Nowcasting Supplemental Poverty Measure thresholds

A pre-committed 2025 estimate from public microdata

Max Ghenis

2026-07-18

Abstract

On July 17, 2026, the Bureau of Labor Statistics reissued six years of Supplemental Poverty Measure thresholds after finding errors in the code that produced them. The correction moved every published threshold by less than 1.6 percent. Auditing our own software against it found hand-entered threshold values off by up to 7.4 percent — and both errors had survived for years because no one was diffing published artifacts. The paper makes three contributions. It quantifies the correction and documents a replication of the BLS threshold methodology from public Consumer Expenditure microdata, which reproduces the corrected thresholds within 1.5 to 2.1 percent per year and surfaces six pitfalls in the public files that face any CE researcher. It shows that CPI-U threshold indexation, common in microsimulation practice including our own, missed actual threshold growth by 2.2 percent per year in absolute value over 2020–2024, understating it in four of five years, because the SPM is quasi-relative: BLS re-estimates each year from consumption data. And it commits to a falsifiable nowcast of the unpublished 2025 thresholds — $41,036.34 for owners with mortgages, $34,135.99 for owners without, and $40,755.98 for renters — from a 50/50 blend of consumption growth and composite price aging (backtest mean absolute error 0.76 versus 2.23 percent per year for CPI-U aging; the consumption component alone scores 0.41). The nowcast was amended once, before any BLS release, after adversarial cross-model review caught a construction error in the composite index; the original values remain in the repository history. BLS publishes the actual 2025 thresholds around September 2026; a revision of this paper will report the realized errors, whatever they are.

1 Introduction

On July 17, 2026, the Census Bureau announced it would re-release Supplemental Poverty Measure estimates for 2019 through 2024, after the Bureau of Labor Statistics found errors in the thresholds that anchor them (U.S. Census Bureau 2026). BLS reissued the corrected thresholds the same day, attributing the errors to the computer code introduced with its September 2021 methodology change, and re-anchored the series at 82 percent of median expenditures, down from 83 percent, to minimize the break (U.S. Bureau of Labor Statistics 2026b). Every published SPM poverty statistic for those six years — national rates, state rates, demographic breakdowns, and the research literature built on them — used thresholds that were wrong by up to 1.6 percent.

We maintain open-source software that carries these thresholds into policy microsimulation, so we audited it against the correction. The audit found a larger problem on our side: threshold values transcribed by hand into our own package matched no BLS or Census publication for four of the six affected years, with errors up to 7.4 percent. The two failure modes are the same failure mode. Reference data that people transcribe, in agencies or in open-source projects, drifts silently unless something mechanical checks it against the source.

That observation sets the paper's agenda. Section 2 describes how BLS produces the thresholds and their revision history, which now includes two code-correction episodes. The 2026 correction is quantified in Section 3 against the values Census actually published, with our own package's larger errors alongside it. Section 4 documents an independent replication of the BLS methodology from public Consumer Expenditure microdata: its measured fidelity, the bugs the exercise found in our replication code, six pitfalls in the public CE files that face any researcher who uses them, and a negative result — replication at survey level cannot detect a sub-2-percent error in the official pipeline, though mechanical artifact-diffing catches it in a week. Because BLS re-estimates thresholds each year from a rolling consumption window, the series is quasi-relative, and Section 5 shows what that costs the price indexation used in microsimulation practice, ours included; a backtest scores replacement rules against five years of corrected actuals. Section 6 applies that rule to data that already exist and commits, before publication, to a numeric estimate of the 2025 thresholds that BLS will publish around September 2026. Section 7 draws the implications for poverty measurement infrastructure.

2 How the thresholds are produced

The Supplemental Poverty Measure compares a resource unit's post-tax, post-transfer resources to a threshold that varies by housing tenure and local housing costs (Interagency Technical Working Group on Developing a Supplemental Poverty Measure 2010). BLS produces the national thresholds from the Consumer Expenditure (CE) Interview Survey; Census applies geographic adjustments, computes resources, and publishes the poverty statistics (Garner and Munoz 2021).

Under the methodology the SPM Interagency Technical Working Group approved in September 2020 (Burns and Fox 2021), the threshold for target year T uses five years of CE Interview quarters lagged one year — collection quarters $(T-5)$Q2 through TQ1 (U.S. Bureau of Labor Statistics 2026b, 2026c). Within the estimation sample of consumer units with children, BLS computes expenditures on food, clothing, shelter (including owners' mortgage-principal payments), utilities, telephone, and internet — FCSUti — plus imputed in-kind benefits from broadband subsidies, the Low Income Home Energy Assistance Program, the National School Lunch Program, the Special Supplemental Nutrition Program for Women, Infants, and Children, and rental assistance. It equivalizes them to a two-adult, two-child reference unit with the three-parameter scale of Betson (1996), updates each quarter to threshold-year dollars with a composite FCSUti price index, and applies the formula

$$\mathrm{SPM}_h = s \times \left(1.2 \times \overline{\mathrm{FCSUti}}_E - \overline{\mathrm{SU}}_E + \overline{\mathrm{SU}}_{Eh}\right),$$

where E is the subsample inside the 47th–53rd percentile band of equivalized FCSUti, SU is shelter plus utilities excluding telephone, h indexes three housing-tenure groups, the 1.2 multiplier covers other basic goods and services, and the anchor share s is 82 percent as of the 2026 correction, 83 percent before it.

Three properties of this design matter for what follows. First, the thresholds are re-estimated, not indexed: each year's value reflects the consumption distribution in its window, so the series moves with living standards as well as prices. The National Academy of Sciences panel that proposed this architecture intended exactly that (Citro and Michael 1995). Second, the pipeline is complex enough to fail quietly. The 2026 reissue is its second code-correction episode: the 2022 thresholds already carried "corrections in

the computer code used to model" in-kind benefits, disclosed in a footnote of the 2022 poverty report (U.S. Census Bureau 2023); a BLS decomposition attributes 2.5 to 2.9 percentage points of that year's unusually large threshold increase to the code improvement and data updating combined, against 7.2 points from inflation (Schild 2023). Census has separately explored simplifying the design, estimating that a decennially-anchored variant would have put 2024 SPM poverty at 11.3 percent rather than the published 12.9 (Creamer 2026). Third, everything BLS consumes is public — the CE microdata, the CPI component series, the methodology documents — so the pipeline can be replicated, and its outputs checked, by anyone.

3 The 2026 correction, quantified

The corrected workbook publishes full-precision thresholds for 2005–2024, with standard errors and tenure population shares, and states that these are the values Census uses to produce SPM statistics (U.S. Bureau of Labor Statistics 2026b).[1] The pre-correction values were removed from the per-year BLS threshold pages at the correction, so we transcribed them from the five annual P60 reports that published them — each year from 2020 on cross-verified against the two consecutive reports that print it, and 2019 from its single source, the 2020 report (U.S. Census Bureau 2021, 2022, 2023, 2024, 2025).[2] Table 1 compares the two series for two-adult, two-child units.

Table 1: The 2026 correction: thresholds as published in Census P60 reports versus the corrected BLS workbook, two-adult, two-child consumer units, nominal dollars of the threshold year.

Year	Tenure	Published	Corrected	Change
2019	Owners with mortgages	29,080	29,076.17	-0.0%
2019	Owners without mortgages	24,413	24,514.95	+0.4%

[1]The Census statement singles out 2019 and 2020 as the years the code errors affected; the workbook nonetheless reissues all six years 2019–2024, and every year's values changed. This paper treats the full reissued range as revised.

[2]The removal has since become inconsistent: as of August 6, 2026, the BLS page for the 2024 thresholds again serves the pre-correction values — $39,430 for renters, $39,068 for owners with mortgages, $32,586 for owners without — with no correction notice, while the 2023 page remains empty and the corrected workbook carries different 2024 values. A reader who lands on the per-year page today receives superseded numbers, 0.4 to 0.9 percent from the corrected ones: the vintage-consistency problem this paper describes, on the producer's own site.

Year	Tenure	Published	Corrected	Change
2019	Renters	29,194	28,913.05	-1.0%
2020	Owners with mortgages	29,959	29,814.55	-0.5%
2020	Owners without mortgages	25,222	25,249.14	+0.1%
2020	Renters	30,150	29,978.90	-0.6%
2021	Owners with mortgages	31,107	30,983.24	-0.4%
2021	Owners without mortgages	26,279	26,055.40	-0.9%
2021	Renters	31,453	31,216.69	-0.8%
2022	Owners with mortgages	34,235	33,978.49	-0.7%
2022	Owners without mortgages	28,909	28,454.92	-1.6%
2022	Renters	34,518	34,140.06	-1.1%
2023	Owners with mortgages	36,915	36,966.24	+0.1%
2023	Owners without mortgages	30,870	30,587.73	-0.9%
2023	Renters	37,482	37,230.75	-0.7%
2024	Owners with mortgages	39,068	39,230.99	+0.4%
2024	Owners without mortgages	32,586	32,878.59	+0.9%
2024	Renters	39,430	39,219.89	-0.5%

Every change is within ±1.6 percent, with mixed signs: renter thresholds fall in five of six years, owner thresholds move both ways. 2022 sees the largest downward revision across all tenures. Census will quantify the effect on poverty rates in a working paper before the September 2026 annual report. Most published year-to-year rate comparisons have room to absorb a revision of this size — our microsimulation puts the level effect near 0.2 percentage points — but not all: the 2024 SPM rate printed at 12.9 percent, statistically unchanged from 2023 (U.S. Census Bureau 2025), and a flat comparison can change direction under even a small revision. We leave the rate arithmetic to the Census working paper.

Our own software did worse. The spm-calculator package, which supplies threshold parameters to the PolicyEngine US microsimulation model, carried a hand-entered table of "published BLS thresholds." Table 2 compares that table to what Census actually published.

Table 2: Threshold values shipped in spm-calculator through version 0.3.1, against the values Census published. The 2019–2020 rows appear to mix misattributed vintages; 2022–2023 match no located BLS or Census publication.

Year	Tenure	Package $ $0.3.1	Published	Error
2019	Owners with mortgages	27,172	29,080	-6.6%
2019	Owners without mortgages	22,600	24,413	-7.4%
2019	Renters	27,515	29,194	-5.8%
2020	Owners with mortgages	28,533	29,959	-4.8%
2020	Owners without mortgages	23,948	25,222	-5.1%
2020	Renters	28,881	30,150	-4.2%
2021	Owners with mortgages	31,089	31,107	-0.1%
2021	Owners without mortgages	26,022	26,279	-1.0%
2021	Renters	31,453	31,453	+0.0%
2022	Owners with mortgages	32,949	34,235	-3.8%
2022	Owners without mortgages	27,679	28,909	-4.3%
2022	Renters	33,402	34,518	-3.2%
2023	Owners with mortgages	36,192	36,915	-2.0%
2023	Owners without mortgages	30,347	30,870	-1.7%
2023	Renters	36,606	37,482	-2.3%
2024	Owners with mortgages	39,068	39,068	+0.0%
2024	Owners without mortgages	32,586	32,586	+0.0%
2024	Renters	39,430	39,430	+0.0%

Errors reach −7.4 percent, several times the BLS correction. Only 2024 and the 2021 renter value were right. The releases that carried these values passed continuous integration on every commit: the test suite asserted the code returned the packaged values, and the packaged values were the thing that was wrong. The fix removes the transcription step entirely: values are now generated from the official workbook, which ships alongside them with its SHA-256 recorded, and a scheduled comparison against the live BLS file flags any divergence — a silent-drift risk becomes a detectable event. Against the live BLS site, twenty years of packaged thresholds now match within half a cent.

4 Replicating the thresholds from public microdata

The same package includes a from-scratch implementation of the BLS methodology over raw CE Interview public-use microdata: it downloads the quarterly FMLI files, constructs FCSUti, equivalizes, deflates with a composite index built from CPI component series, and applies the threshold formula. Before this work, that implementation had never been validated against official values in earnest — its one comparison test allowed 5 percent tolerance and was skipped by default. Benchmarking it against both the pre-correction and corrected series, with the anchor share matched to each (83 percent versus published values, 82 percent versus corrected, so anchor choice cannot manufacture a fit), surfaced four errors in our replication code:

1. **Annualization off by a factor of two.** The FMLI expenditure variables report a three-month recall window split across two calendar quarters (`*PQ` and `*CQ`); the code treated the pair as six months of spending and multiplied by two instead of four.
2. **Telephone double-counted.** The FMLI `UTIL` summary already contains `TELEPH` (utilities are defined as natural gas, electricity, fuels, telephone, and water); the code added telephone again.
3. **Phantom columns.** Two columns the code referenced — mortgage principal and home internet — do not exist as FMLI summaries; both silently contributed zero.[3]
4. **The wrong formula shape.** The code took per-tenure percentiles of the FCSUti distribution. The BLS formula pools the estimation band, applies the 1.2 other-necessities multiplier, and swaps a tenure-specific shelter-utilities average — a different object.

After the fixes, the faithful variant reproduces the corrected thresholds with mean absolute deviation across the three tenures of 1.5 to 2.1 percent per year (1.0 to 1.6 percent against the superseded published series), with no imputation of in-kind benefits. For scale, the workbook's published sampling standard errors run 0.7 to 2.1 percent of the corresponding threshold across 2019–2024, so the replication sits within roughly one to three standard errors of the official values. Table 3 reports signed deviations.

[3] Principal outlays live in the `EMRTPNO*` and `MRTPRNO*` outlay columns; home internet has no FMLI summary variable at all.

Table 3: Replication deviations from official thresholds (owners with mortgages / owners without / renters), with the anchor share matched to each reference series.

Year	vs published (83% anchor)	vs corrected (82% anchor)
2019	-1.2% / -0.5% / -1.7%	-2.4% / -2.1% / -1.9%
2020	-1.5% / +0.6% / -1.7%	-2.2% / -0.7% / -2.4%
2021	-1.4% / -0.7% / -1.5%	-2.2% / -1.0% / -1.9%
2022	-1.6% / +0.1% / -1.9%	-2.1% / +0.4% / -2.0%
2023	-1.1% / +0.4% / -1.6%	-2.4% / +0.1% / -2.1%
2024	-1.4% / +2.4% / -1.2%	-3.0% / +0.2% / -1.8%

The deviations are predominantly negative and roughly stable across years: the replication imputes no in-kind benefits and observes no internet expenditures, and that shows up as a level deficit — up to three percent for owners with mortgages — that persists across estimation windows rather than growing or shrinking with any one year. Two conclusions follow. The negative one: the BLS correction moved thresholds by at most 1.6 percent, inside this replication's noise, and neither reference series fits systematically better in the affected years — so survey-level replication could not have caught the BLS bug, and we do not claim otherwise. What catches that class of error is mechanical comparison of published artifacts, which now runs weekly. The positive one: a level deficit that is stable across adjacent, largely overlapping windows is exactly the error structure that cancels when the replication is used in ratio form, which is how Section 5 and Section 6 use it.

4.1 Pitfalls in the public CE files

Six properties of the public data, encountered while building the replication and benchmark, will face any researcher who works with these files. None is documented in one place elsewhere.

1. **Recall-window semantics.** FMLI `*PQ`/`*CQ` pairs split one three-month recall period across calendar quarters; on average the previous-quarter share is roughly twice the current-quarter share. Annualization multiplies the pair's sum by four.
2. **`UTIL` contains `TELEPH`.** Summing summary variables without consulting the variable hierarchy double-counts telephone service.

3. **The 2023 food redesign changes schema mid-window.** Files from 2024Q2 onward drop the `FOOD` and `FDHOME` summaries for `GROCER` (all grocery purchases, food and nonfood); BLS's errata allocates 80 percent of `GROCER` to food at home (U.S. Bureau of Labor Statistics 2026a, 2026c), and BLS applies the same allocation in the official thresholds. Pooled multi-year windows mix the two schemas row by row, so the construction must be per record — a frame-level column check zeroed food for the newest quarters and moved replicated 2025 thresholds four to five percent.
4. **Bundle layout migrations.** CE moved its CSV year-bundles from one directory layout to another with the 2022 release (U.S. Bureau of Labor Statistics 2026a), and interior packaging varies by vintage; per-quarter archive paths from pre-2023 documentation no longer resolve.
5. **The BLS API truncates silently.** Unregistered requests to the public data API cap at ten-year spans and return the *first* ten years of a longer request rather than erroring. A 2005–2025 request comes back as 2005–2014, and downstream code sees a valid, wrong series.
6. **October 2025 CPI does not exist.** The scheduled release covering October 2025 was not published during the federal government shutdown, and the month is absent from the CPI series; 2025 annual averages must be computed over eleven months.

5 Projecting thresholds past the published years

BLS publishes thresholds with roughly a fourteen-month lag: the 2024 thresholds arrived in September 2025 (U.S. Census Bureau 2025), and 2025 thresholds are expected in September 2026. Any analysis of current-law poverty — which for microsimulation models is most analyses — needs thresholds for years BLS has not published. A common approach, and ours until this work, ages the latest published threshold by projected growth in the Consumer Price Index for All Urban Consumers (CPI-U).

That practice has company, and no apparent exception. When Urban Institute researchers projected 2021 poverty during the pandemic, they took the 2019 BLS thresholds and adjusted them for projected inflation (Wheaton et al. 2021). Columbia's monthly poverty framework, which produces sub-annual SPM estimates, declines to project thresholds at all: it carries the latest published ASEC thresholds into subsequent months, on the stated grounds that projecting new ones "requires more timely consumption data" — so its 2020–2021 estimates ran on 2019 thresholds (Parolin et

al. 2022); a 2024 revision updated the framework's base data and income imputation without changing how thresholds are handled (Curran et al. 2025). Han, Meyer, and Sullivan's real-time series prices official-measure thresholds forward with monthly CPI, which is exact for that measure because official thresholds are CPI-indexed by construction (Han et al. 2020; Fisher 1992) — precisely what the SPM's are not. Census's own analysis of the 2024 estimates quantifies what indexation conceals: thresholds re-estimated from consumption rose 5.2 to 5.8 percent that year (on the pre-correction vintage then published) against 2.9 percent for the official measure's CPI-U update, and price-updating the SPM thresholds instead would have put 2024 SPM poverty half a percentage point below the published rate (Creamer 2025). A related tradition price-indexes deliberately: anchored-SPM series hold a reference-year threshold and move it by CPI precisely to separate policy effects from drift in living standards, in historical work (Wimer et al. 2016) and in the National Academies' child-poverty simulations (National Academies of Sciences, Engineering, and Medicine 2019) — a design choice about what to measure, not an estimate of what BLS will publish. We are not aware of prior work that projects the SPM threshold itself by re-estimating it from the consumption microdata — the path Columbia's caveat names — or that commits to numeric threshold values ahead of publication.

Because the thresholds are re-estimated from consumption data each year, price indexation is mismatched in principle, and Table 4 measures the cost in practice. For each year 2020 through 2024 we stand at the prior year's corrected threshold and project one year ahead under four rules: All-Items CPI-U aging; aging by a composite of the CPI components matching the FCSUti basket; scaling by the growth of our replicated thresholds; and a 50/50 blend of the last two. All rules are scored against the corrected actuals.

Table 4: Mean absolute projection error across the three tenure thresholds, by target year, projecting each year from the prior year's corrected values.

Rule	2020	2021	2022	2023	2024	Mean
All-Items CPI-U aging	1.78%	0.92%	1.29%	3.99%	3.16%	2.23%

Rule	2020	2021	2022	2023	2024	Mean
FCSUti-composite CPI aging	0.90%	0.43%	1.41%	2.52%	2.59%	1.57%
CE replication growth ratio	0.65%	0.25%	0.57%	0.26%	0.33%	0.41%
50/50 blend (FCSUti CPI + replication)	0.45%	0.09%	0.54%	1.39%	1.32%	**0.76%**

The direction of the error is the clearest result. All-Items CPI-U aging misses by 2.23 percent per year in absolute value — projection error against the realized thresholds, not the much smaller pure price-index gap between the FCSUti basket and the all-items index — and the misses are mostly on one side: it understated actual threshold growth in four of five years — most sharply in 2023, when the CE window's shelter costs compounded faster than the all-items index — and overstated it only in 2021, for a signed mean error of −1.9 percent per year. The composite price index removes about thirty percent of that error — the FCSUti basket is nearly half shelter — but runs low in all five years. The replication growth ratio, the only rule that tracks consumption rather than prices, posts the lowest mean error, 0.41 percent, and is nearly unbiased, with a signed mean of +0.1 percent per year. The 50/50 blend inherits half of the composite's downward bias and scores 0.76 percent. We fix the blend weights at one half each rather than fitting them: with five backtest years, fitted weights would mostly fit noise, and the equal blend's rationale is structural — the two components fail for different reasons.

Table 4 is itself an amended result, and the amendment bears on how the

blend row should be read. As first released, the composite index behind the second and fourth rows averaged CPI component series at their raw published levels; the components carry different index reference bases, so this weighted them by base level as well as by budget share — shelter received an effective weight near 55 percent rather than the stated 47. Adversarial cross-model review caught the error the day after the 2025 nowcast was first committed. Repairing it left the CPI-U row unchanged, moved the composite rule from 1.40 to 1.57 percent, and — because the replication deflates CE expenditures with the same composite — improved the replication ratio from 1.58 to 0.41 percent and the blend from 1.35 to 0.76. The ranking inverted: the blend had ranked first under the erroneous construction, which is why it is the committed rule; the repaired table puts the replication ratio first. We retain the blend as the primary estimate rather than switching to the new winner, because re-selecting a rule after a second look at five noisy backtest years would be selection on noise, and the blend's structural rationale is unchanged. The September evaluation scores every rule in Table 4 against the realized 2025 thresholds, so the choice costs nothing in falsifiability; the pre-repair table remains in the repository history. The repair then demonstrated the paper's thesis once more: colleague review caught the pre-repair numbers surviving in the packaged nowcast's own method and caveat strings after every table and prose figure had been regenerated — a stale value persisting in exactly the provenance field meant to prevent staleness. Those strings are now pinned by the pre-render check too.

Two caveats on reading the table. First, the two price rules use realized index values — information unavailable when a true forecast is made; a 2022 projection built from consensus 2021 inflation forecasts would have erred by several points. The replication ratio uses CE microdata that BLS publishes well before the corresponding thresholds, so in the nowcasting position the data it needs genuinely exist; that it also scores best in backtest counts in its favor, but five backtest years cannot settle rankings this close. Second, the price rules share a residual: both run low in nearly every year, reflecting real consumption growth beyond what a five-year averaging window passes through and in-kind benefit expansions that prices do not capture. The replication ratio, which tracks consumption, shows no such bias — its annual signed errors range from $+0.5$ to -0.3 percent. With five backtest years we leave the price rules' bias unmodeled rather than fit a constant to it.

6 A pre-committed 2025 nowcast

The 2025 thresholds are a nowcasting target, not a forecasting one — estimation of a present quantity from data already in hand, in the macroeconomic sense of the term (Giannone et al. 2008): their entire CE window (2020Q2 through 2025Q1) and eleven of twelve months of 2025 CPI are published. We apply the blend rule to the corrected 2024 base. The replication component computes thresholds for 2024 and 2025 with identical code over their respective windows and takes the ratio, per tenure — level biases that persist across adjacent, 80-percent-overlapping windows cancel, and the anchor share cancels exactly. The price component is the FCSUti-composite ratio of 2025 to 2024 annual averages, with 2025 averaged over its eleven published months. Table 5 states the estimate.

Table 5: The pre-committed nowcast of 2025 SPM thresholds, two-adult, two-child consumer units, nominal dollars, with per-tenure components. These are model estimates, not BLS publications.

Tenure	Replication ratio	FCSUti CPI ratio	Blend	Nowcast 2025
Owners with mortgages	1.0599	1.0321	1.0460	$41,036.34
Owners without mortgages	1.0443	1.0321	1.0382	$34,135.99
Renters	1.0462	1.0321	1.0392	$40,755.98

The consumption signal moves the estimate materially: replicated thresholds grew 4.4 to 6.0 percent by tenure against 3.2 percent composite price growth, so the blend projects 2025 threshold growth of 3.8 to 4.6 percent, against 2.6 percent realized All-Items CPI-U growth over the same period. Applied in the PolicyEngine US microsimulation model, replacing CPI-U-aged 2025 thresholds with the nowcast raises projected 2025 SPM poverty by 0.4 percentage points (simulation output pinned in the repository); we treat that as a sensitivity, not a headline, since modeled baseline levels carry their own calibration error. To be explicit about adoption: the companion policyengine-us change takes up only the corrected historical series, and that model still ages thresholds past 2024 by projected CPI-U — so its current-law poverty projections carry the understatement quantified in Section 5 until the nowcast is wired into its forecast path, at the measured 2025 cost of those 0.4 points.

These are amended values. The nowcast as first committed was built on the erroneous composite index described in Section 5; repairing the index

moved each tenure's estimate by 0.1 to 0.3 percent — owners with mortgages from $41,099.57 to $41,036.34, owners without from $34,250.70 to $34,135.99, renters from $40,791.72 to $40,755.98. The amendment predates any BLS publication of 2025 thresholds, the original values remain in the repository history, and the September evaluation will report realized errors for both sets.

Evaluation is committed here, in advance — the commitment device of preregistration (Nosek et al. 2018), applied to a numeric estimate rather than a hypothesis test. When BLS publishes the actual 2025 thresholds, expected around September 2026, a revision of this paper will report the signed percent error per tenure and the mean absolute error of this nowcast, alongside the same errors for All-Items CPI-U aging and each component rule, all computed from the corrected 2024 base as committed in the repository — regardless of any subsequent BLS revision to 2024. The packaged estimate is fixed with its method, components, and hash. The blend's annual signed errors in backtest, averaged across tenures, ranged from -0.1 to -1.4 percent; that range is our stated expectation of the nowcast's accuracy. If the nowcast misses by more than the rules scored alongside it, that result publishes too.

7 Discussion

The 2026 correction cost the SPM six years of restated statistics, and the error is less notable for how it arose than for taking five years to surface. The threshold pipeline's inputs are public, its methodology is documented, and its outputs are three numbers a year. Nothing checked them. Our package's larger errors have the same anatomy: values a person typed, tests that enshrined the typing, and no process comparing the artifact to its source. Both failures were detectable by a scheduled diff against the published artifact, which has no false-negative mode for this error class short of the source itself being wrong.

Replication plays a narrower role than adjudication. At its measured fidelity, an independent implementation cannot resolve sub-2-percent errors in the official series, and claiming otherwise would overreach. What it can do is structural: confirm the formula, the window, and the re-anchoring behave as documented; catch transcription-scale errors instantly; and measure the consumption growth that price indexation misses — the use this paper puts it to, where a level bias that cancels in ratios stops being a disqualification. The nowcast is the constructive payoff of infrastructure built for verification.

The verification infrastructure itself needed verifying: the composite index at the heart of the backtest carried a construction error until adversarial review by a model from a different family caught it. The checking layer is code too, with the same failure modes as what it checks.

The projection result has the most direct consequence for microsimulation practice. Aging SPM thresholds by projected CPI-U, as our own model did, understated threshold growth by an average of 1.9 percent per year over 2020–2024, which flows directly into understated current-law poverty projections. Effects on modeled eligibility and poverty transitions near program cliffs are plausible but outside this paper's scope. The blend rule is implementable from public data with a published, tested codebase, and it degrades gracefully: past the CE data horizon it reduces to composite-price aging, which alone removed about thirty percent of CPI-U aging's error in backtest.

The pattern generalizes past poverty measurement. Statistical agencies publish reference series; models embed them; both sides revise — the macroeconomic real-time-data literature documents how consequential vintage differences are for downstream conclusions (Croushore and Stark 2001). Every embedded copy is a place where the numbers can silently diverge from the source, and the failure mode is not unique to official statistics: a spreadsheet error in a widely cited debt-and-growth analysis likewise survived until outside replication (Herndon et al. 2014). The remedy — bundle the source artifact, generate the values from it, diff it on a schedule — required one scheduled job and a parsing script in our case; we do not measure its cost against manual auditing across settings. We found errors on both sides of that interface in the same week. From a single case we cannot say how common such divergences are; our hypothesis is that the interface is ordinary and rarely checked.

8 Reproducibility

Everything in this paper regenerates from public data. The spm-calculator implementation lives on PolicyEngine/spm-calculator pull request #32 (branch `max/bls-2026-threshold-correction`; this revision pins commit `d2746cf`), to be released as version 0.4.0 on merge. It contains the corrected and pre-correction threshold series with the official workbook and its SHA-256; the replication, benchmark, backtest, and nowcast scripts; and the weekly drift-check workflow — which, as a scheduled job, does not run until the branch

merges; its workbook comparison was last run manually on August 6, 2026, and matched all twenty packaged years within half a cent. Regenerating the artifacts from scratch requires the public CE PUMD year bundles (roughly 700 MB, cached locally by the scripts) and the BLS public data API; bls.gov's bot filtering requires a browser-impersonating TLS client, which the scripts document.

The paper repository pins the exact artifacts behind every table and the poverty-sensitivity simulation output, with SHA-256 sums, and bundles the corrected BLS workbook itself (hash matching the file bls.gov served on August 6, 2026), so the correction table is checkable without fighting the bot filtering. Each table is generated from the artifacts by script, and a pre-render check regenerates the tables, verifies the artifact hashes, and re-derives the load-bearing prose figures — the paper does not build if any of them drifts from the data; continuous integration runs the same check on every push. The nowcast's in-advance commitment is externally timestamped, all before any BLS 2025 publication: the artifact checksum file is stamped via OpenTimestamps (proof committed alongside it), the original and amended nowcasts are tagged releases in the repository history, and the Internet Archive holds a capture of the published page from August 7, 2026. The history documents one amendment: the composite-index construction error caught in cross-model review on July 18, 2026, with every downstream artifact, table, and pinned prose figure regenerated in the same change.

References

Betson, David M. 1996. "'Is Everything Relative?' The Role of Equivalence Scales in Poverty Measurement." In *Poverty Measurement Working Papers*. U.S. Census Bureau. https://www.census.gov/library/working-papers/1996/demo/betson-02.html.

Burns, Kalee, and Liana E. Fox. 2021. *Improvements to the Census Bureau's Supplemental Poverty Measure for 2021*. SEHSD Working Paper Nos. 2021-17. U.S. Census Bureau. https://www.census.gov/library/working-papers/2021/demo/SEHSD-WP2021-17.html.

Citro, Constance F., and Robert T. Michael, eds. 1995. *Measuring Poverty: A New Approach*. National Academies Press. https://doi.org/10.17226/4759.

Creamer, John. 2025. *How Updating Annual Poverty Thresholds Impacts Poverty Rates*. U.S. Census Bureau Research Matters blog. https://www.census.gov/newsroom/blogs/research-matters/2025/09/supplemental-poverty-measure.html.

Creamer, John. 2026. *Simplifying SPM Poverty Thresholds: A Ten-Year Anchoring Approach*. SEHSD Working Paper Nos. 2026-09. U.S. Census Bureau. https://www.census.gov/library/working-papers/2026/demo/sehsd-wp2026-09.html.

Croushore, Dean, and Tom Stark. 2001. "A Real-Time Data Set for Macroeconomists." *Journal of Econometrics* 105 (1): 111–30. https://doi.org/10.1016/S0304-4076(01)00072-0.

Curran, Megan, Jiwan Lee, and Zachary Parolin. 2025. *Changes to Our Methodology for Estimating Monthly Poverty Rates in 2024: A Technical Note*. Center on Poverty; Social Policy, Columbia University. https://povertycenter.columbia.edu/sites/povertycenter.columbia.edu/files/content/Data/Monthly/%20SPM/%20Technical/%20Note/%20for/%202024_updated/%2010/%2013/%2025.pdf.

Fisher, Gordon M. 1992. "The Development and History of the Poverty Thresholds." *Social Security Bulletin* 55 (4): 3–14. https://www.ssa.gov/policy/docs/ssb/v55n4/v55n4p3.pdf.

Garner, Thesia I., and Juan D. Munoz. 2021. *Choices in Defining and Estimating Poverty Thresholds: Focus on the U.S. Supplemental Poverty Measure*. U.S. Bureau of Labor Statistics, Division of Price; Index Number Research. https://www.bls.gov/pir/spm/garner_spm_choices_03_15_21.pdf.

Giannone, Domenico, Lucrezia Reichlin, and David Small. 2008. "Nowcasting: The Real-Time Informational Content of Macroeconomic Data." *Journal of Monetary Economics* 55 (4): 665–76. https://doi.org/10.1016/j.jmoneco.2008.05.010.

Han, Jeehoon, Bruce D. Meyer, and James X. Sullivan. 2020. "Income and Poverty in the COVID-19 Pandemic." *Brookings Papers on Economic Activity* 2020 (2): 85–118. https://doi.org/10.1353/eca.2020.0007.

Herndon, Thomas, Michael Ash, and Robert Pollin. 2014. "Does High Public Debt Consistently Stifle Economic Growth? A Critique of Reinhart and Rogoff." *Cambridge Journal of Economics* 38 (2): 257–79. https://doi.org/10.1093/cje/bet075.

Interagency Technical Working Group on Developing a Supplemental Poverty Measure. 2010. *Observations from the Interagency Technical Working Group on Developing a Supplemental Poverty Measure*. https://www.census.gov/content/dam/Census/library/working-papers/2010/demo/SPM_Wkg-Grp.pdf.

National Academies of Sciences, Engineering, and Medicine. 2019. *A Roadmap to Reducing Child Poverty*. National Academies Press. https://doi.org/10.17226/25246.

Nosek, Brian A., Charles R. Ebersole, Alexander C. DeHaven, and David T. Mellor. 2018. "The Preregistration Revolution." *Proceedings of the National Academy of Sciences* 115 (11): 2600–2606. https://doi.org/10.1073/pnas.1708274114.

Parolin, Zachary, Megan Curran, Jordan Matsudaira, Jane Waldfogel, and Christopher Wimer. 2022. "Estimating Monthly Poverty Rates in the United States." *Journal of Policy Analysis and Management* 41 (4): 1177–203. https://doi.org/10.1002/pam.22403.

Schild, Jake. 2023. *Decoding the Surge in the 2022 SPM Threshold: The Influence of Inflation and Other Factors*. U.S. Bureau of Labor Statistics. https://www.bls.gov/pir/journal/2021_2022_spm_analysis.pdf.

U.S. Bureau of Labor Statistics. 2026a. *Consumer Expenditure Surveys Public Use Microdata*. https://www.bls.gov/cex/pumd.htm.

U.S. Bureau of Labor Statistics. 2026b. *Corrected Supplemental Poverty Measure Thresholds*. https://www.bls.gov/pir/spm/spm_thresholds_2024_correction.htm.

U.S. Bureau of Labor Statistics. 2026c. *Experimental Poverty Measures: Research Supplemental Poverty Measure Thresholds*. https://www.bls.gov/pir/spmhome.htm.

U.S. Census Bureau. 2021. *The Supplemental Poverty Measure: 2020*. Current Population Reports Nos. P60-275. U.S. Census Bureau. https://www.census.gov/library/publications/2021/demo/p60-275.html.

U.S. Census Bureau. 2022. *Poverty in the United States: 2021*. Current Population Reports Nos. P60-277. U.S. Census Bureau. https://www.census.gov/library/publications/2022/demo/p60-277.html.

U.S. Census Bureau. 2023. *Poverty in the United States: 2022*. Current Population Reports Nos. P60-280. U.S. Census Bureau. https://www.census.gov/library/publications/2023/demo/p60-280.html.

U.S. Census Bureau. 2024. *Poverty in the United States: 2023*. Current Population Reports Nos. P60-283. U.S. Census Bureau. https://www.census.gov/library/publications/2024/demo/p60-283.html.

U.S. Census Bureau. 2025. *Poverty in the United States: 2024*. Current Population Reports Nos. P60-287. U.S. Census Bureau. https://www.census.gov/library/publications/2025/demo/p60-287.html.

U.S. Census Bureau. 2026. *Census Bureau Statement on Supplemental Poverty Measure*. Press release CB26-RTQ.02. https://www.census.gov/newsroom/press-releases/2026/statement-on-supplemental-poverty-measure.html.

Wheaton, Laura, Linda Giannarelli, and Ilham Dehry. 2021. *2021 Poverty Projections: Assessing the Impact of Benefits and Stimulus Measures*. Urban Institute. https://www.urban.org/research/publication/2021-poverty-projections-assessing-impact-benefits-and-stimulus-measures.

Wimer, Christopher, Liana Fox, Irwin Garfinkel, Neeraj Kaushal, and Jane Waldfogel. 2016. "Progress on Poverty? New Estimates of Historical Trends Using an Anchored Supplemental Poverty Measure." *Demography* 53 (4): 1207–18. https://doi.org/10.1007/s13524-016-0485-7.