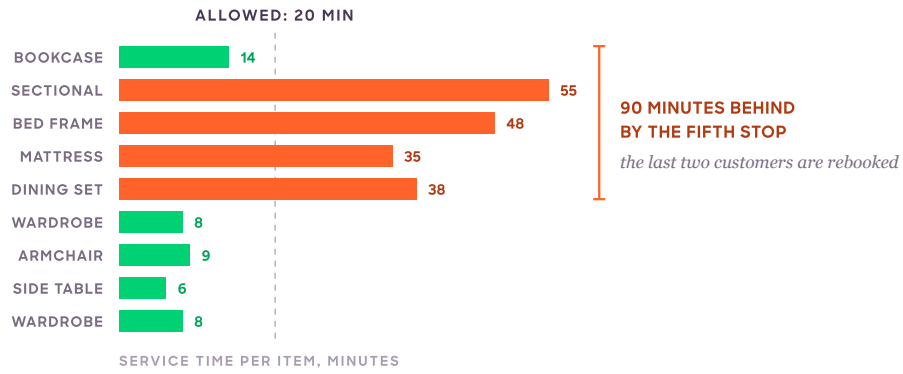
Sign 1 Both flows, acute in consumer white glove

Service time belongs to the item

Fig. 1 Service time by item against the flat average the router assumed

The router allowed twenty minutes a stop. By the fifth stop the crew was ninety minutes down, and nothing in the plan could have said so.

— SCROLL TO EXPLORE →



The plan was not optimistic. It was solving a different problem.

Illustrative model. Detention figure from ATRI, Costs and Consequences of Truck Driver Detention, 2024 ↗

What it looks like. Utilization looks healthy and the route looked feasible when it was published. By mid-morning the crew is behind, and the gap is widening rather than holding, because each complex item adds its overrun to every stop that follows it.

> **What we tell ourselves.** *"You cannot predict how long a delivery takes. Every building is different and every customer is different, so you build in slack and accept some slippage."*

What it actually is. Service time in this category is a property of what is being delivered. Some items can be assembled in advance and loaded built. Others physically cannot be assembled before they reach the room they are destined for. A single driver handles some items, others require two people, and white glove adds room-of-choice placement, assembly, packaging removal and old-item removal to the same stop. When a system models service time as a per-stop constant, or as a constant per vehicle type, it is not making a small approximation. It is solving a different problem from the one the crew will meet, and the route it produces was never completable. Item-level transaction time has to be a modeled attribute, configurable per SKU, and it has to flow into both the route sequence and the window the customer is given.

Two further variables compound it. Crew composition is a planning input rather than a dispatch afterthought, because whether a stop needs one person or two is decided by the item and not by the route. And whether a driver assists with unloading varies by customer and by stop type, which means the same address can carry two different service times depending on the arrangement. On the wholesale side the waiting itself is significant and largely invisible: ATRI records drivers detained at 39.3% of all stops, losing between 117 and 209 hours a year. A system holding one service time per address cannot represent any of this, which is why the error is never a small one.

THE OPERATING REALITY

Nine items across seven stops, twenty minutes allowed for each. Two flat-pack wardrobe drops take eight. A sectional with room-of-choice placement takes fifty-five. A mattress delivery with old-item removal takes thirty-five. The crew is ninety minutes down by the fifth stop and the last two customers are called and rebooked. Both had taken the morning off work.

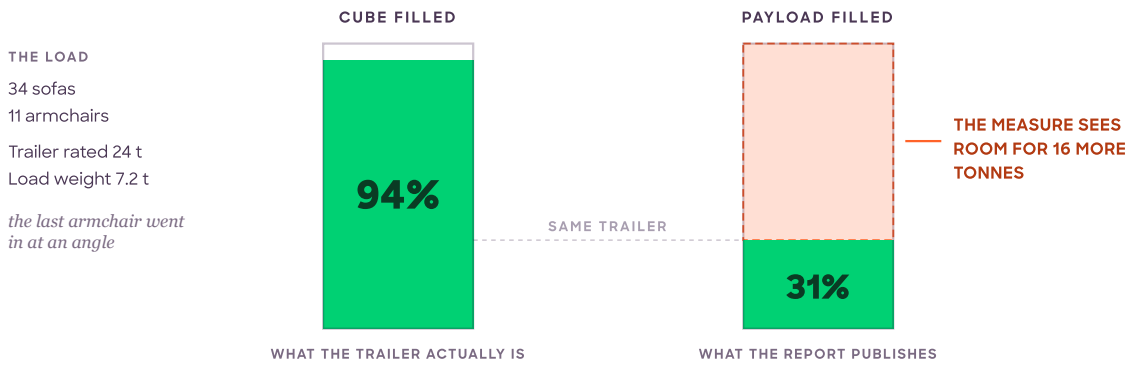
Sign 2 Wholesale multi-stop truckload

The trailer cubes out before it weighs out

Fig. 2 One trailer, measured two ways

Ninety-four percent of the cube. Thirty-one percent of the payload. Only one of those numbers reaches the utilization report.

— SCROLL TO EXPLORE →



The trailer is full. The measure cannot say so.

Eurostat, road freight transport by journey characteristics ↗. Trailer figures are our arithmetic on a representative furniture load.

What it looks like. Trailers leave visibly full. The utilization report says there is room. A planner is asked why fleet utilization sits in the thirties

and has no answer the reporting system can represent.

> **What we tell ourselves.** *"Furniture is light and bulky, everybody knows that. Our utilization numbers have always looked low for the category, and we benchmark against ourselves anyway."*

What it actually is. Furniture consumes cubic capacity long before it consumes payload, so planning against weight limits alone leaves trailers physically full and nominally under-loaded. This is not a local reporting quirk, it is how the category's official statistics are built. Eurostat reports EU road freight load factors in tonnes, and records that 21.6% of total EU road freight vehicle-kilometers were driven empty in 2024, with national transport running emptier than international at 24% against 13%. A tonnage-based measure is structurally incapable of seeing a full furniture trailer, which means an operator optimizing against it is optimizing against a number that cannot describe the constraint. Planning against modeled volume is what turns the invisible loss into fewer trips, and in a category where every trip carries a crew as well as a vehicle, removing a trip is worth considerably more than shortening one.

Two consequences follow from the same physical fact. Mode eligibility is an item property: above certain dimensional and weight thresholds an order stops being parcel-eligible and has to move to LTL or dedicated capacity, and that threshold can be crossed by adding one line to a basket. And network design is an item property too. Furniture is generally not packaged for shared-pallet freight environments, and damage in transit is expensive on high-value goods with subjective acceptance criteria. That single fact is why many furniture operators run their own multi-stop truckload models rather than tendering to shared networks. The network exists because of the product, not because of the geography.

THE OPERATING REALITY

Thirty-four sofas and eleven armchairs on a trailer rated to 24 tonnes. The load weighs a little over seven. The cube is at ninety-four percent and the last armchair went in at an angle. The utilization report shows the trailer at thirty-one percent and flags the region for consolidation.

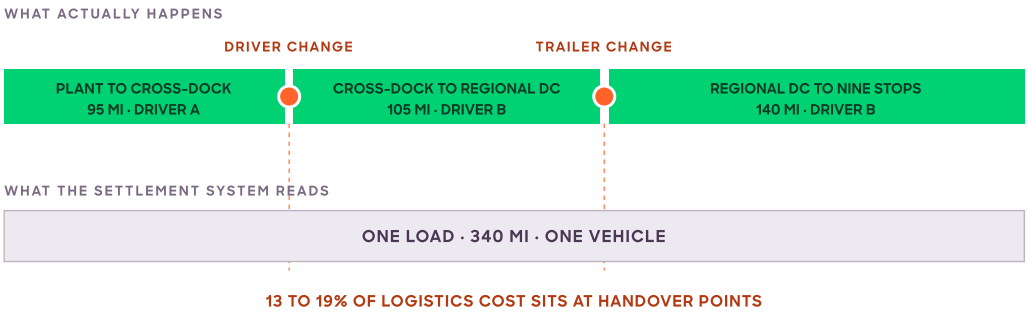
Sign 3 Wholesale, and any operation running relays

Nobody drove the load

Fig. 3 Where the cost of a relayed load actually sits

Two drivers, three segments, one trip total. The pay engine reads a number nobody drove.

— SCROLL TO EXPLORE →



The handover is where the cost is, and where the record stops.

McKinsey & Company, Digitizing mid- and last-mile logistics handovers to reduce waste ↗

What it looks like. Settlement disputes that recur every period and get resolved by hand. Cost per load figures nobody trusts enough to price from. A modernization business case that stalls because no two people agree on what a load actually cost.

> **What we tell ourselves.** *"The pay model is complicated because the operation is complicated. We will clean up attribution after the system goes in."*

What it actually is. Relay is common in furniture wholesale. A first driver takes a load to a transfer point and a second completes the multi-stop run. Combine that with unassigned vehicles, where drivers are not tied to a specific truck, and the attribution model most systems rely on stops working. If pay and cost derive from miles driven, and a load's journey splits across two or three drivers, then attribution has to be per driver and per segment. It cannot be per load, because no single driver drove the load. It cannot be per vehicle, because no single driver drove the vehicle. A system reading a trip total or a vehicle odometer produces numbers that are structurally wrong rather than slightly imprecise. McKinsey costs the general case at 13 to 19% of logistics costs, up to $95 billion a year in the US economy alone, arising from inefficient interactions at handover points, in what it calls blind handoffs. A relay without segment-level attribution is a blind handoff with a payroll attached.

Driver segmentation compounds it. Furniture wholesale fleets typically segment drivers by role: fixed shuttle runs between two facilities, long-haul point-to-point with drop-and-hook, and multi-stop delivery runs. Each carries different qualification requirements, different pay treatment, and different eligibility for particular work. A planning system either models those groups as constraints, in which case allocation is automatic, or it does not, in which case a dispatcher holds the segmentation in their head and the operation depends on that person being at work.

This is worth dwelling on, because it is frequently the thing that stalls modernization entirely. Operators discover that consolidating systems requires resolving attribution first, and attribution turns out to be a compensation design question rather than a software one. Recognizing that early, and treating simplification of the pay model as part of the technology program rather than a prerequisite to it, is what separates

the operations that get through this from the ones that stop at the business case.

THE OPERATING REALITY

A load moves 340 miles across three segments. One driver runs the shuttle from the plant to the cross-dock and back. Another takes the trailer out on a nine-stop run, in a truck she has not driven before. The settlement system sees one load, 340 miles, one vehicle. It pays one driver correctly and the other approximately, and the cost per stop it reports for that region is wrong by enough to misprice the account at renewal.

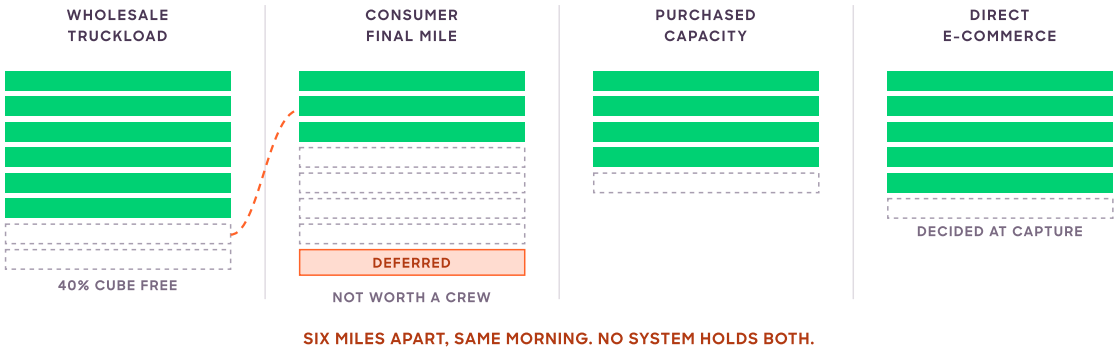
Sign 4 Both flows, and the boundary between them

Four modes that cannot see each other

Fig. 4 The load that could have combined

Four planning systems, one morning, six miles apart. Each decision was correct inside its own system.

— SCROLL TO EXPLORE →



Nothing here is a bad decision. There is just no system that could make the good one.

Gartner, Future of Supply Chains 2026 ↗

What it looks like. A load that cannot be combined across an artificial regional boundary. Under-capacity volume pushed to a later date rather than consolidated with adjacent work. Nobody able to answer whether the same driver could serve a wholesale stop and a consumer delivery on one run, because no system holds both.

> **What we tell ourselves.** *"Those are genuinely different businesses with different service models. Keeping them on separate systems reflects how the operation actually works."*

What it actually is. Most furniture operators run several distribution models at once, each with different planning logic. Wholesale multi-stop truckload moves manufactured goods to distribution centers and on to trade customers, built the day before and batched by region. Consumer final mile serves doors from regional fulfillment centers with scheduled white glove or threshold delivery: fewer stops, longer service times, mandatory customer presence, and a customer experience component absent from the B2B flow. Purchased capacity adds a commercial layer alongside the operational one. Direct e-commerce needs rate and mode decisions at the moment of order capture rather than in a planning cycle.

Each usually sits on a separate platform, and that separation is the root cause underneath most of the specific complaints operators raise. The industry-wide version of the gap is documented: Gartner finds that 95%

of supply chains must react quickly to change while only 7% can execute decisions in real time, and expects investment in real-time decision execution to rise fivefold by 2028. Four systems that cannot see each other cannot re-decide together, whatever any one of them manages alone. What is needed is not one screen showing four systems, but one system making allocation decisions with visibility of all available capacity, whether owned, contracted or purchased.

THE OPERATING REALITY

A consumer route in one postcode has three deliveries on a Tuesday and is held over, because it is not worth sending a crew. Six miles away, on the same morning, a wholesale trailer leaves for a trade customer with forty percent of its cube free. Both decisions were correct inside their own system. Together they are a crew day and a trailer, spent to move work that could have moved once.

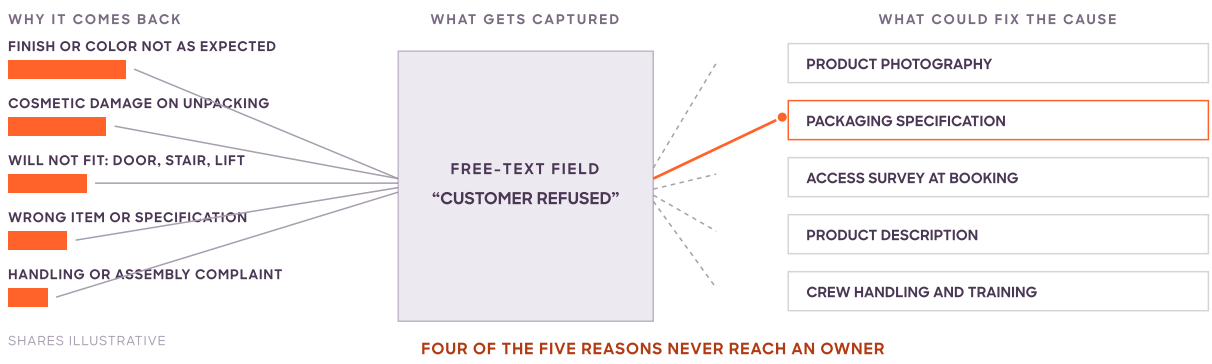
Sign 5 Both flows, acute in consumer

The delivery that comes back

Fig. 5 Refusal reasons, and what each one should have changed

Five reasons a furniture delivery comes back. Four are not operational faults, and none of them reach the team that could fix the cause.

— SCROLL TO EXPLORE →



The freight comes back. The reason does not.

Illustrative composition. Returns baseline from National Retail Federation with Happy Returns, 2025 Retail Returns Landscape ↗

What it looks like. Freight arriving back at a warehouse that was not expecting it. A refusal rate everyone has a theory about and nobody can decompose. Discount-or-return decisions made differently by every branch.

> **What we tell ourselves.** *"Some customers change their minds. You cannot design around taste."*

What it actually is. Furniture refusal at the door is structurally different from a parcel return, and the difference is subjective acceptance. A customer can refuse because the finish is not the shade they expected, or because of minor cosmetic damage invisible until unpacking. Neither is a fault in the operational sense, and both produce the same outcome, which is freight going back. Returns are already a material line across retail: the National Retail Federation puts 15.8% of annual retail sales as returned in 2025, totaling $849.9 billion, rising to 19.3% for online sales. Furniture sits on top of that baseline with a mechanism the baseline does not describe.

Access failure is the same problem arriving earlier. A validated address tells you where a building is. It does not tell you whether a particular item fits through the doorway, up the stairwell, or into the lift. For furniture the item and the building have to be evaluated together, and

neither alone is sufficient, which makes an access refusal less a delivery failure than a planning input nobody collected.

Three requirements follow, and most operations meet none of them systematically. Reason codes captured at the door, structured rather than free text, because the mix of refusal reasons is what tells you whether to fix packaging, photography, description or crew handling. Warehouse notification on refusal, immediately, because inbound freight nobody expected creates a receiving problem and delays whatever resolution the customer is waiting for. And retention economics as a modeled decision, because the cost of recovery, inspection and resale on a used item frequently exceeds the discount required to keep it in place. That calculation deserves to be a rule the system applies with visibility, rather than a judgment made differently by every branch.

THE OPERATING REALITY

A walnut sideboard is refused because the finish photographed lighter than it looks in the customer's hallway. The crew takes it back. The warehouse learns it is coming when it arrives. The reason is typed into a free-text field as “customer refused.” Nine more come back that quarter for the same reason, the photography is never changed, and no report exists that could have connected the two.

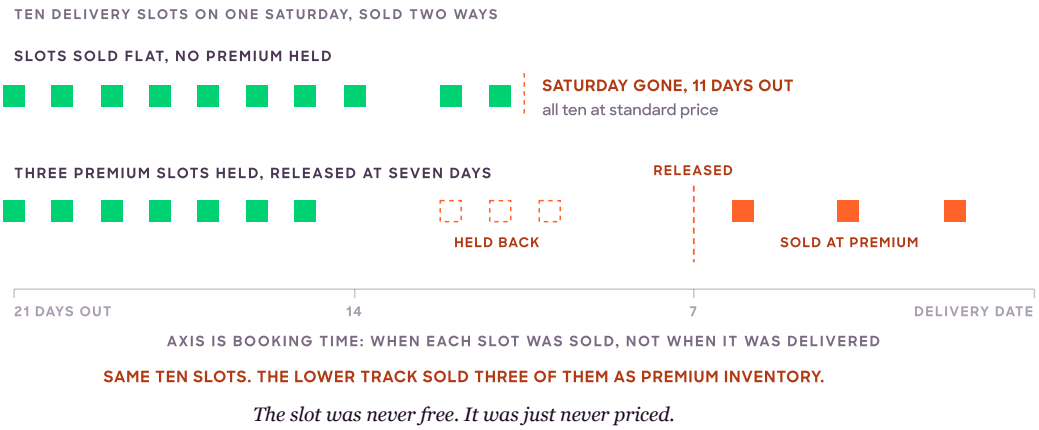
Sign 6 Consumer is the revenue case, wholesale the retention case

The window is inventory

Fig. 6 Slot capacity, held and released

A Saturday slot booked three weeks out at no premium is inventory sold at the wrong price.

— SCROLL TO EXPLORE →



The slot was never free. It was just never priced.

Illustrative model. Consumer preference from Capgemini Research Institute, The Last-Mile Delivery Challenge ↗

What it looks like. Delivery slots offered as a courtesy. A Saturday morning that fills three weeks out with customers who would happily have taken a Tuesday. Trade customers asking for narrower windows and being told what the operation can do rather than what it would cost.

> **What we tell ourselves.** *"Charging for delivery windows is a retail pricing decision, not an operations one. And our customers would not pay it."*

What it actually is. Two expectation shifts are reshaping requirements faster than most systems can accommodate. Trade customers have adopted consumer standards. Buyers who track their own consumer parcels to a two-hour window no longer accept "sometime that morning" from a truckload delivery, and the comparison is unavoidable. McKinsey's 2026 survey of roughly 4,000 B2B decision-makers found that 51% say a lack of customer tracking across channels is an impediment to doing business, and 54% of those likely to switch suppliers cite poor digital customer experience as a reason. Wholesale delivery precision has become a differentiator in a channel where it was never previously measured.

And consumers will pay to influence the window, which means slots have a price. Capgemini's research found that 73% of customers consider a convenient delivery time more important than a fast one. A window that suits the route is worth offering at a discount. A window that does not is worth charging for. The capability required is more subtle than pricing: capacity has to be reserved for premium slots against forecast demand and released dynamically as the date approaches, so early standard bookings do not consume the inventory premium bookings would have paid for. Very few furniture operations do this today, and it is one of the only operational changes in this category that generates revenue rather than reducing cost.

THE OPERATING REALITY

A customer books a Saturday morning delivery nineteen days out, free, because Saturday was in the list. Eleven days later another customer asks for that same Saturday and would have paid a premium for it. The slot is gone. The first customer, asked afterward, says Tuesday would have been fine.

What changes

Model the item, reconcile the modes, re-decide during the day

Item-level constraint modeling. SKU-level transaction time, crew requirement, dimensional and volumetric attributes, mode eligibility and access requirements as first-class planning inputs rather than notes on a manifest. This single capability addresses the largest source of plan failure in the category.

One decisioning layer across modes. Not one screen showing four systems, but one system making allocation decisions with visibility of all available capacity, whether owned, contracted or purchased. That is what allows loads to combine across artificial boundaries, and under-utilized capacity to be filled rather than deferred.

Continuous re-decisioning with human control. Long service times and high variability mean plans need revision during execution, and revision should be scoped to affected routes rather than rebuilding the network. Equally important, and often understated by technology vendors: operators in this category generally do not want fully autonomous decisioning. They want a system that surfaces options, explains its reasoning, and lets an experienced dispatcher choose, with autonomy extended by decision class once the pattern has proven itself. That preference is worth stating plainly, because the analyst trajectory points elsewhere. Gartner predicts 60% of supply chain disruptions will be resolved without human intervention by 2031. The distance between that destination and what a dispatcher will accept in the next planning cycle is exactly why governed autonomy is a design position rather than a hedge.

THE SEQUENCING POINT

There is a fourth requirement that is less about capability than about order of operations. Attribution and pay complexity will surface during any consolidation attempt in this category. Treating it as part of the program, rather than as something to resolve before starting, is the difference between a modernization that completes and one that stalls at the business case.

NOTES & SOURCES

Notes

Not every furniture operation has this problem. An operator running a single distribution model, over predictable lanes, with a narrow SKU range whose handling requirements barely vary, gets adequate performance from stop-based planning and should not rebuild anything. The failure surfaces where several distribution models run at once, where service time and access vary widely by item, and where the delivery margin is thin enough that a wasted crew day is material.

Sign 2 cites European data to make a point about measurement, not about fleets. Eurostat's load factors are reported in tonnes, and that is precisely why they are quoted here: the category's official statistics are weight-based, so they cannot describe a trailer that is full by volume. The figures are not offered as a benchmark for a furniture fleet, and volumetric fill for furniture specifically is not published anywhere we could verify.

The fix is also imperfect. Consolidating four planning systems onto one decisioning layer, on top of item master data that is incomplete or inconsistent, reproduces the original problem at higher cost. Item-level planning is only as good as the SKU attributes feeding it, and populating those attributes is a data program before it is a software one. The shift to item-level, volume-aware, continuously re-decided planning is real. So is the work of getting the item data clean enough for it to mean anything.

Sources

Framing (market growth): Armstrong & Associates with the National Home Delivery Association, “Recalibrating: Big and Bulky Last-Mile Delivery in the United States, 2026,” July 2026 ↗. US third-party big-and-bulky last-mile market at $10.63 billion in 2025; 10.6% CAGR 2017 to 2025 against 5.1% projected 2025 to 2027. Scope: US third-party residential delivery of oversized goods, so it excludes in-house fleets and non-US operations.

Sign 1 (Service time belongs to the item): ATRI, “Costs and Consequences of Truck Driver Detention,” September 2024 ↗. Drivers detained at 39.3% of all stops, losing between 117 and 209 hours a year. Scope: ATRI measures truckload carriers, so it frames the wholesale multi-stop flow and does not describe consumer final mile. Service-time distribution by SKU is not published by any research body we could verify, so the figures in Fig. 1 are an illustrative model rather than data, and the argument is mechanical rather than statistical.

Sign 2 (The trailer cubes out): Eurostat, "Road freight transport by journey characteristics" ↗. 21.6% of total EU road freight vehicle-kilometers driven empty in 2024; 24% national against 13% international. Scope: EU road freight across all commodities, measured by weight. The 94% cube against 31% payload trailer is our arithmetic on a representative furniture load, not a Eurostat figure.

Sign 3 (Nobody drove the load): McKinsey & Company, "Digitizing mid- and last-mile logistics handovers to reduce waste" ↗. 13 to 19% of logistics costs attributable to inefficient interactions at handover points; up to $95 billion annually in the US economy; waste reducible by up to 40%. Scope: mid- and last-mile logistics handovers generally. Applying it to driver relay and segment-level attribution is our extension of the mechanism, not McKinsey's stated case.

Sign 4 (Four modes): Gartner, "Future of Supply Chains 2026" ↗. 95% of supply chains must react quickly to change; 7% can execute decisions in real time; investment in real-time decision execution expected to rise fivefold by 2028. Scope: cross-industry supply chain, not furniture-specific.

Sign 5 (The delivery that comes back): National Retail Federation with Happy Returns, "2025 Retail Returns Landscape" ↗. 15.8% of annual retail sales returned in 2025, totaling $849.9 billion; 19.3% of online sales. Scope: all US retail. Cited as the baseline furniture sits on, not as a furniture figure. No credible research organization publishes a furniture-specific refusal or damage rate; figures do circulate, but they originate in vendor marketing material and contradict each other, so Fig. 5 shows an illustrative composition rather than measured shares.

Sign 6 (The window is inventory): McKinsey & Company, "2026 Global B2B Pulse Survey," roughly 4,000 decision-makers across 13 countries ↗ · Capgemini Research Institute, "The Last-Mile Delivery Challenge" ↗. 51% cite absent cross-channel tracking as an impediment to doing business; 54% of likely switchers cite poor digital customer experience. 73% of consumers rate a convenient delivery time above a fast one; last-mile delivery at 41% of total supply chain cost. Scope caveat: the Capgemini consumer survey was fielded in late 2018, so treat the 73% as directional rather than current. The B2B Pulse figures are 2026.

What changes: Gartner press release, March 2026 ↗. 60% of supply chain disruptions predicted to be resolved without human intervention by 2031.