

U.S. Department of the Treasury



Gold Standard Science

Initial Implementation Report

Introduction

The mission of the U.S. Department of the Treasury (Treasury) is to maintain a strong economy, to create economic and job opportunities through economic growth and stability at home and abroad, to strengthen national security, and to manage the U.S. Government's finances and resources effectively. In short, Treasury's role is to be a good steward of the U.S. economy and financial system and to promote the United States's influence in the world economy.

On March 23, 2025, President Trump signed Executive Order (EO) 14303, "Restoring Gold Standard Science."¹ The order seeks to "restore the American people's faith in the scientific enterprise and institutions that create and apply scientific knowledge in service of the public good." The order tasked federal agencies to report to the White House's Office of Science and Technology Policy (OSTP) on how they would address each tenet of Gold Standard Science (GSS) as laid out in EO 14303 for the conduct and management of their respective scientific activities. It is imperative that agencies produce accurate and reliable data and that they exercise rigor, transparency, and impartiality in consequent analyses. This will build public confidence in policy decisions that have real impacts on the lives of Americans.

Although Treasury's mission does not directly relate to applied research or experimental development (as defined in OMB Circular No. A-11, it uses basic research to fulfill its role as steward of the U.S. economy and financial systems. Treasury has been committed to, and reaffirms its commitment to, applying gold standard principles to its functions. The purpose of this report is five parts:

- (i) describe how Treasury is addressing each of the nine tenets of GSS as laid out in EO 14303, including how the tenets are reflected in Treasury's culture, funding opportunities, and award selection,
- (ii) detail how Treasury could standardize metrics and evaluation mechanisms to address adherence to GSS tenets,
- (iii) relate plans to provide training and resources to ensure Treasury's personnel understand and adhere to the tenets of GSS
- (iv) discuss how technology will be leveraged for implementing GSS, and
- (v) describe potential challenges when implementing GSS.

Tenets of Gold Standard Science

Executive Order 14303 outlines nine tenets to ensure federal data, research, and analyses are compatible with Gold Standard Science. Together these nine tenets promote the generation of knowledge and will restore the public's trust in science and its subsequent application to public policy:

- (i) replicable

¹ Federal Register, Vol. 90, No. 102. Executive Order 14303 of May 23, 2025. "Restoring Gold Standard Science." Available: <https://www.govinfo.gov/content/pkg/FR-2025-05-29/pdf/2025-09802.pdf>



- (ii) transparent
- (iii) communicative of error and uncertainty
- (iv) collaborative and interdisciplinary
- (v) skeptical of its findings and assumptions
- (vi) structured for falsifiability of hypotheses
- (vii) subject to unbiased peer review
- (viii) accepting of negative results as positive outcomes; and
- (ix) without conflicts of interest.

1) Research must be reproducible. The GSS tenet of reproducible science reflects two key concepts: replicability and reproducibility. Replicability is the ability of the same researchers and external reviewers to perform the same analysis and achieve the same results from the same methods, while reproducibility is the ability to use multiple methods to consistently achieve results that confirm or refute prior results. Together, these concepts increase confidence in the integrity and precision of the results and can validate broader scientific claims. To promote this tenet, OSTP has directed that agencies should require of their research clear, standardized, and justifiable protocols; comprehensive documentation; robust statistical methods; adequate sample sizes; validated methodologies; and appropriate controls.

2) Research and results must be transparent. This tenet entails the open, accessible, and comprehensive sharing of all components of the research process—methodologies, data, analytical tools, and findings—to enable stringent scrutiny, validation, and reuse by research peers and the public. OSTP has charged agencies to prioritize clear, detailed reporting of methodologies, make raw data and analytical tools publicly available when feasible and lawful, and disclose funding sources or conflicts of interest.

3) Results must acknowledge and communicate the limitations of assumptions and methodologies within research. That is, research should not claim too broad conclusions—though subsequent extensions of previous research may build on and expand the applicability of a study’s results. Communicating error and uncertainty in science requires the clear, precise, and accurate disclosure of limitations, variability, and potential sources of error, as well as acknowledging any limitations in data, measurements, or research findings. OSTP has instructed agencies to ensure research reporting includes quantitative measures of uncertainties—such as confidence intervals, error margins, or sensitivity analyses—alongside clear explanations of methodological constraints and assumptions and the intended scope of the research, including what the scientific findings do and do not establish. Moreover, agencies should promote cautious, evidence-based language in reports, publications, and public communications and discourage speculative claims or extrapolations that extend well beyond the data’s scope, especially when science is used in an operational or regulatory context.

4) GSS also encourages collaborative and interdisciplinary science. Including multiple disciplines in the creation and review of research can increase insight into topics of study, and the integration of a wide range of expertise, methodologies, and perspectives can catalyze transformative discoveries. OSTP says that agencies shall prioritize collaborative and interdisciplinary approaches in scientific research to accelerate discovery and innovation. These approaches include



recognizing limitations in an individual's or an agency's expertise and engaging other divisions within an agency, or other agencies, for complementary expert support when appropriate to address cross-disciplinary problems.

5) In addition, researchers should have a healthy skepticism of assumptions and consequent findings. Critical evaluation of underlying assumptions, methodologies, and findings—as well as openness to alternative assumptions, model specifications, interpretations—promotes the validity, robustness, and reliability of research. As a result, the scientific process is strengthened. One key aspect of this tenet is to be aware of potential confirmation bias and avoid it. OSTP has mandated that agencies foster this GSS tenet through policies and programs that emphasize critical evaluation, transparency, and objectivity. One such way is through encouraging constructive collaborations among researchers with differing viewpoints to critically assess the reliability of research results.

6) Gold Standard Science should also be structured that hypotheses can be disproven. The GSS tenet of falsifiability requires designing studies to so that hypotheses can be carefully tested and potentially disproven empirically. This approach promotes rigor and the generation of new knowledge by preventing the perpetuation of unproven assumptions. OSTP has directed that federal agencies prioritize research that is structured for falsifiability of hypotheses and advances knowledge through careful testing and is transparent about null or negative results in publications.

7) Another tenet of GSS is to ensure that research is subject to unbiased peer review at various stages, including but not limited to the proposal, draft, and publication submission stages. Ensuring impartial and independent evaluation by qualified experts increases the credibility of research, minimizes bias, ensures rigor, and promotes objective scrutiny. OSTP has charged agencies to prioritize unbiased peer review to advance sound science in the review, selection, and awarding of Federal grants and contracts, including competitive and discretionary awards.

8) It is also imperative that research accept negative results as beneficial to the goal of increasing knowledge. A study's failure to confirm a hypothesis or its production of counterintuitive results still adds to the understanding of the world. Moreover, it leads to further inquiry to confirm the results or explore why the hypothesis was rejected. This tenet counters publication bias, encourages comprehensive reporting, and provides valuable insights into ineffective approaches, thereby guiding future research directions and avoiding redundant efforts. OSTP has instructed agencies to foster integrity and research innovation by recognizing that negative or null results are valuable contributions to scientific knowledge, such that research projects transparently report all outcomes, including null or negative results, in publications and publicly accessible data repositories, accompanied by clear, detailed documentation of methods, analyses, and limitations.

9) The final tenet of GSS requires that those performing research do so without conflicts of interest. Research should be designed, executed, reviewed, and reported free from financial, personal, or institutional influences that could bias outcomes or undermine objectivity. This upholds scientific integrity, fosters public confidence, and ensures that results reflect evidence rather than external agendas. OSTP has says that agencies must require researchers, reviewers, and managers to disclose all relevant affiliations, funding sources, and relationships and adhere to stringent ethical



standards that are supported by strong institutional oversight, transparent reporting systems, and independent expert review mechanisms.

Research and Analysis at Treasury

Treasury's mission is as a steward of the U.S. economy and financial systems. It is not a scientific research agency, either directly or indirectly. However, GSS principals can be applied to several of the support activities that enable Treasury to maintain a strong economy and create economic and job opportunities. By applying GSS tenets, Treasury may be better able to promote the conditions that enable economic growth and stability at home and abroad, strengthen national security, protecting the integrity of the financial system, and manage the U.S. Government's finances and resources effectively.

Policy Analysis

In furthering its mission of ensuring economic and financial stability, Treasury routinely uses economic, tax, and financial analyses and research to inform and guide principals as to the likely impacts, tradeoffs, and effectiveness of potential policy proposals and developments in the economy and financial sector. Staff in various offices—including, but not limited to, the Office of Economic Policy, Office of Tax Policy, and Office of Financial Research—pursue different lines of inquiry in support of principals.

Treasury is a department that works in a quickly changing economic and financial environment, and the feasibility of applying GSS tenets in analysis depends on the timeline. When prompt action is needed, the application of some GSS tenets—such as interdisciplinary collaboration—may be outweighed by the necessity of timeliness. At the same time, accuracy and validity of the models used in analysis must not be sacrificed for timeliness—after all, Treasury's principals need quality analysis and an accurate assessment of the economic and financial environment to make decisions that best serve the American public. For this type of quick-turn analysis, Treasury offices rely on proven models, validated methodologies, and intraoffice review to provide quality analysis that will best support principals' needs.

When timeliness is less paramount, longer form policy research and post-hoc analyses can better implement GSS tenets in support of Treasury's mission. Treasury's research and research-adjacent offices will seek to implement standards that comply with OSTP's guidance on EO 14303.

Grant Evaluation

The U.S. Treasury administers many grants and financial assistance programs through formula grants, block grants, loans, direct payments for specified uses, and cooperative agreements. While much of this financial assistance is disbursed by formula, legislation, and regulation, one of Treasury's grant programs signals how GSS tenets may help guide public policy.

In 2018, Congress passed the Social Impact Partnerships to Pay for Results Act (SIPRA)—a landmark law that focuses federal spending on policies that can be proven to work—and assigned Treasury to run the grant program through 2033. Unlike many federal programs that pay for specific benefits or services, this \$100 million program is connected to realized outcomes. States and local



governments identify social problems they would like to address; an independent evaluator measures whether the goals were achieved; and Treasury only pays if the program delivers results.²

The SIPPPRA statute requires projects to demonstrate that outcomes “have been achieved as a result of the intervention.” The independent evaluator must use rigorous methods that can reliably show a direct link.³ The main goal of the project's independent evaluation is to determine how strongly the project can say that it caused the observed outcomes, and not some other unaccounted for variable.

Some projects use randomized controlled trials (RCTs), which are generally considered to be the most rigorous type of experimental design. In RCTs, a sample is randomly split into two groups – treatment and control. One group will receive the intervention and the other will continue as normal. Because people are randomly assigned to these groups, RCTs help minimize the chance that any differences in outcomes we see are due to other factors, rather than the project itself.

If randomization is not feasible, Treasury has accepted other reliable, evidence-based quasi-experimental designs. These methods compare the outcomes of the group receiving the project to a similar group that did not. While they do not use random assignment, these designs use careful planning and analysis to create a comparison group that is as similar as possible to the project group. When using quasi-experimental designs, grant applicants must clearly detail why randomization is not feasible and show how controls will provide robust evidence that alternative variables—such as who was selected for the project, other policies that were in place, changes in the economy, or other factors—could not explain sufficiently the results of the program. Both RCT and quasi-experimental designs must use statistical significance tests to show that the results are likely due to a causal effect and not chance.

In short, the SIPPPRA program seeks to use many of the GSS principles, thereby ensuring that federal grants will achieve beneficial results and yield savings for the federal government. Treasury’s experience with implementing and managing SIPPPRA could serve as a guide for future grant programs.

Regulatory Analysis

Another primary stream through which Treasury can implement the Administration’s GSS agenda is through its regulatory and deregulatory efforts to ensure changes in regulations are tailored to minimize unnecessary compliance costs on companies while also ensuring the stability of the financial system.

In July 2025, Congress passed the Genius Act in July 2025, which intends to create a comprehensive regulatory framework for stablecoins. Treasury is responsible for issuing regulations to implement the Genius Act that encourage financial innovation while also checking illicit finance, protecting consumers, and addressing financial stability risks. Various Treasury

² The SIPPPRA statute defines the term “State” to mean each State of the United States, the District of Columbia, each commonwealth, territory, or possession of the United States, and each federally recognized Indian tribe. See 2 CFR Part 200 for definitions of State, local government, or federally recognized Indian tribe.

³ [Section 2056\(a\) of SIPPPRA Legislation.](#)



offices will support the creation of this regulatory regime through regulatory impact analyses (RIAs). These analyses will use clear assumptions and will detail methodologies so that the regulations can withstand scrutiny. Treasury’s regulatory efforts with respect to the Genius Act—as well as other legislation that will require financial regulation—should result in the correct degree of regulation such that burden of compliance is not onerous and is able to achieve the multiple goals of the specific legislation. Complementary to the regulatory efforts is using analysis to ensure deregulatory efforts are properly targeted to mitigate compliance costs while protecting consumers, the economy, and the financial system.

Metrics and Mechanisms of GSS Implementation

Treasury is still exploring the methods by which it will evaluate GSS implementation. Research and research-adjacent offices will collaborate the support and evaluation offices, like the Office of Strategic Planning and Performance Improvement, to ensure implementation of GSS principals and report to the OSTP in the annual reports due on September 1 each year starting in 2026.

Emerging Technology Application

Artificial Intelligence (AI) has the potential to transform society, and Treasury is in the process of embracing the potential gains from AI to better serve the American people. Treasury’s AI Strategy reflects both the department’s mission and the moment by applying AI to modernize Treasury operations, strengthen American economic security, and reinforce the Department’s global leadership in financial innovation and governance.

Treasury’s AI Strategy is anchored in three core objectives:

- Apply AI to improve Treasury operations, deploying AI to enhance analytic capabilities, to improve service delivery, and to increase the efficiency of Treasury programs — while safeguarding classified and sensitive information.
- Advance responsible AI adoption in the financial sector by working with financial institutions, technology partners, and regulators to support innovation while safeguarding market integrity and operational resilience.
- Promote U.S. leadership through global and federal AI alignment. Treasury will engage with interagency and international partners to shape the standards and frameworks that govern AI adoption across financial systems around the world.

Treasury’s AI Governance Board is working with career and appointees to explore how AI implementation across the department can improve consistency, coordination, and accountability of Treasury operations and analyses used for policymaking.



Challenges

Economic, tax, and financial analyses and research are often not experimental. Rather they routinely use existing data and apply accepted statistical techniques to ascertain correlation and causality, magnitude, uncertainty, etcetera. This can lead to several challenges in applying GSS tenets.

- The biggest challenge for applying GSS at Treasury is that the department works in a quickly changing economic and financial environment that may require prompt action. The usefulness of analysis to inform public policy is constrained by the necessity of timeliness. At the same time, accuracy and validity of the models used in analysis must not be sacrificed for timeliness, but the application of some GSS tenets is outweighed by the needs of the moment.
- Another challenge for applying GSS at Treasury lies in the very nature of data used for policy analysis. For most economic and financial research, Treasury does not collect its own data, but rather relies on data from various sources, such as surveys conducted by the Census Bureau, Bureau of Labor Statistics, Federal Reserve Board and regional banks, etcetera as well as private sector data. The results of analysis can be limited by unknown bias within the data, which may not be apparent a priori and for which analysts may not sufficiently control. As a result, conclusions can change based on (sometimes large) data revisions.
- Data availability is another key limitation to the pursuit of longer form social policy research at Treasury. While aggregate survey data are publicly available from federal statistical agencies, data are not usually disaggregated finely enough to provide confidence in results—thus limiting the implementation of some GSS tenets. Researchers can apply to use confidential microdata from Federal Statistical System Research Data Centers (RDC), the cost is prohibitively expensive—especially in a budget constrained environment. Increased access by qualified Treasury staff to RDCs would provide for better analysis of discrete policy questions for principals, and it would provide Treasury increased ability to provide interdepartmental review of similar research performed by other agencies.
- Data that are more disaggregated are often released on a lagged timeline. For example, the latest Statistics of Income (SOI) data were released in 2025 and pertained to tax returns in the 2022 tax year. This adds uncertainty as to how applicable analysis using the 2022 SOI data are to current policy. Replicability, however, is not affected if analysis is repeated annually. Repeated analysis with new data can confirm or cast doubt on previous estimates and related error bands, updating the knowledge research provides for principals in pursuing public policies.

