The Drain in the Bathtub Curve

Why technological obsolescence changes how electronics reliability should be managed

Kirk A. Gray
Original article: May 23, 2012 | Revised and edited draft July 9. 2026

Central argument: For many electronics products, the dominant reliability cost is paid before intrinsic wear-out ever becomes the controlling failure mechanism. The practical issue is not how long the parts might last in theory, but whether the design, manufacturing process, and use environment create early field failures. At the same time, the product is still commercially relevant.

Most reliability engineers recognize the life-cycle bathtub curve: a high early failure rate, a relatively flat useful-life region, and an increasing wear-out region at the end of life. The curve remains a useful teaching aid, but it can also lead reliability programs toward the wrong priorities when it is treated as a literal description of modern electronics field performance.

The important question is not whether wear-out can occur. It can. The question is whether wear-out is the primary business and customer problem for the product being developed. In many electronic systems, especially consumer products and IT hardware, the answer is often no. The product is replaced, upgraded, or made obsolete before the theoretical wear-out region dominates the cost of unreliability.

The traditional bathtub curve

The classic bathtub curve describes failure risk over time. The left side represents early-life failures, often called infant mortality. The middle represents a lower, approximately constant rate of failure during useful life. The right side represents increasing failures due to aging and wear-out.

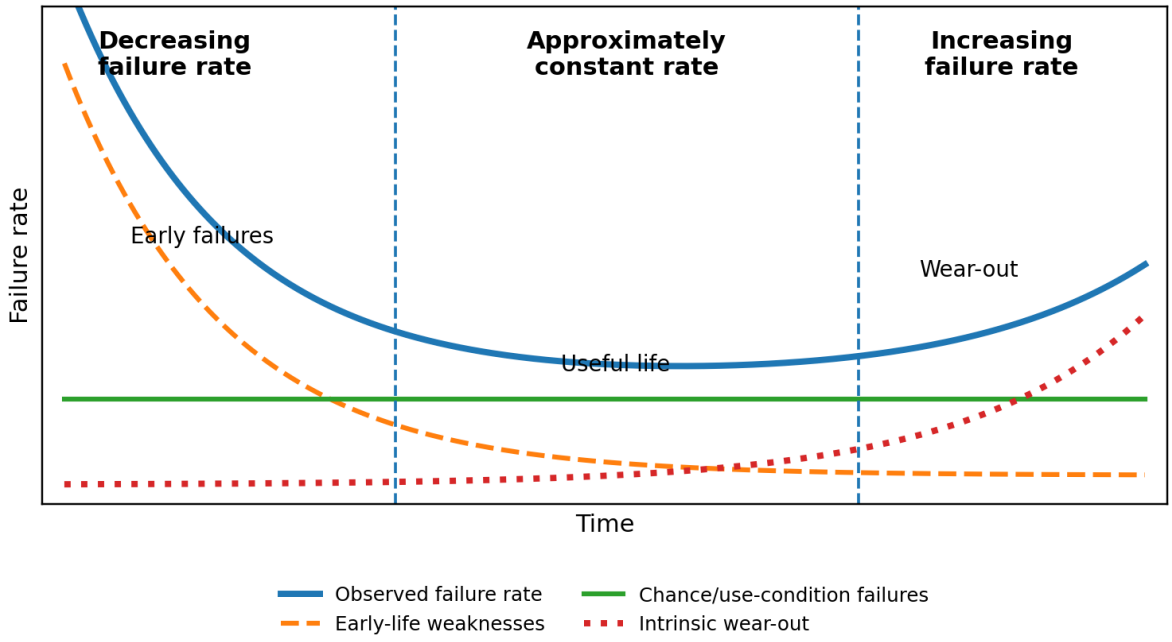

Figure 1. The traditional bathtub curve is a helpful concept, but it can become misleading if it causes teams to over-focus on the wear-out region for products that will be replaced before wear-out dominates.

The origin of the bathtub-curve analogy is not entirely clear. Still, it resembles human mortality experience: high vulnerability early in life, a long period of lower risk, and increasing risk late in life. The analogy also made sense in the era of vacuum tubes. Vacuum tubes and some early electronic technologies had inherent wear-out mechanisms that could limit the practical life of a system. In that context, life prediction and end-of-life behavior were central reliability concerns.

Modern solid-state electronics changed the balance. Many semiconductors and passive components can have intrinsic life capability far beyond the technologically useful life of the system. That does not mean systems will not fail. It means that the most important failures are often not caused by intrinsic component wear-out. They are more often caused by weak design margins, manufacturing variation, assembly defects, environmental exposure, power and signal integrity problems, firmware interactions, handling damage, or misapplication in the field.

Why reliability prediction can misdirect resources

Traditional electronics reliability programs have often relied on failure prediction methodology, including component-rate summations and handbook-style models. These methods depend on major assumptions: capable manufacturing processes, representative stress distributions, known use environments, stable component quality, and a relationship between calculated component failure rates and actual system-level field failure rates.

Those assumptions are rarely true in the simple way a model requires. A system-level failure is not just a component failure. A component can degrade without causing a system failure, and a system can fail even when no component has reached intrinsic end of life. The field failure may arise from the way the component is applied in-circuit, the design's operating margin, a marginal solder joint, contamination, a timing interaction, a software response to hardware behavior, or a user environment that was never adequately represented during development.

This is why a calculated mean time between failures can create false confidence. It can imply precision even when the dominant risks are uncertainty, variation, and untested margins. A reliability program can spend considerable effort estimating the far-right side of the bathtub curve while overlooking the weaknesses that will lead to warranty returns, no-fault-found events, customer dissatisfaction, and lost future sales during the first few years of use.

The drain: technological obsolescence

The bathtub analogy is incomplete for many electronics products because there is a drain in the tub. That drain is technological obsolescence. It is the point at which a product is replaced because newer technology, better performance, improved features, lower cost, security requirements, software support, or customer expectations make the existing product economically or practically obsolete.

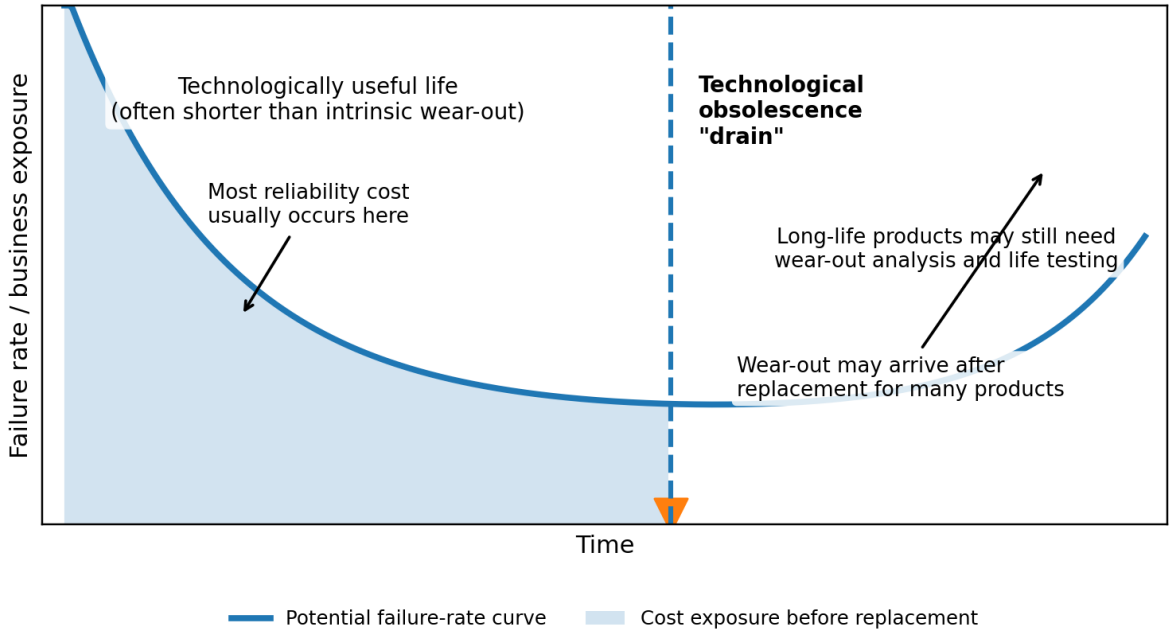

Figure 2. The obsolescence drain can occur before intrinsic wear-out dominates. For many products, reliability investment should therefore focus on early weaknesses and useful-life robustness, not only on theoretical end-of-life prediction.

For consumer electronics and much IT hardware, this drain can arrive well before the back end of the bathtub curve. A product may be technologically useful for five to seven years, while the wear-out mechanisms being modeled may not become significant until much later. In long-life applications - such as energy systems, infrastructure, aerospace, medical equipment, industrial controls, and other products expected to operate for decades - wear-out analysis and life testing remain important. The point is not to ignore wear-out. The point is to align the reliability program with the product mission and the actual cost profile.

When the drain occurs before the wear-out region, the business impact of reliability is concentrated at the front end and during useful life. Early failures are the failures customers remember. They are the failures that drive returns, service cost, brand damage, lost confidence, and lost follow-on purchases. Warranty cost is only the portion that is easiest to measure. The cost of lost trust is often higher, but much harder to quantify.

Where the real cost usually occurs

Most early-life and useful-life failures in electronics are assignable-cause failures. They usually come from something the organization could have discovered, understood, and corrected before shipment. Common contributors include overlooked design margins, inadequate derating, weak thermal paths, marginal interconnects, manufacturing process variation, supplier changes, contamination, poor handling, inadequate screening, immature diagnostics, and insufficient test coverage at realistic operating stresses.

This is the reliability cost that matters most to many manufacturers. A product that fails early may never reach the age at which wear-out becomes relevant to the customer. A company that repeatedly ships products with early field problems may lose market share before long-term durability can become a competitive advantage.

The practical implication is simple: reliability engineering should not be dominated by the attempt to calculate an impressive life number. It should focus on finding and removing weaknesses before customers do.

What should reliability engineering do instead?

A better use of reliability-development resources is to shift effort from prediction to discovery, correction, and verification. The goal is to expose design and process weaknesses early enough to be fixed economically. This requires tests and measurements that distinguish a robust product from a marginal one.

A more effective program should include:

- Review of field failures and lessons learned from predecessor products, with clear evidence that known mechanisms have been eliminated or mitigated.
- Design-margin analysis for thermal, voltage, timing, vibration, humidity, contamination, power quality, and realistic use conditions.
- HALT or other step-stress methods to expose operating and destruct limits and identify weak links before production.
- Failure analysis and root-cause discipline treats every significant failure as information, not as a statistical inconvenience.
- Production screening or stress auditing, when justified, to detect workmanship and process escapes without consuming useful product life.
- Better reliability discriminators, diagnostics, and prognostics that reveal degradation, intermittent behavior, and margin loss before a customer-visible failure occurs.
- Closed-loop corrective action so that design, supplier, manufacturing, and test processes are improved based on evidence.

This approach does not reject reliability modeling when appropriate and verified by field data and physics of failure. Verified correlating models can help compare design alternatives, understand physics-of-failure mechanisms, estimate acceleration factors, and plan tests. The problem occurs when the model becomes a substitute for evidence. Field reliability comes from margins, process control, diagnostics, root-cause learning, and corrective action. It is not created by a spreadsheet prediction.

Conclusion

The bathtub curve remains a useful picture, but it should not be allowed to define the reliability strategy for every electronic product. For many systems, technological obsolescence drains the bathtub before intrinsic wear-out dominates. The customer and business costs of unreliability are therefore concentrated in early-life and useful-life failures - the failures that are most often caused by design, manufacturing, application, and margin weaknesses.

Reliability engineering delivers the greatest value when it helps teams find those weaknesses before market introduction, understand their root causes, and remove them from the design and process. The objective is not to predict a distant theoretical life with false precision. The objective is to make the product robust enough that customers never have to discover the weak links themselves.