From Datasheet to ATE: A Systematic Method for Power-IC Test Program Development

A SiliconEng white paper · October 2026 · Version 2.7

1. The problem

Every analog test engineer knows this project. A new power-management part tapes out, the datasheet runs forty pages, and someone has six weeks to turn it into a production test program: choose the tests, pick the conditions, set the limits, write the code, correlate it against the bench, and defend every decision in the first production meeting.

Six weeks later, the program ships — and the arguments begin. Yield falls out on a corner nobody tested. The datasheet's "typical" PSRR curve doesn't match what the ATE measures at 1 MHz. The current-limit test fails parts that the bench says are fine, or passes parts the customer sends back. Each of these is, at root, the same failure: a judgment call made somewhere between the datasheet and the tester, never written down, now impossible to revisit.

The cost of getting this wrong is not measured in test time. It is measured in test escapes, in qualification failures during AEC-Q100, in the re-spin that a test hook could have caught — six months and a mask set. The industry treats test program development as craft: experienced engineers, tribal knowledge, programs that work because the person who wrote them remembers why. Craft is valuable. Craft that isn't written down doesn't scale, doesn't transfer, and doesn't survive the engineer leaving.

This paper describes a systematic method for developing power-IC test programs — specifically, for turning a datasheet into a production ATE program with defensible limits and a written record of every decision. It is the method we use. It is not a tool pitch: there is no product for sale here. It is a discipline, and it can be adopted with nothing more than a spreadsheet and the willingness to write things down. A companion paper, *Analog and Power IC Test Development: A Field Guide to Pitfalls and Good Practice* (referred to below as Paper 2), covers the measurement traps this method runs into in practice.

2. Why datasheets make bad test plans

Datasheets are not test plans. They are marketing documents with a specifications table attached. Understanding exactly how they fall short is the first step of the method, because each shortfall is a decision the test engineer will have to make — and the method's job is to make sure those decisions are explicit.

Typical values promise nothing. A line-regulation spec of 0.03 %/V "typical" is a characterization of the design, not a guarantee about the part in the tray. Production limits can only be set from guaranteed min/max values, or from characterization data the test house collects itself. Every "typ-only" parameter in the datasheet is a question, not an answer: will you screen it, sample it, or trust it? The datasheet will not tell you.

Test conditions hide in footnotes. Consider a dropout-voltage specification: 140 mV maximum, at 300 mA load, for the 2.5-to-5.5 V output variants (DQN, YCJ, and YCK packages; 145 mV for DBV). That single row encodes at least four conditions — load current, output voltage range, the definition of dropout (output at 95% of nominal), and the temperature range from the EC-table header (TJ = −40 to +125 °C, typicals at 25 °C). Change any one and the number changes. A test plan that copies "140 mV" without reproducing the conditions is testing something the datasheet never specified.

Corners are never shown together. Datasheets specify parameters at convenient points: PSRR at 100 Hz and 20 mA, quiescent current at no load, dropout at full load. Production failures live in the combinations — full load *and* high temperature, dropout *and* line transient. The datasheet's condition matrix is sparse by design; the test plan's condition matrix has to be denser, and choosing where to add points is engineering judgment that must be recorded.

Some parameters are specified at conditions production cannot reproduce economically. A 1-MHz PSRR measurement with a clean stimulus and a settled thermal state is a bench exercise. On an ATE with a 50-ms insertion budget, it is a different test with different error sources. The method does not pretend these are the same measurement — it forces an explicit decision about what production actually screens versus what characterization already proved.

Datasheets omit parameters production still has to screen. Short-circuit behavior, enable-pin hysteresis, power-good deglitch timing — the rows that exist in the silicon but not in the EC table. A test plan derived only from the

specifications table will have holes, and the method's extraction step is designed to find them by reading the whole document: typical curves, application notes in the datasheet body, and the absolute-maximum table, which constrains what the test itself is allowed to do to the part.

3. The method: extraction → normalization → test plan → limits

The method has four stages. Each stage produces a written artifact. The artifacts are the point: they are what gets reviewed, what gets argued about before production instead of during it, and what lets the next engineer understand why the program looks the way it does.

3.1 Extraction: every testable claim becomes a row

Read the entire datasheet — EC table, typical curves, application section, absolute maximums, footnotes — and transcribe every testable claim into a uniform row: parameter name, symbol, conditions, min/typ/max, unit, and the source (table, figure, footnote, page). One row per claim. A curve with five frequency points is five rows, not one. A footnote that changes a condition creates a new row.

Two rules govern extraction. First, **transcribe, don't interpret**: if the datasheet says the dropout condition is "VOUT = 95% × VOUT(NOM)," the row says exactly that, even if you know what it means. Interpretation belongs in the next stage, and keeping the stages separate is what makes the result auditable. Second, **flag every missing condition explicitly**. If a row is specified at 25 °C only, with no limit at temperature — the worked example's shutdown current is exactly such a row — the row records "hot limit: unspecified" rather than silently assuming the 25 °C value holds. The flags from this stage become the agenda for the normalization stage.

For a typical 40-page LDO datasheet, extraction produces on the order of a hundred rows or more once variants, footnotes, and curve points are counted — the TPS7A20's PSRR curves alone contribute ten rows (five frequencies × two loads). That number surprises people exactly once.

3.2 Normalization: one canonical representation

The extracted rows will contradict each other. The EC table says one thing, a typical curve implies another, and a footnote narrows the conditions of both. Normalization resolves every row into a single canonical parameter set: one name, one definition, one condition set per parameter, with contradictions resolved and the resolution recorded.

This is where the missing-condition flags get decided. "Temperature: unspecified" becomes either "assume TA = 25 °C, justify in the notes" or "requires characterization before production limits are set" — a written decision either way. Parameters specified only as typical get classified: screen in production (with limits from characterization), sample at probe, or accept by design. Parameters with no datasheet row at all — the short-circuit current the curves imply, the enable hysteresis the application section mentions — get added as explicit rows marked "not in EC table," so the decision to test or not test them is visible.

The output of normalization is a device specification: the complete, contradiction-free description of what the part is claimed to do. It is the document the test plan is written against, and it is reviewable by someone who has never seen the datasheet.

3.3 Test plan: every row maps to a test — or a written reason not to test

Each normalized parameter maps to exactly one of five dispositions:

1. **Test in production**, with specified force/measure conditions and a sweep or static definition.

2. **Characterize only** — measured during development and qualification, sampled or skipped in production, with the sampling rationale recorded.

3. **Not tested**, with the reason written down: "covered by design," "destructively tested at qual," "no production screen correlates."

4. **Tested via correlated proxy** — the datasheet parameter is covered by a production-screenable measurement proven to track it (bench characterization plus a documented correlation), rather than by a direct measurement the tester cannot make well. See Paper 2 §3.6.

5. **Tested, measurement-limited (guardbanded)** — tested in production with a guardband set by the full measurement uncertainty (U as defined in §5), where the TUR is below target. See Paper 2 §3.7. The worked example's output-accuracy test (§4.2, §5) is one: TUR ≈ 5:1 under the half-width convention defined in §5.

The 'not tested' category is the one that matters most. Every experienced test engineer has a mental list of things they don't test and why. The method requires that list to exist on paper, because the engineer who inherits the program — or the customer who audits it — cannot read minds.

Each production test specifies its full condition set: forced voltages and currents, load, temperature, timing, and the measurement definition. "Measure dropout" is not a test definition. "Force IOUT = 300 mA; sweep VIN from 3.6 V down in 10 mV steps; dropout is VIN − VOUT at the point where VOUT falls to 95% of nominal; datasheet limit 140 mV maximum" is a test definition. (The production limit then applies the §5 guardband.) The difference between those two sentences is where yield arguments come from.

3.4 Limits: guardbanded, documented, defensible

Production limits are set from guaranteed datasheet values where they exist, adjusted by a guardband stack (Section 5), and from characterization data where they don't. Every limit carries its provenance: which datasheet row it came from, what guardband was applied and why, and who approved it. When the yield meeting asks "why is this limit 3.25 V and not 3.24 V," the answer is a document, not a memory.

4. Worked example: a 300-mA low-noise LDO

To make the method concrete, we walk it through a representative part: a 300-mA, low-noise LDO with a 3.3 V output, enable pin, and no power-good output — the TPS7A20, DQN package, using values from TI datasheet SBVS338H (July 2024). The numbers below are real datasheet values from that revision; the test-plan tables are the output the method produces from them. To re-verify any value below, check it against that datasheet revision — later revisions may change limits.

4.1 Extraction highlights

The EC table yields the expected rows — output accuracy, line and load regulation, quiescent and shutdown current, dropout, current limit, PSRR — but the interesting findings are in the flags:

- **Output tolerance splits by output voltage.** For VOUT ≥ 1.85 V the spec is ±1.5% (DQN, YCJ, and YCK packages); below 1.85 V it is ±30 mV absolute. The DBV package splits differently: ±1.5% at or above 2.8 V, ±40 mV below. One datasheet row, two test conditions — extraction records both, because a 0.8 V variant tested against ±1.5% would get limits of ±12 mV, far tighter than the ±30 mV the part is actually specified to.
- **Dropout is specified per output-voltage band**, not as a single number (DQN, YCJ, and YCK packages): 140 mV max for 2.5–5.5 V variants at 300 mA (145 mV for DBV), 200 mV for 1.5–2.5 V (205 mV for DBV), 355 mV for 1.2–1.5 V, 490 mV for 1.0–1.2 V, and 690 mV for 0.8–1.0 V. The condition

"VOUT = 95% × VOUT(NOM)" defines the measurement, and the EC-table header specifies TJ = −40 to +125 °C (typicals at 25 °C) — temperature here is specified, not missing. Note the footnote: the three rows below 1.5 V are "design simulation data only," not guaranteed limits — a textbook case for this paper's typical-versus-guaranteed argument.

- **Current limit has a split condition**: the test forces VOUT to 0.9 × nominal, with VIN = VOUT(NOM) + 0.3 V for VOUT ≥ 1.5 V and VIN = VOUT(NOM) + 0.5 V for VOUT < 1.5 V (the DBV package instead forces VOUT to VOUT(NOM) − 150 mV for VOUT < 1.5 V). The max is package-dependent: 770 mA for YCJ/YCK, 730 mA for DQN (360 mA min, 520 mA typ in all cases). Miss the footnote and the test is wrong for the low-voltage variants and for the DQN package.
- **PSRR is typical-only at every point**, specified at two loads with VIN = VOUT + 1.0 V: at 20 mA, 95/95/75/75/45 dB at 100 Hz/1 kHz/10 kHz/100 kHz/1 MHz; at 300 mA, 65/92/75/60/40 dB. Note the 300 mA curve is genuinely non-monotonic — it peaks at 1 kHz — which is datasheet behavior, not a transcription error. No min/max anywhere. Extraction therefore produces ten rows from the PSRR curves, not five — and every one of them is a characterize-don't-screen decision waiting to be made.
- **Short-circuit current (160 mA typical, VOUT = 0 V) has no min/max.** Another characterization candidate — and a reminder that the EC table alone would leave the part's fault behavior entirely unscreened.
- **Ground current is specified by temperature**: 8.5 µA max at 25 °C, 10 µA max to 85 °C, 15 µA max to 125 °C (VEN = VIN = 6 V, IOUT = 0 mA). A test limit for this parameter must carry its temperature condition with it.
- **Shutdown current is specified at 25 °C only** (0.2 µA max, VEN = 0 V, VIN = 6 V) — no hot limit exists. This is the worked example's genuine missing-condition flag: a production test at temperature cannot be justified from this row alone.

4.2 From rows to tests

Normalization resolves these into the test plan; a selection follows. Limits are for the DQN package at 3.3 V output; production limits apply the §5 guardband method with illustrative terms (replace with measured terms for the actual test). The output-accuracy and current-limit guardbands are shown symmetric (±10 mV, ±10 mA) for illustration only. Per §5, build each side from its own stack. YCJ/YCK differ in current-limit max (770 mA); DBV differs in dropout (145 mV), tolerance split (2.8 V / ±40 mV), and the <1.5 V current-limit forcing point:

Test	Package	Force conditions	Measure	Datasheet limit	Production limit	Disposition rationale
Output accuracy	DQN	VIN = 3.6 V and 6.0 V; IOUT = 1 mA, 150 mA, and 300 mA	VOUT	3.2505–3.3495 V (±1.5%)	3.2605–3.3395 V (10 mV guardband each side)	Datasheet min/max, guardbanded per §5 (TUR ≈ 5:1, half-width convention)
Dropout voltage	DQN	IOUT = 300 mA; sweep VIN 3.6 → 3.0 V in 10 mV steps	VIN − VOUT at VOUT = 95% × 3.3 V	≤ 140 mV (max)	≤ 125 mV (140 − 15 mV stack: 10 mV + 5 mV sweep quantization, §5)	Datasheet max, guardbanded per §5
Current limit	DQN (YCJ/YCK: 770 mA max)	VIN = 3.6 V; force VOUT = 2.97 V (0.9 × nominal)	ILIMIT	360–730 mA	370–720 mA (10 mA guardband each side)	Datasheet min/max, guardbanded per §5; footnote condition reproduced exactly
Ground current	DQN	VIN = 6.0 V; IOUT = 0 mA; EN = VIN	IGND	≤ 15 µA (−40 to +125 °C)	≤ 14 µA (1 µA guardband)	Datasheet max at full temperature range, guardbanded per §5
Shutdown current	DQN	VIN = 6.0 V; EN = 0 V	IIN	≤ 0.2 µA at 25 °C only	≤ 0.18 µA (0.02 µA guardband)	Datasheet max at 25 °C only, guardbanded per §5 (TUR ≈ 10:1); no hot limit exists — the gap is the missing hot limit, not the measurement

Test	Package	Force conditions	Measure	Datasheet limit	Production limit	Disposition rationale
PSRR	—	—	—	—	—	**Characterize only.** Typical-only spec; no production limit is defensible. Sampled at bench during NPI; production relies on design. Reason recorded.
Short-circuit current	—	—	—	160 mA typ (no min/max)	—	**Characterize only.** Fault behavior characterized at qualification; production screens it indirectly through the current-limit test. Reason recorded.

Note what the table does: every row from extraction appears, including the ones that become "characterize only." The PSRR and short-circuit rows are not holes in the plan — they are decisions, with reasons, that a reviewer can challenge before production instead of discovering during a yield excursion.

4.3 The test the datasheet almost got wrong

The current-limit test is the example we show every new engineer. The datasheet's ICL rows read simply enough: 360 mA min, 520 mA typ, and a max that depends on the package — 770 mA for YCJ/YCK, 730 mA for DQN — with VOUT forced to 0.9 × VOUT(NOM). But the footnote changes the input condition for low-voltage variants: VIN = VOUT(NOM) + 0.5 V instead of + 0.3 V, and the DBV package forces VOUT to VOUT(NOM) − 150 mV instead of 0.9 × nominal. A test program written from the 3.3 V part's condition alone — VIN = 3.6 V for all variants — would apply the wrong stimulus to every VOUT < 1.5 V part and judge DQN parts against the wrong limit. Extraction caught it because extraction transcribes footnotes as rows. This is not a hypothetical: it is a recurring class of

error we find when reviewing inherited test programs, and it is invisible until you look for it.

5. Guardbanding: from characterization to shippable limits

A datasheet limit is a statement about the part. A production limit is a statement about the part *as measured by your tester, on your handler, at your temperature*. The difference between them is the guardband, and setting it is where most test programs quietly go wrong — usually by skipping it entirely.

The guardband is a stack of margins, each with a number and a reason:

1. **Tester accuracy.** Every measurement the ATE makes has an uncertainty — the force accuracy of the source, the measurement accuracy of the digitizer, at the specific range and conditions used. This number comes from the tester's specifications, not from optimism. If the digitizer contributes 2 mV on the range used, that 2 mV is the first entry in the stack.

2. **Contact and interface variation.** Probe needles wear. Socket contacts oxidize. A Kelvin connection removes the sense-path error — the sense lines carry no current, so their contact resistance drops out of the measurement — but it does not remove everything. The force-path drop still changes the true voltage at the DUT: a 17 mΩ force path at 300 mA is 5 mV of VIN error when VIN is forced at the tester without remote sense at the DUT pin. With remote sense at the DUT pin, the force loop compensates the drop and this term falls to the residual sense-path error (the 5 mV in the worked example assumes no remote sense). Sense-line leakage still flows, and any non-Kelvin path, such as a shared ground return, still contributes. Contact error dominates high-current, low-voltage measurements without remote sense; it is negligible for low-current measurements with proper Kelvin connections. Characterize the force-path resistance across the interface's life, not just when it's new, and state which part of it your guardband term covers.

3. **Thermal variation.** The part under test is not at the temperature you think it is. Self-heating during the insertion, handler chuck gradients, the difference between the thermal chuck setpoint and the die temperature — for a power IC, these routinely contribute several degrees of error, and every temperature-sensitive parameter (quiescent current, dropout, thermal shutdown) moves with it.

4. **Drift and process margin.** The characterization data came from a few lots. Production will see more. The guardband needs room for the process variation

you haven't measured yet — which is why limits get revisited after the first production lots, not set once and forgotten.

The method for setting a production limit, then, is arithmetic plus judgment:

Production limit (upper) = datasheet max − GB; production limit (lower) = datasheet min + GB

Each term documented, each term revisit-able. GB is the guardband: the measurement uncertainty U, built from the stack above. Where the terms are combined by root-sum-of-squares, U is reported as an expanded uncertainty at coverage factor k = 2 (approximately 95% confidence). The worked example instead combines the terms linearly (2 + 5 + 3 + 0 = 10 mV — tester, force-path, thermal, and drift, where the drift term is 0 now and revisited after the first three production lots) — a worst-case bound that is more conservative than a k = 2 RSS combination, and is used here as GB. One more term joins the stack only where the measurement is a sweep: step quantization. The dropout test sweeps VIN in 10 mV steps, so the dropout voltage is resolved only to 10 mV. The test reports dropout by linear interpolation between the last step above 95% of VOUT(NOM) and the first step below it, which bounds the quantization error at a worst-case ±5 mV (half-step). If the program instead reports the first failing step, the worst case is a full 10 mV, and the stack and limit must be adjusted (140 − 20 = 120 mV). The dropout production limit therefore uses a 15 mV stack (140 − 15 = 125 mV), while the static measurements use the 10 mV stack. State which combination you use, because it changes the guardband. This guardband is the same quantity the companion paper calls measurement uncertainty in its TUR definition (tolerance ÷ U). In both papers, 'tolerance' means the half-width: the ± figure for a two-sided specification, the limit value itself for a one-sided specification (a 0.2 µA maximum has a tolerance of 0.2 µA). A 10:1 TUR means U is a tenth of the tolerance, and a production limit set with GB = U is exactly a guardbanded limit. For the worked example, the 10 mV guardband against the ±49.5 mV output-accuracy tolerance is a TUR of about 5:1 — below the 10:1 target but above the 4:1 minimum, which is why the output-accuracy test is a measurement-limited (guardbanded) test under §3.3, not a full-TUR test. When the yield meeting asks why the output-accuracy production limit is 3.2605–3.3395 V instead of the datasheet's ±1.5% band, the answer is the stack — and if someone believes the contact-variation term is too conservative, the argument is about a number in a table, not about the whole program.

One more rule: **guardbands are asymmetric where the physics is asymmetric.** A dropout measurement that reads high because of contact resistance needs margin on the max side; the min side may need none. Blanket symmetric guardbands are a sign the stack wasn't actually built.

6. Honest limits: what this method doesn't cover

A method earns trust by stating its boundaries. This one has several.

It doesn't replace characterization. The method organizes decisions; it doesn't make measurements. Someone still has to run the corner lots, sweep the temperatures, and collect the data the guardband stack is built from. A beautiful test plan with unmeasured margins is fiction with good formatting.

It doesn't cover ESD, latchup, or reliability. Those are separate disciplines with separate standards (HBM/CDM, JEDEC latchup, and HTOL reliability flows). The method's "not tested, reason recorded" disposition will point at them, but it won't design them. This exclusion covers device-level ESD, latchup, and reliability qualification. It does not excuse test-induced overstress — spikes, sequencing, and relay transients the test program itself inflicts on the part are part of test development, covered in the companion paper's spike check (Paper 2 §8).

It doesn't validate system-level behavior. A part that passes every production test can still fail in the customer's system — line transients the datasheet never specified, startup sequencing interactions, thermal coupling to neighboring parts. Production test screens the part against its specification; it doesn't prove the application.

It doesn't eliminate judgment. The method's job is to make judgment explicit and reviewable, not to replace it. Deciding which corners to add to the condition matrix, how much drift margin to hold, whether a typical-only parameter deserves a production screen — these remain engineering calls. The method just ensures they're written down, so the next engineer can disagree with them productively instead of rediscovering them painfully.

7. The checklist

Everything above compresses into the checklist we run before writing a single line of test code. Eighteen questions; every "no" or "don't know" is a task, not an acceptable answer.

Extraction

1. Has every EC-table row been transcribed with its full condition set, including footnotes?

2. Have the typical curves been converted to rows (one per plotted point)?

3. Is every missing condition explicitly flagged rather than assumed?

4. Have the absolute-maximum ratings been checked against what the test itself will do to the part (the spike check in Paper 2 §8)?

Normalization

5. Are all contradictions between table, curves, and footnotes resolved, with the resolution recorded?

6. Has every typical-only parameter been classified: screen, sample, or accept by design?

7. Have parameters with no datasheet row (implied by curves or application text) been added as explicit rows?

Test plan

8. Does every normalized row map to one of the five dispositions: test in production, characterize only, not tested (with reason), tested via correlated proxy, or tested, measurement-limited (guardbanded)?

9. Does every production test specify its complete force/measure condition set?

10. Are variant-specific conditions (output voltage bands, package options) handled, not averaged?

11. Has the condition matrix been reviewed for the combinations the datasheet never shows together?

Limits

12. Is every limit traceable to a datasheet row or characterization dataset?

13. Has the guardband stack been built per test (tester, contact, thermal, drift — plus step quantization for sweeps)?

14. Are guardbands asymmetric where the physics is asymmetric?

15. Who approved each limit, and when — is it written down?

Measurement confidence

16. Has the TUR been assessed for every precision test (target 10:1, minimum 4:1 with documented guardband; tolerance = half-width)?

17. Has gauge R&R been signed off on the production setup (under 10% of tolerance target, 10–30% conditional, ndc ≥ 5; %GR&R on the full tolerance span, per Paper 2 §7.1)?

18. Has every pin been spike-checked through every transition (see Paper 2 §8)?

Decision log template

Parameter	Disposition	Conditions (if tested)	Limit source	Guardband	Approved by / date	Notes
dropout voltage (example)	Production test	IOUT=300 mA, VIN sweep 3.6→3.0 V, 10 mV steps	Datasheet max 140 mV (DQN, 2.5–5.5 V)	15 mV total (tester 2 mV + force-path 5 mV at 300 mA + thermal 3 mV + drift 0 mV + sweep quantization 5 mV at 10 mV steps) → production limit 125 mV	name, date	Temp specified by EC header (−40 to +125 °C); revisit after 3 lots; dropout reported by interpolation between sweep steps

8. About SiliconEng

SiliconEng is a silicon engineering services company in California, focused on analog, power management, RF, and mixed-signal ICs. We develop ATE test programs (wafer probe and package), run product qualification (AEC-Q100, MIL-STD-883), and support silicon validation, characterization, and datasheet creation. Our work is built on three principles: competence, customer care, and continuous improvement.

For questions about this paper or our services: www.SiliconEng.com, info@siliconeng.com

Sources

Texas Instruments, TPS7A20 datasheet, SBVS338H (July 2024). Values in the worked example are from this revision; later revisions may change limits — re-verify before reuse.

Trademarks and Copyright Notice

© 2026 SiliconEng. All rights reserved.

TPS7A20 is a product of Texas Instruments Incorporated. AEC-Q100 is a standard of the Automotive Electronics Council. MIL-STD-883 is a standard of the United States Department of Defense. JEDEC, HBM, CDM, and HTOL refer to standards and test methods of their respective publishing bodies. Device specifications cited in this paper are from the manufacturer's published datasheet; refer to the current datasheet for the authoritative specifications. All other trademarks are the property of their respective owners.