

Remapping the Input–Output Accounts across the SIC–NAICS Break

What the Bureau of Economic Analysis published, what it told Shaikh privately,
and the choices a Shaikh–Tonak reconstruction is obliged to make

Real Econ Data (RED) — comparative input–output layer
Prepared for Paul Cooney

11 August 2026

Abstract

The productive/unproductive partition of the Shaikh–Tonak apparatus is defined on an industry classification, and that classification breaks in 1997. This report sets out, first, what the Bureau of Economic Analysis (BEA) itself published about carrying its industry accounts from the 1987 Standard Industrial Classification (SIC) onto the 1997 North American Industry Classification System (NAICS), together with the limits BEA stated in its own words; second, what BEA staff told Anwar Shaikh directly in private correspondence, which covers ground the published methodology does not; and third, the concrete choices a reconstruction of the book’s input–output uplift factor has had to make, each recorded as a fork with its evidence and its measured consequence. Two results are new to this report. The BEA conversion-weight workbook that the project holds but had never opened does carry the literal conversion weights in machine-readable form: 61 SIC industries by 61 NAICS industries, eight variable-specific matrices, with column coefficients summing to unity to machine precision — and its weights are exactly constant from 1947 through 1977, confirming BEA’s stated practice by direct measurement rather than by citation. The same workbook’s own read-me states that the currently published historical GDP-by-industry accounts rest on a *different* conversion methodology from the one it documents. The report does not claim the remapping is settled. It sets out the choices, the evidence for each, and three questions on which an input–output specialist’s judgement would change what the reconstruction does next.

Contents

1	The problem	2
1.1	Where the classification enters the measurement	2
1.2	Why the 1997 break is not an ordinary revision	2
1.3	The chain of links a book-period series must traverse	3
2	What BEA published about the conversion	4
2.1	The six-step procedure	4
2.2	The limits BEA states about its own method	5
2.3	The 1992 benchmark: the SIC side of the anchor	6
2.4	What BEA told Shaikh directly	6

3	The reconstruction and its forks	8
3.1	What is being reconstructed	8
3.2	The fork table	8
3.3	The bias decomposition	8
3.4	The structure of the published uncertainty	9
3.5	A measurement made for this report: the conversion weights are machine-readable	10
3.6	A second state change: the SIC-side concordance is now readable	11
4	Open questions for Cooney	12
5	Note on citations and the quote record	13
	Bibliography	14

Conventions. Every empirical statement carries a footnote naming the document and the locus within it. Quotations are reproduced only from passages verified against the source at the time of writing; each is registered in the project’s quote ledger, and quotations newly verified for this report are listed with their identifiers in §6. In quotations taken from electronic mail, hard line breaks in the source body are rendered as spaces; wording, spelling and punctuation are otherwise unaltered, including one misspelling noted in place. Figures described as *re-counted* were recomputed from the underlying files on 11 August 2026 rather than carried from an earlier note.

1 The problem

1.1 Where the classification enters the measurement

In the Shaikh–Tonak framework the aggregates of interest — the Marxian gross final product, the productive share of value added, the rate of surplus value — are not read off a national account. They are constructed by partitioning industries into productive and unproductive activity and then reallocating trade margins, royalties and components of final demand across that partition. The partition is therefore a function defined *on the industry classification*. It is not a function of the underlying economy that happens to be reported by industry; change the classification and the domain of the function changes.

Two objects in the book’s apparatus depend on this directly. The first is the input–output uplift factor, the ratio of the reallocated Marxian gross final product to the plain partition sum of value added. The second is the productive/unproductive filter itself, a categorical assignment of each sector to one side of the line. The uplift factor was last computable, on the classification the book used, for 1989. Carrying either object past 1997 requires a decision about what a 1987-SIC sector *is* in 1997-NAICS terms — and, if a magnitude and not merely a label is to cross, how much of it goes where.

1.2 Why the 1997 break is not an ordinary revision

BEA’s own account is unambiguous that no NAICS-basis source data exist before 1997, and that this absence — not a preference for one system over another — forced a backcast rather than a re-estimation:

“Perhaps an even greater challenge, given the lack of NAICS source data before 1997, was the need to convert historical industry time series data from SIC to NAICS.”¹

NAICS was formally adopted in 1997 and first implemented in the 1997 economic censuses; most industry economic programmes were fully converted only by 2004, seven years after the reference year, and different agencies converted on different schedules using different methods.² The methodology BEA introduced for 1998 forward could not be applied to earlier years, so the pre-1998 estimates on a NAICS basis are the output of an extrapolation procedure, not of a measurement.³

1.3 The chain of links a book-period series must traverse

It is worth stating the chain explicitly, because a single “SIC to NAICS bridge” is not sufficient for input–output work. The book’s sectoring is a BEA *input–output* classification, which is based on the SIC but is not the SIC. BEA states three differences:

“The I-O industry classification system differs from the SIC system in three major ways. First, the I-O industry system redefines some secondary production of some SIC industries to other industries. Second, the I-O industry classification system includes “special industries” that are not considered to be industries in the SIC system. Third, because of data limitations, the I-O industry system includes three industries—agriculture, construction, and real estate—that are defined on an activity basis rather than an establishment basis.”⁴

The chain is therefore: input–output code → 1987 SIC (the volume’s Appendix A concordance)⁵ → 1997 NAICS (BEA’s conversion apparatus). The first link already carries objects with no SIC-side existence, and it carries them in quantities that matter to a Marxian partition. Redefinitions in the 1992 benchmark moved \$48.1 billion into automotive repair and services and \$45.6 billion into eating and drinking places, while removing \$51.0 billion from wholesale trade and \$82.7 billion from retail trade.⁶ BEA’s own summary of the magnitude is that “redefined auto repair output from the retail trade industry makes up almost 40 percent of the I-O auto repair industry output”.⁷ The special industries — federal and state-and-local government enterprises, general government, and the inventory valuation adjustment — have no SIC counterpart at all.⁸

These are precisely the sectors on which a productive/unproductive line is drawn. The first link in the chain is not a formality.

¹Yuskavage, *Converting Historical Industry Time Series Data from SIC to NAICS* (BEA, 2007), text as extracted in the project knowledge base, line 21. Quote identifier QY-001.

²Ibid., lines 41 and 47.

³Ibid., line 71.

⁴BEA, *Benchmark Input–Output Accounts of the United States, 1992* (1998), methodology text as extracted, line 303. Quote identifier QB-001.

⁵Ibid., line 55: Appendix A “provides a concordance between the codes used in the I-O accounts and the 1987 Standard Industrial Classification (SIC) system”.

⁶Ibid., line 316. Quote identifier QB-004.

⁷Ibid., line 165. Quote identifier QB-002.

⁸Ibid., lines 318–322.

2 What BEA published about the conversion

2.1 The six-step procedure

BEA’s published account of how the 1947–97 SIC industry estimates were carried onto NAICS is a six-step procedure, described in detail for 1987–97 and more briefly for earlier years.⁹ Table 1 summarises it.

Table 1: BEA’s published conversion procedure for 1987–97

Step	Content
1	Build a 1997 benchmark concordance between 1997 NAICS and 1987 SIC from the detailed 1997 benchmark input–output accounts (about 850 private industries at approximately six-digit NAICS), each mapped to both a detailed and a publication-level SIC code.
2	Extrapolate that concordance annually back to 1987 using SIC-matched shipments, sales or receipts series, with employment extrapolators from the BLS establishment programme.
3	Cross-tabulate the annual concordances at publication level into 61×61 conversion matrices whose column coefficients sum to 1.0.
4	Convert the published SIC estimates by the matrix product $C_t^k S_t^k = N_t^k$.
5	Extrapolate the 1997 benchmark NAICS levels back to 1987 by chained ratios of the converted values.
6	Scale each extrapolated value-added component and employment to independent NIPA private-sector control totals.

Two features of the design matter here. First, the conversion is a share allocation, not a relabelling:

“Using the cross-tabulated dollar and employment amounts, column coefficients (which sum to 1.0) were calculated that show the percentage of a publication-level SIC industry’s estimate that should be allocated to a specific publication-level NAICS industry.”¹⁰

Second, a matrix is built *per variable per year*, not once.¹¹ Value added is converted through its three components — compensation, taxes less subsidies, gross operating surplus — separately, because “Research has demonstrated that conversion of the three components separately yields better overall results than converting nominal value added directly.”¹² BEA reports no residual for that comparison, so the improvement is stated but not measured.

Government is outside the apparatus entirely: “The revised SIC estimates were used directly for government enterprises and for general government.”¹³ Farms, government enterprises and general government also take their price and quantity indexes directly from the revised SIC estimates in all periods.¹⁴ A Shaikh–Tonak partition that treats general government as unproductive is therefore, on the NAICS side of the break, working with an object that never crossed the bridge.

⁹Yuskavage (2007), lines 93–100 for the step list; lines 104–190 for the detail.

¹⁰Yuskavage (2007), line 128. Quote identifier QY-005.

¹¹Ibid., line 134.

¹²Ibid., line 176. Quote identifier QY-008.

¹³Ibid., line 176. Quote identifier QY-007.

¹⁴Ibid., line 200.

2.2 The limits BEA states about its own method

BEA is explicit about the failure mode of the alternative it rejected:

“Such single-year static concordances are reliable for a limited number of years before the reference year, but they become increasingly unreliable over time as relationships change among the industries.”¹⁵

The paper then quantifies the point twice. Distributing SIC retail-trade employment, the retail share falls from 62.9 to 60.9 per cent and food services and drinking places rises from 34.7 to 36.7 per cent between 1988 and 1997.¹⁶ Holding 1997 shares fixed and applying them to 1988 “would have been overstated in 1988 and would have shown slower growth over the period 1988-97” for food services, with the opposite bias for retail trade.¹⁷ On nominal gross output for 1987-97, fixed versus variable shares change the average annual growth rate by +3.4 percentage points for computer systems design and related services, +2.2 for administrative and support services, +1.4 for publishing including software, and +0.8 for computer and electronic products.¹⁸

The method also degrades sharply with distance from 1997. Table 2 sets out the three regimes.

Table 2: Period-by-period degradation of the published conversion

Period	Conversion matrices	Variables carried	Industry detail
1987-97	annual, variable shares, per variable	gross output, three value-added components, intermediate inputs, employment	61 publication-level private industries; real estimates for 65 detailed industries
1977-86	annual, variable shares, but on the 1972 SIC basis	value added and employment only	65 industries
1947-76	the 1977 matrix held constant	value added and employment only	22 broad industry groups

BEA’s statement of the constant-matrix decision is direct: “The 1977 conversion matrix for value added was held constant for 1947-76 because of the limited availability of SIC-based source data for extrapolation, especially in the non-manufacturing sector.”¹⁹ And: “For the years 1947-86, only estimates of value added and employment were prepared. Components of value added, gross output, and intermediate inputs were not estimated.”²⁰

The defence offered for the constant period is not an accuracy claim but an error-cancellation argument:

“However, because allocations to more than one detailed NAICS industry from a single SIC industry usually fell within the same higher-level NAICS industry group, errors in the allocation matrix tended to cancel one another at the published industry group level.”²¹

¹⁵Ibid., line 77. Quote identifier QY-002.

¹⁶Ibid., lines 138-143; extracted table of retail employment shares.

¹⁷Ibid., line 145. Quote identifier QY-006.

¹⁸Ibid., line 147; extracted Table 4.

¹⁹Ibid., line 194. Quote identifier QY-009.

²⁰Ibid., line 194. Quote identifier QY-010.

²¹Ibid., line 230. Quote identifier QY-012.

The argument licenses aggregate use and, by its own terms, does not license industry-level use. Any series that reads a single pre-1977 NAICS industry is outside the warrant BEA states for it. This bears directly on work in the 1958–77 benchmark window.

Three further limits are on the record. The evaluation was comparative, not independent: “The converted NAICS estimates were evaluated for reasonableness and consistency primarily by comparison with other related estimates.”²² The non-manufacturing problem was never solved — “Unfortunately, similar data for developing SIC to NAICS conversion factors were not readily available for non-manufacturing industries, which accounted for about 80 percent of private-sector GDP in 1997”²³ — which is why other agencies used fixed 1997 factors, and why their conversions are not the same object as BEA’s. And the deflation carries its own stack of seams: double deflation for 1987–97 at about 130 commodities with linear interpolation between benchmark use tables; single deflation for 1977–86, with NAICS price indexes formed as weighted averages of SIC value-added price relatives using the *row* coefficients of the nominal value-added matrix; and, for 1947–76, SIC value-added price indexes last updated in July 1988, before the 1991 change in real-value-added methodology, resting on fixed 1982 relative price weights.²⁴

Finally, the base object of the whole apparatus is unpublished: “This concordance was based on an unpublished version of the 1997 benchmark I-O use table that was adjusted to incorporate the results of the 2003 NIPA comprehensive revision.”²⁵ Step 1 cannot be reproduced from public data. Only the weights derived from it can be audited — which is the subject of §3.5.

2.3 The 1992 benchmark: the SIC side of the anchor

The last benchmark on the SIC basis is the 1992 volume, and it is the natural level anchor for any reconstruction that wants to stay on the book’s own machinery for as long as possible. It presents 498 industries and commodities at the six-digit level, aggregated to 97 industries in summary form.²⁶ Alongside the traditional (redefined) tables, BEA also published an *alternative* 1992 make and use pair more closely based on the 1987 SIC, whose stated purpose is comparison:

“The primary purpose of the new alternative tables is to facilitate comparisons of industry data from the 1992 benchmark accounts with other SIC-based data, such as the gross domestic product by industry and capital stock data that are prepared by BEA and the employment and other industry data that are prepared by the Bureau of the Census and the Bureau of Labor Statistics.”²⁷

Two benchmark tables therefore exist for the same year on two different industry definitions, and they differ by up to 40 per cent for at least one industry that a Marxian partition must classify. Which pair is used is a declared choice, not a detail.

2.4 What BEA told Shaikh directly

The published methodology covers the GDP-by-industry accounts. It does not cover the capital stock, and it does not cover several matters on which Shaikh’s own reconstruction depended.

²²Ibid., line 224. Quote identifier QY-011.

²³Ibid., line 83. Quote identifier QY-003.

²⁴Ibid., lines 202–220 and 234; the last point is quote identifier QY-013.

²⁵Ibid., line 104. Quote identifier QY-004.

²⁶BEA (1998), lines 173 and 69.

²⁷Ibid., line 155. Quote identifier QB-003.

Eight years of correspondence with BEA staff fill part of that gap, and are the only documentary authority for several of the points below. They are cited here as correspondence, never as BEA publications.

On the SIC-to-NAICS concordance for the capital stock, David Wasshausen’s answer names a spreadsheet tab, not a publication:

“Yes, the capital stock estimates were converted from SIC to NAICS several years ago. If you open any of the nonresidential “detailed” fixed asset spreadsheets (e.g. ...detailnonres_stk1.xls), you will find the first tab lists exactly which 1997 NAICS codes are in the capital stock industries we publish. Our industry classifications should be consistent with BEA’s GDP by industry statistics.”²⁸

Three limits follow from that one answer. The mapping is on 1997 NAICS only. It is vintage-bound: it travels inside a data file, so superseding the vintage supersedes the mapping. And nowhere in the corpus is there a statement that the mapping is stable across vintages. The cross-account consistency claim is also modal — “should be” — not a guarantee; and Shaikh had an empirical symptom pointing the other way, a ratio of gross operating surplus to current-cost capital stock in construction “over 90%, far higher than in any other major industry”.²⁹

Two further statements bear on any attempt to reconstruct earlier vintages. The pre-geometric gross stock capability was lost rather than discontinued: the old programs “were written in Fortran which we no longer use and were phased out over 10 years ago. Nobody working on capital stocks back then is still working for BEA.”³⁰ And the documentation simply did not exist: “Unfortunately, I don’t think we’ve done a very good job documenting this type of information in the fixed assets methodology paper.”³¹

Most important for the present question: *no SIC-to-NAICS backcast method for the capital stock appears anywhere in the correspondence*. Nothing describes how pre-conversion SIC-basis capital-stock series were carried onto a NAICS basis, whether industries were split or merged, or where the discontinuities fall.³² Reading the published GDP-by-industry six-step method onto the capital stock would be an unlicensed transfer between two accounts that BEA never linked in writing.

The correspondence also settles two questions about rebuilding requirements matrices that any input–output replication meets immediately. The commodity perimeter must retain scrap, used and secondhand goods and noncomparable imports plus the rest-of-world adjustment; excising them is the documented cause of off-diagonal residuals of 0.113 and 0.09 against BEA’s published total requirements.³³ And BEA applies no adjustment at all: “Currently no adjustments are made for the treatment of scraps or imports.”³⁴ That contradicts the scrap adjustment described in BEA’s own input–output manual, so practice and manual must be distinguished when reproducing published tables.

²⁸David Wasshausen to Anwar Shaikh, 2008; quote Q13 in the project quote ledger, cite key **wasshausen2008a**. The standalone body of this message is missing from the archives; the text survives verbatim inside two later message bodies which agree with one another, so it is second-generation evidence and is described as such wherever cited. The misspelling “classficiations” is in the source.

²⁹Anwar Shaikh to BEA, 2008; quote Q18, cite key **shaikh2008a**.

³⁰David Wasshausen to Anwar Shaikh, 2008; quote Q17, cite key **wasshausen2008c**.

³¹David Wasshausen to Anwar Shaikh, 2010; quote Q21, cite key **wasshausen2010b**.

³²Project note mapping the correspondence findings onto artefacts, finding W-04 and open item O-4 (11 August 2026).

³³Gabriel Medeiros to Shaikh’s group, 2012; quotes Q55 (**medeiros2012e**) and Q53 (**keil2012f**).

³⁴Gabriel Medeiros, 2012; quote Q56, cite key **medeiros2012f**.

3 The reconstruction and its forks

3.1 What is being reconstructed

The object is the input–output uplift factor $k_{IO}(t)$, the ratio of the Marxian gross final product to the plain partition sum of value added, which carries the trade- and finance-margin and royalty reallocation into productive final demand. The book’s apparatus yields a single value for 1989; that constant, 1.5714, had been carried forward frozen. The reconstruction replaces it with a benchmark-anchored annual series for 1990–2024 with an uncertainty band and a per-year flag saying which machinery produced the figure.³⁵

Table 3: The reconstructed uplift factor, selected years

Year	Central	Low	High	Machinery
1990	1.5561	1.5162	1.6200	SIC bridge (illustrative)
1992	1.5695	1.5293	1.6340	native 1992 SIC benchmark (level anchor)
1996	1.5713	1.5310	1.6359	SIC bridge (illustrative)
1997	1.5731	1.4042	1.6978	NAICS benchmark input–output
2007	1.6089	1.4183	1.7362	NAICS benchmark input–output
2017	1.7290	1.4720	1.8660	NAICS benchmark input–output
2024	1.7311	1.4710	1.8682	annual 71-order supply–use

The construction is a classic splice: level from the benchmark, shape from the annual series. The level anchor is the native 1992 SIC benchmark, reproduced by running the book’s own 85-order apparatus on the last SIC benchmark; the shape is the NAICS-era 71-order annual series with margin mass read off the supply tables and royalties restored; the 1990–96 arm is a real-data SIC bridge scaled by $S_b = 1.5696/1.3997 = 1.1213$ to hit the anchor exactly; and the NAICS arm carries $S_c = 1.5731/1.3533 = 1.1624$, continuous at 1997.³⁶

3.2 The fork table

Every consequential decision in the remapping is recorded in the accompanying fork table, which carries 35 rows under the headings `fork_id`, `decision_point`, `option_chosen`, `alternatives`, `evidence_locus`, `consequence_estimate`. The rows fall into four groups: eleven forks in the conversion apparatus itself (F-01 to F-11), six on the input–output side of the SIC anchor (F-12 to F-17), three on matrix construction and perimeter inherited from the BEA correspondence (F-18 to F-20), and fifteen in the uplift reconstruction and the measurement of the weights (F-21 to F-35). Table 4 lists the forks whose consequences are largest or most easily overlooked.

3.3 The bias decomposition

The reconstruction met a level gap in the NAICS era: an uplift near 1.40 computed natively on NAICS data against the 1.57 the book’s machinery gives on the SIC benchmark. The decomposition of that gap is approximately 72 per cent omitted royalties, 23 per cent classification, 7 per cent aggregation, and about 0 per cent genuine change between 1992 and 1997.³⁷

Three things follow, and it is worth separating them.

³⁵Series provenance note for the uplift reconstruction (7 July 2026), section “What this series is”; series file re-counted 11 August 2026: 35 rows, columns `year`, `central`, `lo`, `hi`, `machinery_flag`.

³⁶Series provenance note, construction steps 1–4.

³⁷Series provenance note, bias decomposition; the same figures are carried as claim K-07 in the project’s SIC–NAICS bridge notes (11 August 2026).

Table 4: Selected forks and their measured consequences

ID	Decision	Consequence of the alternative
F-01	Membership map or conversion weights for crossing the break	Retail value added allocated to NAICS retail falls 0.7656 to 0.7511 between 1987 and 1997; food services rises 0.1893 to 0.1964; management of companies rises 0.0370 to 0.0428
F-02	Fixed 1997 shares or annual shares	Up to +3.4 percentage points on a decade’s average annual gross-output growth (computer systems design)
F-03	Pre-1977 industry detail	Weights are literally constant 1947–77; only an aggregate error-cancellation warrant exists
F-06	Gross output with or without auxiliaries	Under one basis 1 of 61 SIC industries sends mass to NAICS 55; under the other, 59 of 61, mean 2.0 per cent and maximum 11.6 per cent
F-13	Traditional or alternative 1992 benchmark tables	Auto repair differs by almost 40 per cent of the industry’s output
F-22	Royalty restoration	About 72 per cent of the NAICS-era level gap
F-29	What the published band covers	The band contains no conversion-mapping term at all
F-34	Business services	SIC 73 splits across fourteen NAICS destinations, none a majority

1. **None of the gap is an economic finding.** The largest term is a correctable omission — royalty and rental flows that the book reallocates and the uncorrected NAICS-era computation had dropped. Restoring them, at the 1992 benchmark’s own share of the gross final product, is a restoration of the book’s method, not an adjustment to it.
2. **Classification is the second-largest term, and it is measured, not assumed.** At about 23 per cent of the gap it is the only quantitative indication anywhere in the project of the size of the classification effect. That is an indication of magnitude, not a bound: it is a decomposition estimate for one level comparison, and it bounds nothing directly.
3. **The residual for genuine change is approximately zero.** This is the most useful diagnostic in the decomposition. If a large real change had been found between two benchmarks five years apart, that would have been evidence that the reconstruction was measuring something other than what it claimed.

3.4 The structure of the published uncertainty

The band on the reconstructed series is not a confidence interval and does not pretend to be. It combines two axes multiplicatively: a level band from the anchor (denominator vintage and royalty scaling) and a perimeter band from two independent crosswalk routes — a narrow partition and an all-productive partition.³⁸ Three properties of it must be stated whenever it is cited.

First, the perimeter arm exists only where NAICS data exist, that is from 1997. The band therefore steps *wider* at the 1996/97 seam, which is the opposite of the visual impression a reader usually takes from a widening band — it is not that the data get worse going forward, but that a second axis of uncertainty becomes measurable. The provenance note states this explicitly.

³⁸Series provenance note, section “Band”; the two routes and the reason for using their spread are set out in the underlying method note, sections 1.3 and 5.

Second, the 1990–96 years carry level-only uncertainty, honestly, because there is no NAICS partition choice to make in them.

Third, and most important for a specialist reader: **the band contains no conversion-mapping error term anywhere.** A statement that the band “covers the classification break” would be false. The honest statement is that the level uncertainty is anchored on the SIC side, the perimeter uncertainty is a NAICS-era partition choice, and the conversion-mapping error is unbanded.³⁹

There is a fourth limit that is not a property of the band but of the period. BEA’s error-cancellation argument licenses group-level use of pre-1977 material and nothing finer. Anything the reconstruction says about a single industry before 1977 is outside BEA’s stated warrant, and is labelled as such.

3.5 A measurement made for this report: the conversion weights are machine-readable

The project holds BEA’s own supplement to the conversion paper — a workbook of conversion weights retrieved with its checksum in July 2026 — but until now nobody had opened it, and the project’s own notes recorded, correctly, that the presence of the literal cell weights in machine-readable form was an unverified manifest claim rather than a data claim.⁴⁰ That check was run for this report on 11 August 2026. The result closes the gap.

The workbook contains a read-me and eight weight sheets in long format, with columns for SIC code, SIC description, NAICS code, NAICS description, year and weight. Its own statement of what the weights are matches the published method exactly: “Weights are sorted by year in ascending order. For any given year, the sum of the weights for each particular SIC code is equal to 1. Any NAICS that does not appear within the conversion table for a particular SIC code can be assumed to have a weight of 0.”⁴¹ Table 5 reports what is in it.

Table 5: Measured contents of BEA’s conversion-weight workbook (11 August 2026)

Sheet	Rows	Years	SIC codes	NAICS codes	Max. deviation of column sums from 1
1947–1987	24,518	1947–1987	61	61	1.2×10^{-15}
GO	1,782	1987–1997	61	61	4.4×10^{-16}
GO AUX	6,589	1987–1997	61	61	5.6×10^{-16}
II	6,578	1987–1997	60	61	4.4×10^{-16}
VA	6,578	1987–1997	61	61	4.4×10^{-16}
COMP	6,578	1987–1997	61	61	5.6×10^{-16}
TXPlxSUB	5,907	1987–1997	60	61	5.6×10^{-16}
GOS	6,556	1987–1997	60	61	6.7×10^{-16}

Five findings follow, each of which changes what the project can say.

1. **The weights exist and are exact.** Column coefficients sum to unity to machine precision across all 2,501 year-by-SIC columns of the historical sheet. The published normalisation is not merely asserted; it holds in the file.
2. **The pre-1977 constancy is directly confirmed.** On the 1947–1987 sheet the weight vector is *identical* across 1947–1977 — maximum absolute difference against 1947 of exactly

³⁹Project bridge notes, derived claim K-10 and operating rule R-6.

⁴⁰Project bridge notes, “What remains unverifiable”, item 2, and the concordance layer index, section C caveat: “nobody has yet opened ... and confirmed the literal 61×61 cell weights are machine-readable”.

⁴¹BEA, *SIC to NAICS Conversion Tables* workbook, sheet ReadMe, Notes. Quote identifier QX-002.

0.0 over 598 non-zero cells — and first moves in 1978 (0.0736), reaching 0.1596 by 1987. BEA’s statement that the 1977 matrix was held constant for 1947–76 is confirmed by measurement, and the constancy in fact extends through 1977 itself. This is now a verified fact about the data rather than a citation.

3. **The matrices are sparse.** 598 of a possible $61 \times 61 = 3,721$ cells are non-zero, both in the constant historical matrix and in the 1997 value-added matrix. Off-diagonal mass is concentrated in a small number of industries, which is why aggregating to a common core loses less than a reader might fear — and why the surviving fractional splits, where they occur, are large.
4. **There are eight matrices, not one, and they disagree in scope.** The choice of which variable’s matrix converts a Marxian value-added partition is a real fork (F-07). Private households (SIC 88) appears in the value-added, gross-output and compensation sheets but is absent from intermediate inputs, taxes less subsidies and gross operating surplus, so component-wise conversion and direct value-added conversion cannot agree for that industry by construction.
5. **The auxiliary treatment is a switch with real mass behind it.** The workbook ships gross-output weights both without auxiliaries (“GO”) and with them (“GO AUX”). In 1997, under the first, exactly one of 61 SIC industries sends any mass to NAICS 55, management of companies and enterprises; under the second, 59 of 61 do, with a mean share of 2.0 per cent and a maximum of 11.6 per cent for holding and other investment offices. Average fan-out rises from 2.66 to 9.82 NAICS destinations per SIC industry. Since NAICS 55 collects the central administrative offices that the SIC placed inside the industry served, this switch moves mass across a productive/unproductive line that a Marxian partition draws.

There is one more sentence in the read-me that a vintage-aware reader should not pass over: “Note that the currently published historical GDP-by-industry accounts rely on a different conversion methodology.”⁴² The weights the project holds are the apparatus documented in the 2007 paper. They are not, on BEA’s own statement, the apparatus behind the historical GDP-by-industry series a user downloads today. Any comparison between a series built on these weights and a current BEA vintage is therefore a comparison across two conversion methodologies, and the difference between them has not been measured here.

3.6 A second state change: the SIC-side concordance is now readable

The volume’s Appendix A — the input–output code to 1987 SIC concordance, the first link in the chain of §1.3 — had no validated machine-readable representation in the project as recently as this month’s bridge notes, which called it the layer’s single largest open gap. That was rechecked at the time of writing. The earlier partial and unverified concordance tables have been withdrawn and replaced by six page-level extractions carrying approximately 651 data rows, with one concordance row per page confirmed against the page image before writing. Three input–output codes remain marked as not captured, two between 14.1801 and 14.1900 and one in industry 38.⁴³ The link is therefore usable with a named residual rather than unavailable — a state change worth recording because the bridge notes, correctly at their time of writing, said otherwise.

⁴²BEA, *SIC to NAICS Conversion Tables* workbook, sheet ReadMe, purpose paragraph. Quote identifier QX-001.

⁴³Quality-hold note for the 1992 benchmark extraction, repair record dated 11 August 2026; six page-level tables re-counted on the same day.

4 Open questions for Cooney

The three questions below are the ones on which an input–output specialist’s judgement would change what the reconstruction does next. They are put as questions because the project does not consider the answers settled.

Q1. What is the right basis for fractional splits where a Marxian partition cuts inside an SIC industry?

The measurement in §3.5 establishes that the weights exist and are exact, but not that they are the *right* weights for this purpose. Two problems remain.

First, BEA’s weights are computed on the variable being converted — value added, compensation, gross operating surplus and so on — while a productive/unproductive partition is defined on activity. Where these coincide the choice is immaterial; where they do not, the fork is live. The sharpest case measured here is business services. SIC 73 splits, on value-added weights, into fourteen NAICS destinations: administrative and support services 0.4092 in 1987 and 0.4026 in 1997; other professional, scientific and technical services 0.2234 then 0.1814; computer systems design 0.1483 then 0.1949; publishing 0.0432 then 0.0823. In the constant pre-1978 matrix the same industry splits 0.3511 professional, 0.3453 administrative, 0.0996 computer systems design. No destination holds a majority in any year. A membership map cannot represent this at all; a weighted map represents it, but only under one weighting.

Second, the underlying method note recommended a secondary sensitivity band from the Census bridge’s alternative weightings — establishments, sales, employment, payroll — and that band was specified but never built. The published band is the two-route perimeter spread only.

The question: for the specific sectors where the Shaikh–Tonak line runs inside an SIC industry — business services, real estate and rental, auto repair, eating and drinking places inside retail — which weighting basis should govern the split, and is the right treatment a single chosen basis with a stated convention, or a reported range across bases?

Q2. Should the reconstruction now be rebuilt on BEA’s own weights, and what would that change?

Until this report the weights were a manifest entry. They are now a measured object: eight matrices, 61 by 61, 1947–1997, exact to machine precision. The reconstruction’s NAICS-era arm was built through two crosswalk routes of its own construction, with the spread published as the perimeter band; BEA’s weights were used at the 1997 join but not as the conversion apparatus throughout.

That leaves a concrete and falsifiable next step: rebuild the 1987–97 arm using BEA’s own variable-specific weights, and compare. Three outcomes are informative. If the rebuilt arm lands inside the published band, the band is doing its job and the classification term is smaller than the 23 per cent decomposition share suggests. If it lands outside, the band understates classification uncertainty and must be widened with a stated third axis. If the two routes and BEA’s weights all disagree in the same direction, the disagreement is systematic and identifies which industries carry it.

Two complications are already visible. BEA’s read-me states that the currently published historical accounts use a *different* conversion methodology, so a rebuild on these weights reproduces the 2004–06 vintage, not today’s. And the weights stop at 1997: the five later NAICS benchmarks in the project’s concordance are the project’s construction, not BEA’s, for this

purpose.

The question: is the right target a faithful reproduction of the vintage these weights document — with the vintage difference then measured separately — or a mapping onto the current published vintage, accepting that no BEA-published weight set exists for it?

Q3. What is the provenance of the mappings used in the 1958–77 benchmark work, and can they arbitrate the constant-matrix period?

The pre-1977 period is where the published apparatus is weakest and where a specialist’s own working knowledge is most valuable. Three facts frame the question. BEA held the 1977 matrix constant for 1947–76, which this report confirms by measurement, with constancy running through 1977 itself. Only value added and employment exist on a NAICS basis for those years, at 22 broad industry groups. And BEA’s only defence is error cancellation at industry-group level, offered without a residual, a confidence interval or an alternative estimate — so the 1947–76 backcast cannot be error-bounded from anything published.

Work built directly on the 1958, 1963, 1967, 1972 and 1977 benchmark tables sits inside exactly this window. Such work does not need the NAICS backcast at all — it is on the original tables — which is precisely why it can serve as an external check on the backcast rather than a consumer of it. If the industry mappings used to make those benchmarks comparable to one another are recoverable, they constitute an independent, contemporaneous account of how the classification actually moved across those years, against which the constancy assumption can be tested at group level.

The question: what mappings were used to align the 1958–77 benchmark tables to one another, how were the trade, business-services and real-estate boundaries handled, and are those mappings documented well enough to be recovered? A related and narrower question: are there sectors where you judge the constant-matrix assumption to be plainly wrong over 1947–77, and if so which?

Standing gaps behind the questions

For completeness, the items the project records as not resolvable from what it holds: step 1 of the published procedure rests on an unpublished use table and cannot be reproduced from public data; the capital-stock concordance carrier file is vintage-bound and not held; the two 85-order sector orderings in use are not reconciled by anything held; no NAICS vintage after 1997 is BEA-sanctioned for this purpose; and no conversion-mapping error term exists anywhere in the layer.

5 Note on citations and the quote record

Sources are cited through the project’s structured citation registry wherever a key exists. The registry currently holds 82 entries, and the sources most heavily used in this report are not yet among them: the conversion paper, the 1992 benchmark volume, the conversion-weight workbook, and the 59 correspondence keys carried in the quote ledger. Those are marked *pending registration* in the bibliography below rather than being cited informally, and registering them is a prerequisite for any series citing the findings here.

Quotations reproduced above are of two kinds. Correspondence quotations (Q13, Q17, Q18, Q21, Q53, Q55, Q56) were already verified byte-for-byte against the source message bodies and carried in the project’s quote ledger. Quotations from the conversion paper, the 1992 benchmark volume and the workbook read-me were newly located and extracted programmatically

from those sources at the time of writing; they are recorded as candidates QY-001 to QY-013, QB-001 to QB-004, and QX-001 to QX-002, nineteen in all, in the quote-candidate file accompanying this report, in the ledger schema, for promotion into the ledger.

Bibliography

Registered in the project citation registry.

`shaikh_tonak1994` — Shaikh, Anwar, and E. Ahmet Tonak. 1994. *Measuring the Wealth of Nations: The Political Economy of National Accounts*. Cambridge University Press. DOI 10.1017/cbo9780511528330.

`shaikh2016` — Shaikh, Anwar. 2016. *Capitalism: Competition, Conflict, Crises*. Oxford University Press.

`shaikh_tonak1987` — Shaikh, Anwar, and E. Ahmet Tonak. 1987. “The Welfare State and the Myth of the Social Wage.” In Robert Cherry et al. (eds.), *The Imperiled Economy, Book I*. New York: Union for Radical Political Economics, pp. 184–194.

`shaikh_tonak2000` — Shaikh, Anwar, and E. Ahmet Tonak. 2000. “The Rise and Fall of the U.S. Welfare State.” In R. Baiman, H. Boushey and D. Saunders (eds.), *Political Economy and Contemporary Capitalism*, ch. 29, pp. 247–265. Armonk, NY: M.E. Sharpe.

`tonak1984` — Tonak, Ertugrul Ahmet. 1984. “A Conceptualization of State Revenues and Expenditures: U.S., 1952–1980.” PhD dissertation, New School for Social Research.

`mohun2005` — Mohun, Simon. 2005. “On Measuring the Wealth of Nations: The US Economy, 1964–2001.” *Cambridge Journal of Economics* 29(5): 799–815. DOI 10.1093/cje/bei043.

`mohun2014` — Mohun, Simon. 2013 [publ. 2014]. “Unproductive Labor in the U.S. Economy 1964–2010.” *Review of Radical Political Economics* 46(3): 355–379. DOI 10.1177/0486613413506080.

`moos2019` — Moos, Katherine A. 2019. “Neoliberal Redistributive Policy: The US Net Social Wage in the Early Twenty-First Century.” *Review of Radical Political Economics* 51(4): 581–605. DOI 10.1177/0486613419848097.

`tsoulfidis_tsaliki2011` — Tsoulfidis, Lefteris, and Persefoni Tsaliki. 2011. *Classical Competition and Regulating Capital*. University of Macedonia.

`bea_gdp_by_industry` — U.S. Bureau of Economic Analysis, Industry Economic Accounts. *GDP by Industry*. <https://www.bea.gov/data/gdp/gdp-industry>.

`shaikh2008` — Shaikh, Anwar. U.S. industry rate-of-profit panel, 1988–2005 (BEA NAICS reconstruction), as carried in *Capitalism* Appendix 7.2.

Pending registration (used in this report; a registry key is proposed for each, with `container_type` noted).

`yuskavage2007` (*conference paper*) — Yuskavage, Robert E. 2007. “Converting Historical Industry Time Series Data from SIC to NAICS.” Bureau of Economic Analysis, paper P2007-7; presented to the Federal Committee on Statistical Methodology.

`bea1998_benchmark_io_1992` (*government document*) — U.S. Bureau of Economic Analysis. 1998. *Benchmark Input–Output Accounts of the United States, 1992*. Washington, DC: U.S. Government Printing Office. Incorporating “Benchmark Input–Output Accounts for the U.S. Economy, 1992: Make, Use, and Supplementary Tables” (*Survey of Current Business*, November 1997) and the companion requirements-tables article (December 1997).

`bea_sic_naics_conversion_tables` (*dataset*) — U.S. Bureau of Economic Analysis, Industry Economic Accounts. *SIC to NAICS Conversion Tables* (supplement to Yuskavage 2007). Workbook, 2,837,022 bytes, SHA-256 `ad4388b6...f6d0119d`, retrieved 7 July 2026.

`bea_scb_2005_12_conversion_matrices` (*journal*) — U.S. Bureau of Economic Analysis. 2005. GDP-by-industry NAICS conversion matrices. *Survey of Current Business*, December.

lawson_bersani_fahimnader_guo2002 (*journal*) — Lawson, Ann M., Kurt S. Bersani, Mahnaz Fahim-Nader and Jiemin Guo. 2002. “Benchmark Input–Output Accounts of the United States, 1997.” *Survey of Current Business*, December, pp. 19–108.

moyer_planting_kern_kish2004 (*journal*) — Moyer, Brian C., Mark A. Planting, Paul V. Kern and Abigail M. Kish. 2004. “Improved Annual Industry Accounts for 1998–2003.” *Survey of Current Business*, January, pp. 21–57.

yuskavage_pho2004 (*journal*) — Yuskavage, Robert E., and Yvon H. Pho. 2004. “Gross Domestic Product by Industry for 1987–2000.” *Survey of Current Business*, November, pp. 33–53.

yuskavage_fahimnader2005 (*journal*) — Yuskavage, Robert E., and Mahnaz Fahim-Nader. 2005. “Gross Domestic Product by Industry for 1947–86.” *Survey of Current Business*, December, pp. 70–84.

bayard_klimek2003 (*conference paper*) — Bayard, Kimberly, and Shawn Klimek. 2003. “Creating a Historical Bridge for Manufacturing between the Standard Industrial Classification System and the North American Industrial Classification System.” *Proceedings of the Annual Meeting of the American Statistical Association*, August.

shimberg_detlefsen_davie2002 (*conference paper*) — Shimberg, Michael, Ruth E. Detlefsen and William C. Davie Jr. 2002. “Methods Used to Develop Retail and Wholesale Time Series Under the North American Industry Classification System.” American Statistical Association annual meeting.

morisi2003 (*journal*) — Morisi, Teresa. 2003. “Recent Changes in the National Current Employment Statistics Survey.” *Monthly Labor Review*, June.

fort_klimek2016 (*working paper*) — Fort, Teresa C., and Shawn D. Klimek. “The Effects of Industry Classification Changes on U.S. Employment Composition.”

cooney2017 (*journal*) — Cooney, Paul. 2017. “Competencia vs. Monopolio: Un análisis insumo-producto de las tasas de ganancia y markups en la economía de los EE.UU.: 1958–1977.” *Ensayos de Economía* 27(50): 209–235.

Correspondence keys (*correspondence*) — 59 keys carried in the project quote ledger, of which the following are cited above: wasshausen2008a (Q13, a recovered quotation), wasshausen2008c (Q17), wasshausen2010b (Q21), shaikh2008a (Q18), medeiros2012e (Q55), medeiros2012f (Q56), keil2012f (Q53).

Project documents cited as loci (internal analytical records, not publications): the SIC–NAICS bridge notes and concordance layer index (11 August 2026); the note mapping the BEA correspondence findings onto project artefacts (11 August 2026); the uplift-series provenance note (7 July 2026) and its underlying method note (7 July 2026); and the quality-hold record for the 1992 benchmark extraction (repair entry, 11 August 2026).