

The Annual Plan Reality Check

Most annual revenue plans miss their number before the year has started. Here's what the strongest plans had in common, and why the majority never had a chance.

RESEARCH BY THE ORBYTAL TEAM | April 2026

HEADLINE FINDING

The strongest plans in our dataset hit 91% of quota. The weakest hit 49%. You could see the 43-point gap on the day the plans were deployed — before a single rep had worked a single deal.

WHAT THAT COSTS

\$329,840 in lost attainment per quota-carrying rep per year. On a 30-rep mid-market sales org, that's roughly **\$9.9M** a weak plan leaves on the table compared to a strong one.

The plan itself is rarely on trial. When a team misses the number, the autopsy lands on familiar ground: weak coaching, soft pipeline, market headwinds, a handful of underperforming reps. The plan is treated as the baseline against which performance is measured, not as a variable that might explain it. Which is strange, because the plan is the one thing every rep is working against for twelve months. In this study, it was the variable that mattered most.

We examined 403 annual revenue plans built between 2022 and 2025, scored each one on four attributes measurable on the day the plan was deployed, and matched the scores against year-end attainment. The question was simple: **before the year even started, did the way the plan was built predict whether the company would hit its number?**

It did. The 43-point gap above is not a statistical artifact. It survives every cut of the data we tried. It persists across company size and industry. It holds whether we measure attainment by company average, rep median, or quartile distribution. The uncomfortable implication for

revenue leaders is that the gap was baked into the plans before the year began, rather than created by what happened afterward. The plans were the variable.

How we measured. The dataset comprises 403 annual revenue plans drawn from B2B SaaS companies with 30–300 quota-carrying reps, deployed between 2022 and 2025. One plan per company per year; no company appears twice in the same plan year. Each plan was scored on four dimensions at deploy day. **Territory Equity** measures how fairly opportunity is distributed across territories, calculated as the Gini coefficient of TAM-weighted opportunity per territory. **Coverage Adequacy** is the percentage of territories that met a 3x planned pipeline coverage threshold at deploy. **ICP Coherence** is the percentage of a plan's named accounts that scored above the fit threshold of the plan's stated ICP, scored against the definition active on the day the plan was deployed, not a later version. **Capacity Realism** combines two sub-measures: phantom capacity (the share of quota assigned to reps not yet hired) and unramped capacity (the share of quota assigned to reps below the 5.3-month ramp benchmark from Bridge Group 2024). The four driver scores combine into an equal-weighted composite, the Structural Soundness Score. Year-end attainment data was attached from the same dataset and scored relative to assigned plan quota. Tertile thresholds were computed from the dataset's own distribution rather than set externally. The research team can field specific methodology questions at info@orbytal.ai.

Year-End Attainment by Structural Soundness Decile

403 plans, sorted by deploy-day SSS, grouped into deciles. Each bar = average year-end attainment for that decile.

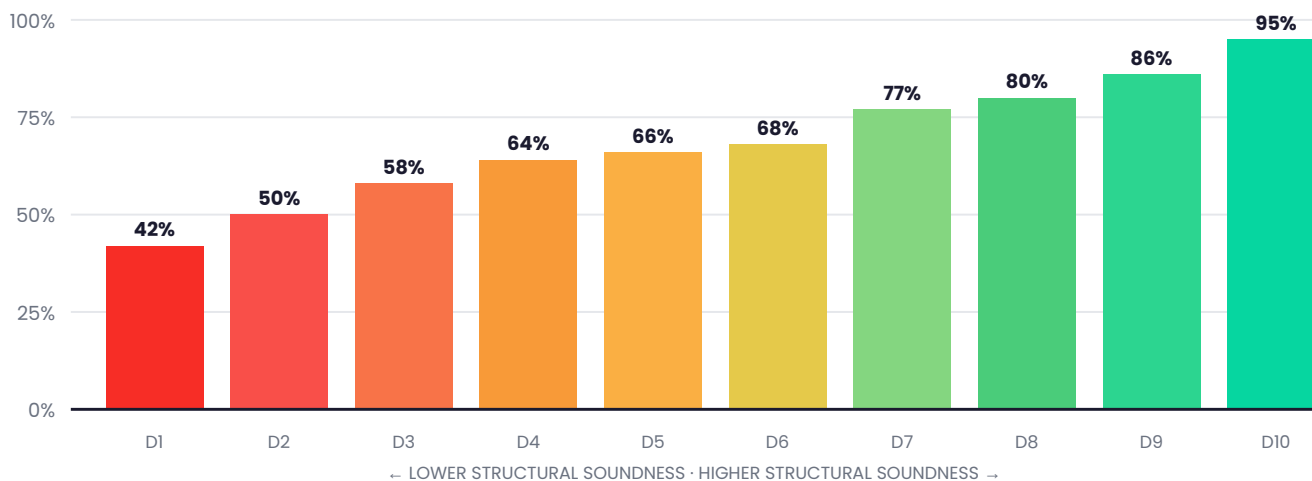


FIG 1 · N = 403 PLANS

THE FOUR DRIVERS

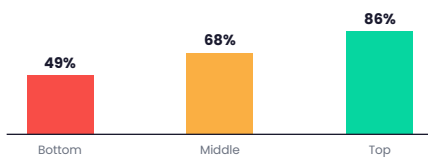
What the strongest plans had in common.

Each of the 403 plans was scored on four independent attributes at deploy day. Each attribute is measurable, defensible, and visible before the year begins. Each one independently predicted attainment. Together, they explain a substantial share of why some plans hit and others didn't. Ordered below by predictive strength, from the driver that explained the most variance to the least.

DRIVER 01 · ICP

ICP Coherence

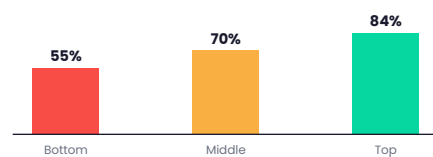
Percentage of named accounts in the plan that matched the deploy-day ICP definition. Top-tertile plans delivered **86%** attainment; bottom-tertile, **49%**. This is the widest spread of any driver, and the strongest individual predictor of the four.



DRIVER 02 · EQUITY

Territory Equity

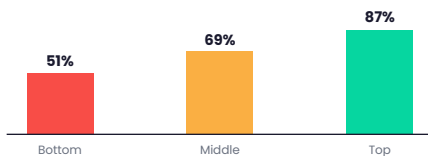
How fairly opportunity was distributed across territories at deploy, measured by the Gini coefficient of TAM-weighted opportunity. Top-tertile plans delivered **84%** attainment against **55%** in the bottom tertile, a 29-point gap.



DRIVER 03 · COVERAGE

Coverage Adequacy

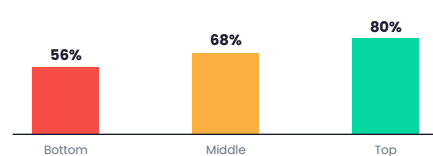
Percentage of territories where planned pipeline coverage met the 3x threshold at deploy. Top-tertile plans hit **87%** attainment; bottom-tertile plans hit **51%**. The coverage built in January predicted the execution seen in November.



DRIVER 04 · CAPACITY

Capacity Realism

A combined measure of phantom capacity (quota assigned to unhired reps) and unramped capacity (reps below the 5.3-month ramp benchmark). Top tertile: **80%**. Bottom: **56%**. The most recoverable of the four drivers mid-year, and still a 24-point gap when it goes unaddressed.



The four drivers are not redundant. Each captures a different failure mode. Equity captures fairness, coverage captures pipeline math, ICP captures targeting quality, and capacity captures the gap between the reps a plan assumes and the reps that actually exist. A plan can fail any one of them in isolation. Plans that fail multiple drivers simultaneously don't just underperform — they collapse.

THE COST · THE STRONGEST LEVER

What this is worth, in dollars and in priority.

A 43-point attainment gap is a research finding. Translated into dollars on a typical mid-market sales org, it becomes a budget conversation. And once we look at which drivers individually carried the most predictive weight, the priority of where to fix things first becomes specific.

THE COST OF A WEAK PLAN

\$329,840

in lost attainment per quota-carrying rep per year, measured against the average AE quota in the dataset of \$760,000. A 30-rep org running a weak plan leaves roughly \$9.9M on the table compared to one running a strong one. The number survives every reasonable assumption you can challenge.

Math: 43-point attainment gap × \$760,000 average AE quota = \$329,840 per rep.

Driver Strength Analysis

Each driver was regressed against year-end attainment individually to identify which attribute carried the strongest predictive power. The category has argued about this on intuition for years. The data has an answer.

DRIVER	STD. COEFFICIENT	R ²	RANK
★ ICP Coherence	0.76	0.59	1
Territory Equity	0.69	0.47	2
Coverage Adequacy	0.64	0.42	3
Capacity Realism	0.57	0.33	4

ICP Coherence is the strongest individual predictor of year-end attainment in the dataset. By itself, it explains 59% of the variance in attainment outcomes. Knowing nothing else about a plan except how well its named accounts matched its stated ICP would let you predict the outcome with substantial accuracy. Territory Equity comes second, Coverage third, and Capacity fourth.

The implication for revenue leaders is direct. The four drivers are not equally important, and the difference is large enough to matter. If your team has the bandwidth to audit one thing about your plan before it deploys this year, audit ICP coherence first. The other three drivers are still worth measuring, and as we'll see below, the cost of ignoring any of them compounds quickly. But ICP is where the largest single share of attainment variance lives.

THE PATTERN

When plans fail, they don't fail one driver at a time.

The four drivers, taken in isolation, each predict a meaningful share of year-end attainment. Taken together, they reveal something more important: the things that go wrong in revenue planning rarely arrive alone, and the attainment cost compounds in a way that no individual driver anticipates by itself.

We grouped the 403 plans by a single number: how many of the four drivers landed in the bottom tertile of the dataset. A plan with zero drivers in the bottom tertile means a plan that was at least middling on equity, coverage, ICP, and capacity. A plan with all four in the bottom tertile means a plan that failed every test simultaneously. The result, plotted below, is the most consequential finding in the entire study.

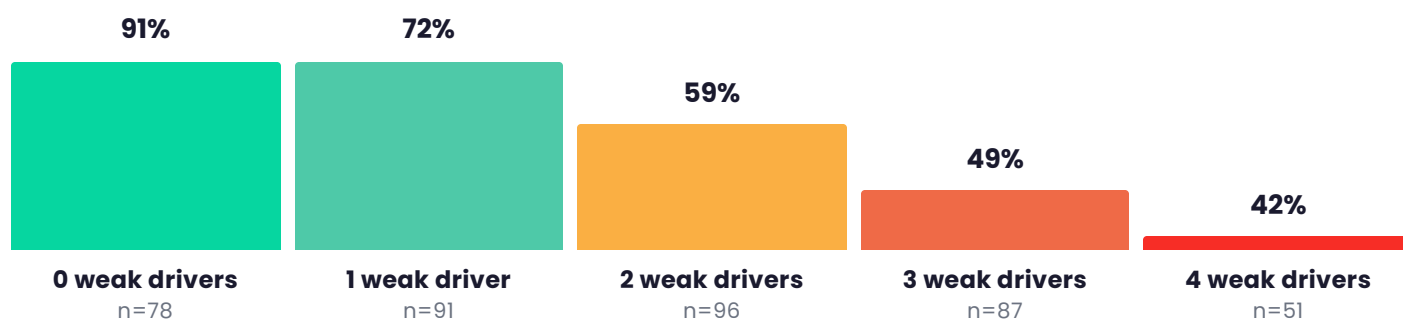


FIG 6 · MEAN ATTAINMENT BY # OF DRIVERS IN BOTTOM TERTILE · N = 403 · Y-AXIS BASELINE AT 40%

The pattern is monotonic and severe. Plans with no drivers in the bottom tertile averaged 91% attainment, a healthy benchmark for a well-run sales org. Plans weak on all four drivers cratered to 42%, a 49-point collapse that is larger than the top-versus-bottom-quartile gap shown above. Taken together, the four drivers predict more of the story than any one of them does alone. They should be treated as a single audit checklist, not as independent diagnostics.

THE FINDING BEHIND THE FINDING

60% of the plans in the dataset deployed with two or more drivers already in the bottom tertile. On the day they launched, the majority of revenue plans were already on the wrong side of the attainment cliff.

The distribution matters as much as the curve. Of the 403 plans, only 78 (19%) deployed with all four drivers above the bottom tertile threshold. The largest single bucket was plans with exactly two drivers already in trouble: 96 plans, nearly a quarter of the dataset. Another 87 plans deployed with three weak drivers, and 51 with all four. The conclusion is uncomfortable but clean: the typical plan in this dataset was already set up to miss its number before the year began, not because of any single structural failure, but because multiple failures had compounded at the moment of deployment.

A revenue leader looking at this curve has a clear operational decision in front of them. Plans with two or more drivers in the bottom tertile are not "slightly worse" plans — they are plans that, in this dataset, lived on the wrong side of an attainment cliff. The threshold at which failure becomes unrecoverable looks closer to two simultaneous failures than to four. Whether that cliff is at exactly two or somewhere between two and three is a refinement question for future research. The qualitative finding, that failures compound rather than add, is robust across every cut of the data we tried.

What this study does not prove. The relationships above are correlational, not causal. We cannot claim that fixing these attributes would have closed the attainment gap, only that they were predictive of it. Market conditions, sales execution skill, product quality, competitive dynamics, and macroeconomic effects are all outside the scope of this analysis. Readers should weigh these limitations when generalizing the findings to their own organization.

WHAT THIS MEANS

Revenue plans are failing earlier than anyone is looking.

The category has spent a decade improving how revenue teams execute against their plans. The data from 403 real plans suggests the larger problem may be how those plans are built — and that the difference between hitting and missing the number is, to a meaningful degree, baked in before the first deal of the year is worked.

If a 43-point attainment gap is sitting in the way plans are built, then the highest-leverage move available to most revenue leaders is not better forecasting, better coaching, or better tooling at the rep level. It is a more honest audit of the plan itself, before deployment, against the four drivers identified above. The audit takes hours. The downstream cost of skipping it, in this dataset, was measured in hundreds of thousands of dollars per rep.

The Self-Audit

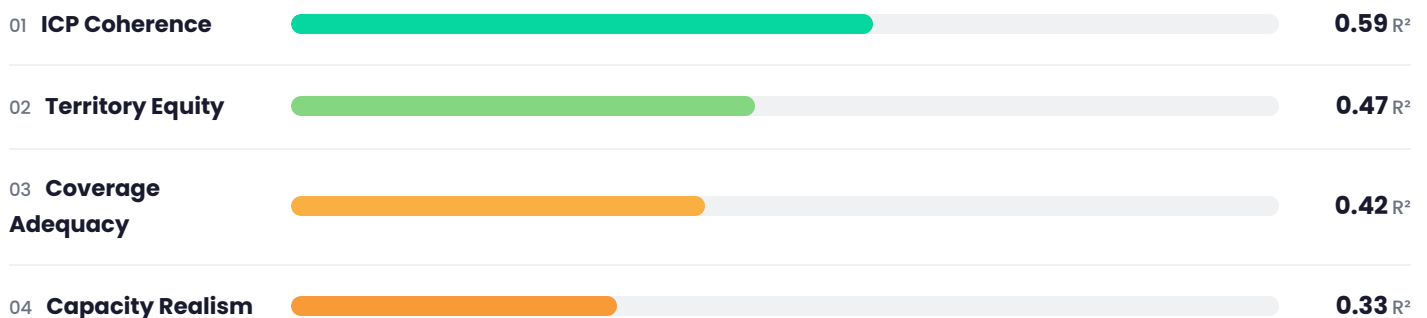
Five questions to ask of your current plan, today.

- 01 Equity:** If you calculated the Gini coefficient on TAM-weighted opportunity across your territories, would it look balanced? Or do your top reps quietly own 40% of the opportunity?
- 02 Coverage:** What percentage of your territories deployed with at least 3x planned pipeline coverage? If the answer is below 70%, your plan is mathematically depending on conversion rate increases that have not been earned.
- 03 ICP fit:** Of the named accounts your reps are working today, what percentage actually score above your ICP threshold? Most teams have never asked this question against their own definition.
- 04 Phantom capacity:** How much of your annual quota number depends on reps you have not yet hired? On reps with less than five months of tenure?
- 05 The compounding test:** Of the four questions above, on how many do you score in the bottom third? Two or more is the threshold at which the data shows attainment falling off a cliff.

The discipline of asking these questions before the plan goes live – rather than after the quarter slips – is a small operational change with a substantial measured payoff. The plans in our dataset that scored in the top quartile on the audit were not built by larger teams or more expensive tools. They were built by revenue leaders who paid attention to the four things that, in retrospect, mattered most.

THE FOUR DRIVERS · RANKED

Share of attainment variance explained by each driver, individually.



TAKE THE REALITY CHECK YOURSELF

Score your plan against the four drivers in five minutes, or book a deeper audit with the Orbytal team.

Score My Plan → (orbytal.ai/plan-reality-check)

OR

Book a 30-Minute Audit (orbytal.ai/contact)

The self-serve scorecard takes about five minutes and gives you a tier rating across all four drivers. The 30-minute audit is a working session with the Orbytal team using your actual territory data. No pitch.

CITE THIS REPORT

Orbytal Research (2026). *The Annual Plan Reality Check: What 400+ Revenue Plans Reveal About Why Most Miss*. Salt Lake City: Orbytal.

orbytal.ai/research/annual-plan-reality-check

Published by **Orbytal**, the continuous revenue planning platform for RevOps and CRO teams.