County Distress Index

Methodology and Scoring

Publication	American Default Research
Prepared by	Ross Kilburn, American Default Research
Version	2.0
Date	April 2026
License	Creative Commons Attribution 4.0
Canonical URL	https://americandefault.org/methodology/cdi/
Contact	press@americandefault.org

Abstract. The County Distress Index (CDI) scores all 3,144 U.S. counties on a 0–100 scale measuring relative household financial distress. It is built from 21 indicators drawn from 14 federal and nonprofit data sources, organized by principal component analysis into five statistically derived dimensions with PCA-proportional weights: Consumer Credit Distress (47.5%), Housing Cost Burden (22.3%), Structural Poverty (13.6%), Economic Vitality (9.2%), and Legal Distress (7.4%). Scoring uses percentile-rank averaging with a small-county adjustment that caps individual indicator percentiles at [5, 95] for counties with populations at or below 20,000. Pre-PCA diagnostics are favorable (KMO = 0.85; Bartlett's test of sphericity rejects the null at $p < 0.001$). Parallel analysis confirms five components, which together explain 70.6% of total variance. Face validation passes on 27 of 28 benchmark counties, and sensitivity analysis across four candidate weighting schemes produces Spearman rank correlations between 0.93 and 0.97. This document specifies the full methodology, data sources, validation, and limitations.

Keywords. County Distress Index, household financial distress, principal component analysis, consumer credit distress, bankruptcy filings, housing cost burden, structural poverty, economic vitality, county-level data, percentile-rank scoring.

Suggested citation. American Default Research. (2026). *County Distress Index: Methodology and Scoring*. https://americandefault.org/methodology/cdi/

Contents

1. Summary

The County Distress Index ranks all 3,144 U.S. counties on household financial distress using 21 indicators across five statistically derived dimensions. Weights are determined by principal component analysis — a standard statistical technique that lets the data reveal which dimensions of distress vary most across counties, rather than relying on editorial judgment. A small-county adjustment caps individual indicator percentiles at [5, 95] for counties with populations at or below 20,000, so statistical noise in tiny counties does not drive the extremes of the national rankings.

The five dimensions and their weights are listed below. A score of 50 is the national median. Counties are classified into five zones: Healthy (< 35), Normal (35–50), Elevated (50–65), Serious (65–80), and Crisis (≥ 80).

Dimension	Weight	What It Measures
Consumer Credit Distress	47.5%	Whether households are falling behind on debt payments
Housing Cost Burden	22.3%	Whether housing costs are crushing household budgets
Structural Poverty	13.6%	Whether a county has deep, persistent economic disadvantage
Economic Vitality	9.2%	Whether wages, business formation, and housing values support stability
Legal Distress	7.4%	Whether households are turning to bankruptcy courts for relief

Data sources: Federal Reserve/Equifax, Bureau of Labor Statistics, Census Bureau, U.S. Courts, HUD, Bureau of Economic Analysis, FHFA, and CFPB. All data is publicly available. No proprietary data is used.

2. The Five Dimensions

The CDI organizes its 21 indicators into five dimensions. The dimensions were not chosen editorially. They emerged from principal component analysis run on the full indicator set across all 3,144 counties — the statistical procedure identifies independent factors of variation, and the weights are proportional to how much of the cross-county variance each factor explains.

2.1 Consumer Credit Distress (47.5%)

This factor captures whether households are actively falling behind on debt payments. It explains the most cross-county variation because debt distress is the sharpest differentiator between struggling and stable counties.

Indicator	Source	Weight	Loading	Vintage
Debt in collections	Urban Institute / Equifax	10.4%	0.915	Aug 2025
Subprime credit population	FRED / Equifax	10.2%	0.906	Q4 2025
Auto loan delinquency	Urban Institute / Equifax	8.5%	0.825	Aug 2025
Credit card delinquency	Urban Institute / Equifax	8.0%	0.804	Aug 2025

Medical debt in collections	Urban Institute / Equifax	6.6%	0.728	Aug 2025
Uninsured rate	Census ACS 5-Year	3.8%	0.553	2019–2023

Why the uninsured rate lands here. *The uninsured rate correlates with credit distress because uninsured households are one medical event from collections. PCA discovered this empirical relationship — the indicator was originally tagged as community vulnerability, but it loads on the credit distress factor alongside the collections, subprime, and delinquency series.*

2.2 Housing Cost Burden (22.3%)

This factor captures whether housing costs — rent and mortgage payments — are consuming an unsustainable share of household income. Homeownership rate is inverted: lower homeownership means more renters and higher exposure to displacement.

Indicator	Source	Weight	Loading	Vintage
Severe renter cost burden >50%	Census ACS 5-Year	8.2%	0.852	2019–2023
Renter cost burden >30%	Census ACS 5-Year	8.1%	0.847	2019–2023
Homeownership rate (inverted)	Census ACS 5-Year	3.5%	0.556	2019–2023
Owner cost burden >30%	Census ACS 5-Year	2.4%	0.461	2019–2023

2.3 Structural Poverty (13.6%)

This factor captures deep, persistent economic disadvantage — counties where poverty, disability, and transfer dependency are structural features rather than cyclical blips. MHI ratio to state median is inverted.

Indicator	Source	Weight	Loading	Vintage
Transfer income (% of personal income)	Bureau of Economic Analysis	3.0%	0.828	2024
MHI ratio to state median (inverted)	Census SAIPE	3.0%	0.823	2023
Disability rate	Census ACS 5-Year	2.0%	0.680	2019–2023
Poverty rate	Census SAIPE	2.0%	0.675	2023
Child poverty rate	Census SAIPE	1.9%	0.667	2023
Unemployment rate	BLS LAUS	1.7%	0.626	Feb 2026

The coal country vs. Delta divergence. *PCA separates acute credit distress (Factor 1) from chronic structural poverty (Factor 3). Appalachian coal counties and tribal reservations score high on Factor 3 but lower on Factor 1 — their populations have less access to credit, so they show less credit-based distress. Mississippi Delta and urban South counties score high on both. The separation is a feature: the CDI can distinguish between counties where people are actively falling behind on payments and counties where many people never had access to credit in the first place.*

2.4 Economic Vitality (9.2%)

This factor captures the earning power and economic dynamism of a county. Wage-to-rent ratio, business applications per 1,000, and FHFA house price YoY change are inverted: higher values indicate less distress.

Indicator	Source	Weight	Loading	Vintage
Wage-to-rent ratio (inverted)	BLS QCEW + HUD FMR	4.3%	0.839	2024
Rent-to-income ratio	HUD FMR + Census SAIPE	2.5%	0.648	FY2026
Business applications per 1,000 (inverted)	Census BFS	1.6%	0.519	2024
House price YoY change (inverted)	FHFA	0.8%	0.367	2024

2.5 Legal Distress (7.4%)

This factor captures a single indicator: the bankruptcy filing rate. It loads on its own factor because bankruptcy is a terminal distress signal that does not correlate cleanly with the other four dimensions. Some high-poverty counties have low filing rates because their residents lack access to bankruptcy counsel. Some moderate-income counties have high rates because Chapter 13 filings are used strategically to preserve homes. PCA correctly identifies this as an independent axis of distress.

Indicator	Source	Weight	Loading	Vintage
Bankruptcy filing rate per 100,000	U.S. Courts Table F-5A	7.4%	0.571	2025

3. Scoring Methodology

The pipeline is thirteen steps. Each county starts with 21 raw indicator values and ends with a single 0–100 composite score, a zone assignment, and per-factor breakdowns.

3.1 Data Collection. 21 indicators from 14 federal and nonprofit sources. All data is freely available. No proprietary data is used.

3.2 Coverage Check. Indicators must cover at least 75% of counties (≥ 2,358 of 3,144) to be included in the composite. Three mortgage-related indicators were evaluated and excluded — see Section 7, Coverage Gap.

3.3 Imputation. Missing values are filled with the state-level median for the indicator. If no state median is available, the national median is used. Total imputations: 1,222 of 66,024 cells (1.9%). The most imputed indicator is FHFA house price YoY change (14.2% imputed), which carries the lowest weight (0.8%).

3.4 Direction Alignment. All indicators are oriented so that higher values mean more distress. Five indicators are inverted: MHI ratio to state median, homeownership rate, wage-to-rent ratio, business applications per 1,000, and FHFA house price YoY change.

3.5 Winsorization. Extreme values are clipped at the 2.5th and 97.5th percentiles of each indicator before standardization. This prevents a single outlier county from distorting the scale of any indicator.

3.6 Z-Score Standardization. Each indicator is converted to Z-scores (mean = 0, standard deviation = 1) across all 3,144 counties, so all indicators are on the same scale regardless of units.

3.7 Principal Component Analysis. PCA is run on the correlation matrix of the 21 Z-scored indicators. Kaiser-Meyer-Olkin sampling adequacy is 0.85 ("meritorious" on Kaiser's 1974 scale; minimum threshold 0.60). Bartlett's test of sphericity rejects the null at $p < 0.001$ ($\chi^2 = 49{,}636$, df = 210). Component count is determined by parallel analysis (1,000 Monte Carlo iterations), which recommends five components. The Kaiser criterion (eigenvalue > 1) and the 70% cumulative variance threshold agree. The five components explain 70.6% of total variance.

3.8 Varimax Rotation. Components are rotated using varimax rotation to maximize interpretability. Each indicator loads strongly on one factor rather than weakly on several. Rotation does not change total variance explained.

3.9 Weight Derivation. Each indicator is assigned to its primary factor (highest absolute loading). Factor weights are proportional to variance explained. Within each factor, indicator weights are proportional to squared loadings. Final weight = factor weight × within-factor weight. All weights sum to 100%.

3.10 Percentile Scoring. Each Z-scored indicator is converted to a percentile rank (0–100) across all 3,144 counties. Percentile ranks are more interpretable than Z-scores for county-level rankings — a percentile of 72 means "more distressed on this indicator than 72% of counties."

3.11 Small-County Adjustment. For any county with a population at or below 20,000, each individual indicator percentile is clipped to the range [5, 95]. See Section 4 for the rationale and a concrete example.

3.12 Weighted Average. The CDI composite score is the weighted sum of clipped indicator percentiles: CDI = Σ (indicator_percentile × indicator_weight).

3.13 Zone Assignment. The composite score is mapped to one of five zones — Healthy, Normal, Elevated, Serious, or Crisis — using fixed thresholds (see Section 5).

4. Small-County Adjustment

Before indicator percentiles are combined into the composite, one more adjustment runs. For any county with a population at or below 20,000, each individual indicator percentile is clipped to the range [5, 95]. The motivation is statistical volatility.

A county of 5,000 people can trip to the 99th percentile on a single indicator because of small-sample noise alone. A handful of extra bankruptcies in a single year, or one unusual ACS housing estimate, can push a tiny county to an extreme reading that would not appear in a larger county with the same underlying conditions. The CDI is meant to measure relative household distress, not to reward or punish counties for the statistical fragility of their data.

The clip prevents any single indicator from dominating the composite in a small-population county. The county can still score at the very top or bottom of the national rankings — but only if many indicators move together, which is precisely what the index is designed to detect. Counties with populations above 20,000 are not adjusted.

4.1 Effect on the Extremes

The most visible effect of the adjustment is on the extremes of the distribution. Tunica County, Mississippi (population 9,234) would score 90.8 without the adjustment and would be the most distressed county in the country under raw percentile scoring. With the adjustment, Tunica scores 88.54 and ranks third nationally, and Richmond County, Georgia (population 205,414, no clip applied) becomes the most distressed county at 89.15. At the opposite end of the distribution, Los Alamos County, New Mexico (population 19,444) scores 11.37 with the clip (versus 9.3 raw), and Hamlin County, South Dakota (population 6,451) holds the least-distressed position at 10.85.

5. Zone Thresholds and Score Distribution

The 0–100 composite is mapped to five fixed zones. Counts below reflect the live distribution across all 3,144 counties.

Zone	Score Range	Counties	Share	Interpretation
Healthy	< 35	672	21.4%	Less distressed than roughly 65% of counties
Normal	35 – 50	894	28.4%	Near the national median
Elevated	50 – 65	937	29.8%	More distressed than roughly half of counties
Serious	65 – 80	579	18.4%	More distressed than roughly 80% of counties
Crisis	≥ 80	62	2.0%	Among the most distressed counties in the country

5.1 Score Distribution

Mean composite score: 50.04. Median: 50.06. Standard deviation: 16.24. Range: 10.85 (Hamlin County, SD) to 89.15 (Richmond County, GA). By construction of percentile-rank scoring, the national median is approximately 50 — half of counties score above and half below.

6. Validation

6.1 Face Validation

Twenty-eight benchmark counties were tested before the index was locked: sixteen known-distressed (Mississippi Delta, Appalachian coal, Black Belt, border counties, tribal reservations) and twelve known-healthy (wealthy suburbs, tech corridors, college towns, resort areas). Fifteen of sixteen known-distressed counties correctly ranked in the top half (94%). All twelve known-healthy counties ranked in the bottom half (100%). The only misclassification was Towns County, Georgia, which ranked 1,710 — just outside the top half.

6.2 Sensitivity Analysis

Four candidate formulas were tested against the PCA-proportional specification. Spearman rank correlations between all pairs range from 0.93 to 0.97. The top 50 most distressed counties overlap between 38 and 50 across specifications. The core distressed counties (Tunica, Bolivar, Sunflower, Richmond) appear in the top 50 across all four formulas. Borderline zone assignments are more sensitive to weighting than the core rankings.

Formula	Method	Spearman ρ vs PCA	Top-50 overlap
A (production)	PCA-proportional weights	1.000	50 / 50
B	Equal indicator weights	0.952	33 / 50
C	Equal domain weights	0.933	38 / 50
D	PCA + geometric mean	0.968	41 / 50

7. Coverage Gap: Mortgage Delinquency

The most significant limitation of the current CDI is the absence of a direct mortgage delinquency indicator. Three mortgage-related indicators were evaluated during construction and excluded because none could cover the 75% threshold required for inclusion in the composite. The exclusions are structural features of the source data, not data staleness.

Indicator	Coverage	Reason excluded
CFPB mortgage delinquency 90+ days	471 (15.0%)	CFPB/FHFA National Mortgage Database is a 5% sample. County estimates are publi
HUD FHA serious delinquency	1,686 (52.7%)	HUD Neighborhood Watch reports serious delinquency only for counties with active l
Student loan delinquency	2,128 (67.7%)	Urban Institute suppresses county-level estimates where the Equifax 4% panel has ins

Mortgage distress is captured indirectly in the composite through the subprime credit population indicator (99.7% county coverage), debt in collections, and credit card delinquency — all of which the national American Distress Index validates as leading indicators of mortgage default. When county-level mortgage data coverage improves, direct mortgage delinquency will be incorporated.

8. Limitations

8.1 Cross-sectional snapshot. The CDI ranks counties relative to each other at a point in time. It does not tell you whether a specific county is getting better or worse over time. Annual refreshes will enable year-over-year comparisons starting in Year 2.

8.2 Data lag varies by indicator. BLS unemployment is roughly two months old; Census ACS housing data is from 2019–2023 (a five-year rolling average). The composite blends indicators of different vintages. All vintages are documented in Section 2.

8.3 Imputation. 1.9% of indicator-county values are imputed from state medians. Counties with imputed values are flagged in the underlying data but are not downweighted in the composite.

8.4 No direct mortgage delinquency signal. The most significant data gap, documented in Section 7.

8.5 PCA captures variance, not welfare. The 47.5% weight on Consumer Credit Distress reflects the fact that debt indicators vary most across counties. This is a statistical observation, not a normative claim that credit distress matters more than structural poverty. Counties with deep poverty but limited credit access may score lower than their lived experience warrants. Factor-level scores are reported alongside the composite specifically so researchers can examine each dimension separately.

8.6 Source concentration in Factor 1. Five of the six highest-weighted indicators in Consumer Credit Distress originate from the Equifax panel via Urban Institute. PCA correctly identifies them as one factor, but shared source methodology means shared measurement artifacts. Cross-validation against the independently sourced subprime credit indicator (FRED/Equifax, separate pipeline) and bankruptcy filing rate (U.S. Courts) confirms the factor structure is not an artifact of single-source measurement.

9. How the CDI Compares to Other County Indices

Several county-level indices exist. Each answers a different question, and the differences are not cosmetic. The CDI is the only one focused specifically on whether households are falling behind on their bills.

	CDI	EIG DCI	CDC SVI	Cty Health Rankings
Question	Are households falling behind financially?	Is the local economy growing?	Is population vulnerable to disaster?	Is population healthy?
Coverage	3,144 counties	ZIP codes aggregated	Census tracts aggregated	~3,000 counties
Credit bureau data	Yes	No	No	No
Direct debt distress	Yes	No	No	No
Weighting	PCA-derived	Equal weights	Percentile ranking (unweighted)	Equal weights
Update frequency	Annual	~5 years	Biennial	Annual

The Economic Innovation Group's Distressed Communities Index measures whether a local economy is producing jobs and income. The CDC's Social Vulnerability Index measures whether a population is exposed to environmental and social hazards. County Health Rankings measures health outcomes. None of them ask the question the CDI asks. A county can have strong job growth (low EIG distress) while its households drown in credit card debt (high CDI distress). The CDI captures that divergence by combining credit bureau panel data with federal housing, labor, and bankruptcy sources — a combination no other public index uses.

10. Data Sources

All CDI data comes from public sources. No proprietary datasets, no paywalled inputs, no survey-based estimates that cannot be independently replicated.

Source	Agency	Indicators	Cadence
Debt in America	Urban Institute / Equifax	Collections, medical, auto, credit card delinquency	Annual
Equifax Subprime	FRED / Equifax	Subprime credit population	Quarterly
LAUS	Bureau of Labor Statistics	Unemployment rate	Monthly
SAIPE	U.S. Census Bureau	Poverty, child poverty, MHI	Annual
ACS 5-Year	U.S. Census Bureau	Housing burden, homeownership, uninsured, disability	Annual rolling
Table F-5A	U.S. Courts	Bankruptcy filings per 100,000	Annual
FMR	HUD	Rent-to-income, wage-to-rent	Annual (FY)
QCEW	Bureau of Labor Statistics	Wage-to-rent ratio	Quarterly
BFS	U.S. Census Bureau	Business applications per 1,000	Monthly
HPI	Federal Housing Finance Agency	House price YoY change	Annual county
Personal Income	Bureau of Economic Analysis	Transfer income %	Annual

11. Reproducibility

The scoring pipeline is implemented in Python and runs against publicly available source data. The production scoring engine is at `scripts/county/compute_county_dci.py`. The full factor analysis pipeline — imputation, winsorization, PCA, rotation, weight derivation, face validation, and sensitivity analysis — is in `scripts/county/cdi_rebuild_analysis.py`, with output artifacts in `docs/cdi_rebuild/`: correlation matrix, rotated loadings, weight derivation, sensitivity tests, face-validation results, and all-county scores under four candidate formulas. Running the rebuild script reproduces every number in this document.

Questions about methodology or requests for the raw data extract should go to press@americandefault.org.

12. Citation

The County Distress Index is released under Creative Commons Attribution 4.0. Please cite American Default Research as the source.

12.1 APA

American Default Research. (2026). *County Distress Index: Methodology and scoring*. https://americandefault.org/methodology/cdi/

12.2 Chicago (author-date)

American Default Research. 2026. “County Distress Index: Methodology and Scoring.” https://americandefault.org/methodology/cdi/.

12.3 BibTeX

```
@techreport{cdi_methodology_2026,
  title = {{County Distress Index: Methodology and Scoring}},
  author = {{American Default Research}},
  year = {2026},
  institution = {{American Default Research}},
  url = {https://americandefault.org/methodology/cdi/},
  note = {Principal component analysis across 21 indicators in 5
               statistically derived dimensions, covering all 3,144
               U.S. counties}
}
```