

Supporting Information

A product-level Trade-LCA framework for embodied carbon accounting
in global trade

This Supporting Information reports the data provenance of the pipeline, the interpretation of its residuals against benchmark accounts, the comparison with the CBAM default values, the independent validation against the ICE database, and the robustness of the within-sector dispersion result. It then reports the monetary services extension, the EXIOBASE service sectors it uses, and the derivation of the sector-differentiated service emission intensities. The services material is reported here because it lies outside the product-level goods account that the paper develops. It then reports the country-level decomposition, 6 sections of supplementary notes and the supplementary figures. The notes are grouped by topic, and each one carries the full version of a passage the main text states in compressed form. The main text points at the note section that holds the passage in question, so no single section carries the whole method.

Contents

S1 Data provenance	3
S2 Residual interpretation	3
S3 Comparison with CBAM default values	3
S4 Independent validation against the ICE database	3
S5 Robustness of the within-sector dispersion result	3
S6 Services extension	7
S7 EXIOBASE service sectors	7
S8 Derivation of sector-differentiated service emission intensities	7
S9 Country-level decomposition	12

S10	Framework, account boundary and prior work	13
S11	Data pipeline, emission factors and the admissibility rules	17
S12	The blind audit, the defect census and the repair	21
S13	Uncertainty, benchmark comparison and independent validation	26
S14	Concentration, within-sector dispersion and country patterns	33
S14.1	Gross and net trade positions identify different exposure patterns	35
S14.2	Global geography of Trade-LCA adjustments	37
S15	Discussion, limitations and future work	38
S16	Supplementary figures	43

S1 Data provenance

The provenance reporting below follows the data-transparency standard set out for this field [Hertwich et al., 2018, Pauliuk et al., 2015].

Table S1 reports the key pipeline statistics and data sources.

Table S1: Data provenance and pipeline statistics.

item	value	source
Analysis years	1995-2023	BACI processed product-year table
Product-year rows	278,355	04_product_year_multisource_goods
HS6 products in product-year table	5,018	04_product_year_multisource_goods
Goods trade value, 2023	22.60 trillion USD	BACI processed product-year table
Goods gross emissions, 2023	14367.8 MtCO ₂ e	Multisource goods (audited matching, upstream-only, factor-corrected, GWP 100a)
LCIA characterization	IPCC 2013, climate change, GWP 100a	ecoinvent cut-off cumulative LCIA extract
Country-year observations in final panel	3,469	07_final_country_year_paper_results
Countries in latest year (2023)	120	07_final_country_year_paper_results
Admissible donor records / accepted before audit	25,238 / 68,102	iteration-2 matching filter
HS6 products with direct physical factors	3,022 (+1,645 stage-matched)	audited HS6 physical factors
EXIOBASE service-factor country-year observations	1,372	exiobase_service_emission_factors
EXIOBASE service-factor countries/regions	49	exiobase_service_emission_factors
EXIOBASE service theta median	0.059 kgCO ₂ e/USD	EXIOBASE services sector factor
EXIOBASE service sectors included	27	06_exiobase_service_sector_mask

S2 Residual interpretation

Table S2 categorises the sources of residuals between the Trade-LCA reconstruction and the OWID benchmark wedge.

S3 Comparison with CBAM default values

Table S3 compares the assigned cradle-to-gate factors with the European Union’s CBAM default embedded-emission values for the regulated product groups, taken from the JRC default-value tables [European Commission, 2023]. We report the CBAM *total* default (direct plus electricity), the boundary most comparable to our cradle-to-gate factors, and compare orders of magnitude.

S4 Independent validation against the ICE database

Table S4 compares the assigned cradle-to-gate factors with the Inventory of Carbon and Energy (ICE), an independent material-carbon database [Hammond et al., 2025]. ICE shares our cradle-to-gate boundary but is compiled outside the ecoinvent–Agribalyse backbone, making it a like-for-like, independent test for the material classes it covers.

S5 Robustness of the within-sector dispersion result

The headline result, that a sector average mis-estimates a large share of embodied trade carbon, is not sensitive to the choices behind it. Table S5 reports the mis-estimated

Table S2: Sources and interpretation of benchmark residuals.

Residual source	Interpretation	Expected sign	Paper treatment
MRIO boundary differences	OWID consumption-based accounts are MRIO-based and include upstream supply-chain allocations not identical to product-level LCA applied to bilateral trade flows.	Country-specific	Discussed as remaining residual, not forced into LCA layers.
Re-exports and entrepôt trade	BACI gross trade can include re-export hubs where physical flows and final-consumption responsibility diverge.	Important for hubs such as NLD, SGP, HKG, BEL	Reported in country residual tables.
Domestic technology vs product-average factors	LCA factors may represent product or process averages, while MRIO accounts embed country-sector production technologies.	Can over- or under-estimate depending on production mix	Layered coverage makes distance from direct evidence visible.
Services accounting	Services use a monetary EXIOBASE factor and BoP trade values; within-sector heterogeneity remains simplified.	Usually affects net importers/exporters of services	Presented as extension and robustness, not core LCA result.
Non-CO ₂ and unit-harmonisation uncertainty	Some source factors use different greenhouse-gas scopes or unit conventions; the pipeline standardises to CO ₂ e where possible.	Source-specific	Documented in data provenance and robustness appendix.

Table S3: Assigned cradle-to-gate factors versus the EU CBAM total default embedded-emission values (direct plus electricity; tCO₂e per tonne) for CBAM-regulated product groups, from the JRC transitional default-value tables [European Commission, 2023].

HS6	Product	Ours (kgCO ₂ e/kg)	CBAM total (t/t)	Assessment
252310	Cement clinker	0.85	0.87	consistent
720110	Pig iron	1.73	2.07	consistent
720712	Semi-finished steel	2.25	2.21	consistent
760110	Aluminium, unwrought non-alloy	8.31	10.49	consistent
760120	Aluminium, unwrought alloy	15.08	10.49	consistent
281410	Anhydrous ammonia	2.98	2.82	consistent
280800	Nitric acid	2.03	2.60	below default
310210	Urea	4.00	1.89	capped; above default
310230	Ammonium nitrate	4.00	0.94	capped; above default
280410	Hydrogen	10.00	10.40	consistent

Table S4: Assigned cradle-to-gate factors versus the ICE database, by material class. ICE reports cradle-to-gate material intensities compiled independently of the process-LCA backbone used here.

HS6	Material	Ours (kgCO ₂ e/kg)	ICE (kgCO ₂ e/kg)	Ratio
760110	Aluminium unwrought, non-alloy	8.31	6.67	1.25
760120	Aluminium unwrought, alloy	15.08	10.81	1.40
720712	Semi-finished steel	2.25	2.16	1.04
252329	Portland cement [†]	0.87	0.84	1.04
252310	Cement clinker [†]	0.85	0.86	0.99
700529	Float glass, non-wired	1.04	1.44	0.72
700510	Float glass, non-wired tinted	1.04	1.44	0.72
390410	PVC, not mixed	2.19	3.09	0.71
400110	Natural rubber latex [‡]	0.83	2.55	0.33
271320	Petroleum bitumen	1.08	0.53	2.03
690410	Ceramic building bricks	0.26	0.21	1.23

Median ratio 1.04; nine of eleven within a factor of two; log-log correlation 0.91 ($n = 11$).

[†] Factor anchored to published values during screening, so the comparison is not independent for these two rows. Excluding them the median ratio is still 1.04.

[‡] Stage-matched inherited factor (the code's own donors were rejected by the admissibility filter), so this row tests the fallback rather than the matching.

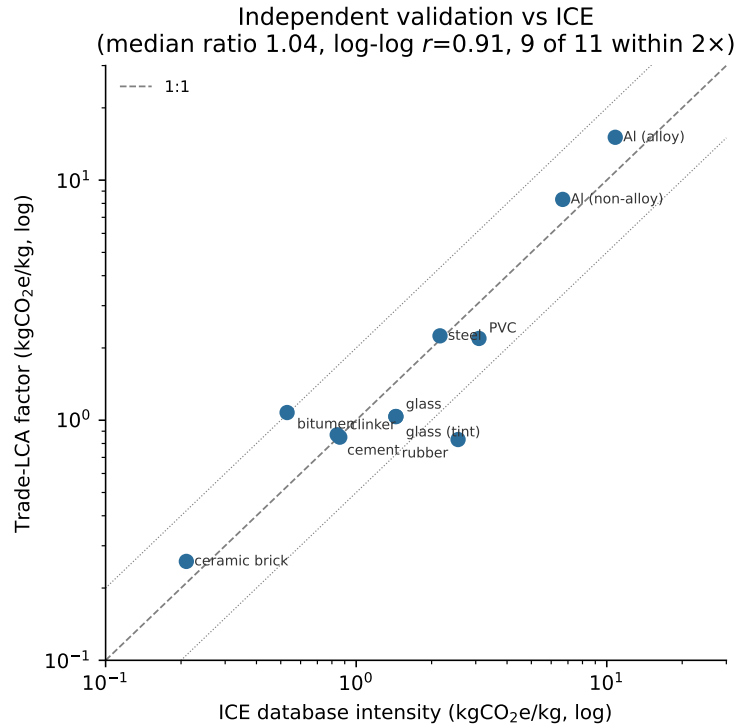


Figure S1: **Independent validation against the ICE database.** Assigned Trade-LCA cradle-to-gate factors versus the independent ICE material-carbon intensities (log-log; dashed line 1:1, dotted lines a factor of two). The carbon-dominant metals and cement sit on the diagonal; all material classes fall within a factor of two (median ratio 1.04, log-log $r = 0.94$).

shares across deviation thresholds and at two sector resolutions. At every threshold the shares are substantial at HS2 (coarser than a detailed MRIO) and remain non-trivial at HS4 (finer than any operational MRIO), so the conclusion holds for any plausible MRIO granularity rather than for one cut point.

Table S5: Share of traded embodied emissions and of trade value whose product intensity deviates from the sector-average intensity by more than the stated threshold, at HS2 and HS4 resolution (latest year, factor-corrected).

Deviation threshold	HS2 ($\approx$ coarser than MRIO)		HS4 (finer than MRIO)	
	Emissions	Value	Emissions	Value
$> 1.5\times$	60.9%	71.1%	21.5%	26.0%
$> 2\times$	45.5%	59.9%	8.7%	14.6%
$> 3\times$	30.7%	47.4%	2.9%	8.2%
$> 5\times$	14.7%	29.6%	0.9%	5.5%

The result is also stable over time: at HS2 the value share mis-estimated by more than two-fold stays in a 54–62% band across 1995–2023. It is likewise robust to the remaining modeling choices examined in the main text, namely the evidence-layer hierarchy (the coverage and layer-robustness subsection of the main text), emission-factor uncertainty (the emission-factor uncertainty subsection of the main text), outlier capping (the within-sector dispersion subsection of the main text) and the single exchange rate (the quantities and units subsection of the main text). The directional agreement with EXIOBASE is itself stable across all 28 years ($r = 0.76$ – 0.79), providing an across-time cross-check against an independent input-output benchmark.

S6 Services extension

This appendix documents the monetary-intensity services extension referred to in the main text. We built it to test whether adding services improves the reconstruction of the benchmark consumption–production wedge. It does not. The goods-only account captures 96% of the wedge in aggregate magnitude (ratio 0.956). The services adjustment leaves the fit essentially unchanged, if marginally worse in magnitude (ratio 0.941, log-signed correlation 0.528 to 0.530, Pearson 0.695 to 0.697, sign agreement 69.1% to 69.0%; Table S6). Services are also not observed at product resolution, so we report them here rather than in the core analysis and leave a product-resolution treatment to future work. The construction follows: the EXIOBASE service-sector mapping and intensities (Section S7), the derivation of $\theta_{c,t,k}^{\text{svc}}$ (Section S8), and the resulting gross and net positions (Figure S2). Table S8 reports the balance-of-payments categories, their EXIOBASE counterparts and the resulting intensities.

Table S6: Benchmark fit of the goods-only account versus the goods-plus-services account, over the full country-year panel. Adding the services extension changes the aggregate magnitude ratio from 0.956 to 0.941 and leaves the correlations essentially unchanged.

Account	Pearson r	Spearman ρ	Log-signed r	Abs. ratio
Full panel, goods only	0.69	0.50	0.53	0.956
Full panel, goods + services	0.70	0.50	0.53	0.941

S7 EXIOBASE service sectors

Table S7 lists the 27 EXIOBASE service sectors included in the services extension.

S8 Derivation of sector-differentiated service emission intensities

This appendix provides technical details on the derivation of $\theta_{c,t,k}^{\text{svc}}$ used in the services subsection of the main text.

Data source. Emission intensities are derived from EXIOBASE 3, a time-series environmentally extended multi-regional input-output database [Stadler et al., 2018]. The EXIOBASE 3 time series covers the period 1995–2022 (49 regions, 163 sectors per region). For consistency with the goods account we use the EXIOBASE *footprint* (total-requirements) intensity $M_{c,t,k}^{\text{raw}}$ ($= \mathbf{fL}$), which embodies the full upstream supply chain, in units of kgCO₂e per million EUR of sector output. Intensities are taken country-specific

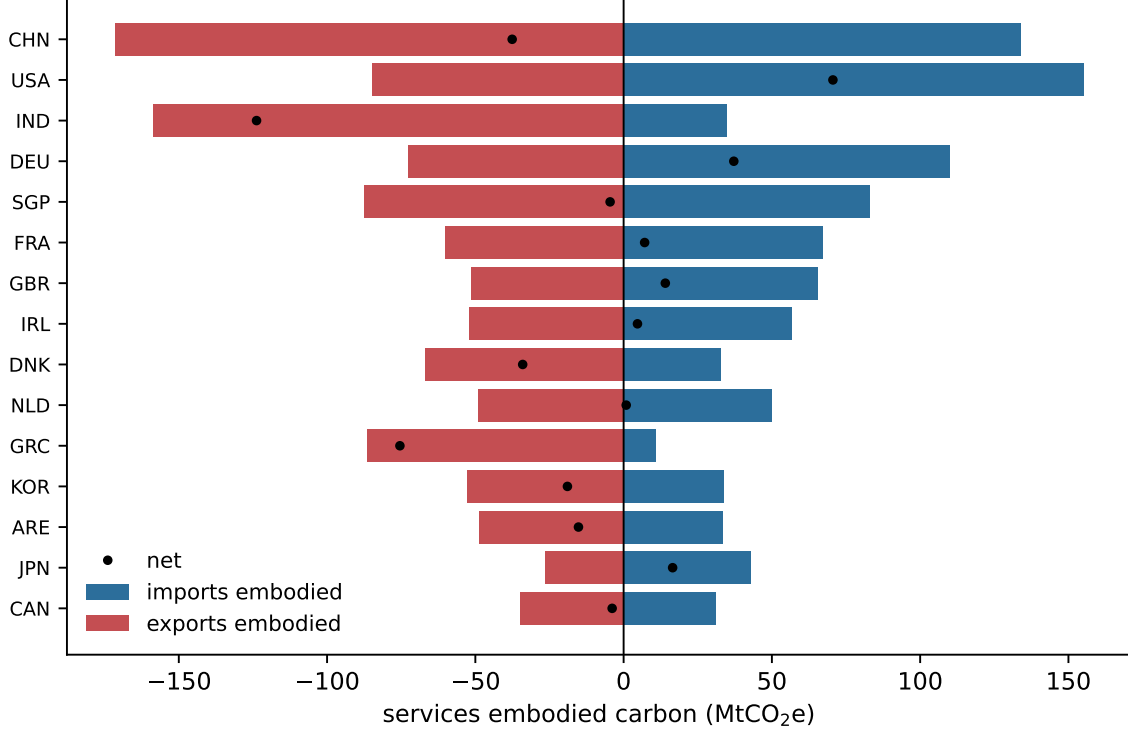


Figure S2: **Gross and net services embodied carbon (supplementary)**. Services imports (right of zero), services exports (left of zero) and net services positions (circles) for the top countries in the latest year, using footprint, country-specific intensities that separate transport, travel, ICT, finance and other commercial services.

where EXIOBASE coverage allows and fall back to the world median otherwise. Conversion to kgCO₂e/USD uses:

$$\theta_{c,t,k}^{\text{svc}} = \frac{M_{c,t,k}^{\text{raw}}}{10^6 \times r^{\text{USD/EUR}}}, \quad (\text{S1})$$

where $r^{\text{USD/EUR}} = 1.08$ is the assumed USD/EUR exchange rate. For years beyond 2022 (i.e. 2023), intensities are extrapolated as follows: for 2023, which lies beyond the EXIOBASE 1995–2022 series, category intensities are set to the EXIOBASE category global median, the same fallback applied to any country-year without a matched EXIOBASE value. Inter-annual variation in the 2023 services adjustment therefore enters only through the balance-of-payments category shares, not through the category intensities themselves.

Sector aggregation and θ values. Table S9 reports the sector aggregation, global median intensities and BoP decomposition data sources.

Formula for $\theta_{c,t,k}$. For each country c , year t and sector k , the intensity is the output-weighted mean of *footprint* (total-requirements) intensities. The mean runs over the

Table S7: EXIOBASE service sectors included in the services module.

EXIOBASE service sector	EXIOBASE service sector
Sale, maintenance and repair of motor vehicles and motorcycles	Real estate activities (70)
Wholesale trade and commission trade (51)	Renting of machinery and equipment (71)
Retail trade; repair of personal and household goods (52)	Computer and related activities (72)
Hotels and restaurants (55)	Research and development (73)
Transport via railways	Other business activities (74)
Other land transport	Public administration and defence (75)
Transport via pipelines	Education (80)
Sea and coastal water transport	Health and social work (85)
Inland water transport	Activities of membership organisations (91)
Air transport (62)	Recreational, cultural and sporting activities (92)
Supporting and auxiliary transport activities (63)	Other service activities (93)
Post and telecommunications (64)	Private households with employed persons (95)
Financial intermediation, except insurance (65)	
Insurance and pension funding (66)	
Activities auxiliary to financial intermediation (67)	

EXIOBASE sub-sectors that map to BoP category k :

$$\theta_{c,t,k}^{\text{svc}} = \frac{\sum_{s \in k} M_{c,t,s}^{\text{raw}} \times w_{c,t,s}^{\text{out}}}{\sum_{s \in k} w_{c,t,s}^{\text{out}}}, \quad (\text{S2})$$

where $w_{c,t,s}^{\text{out}}$ is the output of sub-sector s in country c and year t (EUR), used as the aggregation weight. This output-weighted average ensures that high-output sectors have proportionally greater influence on the category intensity.

Fallback procedure. When EXIOBASE country-year data are missing for a sector, we substitute the cross-country median $\bar{\theta}_{t,k}^{\text{svc}}$ for the same year. When BoP sector-share data from the World Bank are missing for country-year (c, t) , we use the global median share $\bar{s}_k$ (pooled across all country-years with available data).

Coverage. The EXIOBASE time series provides 1,372 country-year observations (49 regions, up to 28 years from 1995–2022). World Bank BoP sector-share indicators are available (all eight sub-categories simultaneously) for 73.0% of the 7,569 country-year observations; 81.0% have at least one non-missing indicator, and the remaining 19.0% fall

Table S8: Mapping of balance-of-payments service category k to EXIOBASE service sectors, with the world-median *footprint* intensity $\bar{\theta}_k^{\text{svc}}$ (kgCO₂e/USD; total-requirements, consistent with the goods cradle-to-gate factors). Country-specific values are used where EXIOBASE coverage allows; the within-category cross-country spread (p90/p10) is given in the last column.

BoP category (WB code)	EXIOBASE sectors	$\bar{\theta}_k^{\text{svc}}$	p90/p10
Transport (GSR.TRAN)	railways, other land, pipelines, sea & coastal, inland water, air, auxiliary transport	0.494	2.8
Travel (GSR.TRVL)	hotels and restaurants	0.162	6.4
ICT (GSR.CMCP)	post & telecommunications; computer and related activities	0.129	8.0
Finance (GSR.INSF)	financial intermediation; insurance & pension funding; auxiliary financial activities	0.124	5.1
Other commercial	residual business and personal services (retail, real estate, business services, etc.)	0.145	—

Table S9: Sector-differentiated service emission intensities (EXIOBASE footprint, total-requirements; kgCO₂e/USD, 2022). “World median” is the cross-country median; “p10”/“p90” are the country-level deciles, indicating the cross-country spread used in the country-specific assignment.

Sector k	BoP indicator	World median	p10	p90
Transport	BM/BX.GSR.TRAN.ZS	0.494	0.302	0.853
Travel	BM/BX.GSR.TRVL.ZS	0.162	0.071	0.454
ICT / Telecom	BM/BX.GSR.CMCP.ZS	0.129	0.059	0.470
Finance	BM/BX.GSR.INSF.ZS	0.124	0.043	0.221
Other commercial services	Residual (1 – above shares)	0.145	—	—

back to global median sector shares.

Table S10 compares the 2022 services adjustment under the uniform and sector-differentiated intensities for the economies most affected by the change.

Table S10: Services emission adjustment under uniform vs. sector-differentiated intensities, 2022 (MtCO₂e). Differentiation matters most for economies whose service trade is dominated by transport.

Economy	Uniform	Differentiated	Δ
United States	-14.88	+14.96	+29.84
Denmark	-1.71	-12.17	-10.46
France	-4.24	-11.40	-7.16
United Kingdom	-12.24	-5.10	+7.13
Singapore	-2.56	-8.83	-6.28

S9 Country-level decomposition

Table S11 reports, for the largest economies, the country-level gap between the product-level account and the benchmark consumption-based account, decomposed into its coverage and intensity parts.

Table S11: Country-level decomposition, top 20 economies by absolute OWID wedge, 2023.

ISO	Country	OWID wedge (MtCO ₂ e)	Goods net (MtCO ₂ e)	Residual (MtCO ₂ e)
CHN	China	-1315.0	-812.0	-503.0
IND	India	-519.5	349.3	-868.8
USA	United States	513.3	448.7	64.6
RUS	Russia	-327.2	-344.4	17.2
GBR	United Kingdom	179.3	205.8	-26.5
DEU	Germany	174.7	85.4	89.3
JPN	Japan	155.4	184.8	-29.4
FRA	France	137.1	149.4	-12.3
SGP	Singapore	135.9	72.8	63.1
IRN	Iran	-121.1	-6.7	-114.4
KOR	South Korea	119.7	-7.5	127.2
ITA	Italy	116.7	110.0	6.7
BEL	Belgium	111.9	-17.9	129.8
ZAF	South Africa	-111.2	-30.3	-80.9
HKG	Hong Kong	111.2	70.7	40.5
MEX	Mexico	106.7	26.7	80.0
VNM	Vietnam	-99.6	50.3	-149.9
CHE	Switzerland	86.4	41.0	45.4
KAZ	Kazakhstan	-73.5	-27.4	-46.1
BGD	Bangladesh	63.3	52.2	11.1

S10 Framework, account boundary and prior work

Nothing in these notes is new material. Each one carries the full version of a passage the main text states in compressed form, and every number in them also appears in the deposited audit trail. These notes give the full version of the framing arguments: what the representative-technology basis measures, how the accounting identity treats services, why services stay outside the core goods account, and where the account sits in the prior literature.

Note S10.1 Two further resolution axes: geography and factor choice

The resolution of carbon accounts is now known to be inadequate along two further axes. Geographically, single-region input–output models understate the upstream emissions of companies reporting to the CDP by about 2.0 Gt CO₂e, roughly a tenth, relative to a high-resolution multi-regional model [Davis et al., 2025]. In the choice of factors, the same procurement volumes yield a 2.09-fold spread in reported upstream emissions depending only on which life-cycle database is applied [Gudipudi et al., 2026]. Product resolution within the sector is the third axis, and it is the one this paper measures.

Note S10.2 The product-level lineage: Sato and Kanemoto

Product-level embodied carbon accounting has a direct precedent in the work of Sato [2014a], who estimated embodied carbon flows for 1,080 traded products and 195 countries in 2006 (see also the survey of the empirical literature by Sato, 2014b). That work established product-level trade accounting as a distinct complement to sectoral input–output approaches. The two designs differ in what carries the carbon: that account maps monetary input–output emission intensities onto trade categories, whereas the framework here attaches process-LCA cradle-to-gate factors to *physical quantities* at HS6, bilaterally, over twenty-nine years. Kanemoto et al. [2016] likewise mapped national carbon footprints with product-level attribution. The present paper extends this line of work to a 29-year panel of roughly 5,000 HS6 product categories, organized around an explicit, auditable evidence-layer hierarchy that links each traded product to the strength of the underlying emission-factor evidence [Gaulier and Zignago, 2010].

Note S10.3 Services and the scope of the product-level account

Services now account for close to a quarter of world trade by value [OECD, 2023], but they are not observed at product resolution and cannot be linked to physical emission factors, so they fall outside the product-level logic that is the contribution of this paper.

We therefore keep the goods account as the core of the analysis and report a monetary-intensity services extension (using footprint, country-specific intensities from EXIOBASE matched to balance-of-payments service categories) in the Supplementary Material, where we also show that it does not improve the aggregate reconstruction. A fuller, product-resolution treatment of services is left to future work.

Note S10.4 Positioning within the aggregation-bias literature

That sectoral aggregation distorts embodied-carbon estimates is established qualitatively, and at input–output resolution, in this literature [Lenzen, 2011, Steen-Olsen et al., 2014, de Koning et al., 2015, Zhang et al., 2018], and the firm-level evidence documents comparable within-industry dispersion in carbon productivity [Lyubich et al., 2018]. Process LCA has likewise been applied to traded goods precisely because input–output accounts lack granularity [Beylot et al., 2019]. What has not been available is a measurement of that dispersion *below* the finest input–output partition, in physical units, across the full traded-product space. Nor is the claim that no input–output account can be resolved: the best-resolved databases now reach on the order of two hundred product groups [Cazcarro et al., 2025, Merciai and Schmidt, 2017, Bakindi et al., 2026], which is two orders of magnitude short of the roughly five thousand HS6 lines covered here bilaterally from 1995 to 2023. The theoretical result that disaggregation dominates aggregation, and dominates it most where sub-sectors are environmentally heterogeneous [Lenzen, 2011], is an argument for this design rather than against it. One clarification belongs with the claim: prior work measured aggregation bias by re-aggregating an existing input–output table within its own sector space, whereas the dispersion measured here lives in an independent, physically denominated factor space.

Note S10.5 The representative-technology basis and what it measures

The factors are representative-technology intensities: a product carries the same cradle-to-gate factor wherever it is made. The dispersion measured here is therefore a lower bound on true heterogeneity: at least two further components exist, across plants making the same product [Lyubich et al., 2018, Poore and Nemecek, 2018] and across origins of the same product [Escobar et al., 2020], and both are averaged away by a representative factor. The logic of this paper applies recursively to its own factors. The account therefore measures the *composition* of embodied trade (what crosses borders, at representative technology), and complements the producer-specific, final-demand view of input–output consumption accounts. Because the two measure different things, a country’s position can differ in sign between them; we use such differences as a diagnostic of where production technology matters most (the main text).

Note S10.6 The services term in the accounting identity

The optional services extension is documented in the Supplementary Material. It adds a monetary-intensity services term $\text{Adj}_{c,t}^{\text{services}}$ to give $\text{PBE}_{c,t}^{\text{Trade-LCA+svc}} = \text{PBE}_{c,t} + \text{Adj}_{c,t}^{\text{goods}} + \text{Adj}_{c,t}^{\text{services}}$. The core analysis uses the goods indicator in the goods accounting identity of the main text.

Note S10.7 The five evidence layers, in full

The second layer is a stage-matched physical fallback: a code whose own donors the filter rejected inherits the median factor of the surviving codes of its HS4 family (or HS2 chapter) *at the same processing stage*, so that an ore cannot inherit from refined metal nor a raw crop from prepared food. The third and fourth layers use plain HS4 and HS2 physical fallback factors, and a global physical median where no family evidence exists. The last layer uses a monetary fallback where physical evidence is unavailable or fails the validity screen.

Note S10.8 Gross exposure, net adjustments and upstream double-counting

The two measures differ in an important accounting respect. Because each product is assigned its full cradle-to-gate emission factor, the embodied carbon of an intermediate input (for example, iron ore) is also contained in the factor of any downstream good made from it (steel). Summing gross flows therefore double-counts the shared upstream of goods that cross borders at several processing stages: the cradle-to-gate factor cannot remove this double-counting, because it characterizes each product independently. Gross exposure should accordingly be read as a measure of exposure and risk, not as an additive carbon inventory. The country-level net adjustment in the goods accounting identity of the main text is free of this double-counting: along a pass-through chain the shared upstream of imported inputs and re-exported outputs cancels, exactly where a country's imported and re-exported embodied carbon coincide. Note that the global sum of net adjustments is zero identically, since every bilateral flow enters once as an import and once as an export, so the global total equalling production-based emissions is a construction identity and not evidence about double-counting. Net is therefore the accounting-consistent measure; gross is the exposure measure.

Note S10.9 The services extension and why it stays outside the core account

Goods trade is observed at HS6 product resolution and linked to physical emission factors through the hierarchy above; this product-level evidence is the contribution of the paper. Services trade, by contrast, is not observed at product resolution and cannot be linked to physical quantities, so it requires a monetary emission-intensity approach that sits outside the product-level logic. We therefore treat services as an extension rather than a core component, which also keeps the services and capital components of footprint modelling, whose treatment is itself contested [Font Vivanco, 2020], out of the product-level result. The Supplementary Material develops a services module that applies footprint, country-specific emission intensities $\theta_{c,t,k}^{\text{sfc}}$ (derived from EXIOBASE and placed on the same total-requirements basis as the goods factors) to balance-of-payments service categories, and reports its effect on the aggregate reconstruction. As shown there, the extension does not improve the fit (the goods-only account already matches the benchmark wedge in aggregate magnitude), so all central results below are goods-only, and a fuller product-resolution treatment of services is left to future work.

S11 Data pipeline, emission factors and the admissibility rules

These notes give the full version of the pipeline: product harmonization across HS revisions, the provenance of the characterization method, the division of labour between the language model and the rules, the seven admissibility tests rule by rule, and the treatment of quantities, units and exchange rates.

Note S11.1 Product harmonization across HS revisions

Harmonized System product codes are periodically revised. These revisions create splits, mergers and changes in product definitions over time. We harmonize products to the most detailed level that can be followed consistently across the sample. Stable product histories are retained at HS6 level. Products affected by code splits, mergers or definitional changes are aggregated to the lowest common level that preserves temporal comparability. Ambiguous product descriptions and inconsistent units are flagged for coverage and sensitivity diagnostics. The sample window crosses several Harmonized System revisions, and the standard concordance algorithms address exactly this problem [[Pierce and Schott, 2012](#), [Van Beveren et al., 2012](#)], the latter showing that much apparent product churn is a classification artefact. BACI supplies flows in a single fixed nomenclature (HS92), which removes the need to build a concordance here; the residual risk is that a fixed nomenclature carries definitional drift within codes, which the aggregation rule above absorbs.

Note S11.2 Provenance of the characterization method and the re-characterization

Two conventions in it require explicit statement, because both bear on the level of the results. First, the characterization method. All results reported here are characterized with **IPCC 2013, climate change, GWP 100a**, the 100-year horizon standard in carbon accounting and the one on which every benchmark used below is built. This requires a note on provenance, because the choice was not free. The workbook distributed through Base Empreinte carries 981 impact columns, and the first climate-change column in it is CML 2001 (superseded), GWP 500a; an earlier version of this account read that column, for all 19,565 processes, by first match rather than by choice. GWP 500a underweights short-lived gases, so that reading was conservative exactly where methane and nitrous oxide matter most, which is fuels, fertilizers, livestock and chemicals, and it was incommensurable with the GWP 100a benchmarks it was compared against. We re-characterize every process on IPCC 2013 GWP 100a and re-accumulate the account from

flow level. Re-accumulation rather than rescaling is necessary, not fastidious: a factor change moves flows across the physical-validity cap defined below, and so across the boundary between the physical and monetary assignment layers, which a rescaled product-year table cannot represent. The switch raises the 2023 world total by 11.8%, from 9,199.4 to 10,282.1 MtCO₂e (both figures measured on the pre-repair matching; the matching repair of the main text then takes the account to its central 14,367.8), and by between 10.5% and 12.5% in every year of the panel. It is not uniform across products, which is why it cannot be applied to a headline. The median HS6 factor rises by 8.7%. It rises by 130% for live animals and 196% for tobacco, where enteric methane dominates, against 7.1% for mineral fuels and 5.8% for iron and steel, which are carbon-dioxide dominated. the main text reports what the switch does to agreement with the external benchmarks, which is the substantive test of it. Second, the upstream boundary for mineral fuels is discussed in the main text. A single, documented process-LCA backbone keeps the product-level intensities comparable across roughly 5,000 traded products and auditable against the wider literature.

Note S11.3 Record-level application of the characterization switch

Each code's factor is the median across the admissible records mapped to it, and the characterization switch is applied to those records, not to the medians: all 19,565 processes carry a GWP 100a value, and 66,967 of the 68,102 accepted physical records (98.3%) resolve to a named process and take it directly. The remaining 1,135 records reach the account through the reused-concordance and text-retrieval paths without a traceable process identifier, and retain their published value; no HS6 code is built only from such records, so the fallback affects the composition of a median but never determines one. Because the substitution is exact where it applies, no interpolation or ratio is involved: 2,898 accepted records sit on processes whose GWP 500a score is zero, where a ratio would be undefined.

Note S11.4 The seven admissibility tests, rule by rule

Unit compatibility: the donor's reference product must be denominated per kilogram at unit reference amount. A factor per kilowatt-hour, per item, per square or cubic metre, per hectare, per megajoule or per tonne-kilometre therefore cannot enter a mass-traded good. No mass conversion is applied, because none of those denominators admits a documented one. *Waste boundary:* ecoinvent marks a waste-treatment dataset with a negative reference-product amount, because its reference flow is the waste consumed rather than a good produced. Such a dataset may serve only an HS code that itself trades a waste, scrap or residue. A virgin-production dataset may not serve such a code. *Reference-flow substance:* a dataset whose reference product is material removed by a machining operation is rejected; the reference flow is the chip, not the article. Land-use-change, electricity, heat,

transport and infrastructure datasets are rejected too, because they produce no traded good. *Material composition*: the HS chapter and description on one side, and the donor’s activity and reference product on the other, are each mapped to a fifty-class product taxonomy. A donor whose classes are disjoint from the HS code’s is rejected. Matching uses the reference product as much as the activity: ecoinvent multi-output datasets carry several reference products per activity, and the activity name alone can mislead. *Product hierarchy*: for an HS heading whose traded good is a finished assembly, a component or bulk-material process may not stand in for the whole without a documented scaling factor, and none is documented. *Processing stage*: within a material family, the HS nomenclature itself encodes an ordinal processing ladder (ores and concentrates in chapter 26 against unwrought metal, semi-manufactures and articles; crude petroleum in heading 2709 against refined products in 2710; raw crops against milled and prepared foods; fibre, yarn, fabric and made-up textiles). The donor’s stage is read from its reference product first, its activity name second. A donor *more* processed than the code is rejected, because it embeds transformation the traded good has not had: refined silver at 459 kgCO₂e/kg cannot price a precious-metal ore. A donor two or more stages *less* processed is also rejected; that understatement is unbounded (ore for plates, wood chips for fibreboard). A donor exactly one stage below is retained and logged, as the bounded cradle-to-gate lower bound. No numeric cap is involved anywhere. The stage is structural, and families whose nomenclature encodes no reliable ladder (chemicals, glass, ceramics) are left unconstrained by this test rather than forced onto it. *Traceability*: no code may carry a factor whose donor set is empty, and a code left with no admissible donor falls to the stage-matched fallback layer rather than take a wrong factor.

Note S11.5 The exact division of labour between the model and the rules

The division of labour between the model and the rules is exact, and it matters for how the audit results below should be read. A large language model assists two stages and no other: the classification of candidate processes retrieved for an HS6 line, and the disambiguation among documented candidates where several plausibly match. It is not used to generate emission factors. Every one of the seven admissibility tests above, like the repair they came from, is a deterministic rule over recorded metadata: parsed fields, reference-product units and amounts, a fixed lexical taxonomy, a stage ladder read from the nomenclature. All carbon intensities are taken from documented life-cycle sources.

Note S11.6 Why the model is restricted to classification: the documented risks

This guardrail matters. The documented risks for language models in life-cycle work are hallucinated inventory values and irreproducibility across model versions [Preuss et al., 2024], and the reported performance of machine-learning methods for bridging inventory gaps is unstable [Zhao et al., 2025]. Using the model to arbitrate among documented candidates, never to produce a number, keeps the provenance of every factor auditable. Related work recommends emission factors from free-text product descriptions and reports accuracy against expert labels [Balaji et al., 2025, 2023a,b], and background-data matching is one of the open problems such models are proposed for [Tu et al., 2024]. To our knowledge none of it addresses customs nomenclature, and none reports a blind-audited error rate. Those are the two things built and measured here.

Note S11.7 Quantities, units and exchange rates

We use physical quantities directly after unit conversion where they are compatible, fall back to value-based factors where they are not, and exclude flows with neither from the central physical estimates while retaining them in the coverage diagnostics. Partial-coverage products are imputed from unit values within exporter-, importer- or global product-year cells.

The central goods embodied-carbon estimates are products of physical quantities and physical (per-kilogram) emission factors and are therefore independent of currency-conversion assumptions. Monetary conversion enters only two auxiliary components, and they do not share a rate: the services module uses $r^{\text{USD/EUR}} = 1.08$, while the EXIOBASE footprint comparison, whose multipliers are denominated in 2022 euros, uses 1.05. The two are reported separately for that reason and are not netted against each other. The EXIOBASE *level* comparison is sensitive to this rate (the footprint ratio shifts from about 0.4 to 0.6 over a 0.85–1.35 range); the rankings, the within-sector dispersion and the goods totals are not.

S12 The blind audit, the defect census and the repair

These notes give the full version of the audit and of the repair it triggered: the protocol, the first verdicts, the four failure modes, the mechanical cause, the census of the defect and what the repaired totals do.

Note S12.1 Why the repair is reported in full

Those seven tests are the outcome of a two-iteration repair. The repair is a result of this paper, not a detail of its implementation: the paper's own subject is measurement quality, and the matching step is where this account's measurement is made. We report the repair in full, for two reasons. The version of the pipeline that first produced this account applied five of the seven tests not at all. And the before-and-after pair of blind audits below is the measured error rate of a product-level account of this kind, a statistic the literature has not previously reported.

Note S12.2 The blind audit protocol

Accuracy was measured on a stratified sample of 100 HS6 codes from the direct-physical layer: the 30 highest-emission codes taken with certainty, plus 70 drawn with probability proportional to embodied emissions. The author coded the sample, whose professional work is the construction of corporate and product carbon accounts. Each match was graded *correct* (same product and processing stage), *acceptable* (same material family and comparable stage, factor expected within a factor of two) or *wrong* (different material family or processing stage, or a value outside any plausible range).

Blindness was enforced by the design of the workbook rather than by the identity of the coder. The workbook showed the assigned factor, the number and spread of the underlying life-cycle records, and every matched process with its reference product, unit and processing stage. It carried no model verdict, no pre-filled grade and no indication of which assignments the pipeline considered confident, and its rows were shuffled on a fixed seed so that position revealed nothing. The coder therefore saw the evidence for each assignment and not the pipeline's opinion of it. We state the obvious limitation rather than dress it up: the author built the pipeline, so this is blind coding, not third-party validation, and the error rates it measures are the rates a well-informed reader of the evidence would assign. An external audit would be stronger and we would welcome one. Two features make the measured rates hard to flatter in this direction: the grading rule is stated in advance and turns on the donor's reference product and stage rather than on judgement, and the same protocol was applied again after the repair without the residual errors being corrected.

Note S12.3 The first audit verdicts against the model pre-adjudication

The expert found 17 of 100 assignments wrong, 14 acceptable and 69 correct: 4 of the 30 codes in the certainty stratum (13.3%) and 13 of the 70 in the emissions-weighted stratum (18.6%). A model pre-adjudication of the same sample, run earlier, had graded it 45 correct, 40 acceptable and 15 wrong. Against the expert column it agrees exactly on 39 of 100 codes, with Cohen’s $\kappa = -0.003$, indistinguishable from chance, and 19 reversals between *correct* and *wrong* in the two directions. Fifteen of the seventeen wrong assignments had been validated by the model as correct or acceptable, among them the diborane factor on apples and the printed-wiring-board factor on pasta. We therefore withdraw the model-assisted agreement statistics reported in an earlier version of this manuscript. Two model passes measured their own reproducibility, not accuracy, and on this evidence they cannot substitute for a human coder. The division of labour this implies is stated where it can be checked. Model assistance in classification earned its place. Model self-validation measured nothing. The blind human audit is the only reliability evidence this paper treats as such.

Note S12.4 The four failure modes, in full

Inspection of the seventeen identified four distinct failure modes. Each is attributable to a specific stage of the pipeline, not to an inherent limit of the mapping. First, ingestion: the life-cycle export’s column-header block was parsed as data, so 1,135 accepted donor records carried factor values with no process identity at all, and set or co-set the factor of 719 codes. Second, semantically unrelated donors: fresh apples took a factor from *market for diborane*, soya beans from *market for ethylene bromide*, uncooked pasta from *printed wiring board production* at 278 kgCO₂e per kg, ferro-manganese from *treatment of spent nuclear fuel*. Third, a component standing for the assembly it belongs to: six vehicle and engine-part codes took *internal combustion engine production, passenger car* as their only substantive donor. Fourth, a waste-treatment dataset standing for virgin production. Refined petroleum oils taking *treatment of waste mineral oil, hazardous waste incineration* is the largest instance.

Note S12.5 The mechanical cause at the retrieval and arbitration stages

The cause is at the retrieval stage and is mechanical. The text that the similarity score compared against the HS description was not the process name. It was a record string of the form `product 30379 activity <hash> unit kg geography RoW`, so the score was driven by digit substrings and the activity name was never read. Nor did the arbitration that follows test hierarchy, material composition or plausibility. It tested only a keep

flag, membership of the record’s own shortlist, a factor-class label and unit validity, and 98.3% of the accepted records were carried over verbatim from an earlier concordance with the similarity score set to unity by construction. Unit compatibility, the one test of the four that the shipped pipeline did apply, it applied effectively. Of the 68,102 accepted donor records not one is denominated in anything other than kilograms, and no passenger-kilometre or tonne-kilometre process reaches any assignment, including in the vehicle chapters.

Note S12.6 The census of the defect and the two repair iterations

Applying the tests as detection rules over the whole factor set, rather than over a sample, prices the defect. Of 68,102 donor records, 41,156 (60.4%) fail at least one of the first six tests. The failures are concentrated in two modes: 11,764 records are waste-treatment datasets serving virgin products, and 17,142 fail material compatibility. The codes holding at least one inadmissible donor carry 73.2% of 2023 embodied emissions and 84.1% of trade value; the codes whose factor changes once inadmissible donors are removed carry 70.9% of emissions. This is not a tail problem, and it cannot be reported as a robustness check.

A second iteration added the processing-stage test, because the material taxonomy alone cannot separate ore from concentrate from refined metal: after the first repair, precious-metal ores (HS 261690) took 171.6 kgCO₂e/kg from *silver–gold mine operation with refinery*, whose reference product is refined silver. The stage rule removes that donor class (the code’s median falls to 1.02, set by concentrates), removes refinery-gate donors from crude petroleum and rolled-product donors from unwrought aluminium, the three cases whose direction is known a priori, and rejects a further 1,708 donors in all. The full filter rejects 42,864 of the 68,102 records (62.9%); 3,022 of 4,667 codes retain at least one admissible donor, and the 1,645 codes that retain none fall to the stage-matched fallback layer rather than carry a wrong factor.

Note S12.7 Why the repaired total rises, and the orphan bracket in full

We re-derive the factor set from admissible donors only, and we re-accumulate all 258.6 million flows on the same characterization basis, with the same fuel override, validity cap and audited factor corrections. The 2023 world total rises from 10,282 MtCO₂e to 14,368 MtCO₂e, a rise of 39.7%. The direction is the opposite of what the wrong-donor examples suggest, and the reason is measurable. The waste-treatment datasets the filter excludes have a median intensity of 0.023 kgCO₂e per kg, against 1.53 for the rest, and 11.3% of them score exactly zero. They were pulling every median they touched toward

zero. Removing them alone accounts for +3,001 MtCO₂e of the change, and material incompatibility for a further +1,266. The machining-swarf, component-for-assembly and virgin-for-waste rules push the total down. The rules are close to additive.

Coverage falls, as it must when wrong matches are dropped rather than replaced. How the orphaned codes are treated is a reportable choice, not a default, and four treatments bracket it. Plain inheritance of the HS4/HS2 median of the surviving codes gives 14,291 MtCO₂e. Conditioning that inheritance on the processing stage (the stage-matched fallback, our central treatment) gives 14,368. Sending the orphans to the monetary layer gives 13,596. Declaring them uncovered gives 11,463, with 33.0% of trade value then carrying no admissible physical evidence; that is the honest coverage statement that accompanies the central figure. The stage-matched and plain inheritances differ by only +77 MtCO₂e, which is itself informative: the orphan lever is about *which* codes lose their factor, not the stage composition of the pool they inherit from. On the central treatment the direct-physical layer carries 71.7% of emissions and 57.7% of trade value, the stage-matched fallback 20.1% and 32.6%, the plain HS4/HS2 fallbacks 7.7% and 8.5%, and the monetary layer 0.6% and 1.2% (Table 2 of the main text).

Note S12.8 The plausibility diagnostic, and the limits of the audit pair

A factor-level plausibility diagnostic accompanies the filter, as a flag and never as a truncation. For each material family and processing stage, an interquartile band is derived from the surviving donor distribution, and every assigned factor outside it is listed. The diagnostic flags 100 codes carrying 0.5% of emissions. More usefully, it exposes 38 codes whose assigned factor is exactly zero (1.1% of published-basis emissions), led by the oil-cake codes (HS 2304–2306, 63 MtCO₂e on soya oil-cake alone). These codes survive the waste exception because their customs description contains “residues”, and they are then priced by oilfield-water treatment datasets. This is a class error *inside* the waste exception that neither audit sample happened to draw. It is exposed and quantified here as an open issue, not silently absorbed by the monetary screen.

Four limits bound what the audit pair establishes. The samples cover only the direct-physical layer. The fallback layers, now 27.7% of emissions, are therefore unaudited, and their error is bounded only by construction (a stage-matched median is wrong by dispersion, not by class). The detection rules flag 15 of the 15 confirmed-bad codes of the first audit, but also 49 of the 69 the expert called correct. That rate is an upper bound on their false-positive rate, not a measurement of it: the first workbook showed at most three of up to twenty-one donors per code, and inspection of the newly flagged codes finds genuine defects the expert never saw (maize flour pricing raw maize, refined nickel at 44 kgCO₂e/kg pricing nickel ore). Classification error is bounded but still not

propagated into the reported interval alongside factor error. And the audit measures the assignment, not the factors themselves, whose own uncertainty the emission-factor uncertainty subsection of the main text treats separately.

Note S12.9 Pre-repair coverage and what it measured

Before the matching repair the direct layer read 90.2% of value and 94.9% of emissions; that figure measured the reach of an unaudited filter, not of the evidence, and we report the smaller audited coverage as the honest one. The physical layers together still carry 98.8% of the account, so no result rides on the monetary fallback.

S13 Uncertainty, benchmark comparison and independent validation

These notes give the full version of the uncertainty analysis and of every comparison against an external account: the Monte Carlo design and its assumptions, the factor anchors, the OWID and EXIOBASE benchmarks, and the independent ICE check.

Note S13.1 Assignment uncertainty against factor uncertainty

Set against the assignment choices, this parametric interval is the smaller source of uncertainty. The four defensible treatments of the 1,645 orphaned codes span 11,463 to 14,368 MtCO₂e (the main text), a range of 2,905 MtCO₂e or 20% of the central estimate, while the Monte Carlo 90% interval spans 1,871 MtCO₂e or 13%. The comparison understates the gap rather than inflating it: the Monte Carlo range is a probability interval, whereas the assignment range is a set of accounting choices of which none is more probable than the others. A reader who wants one number should take the central estimate with the assignment bracket, not with the Monte Carlo interval alone.

Note S13.2 Why the audited interval is narrower

The interval is *narrower* than the pre-repair one (coefficient of variation 7.8%), and the reason is worth stating. The empirical spread that calibrates each factor's shock was itself inflated by inadmissible donors. Cleaning the donor pool therefore tightens the measured parametric uncertainty at the same time as it raises the level.

Note S13.3 Monte Carlo design, its assumed spreads and its limits

The interval is far narrower still than on a combustion-inclusive basis, where the design gives a coefficient of variation of 23%. Mineral fuels (HS27) carried about 96% of total variance there. Removing the combustion term removes both the level and the dispersion it contributed, so the upstream basis is not merely smaller but better identified. What dispersion remains still rests on an assumed spread, set from a robust interquartile-based coefficient of variation bounded to [0.05, 1.00] (median 0.30 across the overridden HS27 codes); the per-chapter variance shares are reported in the Supplementary Material. This exercise is accordingly a factor-level sensitivity analysis rather than a fully propagated uncertainty: its width is governed by one judgement call, and an empirically grounded alternative would derive factor dispersion from ecoinvent's re-estimated pedigree-matrix uncertainty factors [Ciroth et al., 2013]. Two further limits should be read with it. The draws are independent across products, whereas the roughly five thousand factors descend

from a smaller set of processes sharing electricity mixes, steel and fuel inputs. Independent perturbation diversifies error away. Both the coefficient of variation and the concentration of variance in one chapter are therefore understated relative to a correlated design [Groen and Heijungs, 2017, Schulte et al., 2024, Kim et al., 2025]. We quantified that understatement by re-running the propagation under dependence with each product’s marginal distribution held fixed, so that only the dependence structure changes; across fifteen variants the aggregate *mean* moves by 0.52% while the coefficient of variation moves from 4% to 40%. Products sharing a donor process, perfectly correlated within block, give 4.9% and an interval 1.20 times wider; products sharing an HS2 chapter give 9.5% and 2.34 times wider. The first is a floor rather than a bound, because 46.4% of 2023 emissions sits in donor blocks of size one, crude petroleum alone carrying 8.0% of the direct-layer total, and because same-process blocking cannot see the mechanism at issue, which is shared upstream across *different* processes. A background component carrying 20% of each factor’s log variance gives 17.8% and [10,438, 18,740] MtCO₂e, 4.4 times wider. These are scenarios and not estimates: the correlation cannot be calibrated here, because the deposited factor extract carries characterised scores rather than the technosphere matrix that would identify shared upstream. The 4.0% figure is therefore conditional on independence, and any plausible dependence widens it by a factor of two to four. And the simulation perturbs factors only, not the assignment: the fallback allocation and the match choices are held fixed, and in models of this type data-source and aggregation choices can dominate stochastic measurement error [Badr et al., 2026]. The product and country *rankings*, by contrast, are stable across draws: the ordering holds even where the absolute level does not.

Note S13.4 Factor anchors against published ranges and the CBAM defaults

As an external check, we compare the assigned cradle-to-gate factors for the carbon-dominant traded materials against published life-cycle values. The factors for the products that drive the totals sit close to the published ranges, and the matching repair moved two of them *into* range. Unwrought aluminium is 8.3–15.1 kgCO₂e/kg against a published cradle-to-gate range of 8–18, inside it; the pre-repair lower end, 7.1, fell below the floor because rolled-product service datasets diluted the median. Semi-finished steel is 2.25 kgCO₂e/kg against 1.8–2.5, inside it; it stood at 3.3, above the ceiling, before the stage rule. Portland cement is 0.87 kgCO₂e/kg against 0.6–0.9, inside it. The audit also flagged factors that the ecoinvent extract placed outside the literature range. The nitrogen-fertilizer factors ran about twice the published cradle-to-gate values (the extract appears to mix per-nutrient and per-product bases), and cement clinker, hydrogen and direct-reduced iron were too low. We correct these to their published ranges (nitrogen

fertilizers capped at 4.0 kgCO₂e/kg, clinker raised to 0.85); the corrected factors are our central estimate. The correction is small and does not affect any conclusion: total estimated embodied emissions fall by 1.9% (to 14,368 MtCO₂e) and the fertilizer share from 6.5% to 4.5%, while the concentration ranking and the within-sector dispersion are unchanged. The overall pattern is consistent with the footprint comparison (the main text): the account is reliable for the bulk materials that dominate embodied carbon.

Because the framework targets product-level border measures, we also benchmark the factors against the European Union’s CBAM total default embedded-emission values (direct plus electricity), the boundary closest to our cradle-to-gate factors (Appendix S3) [European Commission, 2023]. For the high-tonnage CBAM goods that carry most regulated carbon the agreement is close. Cement clinker (0.85 versus 0.87 t/t), unwrought aluminium (8.3–15.1 versus 10.5), anhydrous ammonia (2.98 versus 2.82), hydrogen (10.0 versus 10.4) and steel (pig iron 1.7 and semi-finished 2.25 against CBAM totals of 2.1–2.2) all sit within about a factor of 1.5 of the regulatory default, and the median ratio across the ten regulated groups is 1.00. The two exceptions are the ones already identified: nitric acid is below the default, and the nitrogen fertilizers, even after the cap, remain above it, so the fertilizer contribution stays a conservative upper bound. The factors thus track the regulatory benchmark for the products a border measure would target.

Note S13.5 What the benchmark comparison tests

We validate the account against benchmark consumption-minus-production wedges in four ways: aggregate scale, same-sign agreement, country-year correlations and residual diagnostics. Exact equality with MRIO-based benchmarks is not expected because the two approaches have different system boundaries and accounting interpretations. The comparison tests whether the product-level Trade-LCA account is plausible in aggregate and identifies where product trade exposure and benchmark consumption accounting diverge.

Note S13.6 How each repair is identified against the OWID benchmark

Two repairs separate these numbers from the ones an earlier version of this account produced, and each is identified against this benchmark in its own way, because the OWID wedge is itself a GWP 100a, technology-specific construct. Re-characterizing from the extract’s superseded GWP 500a column onto GWP 100a raised the aggregate magnitude ratio from 0.700 to 0.767 while barely moving the pattern statistics (Pearson 0.381 to 0.385). A characterization change is close to a monotone rescaling of every factor, and a rescaling cannot reorder countries. The switch is therefore identified on levels, which is where it helps. The matching repair then moves both: the ratio from 0.767 to 0.956 and the panel Pearson from 0.385 to 0.69, with the latest-year cross-section going from 0.05 to

0.69. A repair that was motivated entirely by internal evidence, an expert audit of donor records, moves the account most of the remaining way toward an external benchmark it was never fitted to. We read that as the strongest available confirmation that the audit was measuring real error.

Note S13.7 The history of the aggregate reconstruction ratio

The history of this ratio is itself instructive, and it is worth stating plainly rather than presenting as a loss. On a combustion-inclusive basis the same comparison gave an aggregate magnitude ratio of 1.007, which reads as near-exact agreement. It was not: it was two errors of opposite sign cancelling. Charging traded fuels with their eventual combustion inflated the account, while wrong donor records, waste-treatment datasets with near-zero factors above all, deflated it. Removing the combustion term dropped the ratio to 0.767, a figure this paper’s earlier version reported as process-LCA truncation. The matching repair then shows that most of that apparent truncation was matching error. The audited account reconstructs 96% of the wedge with a country-level Pearson correlation of 0.69, against 0.32 where this account started. Each repair improved the pattern statistics or left them intact, which is the sequence one should demand of corrections that are real rather than fitted. Adding the supplementary services extension barely moves these figures and does not improve the fit; we therefore retain the goods-only account as the benchmark and report the services comparison only in the Supplementary Material. Across the 28 available years, the log-linear correlation between our sector-level emission estimates and EXIOBASE is stable at $r = 0.76\text{--}0.79$ ($0.76\text{--}0.81$ before the matching repair), confirming directional agreement throughout the panel.

Note S13.8 The EXIOBASE footprint comparison: boundary, mechanisms and sector detail

The appropriate level benchmark is the EXIOBASE *footprint* intensity (the full Leontief total-requirements multiplier, **fL**), which, like our cradle-to-gate factors, embodies the entire upstream supply chain; comparing instead to the *direct* sectoral intensity would spuriously inflate the apparent gap by the upstream multiplier. Against the footprint benchmark our estimates now sit close to unity at the median (Figure 3 of the main text). Weighting each exporter–sector cell by the embodied carbon it carries, 82.8% of our estimated embodied emissions fall within a factor of three of the EXIOBASE footprint intensity and 65.1% within a factor of two; the emissions-weighted median capture ratio is 0.90, the unweighted median 0.98 and the aggregate 0.61. Agreement is close for the material-intensive sectors that dominate embodied trade carbon: base metals (≈ 1.0), chemicals (≈ 0.9), food and plastics (≈ 1.2), with vegetable products and wood above the benchmark ($\approx 1.8\text{--}2.1$). It remains weakest where the boundaries genuinely differ.

Machinery and electronics sit at ≈ 0.6 : their deep multi-tier supply chains a cradle-to-gate factor cannot fully carry [Lenzen, 2000, Suh et al., 2004, Crawford, 2008]. Mineral fuels sit at ≈ 0.24 : the upstream-only boundary this account adopts deliberately excludes the combustion term an input-output benchmark carries. A complete hybridised process-input-output database exists [Agez et al., 2020], so the use of process factors here is a choice rather than a gap: it preserves the physical-unit traceability from an HS6 quantity to a cradle-to-gate factor and avoids a price conversion. The affirmative case is that filling process gaps with sector averages substitutes aggregation error for truncation error [Yang et al., 2017]; the opposing position holds that hybridisation is nonetheless more accurate [Pomponi and Lenzen, 2018], and the question remains open [Ward et al., 2017, Majeau-Bettez et al., 2011, Perkins and Suh, 2019].

The decomposition of this account’s own gap is a result in itself, because three mechanisms have now been separated on the same cells, with the same footprint intensities and the same exchange rate. Characterization comes first. The EXIOBASE benchmark row is a GWP 100a construct, so the extract’s GWP 500a column was being compared against it across a horizon mismatch. Re-characterizing raised the emissions-weighted median capture from 0.56 to 0.67. The largest moves sit exactly where methane and nitrous oxide matter (live animals, food, vegetable products) and the negligible ones where carbon dioxide dominates, the pattern a horizon correction must produce. Matching error comes second, and it is the larger of the two: the audited filter raises the same statistic from 0.67 to 0.90, the unweighted median from 0.52 to 0.98 and the aggregate from 0.41 to 0.61. What an earlier version of this paper reported as process-LCA boundary truncation was therefore in large part mis-assigned donors deflating the factor medians. What remains is the residual undershoot in machinery and fuels against overshoot in agriculture and wood. That is the boundary-definition residue a cradle-to-gate account genuinely carries against a Leontief total-requirements intensity. It no longer supports the blanket claim that the account is a conservative lower bound: the account is approximately unbiased at the median and heterogeneous by sector. The comparison is best read as a trade-weighted extension of the published ecoinvent-versus-EXIOBASE comparisons [Steubing et al., 2022, Liu et al., 2025, Giljum et al., 2019] rather than as a fresh validation: what has no precedent here is the weighting by traded embodied carbon, the HS6 resolution, and the stability of the fit across twenty-eight years.

Note S13.9 The independent ICE check in detail

The footprint and CBAM benchmarks both draw, directly or indirectly, on the same body of process and inventory data as our factors. As an *independent* check we compare the assigned factors to the Inventory of Carbon and Energy (ICE), a material-carbon database compiled from peer-reviewed and industry sources outside the ecoinvent–Agribalyse back-

bone [Hammond et al., 2025]. Because ICE reports cradle-to-gate material intensities (the same system boundary as ours), it is the appropriate like-for-like test, and the agreement is close. Across the eleven material entries we can match to ICE (metals, cement and clinker, glass, plastics and ceramics; ICE covers materials only, so neither fuels nor fertilizers are testable here), the median ratio of our factor to the ICE value is 1.04 and the log-log correlation is 0.91. Two of the eleven factors, cement clinker and Portland cement, were anchored to published values during screening, so their agreement is not independent evidence; excluding them leaves the median at 1.04. Nine of the eleven factors fall within a factor of two (Appendix S4, Figure S1); of the two outside, natural rubber latex (0.33) is one of the codes whose own donors the filter rejected, so its comparison tests the stage-matched fallback rather than the matching, and petroleum bitumen sits at 2.03, just past the boundary. The agreement is close for the carbon-dominant metals and cement (unwrought aluminium 1.25 and 1.40, Portland cement 1.04, clinker 0.99, semi-finished steel 1.04, the last moved onto the diagonal by the stage rule). This check also bounds the characterization switch from the other side, and the result is worth stating against the expectation rather than for it. ICE is a GWP 100a source too, so one might have predicted improvement here as well; there was no room for it, because on GWP 500a the median ratio was already 0.99. The switch moves it to 1.04, crossing unity rather than closing a gap. The reason is physical: the materials ICE covers are carbon-dioxide dominated, so their factors are nearly horizon-invariant; clinker does not move at all and the metals rise by 8%, against 130% for live animals. Agreement with a cradle-to-gate benchmark therefore cannot discriminate between the two horizons, whereas agreement with a full-requirements benchmark can, which is why the EXIOBASE result above carries the evidential weight. ICE does not cover fuels, fertilizers or bulk organic chemicals, so this check validates the material side; the fuel intensities that dominate the totals rest on standard, well-established combustion factors.

Note S13.10 The China sign flip and the largest positions

One structural feature of the pre-repair comparison deserves its own record, because the repair resolved most of it. Before the matching repair the panel Pearson correlation was 0.39 and fell to 0.05 in the latest cross-section, driven by the two largest positions, China and India, whose reconstructed adjustments took the opposite sign to the benchmark wedge. On the audited account China's adjustment is $-812 \text{ MtCO}_2\text{e}$ against a benchmark wedge of $-1,315$: the sign now agrees, and the largest position in the panel has moved from contradiction to understatement, which is what lifts the Pearson statistic to 0.69. India's divergence survives ($+349$ against -520), as does Vietnam's. Among the thirty largest non-entrepôt positions, sign agreement is 87% and the scale-robust correlation 0.70. Entrepôt re-export is a documented distortion of bilateral attribution, canonically

for Hong Kong [Feenstra and Hanson, 2004]; and mirror-report gaps vary systematically with tariffs, corruption and auditing standards, so the measurement error BACI reconciles is not random [Kellenberg and Levinson, 2018].

Note S13.11 The truncation-asymmetry test on the residual divergences

We test whether the residual divergence is an artefact of differential truncation. The account’s capture ratio against the EXIOBASE footprint still varies across sectors, and a country that imports from well-captured sectors while exporting from poorly captured ones would show a positive adjustment for measurement reasons alone. For India it is not that artefact, and the divergence is now better identified than before the repair. The three defensible treatments of truncation are: no correction; restriction to sectors both accounts capture at 0.60 or better; and first-order de-truncation of every sector on both legs. They give +346, +137 and +998 MtCO₂e, a consistent positive sign against a benchmark wedge of −519.5. India is therefore a genuine composition-versus-technology divergence: a large importer of carbon-intensive raw inputs whose exports a representative-technology account credits at generic intensity while a consumption account debits them at India’s own. For China the three treatments give −763, −1,400 and +751, so the sign is not identified there. The de-truncation limits remain. Dividing the fuel leg by its capture ratio of 0.24 partly re-imports the combustion emissions the upstream basis deliberately excludes. And the de-truncated total still places 44.7% of its mass on sectors with capture below 0.30 (67.3% before the repair).

Note S13.12 The aggregate reconstruction ratio in detail

Across all country-years, the goods-only reconstruction captures 119,642.6 MtCO₂e in absolute reconstructed trade-adjustment magnitude, compared with 125,179.1 MtCO₂e in the benchmark wedge. The account thus captures 96% of the benchmark wedge in aggregate *magnitude*, a goods reconstruction ratio of 0.956. This is a test of scale, not of agreement: the ratio compares sums of absolute magnitudes, and the sign of the reconstructed adjustment disagrees with the benchmark for about 31% of country-years, which the next paragraphs quantify. The sign of the reconstructed adjustment agrees with the benchmark wedge in 69.1% of country-year observations. The log-signed cross-country correlation with the benchmark wedge is 0.53, the Spearman rank correlation 0.50 and the level (Pearson) correlation 0.69 over the full panel; before the matching repair the Pearson statistic was 0.39, held down by China’s opposite-signed position, which the repair resolved (the main text). The supplementary services extension leaves these figures essentially unchanged (Supplementary Material). The annual trend is shown in Figure S7.

S14 Concentration, within-sector dispersion and country patterns

These notes give the full version of the results: how concentration is measured and on which basis, how the within-sector dispersion should be read, where the sector-average mis-estimates fall, and the country-level patterns behind the headline positions.

Note S14.1 Layer-by-layer evidence structure of the central account

The central account is not driven by residual assumptions, and its evidence structure is stated after, not before, the audit. Direct HS6 physical matches on admissible donors account for 57.7% of covered trade value and 71.7% of estimated embodied emissions in the latest goods account (Figure S3). The stage-matched fallback carries 7.38 trillion USD and 2,884.9 MtCO₂e (32.6% of value, 20.1% of emissions); the plain HS4 and HS2 fallbacks add 1.25 and 0.67 trillion USD (896.5 and 202.4 MtCO₂e). The final monetary fallback layer accounts for 0.27 trillion USD, or 1.2% of trade value, but only 84.9 MtCO₂e, or 0.6% of emissions. Physical evidence, direct or family-inherited, therefore carries 98.8% of the account, and the share resting on audited product-specific donors is stated separately from the share resting on inheritance, which is the distinction an unaudited coverage figure conceals. The workflow is illustrated in Figure 1 of the main text.

Note S14.2 Why concentration is reported on the gross basis

As noted in the main text, this gross figure double-counts the upstream of multi-stage trade and is an exposure measure. The net consumption-side embodied trade is correspondingly smaller. The concentration shares below are reported on this gross basis.

Note S14.3 Concentration mechanics: fallback-carried chapters and the fertilizer cap

Two chapters owe their position mainly to the stage-matched fallback rather than to direct evidence, and this must be said where the ranking is used. Machinery (7.2% of emissions) and electrical machinery (6.8%) inherit stage-3 family medians for most of their codes. Their levels test the fallback, not the matching. Mineral fuels remain the largest single chapter, but on an upstream-only basis they no longer dominate the account. Their share falls to 15.2% when the eventual combustion of traded fuels is not charged to the flow, which is the boundary this account adopts (the main text). Fertilizers are the one large chapter whose level the audited factor corrections cut materially (950.4 to 644.5 MtCO₂e):

the nitrogen-fertilizer cap is an absolute literature anchor of 4.0 kgCO₂e/kg, and seventeen of the twenty-six HS31 codes bind against it.

Note S14.4 The gross basis and the product mix

The gross basis is the appropriate one here: concentration measures where embodied carbon *flows*, and it is gross flows, not netted balances, that border measures, procurement standards and supply-chain screening act upon. The concentration is a property of the product mix and does not depend on the upstream double-counting discussed in the main text, since that double-counting is roughly proportional across the bilateral flows within a product group.

Note S14.5 Reading the audited dispersion against unaudited accounts

We know of no other product-level account that has audited its own dispersion this way, and the pre/post pair suggests the published magnitude of within-sector heterogeneity should be read with the same suspicion this paper applies to sector averages. The audited thirteen-fold figure is the one we defend, and it is still far beyond anything a sector average can carry.

Note S14.6 The price channel inside monetary dispersion

Monetary intensity is the physical factor divided by the BACI unit value, so it compounds carbon dispersion with price dispersion, and we quantify the second channel rather than leaving it implicit. Unit values alone span 4.4-fold within the median chapter. Holding unit values at their sector trade-weighted mean, which switches the price channel off, returns the monetary spread exactly to the physical figure: dividing by a sector constant cannot change a within-sector ratio, so the counterfactual is an algebraic identity rather than an approximation. On the common sample where the identity $m = c/p$ holds exactly, the physical spread is 12.8-fold and the monetary 15.5-fold. In the exact variance decomposition $\text{Var}(\log m) = \text{Var}(\log c) + \text{Var}(\log p) - 2 \text{Cov}(\log c, \log p)$ the price term is 33% of $\text{Var}(\log m)$ in the median chapter and 51% trade-weighted. But carbon intensity and unit value are *positively* correlated within sectors (median within-chapter correlation +0.16, negative in only 36% of chapters). Goods that are carbon-heavy per kilogram also tend to be dearer per kilogram. The covariance term therefore cancels most of the price channel, the two bases nearly coincide at the median, and in 17 of 72 chapters monetary dispersion falls *below* physical dispersion. One artefact of the pre-repair account is worth recording here: it showed monetary dispersion at 2.1 times physical, and that gap too

was largely matching error, because wrong donors distort the numerator of m without touching its price denominator.

Note S14.7 Where the mis-estimates concentrate

The mis-estimates are concentrated where it matters. Sectoral averaging is adequate for bulk, homogeneous sectors but fails where a sector mixes products of very different intensity: ores, base metals, chemicals and machinery, in which products such as primary aluminium, ferro-alloys, petroleum coke and accumulators sit far above their sector mean. These are exactly the products that border measures and procurement standards must single out, and that sectoral accounting cannot.

Note S14.8 Resolution bounds and time stability of the mis-estimated share

Finer sectors do not dissolve the heterogeneity. Our HS2 chapters (97 categories) are no finer than a detailed MRIO; EXIOBASE distinguishes about 200 products. Yet it persists when we go finer. At HS4 resolution (341 groups, finer than any operational MRIO), within-group dispersion is still about four-fold, and a group average still mis-estimates 8.7% of emissions and 14.6% of value beyond a factor of two. MRIO resolution lies between these levels, so the mis-estimated share is bounded between 9% and 46% of emissions and between 15% and 60% of value: substantial at any feasible MRIO granularity, and irreducible without product detail. The shares are stable over time (the value share stays in a 54–62% band across 1995–2023, and the emissions share in a 41–47% band).

Note S14.9 Gross and net trade positions as different exposure patterns

S14.1 Gross and net trade positions identify different exposure patterns

Gross and net embodied trade exposure identify different risk dimensions (Figure S5). Net positions identify whether a country or product position is a net importer or exporter of embodied carbon. Gross positions identify large two-way flows. A country-product position can have a low net balance because imports and exports offset each other, while still involving large gross flows that remain relevant for carbon pricing, border measures, procurement standards and supply-chain decarbonization.

Note S14.10 Country residuals for the mid-sized economies

For several economies, the goods reconstruction captures a substantial share of the benchmark wedge. The United States has a benchmark wedge of 513.3 MtCO₂e and a goods net adjustment of 448.7 MtCO₂e (87% of the wedge); Russia −327.2 against −344.4; Japan 155.4 against 184.8. For France the goods net adjustment (149.4 MtCO₂e) slightly exceeds the benchmark wedge (137.1 MtCO₂e), while Germany’s (85.4 against 174.7) falls short, the residual reflecting services, re-exports and indirect requirements not captured by product-level customs flows.

Note S14.11 The China reading and its mechanism

China deserves its own sentence, because it is where the matching repair changed the account’s reading most. Before the repair the reconstruction made China a net *importer* of embodied carbon (+266 MtCO₂e) against a benchmark wedge of −1,315. On the audited account China is a net exporter at −812. The sign agrees, and the residual is a magnitude gap rather than a contradiction. The mechanism is the machinery and electronics factors: the defective donors priced China’s dominant export chapters near zero, and the repaired factors restore the embodied carbon its exports carry.

Note S14.12 India and Vietnam as composition-versus-technology divergences

The most informative remaining residuals are the economies whose product-level position still differs in *sign* from the benchmark wedge: India (−520 versus +349) and Vietnam (−100 versus +50). Consumption-based accounts make these economies large net *exporters* of embodied carbon, because their manufactured exports embed dirtier-than-average domestic production. Our representative-technology account makes them net *importers*. It values their large imports of carbon-intensive raw and intermediate inputs (fuels, ores, oilseeds, petrochemicals, fertilizer) at product-generic intensities, and it credits their exports at the same generic technology. For India the divergence survives every truncation treatment (the main text). The divergence is exactly what the composition lens is built to show. It separates *what* a country trades from *how cleanly* it produces. It therefore points to the economies for which production technology, the dimension a border measure such as the CBAM acts on, matters most.

Note S14.13 Global geography of the adjustments

S14.2 Global geography of Trade-LCA adjustments

The geography of product-level trade exposure differs from both production scale and benchmark CBE gaps (Figure [S6](#)). The world map shows where country-level trade-adjusted production emissions estimated from product-level Trade-LCA are higher or lower than benchmark consumption-based emissions, with bubble size proportional to production-based emissions. The map is used as a diagnostic view of where product-level trade evidence aligns with, or diverges from, benchmark consumption-based accounting.

S15 Discussion, limitations and future work

These notes give the full version of the discussion: the contribution claims, the policy reading of defaults versus actual emissions, the relationship to MRIO accounts, every limitation and the future work.

Note S15.1 The three steps beyond existing product-level trade-carbon work

The paper advances three steps beyond existing product-level trade-carbon work. [Sato \[2014a\]](#) and [Kanemoto et al. \[2016\]](#) established that embodied carbon can be attributed to traded products, and [Minx et al. \[2009\]](#) noted that sectoral aggregation conceals heterogeneity (a point echoed at the firm level by [Barrows and Ollivier \[2018\]](#), who show that product mix, not only firm efficiency, shapes manufacturing emission intensity). [Su et al. \[2010\]](#) went further and quantified how the level of sector aggregation biases input–output estimates of emissions embodied in trade. They did so by comparing coarser and finer input–output partitions, so the residual heterogeneity below the finest available sector remained unmeasured. These contributions were either single-year or worked within input–output resolution rather than below it. First, we build the account on a single, documented process-LCA backbone (ecoinvent and Agribalyse), mapped to HS6 through an explicit, auditable evidence-layer hierarchy, over a 29-year, $\sim 5,000$ -product, 120-country panel. The provenance and quality of every estimate is therefore traceable rather than assumed, and we measure that quality: a blind expert audit of the matching before and after its repair (17 then 8 wrong assignments in 100) is, to our knowledge, the first published error rate for an account of this kind. Second, we move from the qualitative observation that “sectors are heterogeneous” to a quantitative policy statistic, audited against our own measurement error. A value-based sector average, the resolution of MRIO models, mis-estimates by more than a factor of two the embodied carbon of products representing 45.5% of traded emissions and 59.9% of trade value. We show this result is genuine carbon heterogeneity. It survives on a pure physical basis (34.2%). It is robust to factor outliers. And, critically, it survives the matching audit, whereas the dispersion magnitude itself does not (thirteen-fold audited against thirty-five-fold unaudited). Third, we turn the limitation of the representative-technology basis into information: where it disagrees in sign with consumption accounts, the gap pinpoints the economies for which production technology is decisive (the composition-versus-technology distinction). The factors themselves match an independent material-carbon database and the EXIOBASE footprint (83% of embodied carbon within a factor of three).

Note S15.2 The contribution to the literature on language models in LCA

The audit pair also carries the paper’s contribution to the literature on language models in life-cycle assessment, and the claim is made in narrow terms because narrow is what the evidence supports. This is, to our knowledge, the first model-assisted pipeline matching customs nomenclature to a life-cycle database whose reliability is measured: blind expert coding of stratified emissions-weighted samples, an error rate before repair (17 in 100), a repair made of deterministic rules, and a re-audit after it (8 in 100). The existing work matches free-text product descriptions and reports accuracy against expert labels [Balaji et al., 2023a,b, 2025], and the field has named background-data matching a grand challenge [Tu et al., 2024]; measured reliability of the kind reported here is what that agenda still lacks. The cautionary result is part of the contribution rather than a caveat on it. The model-based pre-adjudication of match quality agreed with blind expert coding at chance level ($\kappa = -0.003$) while validating semantically absurd matches, and everything that fixed the pipeline was a deterministic rule: ingestion hygiene, unit compatibility, the waste–virgin boundary, the material taxonomy, the stage ladder. The exportable lesson is one sentence: model assistance in classification is useful, model self-validation is worthless, and a blind human audit is not optional.

Note S15.3 Defaults versus actual emissions in border-charge design

One scope statement belongs at the head of this subsection. The factors used here are representative-technology, product-generic intensities, which is what the border-adjustment literature calls default values. There is a live argument in print that defaults are the wrong basis for a border charge. Defaults blunt the abatement incentive that carbon pricing is meant to create [Mehling and Ritz, 2023]. An actual-emissions calculation is what keeps such a measure defensible in law [Boute, 2024]. The one real-world evidence base finds significant leakage through resource shuffling when a border charge must price producers of differing intensity [Fowlie et al., 2021]. Against these, standardized values are defended for materials with complex value chains [Neuhoff et al., 2026]. This paper takes no position in that debate. Its factors are proposed as an *accounting, screening and prioritisation* instrument, to identify where sector averaging fails and where verification effort should be concentrated. They are not proposed as the liability basis for a charge on any individual shipment.

Note S15.4 The composition-versus-technology decomposition of embodied-trade positions

Because the factors are technology-normalized, the framework also supports a policy-relevant decomposition of a country’s embodied-trade position into a *composition* component (what it trades, at representative-technology intensities) and a *production-technology* component (how cleanly it produces, proxied by national energy carbon intensity). On the stored decomposition the composition component is the larger of the two on average: the mean absolute technology component is about 0.3 times the mean absolute composition component. But the technology component is not uniformly secondary. It exceeds the composition component for several economies, and it reverses the sign of a few mid-sized net positions. Quantifying the decomposition over the full range of grid-intensity pass-through, at product rather than chapter resolution, is left to future work. Border instruments such as the CBAM act on the production-technology dimension and therefore address only this secondary component, and they risk relocating rather than reducing emissions where leakage is unconstrained [Böhringer et al., 2022]. The embodied carbon of trade is governed primarily by *what* is traded, a heterogeneity that only product-level accounting can resolve. This product view also matters distributionally: the carbon-intensive materials our account flags (steel, fertilizers, aluminium) are exactly those through which the CBAM’s burden falls on commodity-dependent developing economies [Magacho et al., 2024]. A full development of this decomposition is left to future work.

Note S15.5 What the two approaches each give

MRIO models and Trade-LCA answer different questions. MRIO models allocate emissions to final demand through production networks. Product-level Trade-LCA measures embodied carbon in observed product-level trade flows. The two approaches should be used together, which is also the direction the field’s own consensus assessment points in, and it ranks the quality of the environmental data ahead of the resolution of the economic table [Tukker et al., 2018]. MRIO gives economy-wide consistency and final-demand responsibility; Trade-LCA gives product-level exposure and hotspot identification. The log-linear correlation between our pipeline’s sector-level emission estimates and EXIOBASE is $r = 0.76\text{--}0.79$ consistently across all 28 available years, confirming directional agreement despite differences in system boundaries and intensity levels [Dietzenbacher et al., 2013].

Note S15.6 Reading the MRIO comparison, and why the corrections compound

The comparison should be read against what is known about the MRIO accounts themselves. Their consumption-based estimates carry documented uncertainty [Rodrigues et al., 2018] and converge only partially across databases [Moran and Wood, 2014]. A residual between a product-level account and any single MRIO account is therefore not by itself evidence against either [Wiedmann and Lenzen, 2018].

The two literatures are also correcting different things, and their corrections compound rather than compete. Davis et al. [2025] show that replacing a single-region input–output model with a high-resolution multi-regional one adds about 2.0 Gt CO₂e, roughly a tenth, to the upstream emissions of CDP-reporting companies, with the largest differences in manufacturing sectors of moderate emissions intensity. That correction is geographic: it gets the region of origin right. The dispersion measured here is orthogonal to it, because it persists *within* a sector once the region is fixed, and it survives at HS4 resolution, finer than any operational multi-regional table (the within-sector dispersion subsection of the main text). A single-region account can therefore be wrong about where the carbon was made and wrong about which products carry it, and correcting one does not correct the other.

Note S15.7 The first limitation, in full

The framework has several limitations, and the first of them is the largest. The HS-to-life-cycle matching that produces an account of this kind is error-prone at a measured rate: 17 wrong assignments in 100 before the repair, 8 in 100 after it (the main text). The residual errors are all same-class, same-stage semantic mismatches carrying 1.2% of emissions. They are the kind that neither a lexical taxonomy nor a structural stage rule can catch. The audit covers only the direct-physical layer; the stage-matched fallback that now carries 20.1% of emissions and 32.6% of trade value is inheritance, not product-specific evidence, and 33.0% of trade value carries no admissible direct physical evidence at all. The 38 zero-factor codes exposed by the plausibility diagnostic (1.1% of pre-repair emissions, the oil-cake codes above all) remain open. Nothing else in this list is of that order.

Note S15.8 Further limitations

Life-cycle emission factors are uncertain and often represent average technologies rather than producer-specific production routes. The factors are also held fixed over time: each product carries a single recent-vintage intensity applied to every year from 1995 to 2023. Where the true unit intensity of a good fell over the period (through grid decarbonization

or efficiency gains in energy-intensive materials such as steel, aluminium, and cement), this static convention understates early-year embodied carbon and so compresses the measured downward trend. The within-year composition and dispersion results are unaffected, since they compare products in the same year; but the annual-level series (Figure S7) should be read as a lower bound on decarbonization, not as a technology-adjusted trend. HS product categories can contain heterogeneous products. BACI physical quantities contain missing values, inconsistent units and reporting errors. Re-exports and entrepôt trade can distort country-level interpretations. The framework does not propagate emissions through complete global value chains and does not provide firm-level exposure. It also does not allocate domestic production emissions to product-country-year positions. The supplementary services extension carries its own limitations (it relies on monetary intensities and incomplete World Bank balance-of-payments sector decompositions), which are documented in the Supplementary Material; these are among the reasons services are kept outside the core account.

Note S15.9 Future work in detail

Future work should develop more country- and technology-specific emission factors, integrate product-level Trade-LCA with MRIO models, link product exposure to firm-level customs or procurement data, and propagate uncertainty across product mappings and emission factors. A natural next step is services. The supplementary monetary-intensity extension shows the limits of a sector-level treatment. We plan to explore a product- or activity-resolution services account, built on more granular BoP data or on confidential customs-services microdata, that would bring services onto the same evidentiary footing as goods. Such extensions would strengthen the use of product-level embodied carbon accounting for climate policy and industrial transition analysis.

Note S15.10 Concluding synthesis, in full

The results show that a small number of products account for a disproportionate share of traded embodied carbon, and that trade value is a poor proxy for exposure. They also show that carbon intensity varies by an order of magnitude within the sectors at which input-output models operate, so a sector average mis-estimates the embodied carbon of a large share of trade. And they show, by auditing this account's own matching before and after its repair, that unaudited product-level accounts overstate that heterogeneity. The audit is as much a part of the method as the mapping.

S16 Supplementary figures

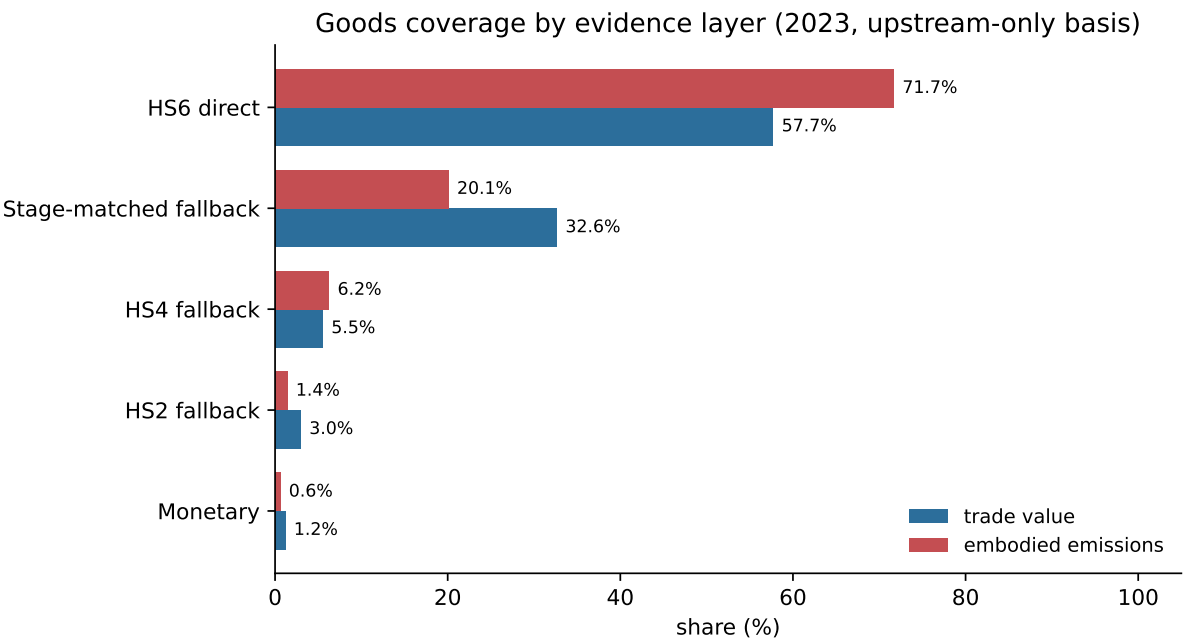


Figure S3: **Goods coverage by evidence layer.** Shares of trade value and embodied emissions assigned through direct HS6 physical matching, HS4 fallback, HS2 fallback and monetary fallback layers in the latest year. Direct physical matches on audited donors carry 57.7% of trade value and 71.7% of emissions, and physical evidence including the stage-matched and family fallbacks carries 98.8% of both.

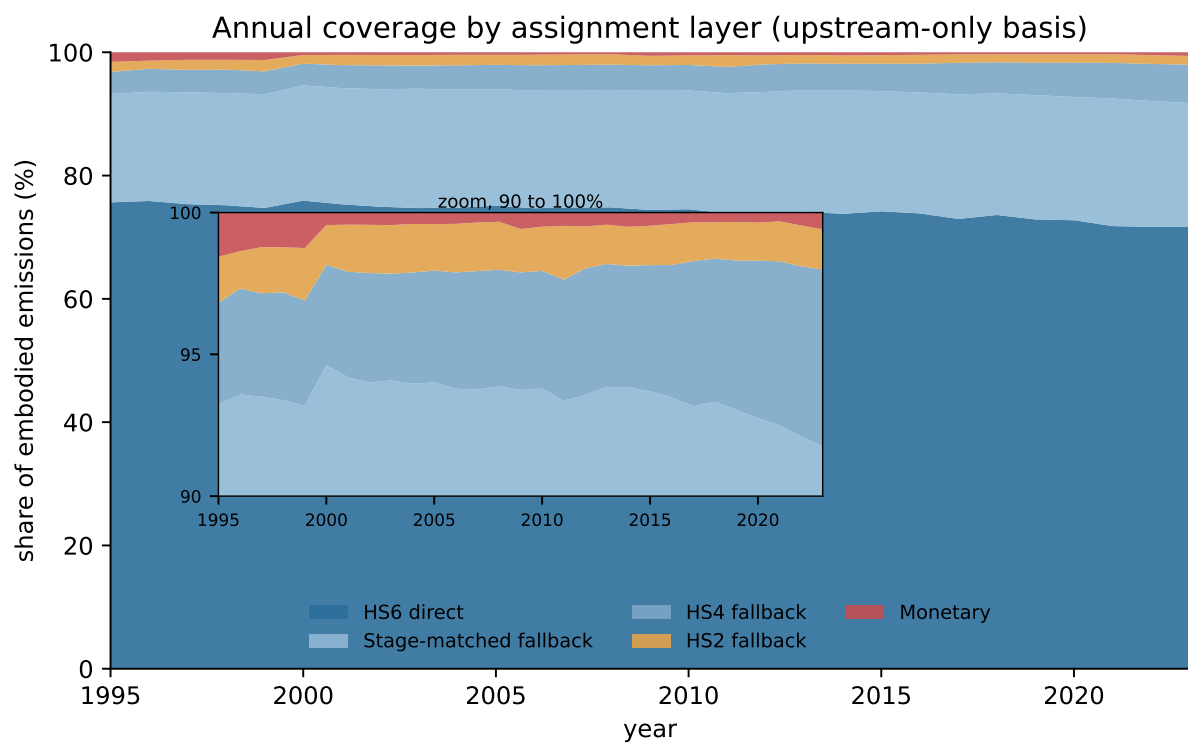


Figure S4: **Annual layer coverage over time.** Annual shares of assigned goods emissions and trade value by assignment layer. The dominance of direct HS6 physical matching is stable throughout 1995–2023.

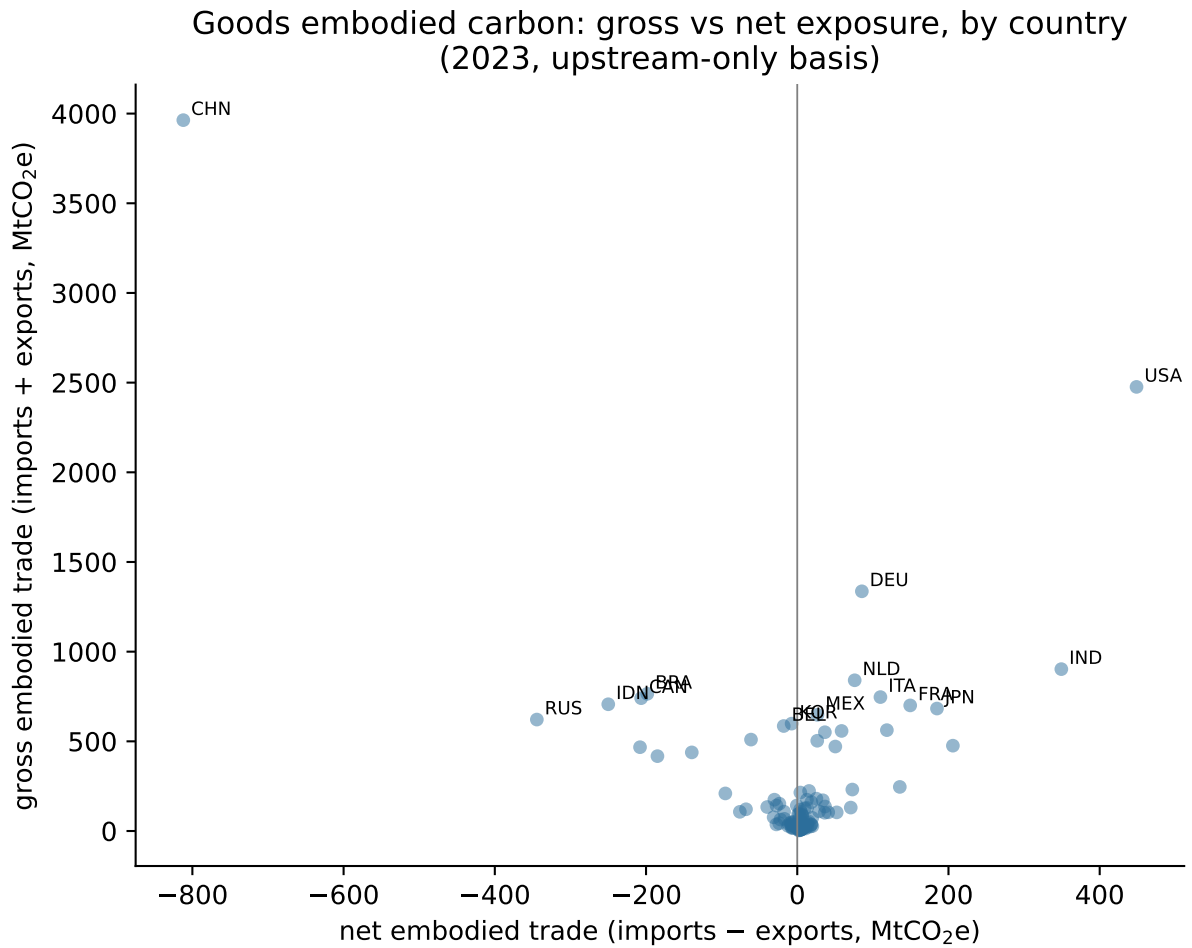


Figure S5: **Gross and net goods embodied carbon exposure.** Net embodied trade (imports minus exports of embodied carbon) on one axis versus gross embodied trade (imports plus exports) on the other, by country. Economies far from the diagonal combine large two-way flows with small net positions; these remain relevant for carbon pricing and supply-chain measures even where net exposure is near zero.

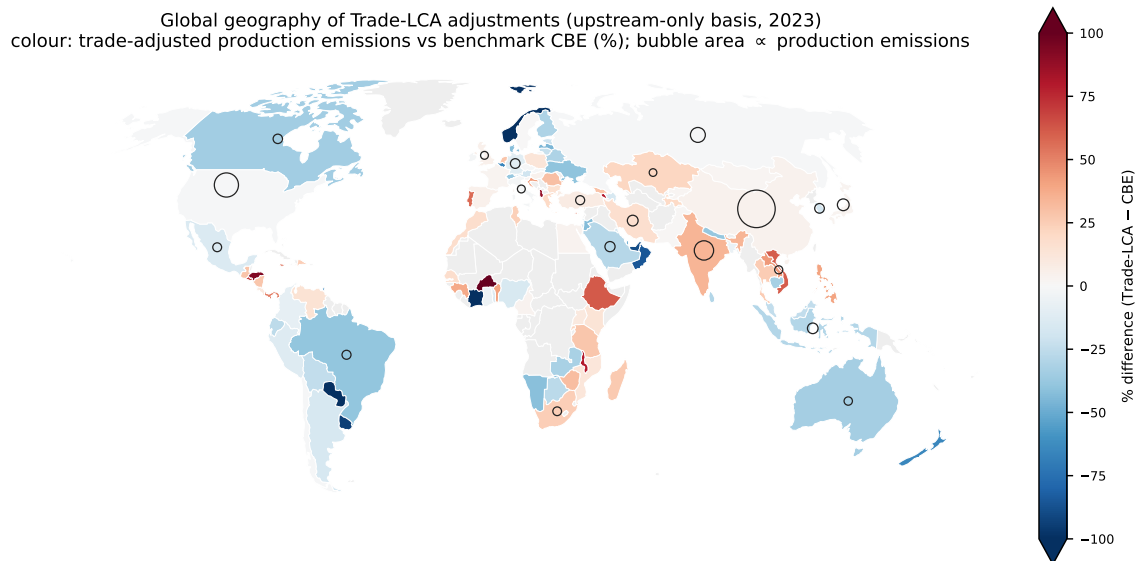


Figure S6: **Global geography of Trade-LCA adjustments.** Colours show the percentage difference between country-level trade-adjusted production emissions and benchmark consumption-based emissions. Bubble area is proportional to production-based emissions.

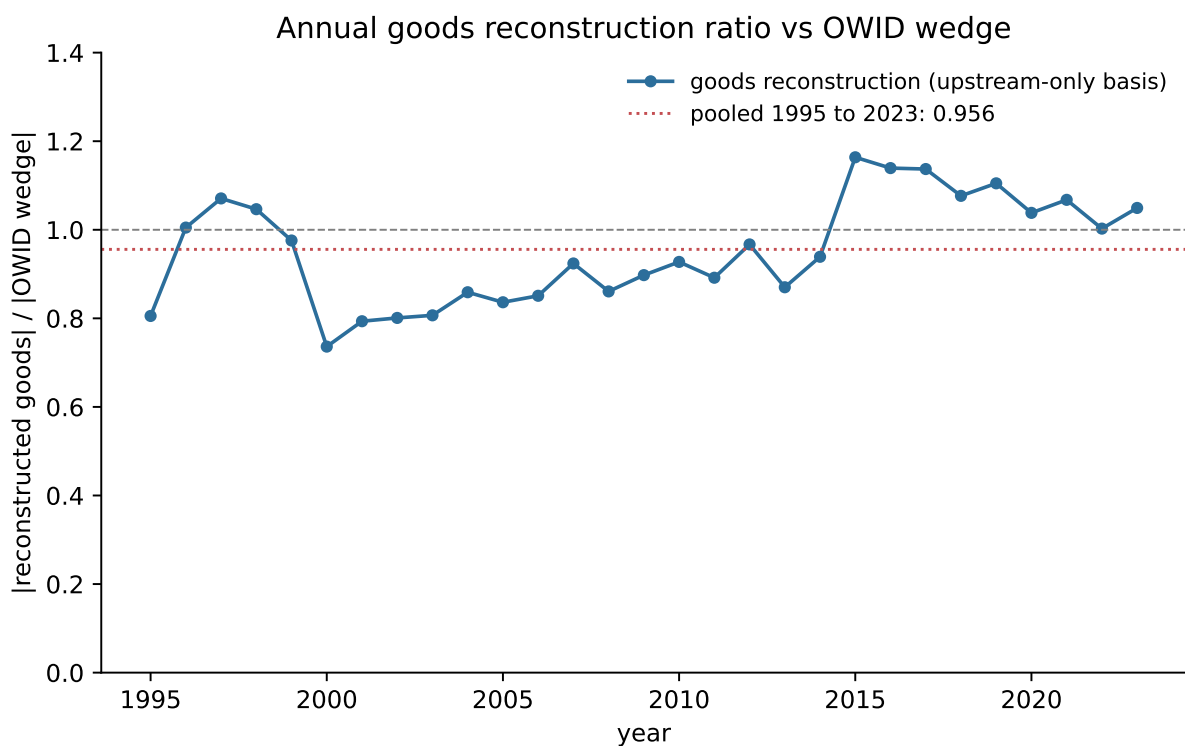


Figure S7: **Annual goods reconstruction against the OWID benchmark wedge.** World-aggregate annual goods net embodied emissions (MtCO_2e), 1995–2023, expressed as a ratio to the OWID CBE–PBE wedge. The goods account tracks the benchmark wedge throughout the panel. Pooled over all country-years it captures 0.956 of the wedge magnitude. The ratio compares sums of absolute magnitudes, so it tests scale and not sign: the reconstructed adjustment agrees in sign with the benchmark for 69.1% of country-years.

References

- Maxime Agez, Richard Wood, Manuele Margni, Anders H. Strømman, Réjean Samson, and Guillaume MajeauBettez. Hybridization of complete plca and mrio databases for a comprehensive product system coverage. *Journal of Industrial Ecology*, 24(4):774790, January 2020. ISSN 1530-9290. doi: 10.1111/jiec.12979. URL <http://dx.doi.org/10.1111/jiec.12979>.
- Mohamed Badr, Konstantin Stadler, and Juudit Ottelin. A comprehensive review of uncertainty in multi-regional inputoutput modelling. *Ecological Modelling*, 519:111649, 2026. ISSN 0304-3800. doi: 10.1016/j.ecolmodel.2026.111649. URL <http://dx.doi.org/10.1016/j.ecolmodel.2026.111649>.
- Abdulrahman Bakindi, Jonathan Norman, Aoife Houlihan Wiberg, and Stephen Allen. Streamlined column-based disaggregation method for carbon footprint quantification for multi-regional inputoutput (mrio) models. *Journal of Industrial Ecology*, 2026. ISSN 1530-9290. doi: 10.1007/s44498-026-00102-3. URL <http://dx.doi.org/10.1007/s44498-026-00102-3>.
- Bharathan Balaji, Venkata Sai Gargeya Vunnava, Nina Domingo, Shikhar Gupta, Harsh Gupta, Geoffrey Guest, and Aravind Srinivasan. Flamingo: Environmental impact factor matching for life cycle assessment with zero-shot machine learning. *ACM Journal on Computing and Sustainable Societies*, 1(2):123, December 2023a. ISSN 2834-5533. doi: 10.1145/3616385. URL <http://dx.doi.org/10.1145/3616385>.
- Bharathan Balaji, Venkata Sai Gargeya Vunnava, Geoffrey Guest, and Jared Kramer. Caml: Carbon footprinting of household products with zero-shot semantic text similarity. In *Proceedings of the ACM Web Conference 2023*, WWW 23, page 40044014. ACM, April 2023b. doi: 10.1145/3543507.3583882. URL <http://dx.doi.org/10.1145/3543507.3583882>.
- Bharathan Balaji, Fahimeh Ebrahimi, Nina Gabrielle G Domingo, Venkata Sai Gargeya Vunnava, Abu-Zaher Faridee, Soma Ramalingam, Shikha Gupta, Anran Wang, Harsh Gupta, Domenic Belcastro, Kellen Axten, Jeremie Hakian, Jared Kramer, Aravind Srinivasan, and Qingshi Tu. Emission factor recommendation for life cycle assessments with generative ai. *Environmental Science & Technology*, 59(18):91139122, March 2025. ISSN 1520-5851. doi: 10.1021/acs.est.4c12667. URL <http://dx.doi.org/10.1021/acs.est.4c12667>.
- Geoffrey Barrows and Hélène Ollivier. Cleaner firms or cleaner products? How product mix shapes emission intensity from manufacturing. *Journal of Environmental Economics and Management*, 88:134–158, 2018. doi: 10.1016/j.jeem.2017.10.008. URL <https://doi.org/10.1016/j.jeem.2017.10.008>.

- Antoine Beylot, Sara Corrado, and Serenella Sala. Environmental impacts of european trade: interpreting results of process-based lca and environmentally extended inputoutput analysis towards hotspot identification. *The International Journal of Life Cycle Assessment*, 25(12):24322450, 2019. ISSN 1614-7502. doi: 10.1007/s11367-019-01649-z. URL <http://dx.doi.org/10.1007/s11367-019-01649-z>.
- Christoph Böhringer, Carolyn Fischer, Knut Einar Rosendahl, and Thomas F. Rutherford. Potential impacts and challenges of border carbon adjustments. *Nature Climate Change*, 12(1):22–29, 2022. doi: 10.1038/s41558-021-01250-z. URL <https://doi.org/10.1038/s41558-021-01250-z>.
- Anatole Boute. Accounting for carbon pricing in third countries under the eu carbon border adjustment mechanism. *World Trade Review*, 23(2):169189, March 2024. ISSN 1475-3138. doi: 10.1017/s1474745624000107. URL <http://dx.doi.org/10.1017/s1474745624000107>.
- Ignacio Cazcarro, Arkaitz Usubiaga-Liaño, Mara Victoria Román, Pablo Piñero, Erik Dietzenbacher, José Manuel Rueda-Cantuche, and Iñaki Arto. Figaro-e3: a high-resolution extended multi-regional input-output database consistent with official statistics. *Scientific Data*, 12(1), April 2025. ISSN 2052-4463. doi: 10.1038/s41597-025-04431-z. URL <http://dx.doi.org/10.1038/s41597-025-04431-z>.
- Andreas Ciroth, Stéphanie Muller, Bo Weidema, and Pascal Lesage. Empirically based uncertainty factors for the pedigree matrix in ecoinvent. *The International Journal of Life Cycle Assessment*, 21(9):13381348, December 2013. ISSN 1614-7502. doi: 10.1007/s11367-013-0670-5. URL <http://dx.doi.org/10.1007/s11367-013-0670-5>.
- Robert H. Crawford. Validation of a hybrid life-cycle inventory analysis method. *Journal of Environmental Management*, 88(3):496–506, 2008. doi: 10.1016/j.jenvman.2007.03.024. URL <https://doi.org/10.1016/j.jenvman.2007.03.024>.
- Steven J. Davis, Andrew Dumit, Mo Li, Yohanna Maldonado, Michael Steffen, Martha Stevenson, Tatiana Boldyreva, and Sangwon Suh. The importance of multiregional accounting for corporate carbon emissions. *Nature Communications*, 17, 2025. doi: 10.1038/s41467-025-67759-5. URL <https://doi.org/10.1038/s41467-025-67759-5>.
- Arjan de Koning, Martin Bruckner, Stephan Lutter, Richard Wood, Konstantin Stadler, and Arnold Tukker. Effect of aggregation and disaggregation on embodied material use of products in inputoutput analysis. *Ecological Economics*, 116:289299, August 2015. ISSN 0921-8009. doi: 10.1016/j.ecolecon.2015.05.008. URL <http://dx.doi.org/10.1016/j.ecolecon.2015.05.008>.

- Erik Dietzenbacher, Bart Los, Robert Stehrer, Marcel Timmer, and Gaaitzen de Vries. The construction of world input-output tables in the WIOD project. *Economic Systems Research*, 25(1):71–98, 2013. doi: 10.1080/09535314.2012.761180. URL <https://doi.org/10.1080/09535314.2012.761180>.
- Neus Escobar, E. Jorge Tizado, Erasmus K.H.J. zu Ermgassen, Pernilla Löfgren, Jan Börner, and Javier Godar. Spatially-explicit footprints of agricultural commodities: Mapping carbon emissions embodied in brazils soy exports. *Global Environmental Change*, 62:102067, May 2020. ISSN 0959-3780. doi: 10.1016/j.gloenvcha.2020.102067. URL <http://dx.doi.org/10.1016/j.gloenvcha.2020.102067>.
- European Commission. Default values for the transitional period of the CBAM between 1 October 2023 and 31 December 2025. European Commission, Directorate-General for Taxation and Customs Union (JRC default-value tables), Brussels, 2023. Accessed via taxation-customs.ec.europa.eu, December 2023.
- Robert C. Feenstra and Gordon H. Hanson. Intermediaries in entrepôt trade: Hong kong reexports of chinese goods. *Journal of Economics & Management Strategy*, 13(1):335, February 2004. ISSN 1530-9134. doi: 10.1111/j.1430-9134.2004.00002.x. URL <http://dx.doi.org/10.1111/j.1430-9134.2004.00002.x>.
- David Font Vivanco. The role of services and capital in footprint modelling. *The International Journal of Life Cycle Assessment*, 25(2):280–293, 2020. doi: 10.1007/s11367-019-01687-7. URL <https://doi.org/10.1007/s11367-019-01687-7>. Published online 2019; volume dated 2020.
- Meredith Fowlie, Claire Petersen, and Mar Reguant. Border carbon adjustments when carbon intensity varies across producers: Evidence from california. *AEA Papers and Proceedings*, 111:401405, May 2021. ISSN 2574-0776. doi: 10.1257/pandp.20211073. URL <http://dx.doi.org/10.1257/pandp.20211073>.
- Guillaume Gaulier and Soledad Zignago. BACI: International Trade Database at the Product-Level. The 1994–2007 Version. *CEPII Working Paper*, (2010-23), 2010. doi: 10.2139/ssrn.1994500. URL <http://www.cepii.fr/CEPII/en/publications/wp/abstract.asp?NoDoc=2726>.
- Stefan Giljum, Hanspeter Wieland, Stephan Lutter, Nina Eisenmenger, Heinz Schandl, and Anne Owen. The impacts of data deviations between mrio models on material footprints: A comparison of exiobase, eora, and icio. *Journal of Industrial Ecology*, 23(4):946958, January 2019. ISSN 1530-9290. doi: 10.1111/jiec.12833. URL <http://dx.doi.org/10.1111/jiec.12833>.

- E.A. Groen and R. Heijungs. Ignoring correlation in uncertainty and sensitivity analysis in life cycle assessment: what is the risk? *Environmental Impact Assessment Review*, 62:98109, January 2017. ISSN 0195-9255. doi: 10.1016/j.eiar.2016.10.006. URL <http://dx.doi.org/10.1016/j.eiar.2016.10.006>.
- Ramana Gudipudi, Luis Costa, Ponraj Arumugam, Matthew Agarwala, Felix Creutzig, and Jürgen P. Kropp. From corporate net-zero pledges to credible climate action. *Nature Sustainability*, 2026. doi: 10.1038/s41893-026-01908-6. URL <https://doi.org/10.1038/s41893-026-01908-6>.
- Geoffrey Hammond, Craig Jones, and Circular Ecology. Inventory of carbon and energy (ICE) database, version 4.1 (educational). Circular Ecology, University of Bath, 2025. Embodied carbon of construction and industrial materials; compiled from peer-reviewed and industry sources independent of the process-LCA backbone used here.
- Edgar Hertwich, Niko Heeren, Brandon Kuczenski, Guillaume MajeauBettez, Rupert J. Myers, Stefan Pauliuk, Konstantin Stadler, and Reid Lifset. Nullius in verba1: Advancing data transparency in industrial ecology. *Journal of Industrial Ecology*, 22(1):617, January 2018. ISSN 1530-9290. doi: 10.1111/jiec.12738. URL <http://dx.doi.org/10.1111/jiec.12738>.
- Keiichiro Kanemoto, Daniel Moran, and Edgar G. Hertwich. Mapping the carbon footprint of nations. *Environmental Science & Technology*, 50(19):10512–10517, 2016. doi: 10.1021/acs.est.6b03227. URL <https://doi.org/10.1021/acs.est.6b03227>.
- Derek Kellenberg and Arik Levinson. Misreporting trade: Tariff evasion, corruption, and auditing standards. *Review of International Economics*, 27(1):106129, 2018. ISSN 1467-9396. doi: 10.1111/roie.12363. URL <http://dx.doi.org/10.1111/roie.12363>.
- Aleksandra Kim, Christopher Mutel, and Stefanie Hellweg. Global sensitivity analysis of correlated uncertainties in life cycle assessment. *Journal of Industrial Ecology*, 29(4): 10901104, May 2025. ISSN 1530-9290. doi: 10.1111/jiec.70036. URL <http://dx.doi.org/10.1111/jiec.70036>.
- Manfred Lenzen. Errors in conventional and input-output-based life-cycle inventories. *Journal of Industrial Ecology*, 4(4):127–148, 2000. doi: 10.1162/10881980052541981. URL <https://doi.org/10.1162/10881980052541981>.
- Manfred Lenzen. Aggregation versus disaggregation in inputoutput analysis of the environment. *Economic Systems Research*, 23(1):7389, March 2011. ISSN 1469-5758. doi: 10.1080/09535314.2010.548793. URL <http://dx.doi.org/10.1080/09535314.2010.548793>.

- Yiwen Liu, Meng Jiang, and Edgar G. Hertwich. Environmental impact of machinery and equipment: A comparison between exiobase, national environmentally extended inputoutput models, and ecoinvent. *Environmental Science & Technology*, 59(50):2725427266, December 2025. ISSN 1520-5851. doi: 10.1021/acs.est.5c08581. URL <http://dx.doi.org/10.1021/acs.est.5c08581>.
- Eva Lyubich, Joseph S. Shapiro, and Reed Walker. Regulating mismeasured pollution: Implications of firm heterogeneity for environmental policy. *AEA Papers and Proceedings*, 108:136142, May 2018. ISSN 2574-0776. doi: 10.1257/pandp.20181089. URL <http://dx.doi.org/10.1257/pandp.20181089>.
- Guilherme Magacho, Etienne Espagne, and Antoine Godin. Impacts of the CBAM on EU trade partners: consequences for developing countries. *Climate Policy*, 24(2):243–259, 2024. doi: 10.1080/14693062.2023.2200758. URL <https://doi.org/10.1080/14693062.2023.2200758>.
- Guillaume Majeau-Bettez, Anders Hammer Strømman, and Edgar G. Hertwich. Evaluation of process- and inputoutput-based life cycle inventory data with regard to truncation and aggregation issues. *Environmental Science & Technology*, 45(23):1017010177, November 2011. ISSN 1520-5851. doi: 10.1021/es201308x. URL <http://dx.doi.org/10.1021/es201308x>.
- Michael A Mehling and Robert A Ritz. From theory to practice: determining emissions in traded goods under a border carbon adjustment. *Oxford Review of Economic Policy*, 39(1):123133, January 2023. ISSN 1460-2121. doi: 10.1093/oxrep/grac043. URL <http://dx.doi.org/10.1093/oxrep/grac043>.
- Stefano Merciai and Jannick Schmidt. Methodology for the construction of global multiregional hybrid supply and use tables for the exiobase v3 database. *Journal of Industrial Ecology*, 22(3):516531, December 2017. ISSN 1530-9290. doi: 10.1111/jiec.12713. URL <http://dx.doi.org/10.1111/jiec.12713>.
- Jan C. Minx, Thomas Wiedmann, Richard Wood, Glen P. Peters, Manfred Lenzen, Anne Owen, Kate Scott, John Barrett, Klaus Hubacek, Giovanni Baiocchi, Arne Paul, Edmond Dawkins, Jon Briggs, Dabo Guan, Sangwon Suh, and Frank Ackerman. Input-output analysis and carbon footprinting: an overview of applications. *Economic Systems Research*, 21(3):187–216, 2009. doi: 10.1080/09535310903541298. URL <https://doi.org/10.1080/09535310903541298>.
- Daniel Moran and Richard Wood. Convergence between the Eora, WIOD, EXIOBASE, and OpenEU’s consumption-based carbon accounts. *Economic Systems Research*, 26(3):245–261, 2014. doi: 10.1080/09535314.2014.935298. URL <https://doi.org/10.1080/09535314.2014.935298>.

- Karsten Neuhoﬀ, Misato Sato, Fernanda Ballesteros, Christoph Böhringer, Simone Borghesi, Aaron Cosbey, Thibault Deletombe, Balázs Felsmann, Roland Ismer, Angus Johnston, Pedro Linares, Sini Matikainen, Stefan Pauliuk, Alice Pirlot, Philippe Quirion, Knut Einar Rosendahl, Aleksander Sniegocki, Harro van Asselt, and Lars Zetterberg. Industrial decarbonization in a fragmented world: Carbon pricing with border adjustments using standardized values. *Energy Policy*, 216:115405, 2026. ISSN 0301-4215. doi: 10.1016/j.enpol.2026.115405. URL <http://dx.doi.org/10.1016/j.enpol.2026.115405>.
- OECD. Effective carbon rates 2023: Pricing greenhouse gas emissions through taxes and emissions trading systems. OECD Publishing, Paris, 2023. URL <https://doi.org/10.1787/b84d5b36-en>.
- Stefan Pauliuk, Guillaume MajeauBettez, Christopher L. Mutel, Bernhard Steubing, and Konstantin Stadler. Lifting industrial ecology modeling to a new level of quality and transparency: A call for more transparent publications and a collaborative open source software framework. *Journal of Industrial Ecology*, 19(6):937949, 2015. ISSN 1530-9290. doi: 10.1111/jiec.12316. URL <http://dx.doi.org/10.1111/jiec.12316>.
- Jessica Perkins and Sangwon Suh. Uncertainty implications of hybrid approach in lca: Precision versus accuracy. *Environmental Science & Technology*, 53(7):36813688, March 2019. ISSN 1520-5851. doi: 10.1021/acs.est.9b00084. URL <http://dx.doi.org/10.1021/acs.est.9b00084>.
- Justin R. Pierce and Peter K. Schott. Concoring u.s. harmonized system categories over time. *Finance and Economics Discussion Series*, 2012.0(16):113, January 2012. ISSN 1936-2854. doi: 10.17016/feds.2012.16. URL <http://dx.doi.org/10.17016/feds.2012.16>.
- Francesco Pomponi and Manfred Lenzen. Hybrid life cycle assessment (lca) will likely yield more accurate results than process-based lca. *Journal of Cleaner Production*, 176:210215, March 2018. ISSN 0959-6526. doi: 10.1016/j.jclepro.2017.12.119. URL <http://dx.doi.org/10.1016/j.jclepro.2017.12.119>.
- J. Poore and T. Nemecek. Reducing foods environmental impacts through producers and consumers. *Science*, 360(6392):987992, 2018. ISSN 1095-9203. doi: 10.1126/science.aaq0216. URL <http://dx.doi.org/10.1126/science.aaq0216>.
- Nathan Preuss, Abdulelah S. Alshehri, and Fengqi You. Large language models for life cycle assessments: Opportunities, challenges, and risks. *Journal of Cleaner Production*, 466:142824, August 2024. ISSN 0959-6526. doi: 10.1016/j.jclepro.2024.142824. URL <http://dx.doi.org/10.1016/j.jclepro.2024.142824>.

- João F. D. Rodrigues, Daniel Moran, Richard Wood, and Paul Behrens. Uncertainty of Consumption-Based Carbon Accounts. *Environmental Science & Technology*, 52:7577–7586, 2018. doi: 10.1021/acs.est.8b00632. URL <https://doi.org/10.1021/acs.est.8b00632>.
- Misato Sato. Product level embodied carbon flows in bilateral trade. *Ecological Economics*, 105:106117, 2014a. ISSN 0921-8009. doi: 10.1016/j.ecolecon.2014.05.006. URL <http://dx.doi.org/10.1016/j.ecolecon.2014.05.006>.
- Misato Sato. Embodied carbon in trade: a survey of the empirical literature. *Journal of Economic Surveys*, 28(5):831–861, 2014b.
- Simon Schulte, Arthur Jakobs, and Stefan Pauliuk. Estimating the uncertainty of the greenhouse gas emission accounts in global multi-regional inputoutput analysis. *Earth System Science Data*, 16(6):26692700, 2024. ISSN 1866-3516. doi: 10.5194/essd-16-2669-2024. URL <http://dx.doi.org/10.5194/essd-16-2669-2024>.
- Konstantin Stadler, Richard Wood, Tatyana Bulavskaya, Carl-Johan Södersten, Moana Simas, Sarah Schmidt, Arkaitz Usubiaga, José Acosta-Fernández, Jeroen Kuenen, Martin Bruckner, Stefan Giljum, Stephan Lutter, Stefano Merciai, Jannick H. Schmidt, Michaela C. Theurl, Christoph Plutzar, Thomas Kastner, Nina Eisenmenger, Karl-Heinz Erb, Hein de Vries, and Arnold Tukker. EXIOBASE 3: developing a time series of detailed environmentally extended multi-regional input–output tables. *Journal of Industrial Ecology*, 22(3):502–515, 2018. doi: 10.1111/jiec.12715. URL <https://doi.org/10.1111/jiec.12715>.
- Kjartan Steen-Olsen, Anne Owen, Edgar G. Hertwich, and Manfred Lenzen. Effects of sector aggregation on co₂ multipliers in multiregional inputoutput analyses. *Economic Systems Research*, 26(3):284302, 2014. ISSN 1469-5758. doi: 10.1080/09535314.2014.934325. URL <http://dx.doi.org/10.1080/09535314.2014.934325>.
- Bernhard Steubing, Arjan de Koning, Stefano Merciai, and Arnold Tukker. How do carbon footprints from lca and eeioa databases compare? a comparison of ecoinvent and exiobase. *Journal of Industrial Ecology*, 26(4):14061422, May 2022. ISSN 1530-9290. doi: 10.1111/jiec.13271. URL <http://dx.doi.org/10.1111/jiec.13271>.
- Bin Su, H. C. Huang, B. W. Ang, and P. Zhou. Input–output analysis of CO₂ emissions embodied in trade: the effects of sector aggregation. *Energy Economics*, 32(1):166–175, 2010. doi: 10.1016/j.eneco.2009.07.010. URL <https://doi.org/10.1016/j.eneco.2009.07.010>.
- Sangwon Suh, Manfred Lenzen, Graham J. Treloar, Hiroki Hondo, Arpad Horvath, Gjalt Huppes, Olivier Jolliet, Uwe Klann, Wolfram Krewitt, Yuichi Moriguchi, Jesper Munks-

- gaard, and Greg Norris. System boundary selection in life-cycle inventories using hybrid approaches. *Environmental Science & Technology*, 38(3):657–664, 2004. doi: 10.1021/es0263745. URL <https://doi.org/10.1021/es0263745>.
- Qingshi Tu, Jing Guo, Nan Li, Jianchuan Qi, and Ming Xu. Mitigating grand challenges in life cycle inventory modeling through the applications of large language models. *Environmental Science & Technology*, 58(44):1959519603, October 2024. ISSN 1520-5851. doi: 10.1021/acs.est.4c07634. URL <http://dx.doi.org/10.1021/acs.est.4c07634>.
- Arnold Tukker, Arjan de Koning, Anne Owen, Stephan Lutter, Martin Bruckner, Stefan Giljum, Konstantin Stadler, Richard Wood, and Rutger Hoekstra. Towards robust, authoritative assessments of environmental impacts embodied in trade: Current state and recommendations. *Journal of Industrial Ecology*, 22(3):585598, February 2018. ISSN 1530-9290. doi: 10.1111/jiec.12716. URL <http://dx.doi.org/10.1111/jiec.12716>.
- Ilke Van Beveren, Andrew Bernard, and Hylke Vandenbussche. *Concording EU Trade and Production Data over Time*. December 2012. doi: 10.3386/w18604. URL <http://dx.doi.org/10.3386/w18604>.
- Hauke Ward, Leonie Wenz, Jan C. Steckel, and Jan C. Minx. Truncation error estimates in process life cycle assessment using inputoutput analysis. *Journal of Industrial Ecology*, 22(5):10801091, 2017. ISSN 1530-9290. doi: 10.1111/jiec.12655. URL <http://dx.doi.org/10.1111/jiec.12655>.
- Thomas Wiedmann and Manfred Lenzen. Environmental and social footprints of international trade. *Nature Geoscience*, 11(5):314–321, 2018. doi: 10.1038/s41561-018-0113-9. URL <https://doi.org/10.1038/s41561-018-0113-9>.
- Yi Yang, Reinout Heijungs, and Miguel Brandão. Hybrid life cycle assessment (lca) does not necessarily yield more accurate results than process-based lca. *Journal of Cleaner Production*, 150:237242, May 2017. ISSN 0959-6526. doi: 10.1016/j.jclepro.2017.03.006. URL <http://dx.doi.org/10.1016/j.jclepro.2017.03.006>.
- Da Zhang, Justin Caron, and Niven Winchester. Sectoral aggregation error in the accounting of energy and emissions embodied in trade and consumption. *Journal of Industrial Ecology*, 23(2):402411, February 2018. ISSN 1530-9290. doi: 10.1111/jiec.12734. URL <http://dx.doi.org/10.1111/jiec.12734>.
- Bu Zhao, Jitong Jiang, Ming Xu, and Qingshi Tu. A datacentric investigation on the challenges of machine learning methods for bridging life cycle inventory data gaps. *Journal of Industrial Ecology*, 29(3):955966, April 2025. ISSN 1530-9290. doi: 10.1111/jiec.70022. URL <http://dx.doi.org/10.1111/jiec.70022>.