EARL
Economic Accounts
Research Lab

TECHNICAL NOTE

Technical Note 01

Historical Industry Reconciliation

Methodology, Crosswalk, Provenance, and Validation

Historical BEA Evidence, Industry Mapping, the 1997 Splice, and Validation

October 2026
Methodology validated 2 September 2026
Version 1.0

Contents

Executive Summary 2
1 Historical Evidence and Scope 2
2 Historical Industry Crosswalk 3
2.1 Mapping outcomes and publication scope 3
2.2 Combined historical categories and allocation restrictions 4
3 Reconciliation and Provenance 4
3.1 Current-dollar splice and index re-anchoring 4
3.2 Source construction and observation lineage 5
4 Validation and Reproducibility 5
4.1 Historical checks and accounting tolerances 5
4.2 Interpretation limits and supporting evidence 5
References and Methodological Record 6
A Companion Crosswalk and Supporting Evidence 7
A.1 Companion CSV field dictionary 7
A.2 Historical reconciliation evidence inventory 7

Executive Summary

The Economic Accounts Research Lab (EARL) constructs long-run annual Industry Income Statement series by reconciling retained official historical Bureau of Economic Analysis (BEA) Industry Accounts evidence with the current Industry publication structure. Historical and current releases differ in industry classification, available detail, and the reference basis of price and quantity indexes. Historical observations therefore enter the annual series through a reviewed crosswalk and explicit source-precedence rules.

Historical reconciliation follows four governing rules: the current BEA publication taxonomy provides the classification reference; verified historical current-dollar observations are used through 1996, with current BEA observations taking precedence from 1997; eligible historical price and quantity indexes are re-anchored to the current 1997 level without changing their growth paths; and combined historical categories are not allocated to modern child industries without official source evidence.

The companion crosswalk records **101 reviewed mapping outcomes**: 97 ordinary current public-entity mappings, two current-taxonomy references, and two historical combined-category bridges. The mapping inventory is distinct from EARL's **99 annual public Industry entities**. Ninety-nine crosswalk outcomes require no allocation; two explicitly prohibit unsupported allocation. This note documents the source evidence, reconciliation decisions, coverage limitations, and validation record behind those results.

1 Historical Evidence and Scope

BEA's retained Historical Industry Accounts workbooks were released on **November 30, 2021**. Their classification and published detail differ from EARL's current analytical structure, which follows the documented **2017 NAICS-based BEA Industry publication taxonomy**. A historical line may correspond to a current published industry, a supported aggregation of current industries, a renamed concept, or a combined category that the current publication reports separately. These relationships require an explicit source-based mapping rather than a presumed one-to-one correspondence.

The historical and current evidence overlap in **1997**, creating a current-dollar source-precedence decision. Historical price and quantity indexes use a **2012 = 100** reference basis and require a separate re-anchoring to comparable current 1997 levels. Classification, source precedence, and index normalization are therefore distinct methodological questions.

The retained source comprises the four BEA workbooks in Table 1.1. Filenames are reproduced as technical source identifiers, not as alternative publication titles.

Table 1.1 — Retained BEA historical workbooks

Historical BEA account	Retained workbook / dataset	Source years	Reconciliation role
Gross Output	`GDPbyInd_GO_1947-1997.xlsx`	1947–1997	Gross Output levels and historical price/quantity indexes
Intermediate Inputs	`GDPbyInd_II_1947-1997.xlsx`	1947–1997	Intermediate Inputs levels and historical price/quantity indexes
KLEMS / input composition	`GDPbyInd_KLEMS_1963-1997.xlsx`	1963–1997	Energy, Materials, and Services components
Value Added	`GDPbyInd_VA_1947-1997.xlsx`	1947–1997	Gross Value Added and supported value-added components

Note: Source years describe workbook coverage, not complete availability for every series or industry. The current-dollar splice uses historical observations only through 1996.
Source: U.S. Bureau of Economic Analysis, Historical Industry Accounts, November 30, 2021 release; retained EARL source manifest.

Historically supported annual Gross Output, Intermediate Inputs, and Gross Value Added begin as early as 1947; input composition begins in 1963; and selected lower value-added components begin in 1987. Those later components are Compensation of Employees, Taxes on Production and Imports less Subsidies, and Gross Operating Surplus. Availability remains specific to the entity and series; coverage of an aggregate does not imply coverage of every current child industry.

The source concepts remain BEA **Gross Output**, **Intermediate Inputs**, and **Gross Value Added**, even when EARL's Industry Income Statement presents the same observations as **Gross Revenue**, **Cost of Goods Sold**, and **Gross Profit**. Those business-style labels change the analytical presentation, not the source definition or numerical value.

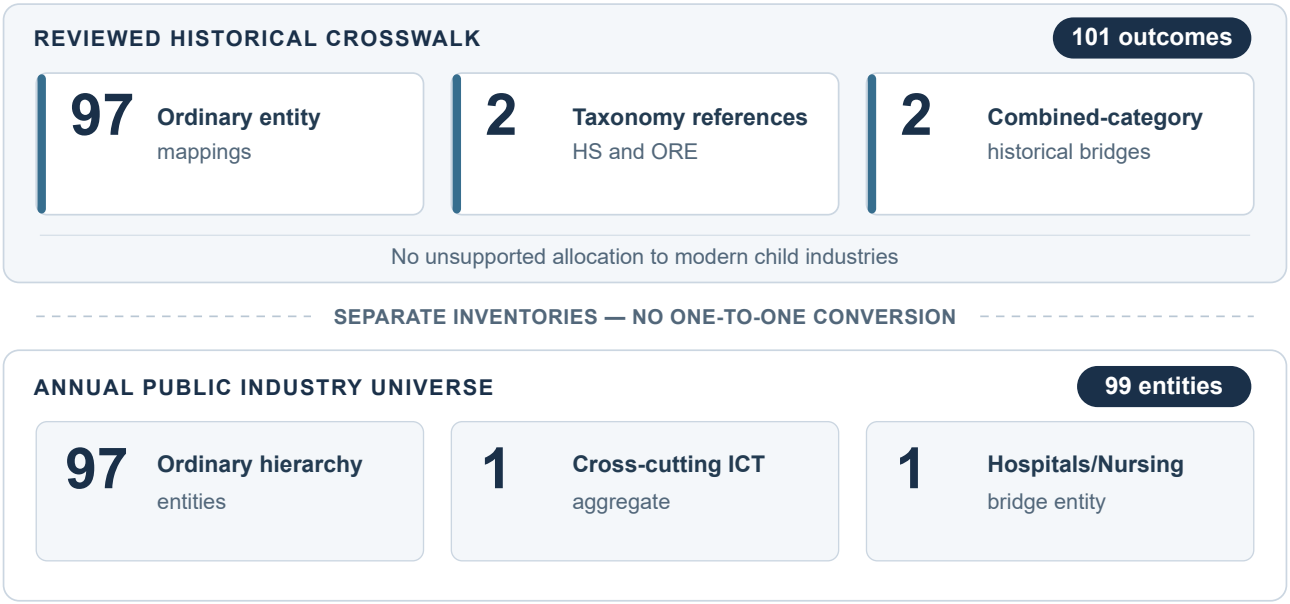
2 Historical Industry Crosswalk

2.1 Mapping outcomes and publication scope

The current BEA publication taxonomy is the classification reference, but the crosswalk and public entity universe are different inventories. The companion CSV contains **101 reviewed mapping outcomes**. Its `target_scope` field identifies 97 rows mapping to ordinary current public entities, two current-taxonomy reference concepts—Housing (`HS`) and Other real estate (`ORE`)—and two historical combined-category bridges.

EARL's annual public Industry universe instead contains **99 entities**: 97 ordinary current-hierarchy entities, a cross-cutting information and communication technology (ICT) aggregate, and a dedicated Hospitals/Nursing historical bridge entity. The 101 mappings must therefore not be counted as 101 published entities. Figure 2.1 displays these separate inventories and the constraint on unsupported allocation.

Figure 2.1 — Separate historical-to-current mapping outcomes and annual public Industry entities



Note: The inventories serve different purposes and are not a row-by-row conversion. The 97 ordinary mappings are one part of the crosswalk; the ICT aggregate and the dedicated Hospitals/Nursing bridge belong to the separate annual public-entity universe.
Source: EARL Technical Note 01 reviewed companion crosswalk and annual Industry entity registry.

The relationship groups in Table 2.1 summarize the crosswalk's nine machine-readable `mapping_type` values.

Table 2.1 — Historical mapping relationships

Historical-to-current relationship	Rows	Interpretation
Direct current description	80	Historical concept corresponds to a current published description.
Current-children aggregate	8	Historical line maps to a supported aggregate of current children.
Contextual current summary	3	Historical summary corresponds to the current summary node.
Supplemental aggregate	3	Historical aggregate remains available as supplemental analytical information.
Renamed current description	2	Economic concept continues under a different current description.
Historical combined category; current split unavailable	2	Historical line combines modern children; no unsupported split is constructed.
Other exact analytical/current mapping	3	Exact mapping to an analytical total, aggregate, or current sector.
Total	**101**	**All reviewed crosswalk outcomes.**

Note: Seven reader-facing relationship groups summarize nine `mapping_type` values. Counts refer to crosswalk outcomes, not the 99 annual public Industry entities.
Source: EARL Technical Note 01 companion crosswalk, reviewed mapping inventory.

Renamed mappings preserve the historical source description and identifier while using the corresponding current term for interpretation. A mapping does not imply that classifications, source detail, or industry coverage were identical in every historical year.

2.2 Combined historical categories and allocation restrictions

The mapping decision and allocation policy are separately recorded. Of the 101 mapping rows, **99 have allocation policy `NONE`**, while the two combined historical categories in Table 2.2 carry **`PROHIBITED_WITHOUT_OFFICIAL_SOURCE`**. The historical BEA publications do not identify the modern child allocations needed to divide those combined concepts.

Table 2.2 — Combined historical categories requiring bridge treatment

Historical BEA category	Current BEA structure	Reconciliation treatment
Real estate	Housing and Other real estate (`HS|ORE`)	Preserve the combined historical concept; do not allocate retrospectively between current children.
Hospitals and nursing and residential care facilities	Hospitals and Nursing/residential care (`622|623`)	Retain the historical bridge concept; do not infer separate historical child values.

Note: `HS|ORE` and `622|623` identify the modern components; neither is an instruction to allocate historical values. The two bridge rows intentionally have no single `target_entity_id`.
Source: Retained BEA historical workbooks and the validated EARL historical crosswalk.

3 Reconciliation and Provenance

3.1 Current-dollar splice and index re-anchoring

For eligible current-dollar Industry series, EARL uses verified historical observations through **1996** and current BEA observations from **1997 onward**. The overlap is not averaged or interpolated to produce a smoother series. This precedence applies at the supported entity and series scope. The same overlap year serves a different role for historical indexes, as Table 3.1 distinguishes.

Table 3.1 — Historical/current splice and 1997 overlap

Historical/current interval	Current-dollar treatment	Index treatment
Through 1996	Use verified historical observations where supported.	Retain eligible historical growth, rescaled to the current 1997 level.
1997 overlap	Current observation takes precedence.	Compare historical and current 1997 indexes to establish an eligible anchor.
From 1998 onward	Use current BEA observations.	Use current BEA index observations.

Note: A historical price or quantity index is re-anchored only when an exact comparable historical/current 1997 pair exists. The table does not imply full historical coverage for every entity and series.
Source: Validated EARL historical Industry reconciliation rules.

For a historically supported price or quantity index I_t with a comparable 1997 pair, EARL preserves the historical index path while expressing its level on the current reference basis:

$$I_t^{\mathrm{EARL}} = I_t^{\mathrm{hist}} \left(\frac{I_{1997}^{\mathrm{current}}}{I_{1997}^{\mathrm{hist}}} \right), \qquad t < 1997. \tag{1}$$

Here I_t^{hist} is the published historical index. The multiplier is fixed across the historical segment, so it **changes the index level but not period-to-period growth**. The retained reference inventory contains **878 verified 1997 index-anchor records** across the eligible historical price and quantity series.

The Hospitals/Nursing bridge has no exact current 1997 anchor for four index series associated with Gross Output and Intermediate Inputs. Consequently **136 observations**, comprising four series across 34 years from **1963 through 1996**, are deliberately omitted from the re-anchored history. An index for a combined historical category is neither allocated to children nor assigned a synthetic 1997 level when that anchor is missing.

3.2 Source construction and observation lineage

Classification differences also limit other historical series. A published historical parent does not establish data for a modern child. EARL does not interpret missing children as zero, interpolate unpublished industry shares, or substitute unsupported allocations. Comparisons over long horizons must respect the entity, source concept, and period for which an observation exists.

Historical mappings do not independently authorize construction of new public hierarchy totals. Where directly reported BEA Gross Value Added exists, EARL retains it. Only for the specified all-industries scope, when the required aggregate is not directly available in the relevant source table, the validated source-side construction permits

$$\text{Gross Value Added} = \text{Gross Output} - \text{Intermediate Inputs}. \tag{2}$$

The resulting observation carries **`derived_fallback`** provenance rather than being described as directly BEA-reported. EARL presents Gross Value Added as Gross Profit without altering the identity. This calculation does not authorize reconstruction of unavailable historical child industries.

Observation-level provenance distinguishes retained historical BEA observations (`bea_historical_2021`) from current BEA observations (`bea_api_current`). Source files, acquisition records, build metadata, and stable `dataset_id`, `entity_id`, and `series_id` fields support tracing across the 1997 boundary. The reviewed crosswalk records classification decisions separately from the observation series, allowing users to distinguish mapped history, source coverage limitations, and documented exclusions.

4 Validation and Reproducibility

4.1 Historical checks and accounting tolerances

The validated historical suite passed **nine core checks** covering observation-key uniqueness and completeness; Gross Output–Intermediate Inputs–Gross Value Added relationships; Intermediate Inputs' Energy, Materials, and Services decomposition where supported; the lower value-added relationship among Compensation, Taxes less Subsidies, and Gross Operating Surplus; corresponding Gross Operating Surplus reconciliation; index-level normalization; historical index-growth preservation; and source provenance.

In addition, **11 selected historical Gross Output observations** were directly compared with the retained official BEA workbooks; all matched exactly. These selected checks do not represent a claim that every source observation received a manual workbook comparison.

Most selected historical production and input-decomposition identities use a **$150 million absolute tolerance** to accommodate BEA source rounding. Two aggregate, cross-table Gross Value Added checks use a **$250 million threshold** for independently published inputs; the largest residual recorded in those checks was **$200 million** on a multi-trillion-dollar aggregate. The thresholds apply to the specified identities, not to general data adjustment or unsupported industry allocation.

4.2 Interpretation limits and supporting evidence

The reconciliation preserves the source evidence's definitions and scope. Historical descriptions can differ from current names; combined historical categories may lack modern child detail; and re-anchoring an index changes its level without creating historical observations. Changing the BEA taxonomy, a mapping, an index anchor, a splice decision, or a validation rule requires separately documented methodological review and renewed validation.

The companion file, **`EARL-Historical-Industry-Crosswalk.csv`**, contains the 101 historical mapping outcomes, source and target descriptions, allocation policies, and review statuses. It is a **classification reference**, not a second observation dataset. Appendix A defines its CSV fields and inventories the associated source and validation evidence, distinguishing reader-facing materials from retained records. The historical QA results and selected source checks are distinct from downstream publication or repository-deployment checks.

References and Methodological Record

Primary source. U.S. Bureau of Economic Analysis, *Historical Industry Accounts*, retained workbooks from the November 30, 2021 release; current GDP by Industry publication evidence as identified in the relevant EARL build.

Companion reference. EARL, `EARL-Historical-Industry-Crosswalk.csv`, reviewed 101-row historical-to-current mapping reference for Technical Note 01.

Related methodology. EARL, *EARL Methodology: Constructing Business-Style Economic Accounts from U.S. Economic Data*, Version 1.0, October 2026, Section 11.

Suggested citation. Economic Accounts Research Lab (2026), *Historical Industry Reconciliation: Methodology, Crosswalk, Provenance, and Validation*, EARL Technical Note 01.

The historical-reconciliation methodology was validated **September 2, 2026**. This validation date and Technical Note Version **1.0** identify the documented methodology; a presentation-only revision does not revise the historical validation evidence.

A Companion Crosswalk and Supporting Evidence

Appendix A provides the field definitions and reference inventory required to interpret the companion CSV. Field names are preserved as published machine identifiers; their roles are described in the main text.

A.1 Companion CSV field dictionary

Table A.1 defines the companion's fields. Empty `target_entity_id` values are intentional for current-taxonomy references and historical bridges without a single current public-entity target.

Table A.1 — Companion crosswalk field definitions

CSV column	Definition
`historical_line`, `historical_title`	Historical BEA line number and original industry description.
`go_earliest`, `go_latest`, `go_n_years`	Reported Gross Output coverage metadata for the historical line.
`source_release`, `source_reference_year`	Retained source-release date and historical index reference basis.
`target_classification`, `target_taxonomy_level`	Reference BEA taxonomy and target classification level.
`target_scope`	Current public entity, current-taxonomy reference concept, or historical bridge.
`target_code`, `target_description`	Target code and reader-facing mapped description.
`target_entity_id`	Current public Industry entity identifier, blank when no one-to-one entity exists.
`legacy_semantic_code`	Historical implementation identifier preserved for traceability.
`mapping_type`	Recorded historical-to-current relationship category.
`allocation_policy`	Whether a historical allocation is required or prohibited.
`historical_comparability`	Documented comparability of the historical concept.
`review_status`	Status of the reviewed historical mapping.

Note: Published field names are retained verbatim. This dictionary describes the 101-row crosswalk, not the canonical observation schema.
Source: EARL Historical Industry Crosswalk, Technical Note 01 companion CSV.

A.2 Historical reconciliation evidence inventory

Table A.2 distinguishes public reader-facing materials from retained validation resources. Listing an evidence artifact does not imply that it is separately distributed with the Technical Note.

Table A.2 — Reconciliation documentation and evidence inventory

Resource	Evidentiary role	Publication treatment
Technical Note 01 and 101-row crosswalk	Methodology and reviewed classification outcomes	Reader-facing companion materials
Historical BEA workbooks and source manifest	Official source values, release identification, and file hashes	Retained historical evidence
Current BEA taxonomy and Industry registry	Current entity descriptions and hierarchy	Governed taxonomy/reference record
Verified 1997 anchor inventory	Comparisons used for eligible index re-anchoring	Retained reconciliation evidence
Documented index omissions	Exclusions without a comparable current anchor	Retained reconciliation evidence
Observation provenance and build metadata	Source route and construction identity of canonical observations	Relevant governed public metadata and build records
Historical QA and direct source-comparison records	Identity, growth, and selected workbook validation results	Retained validation evidence

Note: Inclusion in the inventory records evidentiary role, not a commitment to distribute each retained technical artifact.
Source: EARL Technical Note 01 historical-reconciliation source and validation inventory.