

Coding ledger: replication-package classifications in the 2023 and 2025 ASR audits

Prepared by Luke Ford. All repository checks performed August 24, 2026.

Purpose

This ledger recodes the quantitative articles in two of Philip N. Cohen's audits of the *American Sociological Review* under a single stated rule, so that the 2023 and 2025 figures can be compared on a constant standard. It is sent to Cohen before publication so that any classification he disputes can be corrected first.

The two rules being compared

Cohen's 2023 criterion, in his words, was "data and code sufficient to replicate the papers (at a glance; I didn't test their code)." He classified eight of twenty-five quantitative articles as replication packages and separately identified papers that provided something less.

Cohen's 2025 criterion, in his words, counted articles that "provide packages of code and data, or at least information about getting the data." He counted material hosted on authors' personal websites and excluded material described as available on request. Twelve of eighteen quantitative articles met that standard.

The second is broader than the first. This ledger applies one rule to both years.

The harmonized rule

Count a quantitative article as having a full replication package when publicly accessible analysis code appears, without executing it, to cover the article's quantitative results, and the underlying data are either included or obtainable by eligible third-party researchers through a documented procedure administered independently of the article's authors.

Do not count data whose availability depends on the discretion of the original authors.

Classify separately, and do not count in the numerator:

- Packages covering only part of an article's analyses. Ground: coverage.
- Data reachable only at the authors' discretion. Ground: access.
- Materials not publicly accessible on the audit date. Ground: accessibility.

Restricted data are not disqualifying. The question is whether a third party has a documented route to them. Six cases show the line being drawn. Lersch, Haupt, Engzell and Mood, Mize and Kincaid, Zhao, and Ferry et al. all supply code without bundled data, and all pass, because each identifies an archive with a published application procedure open to outside researchers. Brady, Luft and Zuckerman Sivan and Smith et al. also supply code without data, and both fail, because in each the data reach the reader only if the original authors or a third party grant permission case by case.

No code or replication package was executed in producing any classification below. This is an availability and apparent-completeness audit, matching Cohen's own stated method.

Part one: recoding Cohen's eight 2023 positives

All eight pass under the harmonized rule. The 2023 numerator stands at 8 of 25.

1. Best and Arseniev-Koehler, "The Stigma of Diseases." Cohen: Full package. Harmonized: pass. The article's data-availability statement identifies replication files and instructions for accessing the data on GitHub and OSF. Source: <https://journals.sagepub.com/doi/10.1177/00031224231197436>
2. Biegert, Kühhirt and Van Lancker, "They Can't All Be Stars." Cohen: Full package. Harmonized: pass. The article states that replication files are deposited at OSF, DOI 10.17605/OSF.IO/NTWDY. Van Lancker's data page identifies Stata do-files, CSV files, and a README. Source: <https://journals.sagepub.com/doi/10.1177/00031224231159139>
3. Lersch, "Change in Personal Culture over the Life Course." Cohen: Full package. Harmonized: pass on the documented-route criterion. The article states that computer code for the complete analysis is at OSF, DOI 10.17605/OSF.IO/UEY5H, and identifies the six underlying panel surveys and their institutional sources. The data are restricted; the route is documented. Source: <https://journals.sagepub.com/doi/10.1177/00031224231156456>
4. Engzell and Mood, "Understanding Patterns and Trends in Income Mobility through Multiverse Analysis." Cohen: Full package. Harmonized: pass on the documented-route criterion. Engzell's own data page labels the associated OSF project as replication material, and the project is public. The underlying Swedish register data are accessible through the holding agencies rather than from the authors. Sources: <https://journals.sagepub.com/doi/10.1177/00031224231180607> and <https://perengzell.com/datacode/>
5. Cha, Weeden and Schnabel, "Is the Gender Wage Gap Really a Family Wage Gap in Disguise?" Cohen: Full package. Harmonized: pass. The article's data statement points to OSF replication materials. In Cohen's own comment thread Weeden states that the package runs from the raw public-use SIPP file through the article, footnotes, and supplement, and Cohen then added the paper to his count. Source: <https://journals.sagepub.com/doi/10.1177/00031224231212464>
6. Huang and Butts, "Rooted America." Cohen: Full package. Harmonized: pass. The article's data statement identifies replication data and code in Harvard Dataverse, DOI 10.7910/DVN/I7HT9T. Source: <https://journals.sagepub.com/doi/10.1177/00031224231212679>
7. Haupt, "Who Profits from Occupational Licensing?" Cohen: Full package. Harmonized: pass on the documented-route criterion. The article's data statement links syntax files and the U.S. occupational data at OSF. The German component, the 2018 BIBB/BAuA survey, is available as a Scientific Use File through BIBB and GESIS under a published application procedure. Source: <https://journals.sagepub.com/doi/10.1177/00031224231207395>
8. Zhang and Johnson, "Hierarchies in the Decentralized Welfare State." Cohen: Full package. Harmonized: pass. The Harvard Dataverse deposit is public, titled as replication code and data, and contains R scripts, data, a README, and acquisition code for auxiliary sources. Cohen coded this as a full package after Zhang pointed him to it. Source: <https://doi.org/10.7910/DVN/03YYII>

Part two: the eighteen 2025 quantitative articles

Order follows Cohen's table. Eight pass. The numerator under the harmonized rule is 8 of 18.

1. O'Brien, Schechtl and Parolin, "Fiscal Centralization and Inequality in Children's Economic Mobility." Cohen: Nothing. Harmonized: fail. Independently checked. The journal record carries no data-availability statement. The single supplementary file runs eighteen pages of additional figures, tables, and robustness analyses and contains no analysis code. Title and author searches against OSF, Harvard Dataverse, GitHub, openICPSR, Zenodo, and Figshare located no qualifying material as of August 24, 2026.
2. Sendroiu, Álvarez-Benjumea and Winter, "Time's Up? How Temporal Maps of Climate Change Shape Climate Action." Cohen: Package, OSF. Harmonized: pass. The article states that all data and materials are at OSF. Source: <https://osf.io/jsypa/>. Checked August 24, 2026.
3. Mahutga, Gao and Pandian, "Global Value Chains and Union Decline in Rich Democracies." Cohen: Nothing. Harmonized: fail. Independently checked. The journal record carries no data-availability statement. The seven-page supplement holds additional regression tables and no code. Mahutga maintains a public data page that does not list materials for this article. No qualifying material located as of August 24, 2026.
4. Mize and Kincaid, "Role-Accumulation and Mental Health across the Life Course." Cohen: Package, author Google Drive. Harmonized: pass on the documented-route criterion. The article states that the replication files reproduce all results, and the author's page links the replication folder. HILDA is not open-download data, but the Australian Department of Social Services operates a formal application procedure for approved researchers, including overseas applicants for the General Release. Sources: <https://journals.sagepub.com/doi/10.1177/00031224241313394> and https://www.trentonmize.com/research/mize_kincaid_asr_2025. Checked August 24, 2026.
5. O'Brien and Noy, "Religion, Perceptions of Scientists' Moral Culture, and Support for Science in the United States." Cohen: Nothing, data available from the author on reasonable request. Harmonized: fail. Ground: access. Availability depends on the author's discretion.
6. Silva, Bloemraad and Voss, "Frame Backfire." Cohen: Package, OSF. Harmonized: fail. Ground: accessibility. Cohen's table records that the OSF project required an access request when he checked it, and timestamps his request December 9, 2025 at 12:30 p.m. The barrier persists. An anonymous request to the project URL on August 25, 2026 returned HTTP 403, and a Wayback Machine capture made August 26, 2026 shows the URL redirecting to OSF's sign-in page rather than to a project record. Neither establishes the project's permission settings; both establish that a reader following the article's link without an OSF account does not reach the materials. The Internet Archive recorded that capture as the first ever made of this URL, so no snapshot predating Cohen's check exists. Counting this article anyway, on the view that accessibility should be disregarded, produces the 9 of 18 upper bound. Sources: <https://journals.sagepub.com/doi/10.1177/00031224251333087>, <https://osf.io/amxdb/>, and the capture at <https://web.archive.org/web/20260826024652/https://osf.io/amxdb/>
7. Abraham, Botelho and Carter, "(Not) Getting What You Deserve." Cohen: Package, OSF. Harmonized: fail. Ground: coverage. The data statement says the deposited materials cover the main results of

Studies 2 and 3. The article contains Study 1, Studies 2 and 3, and two further generalizability studies. Data access is not the issue. Source:

<https://journals.sagepub.com/doi/10.1177/00031224251318051>

8. Aksoy and Szekely, "Making Sense of Honor Killings." Cohen: Package, OSF with a route to Pew data. Harmonized: pass. The data statement identifies analysis code, data, and instructions for obtaining the publicly available source data. Source: <https://osf.io/smjnu/>. Checked August 24, 2026.
9. Zhao, "Uneven Mixing, Network Segregation, and Immigrant Integration." Cohen: Package, OSF with data from source. Harmonized: pass on the documented-route criterion. Code at <https://osf.io/5qrjg>. The CILS4EU data are archived at GESIS, with the full version available through the Secure Data Center and reduced versions under a published off-site procedure. Checked August 24, 2026.
10. Cole, "Abortion Policies in a Polarizing World Society, 1970 to 2020." Cohen: Nothing, data and code upon request. Harmonized: fail. Ground: access.
11. Kim, "Culture and Immigrant Selectivity in Shaping Asian American Education." Cohen: Package, OSF with a route to Census data. Harmonized: pass. Stata replication code at <https://osf.io/adwzn>, with the raw data identified as available through IPUMS USA. Checked August 24, 2026.
12. Rohr and Martin, "The State and the Emergence of the First American Party System." Cohen: Package, Dataverse. Harmonized: pass. The article identifies a replication package containing data and code at Harvard Dataverse, DOI 10.7910/DVN/JLUVYU. Checked August 24, 2026.
13. Jiang, "The Cultural Devaluation of Feminized Work." Cohen: Package, GitHub. Harmonized: pass. The repository is public and contains intermediate data and the analysis code, including the merged dataset for the principal analyses. The README identifies public routes to the Google Ngram data and purchase routes for the copyrighted COHA and COCA corpora. Source: <https://github.com/wenhaojiangsoc/devaluation>. Checked August 24, 2026.
14. Hsu and Sharkey, "Same Policy, No Standardized Outcome." Cohen: Nothing. Harmonized: fail. Independently checked. The journal record carries neither a data-availability statement nor a replication supplement. Title and author searches against the six repositories located no qualifying material as of August 24, 2026.
15. Smith, Kennard, Du and McFarland, "How Values and Uncertainty Shape Scientific Advance in Peer Review." Cohen: Package, GitHub, noted as not including data. Harmonized: fail. Ground: access. The article describes a package permitting independent verification of the machine-learning benchmarks and full replication of the main inferential analyses. The repository README states the opposite on the first point: that users cannot independently reproduce the benchmarks, that the code generating bias metrics is excluded, that the package contains no data, and that the anonymized dataframes, model weights, and some underlying datasets must be requested from the authors, with further third-party approval required in places. The classification here follows the repository. The discrepancy between the two descriptions is recorded rather than resolved. A Wayback Machine capture made August 26, 2026 shows the repository public, carrying three commits: a licence file from roughly two years earlier, and the three analysis folders and README added about ten months earlier

under the tag v1, near the article's October 2025 publication. Nothing had been added in the eight months since Cohen named it. Sources:

<https://journals.sagepub.com/doi/10.1177/00031224251362254>,

<https://github.com/danielscottsmith/valuesInScience>, and the capture at <https://web.archive.org/web/20260826024928/https://github.com/danielscottsmith/valuesInScience>

16. Brady, Luft and Zuckerman Sivan, "How Does Culture Matter for Attainment, and How Would We Know If It Did?" Cohen: Package, author website, code as PDF. Harmonized: fail. Ground: access. The file is a thirteen-page printout of Stata code dated December 20, 2024, captured in the Wayback Machine on August 26, 2026. Line 6 states that because the authors of the original study shared potentially confidential data, replication data cannot be provided to others. Lines 9 and 30 carry author-specific Windows paths in place of file locations, and line 29 merges against a dataset not supplied, so an outside reader could not execute it even with the data in hand. Code without bundled data is not itself disqualifying. Lersch, Engzell and Mood, Haupt, Mize and Kincaid, Zhao, and Ferry et al. all pass on that pattern, because in each the restricted data have a published institutional access route. Sources: <https://bradydave.wordpress.com/wp-content/uploads/2024/12/replication-code-bradyluftzuckerman.pdf>, the capture at <https://web.archive.org/web/20260826025144/https://bradydave.wordpress.com/wp-content/uploads/2024/12/replication-code-bradyluftzuckerman.pdf>, and <https://journals.sagepub.com/doi/10.1177/00031224251344634>
17. Ferry, Bouchet-Valat, Drouhot, Ichou and Obuina, "Shades of Égalité." Cohen: Package, OSF with links to data. Harmonized: pass on the documented-route criterion. R code at DOI 10.17605/OSF.IO/NZAVQ. The article supplies two institutional routes to the TeO2 survey, the CASD secure access center and a downloadable version in the Quetelet Progedo archive. Checked August 24, 2026.
18. Marteleto, Kumar, Sereno and Gori Maia, "Who Can Have a Baby? Social Norms and the Right to Reproduce." Cohen: Nothing, data available upon request. Harmonized: fail. Ground: access. The published article carries a Data Note on page 24 stating that data are available upon request. Repository searches located no independent public package as of August 24, 2026. Source: <https://journals.sagepub.com/doi/pdf/10.1177/00031224251387273>

Results

Under the harmonized rule:

2023: 8 of 25, or 32.0 percent.

2025, strict: 8 of 18, or 44.4 percent. Two-sided Fisher exact against 8 of 25, $p = .526$.

2025, counting Silva despite an access barrier documented on both the audit date and August 2026: 9 of 18, or 50.0 percent. $p = .344$.

2025, under the rule stated in the 2025 audit: 12 of 18, or 66.7 percent. $p = .033$.

The reported rise from 2023 to 2025 is 34.7 percentage points. On a constant rule it is 12.4 points. Roughly two-thirds of the reported increase follows from the change in what was counted.

The p-values are calculations from the tables above, not figures drawn from any source. They are offered as an exploratory comparison. These are complete annual volumes rather than random samples, so the sampling interpretation is imperfect. The point does not depend on any threshold: the same observed comparison moves from $p = .033$ to $p = .526$ when the rule is held constant.

The four disputed cases can be added back one at a time. Granting one gives 9 of 18, $p = .344$. Granting two gives 10 of 18, $p = .210$. Granting three gives 11 of 18, $p = .071$. Only when all four are counted does the comparison reach 12 of 18 and $p = .033$. The reported result therefore requires every one of the broadened classifications, and does not survive the loss of a single one.

Limitations

The six 2025 articles Cohen classified as Nothing were independently rechecked on August 24, 2026, and no qualifying public replication package was located for any of them. Three are affirmatively resolved from the articles themselves: O'Brien and Noy, Cole, and Marteleto et al. each publish a statement making materials available on request from the authors, which fails the rule on its face. The remaining three rest on negative searches and should be read as such.

Negative searches carry a known limit. In the 2023 volume Cohen classified Cheng et al. as Nothing while a Harvard Dataverse deposit associated with the article's appendix had been public since August 11, 2022, DOI 10.7910/DVN/CQCKBF, and the article did not link to it. Audits of this kind detect what an article and its links reveal, and this one shares that limit with Cohen's.

Three of the six, O'Brien, Schechtl and Parolin, Mahutga et al., and Hsu and Sharkey, carry no data-availability statement of any kind. Whatever accounts for the change between 2023 and 2025, a uniformly enforced journal requirement to publish such a statement does not appear to be operating across Volume 90.

Repository checks were made on August 24, 2026. Except where Cohen recorded a contemporaneous observation, as with Silva, this ledger cannot establish what any repository looked like on December 9, 2025. The Wayback Machine holds no capture of the Silva project predating that date.

Two of the four cases behind the difference between 12 of 18 and 8 of 18 have been examined in their primary form rather than through a repository listing. The Brady, Luft and Zuckerman Sivan file is a thirteen-page printout of Stata code dated December 20, 2024. Beyond the absence of data, it carries author-specific Windows paths in place of file locations and merges against a dataset not supplied, so it could not be executed by an outside reader even with the data in hand. The Smith et al. repository holds three commits across three folders and a README stating the package contains no data.

No code was executed. Classifications rest on inspection of articles, data-availability statements, repositories, and archive access procedures.

Request

If any classification here is wrong, it will be corrected before publication. The four entries that produce the difference between 12 of 18 and 8 of 18 are numbers 6, 7, 15, and 16.