

The 2020 Census, Differential Privacy, and the Integrity of Representation

An Evidence-Based Examination of Enumeration Accuracy, Disclosure Avoidance, Geographic Allocation, and Citizenship-Neutral Apportionment

Anthony Lewis Jerdine

Institutional Research Monograph

Miami Wealth Capital

Version 1.0 — ADOPTED — CANONICAL

20 August 2026

Controlling principle: *The integrity of representation depends not only on counting people, but on preserving a transparent chain from enumeration to legally consequential public data. Enumeration accuracy, privacy protection, geographic allocation, and citizenship-based analysis are related, but they are not interchangeable.*

Editorial and Evidentiary Note

This monograph expands and critically evaluates the Center for Renewing America paper **The 2020 Census Fraud: How Differential Privacy Skewed Representation** by Wade Miller and Andrew White.[1] The CRA paper raises legitimate questions about documented coverage errors, local geographic utility, group-quarters disruptions, correction mechanisms, and the political consequences of census data. It also advances claims of fraud, partisan manipulation, and citizenship concealment that are not established by the evidence presently available.

The purpose of this monograph is therefore not to dilute the accusation or to adopt it uncritically. It is to force every proposition into the correct evidentiary category: **established finding, supported inference, open question, or unsupported allegation**. That distinction is essential in any work intended to survive technical, legal, congressional, and public scrutiny.

The record establishes that the 2020 Census had statistically significant state-level coverage errors in fourteen states, that COVID-19 created serious operational disruption, and that the Census Bureau's Disclosure Avoidance System altered detailed public tabulations through a differential-privacy framework.[2][3] Peer-reviewed research found that the Bureau's demonstration implementation could introduce consequential error into small-area redistricting analysis, including systematic effects in racially heterogeneous and mixed-partisan precincts.[4]

The record does **not**, on its own, establish that differential privacy caused the state-level enumeration errors, that the Bureau designed the system to hide noncitizens, or that officials intentionally manipulated census data for partisan advantage. Those are allegations that require evidence of causal mechanism and intent.

A further current development matters for 2030 planning: the Census Bureau now marks its earlier differential-privacy guidance as no longer current and states that it is evaluating alternative approaches in response to a Department of Commerce administrative order prohibiting noise infusion.[6] That policy shift does not retroactively prove wrongdoing in 2020, but it makes the governance lessons of the 2020 system immediately relevant to the next decennial census.

Abstract

The decennial census is the constitutional foundation for apportionment of the U.S. House of Representatives and, indirectly, the Electoral College. It also supplies the geographic population data used by states to draw legislative districts and by governments to allocate public resources. Because the census simultaneously serves constitutional, administrative, statistical, privacy, and redistricting functions, methodological decisions can produce consequences extending far beyond statistical publication.

The 2020 Census encountered extraordinary operational disruptions. The COVID-19 pandemic altered household residence patterns, interrupted field operations, displaced students from group quarters, restricted access to nursing homes and correctional facilities, and complicated nonresponse follow-up. The Census Bureau subsequently reported statistically significant net undercounts in six states and net overcounts in eight states. GAO later concluded that there was no statistically significant national net coverage error while emphasizing significant subnational and demographic coverage problems.[2][3]

The 2020 Census also introduced a Disclosure Avoidance System based on differential privacy. The system was designed to reduce the risk that published statistical tables could be used to reconstruct confidential individual responses. It altered detailed geographic and demographic tabulations while preserving specified higher-level invariants, including state-level apportionment totals. Academic studies found that the system could affect redistricting, voting-rights analysis, racial-geographic measurement, and the interpretation of small-area population data.[4]

The central analytical error in many criticisms of the 2020 Census is the collapse of distinct questions into a single allegation of fraud. Statistical perturbation is not equivalent to fabrication for an illicit purpose; disclosure avoidance is not equivalent to partisan intent; and the inclusion of noncitizens in the apportionment population is not an algorithmic concealment scheme but a consequence of the governing constitutional and statutory framework.[5]

A sound reform agenda must preserve numerical integrity where law requires exactness, strengthen auditability and data lineage, protect confidential responses, improve correction pathways, and publish citizenship-related measures as clearly labeled analytical products rather than confusing them with the constitutional apportionment population.

Executive Summary

The central findings of this monograph are:

- The 2020 Census had documented accuracy problems. The Census Bureau estimated net undercounts in Arkansas, Florida, Illinois, Mississippi, Tennessee, and Texas, and net overcounts in Delaware, Hawaii, Massachusetts, Minnesota, New York, Ohio, Rhode Island, and Utah.[2][3]
- The reported state-level coverage errors were measured through the Post-Enumeration Survey and should not automatically be attributed to differential privacy. The PES evaluates the enumerated population; the Disclosure Avoidance System principally transforms released statistical products.[3]
- Differential privacy materially changed detailed published data, particularly at small geographic levels. This can affect redistricting and voting-rights analysis even when state-level apportionment totals are held invariant.[4]
- Peer-reviewed research found systematic undercounts in mixed-race and mixed-partisan precincts in the demonstration data studied, producing unpredictable partisan and racial effects and creating concerns for one-person, one-vote analysis.[4]

- The evidence presently available does not establish that differential privacy caused the state-level undercounts and overcounts.
- The evidence presently available does not establish that differential privacy was designed to conceal noncitizens or produce partisan advantage.
- Noncitizens were included in the 2020 apportionment population because the governing census framework counts persons residing in the states rather than citizens alone. Differential privacy did not create that rule.[5]
- *Trump v. New York* did not resolve the merits of whether unlawfully present noncitizens may constitutionally be excluded from apportionment; the Supreme Court held that the dispute was not ripe in the posture presented.[5]
- The strongest institutional critique concerns **auditability, documentation, geographic utility, correction authority, independent review, and the separation of legally controlling totals from privacy-protected statistical products.**
- For 2030, Congress and the Department of Commerce should establish a transparent hierarchy of census products: legally controlling population totals, validated redistricting data, privacy-protected demographic tables, citizen and citizen-voting-age measures, uncertainty information, and machine-readable audit records.

Part I — Constitutional Function and Political Consequence

1. The Enumeration Clause and the Allocation of Political Power

Article I, Section 2 requires an actual enumeration at ten-year intervals and gives Congress authority to direct the manner in which the census is conducted. The Fourteenth Amendment replaced the original apportionment formula with representation based on the whole number of persons in each state, subject to its constitutional terms.

The census performs at least three distinct governmental functions:

1. **Apportionment:** establishing the population base used to allocate seats in the House of Representatives among the states.
2. **Redistricting:** supplying granular population data used to construct congressional, state legislative, and local districts.
3. **Statistical administration:** producing demographic and geographic information used for public administration, research, planning, and funding formulas.

These functions overlap but are not identical. A statewide total may be sufficient for apportionment yet inadequate for block-level redistricting. A detailed demographic table may be valuable for voting-rights analysis but still be inappropriate as the legally controlling apportionment record.

The governing principle is therefore functional:

Census accuracy must be evaluated against the governmental use to which a particular dataset is put.

The relevant questions are not simply whether a published cell was altered. They are: which data were altered; at what geographic level; for what legal purpose; which totals were invariant; what uncertainty was introduced; and whether affected governments could independently audit the result.

2. Apportionment Population and Citizenship

The 2020 apportionment population included citizens and noncitizens residing in the states under the applicable census residence rules. That was not an accidental product of differential privacy. It reflected the constitutional and statutory framework used for the certified apportionment count.

The population concepts should remain separate:

Measure	What it generally counts	Principal use
Total resident population	Persons residing in a jurisdiction under census residence rules	Constitutional apportionment and general population analysis
Citizen population	Persons identified as U.S. citizens	Citizenship analysis and policy planning
Voting-age population	Persons age 18 and older	Electoral and demographic analysis
Citizen voting-age population	Citizens age 18 and older	Voting Rights Act and redistricting analysis
Eligible voting population	Persons legally eligible to vote under applicable law	Election administration and turnout analysis

The political argument for citizen-based representation is legitimate as a constitutional-policy question. It should be debated directly. It should not be smuggled into a technical claim that the disclosure-avoidance algorithm created or concealed the underlying legal population definition.

A change from total resident population to citizen-based apportionment would require confronting constitutional text, federal statutes, precedent, administrative feasibility, and consequences for children, mixed-status households, military families, and other categories not captured by simple citizenship labels.

Part II — The Census Operational System

3. Geographic Architecture

The Census Bureau organizes the country through nested census geography: states, counties, tracts, block groups, and blocks. Political geography—voting districts, municipalities, precincts, and legislative districts—does not necessarily nest in the same way.

Census blocks are therefore critical. They are the smallest building blocks from which many political geographies are assembled. A person can be counted in the correct state yet assigned to the wrong block. That would not necessarily change statewide apportionment, but it could affect a precinct, municipality, school district, or legislative map.

For redistricting, **geographic placement is often as consequential as total population.**

4. Address Frames and Housing Units

Before enumeration, the Bureau builds and updates an address frame using Postal Service information, state and local records, tribal sources, field operations, and other administrative data.

Potential failure modes include omitted housing units, duplicates, incorrect geographic placement, vacant units treated as occupied, occupied units treated as vacant, omitted group-quarters facilities, and facilities assigned to the wrong block.

These failure modes should be categorized because they produce different legal and statistical effects. A missing apartment complex can affect the population total. A correctly counted complex placed in the wrong block can affect redistricting without changing the state count. A duplicated address can create duplicate-person risk unless resolved downstream.

5. Household Enumeration and Nonresponse

Households may respond online, by mail, by telephone, or through in-person follow-up. Nonresponse follow-up can rely on direct interviews, proxy responses, administrative records, and imputation.

The system must resolve difficult residence questions involving students, military personnel, seasonal housing, shelters, temporary displacement, multiple submissions, vacancies, and households that cannot be reached.

GAO reported that 2020 involved higher rates of missing and imputed demographic information than 2010 and that proxy responses were a particular quality concern.[3] This is operationally important because an exact privacy mechanism applied to a weak input record cannot repair an enumeration failure that already occurred upstream.

6. Group Quarters

Group quarters include dormitories, nursing facilities, military installations, correctional facilities, shelters, and other institutional or communal settings. The pandemic sharply complicated the residence and reporting environment for many of these facilities.

Relevant disruptions included universities closing dormitories, students returning to family homes, nursing facilities restricting access, institutional populations changing rapidly, and administrative records reflecting different dates and residence concepts.

The Post-Census Group Quarters Review identified errors and corrections. Some corrections were not used to revise apportionment or redistricting outputs. That does not establish illegality, but it does expose a governance problem: **what should happen when an official post-certification review identifies a material error in a constitutionally consequential dataset?**

A future statute should distinguish at least four correction stages:

- Before apportionment certification
- After apportionment but before redistricting release
- After redistricting has begun or maps have been enacted
- Corrections affecting statistical products but not constitutional apportionment

Each stage should have defined materiality thresholds, authority, notice requirements, and judicial review.

Part III — What the Evidence Establishes About 2020 Accuracy

7. State-Level Coverage Errors

The Census Bureau's 2020 Post-Enumeration Survey estimated statistically significant net undercounts in six states and statistically significant net overcounts in eight.[2][3]

State	Direction
Arkansas	Net undercount
Florida	Net undercount
Illinois	Net undercount
Mississippi	Net undercount
Tennessee	Net undercount

State	Direction
Texas	Net undercount
Delaware	Net overcount
Hawaii	Net overcount
Massachusetts	Net overcount
Minnesota	Net overcount
New York	Net overcount
Ohio	Net overcount
Rhode Island	Net overcount
Utah	Net overcount

GAO emphasized an important limitation: the PES estimated **no statistically significant national net coverage error**, while still identifying significant subnational and demographic errors.[3] The PES also excludes group-quarters populations and remote areas of Alaska from its national matching universe, which matters when attempting to extrapolate the PES directly into alternative apportionment calculations.[3]

The evidence therefore supports the following propositions:

- The 2020 Census was not uniformly accurate across states.
- Some demographic groups remained harder to count than others.
- The pandemic and operational disruptions created unusual risk.
- A nationally acceptable aggregate result can coexist with consequential subnational error.

The evidence does not, standing alone, establish intentional falsification, partisan manipulation, or deliberate deprivation of representation.

8. Enumeration Error Is Not Disclosure Error

The central causal question is often framed incorrectly.

The Post-Enumeration Survey evaluates whether people were correctly covered by the census—missed, duplicated, or mislocated. Differential privacy was applied later to published products through the Disclosure Avoidance System.

A useful error taxonomy is:

- **Enumeration error:** a person is missed, duplicated, or classified at the wrong residence.
- **Coverage-measurement error:** the independent survey or matching process estimates coverage incorrectly.
- **Geographic-allocation error:** a person or household is placed in the wrong geographic unit.
- **Disclosure-transformation error:** a confidential count is changed in the public release process.
- **Analytical error:** a downstream user treats a privacy-protected value as though it were an exact enumeration.

A proper investigation requires access to the confidential pre-disclosure population records, final Disclosure Avoidance System inputs, geographic crosswalks, group-quarters files, CQR and PCGQR records, PES matching files, certified apportionment calculations, and production-version histories.

Without that chain, a categorical statement that differential privacy caused the state-level undercounts is not demonstrated.

9. The Political-Pattern Argument

The CRA paper argues that the distribution of state errors produced political benefit for Democrats. [1] A partisan pattern can justify investigation. It does not, without more, establish partisan intent.

Political asymmetry can arise from neutral demographic and operational variables: housing tenure, race and ethnicity, age, urban density, student mobility, group-quarters exposure, response rates, internet and mail access, proxy-response frequency, and administrative-record quality.

The proper statistical inquiry is multivariate and causal. It should ask whether errors remain correlated with partisan variables after controlling for demographic and operational factors; whether the disclosure system itself contained a politically selective mechanism; whether comparable effects appear in simulated nonpartisan datasets; and whether any measured effect was large enough to alter legally consequential outcomes.

The correct standard is not “who benefited?” The correct standard is **what mechanism produced the effect, and was that mechanism neutral, negligent, reckless, or intentional?**

Part IV — Differential Privacy and the Disclosure Avoidance System

10. What Differential Privacy Is

Differential privacy is a formal mathematical framework designed to limit what an outside observer can infer about any one individual from a released statistical dataset.

A common formulation is:

$$|Pr[M(D) \in S]| \leq e^{\epsilon} |Pr[M(D') \in S]| + \delta$$

where D and D' are neighboring datasets that differ by one person's record, M is the release mechanism, S is a possible set of outputs, ϵ is the privacy-loss parameter, and δ permits a small probability of exceeding the principal privacy bound.

In plain language: the published output should not change enough when one person's record is added or removed to allow a user to confidently determine whether that individual was in the confidential source data.

Differential privacy is therefore a **privacy guarantee**, not a guarantee that every released local cell is exact.

11. The TopDown / DAS Pipeline

The 2020 Disclosure Avoidance System did not simply “recount” the population with an algorithm. The analytical pipeline is more accurately understood as:

Confidential tabulations → privacy-protected noisy measurements → constrained optimization / postprocessing → internally consistent released tables.

The postprocessing stage matters. Raw noise can create impossible outputs such as negative counts or inconsistent hierarchies. The Bureau therefore imposed constraints and invariants to produce usable published tables.

The consequence is that the public redistricting and demographic data are **derived statistical products**, not a literal transcription of the confidential enumeration at every block and characteristic.

12. Accuracy and Privacy Are Competing Objectives

The census operates under simultaneous objectives:

- Maximize population accuracy
- Preserve geographic utility
- Preserve demographic utility
- Protect respondent confidentiality
- Maintain public trust
- Produce data on statutory deadlines
- Provide legally usable redistricting data

These goals can conflict. Publishing exact tiny-cell demographic values can increase reconstruction risk. Injecting noise can reduce disclosure risk while degrading local accuracy.

The governance question is therefore not whether privacy should exist. Title 13 requires confidentiality. The governance question is **how privacy protection is allocated, documented, validated, and constrained when the resulting product is used to distribute political power.**

13. Redistricting and Voting-Rights Effects

Kenny and coauthors studied the Bureau's demonstration data and found that the Disclosure Avoidance System systematically undercounted population in mixed-race and mixed-partisan precincts, yielding unpredictable partisan and racial biases. They also found potentially consequential effects on redistricting and voting-rights analysis.[4]

Their work is important but must be characterized precisely. It studied demonstration data associated with the developing DAS, not every final production outcome for every jurisdiction. The research therefore establishes a serious class of risk and measurable distortion; it does not, by itself, prove that final 2020 data intentionally favored a political party.

The strongest institutional conclusion is:

A redistricting dataset should not be treated as exact at the block level when its production method intentionally introduces and redistributes statistical error. The release should disclose the error structure, the validation results, and the legal-use limitations.

14. Is the System “Open”?

The mathematical framework, significant technical documentation, demonstration data, and portions of the implementation have been publicly described. But the exact confidential input data are protected by Title 13, and public knowledge of the mechanism does not automatically make the full production system independently reproducible.

Meaning of “open”	Status
Mathematical framework publicly documented	Largely yes
Technical descriptions and demonstrations available	Yes, to a substantial degree
Confidential input records public	No
Full production pipeline independently reproducible from public data alone	Not necessarily

The correct governance demand is therefore not unrestricted release of confidential microdata. It is **secure independent reproducibility by authorized reviewers** and public disclosure of sufficient aggregate evidence to validate the system.

Part V — Data Lineage and Invariance

15. Confidential Inputs and Public Outputs

A defensible audit begins by separating the data layers:

Layer	Description	Primary integrity question
Confidential response data	Household and group-quarters information protected by Title 13	Was each person counted once and in the correct residence?
Edited / resolved census file	Deduplicated and processed enumeration records	Were conflicts and missing data resolved lawfully and accurately?
Apportionment file	State population totals used to allocate House seats	Were constitutionally controlling state totals accurate?
Redistricting file	P.L. 94-171 detailed population and housing data	Are geographic and demographic values fit for districting?
Disclosure-avoidance output	Public data transformed to reduce reconstruction risk	Is privacy effective and uncertainty adequately disclosed?
Analytical estimates	Derived measures such as CVAP	Are estimates labeled, modeled, and fit for use?

A defect in one layer does not automatically establish a defect in every other layer.

A future census should publish a formal **data-lineage statement** identifying the source of every official total, the completion date of each processing stage, the transformations applied at each geography, the invariants imposed, the difference between confidential and public aggregate values, and the uncertainty associated with released estimates.

16. Invariance Is Not Accuracy

The privacy system was designed to preserve specified higher-level totals. That distinction can be expressed simply:

$$P_s^{\{\text{published}\}} = P_s^{\{\text{confidential}\}}$$

for state s when the state total is an invariant.

This establishes preservation through the privacy transformation. It does **not** establish:

$$P_s^{\{\text{confidential}\}} = P_s^{\{\text{true}\}}$$

Preserving a number does not prove the number was correct before it was preserved.

This explains why the differential-privacy system is not the demonstrated cause of the PES state-level findings. Enumeration error can already exist upstream; an invariant can carry that error forward without creating it.

17. A Better Causal Model

A conceptual decomposition is:

$$E_{\{\text{observed}\}} = E_{\{\text{enumeration}\}} + E_{\{\text{coverage}\}} + E_{\{\text{geography}\}} + E_{\{\text{disclosure}\}} + E_{\{\text{analysis}\}}$$

The terms should not be assumed to be arithmetically independent. The equation is a governance model: each source of error must be identified and measured separately.

The record supports these conclusions:

- State-level coverage errors occurred.
- Pandemic disruption and long-standing hard-to-count conditions were consequential.
- The Disclosure Avoidance System changed detailed published geography and demographics.
- Those changes can affect redistricting analysis.
- The evidence presently available does not establish that disclosure avoidance created the underlying state-level coverage errors.

Part VI — Count Question Resolution and Post-Census Correction

18. The Function and Limits of CQR

Count Question Resolution allowed state, local, and tribal governments to challenge certain housing-unit and geographic errors after data release. It was not a universal recount and did not authorize local governments to inspect protected individual responses.

Differential privacy complicated the diagnostic environment because a public-cell discrepancy could reflect an underlying census error, geographic allocation, or the privacy transformation itself.

That does not make CQR meaningless. Local governments can still identify independently verifiable facts: a residential building existed; a facility was in another block; a structure was duplicated; an address was omitted; a facility was incorrectly classified.

The more precise conclusion is that **privacy protection weakened public diagnostic clarity without eliminating every correction pathway.**

19. A Two-Track CQR Model

A better system would separate public review from confidential verification:

Track	Evidence	Output
Public-data review	Released tables, maps, address lists, documented local discrepancies	Correction of housing, boundary, and frame errors
Secure-data review	Title 13-protected review by sworn officials or auditors	Structured determination on confidential assignment and processing issues

The secure review need not reveal individual responses. It can return findings such as confirmed omission, confirmed duplicate, confirmed geographic misplacement, confirmed group-quarters error, no error found, or unable to determine without impermissible disclosure.

20. Post-Certification Correction Authority

Congress should establish statutory materiality thresholds and procedures governing corrections discovered after certification. The framework should specify:

- What constitutes a material error
- Who can initiate review
- Which evidence is admissible
- Whether an error can change statistical products only or apportionment itself
- What notice states receive
- What judicial review applies
- What happens if redistricting has already occurred

The absence of a clear correction architecture turns finality into a matter of institutional discretion. Constitutional data should not depend on an improvised remedy system.

Part VII — Citizenship and Representation

21. Noncitizens and the Apportionment Population

The inclusion of noncitizens in the apportionment population is legally and analytically separate from differential privacy. The relevant census framework uses residence and personhood for the apportionment base, not voting eligibility.

That distinction creates several separate legal questions:

Question	Governing issue
Who is counted for apportionment?	Constitutional text, federal statute, and precedent
Should citizenship information be collected?	Congressional authority, administrative design, privacy, response effects
What population measure should states use for district equality?	Equal-protection doctrine and state law
Should citizen population replace total population for federal apportionment?	Constitutional and statutory change / judicial interpretation
How should CVAP be measured?	Surveys, administrative records, statistical modeling

Treating these as one question produces both technical and constitutional confusion.

22. *Trump v. New York*

The Supreme Court's 2020 decision did not resolve the underlying merits of whether unlawfully present noncitizens could be excluded from the apportionment base. The Court vacated and remanded because the challenge was not ripe: the implementation of the presidential memorandum and its concrete effects were uncertain.[5]

The case therefore should not be cited for the proposition that the Court definitively required inclusion or definitively authorized exclusion. It left the merits unresolved in that posture.

23. Citizenship Data Without Citizen-Only Apportionment

Citizenship information can serve legitimate governmental purposes even if total resident population remains the apportionment base. A sound statistical architecture can publish multiple clearly labeled measures:

- Total resident population
- Citizen population
- Noncitizen population
- Voting-age population
- Citizen voting-age population
- Unknown or unverified citizenship status

The last category is important. Where administrative or survey records cannot verify status, the government should report uncertainty rather than convert missing information into false precision.

Part VIII — Claim Assessment and Terminology

24. Claim Assessment Matrix

Claim	Assessment	Canonical formulation
The 2020 Census had serious errors	Supported	Official review found significant state-level undercounts and overcounts.
COVID-19 disrupted enumeration	Supported	Field, group-quarters, and nonresponse operations faced extraordinary disruption.
Differential privacy changed block-level public data	Supported	DAS transformed detailed public outputs.
Differential privacy affected redistricting analysis	Supported with scope limitations	Research found material effects in small-area and demonstration-data analysis.
Differential privacy caused the state coverage errors	Not established	Evidence does not demonstrate that public-data noise caused underlying coverage error.
Differential privacy was partisan manipulation	Not established	Requires evidence of mechanism and intent not presently shown.
Differential privacy was designed to conceal noncitizens	Unsupported as stated	The stated design purpose was confidentiality protection generally.
The 2020 Census committed fraud	Not established	Error, opacity, or disputed methodology does not by itself prove intentional fraud.
CQR became meaningless	Overstated	Diagnostic clarity was reduced; some correction pathways remained viable.
Trump v. New York required inclusion of noncitizens	Incorrect	The Court resolved ripeness / justiciability, not the merits.
Citizen-only representation is constitutionally required	Unresolved / contested	Requires direct constitutional and statutory analysis.

25. Preferred Terminology

The following terms are technically preferable in the canonical publication:

- Privacy-protected output
- Disclosure-avoidance transformation
- Statistically perturbed table
- Public tabulation derived from confidential inputs
- Privacy-induced uncertainty
- Geographic allocation distortion
- Post-enumeration coverage error
- Administrative-record imputation
- Constraint displacement

“False” should be reserved for a value demonstrated to be factually incorrect relative to the controlling source. “Fraud” should be reserved for an evidentiary showing of material misrepresentation, knowledge or reckless disregard, intent, and legally cognizable harm. A disclosed privacy transformation does not satisfy that standard by itself.

Part IX — The Five-File Verification Model

26. Core Audit Architecture

A future census integrity review should compare five principal files:

File	Function
Enumeration file	Person-level census records before disclosure transformation
Coverage file	Independent PES and matching results
Geographic file	Housing-unit and group-quarters location assignments
Apportionment file	Certified state totals and equal-proportions calculations
Public release file	Privacy-protected redistricting and demographic outputs

For geography g , the disclosure transformation can be measured as:

$$|\Delta_g = P_g^{\{\text{public}\}} - P_g^{\{\text{confidential}\}}|$$

Coverage error is conceptually distinct:

$$C_g = P_g^{\{\text{enumerated}\}} - P_g^{\{\text{estimated true}\}}$$

The first measures the public-release transformation. The second measures estimated census coverage. They must never be collapsed into one statistic.

27. Required Audit Questions

A complete audit should answer:

- What was the confidential count at each relevant geography?
- What was the public count after disclosure avoidance?
- Which invariants constrained the transformation?
- How much population was displaced between blocks or characteristics?
- How frequently did demographic distributions change while total population remained constant?
- Were distortions concentrated in particular racial, socioeconomic, geographic, or partisan environments?
- Did the release affect district population equality?
- Did it materially affect voting-rights analyses?
- How did group-quarters corrections compare with the original enumeration?
- Which CQR and PCGQR challenges were accepted, rejected, or unresolved?
- What legal standard governed the decision not to revise certified products?
- Can an independent authorized auditor reproduce the production output from the confidential source and documented pipeline?

Part X — Independent Review and Governance Reform

28. Census Data Integrity Review Board

Congress should establish an independent **Census Data Integrity Review Board** with:

- Statutory independence from the Bureau's operating chain

- Title 13 sworn personnel
- Statistical, demographic, geographic, legal, privacy, election-administration, and cybersecurity expertise
- Authority to inspect confidential processing records
- Authority to commission replication studies
- Public reporting authority subject to confidentiality law
- A duty to distinguish enumeration error from privacy transformation
- Formal response requirements for major external audits and state challenges

The Board should publish aggregate error findings, geographic patterns, confidence intervals, methodological explanations, public code where lawful, unresolved material questions, and an assessment of data fitness for each major legal use.

29. Redistricting Validation Service

Before P.L. 94-171 data are released, Census should run a formal validation program measuring:

- Mean absolute error (MAE)
- Root mean squared error (RMSE)
- Signed bias
- Relative error
- Quantile error
- Constraint displacement
- Off-spine / VTD error
- Population-parity error
- Minority-opportunity-district sensitivity
- Reconstruction-risk performance

No privacy mechanism should be treated as successful merely because it satisfies a privacy parameter. It must also satisfy use-specific accuracy thresholds.

30. Group-Quarters Controls

The 2030 system should include audited group-quarters controls with pre-Census-Day facility validation, redundant administrative sources, date-specific occupancy reconciliation, geographic verification, anomaly detection, and a rapid post-enumeration challenge process.

Group quarters are too politically and statistically consequential to remain a secondary quality-control issue.

31. Reproducibility and Data Lineage

The Bureau should publish machine-readable documentation describing:

- Every major processing stage
- Software and algorithm versions
- Privacy parameters and invariants
- Validation results
- Known limitations
- Correction history
- Data dependencies
- Release dates
- Changes between demonstration and production systems

Authorized auditors should be able to reproduce the public outputs from the protected source environment. Reproducibility is the bridge between confidentiality and institutional accountability.

Part XI — Reform Options for 2030

32. Option One: Exact Constitutional Totals, Privacy-Protected Detail

This model preserves exact or securely verified legally controlling totals while permitting privacy protection for detailed demographic products.

It would include exact state totals for apportionment, validated housing and geographic totals, privacy-protected demographic tables, enhanced secure products for authorized redistricting users, published uncertainty information, and independent validation before release.

33. Option Two: Selective Suppression and Aggregation

Instead of perturbing broad classes of cells, Census could publish exact total population where legally necessary while suppressing or aggregating high-risk characteristic cells.

This model is easier for users to interpret but can reduce demographic detail and must be designed to prevent differencing attacks across overlapping tables.

34. Option Three: Tiered Access

User	Access
General public	Privacy-protected tables, uncertainty, and methodology
State redistricting authorities	Enhanced protected products and validation tools
Sworn auditors	Secure access to confidential processing and exact aggregates
Courts	Sealed access subject to protective orders
Approved researchers	Controlled-access data under research protocols

Tiered access recognizes that different legal users require different levels of precision while Title 13 continues to protect individual responses.

35. Option Four: Separate Citizenship Measurement Architecture

Citizenship analysis should be built as an explicit statistical and administrative system rather than treated as an incidental byproduct of a privacy-protected population table.

The system should define citizenship categories, identify source and date, distinguish verified from inferred status, report unknown status, publish geographic uncertainty, prohibit silent conversion of estimates into official census counts, and protect information according to applicable law.

36. Option Five: Constitutional and Statutory Clarification

Congress should clarify:

- The legally controlling apportionment population
- Whether and how citizenship data must be collected
- Whether CVAP should be an official redistricting product
- Which totals must remain exact
- Materiality standards for post-certification errors
- Correction authority and judicial review
- State and locality review rights
- Disclosure requirements for privacy transformations

Administrative methodology should not carry constitutional questions that Congress has left unresolved.

Part XII — Constitutional and Legal Analysis

37. Enumeration Accuracy

The Enumeration Clause requires an actual enumeration in the manner Congress directs. It does not prescribe a modern statistical architecture. The constitutional purpose, however, is clear: generate the population numbers used to distribute representation.

A privacy mechanism is not automatically inconsistent with an actual enumeration. The stronger constitutional questions are whether the legally controlling apportionment totals were based on the actual enumeration; whether unauthorized estimation or adjustment was used; whether material known errors were ignored arbitrarily; whether redistricting data became unfit for their legal purpose; and whether states retained a meaningful correction pathway.

38. Equal Representation

Congressional districts are subject to a demanding population-equality standard. State legislative and local districts operate under related but different doctrines.

Privacy-protected data create litigation risk where map drawers cannot tell whether deviations reflect population or transformation. That risk does not automatically establish a constitutional violation. It does establish the need for a secure validation channel and use-specific accuracy evidence before states are forced to rely on the data.

39. Voting Rights Act Analysis

Voting Rights Act litigation can depend on racial composition, citizen voting-age population, geographic concentration, and ecological inference. Small-area perturbations can therefore matter even if statewide totals are preserved.[4]

A stronger architecture should provide secure access to exact aggregate data for authorized legal purposes, transparent uncertainty intervals, expert certification of fitness for use, and court-supervised access where necessary.

Part XIII — Evidence Classification

40. Established Facts

Subject to the cited sources, the following may be treated as established:

- The Census Bureau used a differential-privacy-based Disclosure Avoidance System for 2020 detailed data products.[6]
- The system was designed to protect confidentiality.
- The 2020 PES estimated statistically significant state-level undercounts and overcounts in fourteen states.[2][3]
- The national PES estimate did not show a statistically significant net national coverage error.[3]
- COVID-19 and long-standing operational challenges affected census operations.[3]
- Public detailed data differed from confidential source values because of disclosure avoidance.
- Peer-reviewed research found material redistricting and voting-rights risks in DAS demonstration data.[4]
- The Census Bureau is now evaluating alternatives after a Commerce administrative order prohibiting noise infusion.[6]

41. Allegations Requiring Further Proof

The following remain allegations unless additional evidence establishes them:

- The Bureau intentionally distorted data to benefit Democrats.
- Differential privacy was designed to hide noncitizen populations.
- Officials knowingly certified false apportionment totals.
- Empty group quarters were knowingly populated through unlawful adjustment.
- Correction mechanisms were intentionally weakened to prevent discovery of politically damaging errors.
- Congressional representation was deliberately manipulated.

42. Supported Inferences and Open Questions

Reasonable but qualified inferences include:

- Differential privacy reduced the transparency of some small-area population data.
- The system complicated local diagnosis of discrepancies.
- Pandemic group-quarters disruption likely contributed to some accuracy problems.
- Geographic perturbation can affect redistricting and voting-rights analyses.
- Maximum public detail and maximum confidentiality are structurally competing goals.

These should be tested through secure access to the full data lineage rather than converted into categorical conclusions.

Part XIV — Legislative Architecture

43. Proposed Census Algorithm Accountability Act

Congress should consider legislation containing the following elements.

Findings

1. Decennial census data allocate constitutional political power and therefore require exceptional standards of integrity, auditability, and public accountability.
2. Disclosure avoidance is necessary to satisfy confidentiality obligations, but privacy controls used on legally consequential data must be validated for their intended governmental uses.
3. Enumeration error, geographic error, disclosure transformation, and analytical estimation are separate sources of uncertainty and must be measured separately.
4. States and localities require a meaningful mechanism to challenge material errors without receiving protected individual responses.
5. Independent secure technical review can reconcile Title 13 confidentiality with institutional accountability.

Core Requirements

- Data-lineage publication for all constitutionally and legally consequential products
- Independent Title 13-protected technical review
- Redistricting-specific validation before release
- Public accuracy and uncertainty metrics
- Published algorithm and software versioning
- Materiality thresholds for post-certification corrections
- Audited group-quarters controls
- Secure state and court validation channels

- Separate publication of total population, citizen population, CVAP, and uncertainty measures
- Congressional notification before material changes to disclosure-avoidance architecture

Part XV — Current 2026 Context

The Census Bureau now states on its differential-privacy pages that the information is no longer current and that it is evaluating alternative disclosure approaches to comply with a Department of Commerce administrative order prohibiting noise infusion.[6]

This development changes the 2030 policy environment in two ways.

First, it makes the question of whether the 2020 system should be repeated largely obsolete. The more important question is what replaces it and whether the replacement is subject to stronger validation, transparency, and independent review.

Second, it reinforces the need for durable statutory architecture. A disclosure methodology used on constitutionally consequential data should not swing dramatically between administrations without public technical standards, congressional notice, and use-specific validation.

The 2020 experience should therefore be treated as a governance case study rather than a closed historical dispute.

Conclusion

The 2020 Census presents a serious institutional accountability problem, but the problem is more precise than the categorical claim of fraud.

The evidence establishes significant state-level coverage disparities, pandemic-related operational and group-quarters complications, and privacy-induced changes to detailed public data. Peer-reviewed research demonstrates that disclosure-avoidance transformations can affect small-area redistricting and voting-rights analysis. Those facts are sufficient to justify significant reform.[2][3][4]

The evidence presently available does not establish that differential privacy caused the state-level undercounts and overcounts. The Post-Enumeration Survey addresses coverage of the enumerated population; the Disclosure Avoidance System transforms released tables. Connecting those systems causally requires confidential-data analysis that has not been publicly demonstrated.

The evidence also does not establish that differential privacy was created to conceal noncitizens or to favor a political party. The system's stated purpose was respondent confidentiality. Political consequences can and should be measured, but consequence is not identical to intent.

The strongest governing principle is therefore:

Exactness should be preserved and independently verified wherever the law requires a controlling population value. Privacy protection should be transparent, auditable, use-specific, and limited so that it does not prevent meaningful constitutional, redistricting, judicial, or state-level accountability.

A durable census architecture should produce separate products for separate legal purposes:

- A constitutionally controlling apportionment population
- A geographically validated redistricting population file
- Privacy-protected demographic products
- Citizen and citizen-voting-age measures
- Secure exact-count access for authorized auditors and courts
- Public uncertainty, version, and data-lineage documentation

The 2020 Census should be treated as a case study in **governance failure risk at the intersection of enumeration, privacy, geography, and representation**. The answer is not to eliminate statistical uncertainty. It is to make uncertainty visible; preserve exact legal obligations; create credible correction pathways; and ensure that no privacy mechanism, operational failure, or institutional decision prevents the American people and their governments from auditing the data through which political representation is allocated.

Appendix A — Canonical Audit Metrics

The following measures should form the minimum technical dashboard for future census data governance:

Metric	Purpose
Mean Absolute Error (MAE)	Typical absolute difference between protected and reference values
Root Mean Squared Error (RMSE)	Emphasizes larger deviations
Signed Bias	Detects systematic positive or negative displacement
Relative Error	Scales error to population size
Quantile Error	Measures distributional performance across small and large geographies
Constraint Displacement	Quantifies changes forced by postprocessing constraints
Off-Spine / VTD Error	Measures distortion where political geography does not align cleanly with the geographic spine
Population-Parity Error	Tests district-equality implications
Minority-Opportunity-District Sensitivity	Measures whether disclosure treatment changes legally relevant district classifications
Reconstruction Risk	Tests whether published products still permit inference of protected characteristics
CQR Resolution Rate	Measures effectiveness of local correction pathways
Group-Quarters Reconciliation Error	Measures institutional-residence accuracy

Appendix B — Knowledge Graph

Core entities

U.S. Constitution → Enumeration Clause → Fourteenth Amendment → Congress → Department of Commerce → U.S. Census Bureau → 2020 Census → Post-Enumeration Survey → Disclosure Avoidance System → Differential Privacy → TopDown Algorithm → P.L. 94-171 Redistricting Data → State Redistricting Authorities → Voting Rights Act Analysis → House Apportionment → Electoral College → Count Question Resolution → Post-Census Group Quarters Review → GAO → Independent Review → 2030 Census.

Contested-claim branch

Center for Renewing America → Wade Miller → Andrew White → *The 2020 Census Fraud* → alleged partisan skew → alleged citizenship concealment → demand for reform.

Evidence branch

Official coverage findings → pandemic operational disruption → small-area disclosure error → peer-reviewed redistricting research → legal distinction between persons and citizens → unresolved causal / intent claims.

Appendix C — Canonical SEO and Semantic Metadata

Canonical title: The 2020 Census, Differential Privacy, and the Integrity of Representation **Short title:** 2020 Census and Differential Privacy **Primary entity:** U.S. Census Bureau **Primary topics:** 2020 Census; differential privacy; Disclosure Avoidance System; TopDown; census accuracy; apportionment; redistricting; citizenship; Post-Enumeration Survey; election integrity; constitutional governance **Secondary topics:** group quarters; Count Question Resolution; Voting Rights Act; citizen voting-age population; data governance; algorithmic accountability; public-sector AI and algorithms; privacy; administrative law

Suggested description: A source-anchored examination of the 2020 Census that separates documented enumeration errors from differential-privacy effects, evaluates redistricting and citizenship claims, and proposes an auditable governance architecture for the 2030 Census.

Suggested keywords: 2020 Census, Census Bureau, differential privacy, TopDown Algorithm, Disclosure Avoidance System, census undercount, census overcount, apportionment, redistricting, Electoral College, citizenship data, citizen voting age population, Census fraud claims, census accuracy, data governance, election integrity, constitutional governance, Wade Miller, Center for Renewing America, 2030 Census reform

Appendix D — Source Register

[1] Advocacy / policy analysis Wade Miller and Andrew White, *The 2020 Census Fraud: How Differential Privacy Skewed Representation*, Center for Renewing America, August 7, 2025. This source is used for the CRA thesis and its allegations; it is not treated as independent proof of those allegations. <https://americarenewing.com/>

[2] Primary federal source U.S. Census Bureau, *2020 Census Undercounts in Six States, Overcounts in Eight*, May 19, 2022. <https://www.census.gov/library/stories/2022/05/2020-census-undercount-overcount-rates-by-state.html>

[3] Independent federal oversight U.S. Government Accountability Office, *2020 Census: Coverage Errors and Challenges Inform 2030 Plans*, GAO-25-107160, November 21, 2024. <https://www.gao.gov/products/gao-25-107160>

[4] Peer-reviewed research Christopher T. Kenny, Shiro Kuriwaki, Cory McCartan, Evan T. R. Rosenman, Tyler Simko, and Kosuke Imai, "The use of differential privacy for census data and its impact on redistricting: The case of the 2020 U.S. Census," *Science Advances* 7, no. 41 (2021), eabk3283. <https://doi.org/10.1126/sciadv.abk3283>

[5] Primary legal authority *Trump v. New York*, 592 U.S. 125 (2020). https://www.supremecourt.gov/opinions/20pdf/20-366_7647.pdf

[6] Primary federal methodology / current policy notice U.S. Census Bureau, *Understanding Differential Privacy* and related Disclosure Avoidance materials. The Bureau currently states that this guidance is no longer current and that it is evaluating alternatives under a Department of Commerce administrative order prohibiting noise infusion. <https://www.census.gov/programs-surveys/decennial-census/decade/2020/planning-management/process/disclosure-avoidance/differential-privacy.html>

[7] Advocacy explainer Center for Renewing America, *Differential Privacy in the 2020 Census Explained*, August 21, 2025. Used to document the organization's own structural-inaccuracy, opacity, and citizenship-nullification arguments. <https://americarenewing.com/>

[8] Congressional record U.S. House Committee on the Judiciary, Subcommittee on the Constitution and Limited Government, hearing: *Enumeration or Estimation: Why Inaccurate Census Results Hurt American Citizens*, November 19, 2025. Joseph Wade Miller appeared as a witness. <https://docs.house.gov/>

[9] Statutory authority 13 U.S.C. §§ 9, 141; 2 U.S.C. § 2a. Confidentiality, decennial census, apportionment reporting, and allocation framework.

Appendix E — Publication Control

Document class: Canonical Institutional Research Monograph **Status:** ADOPTED — CANONICAL
Version: 1.0 **Date:** 20 August 2026 **Author:** Anthony Lewis Jerdine **Institutional surface:** Miami
 Wealth Capital **Editorial standard:** Evidence-led; primary-source preference; allegation/fact
 separation; reproducible claims; legal and technical distinctions preserved. **Supersession rule:**
 Any future revision must identify changed findings, changed source status, methodology updates,
 and superseded language. Silent alteration of evidentiary classification is prohibited.