Does life-satisfaction inequality measure societal inequality?

A focal-value-rounding critique

Chris Barrington-Leigh
McGill University

http://wellbeing.research.mcgill.ca

2026 STATEC

If life satisfaction is a comprehensive welfare measure. . .

- . . . then its *distribution* matters, and reflects differences in health, relationships, and purpose that differences in income alone miss
- So researchers compute inequality statistics directly on the 0–10 response distribution:
  - SD (Goff, Helliwell, and Mayraz, 2018), Gini (Stevenson and Wolfers, 2008),
  - threshold shares, Shannon (What Works Centre for Wellbeing, 2018), ordinal indices (Cowell and Flachaire, 2017).
- Panasiuk, McCanny, and Cheung (2025): life satisfaction is non-normal in all 148 countries — no mean-plus-dispersion summary suffices, and the choice of aggregate *encodes a position in population ethics*.

Quantifying life quality

Question: **Satisfaction with life (SWL)**

"Taking all things into account, how satisfied are you with your life these days, on a scale from 0 to 10?"

0: very dissatisfied
⋮
10: very satisfied

⟶ ***not**: "How are you feeling at the moment?"*

⟶ ***not**: "What makes for a satisfying life?"*

How do people make this *cognitive evaluation*?

Two claims under test

Goff, Helliwell, and Mayraz (2018): country-mean life satisfaction is *lower* where its SD is *higher* (conditional on income) — read as evidence that wellbeing inequality is itself welfare-reducing.

1. **Measurement:** that cross-country variation in a dispersion statistic reflects genuine underlying wellbeing inequality.
2. **Welfare:** that this inequality is welfare-relevant, as the mean–dispersion association suggests.

This paper adds a *new measurement problem* and asks what it does to both.

What is already known about the measures

- **Boundedness:** the 0–10 scale is bounded, so SD may *not* be comparable across countries with different means; Goff, Helliwell, and Mayraz (2018) conclude this problem is modest
- **Cardinality:** results from numeric labels of an ordinal scale can be reversed by monotone relabellings — but likely not in practice (Kaiser and Vendrik, 2020)
- **The ordinal repair:** build indices on the ordinal response CDF — the Cowell–Flachaire family I_0^D, I_0^U of status-inequality indices (Cowell and Flachaire, 2017; Stephen P. Jenkins, 2021).
- **Fragility:** Grimes, Stephen P Jenkins, and Tranquilli (2023) — GHM's negative association vanishes under country/wave fixed effects, and the ordinal indices give *opposite* slopes for up- vs downward-looking.

Focal-value rounding (FVR)

Barrington-Leigh (2024): a substantial, *predictable* share of respondents simplify the 0–10 scale to $\{0, 5, 10\}$.

- Propensity predicted most strongly by **education**.
- FVR breaks not just cardinality but **ordinality**:
- The distortion sits exactly at the **modal categories** — where a dispersion statistic is most sensitive.

Visible in published histograms (e.g. GJT Fig. 4); never modelled in the life satisfaction-inequality literature.

How large can the damage be? A sharp bound

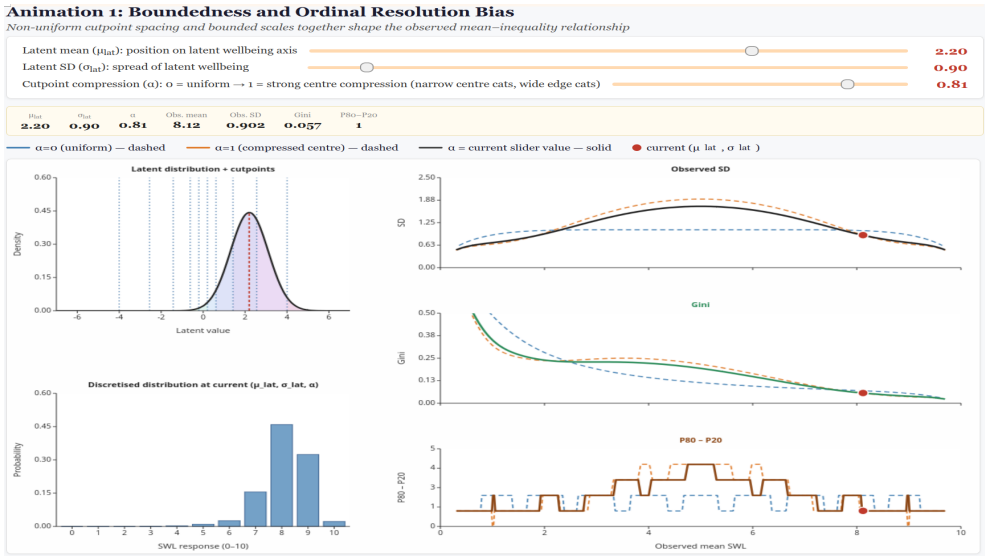
Proposition (Maximum SD bias due to FVR)

A fraction λ of respondents report the nearest element of $\{0, 5, 10\}$ (thresholds 2.5, 7.5) instead of their full-scale integer. Then

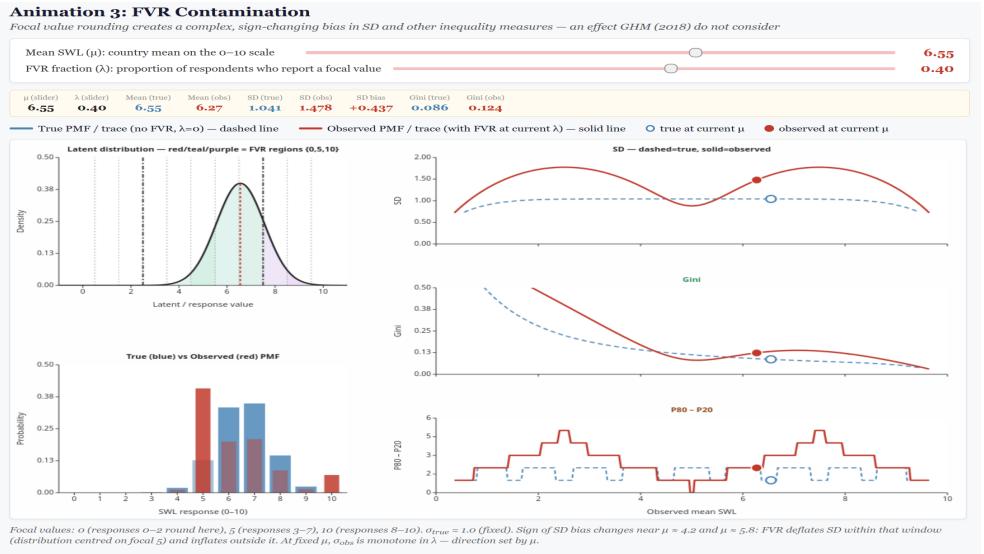
$$|\sigma^{\text{obs}} - \sigma^*| \;\le\; 2\sqrt{\lambda},$$

and the bound is sharp (attained at $\lambda = 1$, in either direction).

- At empirical $\lambda \approx 0.2$–0.4: admissible bias ~ 0.9–1.3 points — comparable to the *entire* cross-country range of measured life satisfaction dispersion.
- The bias can *inflate* or *compress* the SD, depending on where the population sits relative to the thresholds.

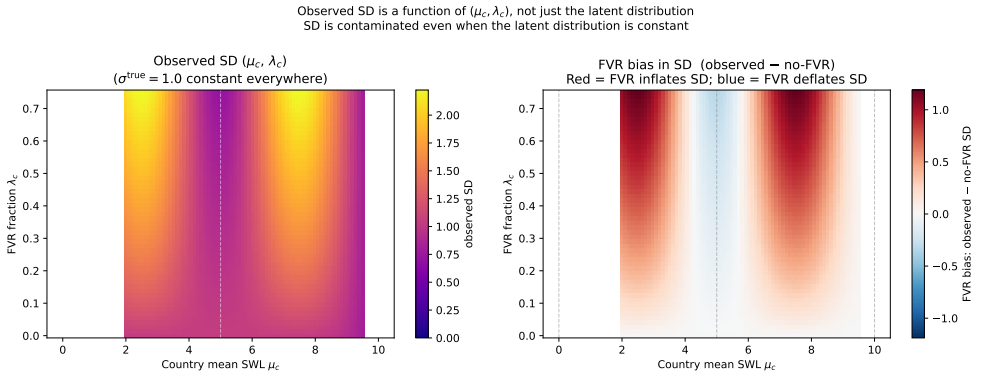


Non-cardinal scales

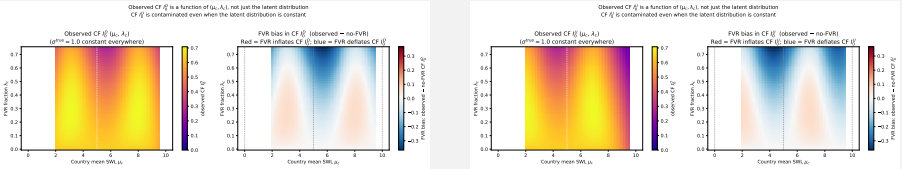


FVR and dispersion

The bias depends on the latent mean — for every statistic



The ordinal indices are *not* protected

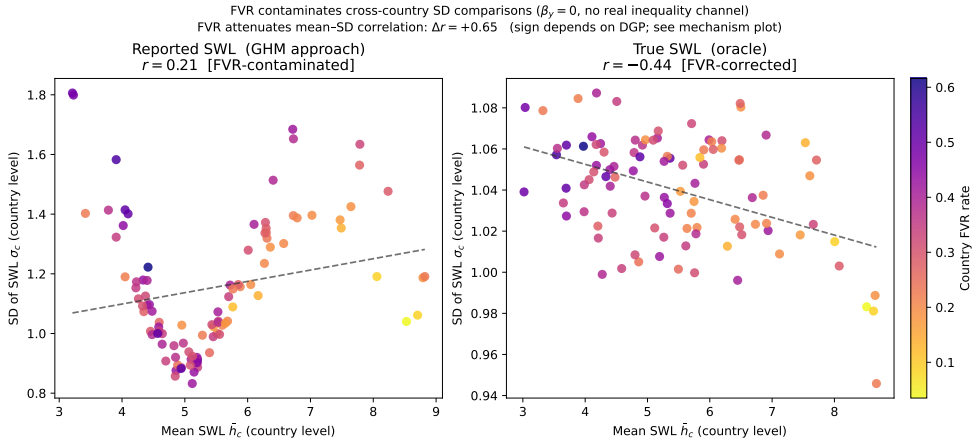


Cowell–Flachaire I_0^D (downward)

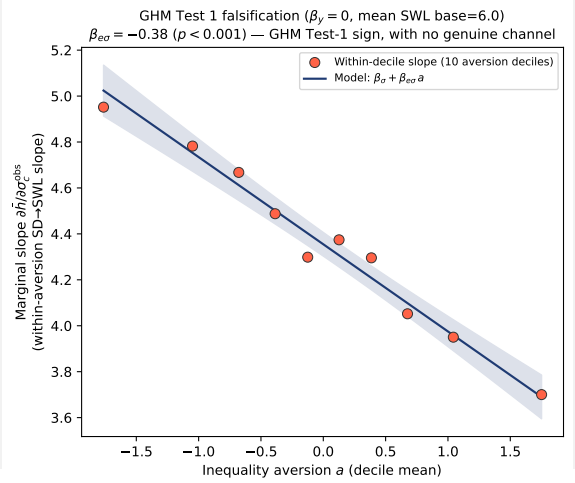
Cowell–Flachaire I_0^U (upward)

Status $s_i = F(u_i)$ is a functional of the CDF F. FVR *reshapes* F at the focal neighbourhoods — so invariance to relabelling buys no invariance to FVR.

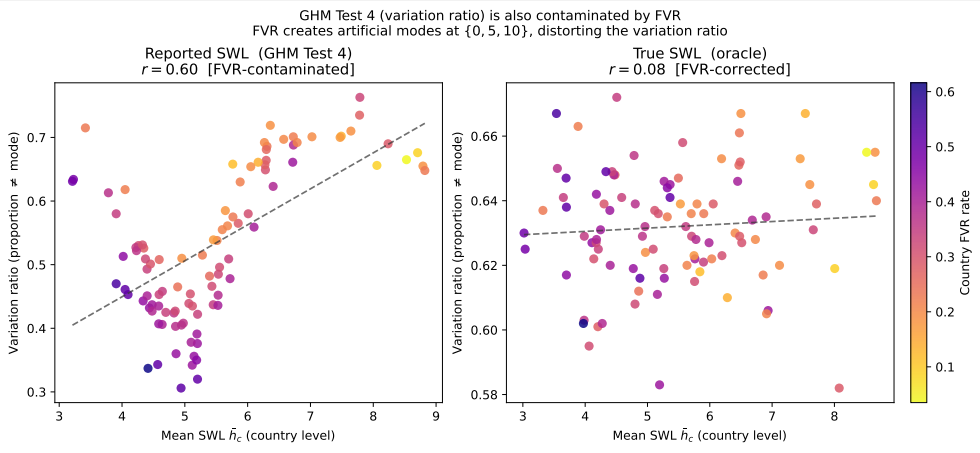
FVR contaminates cross-country SD comparisons



FVR reproduces Goff, Helliwell and Mayraz's *own* robustness checks: Aversion interaction



FVR contamination Goff, Helliwell and Mayraz's *own* robustness checks: Ordinal variation ratio



The reallocation model

An adaptation of Barrington-Leigh (2024) (unpublished form): one rounding propensity per focal value.

- **Latent layer:** ordered probit, latent $\eta_S = X_S^\top \beta_{S,c} + \varepsilon$, with a **hierarchical** per-country scale σ_c.
- **Rounding layer:** for each $f \in \{0, 5, 10\}$, retain the full-scale response near f w.p. $p_i^{H,f}$, else round to f; $p_i^{H,f}$ predicted by education.
- Mass-preserving reallocation; multinomial likelihood; NUTS in PyMC.

Identification: pin the cutpoints

- With *free* cutpoints, the cutpoint bordering a focal value may be incompletely identified from the rounding propensity toward it
- $\rightarrow$ Fix cardinal cutpoints $c_k = k - 5.5$: lets σ_c and the $p^{H,f}$ be read off separately.

Three distributions per country

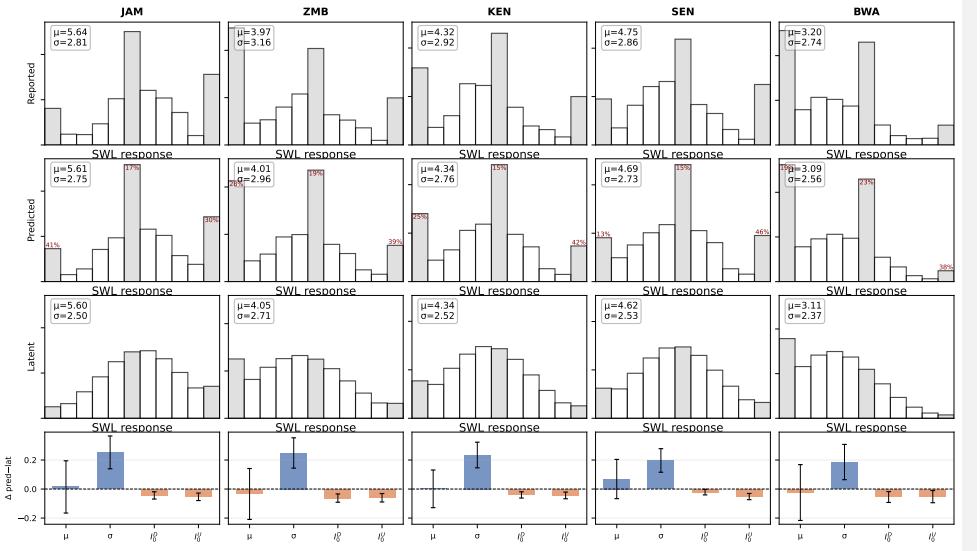
From one posterior:

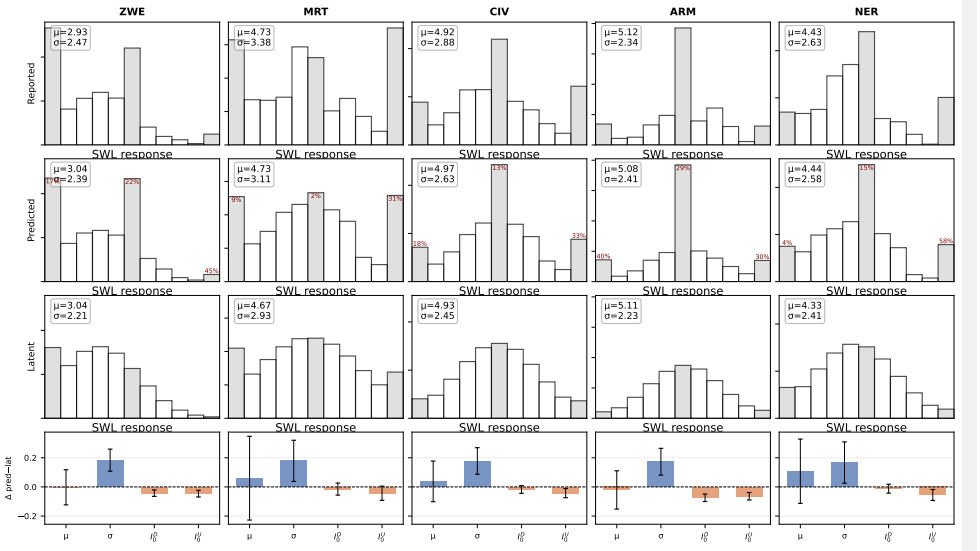
- $\mathbf{P}_c^{\text{obs}}$ — raw empirical PMF.
- $\mathbf{P}_c^{\text{pred}}$ — model's predicted reported PMF.
- $\hat{\pi}_c$ — **latent** PMF: what would be reported if everyone used the full scale ($p^{H,f} \equiv 1$).

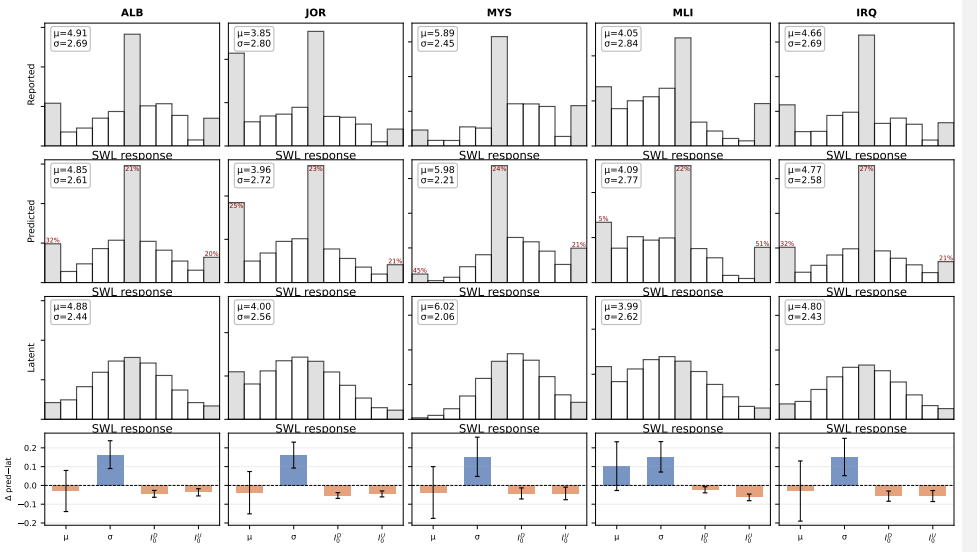
I compare any statistic T on **predicted vs latent** — both from the same posterior, so the comparison isolates the FVR correction and gives exact credible intervals.

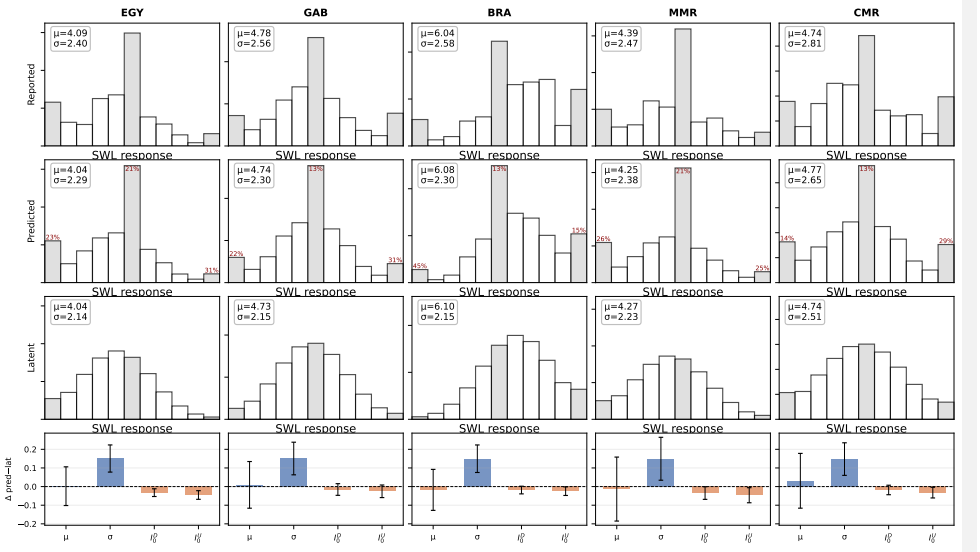
Data

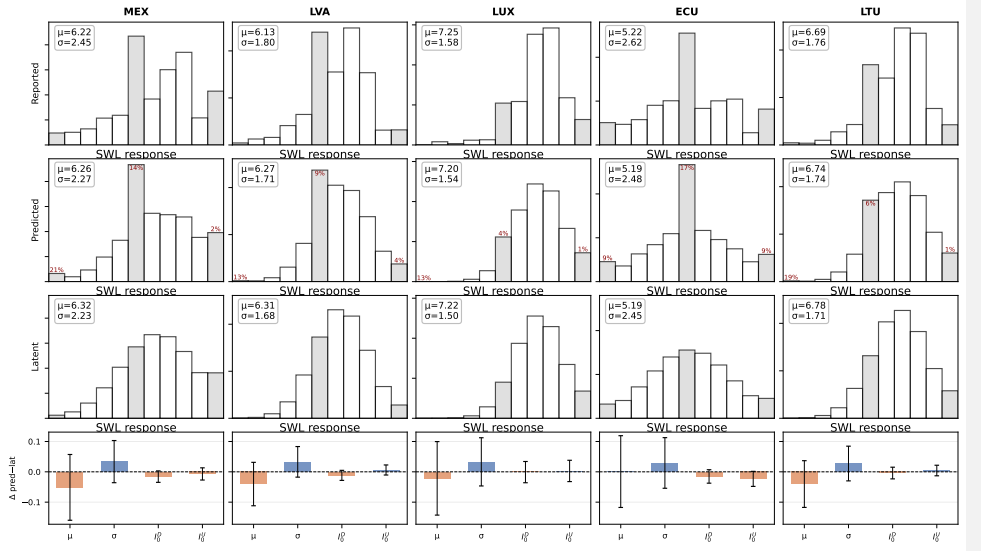
- **Primary:** Gallup World Poll Cantril ladder, 2020–2022 — 135 countries, $N = 300{,}137$.
- **Replications** (appendices, same pattern):
  - GWP satisfaction-with-life item (114 countries, $N = 109{,}375$);
  - Global Flourishing Study life-today & life satisfaction items, one wave, same respondents (22 countries).

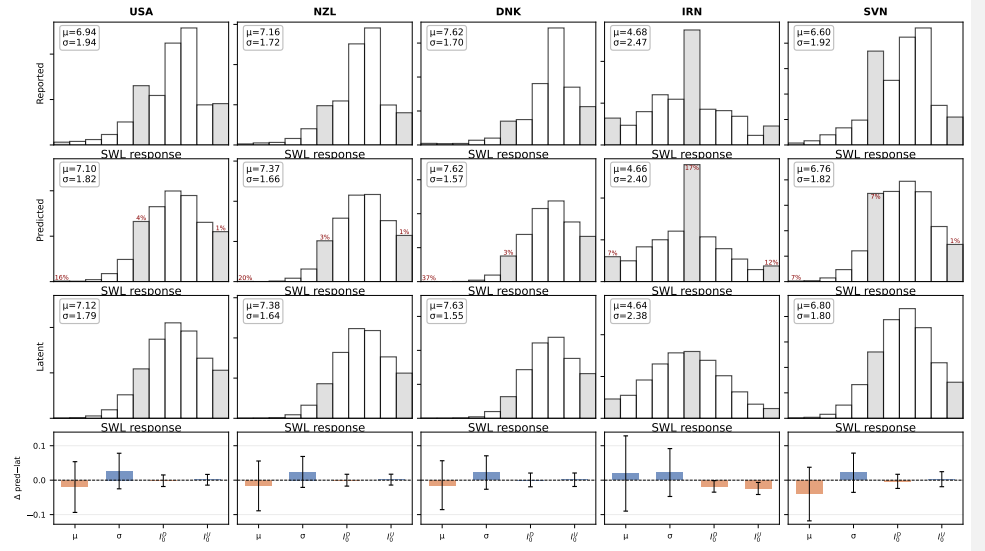

JAM
ZMB
KEN
SEN
BWA
Reported
Predicted
Latent
SWL response
Δ pred−lat

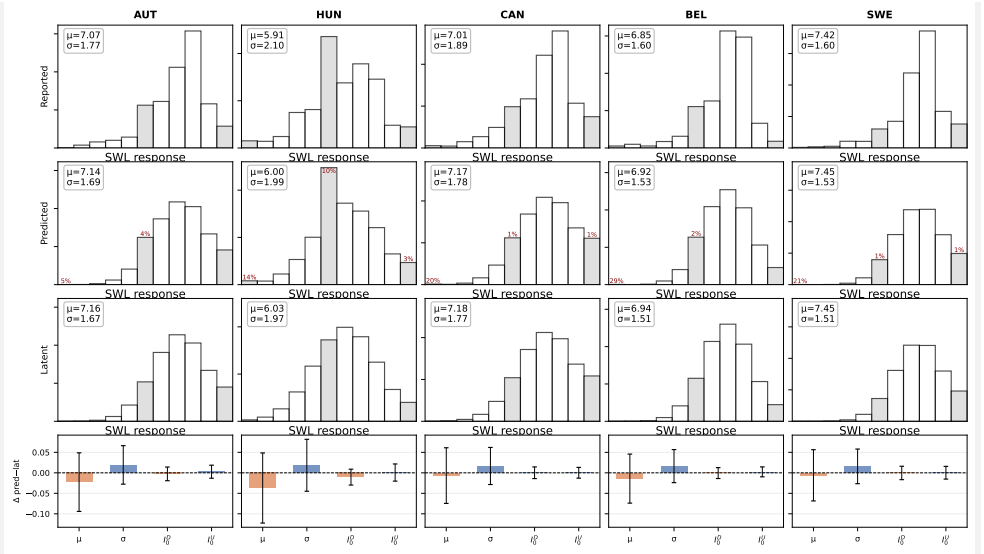

ZWE
MRT
CIV
ARM
NER
Reported
Predicted
Latent
SWL response
Δ pred−lat

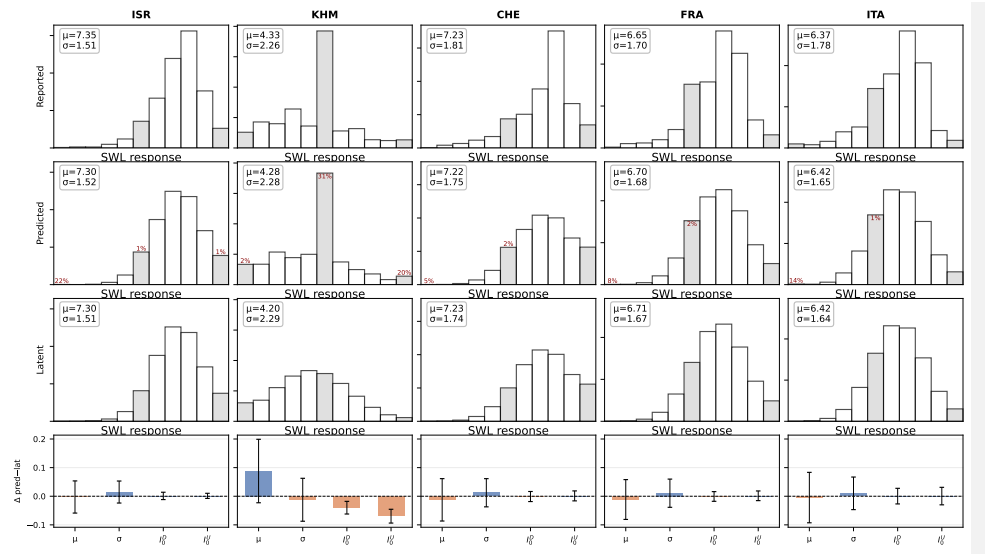

ALB
JOR
MYS
MLI
IRQ
Reported
Predicted
Latent
SWL response
Δ pred−lat

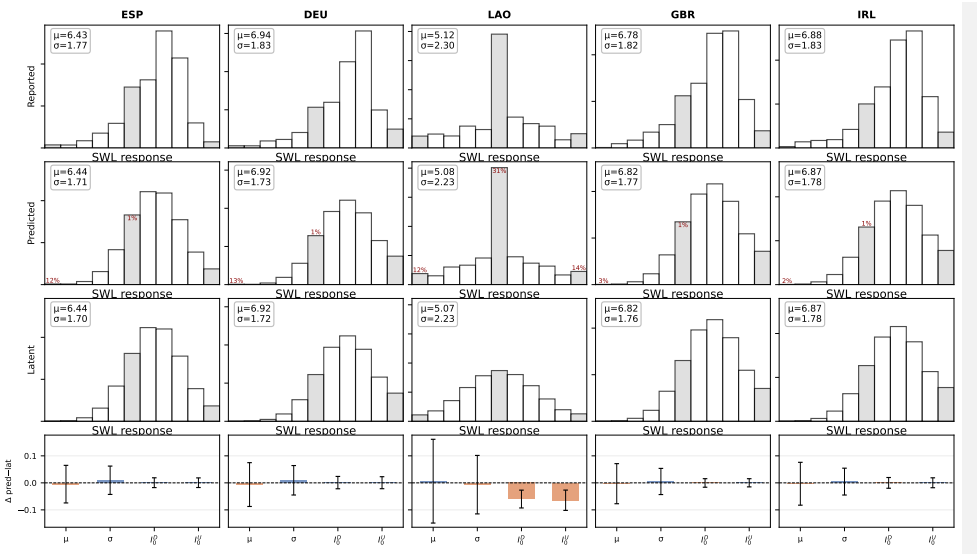

EGY
GAB
BRA
MMR
CMR
Reported
Predicted
Latent
SWL response
Δ pred−lat


MEX
LVA
LUX
ECU
LTU
Reported
Predicted
Latent
Δ pred−lat
SWL response


USA
NZL
DNK
IRN
SVN
Reported
Predicted
Latent
Δ pred−lat
SWL response

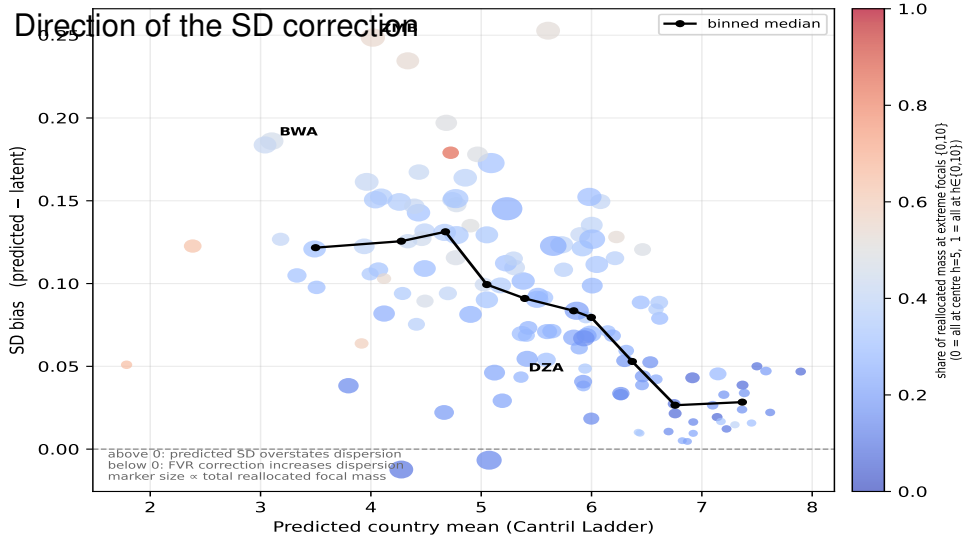

AUT
HUN
CAN
BEL
SWE
Reported
Predicted
Latent
Δ pred−lat
SWL response


ISR
KHM
CHE
FRA
ITA
Reported
Predicted
Latent
Δ pred−lat
SWL response



Result 1 — the correction is material and one-signed

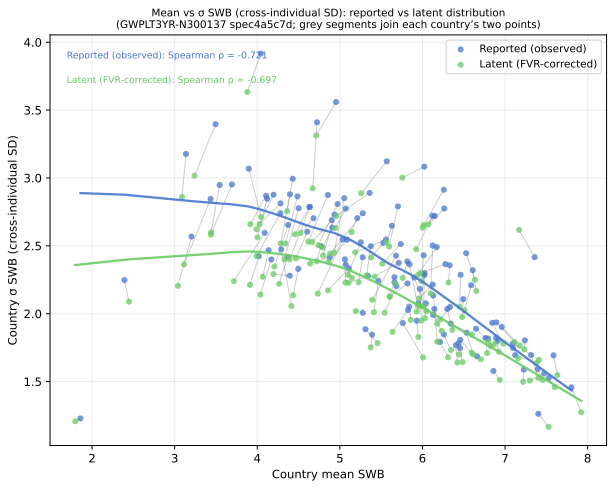
- FVR *inflates* the SD in 133/135 countries; median $|\Delta\sigma| = 0.089$.
- The analytic sign-flip is barely realised: rounding to the extremes (which stretches the distribution) dominates rounding to 5 nearly everywhere.
- The mean is essentially unchanged ($\rho = 0.999$)



Result 2 — the irony of the rankings

- Because the correction is one-signed and partly common, **cardinal rankings survive**:
  - SD: Spearman $\rho = 0.992$ (CI $[0.989, 0.993]$).
- The Cowell–Flachaire **ordinal indices reshuffle the most**:
  - I_0^D: $\rho = 0.875$; I_0^U: $\rho = 0.910$.
- *The statistics advocated for robustness are the least rank-stable under correction* — because they live on the CDF, which FVR deforms at the most-used categories.

Result 3 — the Goff, Helliwell and Mayraz regression survives



Dominance: the correction makes pairs *more* comparable

Stephen P. Jenkins (2021): non-intersection of generalized-Lorenz curves of status ≡ unanimous ranking by the whole Cowell–Flachaire family.

- First-order ranked pairs: 4,778 (predicted) → 6,038 (latent).
- Removing focal spikes stops CDFs crossing ⇒ *more* pairs comparable, with only 0 reversals.

Where dominance holds on the latent distributions, the comparison is robust to every index the theorem covers.

Takeaway

- FVR is real and large in *level*, ubiquitous across surveys, items, and 135-country scale.
- But it does not overturn the cross-sectional story: cardinal **rankings survive**, the Goff, Helliwell and Mayraz coefficient survives with modest attenuation.
- It *does* damage most the **ordinal indices** offered as the cardinality fix — a further reason to distrust *any* single scalar of the reported distribution.

⇒ **Don't scalarise life satisfaction inequality.** Present the distribution; use dominance and targeted group comparisons.

Maybe the problem with trying to find a life satisfaction-inequality scalar is the objective, not the method.

References I

- Barrington-Leigh, C.P. (Feb. 2024). "The econometrics of happiness: Are we underestimating the returns to education and income?" In: *Journal of Public Economics* 230. ISSN: 0047-2727. DOI: https://doi.org/10.1016/j.jpubeco.2023.105052.
- Cowell, Frank A. and Emmanuel Flachaire (2017). "Inequality with Ordinal Data". In: *Economica* 84.334, pp. 290–321. DOI: 10.1111/ecca.12232.
- Goff, Leonard, John F Helliwell, and Guy Mayraz (2018). "Inequality of Subjective Well-being as a Comprehensive Measure of Inequality". In: *Economic Inquiry* 56.4, pp. 2177–2194. DOI: 10.1111/ecin.12582.
- Grimes, Arthur, Stephen P Jenkins, and Florencia Tranquilli (2023). "The relationship between subjective wellbeing and subjective wellbeing inequality: An important role for skewness". In: *Journal of Happiness Studies* 24.1, pp. 309–330. DOI: 10.1007/s10902-022-00591-6.

References II

- Jenkins, Stephen P. (2021). “Inequality Comparisons with Ordinal Data”. In: *Review of Income and Wealth* 67.3, pp. 547–563. DOI: 10.1111/roiw.12489.
- Kaiser, Caspar F and Maarten Vendrik (2020). *How threatening are transformations of happiness scales to subjective wellbeing research?* Tech. rep. IZA Discussion Papers.
- Panasiuk, Sofia L., Anthony McCanny, and Felix Cheung (2025). “Methods Reflect Values: Evaluating the Shortcomings of the Average for Measuring Population Well-Being”. In: *Journal of Personality and Social Psychology* 128.6, pp. 1355–1370. DOI: 10.1037/pspp0000549.
- Stevenson, Betsey and Justin Wolfers (Aug. 2008). “Happiness Inequality in the United States”. In: Working Paper Series 14220. URL: http://www.nber.org/papers/w14220.
- What Works Centre for Wellbeing (2018). *Measuring Wellbeing Inequalities: A How-To Guide*. London: What Works Centre for Wellbeing. URL: https://whatworkswellbeing.org/.