EARL
Economic Accounts
Research Lab

METHODOLOGY PAPER

EARL Methodology

Constructing Business-Style Economic Accounts from U.S. Economic Data

Data Sources, Harmonization, Statement Construction, Historical Reconciliation, Provenance, and Validation

October 2026
Version 1.0

Contents

Abstract 4
Methodological Status and Versioning 4
Scope of this version 4
Documented data build 4
Source vintages 4
Relationship to Technical Note 01 4
Project naming and revision policy 4
1 Introduction and Scope 5
1.1 Motivation and analytical contribution 5
1.2 Scope of EARL 5
1.3 Methodological approach 6
1.4 Coverage and interpretation boundaries 6
2 Conceptual Framework and Terminology 6
2.1 Analytical translation 6
2.2 Concept classes, measurement basis, and provenance 7
2.3 Presentation and terminology 7
3 Source Data and Metadata Governance 8
3.1 Official sources 8
3.2 Metadata governance 8
3.3 Source vintages, historical evidence, and change control 9
4 Canonical Data Architecture 9
4.1 Architecture and source evidence 9
4.2 Harmonization and semantic construction 9
4.3 Public data and publication products 10
5 Data Acquisition and Raw Evidence 10
5.1 Retrieval and raw captures 10
5.2 Acquisition metadata and source integrity 10
6 Source Harmonization and Reconciliation 10
6.1 Matching and normalization 10
6.2 Missingness, source precedence, and source-side constructions 11
6.3 Historical reconciliation and harmonized output 12
7 Semantic Statement Construction 12
7.1 Semantic mapping and formula evaluation 12
7.2 Provenance, presentation, and quality assurance 12
8 National Income Statement Methodology 13
8.1 Scope and production block 13
8.2 Operating expenses and EBITDA 13
8.3 EBIT, EBT, margins, and supplemental income 14
8.4 Sources and availability 14
9 Corporate Income Statement Methodology 14
9.1 Scope and sources 14
9.2 Operating expenses, EBITDA, EBIT, and EBT 15
9.3 Net income, valuation adjustments, and ratios 15
9.4 Frequency and availability 16
10 Industry Income Statement Methodology 16
10.1 Scope and entity architecture 16
10.2 Production, input structure, and operating surplus 16
10.3 Labor, financing, and aggregation 17

10.4 Historical extension, frequency, and availability . . . 17
11 Historical Industry Reconciliation . . . 18
11.1 Historical source and taxonomy . . . 18
11.2 Current-dollar splice and index re-anchoring . . . 19
11.3 Unsupported splits, lower-statement history, and validation . . . 19
12 Frequency, Units, Availability, and Comparability . . . 20
12.1 Frequency, units, and availability . . . 20
12.2 Cross-frequency and historical comparability . . . 21
13 Public Data Structure and Metadata . . . 21
13.1 Public observations and scope . . . 21
13.2 Metadata, statement layout, and provenance . . . 22
13.3 Public package and deterministic output . . . 22
14 Provenance, Vintages, and Revisions . . . 22
14.1 Source release, acquisition, build, and publication . . . 22
14.2 Build identity and observation provenance . . . 23
14.3 Revisions and historical provenance . . . 23
15 Validation and Quality Assurance . . . 23
15.1 Validation architecture and source controls . . . 23
15.2 Accounting and historical validation . . . 24
15.3 Revision, public-data, and publication validation . . . 24
16 Interpretation, Reproducibility, and Governance . . . 24
16.1 Interpretation boundaries . . . 24
16.2 Revisions and reproducibility . . . 25
16.3 Governance, identifiers, and scope . . . 25
17 Conclusion . . . 25
A Complete Public Series Dictionary . . . 27
A.1 Complete dictionary . . . 27
A.2 Interpretation notes . . . 35
B Source Table and Line Crosswalk . . . 36
B.1 Public source-reported series crosswalk . . . 36
B.2 Matching rules . . . 42
B.3 Source families used by the methodology . . . 42
C Formula Reference . . . 43
C.1 Formula language . . . 43
C.2 Public calculated formulas . . . 43
C.3 Source-side constructions and semantic formulas . . . 44
C.4 Formula validation . . . 44
D Statement Layouts . . . 45
D.1 Complete layout contract . . . 45
D.2 Intentional repeated rows . . . 51
E Industry Entity Architecture . . . 52
E.1 Hierarchy interpretation . . . 52
E.2 Complete Industry entity dictionary . . . 52
E.3 Special entities and constructions . . . 55
F Historical Reconciliation Reference . . . 56
F.1 Retained official source evidence . . . 56
F.2 Crosswalk architecture . . . 56
F.3 Current-dollar splice and index re-anchoring . . . 57
F.4 Validation record . . . 58
G Public Data Schema . . . 59
G.1 Observation grain and keys . . . 59
G.2 Observation provenance . . . 59
G.3 Public schema dictionary . . . 60

G.4 Shared metadata files . . . 62
G.5 Absence and sign semantics . . . 62
H Provenance and Vintage Schema . . . 63
H.1 Principal identifiers . . . 63
H.2 Acquisition manifest schema . . . 64
H.3 Current acquisition inventory . . . 64
H.4 Current official source vintage . . . 65
H.5 Documented build identity . . . 65
H.6 Historical provenance branch . . . 65
I Validation Matrix . . . 66
I.1 Consolidated validation matrix . . . 67
I.2 Same-vintage adjudication statuses . . . 67
J Repository and Reproducibility Mapping . . . 69
J.1 Repository authority hierarchy . . . 69
J.2 Shared pipeline modules . . . 70
J.3 Six dataset pipelines . . . 70
J.4 Public data package and publication copies . . . 70
J.5 Rebuild sequence . . . 70
J.6 Documented public file inventory . . . 71
J.7 Reproducibility checklist . . . 71

Abstract

The **Economic Accounts Research Lab (EARL)** organizes selected U.S. economic-account concepts into business-style economic statements for National, Corporate, and Industry analysis. The framework links production, purchased inputs, value added, employee compensation, production taxes, operating surplus, depreciation, interest, profits, and related measures while preserving the definitions and provenance of the underlying official statistics.

Business-style labels in EARL are analytical translations of national-accounting concepts. The methodology distinguishes official source concepts, translated source measures, source-side constructions used to complete specified aggregates, and EARL-calculated analytical measures. It covers six public datasets: National Annual, National Quarterly, Corporate Annual, Corporate Quarterly, Industry Annual, and Industry Quarterly.

Current observations are drawn principally from the U.S. Bureau of Economic Analysis (BEA) National Income and Product Accounts and GDP by Industry accounts. Historical annual Industry series additionally use retained BEA Historical Industry Accounts evidence from November 30, 2021. The historical series are reconciled to the current analytical structure through a reviewed crosswalk, explicit source precedence, and index re-anchoring rules.

The public data package contains 128 dataset-series combinations and 101 entities, together with series definitions, statement layouts, source-vintage information, build provenance, and observation-level provenance. Validation covers source matching, units, formulas, accounting identities, historical reconciliation, public-data consistency, and publication outputs. The resulting framework supports business-style analysis while retaining a documented path from each published measure to its economic definition and source evidence.

Methodological Status and Versioning

Scope of this version

This paper documents the EARL methodology for six public datasets: National Annual (`NIS-A`), National Quarterly (`NIS-Q`), Corporate Annual (`CIS-A`), Corporate Quarterly (`CIS-Q`), Industry Annual (`IIS-A`), and Industry Quarterly (`IIS-Q`). The methodology covers source selection, harmonization, statement construction, historical reconciliation, public-data structure, provenance, and validation.

Documented data build

The manuscript currently documents public-data build `BUILD-20261019T210359Z-ff20569`, generated on **October 19, 2026 at 21:03:59 UTC** from source revision `ff2056952ff08f84a036421c41ca3af0035b0403`. The build contains the six observation datasets and the associated 18-file public data package.

Source vintages

The current BEA acquisitions used by the documented build resolve to `BEA-VIN-2026-09-30`, corresponding to the official BEA release dated **September 30, 2026**. The documented current-source acquisitions were made on **October 7, 2026**. Annual historical Industry reconciliation also uses retained official BEA Historical Industry Accounts workbooks from **November 30, 2021**. Section 14 distinguishes source-release, acquisition, build, and publication dates.

Relationship to Technical Note 01

This paper incorporates the methodology documented in **Technical Note 01 — Historical Industry Reconciliation: Methodology, Crosswalk, Provenance, and Validation**, Version 1.0. The Technical Note remains the specialized reference for row-level historical mappings and the companion crosswalk.

Project naming and revision policy

The current public project name is **EARL — Economic Accounts Research Lab**. Earlier NISP and USEAL identifiers appear only where needed to identify historical or technical artifacts. Future methodology versions should distinguish source-data revisions from changes to mappings, formulas, historical-reconciliation rules, public-series definitions, or validation rules. A new BEA vintage alone does not imply a change in methodology.

1 Introduction and Scope

1.1 Motivation and analytical contribution

The U.S. national economic accounts provide a detailed description of production, income, costs, profits, investment, and sectoral activity. Their concepts are designed for macroeconomic and national-accounting analysis, whereas many users approach economic performance through the language of business statements. Questions about revenue, purchased inputs, gross profit, operating expenses, operating surplus, depreciation, interest, and pretax income therefore require a careful translation between two analytical vocabularies.

EARL (Economic Accounts Research Lab) develops that translation for selected U.S. economic accounts. The project organizes official source concepts into business-style economic statements while retaining the source definitions, measurement conventions, and provenance required for interpretation. At the National and Industry levels, the production identity

$$Gross\ Output - Intermediate\ Inputs = Gross\ Value\ Added$$

is presented as

$$Gross\ Revenue - Cost\ of\ Goods\ Sold = Gross\ Profit.$$

Where the source structure supports a lower statement, EARL organizes employee compensation, production taxes and subsidies, operating expenses, operating surplus, depreciation, interest, and profit concepts into a sequence resembling EBITDA, EBIT, and EBT. The Corporate statement begins at Gross Value Added because the underlying BEA Corporate Business account does not provide a directly comparable Gross Output and Intermediate Inputs layer.

These labels are analytical analogues. Gross Output remains the official national-accounting concept underlying Gross Revenue; Gross Value Added remains the concept underlying Gross Profit; and Gross Operating Surplus remains the source concept underlying the reported EBITDA analogue where EARL uses that presentation. The methodology keeps the official source concept, the EARL label, and any EARL calculation distinct.

1.2 Scope of EARL

The methodology covers six public datasets: National Annual (`NIS-A`), National Quarterly (`NIS-Q`), Corporate Annual (`CIS-A`), Corporate Quarterly (`CIS-Q`), Industry Annual (`IIS-A`), and Industry Quarterly (`IIS-Q`).

National. The National statements describe domestic U.S. economic production and income relationships. They combine production-side concepts from GDP by Industry with income-side concepts from the National Income and Product Accounts (NIPA). National Annual presents BEA Gross Operating Surplus as the primary EBITDA analogue and also provides a calculated EBITDA measure. National Quarterly uses calculated EBITDA because the source structure does not provide a comparable quarterly Gross Operating Surplus series under this methodology.

Corporate. The Corporate statements focus on domestic corporate business. They use the NIPA Corporate Business account and preserve the distinction between profit concepts measured with and without the Inventory Valuation Adjustment (IVA) and Capital Consumption Adjustment (CCAdj). The statement includes calculated EBITDA, source-reported Net Operating Surplus as the EBIT analogue, Corporate Profits with IVA and CCAdj as the principal EBT analogue, after-tax profit, dividends, retained earnings, and selected rest-of-world supplemental series.

Industry. The Industry statements cover 99 public Industry entities. Annual data support a broader statement containing Gross Output, Intermediate Inputs, Gross Value Added, input composition, employee compensation, taxes less subsidies, Gross Operating Surplus, margins, and selected labor and income supplements. Quarterly Industry data support the production block and Gross Margin. The quarterly statement ends there because the available quarterly source data do not support a comparable lower statement.

EARL relies principally on official data from the U.S. Bureau of Economic Analysis (BEA), especially NIPA and GDP by Industry. No single BEA table contains a complete EARL statement. Each series therefore has an explicit source mapping by account, table, line or concept, frequency, and scope. Annual historical Industry series also use retained BEA Historical Industry Accounts evidence from November 30, 2021, reconciled to the current analytical structure under the rules described in Section 11.

1.3 Methodological approach

The source data differ in table structure, frequency, units, Industry classification, historical coverage, seasonal-adjustment treatment, annualization, and revision history. EARL separates the work required to reconcile those differences into four stages:

$$\boxed{Metadata\ Registries \rightarrow Stage\ 1 \rightarrow Stage\ 2 \rightarrow Stage\ 3 \rightarrow Stage\ 4}$$

Stage 1 — Raw acquisition. Required BEA source tables are retrieved and stored in source-preserving tabular form together with the information needed to identify the retrieval and source release.

Stage 2 — Harmonization and reconciliation. Source observations are matched to the documented concepts, units and periods are normalized, source layers are combined under explicit precedence rules, and the historical Industry reconciliation is applied where relevant. Narrow source-side constructions are also performed at this stage when an official aggregate concept can be recovered from specified published components.

Stage 3 — Semantic construction. Harmonized source observations are mapped to stable EARL series. Business-style labels are attached to source concepts, and EARL-calculated measures are evaluated from explicit formulas. Accounting and semantic validation is performed before the processed data are published.

Stage 4 — Public data. The defined public series are selected from the processed data, reshaped into a stable observation schema, joined to the relevant metadata, and written as the public data package. Stage 4 does not introduce new economic calculations.

Metadata define the source mappings, series definitions, formulas, entity relationships, availability rules, and statement layout used across these stages. The public datasets therefore carry both economic values and the information required to distinguish reported, translated, source-side constructed, and EARL-calculated measures.

Historical reconciliation, provenance, and revision treatment are developed separately in Sections 11 and 14.

1.4 Coverage and interpretation boundaries

The paper documents the economic definitions, source mappings, harmonization rules, formulas, historical reconciliation, frequency and unit conventions, public-data structure, provenance, validation, and reproducibility framework for the six datasets. It does not define firm-level representative statements, forecasting or causal methods, unsupported historical allocations, or quarterly Industry lower-statement measures that are absent from the source structure.

Aggregate statement measures require particular care. Gross Output includes transactions among domestic industries and therefore differs from consolidated firm revenue. Gross Value Added represents the value created in production rather than a company-reported gross-profit measure. Gross Operating Surplus and Net Operating Surplus have national-accounting definitions that differ from corporate EBITDA and EBIT. Industry aggregates combine heterogeneous firms and production structures. These differences are part of the interpretation of the EARL statements rather than exceptions to it.

Series-level coverage differs by concept and scope, so dataset-level first and last observations should not be read as balanced coverage for every series or entity.

Sections 2–7 develop the common conceptual and data-construction framework; Sections 8–10 apply it to the National, Corporate, and Industry statements; and Sections 11–16 cover historical reconciliation, comparability, public data, provenance, validation, interpretation, and reproducibility. Appendices A–J provide the detailed technical reference.

2 Conceptual Framework and Terminology

2.1 Analytical translation

EARL separates the economic source concept from its analytical presentation. For example, BEA Gross Output is presented as Gross Revenue, Intermediate Inputs as Cost of Goods Sold, and Gross Value Added as Gross Profit. The numerical source value is unchanged by that translation. The business-style label provides a statement interpretation while the source concept remains available in the metadata.

The central production identity is

$$GO - II = GVA,$$

where GO denotes Gross Output, II Intermediate Inputs, and GVA Gross Value Added. Similar translation principles apply below Gross Value Added. Gross Operating Surplus can serve as a reported EBITDA analogue when that source concept is available for the relevant scope and frequency. Net Operating Surplus can serve as an EBIT analogue, and selected pretax profit concepts can be presented as EBT analogues. The corresponding sections state the exact source basis for each use.

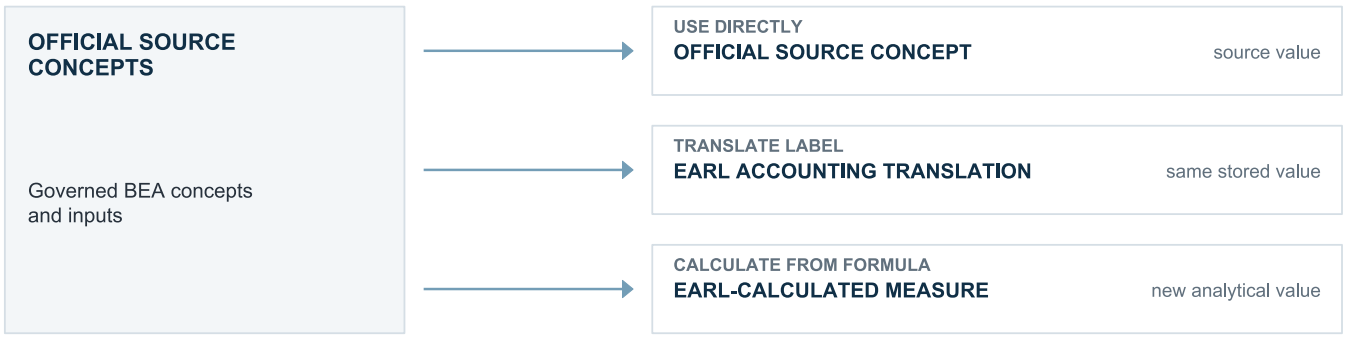
2.2 Concept classes, measurement basis, and provenance

EARL distinguishes three concept classes. An **official source concept** uses the official concept directly. An **EARL accounting translation** presents an official source value under a business-style label. An **EARL-calculated measure** creates a new analytical series from defined inputs using an explicit formula.

Concept class is separate from **measurement basis**. At the series level, a public series is either source-reported or EARL-calculated. Observation-level provenance answers a different question: how a particular value entered the dataset. The five public observation routes are current BEA source data (`bea_api_current`), retained historical BEA data (`bea_historical_2021`), hierarchy rollup (`hierarchy_rollup`), EARL semantic calculation (`derived`), and the aggregate GVA source-side construction (`derived_fallback`).

A source-reported series can therefore include an observation constructed from published source components without changing the measurement basis of the series itself. For example, an Industry parent wage observation can be formed from a complete set of published child wages and retain the meaning of the official wage concept. By contrast, a margin calculated from Gross Value Added and Gross Output is an EARL-calculated series.

Figure 2.1 — From official concepts to EARL measures



2.3 Presentation and terminology

Economic values are stored in their source or formula sign. Presentation metadata can invert a value for statement display, apply percentage formatting, place a series in a statement section, or repeat the same series in more than one analytical context. These operations do not change the underlying observation.

Figure 2.1 summarizes how official concepts become direct, translated, or calculated EARL measures. The terminology used throughout the paper follows the distinctions in Table 2.1. **BEA-reported** refers to an official source concept; **EARL-calculated** refers to a semantic formula; **constructed** refers to the specified source-side construction of an official concept; and **translated** refers to a change in analytical label with no change to the underlying source value. Machine identifiers are shown only when they aid reproducibility.

Table 2.1 — EARL concept and provenance distinctions

Question	Classification	Examples
What kind of economic concept is this?	Concept class	Official source concept; EARL accounting translation; EARL-calculated measure
Is the series source-reported or calculated?	Measurement basis	Source-reported; EARL-calculated
How did this observation enter the dataset?	Observation provenance	Current BEA; historical BEA; hierarchy rollup; derived; aggregate GVA construction
How should the value be displayed?	Presentation metadata	Source sign; display inversion; row type; indentation; display policy

Note: These classifications describe distinct aspects of a published measure. They do not change the underlying economic value or create additional measurement classes.

3 Source Data and Metadata Governance

3.1 Official sources

EARL's current observations draw primarily from two BEA systems: the **National Income and Product Accounts (NIPA)** and **GDP by Industry**. NIPA provides the principal National income-side and Corporate Business concepts. GDP by Industry provides Gross Output, Intermediate Inputs, Gross Value Added, the annual Energy/Materials/Services decomposition, and annual Industry Gross Operating Surplus. Selected Industry labor, interest, and proprietors' income concepts come from Industry-level NIPA tables.

Each EARL series is mapped to its source at the level needed to identify the economic concept unambiguously. Depending on the source, that mapping can include the BEA database, table, line number, Industry code, source concept description, frequency, and unit. A calculated series instead records its formula and input series. Appendix B provides the detailed table and line crosswalk.

3.2 Metadata governance

Table 3.1 summarizes the principal metadata functions. Methodological metadata serve six functions: defining stable economic series; mapping those series to source concepts; defining formulas; specifying the public Industry entity structure; recording explicit inclusions and exclusions; and identifying the applicable source release. The same definitions are used by source acquisition, harmonization, semantic construction, and publication.

Industry metadata require additional structure because the public universe contains ordinary hierarchy entities, a cross-cutting ICT-producing aggregate, and a historical Hospitals/Nursing bridge. Parent-child traversal follows the documented parent relationship. The accompanying hierarchy level is descriptive and does not independently define the graph.

Some methodological choices cannot be inferred from source tables alone. Examples include whether a calculated measure belongs in a particular dataset, whether a source concept should remain supplemental rather than enter the main statement, and whether an Industry parent may be constructed from published children. Such decisions are recorded explicitly so that a change in implementation cannot silently change the economic definition.

Table 3.1 — Principal metadata functions

Metadata category	Methodological role
Series definitions	Stable economic identity, label, unit, measurement basis, and availability
Source mappings	BEA database, table, line/concept, Industry scope, and unit
Formula definitions	Inputs and expression for EARL-calculated measures
Entity metadata	Industry identity, source code, parent relationship, and special-entity status
Inclusion and exclusion rules	Explicit dataset scope and source-side constructions
Source-vintage metadata	Official release identity and evidence supporting the assignment

Source: EARL methodology, Section 3; governed metadata registry.

3.3 Source vintages, historical evidence, and change control

The current build uses BEA vintage `BEA-VIN-2026-09-30`. NIPA provides a machine-readable `LastRevised` date for that release, whereas GDP by Industry requires the release assignment to be linked through the source-vintage metadata. Section 14 documents the evidence and separates source release, API response time, EARL acquisition time, build time, and publication time.

Historical annual Industry reconciliation uses retained BEA Historical Industry Accounts workbooks from November 30, 2021. Those workbooks form a separate historical source record and are reconciled to the current analytical structure under Section 11.

Changes to source mappings, formulas, entity relationships, historical crosswalk rules, availability, or public-series definitions are evaluated against official source evidence and the documented economic interpretation. The corresponding metadata, implementation, validation, and version record are updated together. Earlier outputs can help identify changes across versions, but they do not override the current source evidence or stated method.

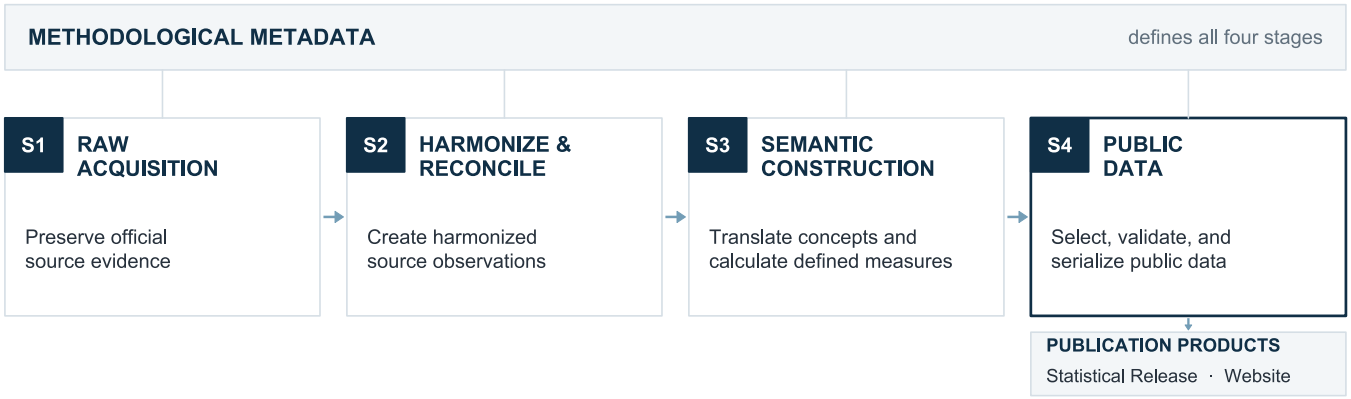
4 Canonical Data Architecture

4.1 Architecture and source evidence

EARL separates data construction into four stages so that source evidence, harmonization, analytical calculation, and publication have distinct responsibilities. Figure 4.1 summarizes the four-stage architecture and the role of methodological metadata.

The metadata used across the pipeline define the source mappings, semantic concepts, formulas, entity relationships, availability rules, and statement structure. This arrangement allows a source change to be distinguished from an EARL formula change and both to be distinguished from a presentation change.

Figure 4.1 — Canonical EARL data architecture



Stage 1 retrieves the official tables required by the source mappings and stores them in source-preserving tabular form. Acquisition metadata identify the BEA system, table request, retrieval time, source revision information where available, and the stored raw artifact. Business-style labels and EARL formulas are not applied at this stage.

Retaining the acquired source evidence makes it possible to repeat later processing against the same source snapshot. This is important because a later live request to BEA can return revised observations even when the request parameters are unchanged.

4.2 Harmonization and semantic construction

Stage 2 matches source observations to the defined economic concepts, normalizes units and periods, applies source precedence, and reconciles historical Industry evidence. It also permits the limited source-side constructions described in Section 6. Stage 3 then maps harmonized observations to stable EARL series and evaluates the analytical formulas described in Section 7. Keeping these responsibilities separate distinguishes recovery of a source concept from creation of an EARL-calculated measure.

4.3 Public data and publication products

Stage 4 selects the defined public series from the processed data, preserves their values and observation provenance, and writes the public observation and metadata files. The public package separates observations from series definitions, entity definitions, statement layout, source-vintage information, build information, and schema documentation. Section 13 describes the resulting structure.

Statements, charts, statistical releases, and the website use this public layer. They can select periods, format ratios, apply statement display signs, or reshape observations for presentation. Economic definitions and formulas remain fixed by the methodology described in Stages 1–3.

5 Data Acquisition and Raw Evidence

5.1 Retrieval and raw captures

The acquisition layer preserves the official observations needed to reproduce later transformations. Required source tables are determined by the source mappings rather than by the needs of an individual chart or release. NIPA and GDP by Industry responses are stored in source-preserving tabular form so that source identifiers, descriptions, periods, units, and values remain available for audit.

A stored raw artifact can contain several table requests from the same BEA system. The acquisition record therefore has a finer grain than the raw file itself: each requested table has its own acquisition identity even when several requests are accumulated in one stored file.

5.2 Acquisition metadata and source integrity

For each table request, EARL records the information needed to establish what was retrieved and when. The acquisition record identifies the EARL dataset, source agency and source system, source table, requested frequency and years, retrieval time, source revision information where available, stored raw artifact, and row count. GDP by Industry requests also record a fingerprint of the ordered Industry-code list so that the requested scope can be verified without duplicating the full list in every acquisition row.

Three time fields have different meanings. EARL retrieval time records when the response was acquired. BEA API production time records when the API generated the response. For NIPA, `LastRevised` identifies the official table-revision date reported by BEA. These timestamps are retained separately. In particular, API production time does not establish the official GDP by Industry release date.

Appendix H provides the detailed acquisition inventory and provenance schema for the documented build. Those build-specific records are refreshed with the publication data build rather than treated as timeless methodological facts.

A required source request must return usable source evidence. Empty or structurally invalid required responses stop processing for the affected acquisition. The methodology does not interpret a failed retrieval as an economic zero or as a fully missing source series.

The same principle applies to source metadata. When a required source identifier, unit, or mapping cannot be established from the documented rules, the issue is resolved at the source or metadata level before the observation proceeds to the harmonized dataset.

6 Source Harmonization and Reconciliation

6.1 Matching and normalization

Stage 2 establishes that each acquired observation corresponds to the intended economic concept. For fixed-line NIPA mappings, the table, line number, and normalized line description must agree. This protects against silent reuse of a line number whose description has changed. A mismatch or ambiguous match stops the mapping for review.

GDP by Industry observations are matched by source table, Industry code or aggregate scope, and the relevant source concept. Multi-concept tables require the specified concept description as well as the Industry code. Selected Industry-level NIPA supplements are matched by normalized entity description because the source tables are organized by Industry rather than by one fixed line per EARL entity; repeated descriptions are resolved using the documented source and registry ordering.

Source units are validated before conversion. GDP by Industry current-dollar series reported in billions are converted to USD millions by multiplying by 1,000. Index series remain indexes. Annual observations are normalized to calendar years. Quarterly observations retain the `YYYYQn` period representation together with their seasonal-adjustment and annualization conventions.

6.2 Missingness, source precedence, and source-side constructions

Table 6.1 — Stage 2 harmonization and source-side construction rules

Rule area	Stage 2 treatment	Guardrail / failure condition	Resulting source treatment
Source matching	Require the governed identifiers appropriate to each source: NIPA table, line, and normalized description; GDP by Industry table, Industry code, and expected concept; NIPA Industry normalized description plus governed occurrence order.	Mismatch or ambiguity blocks processing. Zero NIPA Industry matches block; legitimate source-limited partial entity coverage may remain.	Accepted observations preserve source identity.
Units and time	Validate source units before conversion. GDP by Industry current-dollar billions become USD millions by multiplying by 1,000; indexes remain indexes. Annual periods become calendar years; quarterly periods use `YYYYQn`, preserving seasonal-adjustment and annualization semantics.	Unsupported units or multipliers are not guessed.	Common unit and period semantics for semantic construction and public metadata.
Missingness	Form the governed variable/entity/period grid and represent unavailable observations explicitly.	Missing is not zero. Structural, source-limited, and unresolved-mapping missingness remain distinguishable.	Deterministic layer combination with explicit missing cells.
Source precedence	Apply fixed source priority; lower-priority evidence fills only cells left missing by higher-priority evidence.	Overlapping observations are not averaged. Current source evidence takes precedence from 1997 onward; historical evidence extends eligible series through 1996.	One source observation per defined cell.
Hierarchy rollups	For selected additive series/entity pairs, construct a missing parent from a complete set of published children.	A direct reported parent wins. An incomplete child set yields no parent. Wage per FTE is recomputed from completed parent wage and FTE totals rather than averaged across child ratios.	Source concept and measurement basis are retained; rollup provenance is recorded.
All Industries GVA fallback	If Gross Value Added for All Industries is missing while aggregate Gross Output and Intermediate Inputs are present, fill GVA as GO − II.	A directly reported GVA is preserved; the fallback is restricted to the All Industries aggregate.	Source concept and measurement basis are retained; fallback-derived provenance is recorded.
Historical integration	Annual National and Industry processing may incorporate the retained 2021 historical layer before semantic construction.	Historical data enter as source evidence. Current evidence wins from 1997 onward; taxonomy bridges and index anchoring follow Section 11 and Appendix F.	Eligible history extends backward through 1996 with historical-source provenance.

Note: Stage 2 source-side constructions fill source-layer gaps under narrowly defined conditions while preserving the measurement basis of the source concept. They do not turn the series into an EARL-calculated measure; general semantic and economic calculation belongs to Stage 3. Exact observation-provenance tokens are documented in the public schema and Appendix G.

Source: EARL methodology, Sections 6.2–6.10, and current Stage 2 implementation.

Table 6.1 summarizes the matching, precedence, and source-side construction rules. Harmonization constructs the defined series/entity/period grid and preserves unavailable observations as missing. Missingness can reflect source coverage, scope restrictions, or an unresolved mapping; it is not replaced with zero.

When more than one source layer can populate the same observation, precedence is fixed before the merge. Higher-priority source values are retained, and lower-priority layers fill only missing observations. Overlapping source values are

not averaged. For the historical annual Industry series, current BEA observations take precedence from 1997 onward and the reconciled historical source extends eligible series through 1996.

Selected Industry parent observations can be constructed from published children when the direct parent is unavailable. Construction is limited to defined series and parents, and every required child must be present. A direct reported parent observation always takes precedence.

For labor aggregates, wages and FTE are summed separately across the required children; wage per FTE is then calculated from those parent totals rather than by averaging child wage-per-FTE ratios. Selected interest aggregates use the same complete-child rule. Section 10 identifies the public Industry parents to which these constructions apply.

A separate construction applies to aggregate Industry Gross Value Added. Where All Industries Gross Value Added is unavailable but aggregate Gross Output and Intermediate Inputs are both observed, EARL uses

$$GVA_{All\ Industries} = GO_{All\ Industries} - II_{All\ Industries}.$$

This construction fills only the missing aggregate GVA observation and never replaces a direct source value. Its observation provenance remains distinct from a Stage 3 analytical calculation.

6.3 Historical reconciliation and harmonized output

Annual Industry processing can incorporate the retained historical source before semantic calculations. Historical observations carry their historical-source provenance through the splice and remain distinguishable from current BEA observations. Section 11 describes the taxonomy crosswalk, current-dollar splice, index re-anchoring, bridge concepts, and historical validation in detail.

The output of Stage 2 is a common source representation containing the series identity, entity, period, value, unit, and observation provenance needed for semantic construction. It is an intermediate analytical layer rather than a public dataset.

7 Semantic Statement Construction

7.1 Semantic mapping and formula evaluation

Stage 3 assigns harmonized source observations to stable EARL series. The series identifier is independent of the row's location in a particular statement, which allows the same economic measure to appear in more than one presentation context without creating multiple economic definitions.

Source-reported concepts retain their source measurement basis. EARL accounting translations attach the business-style label while leaving the value unchanged. EARL-calculated series are created only where the methodology defines an explicit formula.

The semantic formula language supports registered series identifiers, numerical constants, arithmetic operators, comparisons, a restricted conditional expression, and guarded division. Unknown inputs, duplicate definitions, circular dependencies, and unsupported expressions stop evaluation.

For ordinary division, a zero or missing denominator yields a missing result. More specific denominator rules are encoded where the interpretation requires them. Corporate Effective Tax Rate, for example, is defined only when adjusted EBT is positive:

$$EffectiveTaxRate = \begin{cases} Taxes/EBT, & EBT > 0, \\ NA, & EBT \leq 0. \end{cases}$$

7.2 Provenance, presentation, and quality assurance

Stage 3-calculated observations receive `derived` provenance. A source-reported series retains its source measurement basis even when a particular observation entered through a hierarchy rollup or aggregate GVA construction. This preserves the distinction between what the series measures and how a particular cell was populated.

Values are stored in their economic sign. Statement-specific sign inversion, indentation, row type, repeated use of a series, and percentage formatting are presentation rules applied after the economic value has been constructed.

Before processed data are accepted, validation checks unique observation keys, required identifiers, finite numerical values, consistency between missingness and provenance, formula dependencies, and the accounting identities defined for

the relevant dataset. The processed data include the public series together with supplemental and reconciliation measures needed for validation; only the series defined for publication enter Stage 4.

8 National Income Statement Methodology

8.1 Scope and production block

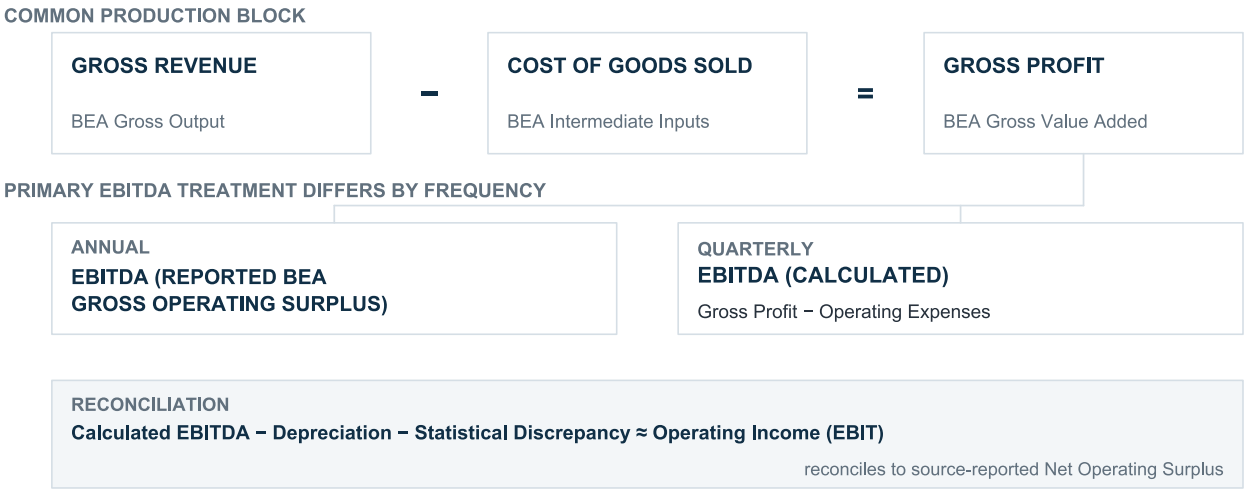
The National Income Statement describes domestic U.S. production and income relationships at the economy-wide level. It combines GDP by Industry production concepts with NIPA income-side concepts so that the production identity and the lower statement can be viewed together.

EARL presents Gross Output as **Gross Revenue**, Intermediate Inputs as **Cost of Goods Sold**, and Gross Value Added as **Gross Profit**. The production identity is

$$GO - II = GVA.$$

Gross Output and Intermediate Inputs come from GDP by Industry, whereas National Gross Value Added is sourced from NIPA. Their agreement is therefore a cross-account check rather than an identity constructed from a single table. Figure 8.1 shows how these production and income concepts enter the National statement.

Figure 8.1 — National statement construction



8.2 Operating expenses and EBITDA

National Operating Expenses are calculated as

$$OpEx = Compensation + Taxes\ on\ Production\ and\ Imports - Subsidies.$$

The public data also expose net production taxes,

$$TaxesLessSubsidies = Taxes - Subsidies.$$

National Annual and National Quarterly use different EBITDA measurement bases. For **National Annual**, the primary EBITDA analogue is source-reported BEA Gross Operating Surplus from GDP by Industry, displayed as **EBITDA (Reported BEA Gross Operating Surplus)**. EARL also calculates

$$EBITDA^{calc} = GVA - OpEx$$

as a supplemental measure. The reported and calculated series remain separate. Their overlap is validated within a $100 million source-rounding tolerance.

For **National Quarterly**, a comparable reported Gross Operating Surplus series is not available under the source structure used by EARL. Calculated EBITDA is therefore the primary quarterly EBITDA measure.

8.3 EBIT, EBT, margins, and supplemental income

Net Operating Surplus is the National EBIT analogue. The production-side bridge to Net Operating Surplus includes the Statistical Discrepancy:

$$NOS \approx EBITDA^{calc} - CFC - StatisticalDiscrepancy,$$

where CFC denotes Consumption of Fixed Capital. The Statistical Discrepancy is stored in its source sign; statement presentation can display the corresponding adjustment with the opposite sign where required for the bridge.

For 2026 Q2, calculated EBITDA is 13,997,289 million dollars, CFC is 5,358,760 million, the source-signed Statistical Discrepancy is −189,498 million, and NOS is 8,828,026 million. The resulting bridge residual is approximately 1 million dollars, consistent with source rounding.

EARL calculates National EBT as

$$EBT = NOS - NetInterest,$$

using the relevant NIPA net-interest-and-miscellaneous-payments concept. This EBT is an EARL analytical measure. The underlying interest concept is a national-accounting net measure and should be interpreted accordingly.

For National Annual, EARL also publishes an output-perspective Net Operating Surplus measure:

$$NOS^{output} = GOS - CFC.$$

This is distinct from source-reported NIPA Net Operating Surplus. Their difference reflects the statistical-discrepancy relationship between the production and income accounts.

National margins are calculated against Gross Output. Gross Margin is

$$GrossMargin = GVA/GO.$$

Annual reported EBITDA Margin uses Gross Operating Surplus over Gross Output, while the quarterly EBITDA Margin uses calculated EBITDA over Gross Output. Operating Margin uses Net Operating Surplus over Gross Output, and EBT Margin uses EARL-calculated EBT over Gross Output.

The annual dataset also provides the main source-reported components of Net Operating Surplus, including interest and miscellaneous payments, net business current transfers, proprietors' income with IVA and CCAdj, rental income with CCAdj, current surplus of government enterprises, and corporate profits with IVA and CCAdj. These series are supplemental rather than a second EARL calculation of NOS.

8.4 Sources and availability

National production concepts use GDP by Industry Tables 15, 20, and 25 together with NIPA Gross Domestic Product/Gross Value Added. Compensation, wages, supplements, production taxes, subsidies, Consumption of Fixed Capital, Net Operating Surplus, interest, supplemental income components, and the Statistical Discrepancy come from the corresponding NIPA tables documented in Appendix B.

Availability differs by concept. National Annual Gross Value Added and several income-side components begin in 1929; Gross Output and Intermediate Inputs begin in 1947; Net Operating Surplus and EBT begin in 1959; Energy, Materials, and Services inputs begin in 1963; and reported Gross Operating Surplus begins in 1997. National Quarterly Gross Value Added and several lower-statement concepts begin in 1947Q1, Net Operating Surplus and EBT in 1959Q1, and Gross Output, Intermediate Inputs, and production-based margins in 2005Q1.

9 Corporate Income Statement Methodology

9.1 Scope and sources

The Corporate Income Statement describes domestic corporate business. Its source structure differs from the National and Industry production accounts because the NIPA Corporate Business account begins naturally at Corporate Gross Value Added. EARL therefore begins the statement at **Gross Profit / Gross Value Added** and does not construct a Corporate Gross Revenue or Cost of Goods Sold layer without a corresponding source basis.

Figure 9.1 summarizes the Corporate statement structure. The principal source is NIPA Table 1.14, supplemented by Table 6.16D for selected rest-of-world receipts and payments. The source provides Corporate Gross Value Added, Consumption of Fixed Capital, employee compensation and its components, taxes on production and imports less subsidies, Net Operating Surplus, net interest and miscellaneous payments, business current transfers, corporate profits, corporate income taxes, after-tax profits, dividends, undistributed profits, and the IVA/CCAdj adjustment series.

Figure 9.1 — Corporate statement construction



9.2 Operating expenses, EBITDA, EBIT, and EBT

Corporate Operating Expenses are calculated as

$$OpEx = Compensation + TaxesLessSubsidies,$$

and Corporate EBITDA as

$$EBITDA^{calc} = GVA - OpEx.$$

The calculated EBITDA measure is independently checked against the source relationship

$$EBITDA^{calc} \approx NOS + CFC,$$

with a $5 million tolerance for published-source rounding.

Corporate EBIT is source-reported Net Operating Surplus. The principal Corporate EBT analogue is also source-reported: **Corporate Profits with IVA and CCAdj**. EARL does not replace that source concept with a calculation from EBIT and interest. A separate reconciliation measure can be formed from Net Operating Surplus, net interest and miscellaneous payments, and net business current transfers to illuminate the relationship among the accounts.

9.3 Net income, valuation adjustments, and ratios

Corporate Net Income is presented from source-reported after-tax corporate profits with IVA and CCAdj. Retained Earnings are represented by source-reported undistributed corporate profits with IVA and CCAdj, and net dividends are retained as a separate source series.

The Corporate accounts publish profit concepts both with and without the Inventory Valuation Adjustment and Capital Consumption Adjustment. EARL preserves this distinction. For the relevant series,

$$With\ Adjustments \approx Without\ Adjustments + IVA + CCAdj,$$

subject to source rounding and the exact BEA definitions of each line. The adjusted series provide the principal statement measures because they align more closely with the national-accounting treatment of inventories and capital consumption.

The Effective Tax Rate is defined as

$$EffectiveTaxRate = \begin{cases} CorporateIncomeTaxes/EBT, & EBT > 0, \\ NA, & EBT \leq 0. \end{cases}$$

The positive-EBT condition prevents negative or undefined denominator periods from being presented as meaningful effective tax rates. In the annual history, 1932 and 1933 therefore have no published Effective Tax Rate.

The Payout Ratio is net dividends divided by adjusted Net Income when the denominator is nonzero and observed. Retention Ratio equals one minus the Payout Ratio. These ratios can fall outside the interval from zero to one when dividends exceed current after-tax profits or when the underlying source concepts have unusual signs; such values are retained when they follow directly from the published inputs.

Annual rest-of-world supplemental series begin in 1998 and quarterly series in 2001Q1. They provide additional context for corporate profit flows but do not alter the domestic Corporate statement identities.

9.4 Frequency and availability

Corporate Annual and Corporate Quarterly each contain 26 public series: 21 source-reported and five EARL-calculated. Their statement structures are intentionally parallel because both frequencies use the same principal NIPA account structure, although period coverage differs by series.

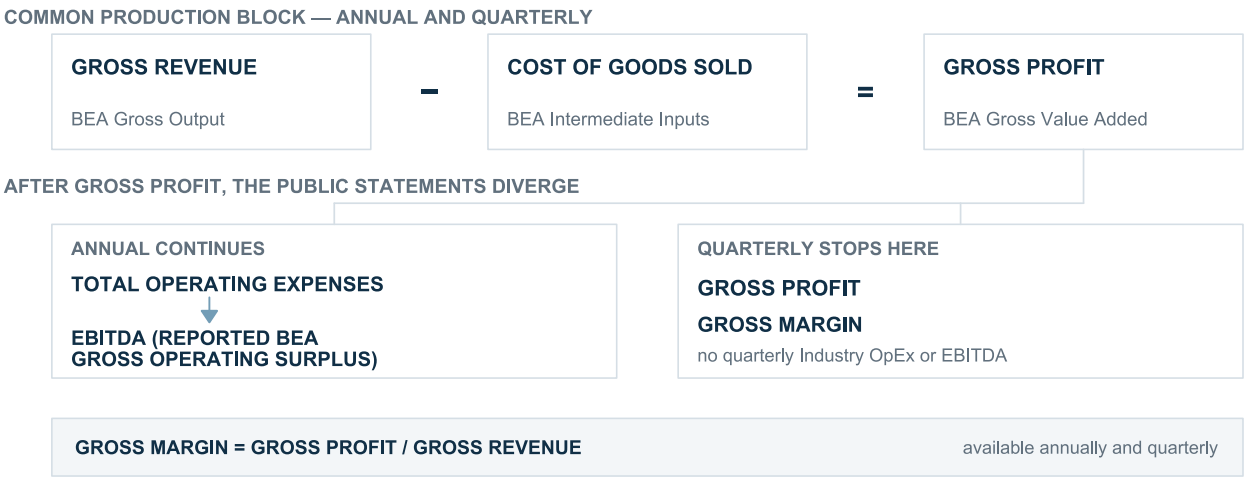
10 Industry Income Statement Methodology

10.1 Scope and entity architecture

The Industry Income Statement applies the production framework across 99 public Industry entities. Most entities belong to the ordinary Industry parent graph rooted at All Industries. The public universe also contains a cross-cutting information-communications-technology-producing aggregate and a historical Hospitals/Nursing bridge entity. The latter supports historical comparability and is not a current BEA taxonomy node.

The parent relationship is the basis for hierarchy traversal. The accompanying hierarchy level is descriptive and can reflect a mixture of BEA and curated levels, so it should not be treated as graph-theoretic depth. Quarterly Industry data cover 97 entities; the cross-cutting ICT aggregate and historical Hospitals/Nursing bridge are absent because the quarterly source structure does not support them. Figure 10.1 shows the resulting annual/quarterly statement boundary.

Figure 10.1 — Industry statement construction



10.2 Production, input structure, and operating surplus

Industry Gross Output, Intermediate Inputs, and Gross Value Added come from GDP by Industry. Current-dollar source values reported in billions are converted to USD millions. The core production identity is

$$GO - II \approx GVA,$$

where small differences can arise from published-source rounding and, for aggregate historical comparisons, from cross-table construction.

Annual Intermediate Inputs are also decomposed into Energy, Materials, and Services where the source provides the components:

$$II \approx Energy + Materials + Services.$$

The source tables and concept mappings used for these measures are listed in Appendix B.

Annual Industry Operating Expenses are calculated as

$$OpEx = Compensation + TaxesLessSubsidies.$$

BEA Gross Operating Surplus is source-reported for the Industry annual statement and presented as the Industry EBITDA analogue. The corresponding relationship is

$$GOS \approx GVA - OpEx,$$

subject to source coverage and published rounding. Gross Margin is Gross Value Added divided by Gross Output, and annual EBITDA Margin is Gross Operating Surplus divided by Gross Output.

Quarterly Industry data publish Gross Output, Intermediate Inputs, Gross Value Added, and Gross Margin. Quarterly Operating Expenses and EBITDA are not published because the corresponding quarterly source concepts required by this methodology are unavailable.

10.3 Labor, financing, and aggregation

Annual Industry supplemental data include wages and salaries, full-time equivalent employees (FTE), wage per FTE, selected interest measures, and proprietors' income. Wages, FTE, and wage-per-FTE concepts are drawn from the relevant NIPA Industry tables. Wage per FTE is expressed in dollars per FTE; FTE is published in thousands.

For selected labor series, EARL forms the public parent aggregates **Finance & Real Estate**, **Professional & Business Services**, **Education & Health**, and **Arts, Entertainment & Food Services** from complete published child sets when the parent is not reported directly. Wage per FTE is calculated from the resulting parent wage and FTE totals. A direct reported parent observation has priority over a rollup. Private-sector interest can also be constructed from its specified children under the same complete-child rule.

Aggregate All Industries Gross Value Added can be filled from Gross Output less Intermediate Inputs when the direct aggregate observation is missing. This source-side construction is limited to the aggregate GVA concept and does not overwrite direct source observations.

10.4 Historical extension, frequency, and availability

Annual Industry history incorporates the retained November 30, 2021 BEA Historical Industry Accounts source. For eligible current-dollar measures, reconciled historical observations extend through 1996 and current observations take precedence from 1997 onward. Energy, Materials, and Services history begins in 1963 where supported. Employee Compensation, production taxes less subsidies, Gross Operating Surplus, Operating Expenses, and EBITDA Margin extend from 1987 where the retained source supports the relevant concepts.

The annual and quarterly statements therefore differ in both time span and economic content. Quarterly Industry data begin in 2005Q1 for the production block and cover the 97 entities supported by the quarterly source. Cross-frequency analysis should compare only measures whose definition and scope are available at both frequencies.

Across the annual public data, broad production measures and Gross Margin are available for all 99 public Industry entities, while some lower-statement and supplemental measures have narrower entity coverage. The current documented build contains annual Energy/Materials/Services for 98 entities, employee compensation and Gross Operating Surplus for 98, Operating Expenses for 94, labor detail for 93, interest for 23, and proprietors' income for 26. Series-level metadata remain the authoritative source for exact availability.

11 Historical Industry Reconciliation

11.1 Historical source and taxonomy

Table 11.1 — Historical Industry reconciliation controls

Control area	Historical reconciliation rule	Guardrail / boundary	Publication effect
Source evidence	Use four retained official BEA Historical Industry Accounts workbooks from the November 30, 2021 release, with file hashes recorded in the source manifest.	Historical extension is based on retained official evidence; the source files are preserved as fixed audit evidence.	Auditable source base for the long-run Industry history.
Taxonomy / crosswalk	Use the current BEA publication taxonomy as the reference structure and map historical concepts only where the relationship is defensible. The reviewed crosswalk contains 101 outcomes: 97 current public-entity mappings, 2 current-taxonomy references, and 2 historical bridges.	Crosswalk outcomes are not a second public-entity universe. A one-to-one current mapping is not inferred where the historical source does not support it.	Historical concepts enter the current publication structure only at defensible scopes.
Current-dollar splice	Use verified historical observations through 1996 and current BEA observations from 1997 onward.	The overlap is not averaged or selected for smoothness; current evidence takes precedence from 1997 onward.	One current-dollar observation per eligible cell.
Index re-anchoring	Preserve the historical price or quantity growth path and re-express its level using the ratio of the current 1997 index to the historical 1997 index.	Re-anchoring requires both historical and current 1997 anchors; the multiplicative factor is constant over the historical segment.	Historical growth is preserved while index levels align to the current overlap basis.
Historical bridges	Preserve combined historical Real estate and Hospitals/Nursing concepts where the current taxonomy publishes separate children.	A modern-child allocation is not created without an official source basis.	Combined historical bridge information remains available only at the defensible aggregate scope.
Index omissions	Omit a historical bridge index when no defensible current 1997 anchor exists. For the Hospitals/Nursing bridge, this affects four price/quantity index series across 1963–1996, totaling 136 observations.	No synthetic anchor or child allocation is introduced to fill the gap.	The affected observations remain deliberately unavailable rather than being treated as processing failures.
Availability / provenance	Extend broad Gross Output, Intermediate Inputs, and Gross Value Added history from 1947 where supported; Energy, Materials, and Services from 1963; and validated lower value-added components from 1987.	Availability remains series- and entity-specific, and retained historical observations preserve their historical-source lineage.	Long-run coverage expands without inventing unsupported periods or detail.

Note: The 1997 overlap serves two distinct purposes: it is the current-dollar precedence boundary and, where an exact match exists, the anchor year for historical price and quantity indexes. Re-anchoring changes index level, not the historical growth path. Where no defensible mapping or anchor exists, EARL preserves the aggregate or leaves the observation unavailable rather than creating unsupported detail.

Source: EARL methodology, Section 11, and the historical reconciliation package.

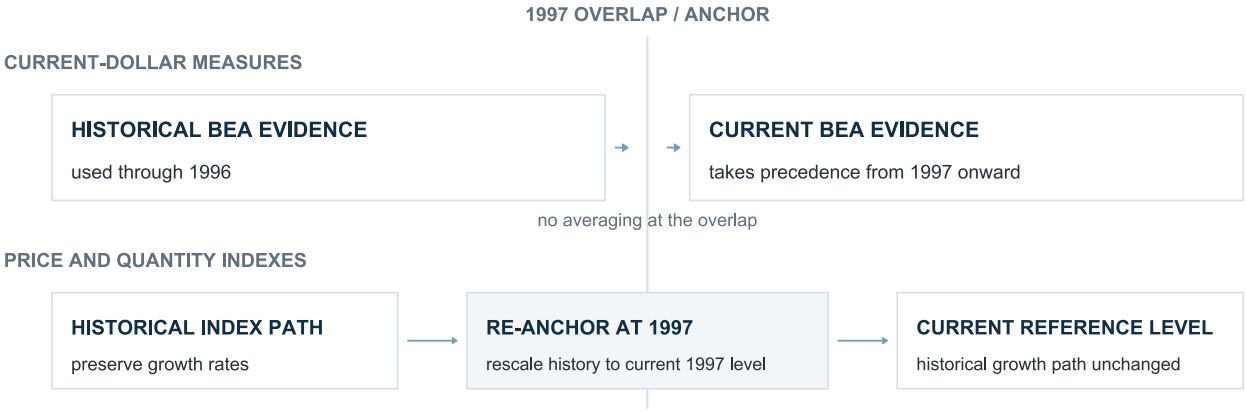
The current Industry publication structure cannot be extended backward by simply appending older files because historical BEA releases use different classifications and, in some cases, broader aggregates. EARL therefore reconciles retained official historical evidence to the current analytical structure before the historical observations enter the semantic calculations.

Table 11.1 summarizes the principal historical reconciliation controls. The retained source consists of four BEA Historical Industry Accounts workbooks released on November 30, 2021: Gross Output, Intermediate Inputs, KLEMS/input composition, and Value Added. The current BEA publication taxonomy serves as the principal reference for public entities. A reviewed 101-row crosswalk records how historical concepts relate to that structure.

The crosswalk contains 97 rows that resolve one-to-one to ordinary current public Industry entities, two current-taxonomy reference concepts used in the historical mapping, and two historical bridge concepts. Figure 11.1 summarizes the reader-facing splice logic. These 101 rows describe reconciliation outcomes; they are not a second count of the public Industry

universe. The public universe contains 99 Industry entities: 97 ordinary current-hierarchy entities, the cross-cutting ICT aggregate, and the dedicated Hospitals/Nursing historical bridge entity.

Figure 11.1 — Historical Industry splice logic



11.2 Current-dollar splice and index re-anchoring

For eligible current-dollar series, EARL uses reconciled historical observations through 1996 and current BEA observations from 1997 onward. Current observations have priority in the overlap; historical and current values are not averaged.

Historical price and quantity indexes use the same 1997 overlap differently. The historical growth path is preserved and rescaled to the current 1997 index level:

$$I_t^{EARL} = I_t^{hist} \times \frac{I_{1997}^{current}}{I_{1997}^{hist}}.$$

The multiplication changes the level of the historical index but leaves its period-to-period growth unchanged. The historical reference package contains 878 verified 1997 anchor observations across the price and quantity series used for re-anchoring.

11.3 Unsupported splits, lower-statement history, and validation

Two historical categories cannot be decomposed into the corresponding current children without introducing unsupported allocation assumptions. The historical Real Estate aggregate cannot be divided retrospectively between Housing and Other Real Estate, and the historical Hospitals and Nursing/Residential Care aggregate cannot be divided between the modern hospital and nursing/residential categories. EARL retains the available combined concept instead of allocating it synthetically.

The same rule applies to historical indexes. The historical Hospitals/Nursing bridge lacks an exact current 1997 anchor for four index series. The affected 136 observations—34 years from 1963 through 1996 across Gross Output and Intermediate Inputs price and quantity indexes—are therefore excluded from the re-anchored index history rather than assigned a synthetic anchor.

The retained historical source also supports selected value-added components. Employee Compensation, Taxes on Production and Imports less Subsidies, and Gross Operating Surplus extend from 1987 where the historical source and crosswalk support the current analytical concept. These observations enter the same annual Industry semantic calculations as current observations after reconciliation.

Historical observations retain `bea_historical_2021` provenance; current API observations retain `bea_api_current`. The splice therefore preserves the source route of each observation even though the final annual series is continuous across the historical/current boundary.

Historical validation covers unique keys, required identifiers, Gross Output/Intermediate Inputs/Gross Value Added relationships, the Intermediate Inputs decomposition, the lower value-added block, index normalization, preservation of historical index growth, and observation provenance. The validated historical release passed nine core checks. Eleven selected historical Gross Output observations were also compared directly with the retained BEA workbooks and matched exactly.

Current-dollar identity tolerances are relationship-specific. Most production and input-decomposition checks use $150 million; two aggregate cross-table value-added checks use $250 million to accommodate the rounding interaction among independently published source tables. These tolerances are validation thresholds for the stated relationships, not general allowances for data error.

12 Frequency, Units, Availability, and Comparability

12.1 Frequency, units, and availability

Table 12.1 summarizes the frequency, unit convention, public series count, entity coverage, and observed span of the six datasets.

Table 12.1 — Dataset publication and comparability profile

Dataset	Monetary basis	Public series (source / calculated)	Entities	Observed span	Comparability point
National annual (`NIS-A`)	USD millions	30 (21 / 9)	1	1929–2025	Annual EBITDA analogue uses source-reported BEA Gross Operating Surplus.
National quarterly (`NIS-Q`)	USD millions SAAR	25 (17 / 8)	1	1947Q1–2026Q2	Quarterly EBITDA is EARL-calculated, so its measurement is not identical to annual.
Corporate annual (`CIS-A`)	USD millions	26 (21 / 5)	1	1929–2025	Annual and quarterly structures are closely aligned; series start dates vary, and some ratios are unavailable in periods where the denominator does not satisfy the measure's definition.
Corporate quarterly (`CIS-Q`)	USD millions SAAR	26 (21 / 5)	1	1947Q1–2026Q2	Annual and quarterly structures are closely aligned; series start dates vary, and some ratios are unavailable in periods where the denominator does not satisfy the measure's definition.
Industry annual (`IIS-A`)	USD millions	17 (14 / 3)	99	1947–2025	Annual includes substantially richer lower-statement and supplemental detail than quarterly; historical extension does not erase classification differences.
Industry quarterly (`IIS-Q`)	USD millions SAAR	4 (3 / 1)	97	2005Q1–2026Q2	Quarterly Industry coverage is limited to the production block and Gross Margin; lower-statement and supplemental detail available annually is not published quarterly. Coverage is 97 entities rather than 99 annual.

Note: Observed span is the union of observations in each dataset; it does not imply that every series or entity is available throughout the full span. Monetary basis describes monetary series; ratios and Industry labor measures retain their series-specific units.

Source: EARL Stage 4 build metadata and series metadata; comparability statements from Section 12.

EARL publishes annual and quarterly datasets. Annual periods represent calendar years. Quarterly periods use `YYYYQn` and correspond to calendar-quarter end dates in the public data.

Annual monetary series are expressed in **USD millions**. Quarterly monetary series are expressed in **USD millions at seasonally adjusted annual rates (SAAR)** where the underlying source uses that convention. Four quarterly SAAR observations should not be summed to obtain an annual flow. Ratios are stored as decimals; a value of `0.25` represents 25 percent. Industry FTE is expressed in thousands and wage per FTE in dollars per FTE.

Availability is defined at the series level and, for Industry data, can also vary by entity. A dataset's first and last period therefore describe the span of observed data in the file, not a guarantee that every series is available throughout that interval.

The National, Corporate, and Industry methodology sections summarize the principal coverage boundaries. The public series metadata provide the exact availability rule and notes for each published series.

12.2 Cross-frequency and historical comparability

Annual and quarterly data can differ in source coverage as well as frequency. National Annual uses source-reported Gross Operating Surplus as the primary EBITDA analogue, while National Quarterly uses calculated EBITDA. Corporate Annual and Quarterly share the same main statement concepts but have different historical start dates for some supplemental series. Industry Quarterly supports only the production block and Gross Margin, while Industry Annual extends further down the statement and includes supplemental labor and income concepts.

Cross-frequency comparisons should therefore use measures with the same conceptual basis at both frequencies and respect the quarterly SAAR convention. Differences in availability are preserved rather than filled solely to make annual and quarterly statements visually symmetric.

Annual Industry series that cross the 1996/1997 boundary combine reconciled historical evidence with the current publication structure under the rules in Section 11. The splice improves long-run comparability but does not imply that every current Industry category has a fully reconstructable history. Bridge concepts and deliberate omissions remain where the historical evidence does not support the modern split.

An unavailable public observation is absent from the observation file. Analysts constructing rectangular panels should introduce explicit missing values after joining their desired entity, series, and period universe. Missing observations should never be interpreted as numerical zero without a separate substantive reason.

13 Public Data Structure and Metadata

13.1 Public observations and scope

Table 13.1 — Stage 4 public observation schema

Field	Type	Meaning	Rule / reference
`dataset_id`	string	Statement family and frequency.	Allowed values: `NIS-A`; `NIS-Q`; `CIS-A`; `CIS-Q`; `IIS-A`; `IIS-Q`
`entity_id`	string	Economic entity or scope represented by the observation.	National, corporate, or Industry entity.
`series_id`	string	Series key within a dataset.	Interpreted jointly with the dataset in series metadata.
`period`	string	Reference period.	Annual: `YYYY`; quarterly: `YYYYQn`.
`period_end`	date	Last day of the reference period.	ISO 8601 date.
`value`	number	Observation in the series' declared unit and natural sign.	Unit from series metadata; display treatment from layout metadata.
`source_type`	string	Observation-level provenance route.	Controlled provenance category; exact allowed values are documented in the public schema.

Note: The observation key combines the dataset, entity, series, and period fields. Missing observations are absent rather than zero-filled. Observation provenance is distinct from series-level measurement basis.

Source: EARL Stage 4 public schema; reader-facing projection of the seven observation-file fields.

The public data package contains six observation datasets and shared metadata describing their economic meaning, presentation, entities, source vintage, build identity, and schema. The current specification contains 128 dataset-series combinations: 97 source-reported and 31 EARL-calculated. Their public roles comprise 93 core-statement series, 33 supplemental-statement series, and two public-data-only series.

The package contains 101 public entities: National, Corporate, and 99 Industry entities. Build-specific observation counts are reported with the public-data provenance record in Appendix G and are refreshed with the publication build.

Table 13.1 defines the seven-field observation schema. Every public observation uses seven fields:

`dataset_id, entity_id, series_id, period, period_end, value, source_type`.

The observation key is the combination of dataset, entity, series, and period. Missing observations are represented by absence from the file. The stored value retains its economic sign; any statement-specific sign inversion belongs to presentation metadata.

Stage 4 selects the defined nonmissing public observations from the processed semantic data. It preserves the Stage 3 value and observation provenance. This boundary ensures that publication changes do not create a second economic calculation layer.

13.2 Metadata, statement layout, and provenance

Series metadata describe the public label, official BEA concept, definition, concept class, measurement basis, formula and inputs where applicable, unit, seasonal-adjustment and annualization treatment, availability, source information, methodology notes, and public role.

Entity metadata describe National, Corporate, and Industry entities and the Industry parent relationship. Statement-layout metadata define where a series appears in a statement, its section and row order, indentation, row type, display sign, and availability policy. The current layout contains 132 display rows covering 126 unique dataset-series combinations. Six additional rows are deliberate repeated uses of the same series in different statement contexts; the two National net-production-tax series published outside the statement layout are public-data-only.

Each observation file has a compact dataset companion describing its seven fields and linking users to the shared metadata. Shared source-vintage metadata identify the official release represented by the current acquisitions. Build metadata identify the public-data build, source revision, observation file, source acquisitions, source vintages, and retrieval window.

The metadata are separated by function so that observation files remain compact and a source release or build identity does not have to be repeated on every row. Appendix G gives the complete relational schema; Appendix H gives the detailed provenance schema.

13.3 Public package and deterministic output

The full package contains 18 CSV files: six observation files, six dataset companions, series metadata, statement layout, entity metadata, source vintages, build metadata, and the public schema. Serialization is deterministic so that unchanged data and metadata produce unchanged public artifacts. Determinism supports reproducibility; it complements, rather than substitutes for, validation of source mappings, formulas, and economic identities.

14 Provenance, Vintages, and Revisions

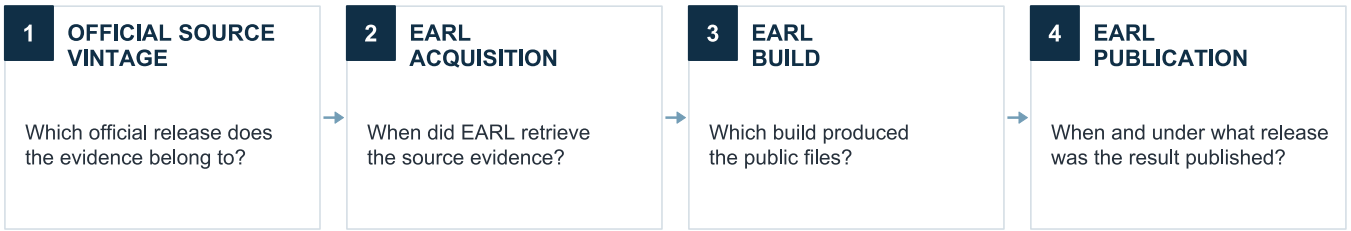
14.1 Source release, acquisition, build, and publication

EARL separates four dates that answer different questions, summarized in Figure 14.1:

1. **Source release/vintage:** when BEA released or revised the source data represented by the observation.
2. **Acquisition:** when EARL retrieved the source evidence.
3. **Build:** when EARL generated the public data package.
4. **Publication:** when a specific EARL publication was released.

The Methodological Status record identifies the source release and build represented by this manuscript; Appendix H provides the detailed acquisition, source-vintage, and build identifiers. Publication has its own release-specific date.

Figure 14.1 — The four provenance clocks



NIPA provides direct machine evidence for its release date through `LastRevised`. GDP by Industry does not expose an equivalent API release-date field, so its vintage assignment is linked through the source-vintage metadata to the applicable official BEA release. Each table retrieval has a separate acquisition identity. Multiple retrievals can therefore belong to the same official source vintage.

14.2 Build identity and observation provenance

Build metadata identify the exact public-data build, source revision, constituent acquisitions, source vintages, and artifact hashes. A Statistical Release has a separate release identity, reference period, version, source cutoff, and publication timestamp; those publication fields complement rather than replace the provenance attached to the data. The exact build record is refreshed immediately before publication as described in Appendix H.

Observation provenance records how each cell entered the public dataset. Appendix G reports the build-specific provenance counts. The provenance field complements the series-level measurement basis: a source-reported series can contain a hierarchy-rollup observation because the observation still represents the official source concept, while an EARL-calculated series uses `derived` provenance because its value is created by a semantic formula.

14.3 Revisions and historical provenance

Published values can change because BEA revises a source, extends a source series, or because EARL changes a mapping, formula, inclusion rule, or implementation. Same-vintage comparisons are used to distinguish source revision from changes introduced within EARL. Unexplained missing observations, unexplained same-vintage value differences, and duplicate keys are treated as errors rather than being attributed automatically to source revision.

Historical annual Industry observations follow a separate provenance branch built from the retained November 30, 2021 source, the reviewed crosswalk, and the 1997 anchor information. They carry `bea_historical_2021` observation provenance and remain distinguishable from current API observations after the annual series are spliced.

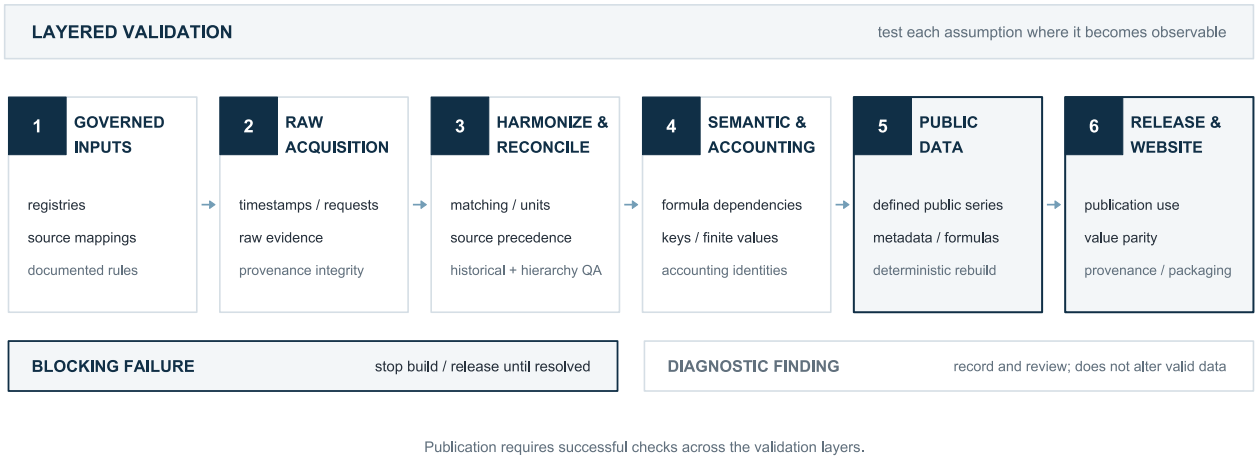
15 Validation and Quality Assurance

15.1 Validation architecture and source controls

EARL validates assumptions at the point where they first become observable. Acquisition checks establish that the required source evidence and source metadata are present. Harmonization checks establish that source concepts, units, periods, and entity mappings are consistent with the documented definitions. Semantic checks validate formulas and accounting relationships. Public-data checks verify that published observations match the processed data and that the metadata describe the published series correctly. Publication checks confirm that statements, charts, and releases use those published values consistently.

A structural or methodological failure blocks the affected build or publication. Diagnostic findings—such as expected missingness, different latest periods across series, or unusual but economically possible ratios—are retained for review without altering otherwise valid observations. Figure 15.1 summarizes these validation gates.

Figure 15.1 — EARL validation gates



Publication requires successful checks across the validation layers.

Source validation checks the identifiers required by each mapping. Fixed-line NIPA series are tested against the expected table, line, and line description. GDP by Industry mappings are tested against table, Industry scope, and source concept. Industry-level NIPA mappings are tested against the entity description and ordering rules used by the source table. Units and period conventions are validated before conversion.

Harmonization validation checks source precedence, complete-child requirements for hierarchy rollups, the limited aggregate GVA construction, and the historical/current splice. Semantic validation checks formula dependencies, guarded division, unique observation keys, finite values, provenance consistency, and the dataset-specific accounting identities.

15.2 Accounting and historical validation

Tolerance levels are relationship-specific and reflect the precision of the underlying published sources. National Annual reported Gross Operating Surplus and calculated EBITDA are compared within $100 million. Corporate calculated EBITDA and the source relationship `NOS + CFC` are compared within $5 million. Annual Industry production and input identities generally use $150 million; selected aggregate cross-table value-added relationships use $250 million. Derived arithmetic in the Statistical Release uses a tighter $1 million tolerance where the values come from the same published inputs.

Historical validation additionally checks index normalization, preservation of historical growth under the 1997 re-anchoring, and direct agreement with selected retained source observations. Section 11 reports the historical validation results; Appendix I records the validation matrix.

15.3 Revision, public-data, and publication validation

Same-vintage adjudication separates changes in official source data from changes in EARL processing. A value difference is classified as a source revision only when the comparison evidence supports that explanation; otherwise it remains unresolved and blocks acceptance.

Public-data validation checks that every published observation corresponds to the defined processed observation with the same value and provenance. The formulas recorded in the public metadata are independently evaluated from public input series. In the documented build, all 26,866 published calculated observations reproduce from those formulas; the largest independently measured floating-point difference is approximately 3.55×10^{-15}.

Entity metadata, statement layout, source-vintage metadata, dataset companions, and the public schema are validated together with the observations. The purpose is to ensure that a numerically correct value is also accompanied by the correct definition, scope, unit, and presentation rule.

The Statistical Release and website are checked as consumers of the public data. Publication validation covers the National and Corporate accounting bridges, displayed values, presentation signs, provenance, release metadata, rendering, and package integrity. Website validation checks that presentation transformations—such as filtering, pivoting, rounding, and percentage formatting—do not create alternative economic values.

Blocking findings include duplicate observation keys, unresolved source mappings, unexplained same-vintage differences, broken accounting identities outside their specified tolerances, invalid observation provenance, and contradictions between the data and their metadata. Diagnostic findings remain visible when they reflect legitimate economic or source characteristics.

16 Interpretation, Reproducibility, and Governance

16.1 Interpretation boundaries

EARL uses business terminology to organize aggregate economic-account concepts. The labels should always be interpreted together with the official source concept and measurement basis. Gross Revenue represents Gross Output, Gross Profit represents Gross Value Added, and the EBITDA and EBIT analogues use the source or calculated measures stated in Sections 8–10. These aggregates do not follow company-reporting definitions under GAAP, IFRS, or tax accounting.

The same business-style label can have different measurement bases across scopes or frequencies. National Annual and Industry Annual present BEA Gross Operating Surplus as the reported EBITDA analogue. National Quarterly and Corporate use EARL-calculated EBITDA. Similar labels therefore do not imply identical source construction.

National and Industry measures are aggregates rather than representative firms. Gross Output includes inter-industry transactions, and Industry margins combine heterogeneous establishments. Missing observations indicate unavailable or unsupported data, not zero. Historical reconciliation extends comparability only where the retained source evidence supports the mapping; combined historical categories remain combined when a current split cannot be reconstructed.

16.2 Revisions and reproducibility

Official economic statistics are revised as source agencies incorporate new information and update classifications or methods. EARL can therefore publish different values for the same historical period under a later source vintage without changing its own methodology. The provenance system identifies the source release, acquisition, build, and publication so that these cases can be distinguished from a change in an EARL formula or mapping.

Reproducibility includes the evidentiary chain as well as executable code. A published observation can be traced through its series definition, source mapping or formula, source evidence, source vintage, acquisition, and build. Stored acquisition evidence permits same-vintage reprocessing even after the live source has been revised.

Deterministic serialization provides an additional check: unchanged data, metadata, and code should produce unchanged public artifacts. Determinism does not establish that a mapping or economic interpretation is correct by itself, so it is used together with source verification, accounting identities, metadata review, and publication validation.

16.3 Governance, identifiers, and scope

A material change to an economic definition, source mapping, formula, Industry relationship, historical crosswalk, splice rule, or public-series scope is documented in the methodological metadata and version record. Implementation and validation are updated with the same change. Presentation-only changes—such as formatting, layout, or branding—do not change the economic methodology when the underlying values and definitions are unchanged.

Methodological metadata, implementation code, source evidence, and validation provide complementary forms of authority. Metadata state the definitions and rules; code implements them; source evidence anchors the official concepts; validation tests whether the resulting data follow the documented method. Publication products use the resulting public data and can add document-specific metadata without redefining the economic series.

Earlier NISP and USEAL identifiers remain in some historical releases, machine fields, repository paths, and test names. They are retained where required for reproducibility or compatibility. Current public terminology uses EARL.

The methodology in this paper applies to the six National, Corporate, and Industry datasets described in Section 1. Extensions to new datasets or new source concepts should be documented through the same source-mapping, formula, provenance, and validation framework before they are incorporated into the public data.

17 Conclusion

EARL provides a consistent way to analyze selected U.S. economic accounts through a business-style statement structure while preserving the official concepts that underlie each measure. Gross Output remains the source concept behind Gross Revenue, Gross Value Added behind Gross Profit, and the EBITDA and EBIT analogues retain the source or formula basis documented for each scope and frequency.

The methodology accommodates differences in the underlying accounts rather than forcing identical statements across datasets. National Annual and Industry Annual can use source-reported Gross Operating Surplus as an EBITDA analogue; National Quarterly and Corporate use calculated EBITDA; Corporate begins at Gross Value Added; and Industry Quarterly ends at Gross Profit and Gross Margin. These differences reflect source availability and measurement structure.

Construction is separated into acquisition, harmonization and reconciliation, semantic calculation, and public-data publication. Historical Industry observations enter that framework through an explicit crosswalk, a 1996/1997 current-dollar splice, and 1997 index re-anchoring. Unsupported historical splits remain combined or unavailable. Provenance records the source route of each observation and separates source release, acquisition, build, and publication dates.

Validation follows the same layered structure, covering source matching, units, formulas, accounting identities, historical reconciliation, public metadata, and publication outputs. Together with stable series and entity definitions, these controls allow a published EARL measure to be traced from its business-style presentation back to its economic definition, calculation rule where applicable, and official source evidence. That traceability is the central requirement for reproducible use of the framework.

APPENDICES

Technical reference material

A Complete Public Series Dictionary

Appendix A is the complete dataset-series dictionary for the Stage 4 public data specification documented in this paper. One row represents one public `dataset_id + series_id` combination. Observed coverage reports the first and last *published observation* in the documented build; it does not imply a balanced panel for every intervening period or, for Industry datasets, every entity. The public specification contains **128** dataset-series combinations: **97 source-reported** and **31 EARL-calculated**. Public-role counts are **93 core statement**, **33 supplemental statement**, and **2 public-data-only**.

A.1 Complete dictionary

Table A.1 is the complete public series dictionary for the documented build.

Table A.1 — Complete public series dictionary

Dataset	Series ID	EARL label	BEA concept	Concept / basis	Unit / coverage	Entities	Public role
NIS-A	gross_output	Gross Revenue	Gross output	EARL accounting translation; source-reported	USD millions · 1947–2025	1	core statement
NIS-A	intermediate_inputs	Cost of Goods Sold	Intermediate inputs	EARL accounting translation; source-reported	USD millions · 1947–2025	1	core statement
NIS-A	intermediate_inputs_energy	Energy Inputs	Energy inputs	official BEA concept; source-reported	USD millions · 1963–2025	1	core statement
NIS-A	intermediate_inputs_materials	Materials Inputs	Materials inputs	official BEA concept; source-reported	USD millions · 1963–2025	1	core statement
NIS-A	intermediate_inputs_services	Purchased Services Inputs	Purchased-services inputs	official BEA concept; source-reported	USD millions · 1963–2025	1	core statement
NIS-A	gross_value_added	Gross Profit	Gross domestic product	EARL accounting translation; source-reported	USD millions · 1929–2025	1	core statement
NIS-A	wages_and_salaries	Wages and Salaries	Wages and salaries	official BEA concept; source-reported	USD millions · 1929–2025	1	core statement
NIS-A	supplements_to_wages_and_salaries	Supplements to Wages and Salaries	Supplements to wages and salaries	official BEA concept; source-reported	USD millions · 1929–2025	1	core statement
NIS-A	compensation_of_employees	Total Employee Compensation	Compensation of employees, paid	official BEA concept; source-reported	USD millions · 1929–2025	1	core statement
NIS-A	taxes_on_production_and_imports	Taxes on Production and Imports	Taxes on production and imports	official BEA concept; source-reported	USD millions · 1929–2025	1	core statement
NIS-A	subsidies	Less: Subsidies	Less: Subsidies	official BEA concept; source-reported	USD millions · 1929–2025	1	core statement
NIS-A	operating_expenses	Total Operating Expenses	—	EARL-calculated measure; EARL-calculated	USD millions · 1929–2025	1	core statement
NIS-A	ebitda_gos	EBITDA (Reported BEA Gross Operating Surplus)	Gross operating surplus	EARL accounting translation; source-reported	USD millions · 1997–2025	1	core statement
NIS-A	consumption_of_fixed_capital	Depreciation	Consumption of fixed capital	EARL accounting translation; source-reported	USD millions · 1929–2025	1	core statement
NIS-A	net_operating_surplus	Operating Income (EBIT)	Net Operating Surplus	EARL accounting translation; source-reported	USD millions · 1959–2025	1	core statement
NIS-A	earnings_before_tax	Income Before Taxes (EBT)	—	EARL-calculated measure; EARL-calculated	USD millions · 1959–2025	1	core statement

Dataset	Series ID	EARL label	BEA concept	Concept / basis	Unit / coverage	Entities	Public role
NIS-A	gross_margin	Gross Margin	—	EARL-calculated measure; EARL-calculated	ratio · 1947–2025	1	core statement
NIS-A	ebitda_margin_gos	EBITDA Margin (Reported-GOS Basis)	—	EARL-calculated measure; EARL-calculated	ratio · 1997–2025	1	core statement
NIS-A	ebit_margin	Operating Margin (EBIT)	—	EARL-calculated measure; EARL-calculated	ratio · 1959–2025	1	core statement
NIS-A	ebt_margin	EBT Margin	—	EARL-calculated measure; EARL-calculated	ratio · 1959–2025	1	core statement
NIS-A	ebitda_calculated	EBITDA (Calculated; reconciliation only)	—	EARL-calculated measure; EARL-calculated	USD millions · 1929–2025	1	supplemental statement
NIS-A	net_operating_surplus_output_perspective	EBIT — Output Perspective	—	EARL-calculated measure; EARL-calculated	USD millions · 1997–2025	1	supplemental statement
NIS-A	statistical_discrepancy	Statistical Discrepancy	Less: Statistical Discrepancy	official BEA concept; source-reported	USD millions · 1929–2025	1	supplemental statement
NIS-A	interest_expense	Interest Expense	Net interest and miscellaneous payments, domestic industries	EARL accounting translation; source-reported	USD millions · 1929–2025	1	core statement
NIS-A	business_current_transfer_payments_net	Business Current Transfer Payments	Business current transfer payments (net)	official BEA concept; source-reported	USD millions · 1929–2025	1	supplemental statement
NIS-A	proprietors_income_with_iva_ccadj	Proprietors' Income	Proprietors' income with inventory valuation and capital consumption adjustments	official BEA concept; source-reported	USD millions · 1929–2025	1	supplemental statement
NIS-A	rental_income_with_ccadj	Rental Income	Rental income of persons with capital consumption adjustment	official BEA concept; source-reported	USD millions · 1929–2025	1	supplemental statement
NIS-A	government_enterprise_surplus	Government Enterprise Surplus	Current surplus of government enterprises	official BEA concept; source-reported	USD millions · 1959–2025	1	supplemental statement
NIS-A	corporate_profits_with_iva_ccadj	Corporate Profit	Corporate profits with inventory valuation and capital consumption adjustments, domestic industries	official BEA concept; source-reported	USD millions · 1929–2025	1	supplemental statement
NIS-A	taxes_on_production_and_imports_less_subsidies	Taxes on Production & Imports, net of Subsidies	—	EARL-calculated measure; EARL-calculated	USD millions · 1929–2025	1	public data only
NIS-Q	gross_output	Gross Revenue	Gross output	EARL accounting translation; source-reported	USD millions, SAAR · 2005Q1–2026Q2	1	core statement

Dataset	Series ID	EARL label	BEA concept	Concept / basis	Unit / coverage	Entities	Public role
NIS-Q	intermediate_inputs	Cost of Goods Sold	Intermediate inputs	EARL accounting translation; source-reported	USD millions, SAAR · 2005Q1–2026Q2	1	core statement
NIS-Q	gross_value_added	Gross Profit	Gross domestic product	EARL accounting translation; source-reported	USD millions, SAAR · 1947Q1–2026Q2	1	core statement
NIS-Q	wages_and_salaries	Wages and Salaries	Wages and salaries	official BEA concept; source-reported	USD millions, SAAR · 1947Q1–2026Q2	1	core statement
NIS-Q	supplements_to_wages_and_salaries	Supplements to Wages and Salaries	Supplements to wages and salaries	official BEA concept; source-reported	USD millions, SAAR · 1947Q1–2026Q2	1	core statement
NIS-Q	compensation_of_employees	Total Employee Compensation	Compensation of employees, paid	official BEA concept; source-reported	USD millions, SAAR · 1947Q1–2026Q2	1	core statement
NIS-Q	taxes_on_production_and_imports	Taxes on Production and Imports	Taxes on production and imports	official BEA concept; source-reported	USD millions, SAAR · 1947Q1–2026Q2	1	core statement
NIS-Q	subsidies	Less: Subsidies	Less: Subsidies	official BEA concept; source-reported	USD millions, SAAR · 1947Q1–2026Q2	1	core statement
NIS-Q	operating_expenses	Total Operating Expenses	—	EARL-calculated measure; EARL-calculated	USD millions, SAAR · 1947Q1–2026Q2	1	core statement
NIS-Q	ebitda_calculated	EBITDA (Calculated)	—	EARL-calculated measure; EARL-calculated	USD millions, SAAR · 1947Q1–2026Q2	1	core statement
NIS-Q	consumption_of_fixed_capital	Depreciation	Consumption of fixed capital	EARL accounting translation; source-reported	USD millions, SAAR · 1947Q1–2026Q2	1	core statement
NIS-Q	net_operating_surplus	Operating Income (EBIT)	Net Operating Surplus	EARL accounting translation; source-reported	USD millions, SAAR · 1959Q1–2026Q2	1	core statement
NIS-Q	earnings_before_tax	Income Before Taxes (EBT)	—	EARL-calculated measure; EARL-calculated	USD millions, SAAR · 1959Q1–2026Q2	1	core statement
NIS-Q	gross_margin	Gross Margin	—	EARL-calculated measure; EARL-calculated	ratio · 2005Q1–2026Q2	1	core statement
NIS-Q	ebitda_margin_calculated	EBITDA Margin (Calculated-EBITDA Basis)	—	EARL-calculated measure; EARL-calculated	ratio · 2005Q1–2026Q2	1	core statement

Dataset	Series ID	EARL label	BEA concept	Concept / basis	Unit / coverage	Entities	Public role
NIS-Q	ebit_margin	Operating Margin (EBIT)	—	EARL-calculated measure; EARL-calculated	ratio · 2005Q1–2026Q2	1	core statement
NIS-Q	ebt_margin	EBT Margin	—	EARL-calculated measure; EARL-calculated	ratio · 2005Q1–2026Q2	1	core statement
NIS-Q	statistical_discrepancy	Statistical Discrepancy	Less: Statistical Discrepancy	official BEA concept; source-reported	USD millions, SAAR · 1947Q1–2026Q2	1	supplemental statement
NIS-Q	interest_expense	Interest Expense	Net interest and miscellaneous payments, domestic industries	EARL accounting translation; source-reported	USD millions, SAAR · 1947Q1–2026Q2	1	core statement
NIS-Q	business_current_transfer_payments_net	Business Current Transfer Payments	Business current transfer payments (net)	official BEA concept; source-reported	USD millions, SAAR · 1947Q1–2026Q2	1	supplemental statement
NIS-Q	proprietors_income_with_iva_ccadj	Proprietors' Income	Proprietors' income with inventory valuation and capital consumption adjustments	official BEA concept; source-reported	USD millions, SAAR · 1947Q1–2026Q2	1	supplemental statement
NIS-Q	rental_income_with_ccadj	Rental Income	Rental income of persons with capital consumption adjustment	official BEA concept; source-reported	USD millions, SAAR · 1947Q1–2026Q2	1	supplemental statement
NIS-Q	government_enterprise_surplus	Government Enterprise Surplus	Current surplus of government enterprises	official BEA concept; source-reported	USD millions, SAAR · 1959Q1–2026Q2	1	supplemental statement
NIS-Q	corporate_profits_with_iva_ccadj	Corporate Profit	Corporate profits with inventory valuation and capital consumption adjustments, domestic industries	official BEA concept; source-reported	USD millions, SAAR · 1947Q1–2026Q2	1	supplemental statement
NIS-Q	taxes_on_production_and_imports_less_subsidies	Taxes on Production & Imports, net of Subsidies	—	EARL-calculated measure; EARL-calculated	USD millions, SAAR · 1947Q1–2026Q2	1	public data only
CIS-A	gross_value_added	Gross Profit	Gross value added of corporate business	EARL accounting translation; source-reported	USD millions · 1929–2025	1	core statement
CIS-A	wages_and_salaries	Wages and Salaries	Wages and salaries	official BEA concept; source-reported	USD millions · 1929–2025	1	core statement
CIS-A	supplements_to_wages_and_salaries	Supplements to Wages and Salaries	Supplements to wages and salaries	official BEA concept; source-reported	USD millions · 1929–2025	1	core statement
CIS-A	compensation_of_employees	Total Employee Compensation	Compensation of employees	official BEA concept; source-reported	USD millions · 1929–2025	1	core statement
CIS-A	taxes_on_production_and_imports_less_subsidies	Taxes on Production & Imports, net of Subsidies	Taxes on production and	official BEA concept; source-reported	USD millions · 1929–2025	1	core statement

Dataset	Series ID	EARL label	BEA concept	Concept / basis	Unit / coverage	Entities	Public role
			imports less subsidies				
CIS-A	operating_expenses	Total Operating Expenses	—	EARL-calculated measure; EARL-calculated	USD millions · 1929–2025	1	core statement
CIS-A	ebitda_calculated	EBITDA (Calculated)	—	EARL-calculated measure; EARL-calculated	USD millions · 1929–2025	1	core statement
CIS-A	consumption_of_fixed_capital	Depreciation	Consumption of fixed capital	EARL accounting translation; source-reported	USD millions · 1929–2025	1	core statement
CIS-A	net_operating_surplus	Operating Income (EBIT)	Net operating surplus	EARL accounting translation; source-reported	USD millions · 1929–2025	1	core statement
CIS-A	interest_expense	Interest Expense / (Income)	Net interest and miscellaneous payments	EARL accounting translation; source-reported	USD millions · 1929–2025	1	core statement
CIS-A	business_current_transfer_payments_net	Business Current Transfer Payments	Business current transfer payments (net)	official BEA concept; source-reported	USD millions · 1929–2025	1	core statement
CIS-A	ebt_with_iva_ccadj	Income Before Taxes (EBT) — with IVA & CCAdj	Corporate profits with IVA and CCAdj	EARL accounting translation; source-reported	USD millions · 1929–2025	1	core statement
CIS-A	net_income_with_iva_ccadj	Net Income — with IVA & CCAdj	Profits after tax with IVA and CCAdj	EARL accounting translation; source-reported	USD millions · 1929–2025	1	core statement
CIS-A	retained_earnings_with_iva_ccadj	Retained Earnings	Undistributed profits with IVA and CCAdj	EARL accounting translation; source-reported	USD millions · 1929–2025	1	core statement
CIS-A	effective_tax_rate	Effective Tax Rate	—	EARL-calculated measure; EARL-calculated	ratio · 1929–2025	1	core statement
CIS-A	payout_ratio	Payout Ratio	—	EARL-calculated measure; EARL-calculated	ratio · 1929–2025	1	core statement
CIS-A	retention_ratio	Retention Ratio	—	EARL-calculated measure; EARL-calculated	ratio · 1929–2025	1	core statement
CIS-A	inventory_valuation_adjustment	Inventory Valuation Adjustment (IVA)	Inventory valuation adjustment	official BEA concept; source-reported	USD millions · 1929–2025	1	supplemental statement
CIS-A	capital_consumption_adjustment	Capital Consumption Adjustment (CCAdj)	Capital consumption adjustment	official BEA concept; source-reported	USD millions · 1929–2025	1	supplemental statement
CIS-A	ebt_without_iva_ccadj	Income Before Taxes (EBT) — without IVA & CCAdj	Profits before tax (without IVA and CCAdj)	EARL accounting translation; source-reported	USD millions · 1929–2025	1	supplemental statement
CIS-A	taxes_on_corporate_income	Income Tax	Taxes on corporate income	official BEA concept; source-reported	USD millions · 1929–2025	1	core statement
CIS-A	net_income_without_iva_ccadj	Net Income — without IVA & CCAdj	Profits after tax (without IVA and CCAdj)	EARL accounting translation; source-reported	USD millions · 1929–2025	1	supplemental statement
CIS-A	net_dividends	Dividends	Net dividends	EARL accounting translation; source-reported	USD millions · 1929–2025	1	core statement

Dataset	Series ID	EARL label	BEA concept	Concept / basis	Unit / coverage	Entities	Public role
CIS-A	retained_earnings_without_iva_ccadj	Retained Earnings — without IVA & CCAdj	Undistributed profits after tax (without IVA and CCAdj)	EARL accounting translation; source-reported	USD millions · 1929–2025	1	supplemental statement
CIS-A	corporate_profits_receipts_from_rest_of_world	Receipts from Rest of World	Receipts from the rest of the world	official BEA concept; source-reported	USD millions · 1998–2025	1	supplemental statement
CIS-A	corporate_profits_payments_to_rest_of_world	Less: Payments to Rest of World	Less: Payments to the rest of the world	official BEA concept; source-reported	USD millions · 1998–2025	1	supplemental statement
CIS-Q	gross_value_added	Gross Profit	Gross value added of corporate business	EARL accounting translation; source-reported	USD millions, SAAR · 1947Q1–2026Q2	1	core statement
CIS-Q	wages_and_salaries	Wages and Salaries	Wages and salaries	official BEA concept; source-reported	USD millions, SAAR · 1947Q1–2026Q2	1	core statement
CIS-Q	supplements_to_wages_and_salaries	Supplements to Wages and Salaries	Supplements to wages and salaries	official BEA concept; source-reported	USD millions, SAAR · 1947Q1–2026Q2	1	core statement
CIS-Q	compensation_of_employees	Total Employee Compensation	Compensation of employees	official BEA concept; source-reported	USD millions, SAAR · 1947Q1–2026Q2	1	core statement
CIS-Q	taxes_on_production_and_imports_less_subsidies	Taxes on Production & Imports, net of Subsidies	Taxes on production and imports less subsidies	official BEA concept; source-reported	USD millions, SAAR · 1947Q1–2026Q2	1	core statement
CIS-Q	operating_expenses	Total Operating Expenses	—	EARL-calculated measure; EARL-calculated	USD millions, SAAR · 1947Q1–2026Q2	1	core statement
CIS-Q	ebitda_calculated	EBITDA (Calculated)	—	EARL-calculated measure; EARL-calculated	USD millions, SAAR · 1947Q1–2026Q2	1	core statement
CIS-Q	consumption_of_fixed_capital	Depreciation	Consumption of fixed capital	EARL accounting translation; source-reported	USD millions, SAAR · 1947Q1–2026Q2	1	core statement
CIS-Q	net_operating_surplus	Operating Income (EBIT)	Net operating surplus	EARL accounting translation; source-reported	USD millions, SAAR · 1947Q1–2026Q2	1	core statement
CIS-Q	interest_expense	Interest Expense / (Income)	Net interest and miscellaneous payments	EARL accounting translation; source-reported	USD millions, SAAR · 1947Q1–2026Q2	1	core statement
CIS-Q	business_current_transfer_payments_net	Business Current Transfer Payments	Business current transfer payments (net)	official BEA concept; source-reported	USD millions, SAAR · 1947Q1–2026Q2	1	core statement

Dataset	Series ID	EARL label	BEA concept	Concept / basis	Unit / coverage	Entities	Public role
CIS-Q	ebt_with_iva_ccadj	Income Before Taxes (EBT) — with IVA & CCAdj	Corporate profits with IVA and CCAdj	EARL accounting translation; source-reported	USD millions, SAAR · 1947Q1–2026Q2	1	core statement
CIS-Q	net_income_with_iva_ccadj	Net Income — with IVA & CCAdj	Profits after tax with IVA and CCAdj	EARL accounting translation; source-reported	USD millions, SAAR · 1947Q1–2026Q2	1	core statement
CIS-Q	retained_earnings_with_iva_ccadj	Retained Earnings	Undistributed profits with IVA and CCAdj	EARL accounting translation; source-reported	USD millions, SAAR · 1947Q1–2026Q2	1	core statement
CIS-Q	effective_tax_rate	Effective Tax Rate	—	EARL-calculated measure; EARL-calculated	ratio · 1947Q1–2026Q2	1	core statement
CIS-Q	payout_ratio	Payout Ratio	—	EARL-calculated measure; EARL-calculated	ratio · 1947Q1–2026Q2	1	core statement
CIS-Q	retention_ratio	Retention Ratio	—	EARL-calculated measure; EARL-calculated	ratio · 1947Q1–2026Q2	1	core statement
CIS-Q	inventory_valuation_adjustment	Inventory Valuation Adjustment (IVA)	Inventory valuation adjustment	official BEA concept; source-reported	USD millions, SAAR · 1947Q1–2026Q2	1	supplemental statement
CIS-Q	capital_consumption_adjustment	Capital Consumption Adjustment (CCAdj)	Capital consumption adjustment	official BEA concept; source-reported	USD millions, SAAR · 1947Q1–2026Q2	1	supplemental statement
CIS-Q	ebt_without_iva_ccadj	Income Before Taxes (EBT) — without IVA & CCAdj	Profits before tax (without IVA and CCAdj)	EARL accounting translation; source-reported	USD millions, SAAR · 1947Q1–2026Q2	1	supplemental statement
CIS-Q	taxes_on_corporate_income	Income Tax	Taxes on corporate income	official BEA concept; source-reported	USD millions, SAAR · 1947Q1–2026Q2	1	core statement
CIS-Q	net_income_without_iva_ccadj	Net Income — without IVA & CCAdj	Profits after tax (without IVA and CCAdj)	EARL accounting translation; source-reported	USD millions, SAAR · 1947Q1–2026Q2	1	supplemental statement
CIS-Q	net_dividends	Dividends	Net dividends	EARL accounting translation; source-reported	USD millions, SAAR · 1947Q1–2026Q2	1	core statement
CIS-Q	retained_earnings_without_iva_ccadj	Retained Earnings — without IVA & CCAdj	Undistributed profits after tax (without IVA and CCAdj)	EARL accounting translation; source-reported	USD millions, SAAR · 1947Q1–2026Q2	1	supplemental statement
CIS-Q	corporate_profits_receipts_from_rest_of_world	Receipts from Rest of World	Receipts from the rest of the world	official BEA concept; source-reported	USD millions, SAAR · 2001Q1–2026Q2	1	supplemental statement
CIS-Q	corporate_profits_payments_to_rest_of_world	Less: Payments to Rest of World	Less: Payments to the rest of the world	official BEA concept; source-reported	USD millions, SAAR ·	1	supplemental statement

Dataset	Series ID	EARL label	BEA concept	Concept / basis	Unit / coverage	Entities	Public role
					2001Q1–2026Q2		
IIS-A	gross_output	Gross Revenue	Gross output	EARL accounting translation; source-reported	USD millions · 1947–1996	99	core statement
IIS-A	intermediate_inputs	Cost of Goods Sold	Intermediate inputs	EARL accounting translation; source-reported	USD millions · 1947–1996	99	core statement
IIS-A	intermediate_inputs_energy	Energy Inputs	Energy inputs	official BEA concept; source-reported	USD millions · 1963–2025	98	core statement
IIS-A	intermediate_inputs_materials	Materials Inputs	Materials inputs	official BEA concept; source-reported	USD millions · 1963–2025	98	core statement
IIS-A	intermediate_inputs_services	Purchased Services Inputs	Purchased-services inputs	official BEA concept; source-reported	USD millions · 1963–2025	98	core statement
IIS-A	gross_value_added	Gross Profit	Varies by industry (see notes)	EARL accounting translation; source-reported	USD millions · 1947–1996	99	core statement
IIS-A	compensation_of_employees	Employee Compensation	Compensation of employees	official BEA concept; source-reported	USD millions · 1987–2025	98	core statement
IIS-A	taxes_on_production_and_imports_less_subsidies	Taxes on Production & Imports, net of Subsidies	Taxes on production and imports less subsidies	official BEA concept; source-reported	USD millions · 1987–2025	94	core statement
IIS-A	operating_expenses	Total Operating Expenses	—	EARL-calculated measure; EARL-calculated	USD millions · 1987–2025	94	core statement
IIS-A	ebitda_gos	EBITDA (Reported BEA Gross Operating Surplus)	Gross operating surplus	EARL accounting translation; source-reported	USD millions · 1987–2025	98	core statement
IIS-A	gross_margin	Gross Margin	—	EARL-calculated measure; EARL-calculated	ratio · 1947–1996	99	core statement
IIS-A	ebitda_margin_gos	EBITDA Margin (Reported-GOS Basis)	—	EARL-calculated measure; EARL-calculated	ratio · 1987–2025	98	core statement
IIS-A	wages_and_salaries	Wages and Salaries	Wages and salaries / wage and salary accruals by industry	official BEA concept; source-reported	USD millions · 1998–2025	93	supplemental statement
IIS-A	fte_employees	Full-Time Equivalent Employees	Full-time equivalent employees by industry	official BEA concept; source-reported	thousand FTE · 1998–2025	93	supplemental statement
IIS-A	wage_and_salary_accruals_per_fte	Wage and Salary Accruals per FTE	Wage and salary accruals per full-time equivalent employee by industry	official BEA concept; source-reported	USD per FTE · 1998–2025	93	supplemental statement
IIS-A	interest_expense	Interest Expense / (Income)	Net interest / net interest and miscellaneous payments	EARL accounting translation; source-reported	USD millions · 1998–2025	23	supplemental statement
IIS-A	proprietors_income_with_iva_ccadj	Proprietors' Income	Nonfarm proprietors' income by industry	official BEA concept; source-reported	USD millions · 1998–2025	26	supplemental statement

Dataset	Series ID	EARL label	BEA concept	Concept / basis	Unit / coverage	Entities	Public role
IIS-Q	gross_output	Gross Revenue	Gross output	EARL accounting translation; source-reported	USD millions, SAAR · 2005Q1–2026Q2	97	core statement
IIS-Q	intermediate_inputs	Cost of Goods Sold	Intermediate inputs	EARL accounting translation; source-reported	USD millions, SAAR · 2005Q1–2026Q2	97	core statement
IIS-Q	gross_value_added	Gross Profit	Varies by industry (see notes)	EARL accounting translation; source-reported	USD millions, SAAR · 2005Q1–2026Q2	97	core statement
IIS-Q	gross_margin	Gross Margin	—	EARL-calculated measure; EARL-calculated	ratio · 2005Q1–2026Q2	97	core statement

Note: One row represents one public `dataset_id` + `series_id` combination in the documented public data package. Concept / basis combines the manuscript's concept class and measurement basis fields; Unit / coverage combines the publication unit with the first and last published observation in the documented build. Coverage does not imply a balanced panel for every intervening period or every Industry entity. The documented public data package contains 128 dataset-series combinations: 97 source-reported and 31 EARL-calculated. Public roles comprise 93 core-statement, 33 supplemental-statement, and 2 public-data-only combinations.

Source: EARL methodology manuscript; Stage 4 `series-metadata.csv`.

A.2 Interpretation notes

- **Source-reported** describes the series measurement basis. An individual observation in such a series can still carry `hierarchy_rollup` or `derived_fallback` provenance when an explicitly sanctioned Stage 2 source-side construction was required.
- **EARL-calculated** identifies a Stage 3 semantic formula. The exact formulas are given in Appendix C.
- `IIS-A` and `IIS-Q` availability is entity-specific. A dataset-level first/last period therefore does not imply that all 99 Industry entities are observed throughout the full interval.
- Quarterly monetary series use USD millions at seasonally adjusted annual rates where specified in series metadata. Ratios are stored as decimals.

B Source Table and Line Crosswalk

Appendix B maps each **source-reported public dataset-series combination** to its governed official-source location. EARL has 97 source-reported dataset-series combinations, supported by 56 unique source-mapping rules in the source registry; repeated dataset-series entries can share the same underlying source rule. Calculated series are excluded here and documented in Appendix C.

B.1 Public source-reported series crosswalk

Table B.1 links every source-reported public dataset-series combination to its governed BEA source mapping.

Table B.1 — Public source-reported series crosswalk

Dataset	Series ID	EARL label	Official concept	Source table	Line / selector	Registered source description
NIS-A	gross_output	Gross Revenue	Gross output	GDP by Industry Table ID 15 Gross Output By Industry	1	All Industries
NIS-A	intermediate_inputs	Cost of Goods Sold	Intermediate inputs	GDP by Industry Table ID 20 Intermediate Inputs by Industry	1	All Industries
NIS-A	intermediate_inputs_energy	Energy Inputs	Energy inputs	GDP by Industry Table ID 25 Composition of Gross Output by Industry	7	Energy inputs
NIS-A	intermediate_inputs_materials	Materials Inputs	Materials inputs	GDP by Industry Table ID 25 Composition of Gross Output by Industry	8	Materials inputs
NIS-A	intermediate_inputs_services	Purchased Services Inputs	Purchased-services inputs	GDP by Industry Table ID 25 Composition of Gross Output by Industry	9	Purchased-services inputs
NIS-A	gross_value_added	Gross Profit	Gross domestic product	NIPA Table ID T10305 Table 1.3.5	1	Gross domestic product
NIS-A	wages_and_salaries	Wages and Salaries	Wages and salaries	NIPA Table ID T11000 Table 1.10	3	Wages and salaries
NIS-A	supplements_to_wages_and_salaries	Supplements to Wages and Salaries	Supplements to wages and salaries	NIPA Table ID T11000 Table 1.10	6	Supplements to wages and salaries
NIS-A	compensation_of_employees	Total Employee Compensation	Compensation of employees, paid	NIPA Table ID T11000 Table 1.10	2	Compensation of employees, paid
NIS-A	taxes_on_production_and_imports	Taxes on Production and Imports	Taxes on production and imports	NIPA Table ID T11000 Table 1.10	7	Taxes on production and imports
NIS-A	subsidies	Less: Subsidies	Less: Subsidies	NIPA Table ID T11000 Table 1.10	8	Less: Subsidies
NIS-A	ebitda_gos	EBITDA (Reported BEA Gross Operating Surplus)	Gross operating surplus	GDP by Industry Table ID 25 Composition of Gross Output by Industry	5	Gross operating surplus

Dataset	Series ID	EARL label	Official concept	Source table	Line / selector	Registered source description
NIS-A	consumption_of_fixed_capital	Depreciation	Consumption of fixed capital	NIPA Table ID T11000 Table 1.10	21	Consumption of fixed capital
NIS-A	net_operating_surplus	Operating Income (EBIT)	Net Operating Surplus	NIPA Table ID T11000 Table 1.10	9	Net Operating Surplus
NIS-A	statistical_discrepancy	Statistical Discrepancy	Less: Statistical Discrepancy	NIPA Table ID T10705 Table 1.7.5	15	Less: Statistical Discrepancy
NIS-A	interest_expense	Interest Expense	Net interest and miscellaneous payments, domestic industries	NIPA Table ID T11000 Table 1.10	11	Net interest and miscellaneous payments, domestic industries
NIS-A	business_current_transfer_payments_net	Business Current Transfer Payments	Business current transfer payments (net)	NIPA Table ID T11000 Table 1.10	12	Business current transfer payments (net)
NIS-A	proprietors_income_with_iva_ccadj	Proprietors' Income	Proprietors' income with inventory valuation and capital consumption adjustments	NIPA Table ID T11000 Table 1.10	13	Proprietors' income with inventory valuation and capital consumption adjustments
NIS-A	rental_income_with_ccadj	Rental Income	Rental income of persons with capital consumption adjustment	NIPA Table ID T11000 Table 1.10	14	Rental income of persons with capital consumption adjustment
NIS-A	government_enterprise_surplus	Government Enterprise Surplus	Current surplus of government enterprises	NIPA Table ID T11000 Table 1.10	20	Current surplus of government enterprises
NIS-A	corporate_profits_with_iva_ccadj	Corporate Profit	Corporate profits with inventory valuation and capital consumption adjustments, domestic industries	NIPA Table ID T11000 Table 1.10	15	Corporate profits with inventory valuation and capital consumption adjustments, domestic industries
NIS-Q	gross_output	Gross Revenue	Gross output	GDP by Industry Table ID 15 Gross Output By Industry	1	All Industries
NIS-Q	intermediate_inputs	Cost of Goods Sold	Intermediate inputs	GDP by Industry Table ID 20 Intermediate Inputs by Industry	1	All Industries
NIS-Q	gross_value_added	Gross Profit	Gross domestic product	NIPA Table ID T10305 Table 1.3.5	1	Gross domestic product
NIS-Q	wages_and_salaries	Wages and Salaries	Wages and salaries	NIPA Table ID T11000 Table 1.10	3	Wages and salaries
NIS-Q	supplements_to_wages_and_salaries	Supplements to Wages and Salaries	Supplements to wages and salaries	NIPA Table ID T11000 Table 1.10	6	Supplements to wages and salaries
NIS-Q	compensation_of_employees	Total Employee Compensation	Compensation of employees, paid	NIPA Table ID T11000 Table 1.10	2	Compensation of employees, paid

Dataset	Series ID	EARL label	Official concept	Source table	Line / selector	Registered source description
NIS-Q	taxes_on_production_and_imports	Taxes on Production and Imports	Taxes on production and imports	NIPA Table ID T11000 Table 1.10	7	Taxes on production and imports
NIS-Q	subsidies	Less: Subsidies	Less: Subsidies	NIPA Table ID T11000 Table 1.10	8	Less: Subsidies
NIS-Q	consumption_of_fixed_capital	Depreciation	Consumption of fixed capital	NIPA Table ID T11000 Table 1.10	21	Consumption of fixed capital
NIS-Q	net_operating_surplus	Operating Income (EBIT)	Net Operating Surplus	NIPA Table ID T11000 Table 1.10	9	Net Operating Surplus
NIS-Q	statistical_discrepancy	Statistical Discrepancy	Less: Statistical Discrepancy	NIPA Table ID T10705 Table 1.7.5	15	Less: Statistical Discrepancy
NIS-Q	interest_expense	Interest Expense	Net interest and miscellaneous payments, domestic industries	NIPA Table ID T11000 Table 1.10	11	Net interest and miscellaneous payments, domestic industries
NIS-Q	business_current_transfer_payments_net	Business Current Transfer Payments	Business current transfer payments (net)	NIPA Table ID T11000 Table 1.10	12	Business current transfer payments (net)
NIS-Q	proprietors_income_with_iva_ccadj	Proprietors' Income	Proprietors' income with inventory valuation and capital consumption adjustments	NIPA Table ID T11000 Table 1.10	13	Proprietors' income with inventory valuation and capital consumption adjustments
NIS-Q	rental_income_with_ccadj	Rental Income	Rental income of persons with capital consumption adjustment	NIPA Table ID T11000 Table 1.10	14	Rental income of persons with capital consumption adjustment
NIS-Q	government_enterprise_surplus	Government Enterprise Surplus	Current surplus of government enterprises	NIPA Table ID T11000 Table 1.10	20	Current surplus of government enterprises
NIS-Q	corporate_profits_with_iva_ccadj	Corporate Profit	Corporate profits with inventory valuation and capital consumption adjustments, domestic industries	NIPA Table ID T11000 Table 1.10	15	Corporate profits with inventory valuation and capital consumption adjustments, domestic industries
CIS-A	gross_value_added	Gross Profit	Gross value added of corporate business	NIPA Table ID T11400 Table 1.14	1	Gross value added of corporate business
CIS-A	wages_and_salaries	Wages and Salaries	Wages and salaries	NIPA Table ID T11400 Table 1.14	5	Wages and salaries
CIS-A	supplements_to_wages_and_salaries	Supplements to Wages and Salaries	Supplements to wages and salaries	NIPA Table ID T11400 Table 1.14	6	Supplements to wages and salaries
CIS-A	compensation_of_employees	Total Employee Compensation	Compensation of employees	NIPA Table ID T11400 Table 1.14	4	Compensation of employees
CIS-A	taxes_on_production_and_imports_less_subsidies	Taxes on Production & Imports, net of Subsidies	Taxes on production and imports less subsidies	NIPA Table ID T11400 Table 1.14	7	Taxes on production and imports less subsidies

Dataset	Series ID	EARL label	Official concept	Source table	Line / selector	Registered source description
CIS-A	consumption_of_fixed_capital	Depreciation	Consumption of fixed capital	NIPA Table ID T11400 Table 1.14	2	Consumption of fixed capital
CIS-A	net_operating_surplus	Operating Income (EBIT)	Net operating surplus	NIPA Table ID T11400 Table 1.14	8	Net operating surplus
CIS-A	interest_expense	Interest Expense / (Income)	Net interest and miscellaneous payments	NIPA Table ID T11400 Table 1.14	9	Net interest and miscellaneous payments
CIS-A	business_current_transfer_payments_net	Business Current Transfer Payments	Business current transfer payments (net)	NIPA Table ID T11400 Table 1.14	10	Business current transfer payments (net)
CIS-A	ebt_with_iva_ccadj	Income Before Taxes (EBT) — with IVA & CCAdj	Corporate profits with IVA and CCAdj	NIPA Table ID T11400 Table 1.14	11	Corporate profits with IVA and CCAdj
CIS-A	net_income_with_iva_ccadj	Net Income — with IVA & CCAdj	Profits after tax with IVA and CCAdj	NIPA Table ID T11400 Table 1.14	13	Profits after tax with IVA and CCAdj
CIS-A	retained_earnings_with_iva_ccadj	Retained Earnings	Undistributed profits with IVA and CCAdj	NIPA Table ID T11400 Table 1.14	15	Undistributed profits with IVA and CCAdj
CIS-A	inventory_valuation_adjustment	Inventory Valuation Adjustment (IVA)	Inventory valuation adjustment	NIPA Table ID T11400 Table 1.14	35	Inventory valuation adjustment
CIS-A	capital_consumption_adjustment	Capital Consumption Adjustment (CCAdj)	Capital consumption adjustment	NIPA Table ID T11400 Table 1.14	36	Capital consumption adjustment
CIS-A	ebt_without_iva_ccadj	Income Before Taxes (EBT) — without IVA & CCAdj	Profits before tax (without IVA and CCAdj)	NIPA Table ID T11400 Table 1.14	32	Profits before tax (without IVA and CCAdj)
CIS-A	taxes_on_corporate_income	Income Tax	Taxes on corporate income	NIPA Table ID T11400 Table 1.14	12	Taxes on corporate income
CIS-A	net_income_without_iva_ccadj	Net Income — without IVA & CCAdj	Profits after tax (without IVA and CCAdj)	NIPA Table ID T11400 Table 1.14	33	Profits after tax (without IVA and CCAdj)
CIS-A	net_dividends	Dividends	Net dividends	NIPA Table ID T11400 Table 1.14	14	Net dividends
CIS-A	retained_earnings_without_iva_ccadj	Retained Earnings — without IVA & CCAdj	Undistributed profits after tax (without IVA and CCAdj)	NIPA Table ID T11400 Table 1.14	34	Undistributed profits after tax (without IVA and CCAdj)
CIS-A	corporate_profits_receipts_from_rest_of_world	Receipts from Rest of World	Receipts from the rest of the world	NIPA Table ID T61600D Table 6.16	6	Receipts from the rest of the world
CIS-A	corporate_profits_payments_to_rest_of_world	Less: Payments to Rest of World	Less: Payments to the rest of the world	NIPA Table ID T61600D Table 6.16	7	Less: Payments to the rest of the world
CIS-Q	gross_value_added	Gross Profit	Gross value added of corporate business	NIPA Table ID T11400 Table 1.14	1	Gross value added of corporate business
CIS-Q	wages_and_salaries	Wages and Salaries	Wages and salaries	NIPA Table ID T11400 Table 1.14	5	Wages and salaries

Dataset	Series ID	EARL label	Official concept	Source table	Line / selector	Registered source description
CIS-Q	supplements_to_wages_and_salaries	Supplements to Wages and Salaries	Supplements to wages and salaries	NIPA Table ID T11400 Table 1.14	6	Supplements to wages and salaries
CIS-Q	compensation_of_employees	Total Employee Compensation	Compensation of employees	NIPA Table ID T11400 Table 1.14	4	Compensation of employees
CIS-Q	taxes_on_production_and_imports_less_subsidies	Taxes on Production & Imports, net of Subsidies	Taxes on production and imports less subsidies	NIPA Table ID T11400 Table 1.14	7	Taxes on production and imports less subsidies
CIS-Q	consumption_of_fixed_capital	Depreciation	Consumption of fixed capital	NIPA Table ID T11400 Table 1.14	2	Consumption of fixed capital
CIS-Q	net_operating_surplus	Operating Income (EBIT)	Net operating surplus	NIPA Table ID T11400 Table 1.14	8	Net operating surplus
CIS-Q	interest_expense	Interest Expense / (Income)	Net interest and miscellaneous payments	NIPA Table ID T11400 Table 1.14	9	Net interest and miscellaneous payments
CIS-Q	business_current_transfer_payments_net	Business Current Transfer Payments	Business current transfer payments (net)	NIPA Table ID T11400 Table 1.14	10	Business current transfer payments (net)
CIS-Q	ebt_with_iva_ccadj	Income Before Taxes (EBT) — with IVA & CCAdj	Corporate profits with IVA and CCAdj	NIPA Table ID T11400 Table 1.14	11	Corporate profits with IVA and CCAdj
CIS-Q	net_income_with_iva_ccadj	Net Income — with IVA & CCAdj	Profits after tax with IVA and CCAdj	NIPA Table ID T11400 Table 1.14	13	Profits after tax with IVA and CCAdj
CIS-Q	retained_earnings_with_iva_ccadj	Retained Earnings	Undistributed profits with IVA and CCAdj	NIPA Table ID T11400 Table 1.14	15	Undistributed profits with IVA and CCAdj
CIS-Q	inventory_valuation_adjustment	Inventory Valuation Adjustment (IVA)	Inventory valuation adjustment	NIPA Table ID T11400 Table 1.14	35	Inventory valuation adjustment
CIS-Q	capital_consumption_adjustment	Capital Consumption Adjustment (CCAdj)	Capital consumption adjustment	NIPA Table ID T11400 Table 1.14	36	Capital consumption adjustment
CIS-Q	ebt_without_iva_ccadj	Income Before Taxes (EBT) — without IVA & CCAdj	Profits before tax (without IVA and CCAdj)	NIPA Table ID T11400 Table 1.14	32	Profits before tax (without IVA and CCAdj)
CIS-Q	taxes_on_corporate_income	Income Tax	Taxes on corporate income	NIPA Table ID T11400 Table 1.14	12	Taxes on corporate income
CIS-Q	net_income_without_iva_ccadj	Net Income — without IVA & CCAdj	Profits after tax (without IVA and CCAdj)	NIPA Table ID T11400 Table 1.14	33	Profits after tax (without IVA and CCAdj)
CIS-Q	net_dividends	Dividends	Net dividends	NIPA Table ID T11400 Table 1.14	14	Net dividends
CIS-Q	retained_earnings_without_iva_ccadj	Retained Earnings — without IVA & CCAdj	Undistributed profits after tax (without IVA and CCAdj)	NIPA Table ID T11400 Table 1.14	34	Undistributed profits after tax (without IVA and CCAdj)
CIS-Q	corporate_profits_receipts_from_rest_of_world	Receipts from Rest of World	Receipts from the rest of the world	NIPA Table ID T61600D Table 6.16	6	Receipts from the rest of the world

Dataset	Series ID	EARL label	Official concept	Source table	Line / selector	Registered source description
CIS-Q	corporate_profits_payments_to_rest_of_world	Less: Payments to Rest of World	Less: Payments to the rest of the world	NIPA Table ID T61600D Table 6.16	7	Less: Payments to the rest of the world
IIS-A	gross_output	Gross Revenue	Gross output	GDP by Industry Table ID 15 Gross Output By Industry	Industry-specific	Varies by industry (see notes)
IIS-A	intermediate_inputs	Cost of Goods Sold	Intermediate inputs	GDP by Industry Table ID 20 Intermediate Inputs by Industry	Industry-specific	Varies by industry (see notes)
IIS-A	intermediate_inputs_energy	Energy Inputs	Energy inputs	GDP by Industry Table ID 25 Composition of Gross Output by Industry	Industry-specific	Energy inputs
IIS-A	intermediate_inputs_materials	Materials Inputs	Materials inputs	GDP by Industry Table ID 25 Composition of Gross Output by Industry	Industry-specific	Materials inputs
IIS-A	intermediate_inputs_services	Purchased Services Inputs	Purchased-services inputs	GDP by Industry Table ID 25 Composition of Gross Output by Industry	Industry-specific	Purchased-services inputs
IIS-A	gross_value_added	Gross Profit	Varies by industry (see notes)	GDP by Industry Table ID 1 Value Added by Industry	Industry-specific	Varies by industry (see notes)
IIS-A	compensation_of_employees	Employee Compensation	Compensation of employees	GDP by Industry Table ID 25 Composition of Gross Output by Industry	Industry-specific	Compensation of employees
IIS-A	taxes_on_production_and_imports_less_subsidies	Taxes on Production & Imports, net of Subsidies	Taxes on production and imports less subsidies	GDP by Industry Table ID 25 Composition of Gross Output by Industry	Industry-specific	Taxes on production and imports less subsidies
IIS-A	ebitda_gos	EBITDA (Reported BEA Gross Operating Surplus)	Gross operating surplus	GDP by Industry Table ID 25 Composition of Gross Output by Industry	Industry-specific	Gross operating surplus
IIS-A	wages_and_salaries	Wages and Salaries	Wages and salaries / wage and salary accruals by industry	NIPA Table ID T60300D Table 6.3D	—	—
IIS-A	fte_employees	Full-Time Equivalent Employees	Full-time equivalent employees by industry	NIPA Table ID T60500D Table 6.5D	—	—
IIS-A	wage_and_salary_accruals_per_fte	Wage and Salary Accruals per FTE	Wage and salary accruals per full-time equivalent employee by industry	NIPA Table ID T60600D Table 6.6D	—	—
IIS-A	interest_expense	Interest Expense / (Income)	Net interest / net interest and miscellaneous payments	NIPA Table ID T61500D Table 6.15D	—	—

Dataset	Series ID	EARL label	Official concept	Source table	Line / selector	Registered source description
IIS-A	proprietors_income_with_iva_ccadj	Proprietors' Income	Nonfarm proprietors' income by industry	NIPA Table ID T61200D Table 6.12D	—	—
IIS-Q	gross_output	Gross Revenue	Gross output	GDP by Industry Table ID 15 Gross Output By Industry	Industry-specific	Varies by industry (see notes)
IIS-Q	intermediate_inputs	Cost of Goods Sold	Intermediate inputs	GDP by Industry Table ID 20 Intermediate Inputs by Industry	Industry-specific	Varies by industry (see notes)
IIS-Q	gross_value_added	Gross Profit	Varies by industry (see notes)	GDP by Industry Table ID 1 Value Added by Industry	Industry-specific	Varies by industry (see notes)

Note: One row represents one source-reported public `dataset_id` + `series_id` combination. Source table combines the source system, governed table ID, and human-readable table label. The label Industry-specific indicates an Industry mapping whose selector depends on governed Industry scope; an em dash for selected annual Industry NIPA sources indicates description-driven entity matching rather than one fixed line. The 97 public rows are supported by 56 unique source-mapping rules because annual and quarterly dataset-series combinations can share one governed A/Q rule. Calculated series are excluded and documented in Appendix C.

Source: EARL methodology manuscript; canonical Stage 4 `series-metadata.csv` source registry `source_mapping.csv`.

B.2 Matching rules

For fixed-line NIPA sources, Stage 2 requires agreement on **table + line + normalized line description**. A familiar line number alone is not sufficient. GDP by Industry mappings require the governed table, Industry code/scope, and the registered concept selector where a table contains multiple concepts.

Annual Industry NIPA sources are matched differently because the source is entity-indexed. EARL orders the source by line, normalizes the source `LineDescription` and governed entity description, pairs matching descriptions, and uses governed occurrence order where repeated normalized descriptions require disambiguation. Zero entity mappings are a structural failure; partial entity coverage is allowed where it reflects the official source.

B.3 Source families used by the methodology

- **National:** GDP by Industry Tables 15, 20, and 25; NIPA Tables 1.3.5 (`T10305`), 1.7.5 (`T10705`), 1.10 (`T11000`), and the registered supplemental lines.
- **Corporate:** principally NIPA Table 1.14 (`T11400`), with rest-of-world receipts/payments from Table 6.16D (`T61600D`).
- **Industry:** GDP by Industry Tables 1, 15, 20, and 25, plus selected Industry NIPA tables `T60300D`, `T60500D`, `T60600D`, `T61200D`, and `T61500D`.

Historical annual Industry evidence is a separate source layer and is documented in Appendix F.

C Formula Reference

Appendix C records the Stage 3 formulas for all EARL-calculated dataset-series combinations. The public data specification contains **31 calculated dataset-series combinations**, representing **13 stable calculated series IDs** and **14 distinct formula definitions** across scopes and frequencies.

C.1 Formula language

The semantic evaluator supports a deliberately constrained expression language: registered series identifiers, numerical constants, supported arithmetic (`+`, `-`, `*`, guarded `/`), comparisons, and restricted conditional `IF`. Unknown inputs, unresolved dependencies, circular dependencies, unsupported functions, and duplicate definitions fail closed. Division by zero or a missing denominator returns missing rather than infinity or `NaN`.

C.2 Public calculated formulas

Table C.1 lists the public calculated measures and their reader-facing formulas.

Table C.1 — Public calculated formulas

Dataset	Calculated measure	Formula	Unit	Public role
NIS-A	Total Operating Expenses	Total Employee Compensation + Taxes on Production and Imports − Subsidies	USD millions	Core statement
NIS-A	Income Before Taxes (EBT)	Operating Income (EBIT) − Interest Expense	USD millions	Core statement
NIS-A	Gross Margin	Gross Profit ÷ Gross Revenue	Ratio	Core statement
NIS-A	EBITDA Margin (Reported-GOS Basis)	EBITDA (Reported BEA Gross Operating Surplus) ÷ Gross Revenue	Ratio	Core statement
NIS-A	Operating Margin (EBIT)	Operating Income (EBIT) ÷ Gross Revenue	Ratio	Core statement
NIS-A	EBT Margin	Income Before Taxes (EBT) ÷ Gross Revenue	Ratio	Core statement
NIS-A	EBITDA (Calculated; reconciliation only)	Gross Profit − Total Operating Expenses	USD millions	Supplemental statement
NIS-A	EBIT — Output Perspective	EBITDA (Reported BEA Gross Operating Surplus) − Depreciation	USD millions	Supplemental statement
NIS-A	Taxes on Production & Imports, net of Subsidies	Taxes on Production and Imports − Subsidies	USD millions	Public data only
NIS-Q	Total Operating Expenses	Total Employee Compensation + Taxes on Production and Imports − Subsidies	USD millions, SAAR	Core statement
NIS-Q	EBITDA (Calculated)	Gross Profit − Total Operating Expenses	USD millions, SAAR	Core statement
NIS-Q	Income Before Taxes (EBT)	Operating Income (EBIT) − Interest Expense	USD millions, SAAR	Core statement
NIS-Q	Gross Margin	Gross Profit ÷ Gross Revenue	Ratio	Core statement
NIS-Q	EBITDA Margin (Calculated-EBITDA Basis)	EBITDA (Calculated) ÷ Gross Revenue	Ratio	Core statement
NIS-Q	Operating Margin (EBIT)	Operating Income (EBIT) ÷ Gross Revenue	Ratio	Core statement
NIS-Q	EBT Margin	Income Before Taxes (EBT) ÷ Gross Revenue	Ratio	Core statement
NIS-Q	Taxes on Production & Imports, net of Subsidies	Taxes on Production and Imports − Subsidies	USD millions, SAAR	Public data only
CIS-A	Total Operating Expenses	Employee Compensation + Taxes on Production & Imports, net of Subsidies	USD millions	Core statement
CIS-A	EBITDA (Calculated)	Gross Profit − Total Operating Expenses	USD millions	Core statement
CIS-A	Effective Tax Rate	If EBT with IVA & CCAdj > 0: Income Tax ÷ EBT with IVA & CCAdj; otherwise missing	Ratio	Core statement
CIS-A	Payout Ratio	Dividends ÷ Net Income — with IVA & CCAdj	Ratio	Core statement
CIS-A	Retention Ratio	1 − Payout Ratio	Ratio	Core statement
CIS-Q	Total Operating Expenses	Employee Compensation + Taxes on Production & Imports, net of Subsidies	USD millions, SAAR	Core statement

Dataset	Calculated measure	Formula	Unit	Public role
CIS-Q	EBITDA (Calculated)	Gross Profit − Total Operating Expenses	USD millions, SAAR	Core statement
CIS-Q	Effective Tax Rate	If EBT with IVA & CCAdj > 0: Income Tax ÷ EBT with IVA & CCAdj; otherwise missing	Ratio	Core statement
CIS-Q	Payout Ratio	Dividends ÷ Net Income — with IVA & CCAdj	Ratio	Core statement
CIS-Q	Retention Ratio	1 − Payout Ratio	Ratio	Core statement
IIS-A	Total Operating Expenses	Employee Compensation + Taxes on Production & Imports, net of Subsidies	USD millions	Core statement
IIS-A	Gross Margin	Gross Profit ÷ Gross Revenue	Ratio	Core statement
IIS-A	EBITDA Margin (Reported-GOS Basis)	EBITDA (Reported BEA Gross Operating Surplus) ÷ Gross Revenue	Ratio	Core statement
IIS-Q	Gross Margin	Gross Profit ÷ Gross Revenue	Ratio	Core statement

Note: Formulas are shown using governed EARL public labels for readability. The retained authority files preserve the corresponding stable series identifiers, explicit input dependencies, and exact machine-evaluable Stage 3 expressions; Appendix A provides the complete public series dictionary. Division is guarded: a zero or unavailable denominator returns missing. Effective Tax Rate is additionally defined only when EBT with IVA & CCAdj is positive. The 31 rows represent 13 stable calculated series IDs and 14 distinct registry formulas. Industry hierarchy rollups and the All Industries GVA fallback (entity ID `All`) are Stage 2 source-side constructions and are intentionally excluded.

Source: EARL revised methodology manuscript; Stage 4 `series-metadata.csv` source registry `source_mapping.csv`.

C.3 Source-side constructions and semantic formulas

Two specified source-side mechanisms are intentionally excluded from the formula inventory above:

1. **Industry hierarchy rollups** for specified parent/series combinations, requiring complete child coverage and yielding to a direct reported parent when available.
2. **All Industries GVA fallback** (`entity_id = All`): where direct aggregate GVA is missing but aggregate Gross Output and Intermediate Inputs are available, `GVA_All = GO_All - II_All` fills only the missing aggregate cell and carries `source_type = derived_fallback`.

These mechanisms preserve the measurement basis of the underlying source concept and are distinguishable from `source_type = derived` Stage 3 calculations.

C.4 Formula validation

Stage 4 independently evaluates the formulas recorded in the public metadata using only public input observations. In the documented build, all **26,866** published calculated observations reproduce their formulas; the largest independently measured floating-point difference is approximately **3.55 × 10^-15**, consistent with ordinary ratio arithmetic. A historical machine alias associated with annual reported GOS does not make reported GOS a calculated series.

D Statement Layouts

The `statement-layout.csv` contains **132 display rows covering 126 unique dataset-series combinations**. Six extra display rows are intentional repeated uses of the same series in different statement contexts. Two public-data-only National series—annual and quarterly `taxes_less_subsidies`—have no statement row. Presentation signs affect display only; stored economic values are unchanged.

D.1 Complete layout contract

Table D.1 records the complete published statement layout.

Table D.1 — Complete statement layout contract

Statement	Order	Series ID	Indent	Row type	Sign	Availability	Notes
NIS-A — Section 10: Income Flow							
Canonical section ID: income_flow · Level: 0							
NIS-A	10	gross_output	0	Line	As stored	Always	—
NIS-A	20	intermediate_inputs	0	Line	Invert sign	Always	—
NIS-A	30	intermediate_inputs_energy	1	Detail	Invert sign	Always	Annual COGS detail; render N/A when unavailable for the displayed period.
NIS-A	40	intermediate_inputs_materials	1	Detail	Invert sign	Always	Annual COGS detail; render N/A when unavailable for the displayed period.
NIS-A	50	intermediate_inputs_services	1	Detail	Invert sign	Always	Annual COGS detail; render N/A when unavailable for the displayed period.
NIS-A	60	gross_value_added	0	Subtotal	As stored	Always	—
NIS-A — Section 20: Operating Expenses							
Canonical section ID: operating_expenses · Level: 0							
NIS-A	70	wages_and_salaries	1	Detail	Invert sign	Always	—
NIS-A	80	supplements_to_wages_and_salaries	1	Detail	Invert sign	Always	—
NIS-A	90	compensation_of_employees	0	Subtotal	Invert sign	Always	—
NIS-A	100	taxes_on_production_and_imports	1	Detail	Invert sign	Always	—
NIS-A	110	subsidies	1	Detail	As stored	Always	Display label should make subtraction explicit (e.g., Less: Subsidies); stored value remains in its source sign.
NIS-A	120	operating_expenses	0	Total	Invert sign	Always	—
NIS-A — Section 30: Operating Profit							
Canonical section ID: operating_profit · Level: 0							
NIS-A	130	ebitda_gos	0	Subtotal	As stored	Always	Public label must explicitly identify reported-source basis: EBITDA (Reported BEA Gross Operating Surplus).
NIS-A	140	consumption_of_fixed_capital	1	Detail	Invert sign	Always	—
NIS-A	150	net_operating_surplus	0	Subtotal	As stored	Always	—
NIS-A	160	interest_expense	1	Detail	Invert sign	Always	—
NIS-A	170	earnings_before_tax	0	Total	As stored	Always	—
NIS-A — Section 40: Margins (%)							
Canonical section ID: margins · Level: 0							
NIS-A	180	gross_margin	0	Ratio	As stored	Always	—
NIS-A	190	ebitda_margin_gos	0	Ratio	As stored	Always	Margin uses reported EBITDA/GOS, not the calculated reconciliation series.
NIS-A	200	ebit_margin	0	Ratio	As stored	Always	—
NIS-A	210	ebt_margin	0	Ratio	As stored	Always	—

Statement	Order	Series ID	Indent	Row type	Sign	Availability	Notes
NIS-A — Section 50: NIPA Reconciliation & Composition (Supplemental)							
Canonical section ID: supplemental_reconciliation · Level: 0							
NIS-A	220	ebitda_calculated	0	Reconciliation	As stored	Always	Reconciliation-only annual calculated EBITDA. Keep distinct from reported BEA Gross Operating Surplus.
NIS-A	230	net_operating_surplus_output_perspective	0	Reconciliation	As stored	Always	Alternative/output-perspective EBIT reconciliation series.
NIS-A	240	statistical_discrepancy	0	Reconciliation	As stored	Always	Preserve source sign.
NIS-A — Section 60: NIPA EBIT Composition							
Canonical section ID: supplemental_ebit_composition · Level: 1							
NIS-A	250	interest_expense	1	Detail	As stored	Always	Repeated intentionally; additive component in NIPA EBIT composition, so display source sign here.
NIS-A	260	business_current_transfer_payments_net	1	Detail	As stored	Always	—
NIS-A	270	proprietors_income_with_iva_ccadj	1	Detail	As stored	Always	—
NIS-A	280	rental_income_with_ccadj	1	Detail	As stored	Always	—
NIS-A	290	government_enterprise_surplus	1	Detail	As stored	Always	Preserve source sign.
NIS-A	300	corporate_profits_with_iva_ccadj	1	Detail	As stored	Always	—
NIS-Q — Section 10: Income Flow							
Canonical section ID: income_flow · Level: 0							
NIS-Q	10	gross_output	0	Line	As stored	Always	—
NIS-Q	20	intermediate_inputs	0	Line	Invert sign	Always	—
NIS-Q	30	gross_value_added	0	Subtotal	As stored	Always	—
NIS-Q — Section 20: Operating Expenses							
Canonical section ID: operating_expenses · Level: 0							
NIS-Q	40	wages_and_salaries	1	Detail	Invert sign	Always	—
NIS-Q	50	supplements_to_wages_and_salaries	1	Detail	Invert sign	Always	—
NIS-Q	60	compensation_of_employees	0	Subtotal	Invert sign	Always	—
NIS-Q	70	taxes_on_production_and_imports	1	Detail	Invert sign	Always	—
NIS-Q	80	subsidies	1	Detail	As stored	Always	Display label should make subtraction explicit (e.g., Less: Subsidies); stored value remains in its source sign.
NIS-Q	90	operating_expenses	0	Total	Invert sign	Always	—
NIS-Q — Section 30: Operating Profit							
Canonical section ID: operating_profit · Level: 0							
NIS-Q	100	ebitda_calculated	0	Subtotal	As stored	Always	Primary quarterly EBITDA is an explicitly EARL-calculated measure: Gross Profit minus Total Operating Expenses. Reported BEA Gross Operating Surplus remains a separate annual/source-supported series and is not overwritten.
NIS-Q	110	consumption_of_fixed_capital	1	Detail	Invert sign	Always	—

Statement	Order	Series ID	Indent	Row type	Sign	Availability	Notes
NIS-Q	120	net_operating_surplus	0	Subtotal	As stored	Always	—
NIS-Q	130	interest_expense	1	Detail	Invert sign	Always	—
NIS-Q	140	earnings_before_tax	0	Total	As stored	Always	—
NIS-Q — Section 40: Margins (%)							
Canonical section ID: margins · Level: 0							
NIS-Q	150	gross_margin	0	Ratio	As stored	Always	—
NIS-Q	160	ebitda_margin_calculated	0	Ratio	As stored	Always	Quarterly EBITDA Margin uses EBITDA (Calculated) / Gross Revenue. Keep distinct from the reported-GOS-based EBITDA Margin used where reported GOS exists.
NIS-Q	170	ebit_margin	0	Ratio	As stored	Always	—
NIS-Q	180	ebt_margin	0	Ratio	As stored	Always	—
NIS-Q — Section 50: NIPA Reconciliation & Composition (Supplemental)							
Canonical section ID: supplemental_reconciliation · Level: 0							
NIS-Q	200	statistical_discrepancy	0	Reconciliation	As stored	Always	Preserve source sign.
NIS-Q — Section 60: NIPA EBIT Composition							
Canonical section ID: supplemental_ebit_composition · Level: 1							
NIS-Q	210	interest_expense	1	Detail	As stored	Always	Repeated intentionally; additive component in NIPA EBIT composition, so display source sign here.
NIS-Q	220	business_current_transfer_payments_net	1	Detail	As stored	Always	—
NIS-Q	230	proprietors_income_with_iva_ccadj	1	Detail	As stored	Always	—
NIS-Q	240	rental_income_with_ccadj	1	Detail	As stored	Always	—
NIS-Q	250	government_enterprise_surplus	1	Detail	As stored	Always	Preserve source sign.
NIS-Q	260	corporate_profits_with_iva_ccadj	1	Detail	As stored	Always	—
CIS-A — Section 10: Income Flow							
Canonical section ID: income_flow · Level: 0							
CIS-A	10	gross_value_added	0	Subtotal	As stored	Always	—
CIS-A — Section 20: Operating Expenses							
Canonical section ID: operating_expenses · Level: 0							
CIS-A	20	wages_and_salaries	1	Detail	Invert sign	Always	—
CIS-A	30	supplements_to_wages_and_salaries	1	Detail	Invert sign	Always	—
CIS-A	40	compensation_of_employees	0	Subtotal	Invert sign	Always	—
CIS-A	50	taxes_on_production_and_imports_less_subsidies	1	Detail	Invert sign	Always	—
CIS-A	60	operating_expenses	0	Total	Invert sign	Always	—
CIS-A — Section 30: Operating Profit							
Canonical section ID: operating_profit · Level: 0							
CIS-A	70	ebitda_calculated	0	Subtotal	As stored	Always	EARL-calculated subtotal. Public label must read EBITDA (Calculated); keep distinct from reported BEA GOS-based EBITDA in National/Industry.
CIS-A	80	consumption_of_fixed_capital	1	Detail	Invert sign	Always	—

Statement	Order	Series ID	Indent	Row type	Sign	Availability	Notes
CIS-A	90	net_operating_surplus	0	Subtotal	As stored	Always	—
CIS-A — Section 40: Pre-Tax Income							
Canonical section ID: pre_tax_income · Level: 0							
CIS-A	100	interest_expense	1	Detail	Invert sign	Always	Public label should allow sign changes: Interest Expense / (Income).
CIS-A	110	business_current_transfer_payments_net	1	Detail	Invert sign	Always	—
CIS-A	120	ebt_with_iva_ccadj	0	Total	As stored	Always	—
CIS-A — Section 50: Net Income							
Canonical section ID: net_income · Level: 0							
CIS-A	130	taxes_on_corporate_income	1	Detail	Invert sign	Always	—
CIS-A	140	net_income_with_iva_ccadj	0	Total	As stored	Always	—
CIS-A — Section 60: Earnings Distribution							
Canonical section ID: earnings_distribution · Level: 0							
CIS-A	150	net_dividends	1	Detail	Invert sign	Always	—
CIS-A	160	retained_earnings_with_iva_ccadj	0	Total	As stored	Always	—
CIS-A — Section 70: Ratios (%)							
Canonical section ID: ratios · Level: 0							
CIS-A	170	effective_tax_rate	0	Ratio	As stored	Always	—
CIS-A	180	payout_ratio	0	Ratio	As stored	Always	—
CIS-A	190	retention_ratio	0	Ratio	As stored	Always	—
CIS-A — Section 80: Alternative Profit Measure — Without IVA & CCAdj (Supplemental)							
Canonical section ID: alternative_profit · Level: 0							
CIS-A	200	inventory_valuation_adjustment	1	Reconciliation	As stored	Always	Preserve source sign.
CIS-A	210	capital_consumption_adjustment	1	Reconciliation	As stored	Always	Preserve source sign.
CIS-A	220	ebt_without_iva_ccadj	0	Total	As stored	Always	—
CIS-A	230	taxes_on_corporate_income	1	Detail	Invert sign	Always	Repeated intentionally to make the alternative EBT-to-Net-Income reconciliation explicit.
CIS-A	240	net_income_without_iva_ccadj	0	Total	As stored	Always	—
CIS-A	250	net_dividends	1	Detail	Invert sign	Always	Repeated intentionally to make the alternative Net-Income-to-Retained-Earnings reconciliation explicit.
CIS-A	260	retained_earnings_without_iva_ccadj	0	Total	As stored	Always	—
CIS-A — Section 90: Rest-of-World Profit Flows (Supplemental)							
Canonical section ID: rest_of_world · Level: 0							
CIS-A	270	corporate_profits_receipts_from_rest_of_world	1	Detail	As stored	Always	—
CIS-A	280	corporate_profits_payments_to_rest_of_world	1	Detail	Invert sign	Always	Display label should make subtraction explicit (e.g., Less: Payments to Rest of World).
CIS-Q — Section 10: Income Flow							
Canonical section ID: income_flow · Level: 0							
CIS-Q	10	gross_value_added	0	Subtotal	As stored	Always	—

Statement	Order	Series ID	Indent	Row type	Sign	Availability	Notes
CIS-Q — Section 20: Operating Expenses							
Canonical section ID: operating_expenses · Level: 0							
CIS-Q	20	wages_and_salaries	1	Detail	Invert sign	Always	—
CIS-Q	30	supplements_to_wages_and_salaries	1	Detail	Invert sign	Always	—
CIS-Q	40	compensation_of_employees	0	Subtotal	Invert sign	Always	—
CIS-Q	50	taxes_on_production_and_imports_less_subsidies	1	Detail	Invert sign	Always	—
CIS-Q	60	operating_expenses	0	Total	Invert sign	Always	—
CIS-Q — Section 30: Operating Profit							
Canonical section ID: operating_profit · Level: 0							
CIS-Q	70	ebitda_calculated	0	Subtotal	As stored	Always	EARL-calculated subtotal. Public label must read EBITDA (Calculated); keep distinct from reported BEA GOS-based EBITDA in National/Industry.
CIS-Q	80	consumption_of_fixed_capital	1	Detail	Invert sign	Always	—
CIS-Q	90	net_operating_surplus	0	Subtotal	As stored	Always	—
CIS-Q — Section 40: Pre-Tax Income							
Canonical section ID: pre_tax_income · Level: 0							
CIS-Q	100	interest_expense	1	Detail	Invert sign	Always	Public label should allow sign changes: Interest Expense / (Income).
CIS-Q	110	business_current_transfer_payments_net	1	Detail	Invert sign	Always	—
CIS-Q	120	ebt_with_iva_ccadj	0	Total	As stored	Always	—
CIS-Q — Section 50: Net Income							
Canonical section ID: net_income · Level: 0							
CIS-Q	130	taxes_on_corporate_income	1	Detail	Invert sign	Always	—
CIS-Q	140	net_income_with_iva_ccadj	0	Total	As stored	Always	—
CIS-Q — Section 60: Earnings Distribution							
Canonical section ID: earnings_distribution · Level: 0							
CIS-Q	150	net_dividends	1	Detail	Invert sign	Always	—
CIS-Q	160	retained_earnings_with_iva_ccadj	0	Total	As stored	Always	—
CIS-Q — Section 70: Ratios (%)							
Canonical section ID: ratios · Level: 0							
CIS-Q	170	effective_tax_rate	0	Ratio	As stored	Always	—
CIS-Q	180	payout_ratio	0	Ratio	As stored	Always	—
CIS-Q	190	retention_ratio	0	Ratio	As stored	Always	—
CIS-Q — Section 80: Alternative Profit Measure — Without IVA & CCAdj (Supplemental)							
Canonical section ID: alternative_profit · Level: 0							
CIS-Q	200	inventory_valuation_adjustment	1	Reconciliation	As stored	Always	Preserve source sign.
CIS-Q	210	capital_consumption_adjustment	1	Reconciliation	As stored	Always	Preserve source sign.
CIS-Q	220	ebt_without_iva_ccadj	0	Total	As stored	Always	—

Statement	Order	Series ID	Indent	Row type	Sign	Availability	Notes
CIS-Q	230	taxes_on_corporate_income	1	Detail	Invert sign	Always	Repeated intentionally to make the alternative EBT-to-Net-Income reconciliation explicit.
CIS-Q	240	net_income_without_iva_ccadj	0	Total	As stored	Always	—
CIS-Q	250	net_dividends	1	Detail	Invert sign	Always	Repeated intentionally to make the alternative Net-Income-to-Retained-Earnings reconciliation explicit.
CIS-Q	260	retained_earnings_without_iva_ccadj	0	Total	As stored	Always	—
CIS-Q — Section 90: Rest-of-World Profit Flows (Supplemental)							
Canonical section ID: rest_of_world · Level: 0							
CIS-Q	270	corporate_profits_receipts_from_rest_of_world	1	Detail	As stored	Always	—
CIS-Q	280	corporate_profits_payments_to_rest_of_world	1	Detail	Invert sign	Always	Display label should make subtraction explicit (e.g., Less: Payments to Rest of World).
IIS-A — Section 10: Income Flow							
Canonical section ID: income_flow · Level: 0							
IIS-A	10	gross_output	0	Line	As stored	Always	—
IIS-A	20	intermediate_inputs	0	Line	Invert sign	Always	—
IIS-A	30	intermediate_inputs_energy	1	Detail	Invert sign	Always	Annual COGS detail; may lag headline Industry series.
IIS-A	40	intermediate_inputs_materials	1	Detail	Invert sign	Always	Annual COGS detail; may lag headline Industry series.
IIS-A	50	intermediate_inputs_services	1	Detail	Invert sign	Always	Annual COGS detail; may lag headline Industry series.
IIS-A	60	gross_value_added	0	Subtotal	As stored	Always	—
IIS-A — Section 20: Operating Expenses							
Canonical section ID: operating_expenses · Level: 0							
IIS-A	70	compensation_of_employees	1	Detail	Invert sign	Always	—
IIS-A	80	taxes_on_production_and_imports_less_subsidies	1	Detail	Invert sign	Always	—
IIS-A	90	operating_expenses	0	Total	Invert sign	Always	—
IIS-A — Section 30: Operating Profit							
Canonical section ID: operating_profit · Level: 0							
IIS-A	100	ebitda_gos	0	Total	As stored	Always	Public label must explicitly identify reported-source basis: EBITDA (Reported BEA Gross Operating Surplus). Industry statement ends its primary accounting chain here.
IIS-A — Section 40: Margins (%)							
Canonical section ID: margins · Level: 0							
IIS-A	110	gross_margin	0	Ratio	As stored	Always	—
IIS-A	120	ebitda_margin_gos	0	Ratio	As stored	Always	Uses reported BEA Gross Operating Surplus as the EBITDA numerator.
IIS-A — Section 50: Labor & Operating Detail (Supplemental)							
Canonical section ID: labor_detail · Level: 0							
IIS-A	130	wages_and_salaries	1	Detail	As stored	Always	—
IIS-A	140	fte_employees	1	Detail	As stored	Always	—
IIS-A	150	wage_and_salary_accruals_per_fte	1	Detail	As stored	Always	—

Statement	Order	Series ID	Indent	Row type	Sign	Availability	Notes
IIS-A — Section 60: Financing & Income Detail (Supplemental)							
Canonical section ID: financing_income_detail · Level: 0							
IIS-A	160	interest_expense	1	Detail	Invert sign	When available	Sparse Industry coverage; public label should allow sign changes: Interest Expense / (Income). No EBITDA-to-EBT bridge is implied.
IIS-A	170	proprietors_income_with_iva_ccadj	1	Detail	As stored	When available	Sparse Industry coverage; supplemental information only, not part of the primary EBITDA chain.
IIS-Q — Section 10: Income Flow							
Canonical section ID: income_flow · Level: 0							
IIS-Q	10	gross_output	0	Line	As stored	Always	—
IIS-Q	20	intermediate_inputs	0	Line	Invert sign	Always	—
IIS-Q	30	gross_value_added	0	Subtotal	As stored	Always	—
IIS-Q — Section 20: Margins (%)							
Canonical section ID: margins · Level: 0							
IIS-Q	40	gross_margin	0	Ratio	As stored	Always	Quarterly Industry statement currently ends at production-side Gross Profit/ Gross Margin; the layout does not imply unsupported quarterly OpEx/EBITDA structure.

Note: The published layout contains 132 display rows covering 126 unique statement-family + series combinations. Section bands identify the statement family, section order and label, section ID, and section level. Order is the defined display position of a row within its statement. Indent is a presentation indentation flag: 0 denotes a base row and 1 a subordinate/detail row. Row type describes the presentation role (line, detail, subtotal, total, ratio, or reconciliation). Sign controls presentation only; “Invert sign” does not alter the stored economic value. Availability states whether a row is always part of the layout or is shown only when data are available. Six extra rows are intentional presentation repeats rather than duplicate observations. The retained `statement-layout.csv` preserves the exact machine-readable layout. Appendix A provides the complete series dictionary. The two National public-data-only net-production-tax series have no statement row.

Source: EARL revised methodology manuscript; Stage 4 `statement-layout.csv` and `series-metadata.csv`.

D.2 Intentional repeated rows

The six additional layout rows arise from intentional reuse of a series: National Annual and National Quarterly each display interest in more than one analytical context; Corporate Annual and Corporate Quarterly each repeat Income Tax and Dividends in their respective flow/reconciliation structures. These are presentation repetitions, not duplicate economic observations.

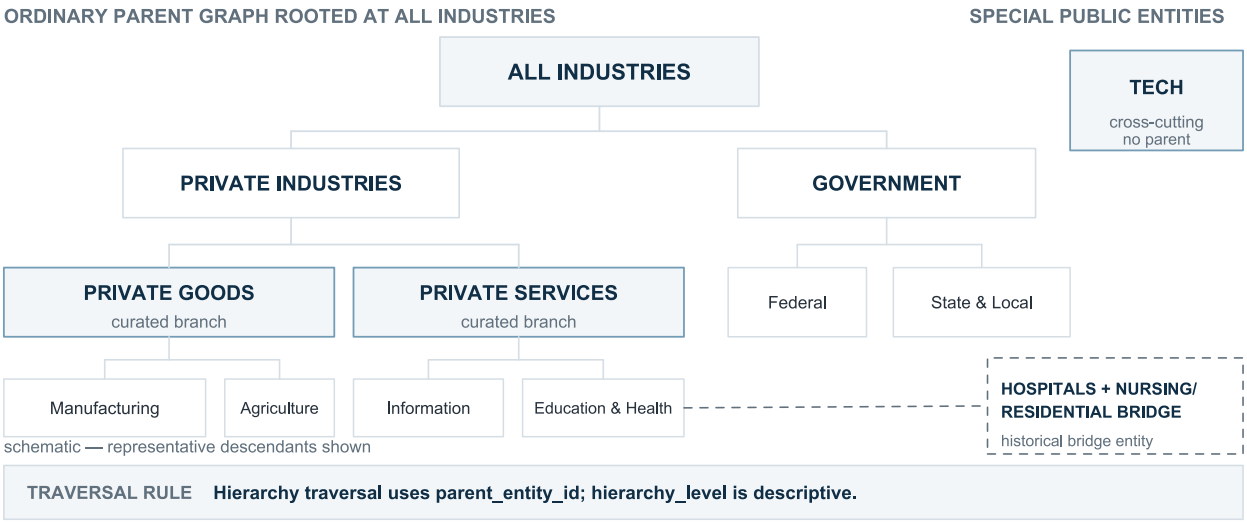
Only selected Industry Annual supplemental rows—principally interest and proprietors’ income—use `when_available`; most statement rows use `always` and render availability according to the data contract.

E Industry Entity Architecture

Figure E.1 provides the visual map of the Industry entity architecture; the appendix tables retain the exact identifiers and parent relationships.

The Stage 4 entity dictionary contains **99 Industry entities**, in addition to National and Corporate. The public Industry universe comprises **97 ordinary current-hierarchy entities**, one **cross-cutting ICT-producing aggregate (`Tech`)**, and one **historical Hospitals/Nursing bridge entity**. `IIS-A` contains all 99 entity IDs; `IIS-Q` contains 97, excluding exactly `Tech` and `EduHth.Hth.HospNurs.Bridge`.

Figure E.1 — EARL Industry entity architecture



E.1 Hierarchy interpretation

`parent_entity_id` is the authoritative public parent-child relation. `hierarchy_level` is descriptive and is **not** a substitute for graph traversal: curated intermediate nodes can be inserted while detailed descendants retain source/BEA depth, so a child can share the same numerical level as its parent. The cross-cutting `Tech` aggregate deliberately has no parent and no hierarchy level.

E.2 Complete Industry entity dictionary

Table E.1 lists all public Industry entities and their parent relationships.

Table E.1 — Complete Industry entity dictionary

Entity ID	Entity label	Parent	Source code	Status / note
All	All Industries	—	II	—
Gov	Government	All	G	—
Fed	Federal	Gov	GF	—
Fed.Gen	General government	Fed	GFG	BEA fallback
Fed.Gen.Def	National defense	Fed.Gen	GFGD	BEA fallback
Fed.Gen.XDef	Nondefense	Fed.Gen	GFGN	BEA fallback
Fed.Enterp	Government enterprises	Fed	GFE	BEA fallback
StLoc	State & Local	Gov	GSL	—
StLoc.Gen	General government	StLoc	GSLG	BEA fallback
StLoc.Enterp	Government enterprises	StLoc	GSLE	BEA fallback
Pvt	Private Industries	All	PVT	—
PvtGds	Private Goods	Pvt	PGOOD	—
Ag	Agriculture	PvtGds	11	—
Ag.Farms	Farms	Ag	111CA	BEA fallback

Entity ID	Entity label	Parent	Source code	Status / note
Ag.Forest	Forestry, fishing, and related activities	Ag	113FF	BEA fallback
Constr	Construction	PvtGds	23	—
Mfg	Manufacturing	PvtGds	31G	—
Mfg.DurGds	Durable goods	Mfg	33DG	BEA fallback
Mfg.DurGds.CompElec	Computer and electronic products	Mfg.DurGds	334	BEA fallback
Mfg.DurGds.Elec	Electrical equipment, appliances, and components	Mfg.DurGds	335	BEA fallback
Mfg.DurGds.FabMetal	Fabricated metal products	Mfg.DurGds	332	BEA fallback
Mfg.DurGds.Furn	Furniture and related products	Mfg.DurGds	337	BEA fallback
Mfg.DurGds.Machinery	Machinery	Mfg.DurGds	333	BEA fallback
Mfg.DurGds.Misc	Miscellaneous manufacturing	Mfg.DurGds	339	BEA fallback
Mfg.DurGds.MotoVeh	Motor vehicles, bodies and trailers, and parts	Mfg.DurGds	3361MV	BEA fallback
Mfg.DurGds.XMetal	Nonmetallic mineral products	Mfg.DurGds	327	BEA fallback
Mfg.DurGds.OthTrans	Other transportation equipment	Mfg.DurGds	3364OT	BEA fallback
Mfg.DurGds.PrimMetal	Primary metals	Mfg.DurGds	331	BEA fallback
Mfg.DurGds.Wood	Wood products	Mfg.DurGds	321	BEA fallback
Mfg.XDurGds	Nondurable goods	Mfg	31ND	BEA fallback
Mfg.XDurGds.Apparel	Apparel and leather and allied products	Mfg.XDurGds	315AL	BEA fallback
Mfg.XDurGds.Chem	Chemical products	Mfg.XDurGds	325	BEA fallback
Mfg.XDurGds.Food	Food and beverage and tobacco products	Mfg.XDurGds	311FT	BEA fallback
Mfg.XDurGds.Paper	Paper products	Mfg.XDurGds	322	BEA fallback
Mfg.XDurGds.PetrolCoal	Petroleum and coal products	Mfg.XDurGds	324	BEA fallback
Mfg.XDurGds.Plast	Plastics and rubber products	Mfg.XDurGds	326	BEA fallback
Mfg.XDurGds.Print	Printing and related support activities	Mfg.XDurGds	323	BEA fallback
Mfg.XDurGds.Text	Textile mills and textile product mills	Mfg.XDurGds	313TT	BEA fallback
Min	Mining	PvtGds	21	—
Min.XOilGas	Mining, except oil and gas	Min	212	BEA fallback
Min.OilGas	Oil and gas extraction	Min	211	BEA fallback
Min.Suppt	Support activities for mining	Min	213	BEA fallback
PvtSrvs	Private Services	Pvt	PSERV	—
ArtAcmdFd	Arts, Entertainment & Food Services	PvtSrvs	7	—
ArtAcmdFd.AcmdFdSrv	Accommodation and food services	ArtAcmdFd	72	BEA fallback
ArtAcmdFd.AcmdFdSrv.Acmd	Accommodation	ArtAcmdFd.AcmdFdSrv	721	BEA fallback
ArtAcmdFd.AcmdFdSrv.FdSrv	Food services and drinking places	ArtAcmdFd.AcmdFdSrv	722	BEA fallback
ArtAcmdFd.Art	Arts, entertainment, and recreation	ArtAcmdFd	71	BEA fallback
ArtAcmdFd.Art.Amuse	Amusements, gambling, and recreation industries	ArtAcmdFd.Art	713	BEA fallback
ArtAcmdFd.Art.PerfArt	Performing arts, spectator sports, museums, and related activities	ArtAcmdFd.Art	711AS	BEA fallback
EduHth	Education & Health	PvtSrvs	6	—
EduHth.Edu	Educational services	EduHth	61	BEA fallback
EduHth.Hth	Health care and social assistance	EduHth	62	BEA fallback
EduHth.Hth.Amb	Ambulatory health care services	EduHth.Hth	621	BEA fallback
EduHth.Hth.Hosp	Hospitals	EduHth.Hth	622	BEA fallback
EduHth.Hth.HospNurs.Bridge	Hospitals and nursing and residential care facilities (historical bridge)	EduHth.Hth	—	Historical bridge

Entity ID	Entity label	Parent	Source code	Status / note
EduHth.Hth.Nurs	Nursing and residential care facilities	EduHth.Hth	623	BEA fallback
EduHth.Hth.SciAst	Social assistance	EduHth.Hth	624	BEA fallback
Fin	Finance & Real Estate	PvtSrvs	FIRE	—
Fin.FinIns	Finance and insurance	Fin	52	BEA fallback
Fin.FinIns.FedBks	Federal Reserve banks, credit intermediation, and related activities	Fin.FinIns	521CI	BEA fallback
Fin.FinIns.Funds	Funds, trusts, and other financial vehicles	Fin.FinIns	525	BEA fallback
Fin.FinIns.Ins	Insurance carriers and related activities	Fin.FinIns	524	BEA fallback
Fin.FinIns.Sec	Securities, commodity contracts, and investments	Fin.FinIns	523	BEA fallback
Fin.REnt	Real estate and rental and leasing	Fin	53	BEA fallback
Fin.REnt.RE	Real estate	Fin.REnt	531	BEA fallback
Fin.REnt.Rental	Rental and leasing services and lessors of intangible assets	Fin.REnt	532RL	BEA fallback
Info	Information	PvtSrvs	51	—
Info.Bcast	Broadcasting and telecommunications	Info	513	BEA fallback
Info.DataProc	Information and data processing services	Info	514	BEA fallback
Info.MotionPic	Motion picture and sound recording industries	Info	512	BEA fallback
Info.Publ	Publishing industries (includes software)	Info	511	BEA fallback
OthSrvsXGov	Other Services	PvtSrvs	81	—
PrfSrv	Professional & Business Services	PvtSrvs	PROF	—
PrfSrv.AdmWst	Administrative and waste management services	PrfSrv	56	BEA fallback
PrfSrv.AdmWst.AdmSup	Administrative and support services	PrfSrv.AdmWst	561	BEA fallback
PrfSrv.AdmWst.WstRem	Waste management and remediation services	PrfSrv.AdmWst	562	BEA fallback
PrfSrv.Mgmt	Management of companies and enterprises	PrfSrv	55	BEA fallback
PrfSrv.Sci	Professional, scientific, and technical services	PrfSrv	54	BEA fallback
PrfSrv.Sci.Comp	Computer systems design and related services	PrfSrv.Sci	5415	BEA fallback
PrfSrv.Sci.Legal	Legal services	PrfSrv.Sci	5411	BEA fallback
PrfSrv.Sci.Misc	Miscellaneous professional, scientific, and technical services	PrfSrv.Sci	5412OP	BEA fallback
Rtl	Retail Trade	PvtSrvs	44RT	—
Rtl.FoodBev	Food and beverage stores	Rtl	445	BEA fallback
Rtl.GenMerch	General merchandise stores	Rtl	452	BEA fallback
Rtl.MotoVeh	Motor vehicle and parts dealers	Rtl	441	BEA fallback
Rtl.Oth	Other retail	Rtl	4A0	BEA fallback
Trans	Transportation & Warehousing	PvtSrvs	48TW	—
Trans.Air	Air Transportation	Trans	481	—
Trans.Oth	Other transportation and support activities	Trans	487OS	BEA fallback
Trans.Pipe	Pipeline transportation	Trans	486	BEA fallback
Trans.Rail	Rail transportation	Trans	482	BEA fallback
Trans.Transit	Transit and ground passenger transportation	Trans	485	BEA fallback
Trans.Truck	Truck Transportation	Trans	484	—
Trans.Warehs	Warehousing and storage	Trans	493	BEA fallback
Trans.Water	Water transportation	Trans	483	BEA fallback
Utl	Utilities	PvtSrvs	22	—
Wholsl	Wholesale Trade	PvtSrvs	42	—
Cross-cutting aggregate				

Entity ID	Entity label	Parent	Source code	Status / note
Tech	Information-communications-technology-producing industries	—	ICT	Outside mutually exclusive hierarchy

Note: Rows are ordered by the authoritative parent graph beginning at All Industries; siblings are alphabetized by entity label, and indentation in Entity ID shows graph depth; Entity labels remain flush left for scanning. Parent is the canonical parent_entity_id. The retained hierarchy_level field is intentionally not printed because it is descriptive source/curation metadata rather than authoritative tree depth and can equal its parent's numerical value. "BEA fallback" marks rows whose label, parent, or retained level uses BEA hierarchy/name fields where curated display fields are unavailable. EduHth.Hth.HospNurs.Bridge remains inside the Health Care branch as a historical bridge and has no standalone BEA source code. Tech is shown separately because it is an official cross-cutting ICT aggregate outside the mutually exclusive parent graph. IIS-A contains all 99 Industry entities; IIS-Q contains 97 and excludes exactly Tech and the historical bridge.

Source: EARL revised methodology manuscript; canonical Stage 4 entity-metadata.csv; IIS-A/IIS-Q entity-presence evidence derived from the canonical observation files.

E.3 Special entities and constructions

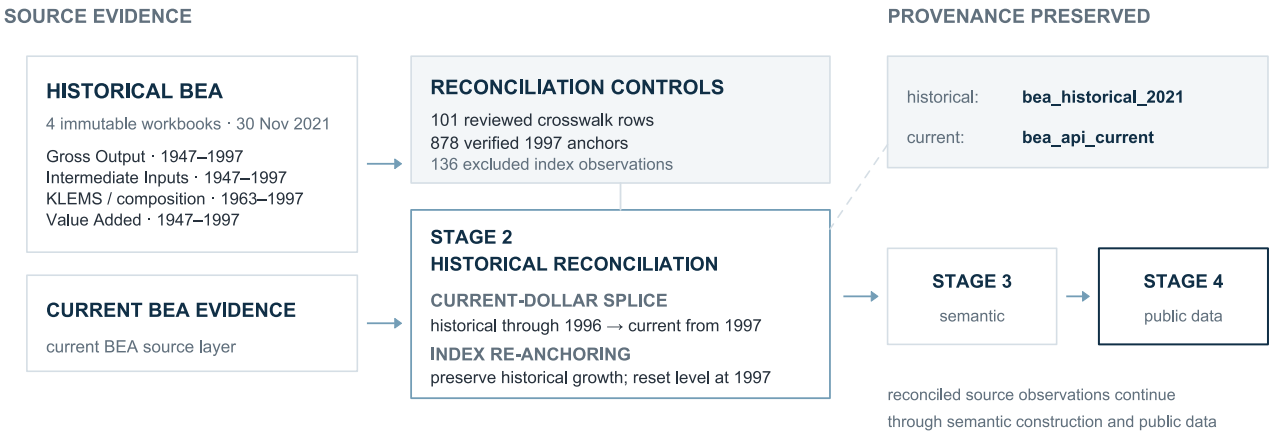
- `PvtGds` and `PvtSrvs` are curated intermediate public branches under Private Industries.
- `Tech` is the official information-communications-technology-producing cross-cutting aggregate and sits outside the ordinary parent graph.
- `EduHth.Hth.HospNurs.Bridge` preserves the historical combined Hospitals and Nursing/Residential concept. It is a public EARL bridge entity, not a current BEA taxonomy node and not a license to allocate historical values into modern children.
- Hierarchy rollups are permitted only for specified series/parent combinations and only with a complete required child set.

F Historical Reconciliation Reference

Figure F.1 shows the evidence and control architecture supporting the historical splice; the tables below provide the exact retained sources, crosswalk classifications, and anchor inventory.

Appendix F is the compact technical reference for the annual Industry historical reconciliation. Section 11 explains the methodological rationale; Technical Note 01 and its machine-readable crosswalk remain the specialized row-level public evidence.

Figure F.1 — Historical reconciliation evidence and control architecture



F.1 Retained official source evidence

Table F.1 identifies the retained official historical workbooks. EARL retains four immutable official BEA Historical Industry Accounts workbooks from **November 30, 2021**. Their cryptographic hashes are part of the historical source manifest.

Table F.1 — Retained official historical source evidence

Source file	Historical coverage
`GDPbyInd_GO_1947-1997.xlsx`	Gross Output, 1947–1997
`GDPbyInd_II_1947-1997.xlsx`	Intermediate Inputs, 1947–1997
`GDPbyInd_KLEMS_1963-1997.xlsx`	KLEMS / Composition of Gross Output, 1963–1997
`GDPbyInd_VA_1947-1997.xlsx`	Value Added, 1947–1997

Note: All four files are immutable official BEA Historical Industry Accounts workbooks from the November 30, 2021 release. EARL records and verifies their full SHA-256 hashes in the historical source manifest; the hashes are omitted here for readability.
Source: EARL historical reconciliation source manifest; BEA Historical Industry Accounts, November 30, 2021 release.

The compact historical reference layer contains **57,520 observations**. Reconciled historical current-dollar observations extend through 1996; 1997 is retained separately as the overlap/anchor year for index reconciliation.

F.2 Crosswalk architecture

Table F.2 summarizes the historical crosswalk outcomes. The released companion crosswalk contains **101 verified rows**: **97** one-to-one current public-entity mappings, **2** current-taxonomy reference concepts, and **2** historical bridges. These are crosswalk outcomes, not a second definition of the 99-entity Stage 4 Industry universe.

Table F.2 — Historical crosswalk classification

Classification	Rows
Panel A. Mapping relationships	
Exact current description	80
Aggregate of current children	8
Contextual current summary	3
Supplemental aggregate	3
Current-description rename	2
Historical combined concept / current split — no allocation	2
Exact analytical total	1
Exact analytical aggregate	1
Exact current sector	1
Panel B. Historical comparability	
Direct or aggregatable	96
Supplemental, nonhierarchical	3
Combined only historically	2

Note: Panels A and B classify the same 101 crosswalk outcomes by mapping relationship and historical comparability, respectively; each panel therefore sums to 101. The two outcomes combined only historically remain subject to the no-allocation rule.
Source: EARL historical reconciliation crosswalk.

The four rows without a one-to-one current public entity are **Real estate** (historical bridge to Housing + Other real estate), **Housing** (current-taxonomy reference), **Other real estate** (current-taxonomy reference), and **Hospitals and nursing and residential care facilities** (historical bridge to Hospitals + Nursing/residential care). Only the two historical bridges carry `PROHIBITED_WITHOUT_OFFICIAL_SOURCE`; EARL does not allocate them synthetically.

F.3 Current-dollar splice and index re-anchoring

Table F.3 summarizes the verified 1997 index anchors used for re-anchoring.

For eligible current-dollar measures, historical evidence is used through 1996 and current BEA evidence takes precedence from 1997 onward. The overlap is not averaged or selected for smoothness.

For a historical price or quantity index, EARL uses:

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

The multiplicative factor changes the reference level but preserves historical growth rates. The anchor file contains **878 verified 1997 historical anchors**. The complete variable-level inventory is:

Table F.3 — 1997 historical index-anchor inventory

Historical index family	Price	Quantity
Gross Output	89	89
Intermediate Inputs	89	89
Energy Inputs	87	87
Materials Inputs	87	87
Services Inputs	87	87

Note: Cells report verified 1997 anchors. Price and quantity indexes each account for 439 anchors, for 878 in total. Each anchor pairs historical and current 1997 levels so the historical index can be re-expressed on the current overlap basis while preserving its historical growth path. Exact variable IDs and anchor values remain in the canonical anchor file.
Source: EARL historical reconciliation 1997 anchor file.

EARL also records **136 historical index observations excluded from the re-anchored history**. All belong to `EduHth.Hth.HospNurs.Bridge` and lack an exact current 1997 anchor; no synthetic child allocation is used.

F.4 Validation record

The validated historical release passed nine core checks covering unique keys, complete identifiers, production identities, input decomposition, value-added decomposition, GOS/operating-expense reconciliation, index normalization, historical growth preservation, and observation provenance. Eleven selected Gross Output observations matched the retained official workbooks exactly. Most historical accounting identities use a **$150 million** tolerance; two aggregate cross-table GVA checks use **$250 million**.

Technical Note 01 version 1.0 records a methodology validation date of **September 2, 2026** and source commit `275aa34f8de950e560f7ae1709020baa42bce7e7`.

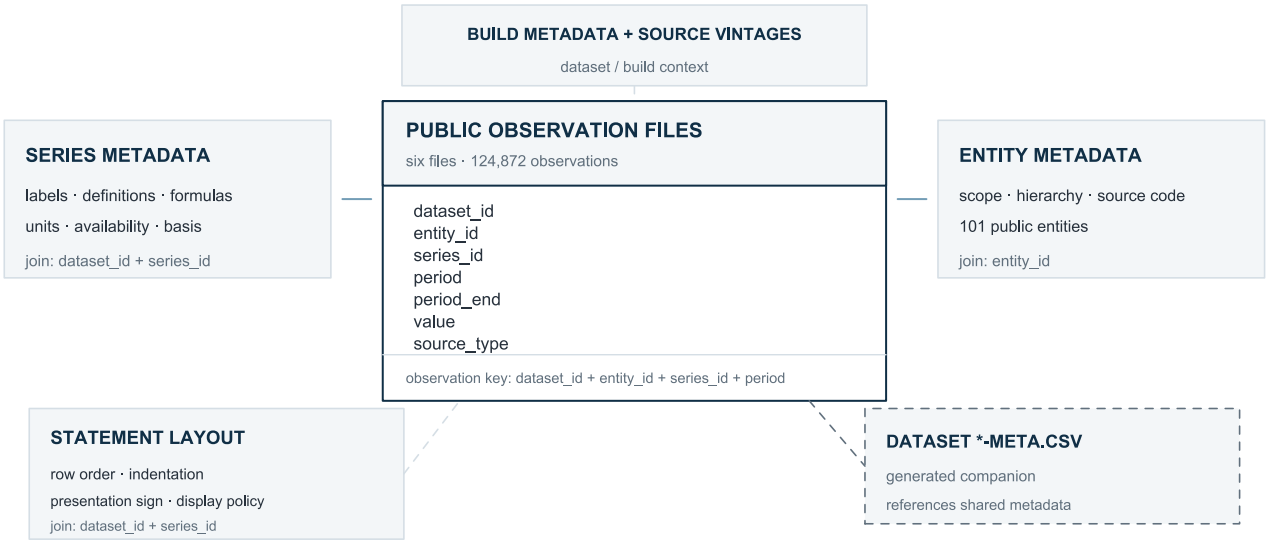
G Public Data Schema

Figure G.1 summarizes the relational structure of the public data package; the tables below document observation provenance and the complete schema fields.

The Stage 4 public data package contains 18 CSV files. Its six observation files contain **124,872 public observations** in the documented build and use the same seven-column schema:

`dataset_id, entity_id, series_id, period, period_end, value, source_type`.

Figure G.1 — EARL public data relational schema



G.1 Observation grain and keys

One row represents one public series for one economic entity in one reference period. The published observation key is `dataset_id + entity_id + series_id + period`; the documented build contains zero duplicate keys. `series_id` resolves jointly with `dataset_id` in `series-metadata.csv`; `entity_id` resolves in `entity-metadata.csv`.

G.2 Observation provenance

Table G.1 reports observation provenance in the documented public data build.

Table G.1 — Observation provenance

Source type	Meaning	Public observations
bea_api_current	Current BEA source observation	73,753
bea_historical_2021	Historical BEA Industry observation (2021 release; pre-1997 history)	23,774
derived	EARL-calculated observation	26,866
hierarchy_rollup	Parent Industry observation constructed from published children	364
derived_fallback	All Industries GVA fallback where directly reported BEA GVA is unavailable	115
Total		**124,872**

Note: source_type records cell-level observation provenance and is distinct from a series measurement basis. A source-reported series can contain hierarchy-rollup or fallback observations; an EARL-calculated series contains derived observations. Counts are recomputed from the six Stage 4 public observation files.

Source: EARL revised methodology manuscript; Stage 4 public observation CSVs and public-schema.csv. `source_type` is cell-level provenance and must not be confused with series-level measurement basis. A source-reported series can contain a `hierarchy_rollup` or `derived_fallback` observation; an EARL-calculated series contains `derived` observations.

G.3 Public schema dictionary

Table G.2 provides the complete public schema dictionary.

The frozen `public-schema.csv` contains 45 field definitions:

Table G.2 — Complete public schema dictionary

Column	Type	Description	Allowed values
Observation files			`nis|cis|iis-annual|quarterly.csv`
dataset_id	string	Statement family and frequency; the key into series-metadata.csv and statement-layout.csv.	NIS-A; NIS-Q; CIS-A; CIS-Q; IIS-A; IIS-Q
entity_id	string	Economic entity or scope represented by the observation: 'national' (NIS), 'corporate' (CIS), or an Industry entity code (IIS; 'All' = All Industries, 'Pvt' = private industries, ...).	—
series_id	string	Series key; joins series-metadata.csv (dataset_id + series_id) for label, definition, formula, unit and availability.	—
period	string	Annual 'YYYY'; quarterly 'YYYYQn'.	—
period_end	date	Last day of the period, ISO 8601.	—
value	number	In the series' unit_id (series-metadata.csv): usd_millions, usd_millions_saar (seasonally adjusted, annual rate), thousand_fte, usd_per_fte, or ratio (a decimal, 0.25 = 25%). Always the reported/natural sign: the statement layout's presentation_sign says how to DISPLAY it. Missing observations are absent, not blank.	—
source_type	string	How this observation was obtained.	bea_api_current; bea_historical_2021; hierarchy_rollup; derived; derived_fallback
Series metadata			`series-metadata.csv`
*	see file	The public series metadata, unmodified apart from removing the byte-order mark and normalizing line endings.	—
Statement layout			`statement-layout.csv`
*	see file	The public statement layout, unmodified apart from removing the byte-order mark and normalizing line endings.	—
Source vintages			`source-vintages.csv`
*	see file	BEA release identities (id, name, date, URL) referenced by build-metadata.csv; rows are copied from the Source/Vintage Registry.	—
Build metadata			`build-metadata.csv`
build_id	string	Identifies the build that produced these files.	—
built_at_utc	datetime	EARL processing/build time (UTC).	—
git_commit	string	Source revision the build ran from.	—
git_dirty	boolean	TRUE if the working tree had uncommitted changes at build time.	—
file	string	Public observation dataset/value file described by this row; build-metadata is dataset-oriented and does not add separate rows for metadata artifacts.	—
dataset_id	string	Dataset the file holds.	—
rows	integer	Observation rows in the file.	—
series	integer	Distinct series in the file.	—
entities	integer	Distinct entities in the file.	—
first_period	string	Earliest period with data.	—
last_period	string	Latest period with data (the data-through period).	—
sha256	string	SHA-256 of the file.	—

Column	Type	Description	Allowed values
contract_series_metadata_sha256	string	SHA-256 of supporting/public_contract/series-metadata.csv as committed; the public series specification used by this build.	—
contract_statement_layout_sha256	string	SHA-256 of supporting/public_contract/statement-layout.csv as committed.	—
source_acquisition_ids	string	EARL retrieval events (';'-separated) the data derive from; each has a row in data/raw/<dataset>/manifest.csv with the retrieval timestamp, request and raw-file hash.	—
source_vintage_ids	string	BEA release vintages the data derive from (';'-separated); resolve in source-vintages.csv.	—
retrieved_from_utc	datetime	Earliest EARL retrieval timestamp among the acquisitions.	—
retrieved_to_utc	datetime	Latest EARL retrieval timestamp among the acquisitions.	—
vintage_basis	string	How each vintage was determined: bea_stated_last_revised (BEA's own response states the date), registry_asserted (BEA states none; the Source/Vintage Registry assigns it), or unresolved.	—
Dataset companion metadata			nis|cis|iis-annual|quarterly-meta.csv
dataset_file	string	Public observation CSV described by this companion file.	—
dataset_id	string	Statement-family/frequency identifier for the public dataset.	NIS-A; NIS-Q; CIS-A; CIS-Q; IIS-A; IIS-Q
column	string	Column of the seven-column public observation schema being described.	—
type	string	Public data type of the column.	—
description	string	Human-readable definition of the column.	—
allowed_values	string	Allowed values where the public data specification defines a closed or documented set.	—
example_value	string	Value from one real observation selected deterministically from the dataset.	—
reference_file	string	Public metadata file containing the authoritative detail for this column.	public-schema.csv; entity-metadata.csv; series-metadata.csv
notes	string	Additional interpretation or example-selection note.	—
Entity metadata			entity-metadata.csv
entity_id	string	Stable entity key used in the public observation files.	—
entity_label	string	Human-readable entity or economic-scope label.	—
entity_type	string	Kind of public entity.	national; corporate; industry
parent_entity_id	string	Parent entity in the public Industry hierarchy; blank for top-level scopes and documented cross-cutting aggregates outside the mutually exclusive hierarchy.	—
hierarchy_level	integer	Descriptive hierarchy-level metadata retained for source/curation purposes. Top-level statement scopes and the Industry root use level 1; blank for documented cross-cutting aggregates outside the hierarchy. parent_entity_id is authoritative for hierarchy traversal.	—
source_code	string	BEA Industry code where applicable; blank for National/Corporate scopes and synthetic historical bridge entities without their own BEA code.	—
notes	string	Additional scope or hierarchy information.	—

Note: Each section band pairs the reader-facing file/file-class category at left with its file or file-class identifier at right; body rows preserve Column, Type, Description, and Allowed values. The printed wording uses current EARL terminology and clarifies hierarchy_level consistently with Appendix E; public-schema.csv is retained unchanged. For source_type, the table lists the five defined identifiers only; their meanings and observation counts appear in Table G.1. An em dash means the public schema does not define a closed allowed-value set.

Source: EARL revised methodology manuscript; Stage 4 public-schema.csv.

G.4 Shared metadata files

`series-metadata.csv` contains 128 rows and 25 fields; `entity-metadata.csv` contains 101 rows and seven fields; `statement-layout.csv` contains 132 rows and 12 fields; `build-metadata.csv` contains six dataset-level rows and 19 fields; `source-vintages.csv` resolves governed official source-release identity. Each of the six dataset-specific `*-meta.csv` companions contains exactly seven rows—one per observation column—and points users to these shared authoritative files rather than duplicating their economics.

G.5 Absence and sign semantics

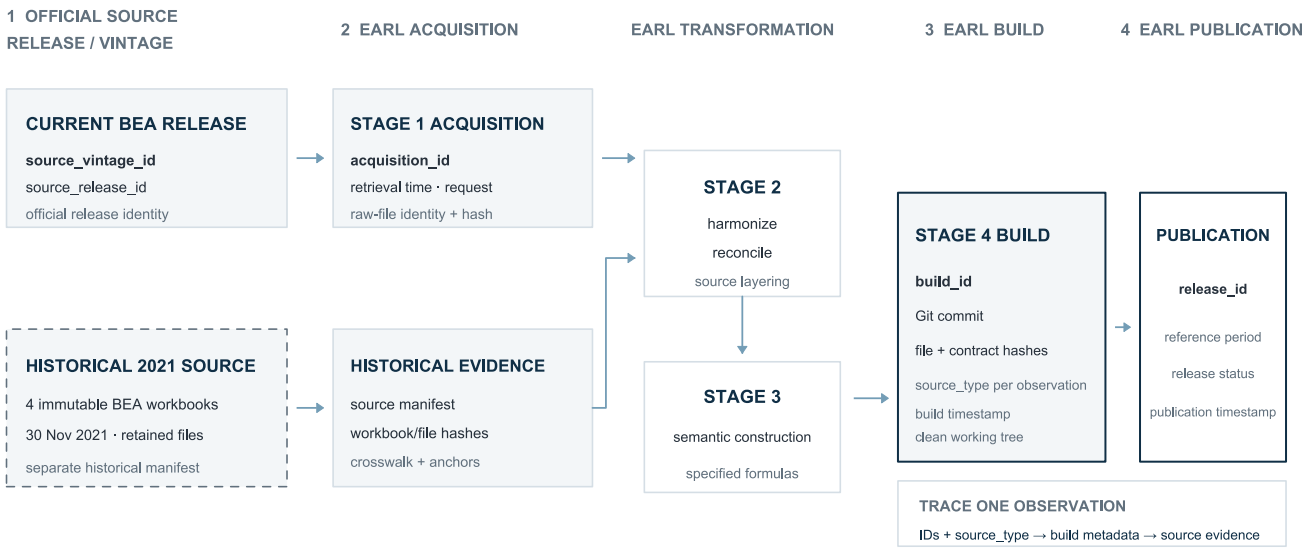
Unavailable public observations are normally absent from the observation file rather than represented as zero-valued or blank rows. The stored `value` retains its natural source/formula sign. `presentation_sign` in the statement layout can invert a value for display without modifying the stored observation.

H Provenance and Vintage Schema

Figure H.1 traces an observation from source evidence through publication; the tables below define the identifiers, fields, acquisition inventory, and documented build.

EARL separates the identity of a **retrieval event**, an **official source release**, a **Stage 4 build**, and an **EARL publication**. These identifiers answer different questions and should not be used as synonyms.

Figure H.1 — EARL end-to-end provenance chain



H.1 Principal identifiers

Table H.1 defines the principal provenance identifiers and their associated clocks.

Table H.1 — Principal provenance fields and clocks

Field	Provenance clock / level	Meaning and distinction
`source_vintage_id`	Official source	Normalized official-source vintage used to group source evidence; distinct from EARL retrieval time.
`source_release_id`	Official source	Identifies the official release event associated with the normalized vintage; not an EARL acquisition.
`acquisition_id`	Acquisition	Identifies one Stage 1 table-level retrieval event; multiple acquisitions can belong to the same source vintage.
`build_id`	Build	Identifies one Stage 4 public-data build and binds outputs to build time, code identity, and provenance; not a publication identity.
`release_id`	Publication	Identifies one EARL Statistical Release; belongs to the publication layer rather than the Stage 4 build.
`source_type`	Observation lineage	Records how an observation entered canonical data; it is not a provenance clock or series-level measurement basis.

Note: EARL separates four clocks: official source release/vintage → acquisition → build → publication. Observation-level provenance is recorded separately through `source_type`. BEA API server-production time, where available, is auxiliary metadata and is neither the official release date nor EARL retrieval time.

Source: EARL Stage 1 acquisition manifests; Stage 4 source-vintage and build metadata; Statistical Release metadata.

H.2 Acquisition manifest schema

Table H.2 records the fields captured for each table-level acquisition.

Table H.2 — Stage 1 acquisition-manifest schema

Field	Meaning	Population / condition
Panel A. Retrieval and request		
`acquisition_id`	Identity of one table-level retrieval event.	Every row; unique.
`dataset`	Canonical dataset driver that requested the source table.	Every row.
`source_agency`	Official source agency.	Every row; BEA in the documented build.
`source_database`	BEA database used for the request.	Every row; NIPA or GDP by Industry.
`table_id`	Source-table identifier within the selected database.	Every row.
`frequency`	Requested annual or quarterly frequency.	Every row.
`year_param`	Year parameter sent with the source request.	Every row.
`industry_param_n`	Number of Industry codes in the request.	GDP by Industry requests only.
`industry_param_sha256`	SHA-256 fingerprint of the requested Industry-code set; records the request without reproducing the list.	GDP by Industry requests only.
Panel B. Timing and response		
`retrieved_at_utc`	EARL-generated UTC timestamp when the retrieval succeeded.	Every row; acquisition clock.
`bea_utc_production_time`	Time at which the BEA server produced the API response.	Where supplied; auxiliary metadata, not a release date.
`bea_last_revised`	BEA-stated table revision date parsed from response notes.	Where BEA states it; NIPA in the documented build.
`bea_table_title`	Source-table title retained from response metadata.	Where captured; NIPA in the documented build.
Panel C. Raw-file integrity		
`raw_file`	Repository-relative path to the source-preserving raw capture associated with the retrieval.	Every row.
`raw_file_sha256`	SHA-256 hash of the exact raw-file bytes written by Stage 1.	Every row.
`raw_file_rows`	Row count of the raw capture.	Every row.

Note: One manifest row represents one BEA table-level retrieval. `retrieved_at_utc` is EARL acquisition time; `bea_utc_production_time` is BEA server-production time and is not a release date. `bea_last_revised` is recorded only when BEA explicitly states a revision date; no date is invented for GDP by Industry. Industry request sets are fingerprinted without storing the code list, and API credentials are excluded from the manifest.

Source: EARL Stage 1 acquisition manifests and canonical provenance helper.

EARL distinguishes **official source release → EARL acquisition → EARL build → EARL publication**. BEA API server-production time is retained where available but is neither the official release date nor EARL retrieval time.

H.3 Current acquisition inventory

Table H.3 summarizes the acquisition inventory for the documented build.

The documented build derives from **58 table-level acquisition records**, all retrieved on October 7, 2026.

Table H.3 — Current acquisition inventory

Dataset	ID	Acquisitions	Retrieval time (UTC), Oct. 7, 2026	Vintage evidence
National annual	NIS-A	17	20:12:08–20:12:13	BEA-stated NIPA revision date + GDP by Industry official-release linkage
National quarterly	NIS-Q	10	20:12:15–20:12:19	BEA-stated NIPA revision date + GDP by Industry official-release linkage
Corporate annual	CIS-A	2	20:12:22	BEA-stated NIPA revision date
Corporate quarterly	CIS-Q	2	20:12:24	BEA-stated NIPA revision date
Industry annual	IIS-A	20	20:12:26–20:12:33	BEA-stated NIPA revision date + GDP by Industry official-release linkage
Industry quarterly	IIS-Q	7	20:12:36–20:12:40	GDP by Industry official-release linkage

Note: The 58 table-level acquisitions shown above resolve to BEA's September 30, 2026 source vintage. NIPA provides a BEA-stated revision date. GDP by Industry responses do not state an equivalent release date, so their vintage linkage uses the governed Source/Vintage Registry and official release coverage.
Source: EARL Stage 1 acquisition manifests; Stage 4 build metadata; governed Source/Vintage Registry.

H.4 Current official source vintage

The current acquisitions resolve to `BEA-VIN-2026-09-30`, associated with BEA release `BEA-REL-2026-09-30`: **GDP (Third Estimate), Industries, Corporate Profits, State GDP, and State Personal Income, 2nd Quarter 2026; State PCE, 2025.** NIPA supplies machine-evidenced `LastRevised = 2026-09-30`; GDP by Industry does not expose an equivalent release-date field, so its assignment is curator-linked through the governed Source/Vintage Registry and official release coverage.

H.5 Documented build identity

Table H.4 summarizes the six public observation datasets in the documented build.

Table H.4 — Documented build dataset summary

Dataset	ID	Public file	Rows	Series	Entities	Observed span
National annual	NIS-A	nis-annual.csv	2,400	30	1	1929–2025
National quarterly	NIS-Q	nis-quarterly.csv	6,414	25	1	1947Q1–2026Q2
Corporate annual	CIS-A	cis-annual.csv	2,382	26	1	1929–2025
Corporate quarterly	CIS-Q	cis-quarterly.csv	7,836	26	1	1947Q1–2026Q2
Industry annual	IIS-A	iis-annual.csv	72,472	17	99	1947–2025
Industry quarterly	IIS-Q	iis-quarterly.csv	33,368	4	97	2005Q1–2026Q2

Note: The six files contain 124,872 observations in total and belong to build `BUILD-20261019T210359Z-ff20569`, generated October 19, 2026 at 21:03:59 UTC with a clean working tree. Observed span reports the first and last period present in each file and does not imply balanced coverage. Full file SHA-256 values, the source commit, and governing specification hashes are retained in `build-metadata.csv` and checked by QA but omitted from the printed table for readability.
Source: EARL Stage 4 build metadata and public observation files.

All six rows share build `BUILD-20261019T210359Z-ff20569`, build time `2026-10-19T21:03:59Z`, source commit `ff2056952ff08f84a036421c41ca3af0035b0403`, and `git_dirty = FALSE`. The governing series-contract hash is `2860c7c4dc5d4852e82879dc2d5146a03cd7b2f257dd971a958324d66d3378af`; the statement-layout hash is `e3465e322fd176e524da7645c48c190ad90d78896af33d3fc9ebc347a4e0162e`.

H.6 Historical provenance branch

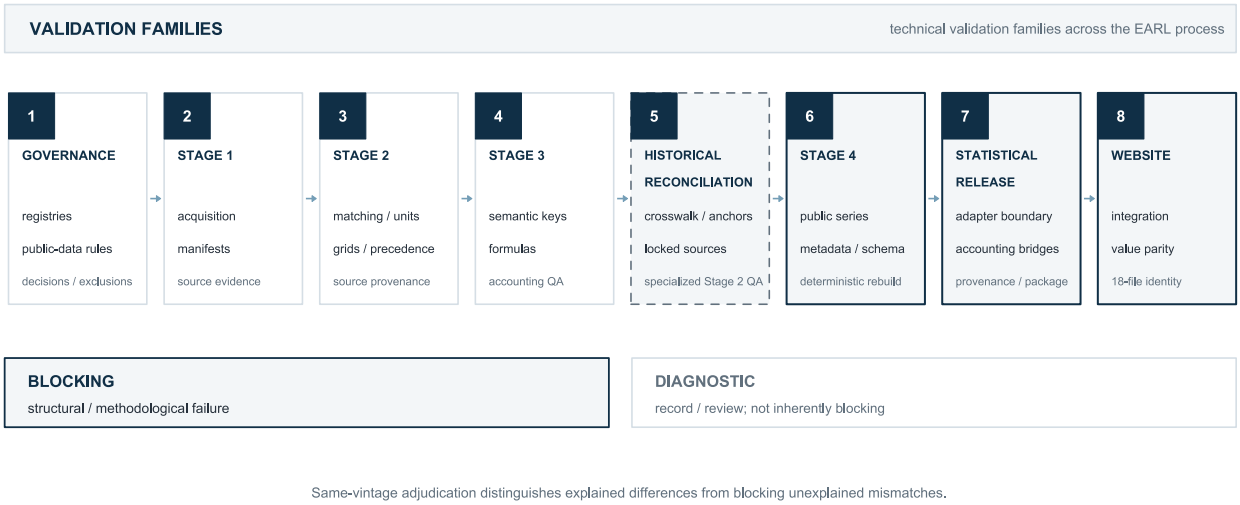
Historical annual Industry observations use the separate immutable November 30, 2021 source-manifest/crosswalk architecture documented in Appendix F and carry `source_type = bea_historical_2021`. The absence of a current-style `BEA-VIN-2021-11-30` row in the frozen current `source-vintages.csv` does not imply missing provenance; the historical workbooks and their hashes are the authoritative evidence.

I Validation Matrix

Figure I.1 summarizes the validation families and outcome classes; the tables below provide the detailed control matrix and same-vintage adjudication statuses.

EARL validates different assumptions at the earliest stage where they become testable. The public validation record distinguishes **blocking controls** from **diagnostic findings**. Blocking failures stop the relevant build or publication step; diagnostic findings are retained for review without changing otherwise valid observations.

Figure I.1 — EARL validation families and outcome classes



I.1 Consolidated validation matrix

Table I.1 records the principal validation controls, scopes, and tolerances.

Table I.1 — Consolidated validation matrix

Layer	Control / evidence	Assumption tested	Scope	Failure behavior	Rule / tolerance
Governance	Methodological decision QA	No unresolved methodological decision remains	EARL methodology	Blocking	No unresolved/ PENDING item
Contract	`CONTRACT-001`	Public data specification and internal registry describe the same methodology	All 6 datasets	Blocking	Exact mapping/formula/ classification agreement
Stage 1	`STAGE1-PROVENANCE-001`	Acquisition provenance is truthful and reproducible	Current API acquisitions	Blocking	Exact provenance/ schema rules
Stage 1	Raw-capture guards	Required source retrieval returns usable evidence	All required source tables	Blocking	Empty/unusable source refused
Stage 2	Source matcher tests	Raw observations match governed source concepts	All source mappings	Blocking	Exact governed table/ line/description/code matching
Stage 2	Unit validation	Source units agree with the documented conversion rules	All source series	Blocking	Supported unit rules only
Stage 2	Merge / precedence controls	Higher-priority source wins; lower-priority fills only missing cells	Historical/current overlap	Blocking	No averaging of vintages
Stage 2	`IND-HIER-001`	Direct All Industries GVA is preserved; fallback closes only specified gaps	Industry	Blocking	Specified fallback only
Stage 3	Formula-engine tests	Formula language and dependency graph are valid	All semantic datasets	Blocking	Unknown/circular/ duplicate definitions refused
Stage 3	Common semantic QA	Keys, names, provenance and values are structurally valid	All 6 processed datasets	Blocking	No duplicates/missing keys/nonfinite values
Stage 3	`STAGE3-EBITDA-001`	EBITDA methodology is implemented as defined	National/Corporate	Blocking	National overlap ≤ $100m; Corporate reconciliation ≤ $5m
Stage 3	`CORP-TRN-001`	Effective Tax Rate uses adjusted EBT and is missing when EBT ≤ 0	Corporate A/Q	Blocking	Exact formula guard
Historical	Historical annual QA	Splice, identities, indexes, and provenance remain valid	Industry Annual history	Blocking	$150m/$250m identities; growth error ≤ 1e-12
Historical	Locked source checks	Historical reference agrees with official retained workbooks	Historical Industry	Blocking	11/11 selected GO values exact
Regression	`QA-ADJUDICATION-001`	Differences from prior outputs are explained or blocked	Migration comparisons	Blocking for unexplained differences	Governed status classification
Stage 4	`STAGE4-PUBLIC-001`	Stage 4 is the exact defined nonmissing Stage 3 subset	All public datasets	Blocking	Exact values/source types; independent formulas
Stage 4	`STAGE4-METADATA-001`	Metadata, hierarchy, schema, companions are valid	Entire 18-file package	Blocking	Exact structural specification
Stage 4	Deterministic rebuild	Same governed inputs reproduce same public artifacts	Public package	Blocking regression control	Byte-for-byte identity
Release	Stage 4 adapter / consumer QA	Release consumes published EARL data instead of recreating economics	National/Corporate release	Blocking	Published schema/IDs/ values
Release	Accounting / bridge QA	Displayed statement identities reconcile	Release display periods	Blocking	Source identity ≤ $100m; derived/bridge ≤ $1m
Release	Provenance / packaging QA	Publication inherits complete Stage 4 provenance and preserves required publication artifacts	Statistical Release	Blocking before publication	Complete resolved provenance and hashes
Website	Integration acceptance	Website presentation equals Stage 4 after presentation-only transformations	Website	Blocks integration acceptance	Published value parity

Note: Blocking findings stop the affected build or publication step. Diagnostic findings remain visible for review when they reflect legitimate source or economic characteristics and do not alter otherwise valid observations. Formal control identifiers are retained where they materially aid reproducibility.

Source: EARL methodology, Section 15 and Appendix I; repository validation evidence.

I.2 Same-vintage adjudication statuses

Table I.2 defines the statuses used to classify same-vintage comparisons.

Table I.2 — Same-vintage adjudication status dictionary

Status	Meaning / evidence condition	Acceptance / reporting
Panel A. Nonblocking		
`MATCH`	Compared values agree; no difference requires adjudication.	Nonblocking; omitted from exception output.
`BOTH_MISSING`	No value is present in the compared outputs.	Nonblocking; omitted from exception output.
`SOURCE_REVISION`	New output agrees with same-vintage comparison evidence while the prior snapshot differs; requires same-vintage evidence.	Nonblocking; reported as an explained difference.
`SOURCE_EXTENSION`	Current comparison evidence supplies a period beyond the older snapshot coverage or fills a gap that the older snapshot did not contain.	Nonblocking; reported as an explained difference.
`NEW_BY_APPROVED_DESIGN`	A new-only value is covered by a documented inclusion rule at the applicable scope.	Nonblocking; reported as an explained difference.
`EXCLUDED_BY_DECISION`	A prior-output-only value is absent by a governed exclusion decision at the applicable scope.	Nonblocking; reported as an explained difference.
Panel B. Blocking		
`UNEXPLAINED_MISSING`	A value appears on only one side of the comparison and no governed addition, exclusion, or extension explains it.	Blocks acceptance; reported cell by cell.
`UNEXPLAINED_VALUE_MISMATCH`	Same-vintage values are both present but differ without a governed explanation.	Blocks acceptance; reported cell by cell.
`DUPLICATE_KEY`	A comparison key is repeated within one source.	Blocks acceptance; reported cell by cell.

Note: Same-vintage evidence is required before a difference can be classified as a source revision; without it, a mismatch remains unexplained and blocks acceptance. Documented inclusion and exclusion rules apply only at their recorded scope and cannot excuse differences for other entities or series.

Source: EARL same-vintage comparison harness; `QA-ADJUDICATION-001` methodology manuscript.

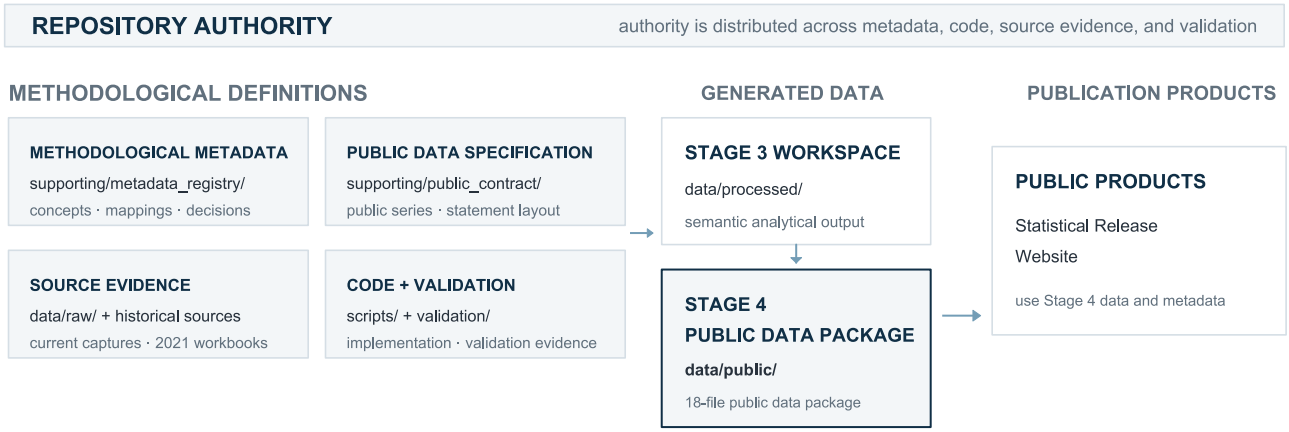
Explained nonblocking statuses are `MATCH`, `BOTH_MISSING`, `SOURCE_REVISION`, `SOURCE_EXTENSION`, `NEW_BY_APPROVED_DESIGN`, and `EXCLUDED_BY_DECISION`. Blocking statuses are `UNEXPLAINED_MISSING`, `UNEXPLAINED_VALUE_MISMATCH`, and `DUPLICATE_KEY`. A source-revision classification requires same-vintage evidence; “BEA probably revised it” is not accepted as a generic explanation.

J Repository and Reproducibility Mapping

Figure J.1 summarizes the repository roles and reproducibility chain; the tables below identify the exact authority classes, shared modules, dataset pipelines, and public package files.

Appendix J maps the methodology to its repository implementation and distinguishes methodological definitions, generated data, publication products, and historical technical identifiers. The repository retains older NISP/USEAL identifiers where exact historical reproducibility requires them.

Figure J.1 — Repository authority and reproducibility map



J.1 Repository authority hierarchy

Table J.1 distinguishes methodological authorities, generated artifacts, and publication products.

Table J.1 — Repository authority hierarchy

Authority class	Principal location	Role
Governed metadata	`supporting/metadata_registry/`	Variables, source mappings, entities, additions/ exclusions, vintages and decisions
Governed public data specification	`supporting/public_contract/`	Public series specification and statement layout
Current source evidence	`data/raw/`	Source-preserving current BEA captures and acquisition manifests
Historical source evidence	`supporting/historical_reconciliation/`	Immutable 2021 workbooks, crosswalk, anchors and omissions
Methodological implementation	`scripts/`	Stage 1–4 implementation
Semantic output	`data/processed/`	Stage 3 analytical workspace
Public data package	`data/public/`	Stage 4 18-file published data package
Validation evidence	`validation/`	Tests and publication/integration evidence
Publication products	`nisp-sandbox/`	Statistical Release and website

Note: Repository location does not by itself establish methodological authority. Governed metadata and the public data specification define the economic concepts and publication structure; generated data and publication products inherit those definitions. Historical `nisp_` identifiers remain where exact repository reproducibility requires them.

Source: EARL methodology, Appendix J; repository structure as documented for this methodology version.

J.2 Shared pipeline modules

Table J.2 identifies the shared implementation modules that support the documented methodology.

Table J.2 — Shared pipeline modules

Module	Layer	Primary responsibility
`scripts/nisp_registry.R`	Shared	Loads and resolves governed metadata used across the pipeline.
`scripts/nisp_fetch.R`	Stage 1	Retrieves BEA responses through the shared dataset-agnostic fetch layer.
`scripts/nisp_provenance.R`	Stages 1 and 4	Generates timestamps, hashes, request fingerprints, and source-vintage/build provenance.
`scripts/nisp_raw_writer.R`	Stage 1	Writes source-preserving raw captures and acquisition manifests.
`scripts/nisp_process.R`	Stage 2	Performs source matching, unit handling, analytical grids, source layering, and reconciliation.
`scripts/nisp_semantic.R`	Stage 3	Constructs semantic series, evaluates documented formulas, and runs common semantic QA.
`scripts/nisp_public.R`	Stage 4	Validates the public data specification and writes public data, metadata, deterministic serialization, and build provenance.
`scripts/Public-Export.R`	Stage 4 entry point	Provides the command-line entry point for the Stage 4 export.

Note: Module filenames retain historical `nisp_` identifiers where exact repository reproducibility requires them. These are internal implementation names and do not alter the public EARL identity. The Layer column reports each module's principal architectural use; shared helpers can support more than one stage.

Source: EARL methodology, Appendix J; repository implementation documented for this methodology version.

J.3 Six dataset pipelines

Table J.3 maps the six datasets to their Stage 1-3 drivers and processed outputs.

Table J.3 — Six dataset pipelines

Dataset	Stage 1	Stage 2	Stage 3	Stage 3 output
National annual	`NationalIncome-Annual-RawCapture.R`	`NationalIncome-Annual-Process.R`	`NationalIncome-Annual-Semantic.R`	`data/processed/national_annual_processed.csv`
National quarterly	`NationalIncome-Quarterly-RawCapture.R`	`NationalIncome-Quarterly-Process.R`	`NationalIncome-Quarterly-Semantic.R`	`data/processed/national_quarterly_processed.csv`
Corporate annual	`CorpIncome-Annual-RawCapture.R`	`CorpIncome-Annual-Process.R`	`CorpIncome-Annual-Semantic.R`	`data/processed/corporate_annual_processed.csv`
Corporate quarterly	`CorpIncome-Quarterly-RawCapture.R`	`CorpIncome-Quarterly-Process.R`	`CorpIncome-Quarterly-Semantic.R`	`data/processed/corporate_quarterly_processed.csv`
Industry annual	`Industry-National-Annual-RawCapture.R`	`Industry-National-Annual-Process.R`	`Industry-National-Annual-Semantic.R`	`data/processed/industry_annual_processed.csv`
Industry quarterly	`Industry-National-Quarterly-RawCapture.R`	`Industry-National-Quarterly-Process.R`	`Industry-National-Quarterly-Semantic.R`	`data/processed/industry_quarterly_processed.csv`

Note: Stage 1, Stage 2, and Stage 3 driver filenames are shown relative to `scripts/`. Stage 3 semantic drivers invoke the corresponding Stage 2 processing functions internally; Stage 2 remains a conceptually distinct layer even though its normal output is passed in memory rather than stored as a separate public artifact. Stage 3 output paths are repository-relative.

Source: EARL methodology, Appendix J; repository implementation documented for this methodology version.

Stage 3 semantic drivers call their Stage 2 processing functions internally. Stage 2 remains a conceptually distinct layer even though its normal output is passed in memory rather than stored as a separate public artifact.

J.4 Public data package and publication copies

Repository-root `data/public/` contains the Stage 4 public data package. Publication systems can copy these files into their own build or download directories, but those copies do not redefine the economic data or metadata.

J.5 Rebuild sequence

A full data refresh is conceptually:

1. run the six `*-RawCapture.R` scripts;
2. run the six `*-Semantic.R` drivers, each invoking Stage 2 internally;
3. run `Rscript scripts/Public-Export.R`;

4. run the applicable validation;
5. publish only after the relevant validation gates pass.

A same-vintage reproducibility exercise can instead reprocess retained Stage 1 raw evidence without calling the live BEA API. Reacquisition and reprocessing answer different questions: the former tests current source state; the latter tests implementation reproducibility.

J.6 Documented public file inventory

Table J.4 lists the 18 files in the public data package.

Table J.4 — Complete Stage 4 public package inventory

File	Scope	Purpose
Panel A. Observation data		
`nis-annual.csv`	National annual	Public observation data for NIS-A.
`nis-quarterly.csv`	National quarterly	Public observation data for NIS-Q.
`cis-annual.csv`	Corporate annual	Public observation data for CIS-A.
`cis-quarterly.csv`	Corporate quarterly	Public observation data for CIS-Q.
`iis-annual.csv`	Industry annual	Public observation data for IIS-A.
`iis-quarterly.csv`	Industry quarterly	Public observation data for IIS-Q.
Panel B. Dataset companions		
`nis-annual-meta.csv`	National annual	Dataset-specific field schema and reference metadata for NIS-A observations.
`nis-quarterly-meta.csv`	National quarterly	Dataset-specific field schema and reference metadata for NIS-Q observations.
`cis-annual-meta.csv`	Corporate annual	Dataset-specific field schema and reference metadata for CIS-A observations.
`cis-quarterly-meta.csv`	Corporate quarterly	Dataset-specific field schema and reference metadata for CIS-Q observations.
`iis-annual-meta.csv`	Industry annual	Dataset-specific field schema and reference metadata for IIS-A observations.
`iis-quarterly-meta.csv`	Industry quarterly	Dataset-specific field schema and reference metadata for IIS-Q observations.
Panel C. Shared metadata		
`public-schema.csv`	Observation schema	Common field definitions and allowed values for the six observation files.
`series-metadata.csv`	Series registry	Series definitions, formulas, source mappings, measurement basis, availability, and public role.
`entity-metadata.csv`	Entity registry	Entity labels, types, hierarchy relationships, and source codes.
`statement-layout.csv`	Statement presentation	Statement sections, row order, indentation, sign treatment, and display policy.
`source-vintages.csv`	Source provenance	Official source-vintage and release identity with supporting date evidence.
`build-metadata.csv`	Build provenance	Build identity, file statistics and hashes, contract hashes, acquisitions, and source vintages.

Note: All filenames are relative to `data/public/`. The Stage 4 public data package contains exactly 18 CSV files: six observation files, six dataset-specific schema companions, and six shared package-metadata files. Repository-root `data/public/` contains the published public data package. Publication or website copies inherit these files and do not redefine the economic data or metadata. Full file hashes are retained and verified in the build/evidence layer rather than printed here.

Source: EARL methodology, Appendix J; documented Stage 4 public data package.

J.7 Reproducibility checklist

To audit a published observation: identify its dataset/entity/series/period/value/source type; resolve the series in `series-metadata.csv`; resolve the entity in `entity-metadata.csv`; obtain the build and hashes from `build-metadata.csv`; trace current direct-source evidence through the relevant `data/raw/<dataset>/manifest.csv`; use Appendix F for historical evidence; use Appendix C for calculated formulas; verify the applicable implementation at the recorded Git revision; and consult validation/release evidence.