CAUSECOMP · RESEARCH

What Actually Satisfies Comparability — and How Far the National Median Misses It

Published 2026-08-23 · CauseComp · causecomp.org/research/comparability-data-in-practice

The §4958 regulations ask boards for pay data from similarly situated organizations — like services, like enterprises, like circumstances. A CauseComp analysis of 732,319 Form 990 filings measures the gap between that standard and the single national median boards most often reach for: in nearly two-thirds of similarly-situated cohorts, the cohort's own median sits more than 25% away from the role's national figure, and the typical gap is 38%.

Legal background: What counts as appropriate comparability data — the CauseComp legal explainer.

Treasury regulations give nonprofit boards a concrete standard for executive pay: compensation is reasonable when it matches what would ordinarily be paid for *like services by like enterprises under like circumstances* (Treas. Reg. §53.4958-4(b)). The rebuttable presumption of reasonableness then asks for appropriate data on that comparison, alongside independent approval and contemporaneous documentation (Treas. Reg. §53.4958-6). What boards frequently rely on in practice is simpler: a single published national figure for the role.

This analysis measures the distance between those two things — the similarly-situated cohort the regulation describes, and the national median a shortcut substitutes for it — across every cohort in the corpus deep enough to measure.

The finding

Define the §4958-shaped cohort the obvious way: same role, same state, same NTEE sector, same budget band. The corpus holds 7,206 such cohorts with at least 20 filings each. **In 63.5% of them, the cohort's own median pay sits more than 25% away from the role's national median. The median gap is 38.0%; one cohort in ten sits more than 139% away.**

The gap shrinks as the cohort definition coarsens — which is exactly the point: each matching dimension a board drops moves it further from the regulation's standard and closer to a number that describes nobody in particular.

Cohort definition	Cohorts (n≥20)	share >10% from national	share >25%	share >50%	median gap
role × state	2,133	54.4%	18.0%	4.6%	11.1%
role × sector	1,173	56.0%	23.3%	8.8%	11.8%
role × budget band	751	75.9%	48.3%	24.6%	23.6%
role × state × sector × band	7,206	84.6%	63.5%	38.4%	38.0%

Read the budget-band row against the state row: dropping geography costs a board far less accuracy than dropping size. A national median is closest to being a fair proxy for a state cohort and furthest from a size-matched one — and size is the dimension the like-enterprises standard most obviously demands.

What this means for a compensation committee

None of this says national figures are dishonest — CauseComp publishes them. It says they answer a different question. “What does this role earn in America?” and “what would like enterprises pay for like services under our circumstances?" differ, in the typical case, by 38% — an error bar wider than most pay decisions' entire range of debate. A committee relying on a single published figure is not wrong to start there; it is wrong to stop there, and the filings themselves say so.

The practice gap, stated plainly

In practice, the comparability files boards assemble rarely resemble what Treas. Reg. §53.4958-6 describes. The common submissions are a single published national figure, a short stack of filings selected after the pay conversation had already settled, or a survey row whose cohort definition no one at the table can state. None of this is negligence in the ordinary sense: assembling a genuinely matched cohort from public filings is slow, and the shortcut looks identical to the real thing until it is challenged. The data above puts a number on the difference — in the typical case the shortcut misses the similarly-situated figure by 38%, an error wider than the raise usually being debated. The working test is one sentence long: a committee that cannot say how its comparables were matched on size, mission and market has documentation that describes a decision, not documentation that supports a presumption.

Method and scope

Scope of every figure on this page: 732,319 Form 990 filings, tax years 2021–2025 (2023 is the last complete filing year; 2024 and 2025 are still filling), covering the 107 curated officer roles CauseComp publishes, filings with a known NTEE sector only. No cohort statistic below 20 filings is reported anywhere.

(a) Cohort medians reflect composition as well as market level: a cohort's distance from the national median blends genuine pay differences with differences in which organizations sit in it. The analysis claims dispersion, not causation.

(b) Gaps are measured between medians of reported total compensation (Schedule J column (E) construction) as filed, unadjusted and unprojected.

(c) The n≥20 floor removes thin cohorts; because thin cohorts are the most dispersed, the reported gaps are, if anything, understated.

(d) A cohort median from this corpus is itself comparability data's raw material, not a substitute for the §4958 process — advance independent approval and contemporaneous documentation are required alongside any data.

Frequently asked questions

Does this mean national medians are useless?

No — they are honest market reads and a sensible starting point. The finding is that they cannot stand in for the similarly-situated cohort the regulations describe: the typical cohort sits 38% away from them.

Why measure against the national median at all?

Because it is what a board without matched data most often reaches for: a single published figure for the role. The analysis quantifies what that shortcut costs in accuracy.

Is a big gap evidence a cohort's pay is wrong?

The opposite — the gaps are real market structure. Budget size, sector and state genuinely move pay, so a cohort matched on them genuinely differs from the national blend.

Free to cite with attribution. For a cohort matched to your organization — budget band, sector, state, with the comparable filers named — see the Executive product.