

The quiet churn

A B2B REVENUE PLAYBOOK

Recurring-revenue accounts rarely cancel loudly. They fade. This white paper sets out how to detect the drift early, rank accounts by revenue at risk, and give the people who can save them the reason and the next step.

Executive summary

In subscription software, churn is an event: someone cancels. In most B2B businesses it is a process. Orders get smaller, service calls stop, the person who championed you changes jobs, a competitor quotes ahead of renewal. By the time the loss shows in revenue, the decision was made months earlier.

That gap is the opportunity. A business that can see the drift while the customer is still a customer has time to act, and the accounts worth acting on are usually a short list. In B2B, a handful of large accounts carry more revenue than the long tail of small cancellations.

This paper sets out a method we use, in seven parts:

1. Which customers suit churn prediction at all, and how to test that on your own data.
2. Which signals to collect, grouped into five families.
3. How to build a score, from simple rules to a model, and what to demand of it.
4. Why to rank accounts by revenue at risk instead of by probability alone.
5. How to route each at-risk account to a person with the reason and a suggested next step.
6. How to prove the whole thing works with a control group, before anything scales.
7. What the first ninety days look like.

The core idea is short. Score every account continuously on its behavior. Rank by money at risk. Hand a person the reason and the next move. Measure against accounts you left alone.

1. Why B2B churn is a process

Retention is worth the effort. Harvard Business Review, citing Fred Reichheld of Bain & Company, reports that increasing customer retention by 5% increases profits by 25% to 95%, and that acquiring a customer can cost anywhere from five to 25 times as much as keeping one, depending on the study and the industry. ([HBR](#))

Those figures describe consumer and business markets alike. The B2B version has two features that change how you work.

First, the loss is slow. A subscription customer clicks cancel. A building owner under a service contract does something quieter. The owner calls a second vendor for a quote, pays invoices a little later each month, books fewer service visits, and lets the renewal date approach without a conversation. Nobody in the business announces the decision. The contract simply doesn't renew.

Second, the loss is lumpy. Ten small accounts leaving is a bad quarter. One large account leaving can be a worse one, and it can be the same size as the ten. A churn program that treats every account alike will spend most of its energy on the accounts that matter least.

Both features point the same way. You need an early warning, a ranking by value, and an intervention that fits the account. A blanket retention discount fails all three tests. It arrives late, it reaches everyone, and it hands margin to accounts that were never leaving.

2. The rhythm test

The method works on customers who come back on a rhythm. Before building anything, find out whether yours do.

A rhythm exists when a customer buys, orders or uses the service at a roughly regular cadence. Service contracts have one. So do consumables, replenishment, leasing, and uniform and linen supply. A hospital's HVAC maintenance, a data center's cooling service, a building's elevator inspections, and a restaurant chain's weekly produce order all tick like a clock.

A rhythm matters because it gives you a baseline. If an account has ordered every thirty days for three years, then day forty-five with no order is information. A customer with no rhythm gives you nothing to measure a break against.

One-time capital purchases usually fail the test. A company that buys a production line every nine years has no cadence to bend, and a model fitted to those accounts produces noise.

How to run the test

Pull three years of orders, service tickets or usage by account. For each account, list the gaps between events. Then ask two questions.

- Is the typical gap stable? As a first cut, an account whose gaps vary by less than about half of their average has a rhythm worth modeling.
- Do enough accounts pass? If most of your revenue sits in accounts with a rhythm, you have a program. If only a small corner does, start there and leave the rest alone.

The result is a map of which part of your book can be scored at all. Many companies find that the part they can score is also the part where the most revenue sits.

When the rhythm is in the contract

Some businesses have the rhythm written into the agreement: a service visit every quarter, a renewal every three years. That helps. The contract calendar becomes a second clock, and the signals that matter most are the ones in the months before the renewal date.

3. Signals

A useful signal changes before the customer leaves and is cheap to collect. Five families cover most of what works.

Volume

How much the customer buys or uses, against their own history.

- Order frequency and order size, compared with the account's trailing average.
- Time since the last order, compared with that account's usual gap.
- Usage or consumption, where you meter it.
- Share of wallet, where you can estimate it: the portion of their spend in your category that comes to you.

Volume is the earliest signal in most books because shrinking comes before stopping. An account that goes from twelve orders a year to nine has told you something, even if it has never complained.

Relationship

How connected your people are to theirs.

- Touchpoints: calls, visits, meetings and emails, and whether they are tailing off.
- Champion change: the person who bought from you left, moved, or stopped replying. This is often the single most predictive event in a contract business, and almost no one tracks it.
- Breadth: how many people at the customer you deal with. One contact is a fragile relationship.

Sentiment

What the customer says and how they say it.

- Complaint count and severity.
- Support tone, read from tickets and emails with a language model and checked by a person.
- Escalations, and how long they took to resolve.

Language models are useful here because tone lives in free text. A ticket that says "second time this month" tells you more than the ticket category does.

Commercial

How the customer behaves with money.

- Days to pay, and whether that is drifting upward.
- Discount requests and renegotiation of terms.

- Disputed invoices.

Slow payment is a classic early tell in service contracts. A customer who is drifting often stops prioritizing your invoices before they stop buying.

Competitive

Signs that someone else is in the conversation.

- Requests for quotes or for competitor price matching.
- Requests to shorten the term, or to remove a renewal clause.
- Where you can see it, a competitor's presence at the customer's site.

The renewal date sits across all five families. Distance to renewal changes how every other signal should be read, because the same weak signal means more at ninety days out than at eighteen months.

Where the data lives

Most of these signals already exist. Orders sit in the ERP. Tickets sit in the service desk. Meetings and emails sit in the CRM. Payment behavior sits in billing. The work is joining them by account and keeping the join current, which is mostly plumbing and rarely glamorous.

4. Building the score

A score is a number per account, refreshed often, that says how likely the account is to leave within a defined window. The window is usually one renewal cycle.

Start with rules

A first version doesn't need a model. A short set of rules built from the rhythm test gets you most of the way and gives the account managers something to react to.

- The account is late by more than an agreed multiple of its usual gap.
- Order volume is down by more than an agreed share against its own trailing average.
- The champion has changed.
- A competitor quote has been requested.

Rules are easy to explain, which matters when you are asking people to change how they spend their week. They also give you a baseline. If a model can't beat the rules, it hasn't earned its complexity.

Then a model

Once the rules are running and the data join is trustworthy, a model can weigh the signals against each other. Two families are common.

- A classification model predicts whether an account leaves within the window. It is simple to train and easy to explain.
- A survival model predicts when. It handles accounts that have not yet reached a renewal, and it uses information a classifier discards. It suits long contracts and short histories well.

Gradient-boosted trees are a sound default for either. A neural network adds cost without adding much on this kind of tabular data.

What to demand of the score

- Explainability. Every score ships with its top reasons in plain words: "order gap is three times normal, champion changed, payment slowed." An account manager will not call a customer because a black box said so.
- Honest windows. Train only on what was knowable at the time. A common failure is to include information that appeared after the decision, which makes the model look brilliant in testing and useless in life.
- Calibration. If the model says 40%, about four in ten such accounts should leave. A score that is only a ranking can still be useful, but money calculations need real probabilities.
- Refresh. Rhythm changes with seasons and with customer circumstances. The score should update as new orders land, not once a quarter.

Few accounts, long histories

Many B2B books have a few hundred accounts, not millions of customers. That is workable. Fewer accounts with long histories can produce good signal, because each account gives you many events. Start with rules and a simple model. Be honest about uncertainty, and treat the score as a prompt for a conversation, not a verdict.

5. Rank by revenue at risk

A score that ranks accounts by probability alone sends effort to the wrong place. A small account with a high probability of leaving matters less than a large account with a moderate one.

The fix is arithmetic. Multiply the probability of leaving by the annual value of the account. The result is revenue at risk, and the work list reorders itself.

A worked example

Take six accounts.

ACCOUNT	ANNUAL VALUE	CHANCE OF LEAVING	REVENUE AT RISK
A	\$180,000	35%	\$63,000
B	\$40,000	80%	\$32,000
C	\$95,000	55%	\$52,250
D	\$22,000	90%	\$19,800
E	\$260,000	12%	\$31,200
F	\$60,000	40%	\$24,000

Rank by probability, and the order is D, B, C, F, A, E. Rank by revenue at risk, and the order is A, C, B, E, F, D.

Suppose an account manager has time for three calls this week. Calling the three most likely to leave covers D, B and C, with \$104,050 of value at risk between them. Calling the top three by revenue at risk covers A, C and B, with \$147,250 at risk. Same three calls, \$43,200 more value in the conversation.

Account D is the instructive one. It is the most likely to leave and the least worth a senior person's week. It can get a lighter touch: a templated check-in, a service credit, a note from support.

Value can mean more than revenue

Annual contract value is the simplest measure. Margin is better where you can get it, because a large low-margin account can be less worth saving than a mid-sized profitable one. Strategic value counts too: a reference customer, an account that opens a new region, one that brings adjacent business. Keep the formula plain and let people add judgment on top of it.

6. Route to a person

The last step is where most programs fail. A score that sits in a dashboard changes nothing. A score that arrives as an actionable brief, in the tool the account manager already uses, changes the week.

Don't automate the save

The right intervention for an at-risk B2B account is a person with context. A discount sent automatically teaches customers to hint at leaving, costs margin on accounts that were never going anywhere, and misses the actual reason. The goal of the system is to make the account manager better prepared, not to replace the call.

The account brief

For each account on the list, the brief carries four things.

- The score and the revenue at risk.
- The top three reasons, in plain words.
- A suggested next step, chosen from a short menu of plays.
- An owner and a date.

It lands inside the CRM, as a task on the account, because that is where the account manager already works. A separate dashboard is one more place to forget to look.

Generative AI is useful for part of this. A language model can draft the briefing from the signals, summarize the last six months of tickets and emails, and propose the wording of an outreach. It does not decide who to call or what to offer. That stays with the score and the person.

A short menu of plays

The plays should match the reasons, not just the risk level.

- Order rhythm broke: a service review, with usage data in hand, to find out what changed.
- Champion left: an executive introduction to the new contact, fast.
- Slow payment and discount requests: a commercial conversation about terms, before price.
- Competitor quote: a prepared comparison and an early renewal offer with a longer term.
- Sentiment dropped: a direct call from someone senior, with the specific issue named.

Discounts come last on the menu, and they come from a person who has seen the reason.

Expansion uses the same signals

The signals that predict loss also predict growth. An account whose orders are rising, whose contacts are widening and whose payments are prompt is ready to hear about more. Run the same system in both directions and the account manager gets one list of conversations worth having.

7. Prove it

A churn program is easy to declare a success. Accounts that were flagged and called mostly renewed. But many of them would have renewed anyway, and a dashboard that counts only flagged accounts will show a win every time.

The only way to know what your program caused is to compare it with accounts you left alone.

Hold out some accounts

Pick a random slice of the at-risk list and don't run the plays on it. Treat everything else the same. At the end of the renewal cycle, compare the two groups on what matters: renewals and revenue retained.

Randomize within value tiers so that one group does not end up with all the large accounts. With a small book, expect the control group to be small too, which means only large effects will be visible. That is fine. A program that can't show a large effect has not earned the right to scale.

Agree the stopping result first

Before the first call, write down what result would make you stop. Without it, ambiguous results get read kindly after launch.

Measure the cost of false alarms

An account manager's time is the scarce resource. Count the calls made, the hours spent and the share of flagged accounts that turned out to be fine. A score that cries wolf teaches people to ignore it.

Keep a small permanent holdout

Customers change, catalogs change and models drift. The gap you measured in the first cycle may not hold in the fifth. A small standing holdout keeps the number honest.

8. Data, systems and governance

The data you need already lives in your systems. The work is bringing it together.

- Sources: ERP for orders and invoices, CRM for contacts and activity, the service desk for tickets, billing for payment behavior.
- Join key: a clean account identifier across systems. This is the most common blocker, and it is worth fixing first.
- History: about three years of orders or tickets per account, plus renewal dates, gives the model enough to learn the rhythm.
- Delivery: scores and briefs written back into the CRM. In a Salesforce shop, that can sit alongside the account-health fields and Agentforce.
- Privacy: signals come from your own commercial relationships and your own systems. Customer emails and tickets contain personal data, so treat them under your privacy policy and your regional law, and keep access narrow.

9. The first ninety days

A first ranked list can reach account managers in six to eight weeks. A plan that gets there looks like this.

Weeks 1 to 2: data and the rhythm test

Pull three years of orders, tickets and renewal dates. Fix the account identifier. Run the rhythm test, and decide which part of the book is in scope.

Weeks 3 to 6: rules, then a first model

Stand up the rules and check them against past churn: would they have flagged the accounts you lost, and how far ahead? Build a first model and compare it with the rules. Rank by revenue at risk.

Weeks 7 to 8: the first list

Deliver the ranked list to account managers as briefs in the CRM, with the reasons and a menu of plays. Agree the holdout and the stopping result.

Weeks 9 to 12: run, measure, adjust

Run the plays. Log what was done and what happened. Review the false alarms. At the end of the first renewal window, compare treated and held-out accounts.

10. Common pitfalls

- Scoring customers with no rhythm. The result is noise.
- Training on information that wasn't known at the time. The model looks great in testing and fails in production.
- Ranking by probability alone. The effort goes to the wrong accounts.
- Sending the score to a dashboard. Nobody looks at it.
- Automating the save. The discount costs margin and misses the reason.
- Skipping the holdout. You can't tell whether the program works.
- Treating the score as a verdict. It is a prompt for a conversation.

Questions to take to your next review

- What share of our revenue sits in accounts with a rhythm?
- Who owns the account identifier across ERP, CRM and billing?

- Which five accounts would we most hate to lose, and what do the signals say about them today?
- What would make us stop this program, and have we written it down?
- How many of the last ten accounts we lost showed a warning at least a quarter ahead?

If you want a second pair of eyes on any of these, the [B2B contract churn field guide](#) covers the method in short, and we are glad to talk it through.