Reconciling Data Years in EEIO Models and Resulting Emission Factors: Alternatives and Best Practices

Dr. Mo Li, Dr. Wesley Ingwersen

June 2026

IIOA Conference

Cornerstone

Why data year reconciliation matters

01

Data Scarcity

Detailed IO tables (IOTs) for a country are often not available for recent, consecutive years. Environmental data may be available more recently but still neither matches the year of spend data used in corporate GHG accounting.

02

Mixed-Year Data Is Common

Pairing economic and environmental data from different years is standard practice in EEIO modeling — but adjustment procedures, their interpretation, and best practices are poorly documented.

03

Temporal Relevance of EFs

Emission factors (EFs) must reflect real changes in GHG intensity and production input structure over time. Using stale data without adjustment can misrepresent real emission trends.

Cornerstone

Objective & Novelty

What this paper does

Describe adjustment procedures
Document common mixed data year integration methods used in EEIO models

Provide clear interpretations
Explain the mathematical and theoretical assumptions behind each approach

Demonstrate effects with real data
Apply all solutions to a US EEIO model time series (2019–2024) across 405 commodities

Establish best practices
Provide guidance for the IO and EEIO community on reconciling data years

Structure: SPSE essay
Situation–Problem–Solution–Evaluation format with quantitative comparisons

The Situation

Goal: Estimate annual GHG emission factors (kg CO2e/USD) for **US** commodities, 2019–2024, using single-region EEIO models in commodity x commodity format.

Relevant datasets available for a US EEIO Model

IO TABLE (Detailed)

Benchmark IOT (Detailed, PRO)

2017 only — released every 5 years with ~5-year delay. Most detailed industry/commodity breakdown available.

IO TABLE (Aggregate)

Aggregate IOT

2017–2024, annual release (~0.8-yr delay). 73 commodities instead of 400

GHG INVENTORY

National GHG Inventory (GHGI)

2017–2023 (annual, ~2-yr delay). GHG totals attributed to industries via sector attribution model to form the E matrix.

Industry Output data

Details Annual Chained-Price Index and Industry Output

2017–2024 (annual, ~0.8m delay) for ~400 industries

The Problem:
No single year provides all the data needed. IOT, GHG inventory, and spend data are each from different years — and none covers the target year (2024).

How do we integrate data from different years without introducing mathematical errors or theoretical inconsistencies?

Matrix/Vector Nomenclature

Selected matrices/vectors relevant for following slides

Matrix or vector	**Definition**	**Basic Derivation**
A	Direct requirements matrix ($/$), commodity-based, industry-tech assumption	Ux-1 Vq-1
Bi	Direct emissions matrix by industry (kg CO2e/$)	E x-1
d	Direct emissions vector by commodity (kg CO2e/$)	Bi Vq-1
E	Environmental (GHG) totals by industry matrix in kg CO2e	Attributed from GHGI to sectors using sector attribution model
e (c)	Environmental totals by commodity in CO2e	d q
L	Leontief total requirements matrix	(I-A)-1
n	Emission factor vector in PUR	BL
q	commodity output vector	iV
U	Use table	given
V	Supply (Make) table	given
x	industry output	given or Vi

Four Solutions to Reconcile Data Years

Solution 1: Common Year Only
Use 2017 data only. No temporal change in EFs. Avoids mixing issues but sacrifices real-world dynamics.

Solution 2: No Adjustment
Latest data used as-is. Annual GHG + annual output, but mixed with 2017 IOT. Creates unit inconsistencies.

Solution 3: Deflation (3a & 3b)
3a: Deflate B to IOT year. 3b: Inflate A to GHG year.

Solution 4: Aggregate Scaling
Scale detailed A, q, x using ratios from annual aggregate IOT.

Solution 3: Deflation — Theory & Math

Solution 3a: Deflate B to IOT year

Use annual GHG data + annual industry output, but deflate output to $2017 before computing B. Then pair B with the 2017 A matrix.

Assumption: Perfect price inelasticity — production input ratios (in $) are constant across years.

Result: n in kg CO2e/USD 2017. Units are consistent. Mathematically valid.

Solution 3b: Inflate A to GHG year

Use annual GHG data + current-year output to build B in current dollars. Deflate A forward to match using price indices.

Assumption: Perfect price elasticity — physical input quantities are held constant; only prices change.

Result: n in kg CO2e/USD current year. Consistent with LCA practice. Also mathematically valid.

Cornerstone

Solution 4: Scaling with Aggregate Data

Uses the annual aggregate IOT to estimate a more current detailed A, q, and x by applying scaling ratios: (aggregate target year / aggregate base year) × detailed base year data.

Step 1: Inflate Aggregate A

Compute scaling ratios (α)

Inflate aggregate A for the target year to base-year ($2017) dollars. Compute ratios α = A_aggregate_target / A_aggregate_base for each cell.

Step 2: Apply to Detailed A

Map to detail schema (O_ds)

Transform α from aggregate to detailed schema using a commodity correspondence matrix O_ds. Multiply element-wise with the 2017 detailed A.

Step 3: Scale V and derive q

Get updated output vector

Apply same scaling to V (make/supply table). Derive the updated commodity output q as column sums of scaled V. Used to compute B = Bi V q^-1 and then n = BL.

Theoretical Evaluation of Solutions/Models

Solutions 1 & 2

Sol 1 (common year): Sound if data scopes match, but sacrifices real trends. Currency conversion may imply false change.

Sol 2 (no adjustment): Mathematically invalid — units of A and B are inconsistent across years.

Solution 3a & 3b

Both are theoretically and mathematically valid. 3a assumes perfect price inelasticity; 3b assumes perfect price elasticity.

Solution 4

Theoretically sound in principle — uses more recent data. But results depend heavily on aggregate table quality. When aggregate data show large inter-annual swings, scaled detailed A inherits those distortions.

Emission Factor Trends

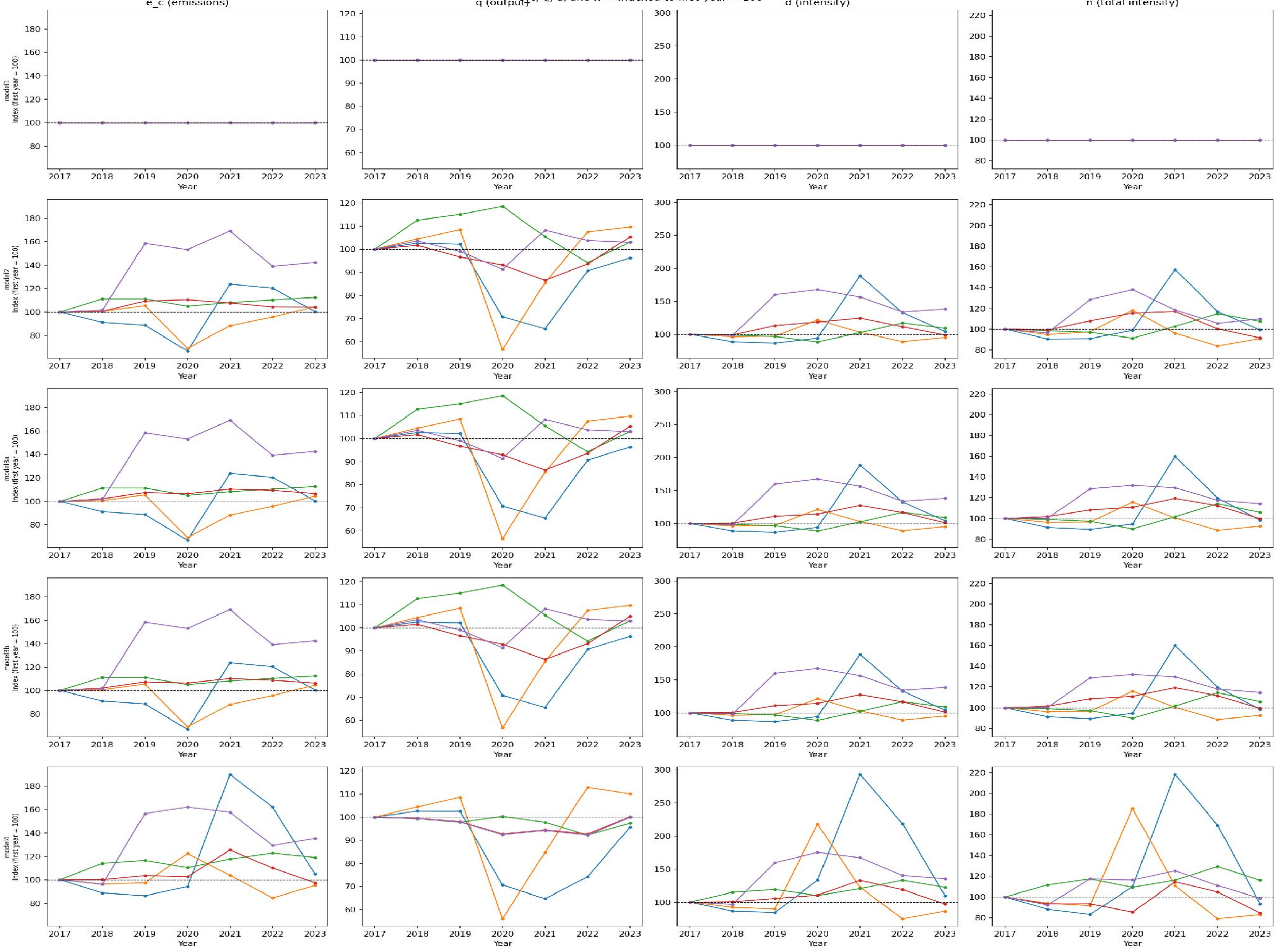
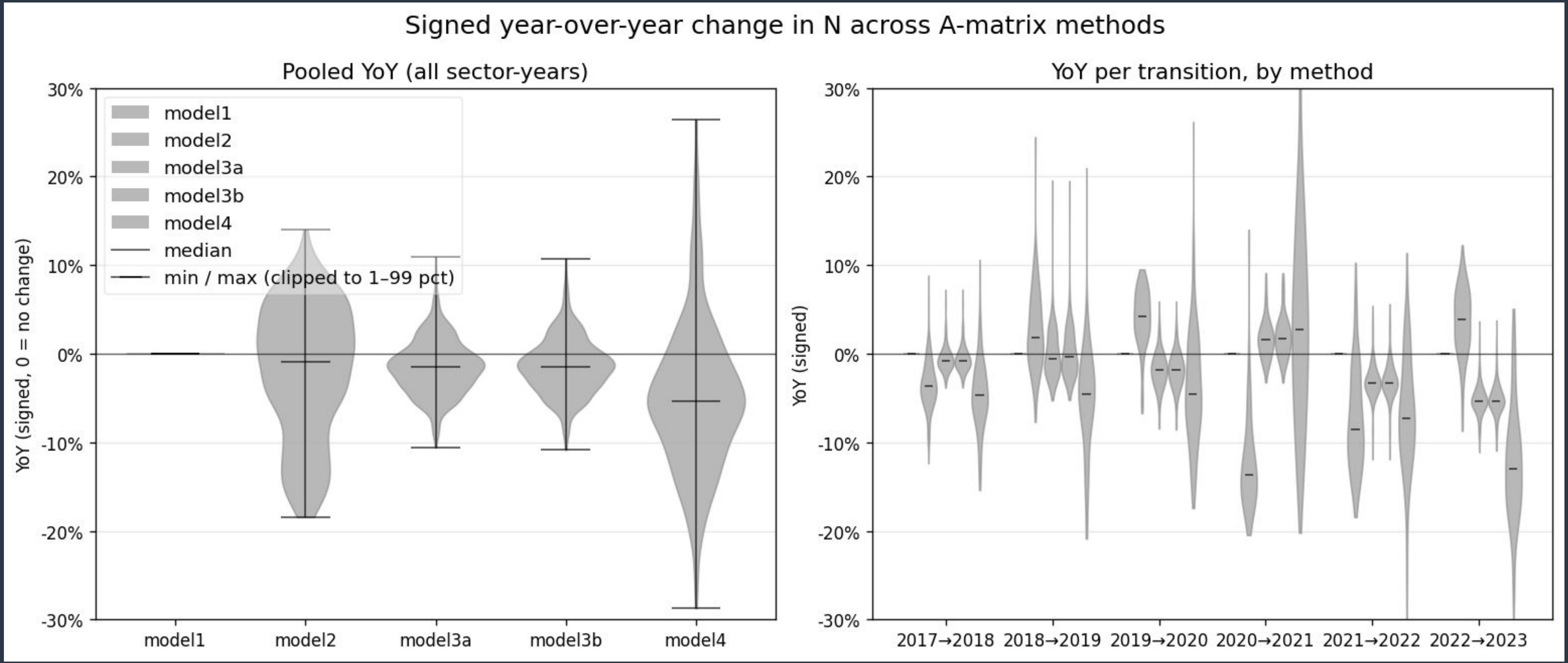


Figure 1 shows e, q, d, and n (indexed to 2017=100) for the 5 commodities with greatest year-over-year variation, across all solutions (2017–2023, expressed in USD 2024).

Stability of EFs



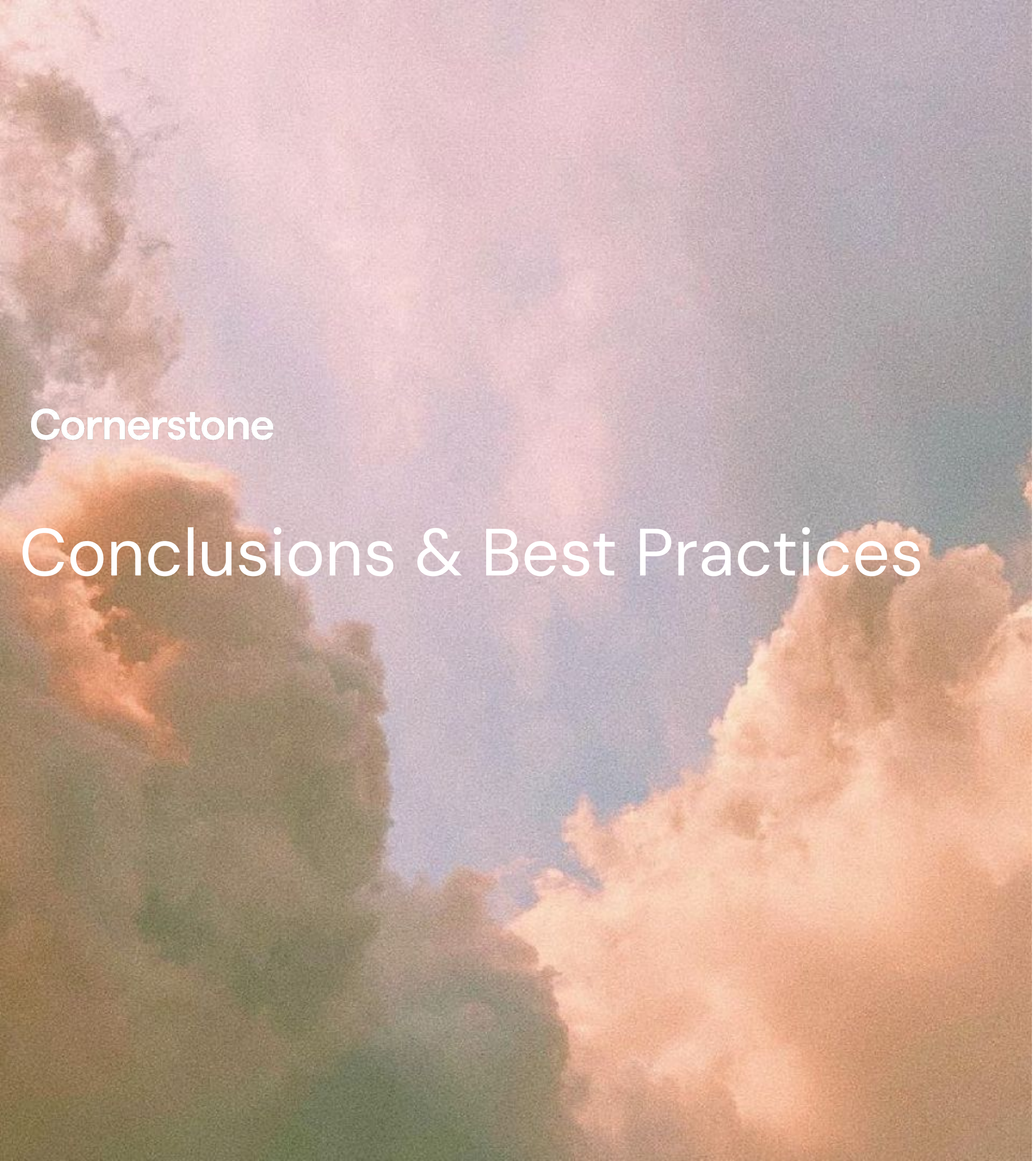
Mean total year-over-year change in *n* (a) by model. Year-over-year total change in *n* in models (b).

Quantitative Evaluation: Summary by Solution

Model 2 (No Adjust.)
Model 2 is mathematically invalid. Mixing 2017 A/L with current-year B creates unit inconsistencies. EFs can be off by up to ~25% vs. the valid models.

Models 3a & 3b
Equivalent results. Deflating B to the IOT year (3a) or inflating A to the GHG year (3b) give the same n vector, despite different theoretical assumptions.

Model 4 (Agg. Scaling)
Instability in some sectors. Air transport doubled in one year; water transport tripled — before reverting. Aggregate data amplifies inter-annual swings.



Cornerstone

Conclusions & Best Practices

Key findings

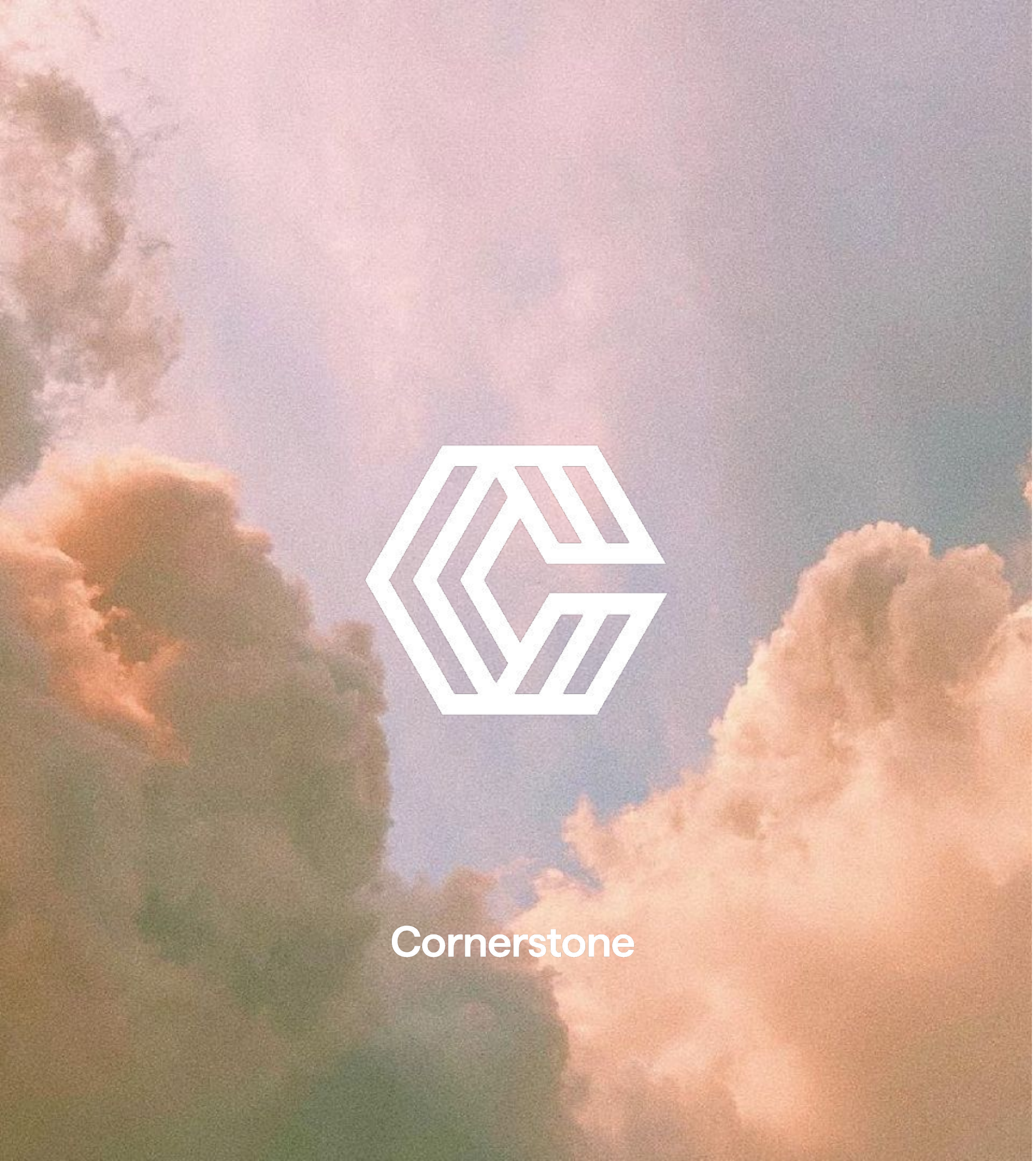
Harmonize data years
Environmental (B) and economic (A) matrices must be in the same monetary year before combining. Failure to do so is mathematically invalid.

Solutions 3a and 3b are equivalent
Both deflation approaches are valid and yield identical n vectors. Choose 3a for EFs in the IOT base year, 3b for EFs in the current GHG year.

Use aggregate scaling with caution
Solution 4 adds contemporary data but can amplify instability when aggregate tables show large inter-annual swings.

Always deflate to reporting year
Final EFs must be converted to the reporting year dollar using commodity-specific price indices before use in organizational footprinting.

Document your approach
Practitioners should clearly communicate which adjustment method was used and what year(s) the EFs represent.



Thank you and stay in touch!

http://cornerstonedata.org

team [at] cornerstonedata.org

https://www.linkedin.com/company/cornerstone-sustainability-data-initiative/

https://github.com/cornerstone-data/