

The Bulrix Score

A Calibrated Standard for the Risk Inside Homes and Buildings

How a stewardship-based score on a permanent building record protects home value and reduces risk for warranty providers, insurers, and lenders

Sebin Joseph • Bulrix Research • Version 2.2 (final) • July 2026 • Model: bulrix-2.1.0 (shipped, conformance-audited)

Executive Summary

For most households the home is the largest asset they will ever own. It is also the collateral under the world's largest consumer credit market. Yet nobody measures its condition in a way that looks forward. Credit risk has FICO. Appraisals grade condition on a six-point scale at the moment of sale^[27]. Imagery scores read the roof from above. Nothing estimates, in advance and in numbers, how likely a particular home is to fail this year.

That gap is expensive. About one insured home in eighteen files a claim each year, and one in sixty-seven files for water or freezing damage at an average severity near \$14,000^[3]. Insurers, lenders, and warranty companies all underwrite the same house, and all of them do it blind. Meanwhile the insurance-availability crisis is pushing property risk straight into mortgage markets^[6,7].

The Bulrix Score closes that gap. It uses the familiar 300–850 scale, anchored so that 700 means a 5% chance of a major unplanned failure within twelve months. A major failure is an unscheduled repair of \$1,000 or more, or an emergency dispatch. Behind the number sits a component-level hazard engine. It begins with engineering survival curves for every major system, then adjusts them for verified maintenance, graded inspection findings, replacement plans, and protective devices. What it measures is stewardship, which the owner controls.

This is not a proposal. The platform is live: contractor tools are in production and the homeowner product is rolling out metro by metro. The engine described here ships as the score the product displays. It is parity-tested against a published reference implementation and covered by an automated invariant suite^[29]. Its priors were published before any outcome data existed, which is what pre-registration means. Calibration is now a test the model can publicly fail, and real homes are scored under it today.

Each stakeholder gets a different instrument. Homeowners get a ranked, priced path to fewer disasters, a defended sale price, and preserved insurability, carried by a record that belongs to the house rather than the app. Warranty providers get the underwriting tool their adverse-selection problem has always lacked, and in return they supply the component-failure outcomes that calibrate the score's frequency side^[18]. That is not one pitch among four. It is the answer to the score's own cold start. Insurers get a governed frequency and severity signal for interior risks no aerial image can see. Lenders get early warning on collateral condition and on borrower payment shock. Section 9 explains why the position compounds, what the moat is made of, and what could break it. The paper closes with a careful argument that a permanent, component-level building record is a first layer of the digital record automated construction will need. That argument draws in part on the author's own field study of automated residential construction, disclosed as such^[26]. It also states plainly what remains unsolved, because a standard that hides its limits never becomes one.

Why Bulrix Wins: the Thesis on One Page

1 WHY NOW

The risk moved inside the walls.

Insurance is withdrawing

1.9M non-renewals 2018-2023; uninsured owners doubled to 12% [6,7]

Deals now die on insurability

1 in 5 surveyed buyers and sellers lost a deal to insurance [7]

The exterior is already priced

imagery roof scores: half of top carriers, 40 states [12]

2 WHY BULRIX

The only assembled four-layer stack, live today. [29]

Permanent building record

every fact anchors to the house, not the app

Owner relationship

permissioned interior; fill-a-fact loop

Contractor rail

evidence is a condition of payment

Governed engine, shipped

anchored, pre-registered, parity-tested

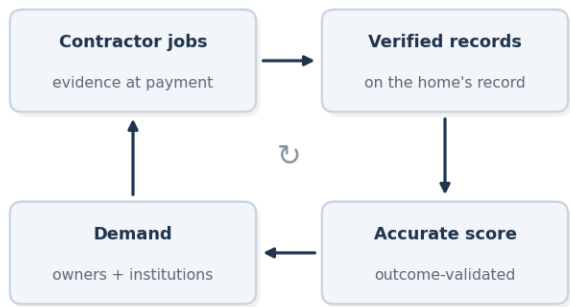
3 WHY INCUMBENTS CAN'T COPY IT

Each holds one layer. The missing three are slow and trust-bound.

	Record	Owners	Rail	Engine
Workflow platforms	-	-	✓	-
Data and imagery giants	✓	-	-	-
Warranty and insurer books	-	-	-	✓
Bulrix	✓	✓	✓	✓

4 HOW THE FLYWHEEL COMPOUNDS

Open standard, earned evidence.



every job sharpens every score

Four answers in under a minute. Depth behind each: timing \$1 · the four-layer stack \$3.1 · incumbents and compounding \$9 · governance of the conflict this model creates \$11.

1. The Problem: Risk Everyone Prices and No One Measures

Every home is a portfolio of aging mechanical systems. Each one moves along a survival curve toward a failure that is expensive, and the cost tends to fall hardest on the households least able to absorb it. In American Home Shield's 2024 survey, 87% of homeowners needed a professional in the prior year and more than a third could not cover an emergency repair above \$1,000^[2]. Roughly 60% of owners say they are deferring repairs because of cost^[4]. Deferral compounds. In one survey year, nearly a quarter of owners spent under \$1,000 on prevention while almost half spent more than \$5,000 on unplanned repairs, led by water and roof failures^[1].

When a market cannot see a risk, it withdraws from it. A U.S. Senate Budget Committee investigation covering 249 million policy records found more than 1.9 million homeowners non-renewed between 2018 and 2023, with the sharpest increases in the highest-risk counties^[6]. Uninsured ownership doubled to 12% between 2015 and 2023, and in one recent survey a fifth of buyers and sellers reported a deal that collapsed over insurance^[7]. Because insurance is a condition of nearly every mortgage, analysts now treat insurance withdrawal as a direct threat to collateral values and to mortgage-market stability^[7,8]. Every party is underwriting the same physical object from a different fragment of information. None of them holds the one number that matters: how likely is this home to fail this year?

2. What the Bulrix Score Is

The score estimates one thing: the probability that a home suffers a major unplanned failure in the next twelve months. A qualifying failure is an unscheduled repair or replacement of \$1,000 or more, or an emergency dispatch. The dollar threshold is nominal and will be indexed to construction costs in later versions. An emergency dispatch means same-day professional work to restore an essential service, which the model defines as water, heat during a freeze, cooling during a heat event, electrical safety, or sewage. Each event is adjudicated from the dispatch record and logged with its class and severity, which is what makes the exchange rate in §3.5 publishable.

Whether a repair counts as unscheduled is settled by documents, not by memory. A repair is scheduled only if a dated, funded plan existed at least sixty days before the event. A plan opened days ahead of a distressed replacement is still a surprise, and an owner's account of events after the fact never converts one.

The scale is the familiar 300–850, with a fixed points-to-double-the-odds of 40 and an anchor at 700, which equals 19-to-1 odds against failure and a 5% twelve-month risk. The resemblance to credit scoring is structural rather than cosmetic. Every score is the log-odds transform of a real probability, so every score is a prediction that can be proved wrong (Figure 1).

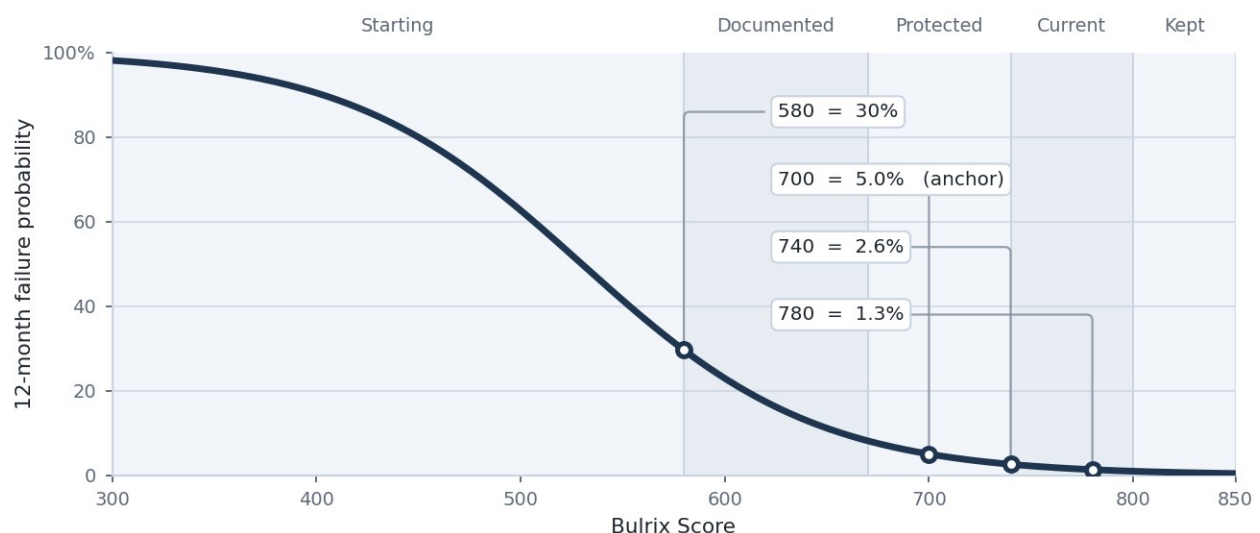


Figure 1. The anchored mapping between score and twelve-month major-failure probability (PDO = 40; 700 $\leftrightarrow$ 5%). The curve defines every score, not the population. Band boundaries shown.

Three invariants separate this from existing property ratings, and each is enforced in the input schema itself.

- Anchored, never curved. Scores map to absolute odds, not to percentile rank. A home's number cannot fall because a neighbor improved, and 720 means the same probability in every year and every market. No synthetic population is ever used to seed or center the distribution.
- Stewardship, not vintage, and never dollars. The chronological age of the house never enters the model. The age and condition of its systems do. A well-kept 1968 home therefore outscores a neglected 2015 build, by design and by automated test. Dollars never enter the probability either: a \$30 verified filter change done by the owner and the same work done professionally are identical to the estimate, because they change the home identically. Two clarifications keep that honest. Funding evidence on a plan is read as a signal of execution, so an escrowed deposit predicts that the replacement actually happens, but the amount is never an input. And the points-per-dollar ranking of Next Moves is a prioritization tool that consumes the score's outputs. Dollars order the to-do list. They never set the probability.
- Physics in, ZIP codes out. Environment enters as physics, because hard water really does shorten tank life and expansive clay really does move slabs. An absolute anchor demands this: 700 has to mean 5% in Phoenix and in Seattle, so identical stewardship under harsher physics can produce a different number. The score reports that rather than hiding it. What geography never becomes is a direct factor. There is no ZIP-code coefficient, no catastrophe-exposure load, and no neighborhood adjustment. Put plainly, environment changes how hard a score is to reach, never what the score means. The paper does not pretend that physics-mediated geography stops being geography. It is geography. The defense is that every environmental multiplier is published, enumerable, and offsettable by maintenance, which was never true of a ZIP-code coefficient. At launch every scored home carries the same published launch-market physics profile, so today no home is differentiated from another by location at all. As coverage expands, per-property physical data such as soil, water chemistry, and freeze and cooling normals will enter through this same published channel, never as a location coefficient.

Scores render in five bands: Starting (300–579), Documented (580–669), Protected (670–739), Current (740–799), and Kept (800–850). Two evidence rules govern them. First, self-reported records alone cap a score at 739, so the top of the scale has to be earned with photographs, sensor data, public records, or professional verification. Second, below a minimum evidence line the engine does not publish a low number at all.

A home becomes scoreable once its evidence covers systems representing at least seventy percent of its prior hazard mass. That is a weighting rather than a count, so a single invoice cannot admit a home with two of nine systems documented. Below that line the two surfaces deliberately differ. Institutions receive a Provisional category, the coverage figure, and a data-confidence grade, and never a numeric floor that an underwriter could quietly adopt as a rating. The homeowner keeps a working number, labeled provisional, next to the ranked facts that would unlock a verified score, because that number is what motivates the climb. Credit scoring's real precedent for missing information is a no-file, not a punishing number. An undocumented home is unknown, not failing. Above the threshold the sparse-record penalty of §3.2 phases in gradually, so no cliff separates unknown from penalized.

What does a typical home score? Simulated against era-weighted U.S. housing stock, an unmanaged mid-age home lands in the mid-500s, deep in Starting and consistent with the base rates in §3.4. The live early cohort, which carries open inspection findings and unknown install years, scores in the 400s and 500s at confidence E and climbs as facts arrive. One real home moved from 505 to 545 when its owner supplied four install years^[29]. The top of the scale is scarce by construction. That is a feature.

3. How the Score Works

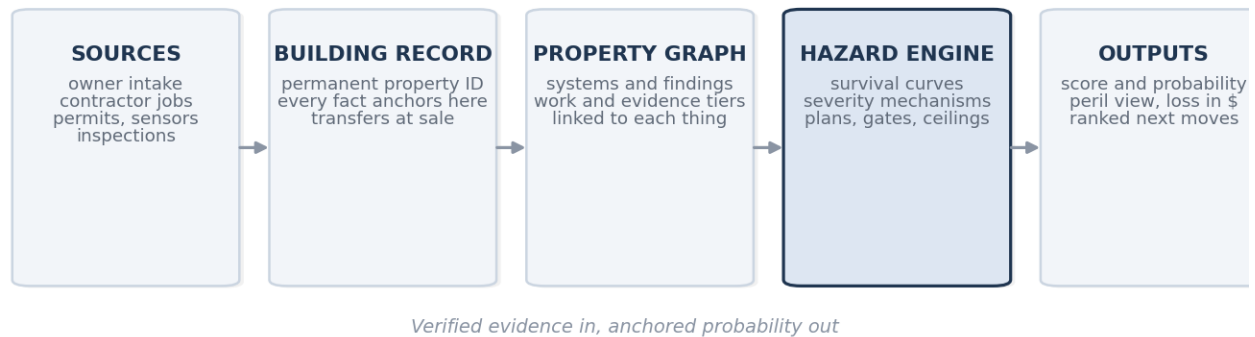


Figure 2. The architecture. Verified sources write to a permanent Building Information Record. The property data graph attaches every event to the thing it concerns. The hazard engine turns the record into an anchored probability, a loss view, and priced interventions.

3.1 A record that runs with the home

Underneath the score is a piece of infrastructure most rating systems skip: the Building Information Record, or BIR. Every property gets a permanent registry entry with an immutable identifier. Every fact anchors to that record rather than to a person or an app account, including systems, findings, work performed, evidence, and score history^[29].

The consequences are structural. An owner's account can be deleted and the personal information removed while the facts about the house survive. At sale the verified record transfers with the property, so the buyer inherits the ledger and its score instead of starting over at Provisional, and the previous owner's personal photos move only with consent. The service-logbook analogy becomes literal. The logbook belongs to the car.

On top of the BIR sits a property data graph: rooms and areas, the systems inside them, and a typed catalog that can be extended. The graph's inventory of major systems is exactly what the engine reads, so every new event attaches to the specific thing it concerns and sharpens every estimate that touches it^[29].

Evidence enters through enforced workflows rather than an honor system. Contractor jobs on the platform move through guarded steps: estimate, acceptance, work, documentation, payment. Capturing evidence at each step is a condition of getting paid, and escrowed funds release only after the work is submitted. That is why platform-completed work arrives at the professional evidence tier. Verification is how the job finishes, not an extra favor asked afterward.

The same loop already runs for owners in the product. Unknown facts are priced conservatively. The owner supplies an install year or a material, and the score recomputes immediately, showing exactly how many points that fact recovered^[29]. The climb is not a promise in a brochure. It is a button in the app.

The engagement burden is attacked by design, because the share of evidence that arrives passively is meant to grow. Permits and assessor records attach to the BIR with no owner action. Platform jobs document themselves as a condition of payment. Device telemetry writes continuously, and utility interval data can act as an early maintenance signal. The target is a mature record that accrues most of its evidence passively, the way credit files always have.

3.2 A hazard engine, not a points checklist

Every major system carries a survival curve by material and equipment type, seeded from published component-life references^[24]. The systems tracked are the water heater, cooling, heating, roof covering, panel, branch wiring, supply piping, sewer lateral, and foundation. The functional form is published rather than proprietary: two-parameter Weibull survival, $S(t) = \exp(-(t/\lambda)^k)$. The full prior table is printed in the implementation reference along with every defect floor and modifier, so a gas tank water heater uses $k = 3.0$ and $\lambda = 13$ years, and an architectural asphalt roof uses $k = 3.5$ and $\lambda = 25$ ^[29].

Equipment with a documented defect history carries a permanent hazard floor that only remediation clears. Federal Pacific and Zinsco panels, polybutylene supply piping, and un-remediated aluminum branch circuits all qualify, and they are the same items already on carrier declination lists. Each system's effective age then departs from its calendar age. Environment changes the physics. Stewardship changes the wear: a credibility-weighted record of completed

maintenance slows effective aging, where self-report counts half, photographs three-quarters, and permits and professionals count in full. An empty record accelerates aging for homes that are already scoreable.

3.3 A severity taxonomy: not all findings are equal

Every finding maps to a canonical issue code with four severity axes: life-safety class, functional impact, progression, and urgency window. Each code also carries exactly one mathematical mechanism. It multiplies a system's hazard, floors its failure probability, changes what a failure costs, or gates the score outright. A burned-out bulb is coded with every axis at zero and moves nothing. The same water heater, meanwhile, produces seven codes with seven different weights.

Water-heater complaint	Model effect	Required closure
Pilot / ignition won't stay lit	$\text{hazard} \times 1.3$	photo
Sediment rumble	$\text{hazard} \times 1.6$	photo (DIY flush)
TPR valve dripping	$\text{hazard} \times 2.0$	licensed invoice
Rust-colored hot water	$\text{hazard} \times 3.5$	pro assessment
Active leak at tank base	$\text{floor } 85\% + \text{incident cap}$	licensed invoice
Capped relief valve	life-safety gate (cap 579)	licensed invoice
Attic tank, no drain pan	$\text{loss severity} \times 1.8$	photo

Component lifetimes carry published provenance. Complaint-level hazard ratios such as $\times 3.5$ have no analog in the literature and are expert-elicited priors. They are the model's least-grounded parameters and the first in line for outcome refit (§3.5).

Issues left open past their urgency window escalate, so deferred maintenance compounds in the math rather than sitting still. Remediation is encoded twice for every issue: who may perform the fix, and what evidence closes it. Gas, refrigerant, panel-interior, and structural work close only on a licensed invoice or an engineer's report. A self-declared fix of that kind of issue keeps a residual risk and a capped score, while do-it-yourself work is fully honored wherever it is appropriate.

Some conditions are disqualifying no matter what the arithmetic says. Any open life-safety hazard caps the score at 579 until licensed repair is verified. Because an erroneous 579 would be an emergency of the system's own making, contested gate findings follow a published dispute protocol. An independent licensed re-verification runs on a five-business-day clock. The gate is annotated as under review but stays in force while the review runs, since a safety cap should not be removable by argument. If the finding is overturned it is expunged, the score is restored retroactively, Bulrix bears the cost of re-verification, and the evidence weight of a repeatedly erroneous reporter is reduced.

3.4 Plans, protections, and the elasticity hypothesis

A funded, scheduled replacement is priced for what it actually is, which is a change in exposure. The old unit is at risk only until the installation date, and the new one carries only infant hazard afterward. That makes a dated, funded plan the cheapest points-per-dollar action in the whole system, which also makes it the most obvious thing to game. So it ships with controls. The sixty-day lead-time floor from §2 applies. Plan credit scales with funding evidence, where an escrowed deposit outranks a saved quote, which outranks a stated intention. If the date passes without completion the plan lapses: credit is revoked, the component reverts to unplanned exposure, and a reason code is issued. Each rescheduling halves the remaining credit, repeated postponement extinguishes it and opens an overdue-replacement issue, and completion must close at the evidence tier the work requires. A plan is a priced commitment, not a promise.

This is also where the paper's load-bearing claim has to be stated at its true epistemic grade. It is not an assumption. It is a falsifiable hypothesis.

Hypothesis 1, on stewardship elasticity: stewardship can compress the 30–50% twelve-month surprise probability that an unmanaged typical home plausibly carries^[1,3] by roughly an order of magnitude. No residential study demonstrates this end to end. The closest evidence is adjacent and comes one peril at a time. Federal facilities guidance reports 8–12% cost savings for predictive over preventive maintenance, and opportunities above 30–40% where operations had been reactive^[91]. Wildfire mitigation moves that peril's risk by 40–75%^[9]. Shutoff devices move water losses^[10]. All of that measures costs, downtime, or a single peril. None of it measures residential twelve-month failure probability. That is exactly why the calibration program pre-registers the test: per-peril elasticity estimated

from cohorts stratified by verified adherence, with confidence intervals published whether or not they flatter the scale.

One channel deserves a candor note. Converting unplanned events into scheduled ones removes them from the outcome by definition. That is deliberate, because surprise is where the excess cost lives, in emergency premiums and consequential damage. The event-to-loss exchange rate in §3.5 is the empirical test of how much of that excess is real rather than merely reclassified. The other two channels are prevention and severity control, and each piece of evidence is filed under the channel it actually supports. The shutoff study^[10] is severity-channel evidence, since the failing tank still counts as an event. Device credits therefore act on the probability that a failure becomes major and on expected annual loss, never on the event probability itself.

3.5 Explanation, calibration, and honest validation

Every score ships with its machinery visible: per-peril probabilities, expected annual loss in dollars, a data-confidence grade, adverse-action-style reason codes, and ranked Next Moves. Each Next Move is a counterfactual in which the engine re-scores the home with that fix completed and reports the exact points recovered per dollar spent.

Calibration follows the path credit scoring took: published expert priors, prediction and outcome logging, and credibility-weighted refits released as numbered model versions. There is no silent drift. One precedent does not transfer, and the program does not pretend otherwise. Imagery scores could be retro-validated against a carrier's book because historical imagery already existed^[12]. Verified maintenance ledgers exist for nobody's book retroactively.

Validation therefore runs forward. Cohorts scored today are measured over 12 to 24 months against published acceptance criteria: a calibration slope of 0.9 to 1.1, a Gini of at least 0.35, and monotone loss-ratio lift across bands. Those criteria bind per peril wherever event counts allow, because a composite can pass while a thin peril is noise. Power is stated in events rather than homes. The slope band is the binding criterion, and it wants on the order of a thousand observed events, which is roughly ten thousand home-years at blended rates.

One wedge is named rather than buried. The score is calibrated to surprise events, while a carrier pays loss dollars. Improvement in the first channel, which reclassifies a failure into a scheduled replacement, moves the event rate further than it moves claims, since claims respond mainly through tail truncation. The per-peril event-to-loss exchange rate is therefore a published calibration deliverable alongside elasticity, and the expected-annual-loss output, not the score, is the carrier's loss-side quantity.

4. For Homeowners: Protecting the Asset

The owner's case builds along three horizons. The first is fewer disasters and cheaper ownership. Prevention economics are lopsided: a quarter of owners spend under \$1,000 preventing the failures on which nearly half spend more than \$5,000^[1]. The score turns that into a ranked, priced work list, and its evidence tiers reward the cheapest valid path, including safe do-it-yourself work^[9].

The second is a defended sale price. Practitioners consistently report that visible deferred maintenance drags a price down by more than the repair would have cost, with one brokerage estimate putting the erosion near 150% of repair cost^[22], because buyers price hidden problems pessimistically^[23]. A verified record inverts that dynamic, and since the record runs with the home (§3.1), the buyer inherits it intact.

The third is insurability, which is becoming part of the asset's value. As non-renewals climb and deals fail over insurance^[6,7], documented evidence that a particular home is not the risk its ZIP code implies is worth money. Onboarding is designed to respect the owner throughout. A new home enters as Provisional rather than as a punishing number, and the score is earned into, in either direction, as facts arrive.

5. The Calibration Partner: Home Warranty

For home warranty providers the alignment is definitional. The industry's own filings describe the product as protecting customers from costly, unplanned breakdowns of essential home systems and appliances^[18], which is the outcome the score predicts, almost word for word. The industry leader alone carries about 2.1 million active warranties and handles on the order of four million service requests a year through roughly 17,000 contractor firms^[18], in a market that covers only about 4 million of 121 million U.S. homes^[19,20].

The score addresses the industry's two structural wounds at once. Adverse selection, where coverage is bought disproportionately on aging systems, becomes surgical underwriting and renewal pricing, and renewals are roughly three-quarters of the leader's revenue^[18]. The industry's ugliest dispute, denial for improper maintenance, becomes a

clean adjudication against a verified ledger. Because home service contracts sit largely outside state insurance codes^[21], adoption needs no rate filings.

Most importantly, the relationship runs both ways. Warranty claims are component-level failure outcomes with costs, ages, and trades attached, and there are millions of them a year^[18]. That is precisely the calibration data the score lacks at launch.

What that data can and cannot calibrate is stated with equal candor. Warranty books are adversely selected, so the industry's defining wound cuts both ways. Their events include service calls below the \$1,000 threshold. They carry no verified-stewardship covariates. Corrected for selection they calibrate failure frequency quickly, but they do not calibrate insured severity, because warranty severity is capped, trade-rate component replacement, while the loss that makes water a \$14,000 peril is the ensuing damage to structure and contents that warranty files systematically lack. So the chicken-and-egg problem is half-broken, and this paper says which half. Frequency calibrates through warranty data. The severity tail still needs carrier claims.

Correcting for selection is specified rather than waved at. Enrollment age enters the survival fits as left truncation, so observation begins at sign-up and absorbs much of the bought-on-aging-systems distortion directly. Enrollment itself is modeled on the observables warranty files carry, which are system age, home age, and plan tier, then reweighted against housing-stock age distributions. Fitted curves must reconcile with published life tables inside credibility bounds, and divergence beyond those bounds is treated as a residual-selection alarm rather than a discovery. Waiting periods screen the sharpest short-horizon selection by construction. The prospective consumer cohort then serves as the unselected holdout that catches whatever bias survives. The residual assumption is selection on observables, and private knowledge is exactly what the holdout exists to detect. The per-peril elasticities of Hypothesis 1 still come only from the slowest channel, which is scored consumer homes with stewardship variation observed over years, and the calibration timeline says so. Even with those limits, the warranty deal is what makes every other deal in this paper closable.

6. For Insurers: The Interior Complement to the Imagery Stack

Carriers have already industrialized the exterior view. Imagery-based roof-condition scores are used by half of the top U.S. insurers and are approved for ratemaking in 40 states^[12], and Moody's acquired the leading provider in 2025^[13]. Composite roof models now blend imagery, claims clustering, and weather forensics^[14]. What no aerial pass can see is the interior: the systems that drive non-weather water losses, which is the one-in-sixty-seven peril averaging \$14,000^[9], and the behavior that decides whether those systems fail catastrophically or get replaced on schedule.

The score supplies that complement in actuarial form, and it is explicit about which evidence supports which output. The frequency signal is the set of per-peril twelve-month probabilities, calibrated on warranty and pilot outcomes. The severity signal is expected annual loss, and here the device literature is strong: shutoff hardware cut paid water-claim frequency by 96% and severity by 72% in a controlled comparison across 1.3 million homes^[10], and precursor-arc sensors are deployed at multi-million-home scale^[11].

Inputs are anchored in independently verifiable tiers, including finaled permits^[15], licensed invoices, sensor telemetry, and job documentation captured by the workflow described in §3.1. Self-report is explicitly discounted rather than trusted. Adoption then follows the industry's normal ladder for third-party data: prevention partnerships and inspection triage first, then renewal monitoring, then filed discounts, then rating. Each step is gated by prospective pilot results, and the vendor-governance file is built for state AI oversight^[16] and for the consumer-reporting duties that attach when a score informs eligibility^[17].

7. For Lenders: Collateral You Can Finally See

Mortgage models are rich on the borrower and nearly blind on the building. The lender is also the stakeholder for whom the score's anchor is exactly right. A carrier pays loss dollars, but the surprise event itself is the payment shock a servicer fears, so the calibration target and the lender's risk are the same variable with no exchange rate in between. The sixty-day floor in §2 is what protects that match.

The building asserts itself through two channels. The first is insurability. With 1.9 million non-renewals documented^[6], uninsured ownership at 12%, reported markets where 30–40% of applications fail on insurance costs, and a fifth of surveyed buyers and sellers reporting a collapsed deal^[7], insurance withdrawal is now a named threat to collateral values and mortgage-market stability^[6,7,8]. Condition and mitigation are the controllable half of insurability^[9].

The second is payment shock. A majority of owners cannot comfortably absorb a \$5,000 emergency repair^[5], so a forward-looking failure probability is an early warning no credit bureau provides. Financing the ranked Next Moves through home-improvement and equity products is the rare case where the bank's product is also the borrower's risk reduction. Deployment needs no regulatory novelty: condition documentation in underwriting files, a portfolio-monitoring overlay, eligibility signals for maintenance-linked lending, and eventually a disclosed collateral attribute in securitization as calibration seasons.

8. A Digital Record of the Built Environment

The claim here has to be precise, because prior art exists. Appraisal condition ratings standardize condition at origination on a six-point scale^[27]. The facility condition index has ordered institutional capital planning for decades. As-built BIM and digital-twin practice describe new buildings in rich detail. The United Kingdom's post-Grenfell golden thread now legislates a digital building record for higher-risk buildings^[28]. What none of them does is estimate forward-looking failure probability from a continuously updated, evidence-tiered operating record. The appraisal is a snapshot, the twin is as-designed, and the golden thread is a compliance file. The BIR is the missing operating layer, published as an open, versioned schema so that it can become a standard rather than a silo, and it extends naturally from houses to small multifamily and light-commercial buildings.

That layer matters beyond risk transfer, because construction automation runs on digital description and is throttled by missing measurement. The United States is short an estimated 3.8 million homes^[26]. A recent field case study of one of the first custom 3D-printed residences, conducted through the author's own build practice and cited with that disclosure, shows a three-person crew printing an entire wall system in about 110 machine-hours directly from a digital model, with the unfinished shell surviving a tornado. The same study documents no applicable building codes, roughly \$25,000 of over-engineering forced by regulatory ambiguity, and no long-term performance data anywhere, and it closes by calling for standardized machine logs and material records^[26].

A calibrated index answers both halves of that problem. It gives durable methods a market price that current standards cannot express, and digitally born homes arrive with day-one, Grade-A records. The connection is operational rather than rhetorical: the engineer-certified entry pathway in §10 is how a printed shell becomes scoreable, its as-built model and machine logs become the BIR's first evidence, and the score becomes the mechanism that pays builders back for documenting their work. Measurement, then automation, then abundance is the sequence that industrialized food and clothing. It is offered here as a direction of travel, not a destiny.

9. Why This Compounds: The Economics of the Standard

A fair investor question deserves a direct answer. Why does an adjacent giant not simply build this? Because the score needs four assets operating together, and each incumbent class holds exactly one of them. Contractor-workflow platforms capture the work, but they anchor it to the contractor's business rather than to a permanent property record, so there is no BIR, no owner relationship, and no actuarial layer. Property-data and imagery incumbents hold records and rooftops at national scale, but their posture cannot summon what the interior requires, which is an owner's ongoing permissioned participation and a contractor rail where evidence is a condition of payment. Warranty and insurance books hold outcomes, but adversely selected ones, and no verification network. Bulrix's position is the assembled stack already operating together: permanent record, owner relationship, enforced evidence workflow, and governed engine^[29]. A challenger has to build the three layers it lacks, and those are always the slow, two-sided, trust-bound ones.

Each side then makes the others better, which is what compounding means here (Figure 3). Every contractor job writes professional-tier evidence to a home's record. Sharper records calibrate every similar component everywhere. A more accurate score earns institutional adoption. Adoption makes the score worth money to owners at sale and at renewal. Owners then prefer contractors who write to the record, and job matching routes on verified standing, so contractors join to win work rather than out of virtue. The switching cost is the record itself. Years of property-anchored, evidence-tiered history cannot be recreated by a rival on day one, and because the record transfers at sale, it grows more valuable with age rather than less.

Figure 3. The flywheel. Each participant's self-interested use of the platform improves the score for everyone else, and the loop closes through job matching that routes on verified standing.

is unequally distributed, since permit digitization and professional-service density track jurisdiction wealth. Physics flows through, so scores will pattern by climate, soil, and water chemistry, which pattern in turn by region and demographics, and the published-multiplier defense in §2 is what makes that channel testable rather than deniable. System count is a mechanical channel of its own: more systems means more chances for one surprise, so a large multi-zone home scores below a condo at identical per-system stewardship. The anchor survives that, because more systems genuinely is more payment-shock exposure, but a per-system stewardship view is published alongside the score so complexity is visible rather than confounded. The mitigations are structural, including equal credit for verified do-it-yourself work, plan credits, financing attached to Next Moves, and published distributional monitoring. Whether they are sufficient is an empirical question to be answered in public.

Novel construction, adoption, and economics. Automated methods produce buildings with no life tables. The printed-home study performed no lab tests on printed samples and notes that no long-term record exists anywhere^[26]. Novel assemblies therefore enter through engineer-certified priors, mandatory instrumentation, and accelerated recalibration, which is cold-start discipline applied to methods instead of components. Commercially, institutions adopt on demonstrated lift and lift requires scale. The warranty path bridges the frequency half of that problem, and the open BIR schema is published so the record layer cannot fragment into proprietary silos, because standard-scale ambition on a proprietary schema is a platform rather than a standard. The economics are stated plainly. Homeowners pay for the maintenance product the score sits on. Institutions license scores and portfolio analytics. Furnishers are compensated through the verification network. Sustaining homeowner engagement is a cost imagery scores never faced, and it is an operating challenge the consumer product has to win on its own merits.

11. Governance, Conflicts, and Trust

The governance model ships with the mathematics. The score is criterion-referenced, so it is meaningful on day one with no reference population, and no synthetic data ever touches calibration evidence. Model changes are numbered, documented releases, and the anchor offset in particular can change only as a new major model version with a republished odds table. Recalibration shrinks parameters toward published priors in proportion to the evidence behind them.

Fairness is structural first: protected characteristics, house vintage, neighborhood, and spending have no input fields to enter through. It is also verified rather than asserted, since the program commits to publishing proxy-discrimination testing across all three difficulty channels in §10. Where scores inform insurance eligibility, the architecture assumes the full set of consumer-reporting obligations, including permissible purpose, accuracy procedures, disclosure, dispute and correction rights, and adverse-action reason codes^[17]. Where they inform credit decisions, the frame widens to fair-lending law, which means Regulation B adverse-action notices and disparate-impact testing on the same channels.

One conflict deserves its own paragraph, because the paper's favorite analogy demands it. Reference layers do become infrastructure, but issuer-pays conflicts also helped turn bond ratings into the cautionary tale of 2008^[30]. The business model here is structurally adjacent, since the score seller also operates a remediation marketplace and the verification rail. So the firewalls are architecture rather than policy. Nothing purchasable moves a score, because the engine's input assembly reads no payment, subscription, or engagement fields, which makes no-pay-for-points a schema guarantee. Next Moves are ranked solely by the engine's counterfactual points-per-dollar, with no paid placement and published ranking logic. Work performed through the marketplace closes issues under exactly the evidence standards required of any outside contractor, the enforced workflow ties payment release to documentation, and closures are audit-sampled at published differential rates that are highest for marketplace work, where an insufficient or fraudulent review reverts the closure and decays that source's evidence weight. Scoring itself operates as a revenue-separated function under annual independent attestation. The scorer must not be for sale, including to itself.

12. Conclusion

Homeowners absorb failure costs they cannot predict^[1,5]. Insurers withdraw from risk they cannot see^[6]. Lenders are learning that collateral quality and insurability are the same question^[7]. Warranty providers price blind against selection^[18]. The Bulrix Score closes that gap with the only design that can earn institutional trust: an anchored probability on a permanent building record, workflow-verified evidence, hard safety gates, explanation and intervention as a single object, and a warranty-first, pre-registered calibration path that repeats how credit scoring earned its authority. Its open problems are demanding but tractable, and none of them is a barrier of principle. The

ambition, read at its true size, is this: reference layers become infrastructure, a score for people underwrote consumer credit and ratings underwrote the bond market, and the aim here is the equivalent reference layer for the built environment, a system of record that others price, lend, insure, warranty, and build against. A home that can prove its own condition is worth more, insures easier, borrows safer, and fails less. The record it leaves behind is the first course of a digitally mapped, automated, and ultimately more abundant built world^[26].

References

- [1] Hippo, "Housepower Report" (3rd annual, survey of 2,000+ U.S. homeowners), 2024–2025; summary reporting 83% experienced unexpected repairs and 46% spent over \$5,000. bluefieldgroup.com/blog/average-homeowner-spent-over-5000-on-repairs-in-2024
- [2] American Home Shield, homeowner survey and maintenance-cost analysis, 2024 (87% required a professional; 33%+ unable to cover a \$1,000 emergency repair). ahs.com/home-matters/homebuyer-hub-resources-and-guides/big-fixes-big-price-tags-the-rising-cost-of-home-maintenance-and-repairs/
- [3] Insurance Information Institute, "Facts + Statistics: Homeowners and Renters Insurance" (ISO claims data, 2019–2023 averages). [iii.org/fact-statistic/facts-statistics-homeowners-and-renters-insurance](https://www.iii.org/fact-statistic/facts-statistics-homeowners-and-renters-insurance)
- [4] Today's Homeowner, national home-repair survey (deferral rates, emergency-repair costs, savings levels). todayshomeowner.com/general/guides/home-repair-survey/
- [5] Clever Real Estate, "The True Cost of Owning a Home," 2026 survey of 1,000 homeowners. listwithclever.com/research/costs-of-owning-a-home-2026/
- [6] U.S. Senate Budget Committee, "Next to Fall: The Climate-Driven Insurance Crisis Is Here — And Getting Worse," December 2024 (249 million policy records; 1.9M+ non-renewals, 2018–2023).
- [7] Levy Economics Institute of Bard College, "A Premium Crisis: Climate Change Threatens Homeowner's Insurance, Housing, and Financial Stability," 2026 (citing First Street 2025 on uninsured share; transaction-failure survey data). levyinstitute.org/publications/a-premium-crisis
- [8] Public Citizen, "Mapping the Home Insurance Crisis" (Federal Insurance Office and Senate Budget Committee data), 2025. citizen.org/article/mapping-the-home-insurance-crisis/
- [9] National Association of Insurance Commissioners, wildfire mitigation research (2020), as summarized in Consumer Reports' homeowners-insurance rights analysis. consumerreports.org/money/homeowners-insurance/proposed-homeowners-insurance-protection-rights-a3867354890/
- [10] LexisNexis Risk Solutions, water-loss correlation study of Flo by Moen shutoff devices (2,306 installed homes vs. ~1.3M control), 2020. risk.lexisnexis.com/about-us/press-room/press-release/20200505-flo-by-moen
- [11] State Farm Newsroom, expansion of the Ting electrical-fire safety program to 2 million homes (Whisker Labs sensors). newsroom.statefarm.com
- [12] Cape Analytics, "Roof Condition for Property Insurers" (Roof Condition Rating adoption and ratemaking approvals). capeanalytics.com/blog/roof-condition-for-property-insurers/
- [13] Moody's Corporation, acquisition of Cape Analytics, January 2025 (public announcement).
- [14] LexisNexis Risk Solutions, "Rooftop" roof loss-propensity model (imagery, claims clustering, weather forensics). risk.lexisnexis.com/products/rooftop
- [15] ATTOM Data Solutions, "The Benefits of Using Building Permit Data in Insurance Underwriting" (200M+ permits, 2,000+ jurisdictions). attomdata.com
- [16] National Association of Insurance Commissioners, "Model Bulletin: Use of Artificial Intelligence Systems by Insurers," adopted December 2023; state adoption tracking through 2026.
- [17] Venable LLP, "FCRA and HIPAA Concerns in Insurtech Underwriting," 2020 (consumer-reporting-agency exposure for third-party scores used in insurance eligibility). venable.com/insights/publications/2020/05/fcra-and-hipaa-concerns-in-insurtech-underwriting
- [18] Frontdoor, Inc., Annual Report (Form 10-K) and quarterly filings, 2024–2025 (active warranties, service-request volume, contractor network, renewal revenue mix). sec.gov
- [19] Warranty Week, "American Home Shield During the Pandemic," June 2021 (market size; ~4M of 121M homes covered). warrantyweek.com/archive/ww20210610.html
- [20] Frontdoor, Inc., completion of the 2-10 Home Buyers Warranty acquisition (\$585M), Business Wire, December 2024.
- [21] National Home Service Contract Association, regulatory statement on state treatment of home service contracts. homeservicecontract.org/about-nhsca/regulatory-statement/
- [22] Berkshire Hathaway HomeServices practitioner estimate of deferred-maintenance value erosion (~150% of repair cost), Coastal Illustrated real-estate report, 2023. coastalillustrated.com
- [23] Houston Association of Realtors (HAR.com), "How Deferred Maintenance Can Impact Your Home's Market Value," 2026. har.com

- [24] InterNACHI, "Standard Estimated Life Expectancy Chart for Homes"; NAHB/Bank of America, "Study of Life Expectancy of Home Components"; ASHRAE equipment service-life data (survival-curve priors).
- [25] ConsumerAffairs, "Home Maintenance Costs: A Breakdown" (1–4%-of-value budgeting guidance; NAHB American Housing Survey routine-maintenance data), 2026. consumeraffairs.com/homeowners/home-maintenance-cost-breakdown.html
- [26] Joseph, S., "Tornado-Resilient 3D-Printed Housing: A U.S. Case Study Validating Cost-Effective Construction Using Conventional Mortar," Inventive X Design Build LLC, SSRN Working Paper No. 5325985 (automated residential construction case study). Author self-citation, disclosed in §8. ssrn.com/abstract=5325985
- [27] Fannie Mae, Selling Guide: uniform appraisal condition ratings C1–C6 (standardized point-in-time condition classification embedded in conforming U.S. mortgage appraisals). fanniemae.com
- [28] United Kingdom, Building Safety Act 2022 and secondary regulations: the 'golden thread' requirement for a digital record of building information on higher-risk buildings.
- [29] Bulrix Research, Bulrix System Documentation: the Building Information Record (BIR) registry, property data graph, enforced contractor job workflows, evidence-tier assembly, parity-tested scoring engine (Python reference + production port), and invariant test suite (implementation reference, 2026). Author self-citation.
- [30] Financial Crisis Inquiry Commission, The Financial Crisis Inquiry Report (2011): findings on credit-rating-agency issuer-pays conflicts of interest.
- [31] U.S. Department of Energy, Federal Energy Management Program / Pacific Northwest National Laboratory, Operations & Maintenance Best Practices Guide, Release 3.0 (predictive maintenance: 8–12% savings over preventive alone; opportunities exceeding 30–40% versus reactive operations). www1.eere.energy.gov/femp/pdfs/om_5.pdf
- [32] Coalition for Content Provenance and Authenticity (C2PA), Content Credentials technical specification (capture-time media provenance). c2pa.org

About this paper — The methodology described here launches as an actuarially structured expert system: its survival parameters and modifier magnitudes are published priors drawn from the component-life and loss literature cited above, replaced by fitted, credibility-weighted estimates as outcomes accrue and released as numbered model versions. The platform, engine, invariant test suite, and building-record architecture described are implemented and operational; the score ships as the product's displayed number, and an evidence-based conformance audit (July 2026) verified the running system against this specification claim by claim, with the report available on request. Three mechanisms are built and tested but await data flow rather than code: live C2PA manifest verification, reporter evidence-weight decay (which activates with a source registry), and the holiday calendar behind the dispute clock. Several market statistics are cited from vendor summaries and survey reports; primary studies are pinned for model-governance review. Hypothesis 1 is stated as falsifiable and pre-registered. Revised (v2.1, final) across four rounds of external review. This paper is informational and is not insurance, lending, legal, or investment advice.