

MERCATUS SPECIAL STUDY



**BEYOND DYNAMIC SCORING**  
ENHANCING CBO'S EVALUATION OF  
PROGRAM INTEGRITY SPENDING

Keith Hall

MERCATUS.ORG



**MERCATUS CENTER**  
George Mason University

*Keith Hall. "Beyond Dynamic Scoring: Enhancing CBO's Evaluation of Program Integrity Spending." Mercatus Research, Mercatus Center at George Mason University, Arlington, VA, June 2024.*

## Abstract

Scorekeeping is the process of developing and recording measures of the budgetary effects of legislation under consideration in Congress. The Congressional Budget Office (CBO) spends time and resources trying to accurately estimate how individuals, firms, and government entities would react to a policy and how this reaction would affect the federal budget. These behavioral responses to legislation are an important part of budget scores, and CBO can use its expert-based judgment on the methodology it employs for these estimates as long as it follows official scorekeeping guidelines and additional direction from the budget committees. Unfortunately, restrictions in the guidelines can interfere with the accurate scoring of spending on program integrity activities (i.e., activities designed to reduce waste, fraud, or abuse). The guidelines often direct CBO to score the change in the cost of program integrity but not the resulting change in the levels of waste, fraud, and abuse, thereby creating a scoring bias in favor of higher levels. A change is needed to correct this bias and free CBO to fully score program integrity.

*JEL* codes: H11, H50, H68

Keywords: national budget, policy making, program integrity

© 2024 by Keith Hall and the Mercatus Center at George Mason University

The views expressed in Mercatus Policy Research are the author's and do not represent official positions of the Mercatus Center or George Mason University.

Congressional scorekeepers rely on budgetary estimates provided by the Congressional Budget Office (CBO).<sup>1</sup> To make accurate and unbiased projections, CBO spends considerable time and resources trying to understand and accurately estimate how individuals, firms, and government entities would react to a policy and how this reaction would affect the federal budget. These so-called behavioral responses to legislation are an important part of budget scores. CBO staff members are free to use their best, expert-based judgment on the methodology they develop and employ for these estimates, as long as they follow the official scorekeeping guidelines and additional direction from the two budget committees. There are, however, two notable ways in which their accuracy is affected by these requirements. One is the well-known (and controversial) issue of dynamic scoring. For much of the time since 1997, congressional rules have asked CBO for dynamic scoring under certain circumstances but for informational purposes only.<sup>2</sup> The other way (and the topic of this paper) is that the scorekeeping guidelines can interfere with the scoring of spending on program integrity activities. Program integrity activities are designed to reduce waste, fraud, or abuse in federal programs. They are vital for a well-functioning government to operate efficiently, especially now that these problems have surged along with the quite unusual level of federal spending brought on by the global pandemic. The Government Accountability Office (GAO) estimates that waste, fraud, and abuse have cost the federal government and taxpayers a trillion dollars since 2020. The challenge for scorekeeping accuracy is that the guidelines often dictate that a change in *spending* on program integrity be estimated and scored, but its resulting effect on waste, fraud, and abuse—the reason for the spending—cannot be. This creates a bias in scoring that

---

1. CBO is directed to use estimates of revenue legislation made by the Joint Committee on Taxation.

2. See Megan S. Lynch and Jane G. Gravelle, “Dynamic Scoring in the Congressional Budget Process” (CRS report R46233, Congressional Research Service, Washington, DC, March 6, 2023), for a discussion of the issue and the current congressional requirements for its use.

works against program integrity and in favor of higher levels of waste, fraud, and abuse. This bias works in both directions; that is, increases in program integrity are not viewed as lowering government spending through a reduction in waste, fraud, and abuse, and decreases in program integrity spending lower overall spending, but there is no accounting of the resulting increase in waste, fraud, and abuse. A change in the guidelines is needed to correct this bias and free CBO to fully score program integrity.<sup>3</sup>

There are 16 scorekeeping guidelines that CBO must follow, but only guidelines 3 and 14 significantly interfere with the scoring estimation work of CBO. Either one can prevent an accurate, unbiased score of program integrity. Guideline 3 is nicknamed the “fingerprint rule.” It applies when changes in mandatory spending are included in an appropriations act. If the mandatory program changes are substantive—as with a new mandatory program or a significant change in an existing program—then all of the spending effects are treated as discretionary changes for purposes of budget enforcement. These provisions are often referred to as CHIMPs (changes in mandatory programs), and they are commonly used as an offset for discretionary spending.<sup>4</sup> The problem for program integrity is that these activities are often funded in one of the annual appropriations bills but are designed to reduce waste, fraud, and abuse in mandatory programs. If spending is added or reduced for an *existing program integrity activity*, then the guideline dictates that the impact on mandatory spending cannot be scored at all. This is essentially a “hole” in the scoring that creates a bias against the funding of program integrity and for a rise in waste, fraud, and abuse. The fix is to simply close the hole. Changes in discretionary spending on existing activities could be treated the same as when the program changes are substantive.

Scorekeeping guideline 14 stipulates that “no increase in receipts or decrease in direct spending will be scored as a result of provisions of a law that provides direct [mandatory] spending for administrative or program management activities.” Program integrity falls into this category of activities. This

---

3. CBO itself has flagged this issue in a 2014 report, “How Initiatives to Reduce Fraud in Federal Health Care Programs Affect the Budget” (Congressional Budget Office, Washington, DC, October 2014), 3. There CBO states that “[t]hose rules were established in large part to avoid crediting uncertain potential savings as offsets against very certain up-front spending (in case the hoped-for savings did not materialize). Nevertheless, those savings, if realized, ultimately reduce federal budget deficits.” Scott Levy also points out the possible distortionary effect of the scorekeeping in some detail in “Spending Money to Make Money: CBO Scoring of Secondary Effects,” *Yale Law Journal* 127, no. 4 (2018): 936–1008.

4. Drew C. Aherne et al., “Congressional Rules Pertaining to Changes in Mandatory Program Spending in Appropriations Bills (CHIMPs)” (CRS report R47705, Congressional Research Service, Washington, DC September 21, 2023).

means that if a bill would provide additional mandatory funding to combat fraud in a program with mandatory funding, only the costs of implementing that provision are counted; any savings that might result are not considered. Note that this rule is strangely asymmetrical in that it does not limit the full scoring of behavioral effects when legislation *decreases* mandatory spending or *raises* revenue for administrative or program management activities.

## Scorekeeping

Scorekeeping is the process of developing and recording measures of the budgetary effects of proposed and enacted legislation. The measures are 10-year estimates of the likely impact of legislation on federal outlays and revenues, compared with what would occur if the proposals were not enacted. This information, of course, is important for Congress to consider in drafting and voting on legislation and, because it is made public, offers transparency about Congress to the public. An additional key purpose of scorekeeping is to attribute likely budgetary effects to the legislation that caused them so that rules and procedures for budget enforcement may be applied. When Congress considers legislation, its official scorekeepers are the House and Senate Committees on the Budget. The committees, however, rely on budgetary estimates provided by the Congressional Budget Office.<sup>5</sup> To make accurate and unbiased projections, CBO staff members make an effort to understand and estimate how individuals, firms, and government entities would react to a policy. These so-called behavioral responses to legislation are an important part of budget scores.

CBO states that its objective is to produce accurate, unbiased, and consistent estimates that can be relied upon by the budget committees and members of Congress in making legislative decisions. This mostly means that CBO staff members have discretion in making estimates of the likely behavioral responses. They put a great deal of time and effort into developing methodologies to do this. While the basic goal of scorekeeping is to have the most accurate projections of budgetary changes, the CBO staff and the various agency scorekeepers must follow a process that has been developed over time and is governed by law,

---

5. The Congressional Budget Act of 1974 established that CBO will assist budget committees in their scorekeeping and requires the office to prepare estimates of the cost of legislation to support this work. It also stipulates that staff of the Joint Committee on Taxation (JCT) will provide the official estimates for all tax legislation. These estimates are typically passed on by JCT staff to CBO to include in the CBO cost-estimate memos. For a detailed description of the process used by CBO, see “CBO Describes Its Cost-Estimating Process” (Budget Primer, Congressional Budget Office, Washington, DC, April 18, 2023).

precedent, and rules. The process is described in the official scorekeeping guidelines. As long as the guidelines and additional budget committee directions are followed, CBO uses its judgment about what methodology to employ in making its estimates.

The scorekeeping guidelines have been developed over time by the House and Senate Budget Committees, CBO, and the other official scorekeeper, the Office of Management and Budget (OMB). The latter records its own official executive branch estimates of the budgetary effects of enacted laws. The guidelines were first published in the conference report for the Budget Enforcement Act of 1990. The act directed CBO and OMB to follow the same official list of guidelines that were intended to standardize scorekeeping estimates and reduce political influence on the methodologies used in the estimates. The act also directed the scorekeepers to review the guidelines every year and allowed the guidelines to be changed by unanimous consensus. There are currently 16 guidelines, but there are generally no official scorekeeping meetings. There were no meetings in the period during which I was the director of CBO (2015–2019) and, to my knowledge, there have been no meetings since. The guidelines are published every year in “Preparation, Submission, and Execution of the Budget.”<sup>6</sup>

### Why Scoring Accuracy Is Important

For the budget process in Congress to work well, scorekeeping estimates of the effects of proposed legislation must be unbiased and accurate. A reputation for meeting these requirements is also important. CBO puts a great deal of effort and resources into improving the accuracy of its estimates and developing as many new methodologies as needed and possible to capture the likely behavioral responses to legislation. The staff members routinely solicit advice from the research community generally, from some very capable advisory committee members, and sometimes from broad calls for research assistance on particular topics.<sup>7</sup> The budget committees hold CBO accountable for the accuracy and unbiasedness of its estimates and provide routine oversight of the agency.<sup>8</sup> CBO

---

6. Office of Management and Budget, “Preparation, Submission, and Execution of the Budget, Appendix A” (Executive Office of the President, Circular A-11, August 2023).

7. See, for example, Noella Duchovny et al., “Estimating the Effects of Federal Policies Targeting Obesity: Challenges and Research Needs,” *CBO Blog*, October 26, 2015, or Joseph Kile, “A Call for New Research on Energy and the Environment,” *CBO Blog*, July 17, 2023.

8. See, for example, Congressional Budget Office, “Answers to Questions for the Record Following a Hearing Conducted by the House Committee on the Budget on CBO Oversight: The Role of Behavioral Modeling in Scoring and Baseline Construction” (report, Washington, DC, July 26, 2018).

also provides a great deal of detail on its methodologies in scoring legislation; regularly evaluates the bias and accuracy of its overall budget projections;<sup>9</sup> and sometimes, but not routinely, evaluates the accuracy of budget scores of individual pieces of enacted legislation.<sup>10</sup> While projecting the impact of legislation may, at times, involve a good deal of uncertainty, CBO explicitly describes its estimates as designed to be “as likely to be too high as too low.” The staff members routinely exercise their professional judgment on the methodology of their estimates with an eye to accuracy and unbiasedness. They will include projections for as many behavioral responses to the legislation as possible but must constantly make judgments about whether their estimate will be significant or accurate enough to include.

For Congress to do its job, an accurate budget score is important. A key purpose of scorekeeping is to attribute likely budgetary effects to the legislation that caused them so that rules and procedures for budget enforcement may be applied. There are often caps on discretionary spending, pay-as-you-go (PAYGO) requirements for mandatory spending, and procedural requirements for bills to advance through the legislative process. PAYGO requires that any legislation estimated to increase mandatory spending or reduce revenues must be offset so that it does not increase the deficit. This issue is discussed in a little more detail later in this study. An example of a bill’s procedural requirement is the use of scoring during the annual budget process to enforce budget resolutions. Budget resolutions set an overall level for spending and revenue as a target for the subsequent detailed appropriations bills. Scoring of the appropriations bills that reveal a violation of the resolution can lead to an individual member of Congress filing a budget “point of order” to block the legislation.

A budget score is so important that, for every major piece of legislation, congressional staff members extensively negotiate behind the scenes with CBO staff members, who make sometimes dozens of draft estimates that the public never sees. Quite often, legislative proposals are even written with predetermined spending or revenue changes in mind. In this case, the crafting of legislation is an effort to “fine-tune” the language so that its estimated cost or saving equals a target score. Budget estimates for legislation are also important politically, as all are made publicly available in detail when they are finalized. Many

---

9. See, for example, Congressional Budget Office, “The Accuracy of CBO’s Budget Projections for Fiscal Year 2022” (report, Washington, DC, January 9, 2023). Many similar reports may be found on the CBO “Accuracy of Projections” website, <https://www.cbo.gov/topics/budget/accuracy-projections>.

10. See, for example, Congressional Budget Office, “A Review of CBO’s Estimate of the Effects of the Recovery Act on SNAP” (report, Washington, DC, December 20, 2018).

scores get press coverage, and the coverage may be quite extensive for high-profile legislation.<sup>11</sup> Scores can clearly affect the chances of a bill's success. All of this makes official scores important for Congress, the president, and the federal agencies receiving the funding.

To understand scoring, one must understand the difference between mandatory and discretionary spending. Mandatory spending is established by prior law and allows spending to continue until the program expires or a subsequent law either terminates the program or reauthorizes it. Most continue automatically throughout the normal 10-year budget period. Discretionary spending, however, must be voted on in the annual appropriations process, typically in an appropriations bill. About one-third of federal spending is discretionary and includes most of the direct activities of the federal government.

The accurate scoring of legislation that changes discretionary spending is important when there are spending caps that place limits on subcategories of discretionary spending (e.g., defense versus nondefense). The first time spending caps were used was in the Budget Enforcement Act (BEA) of 1990 to help restrain appropriations and reduce the budget deficit. Caps were placed on three categories of spending: defense, international, and domestic.<sup>12</sup> If the limits were exceeded on any of those categories, the BEA provided an enforcement mechanism, known as *sequestration*, whereby the president would be required to cancel enough spending to bring total funding back to the cap levels. The caps were to expire in 1995 but were extended and altered until they eventually expired in 2002. Budget caps returned with the Budget Control Act (BCA) of 2011, which placed caps on discretionary spending from 2012 through 2021. Not all discretionary programs were constrained by the caps. Funds deemed “emergency spending” were exempt, as were Overseas Contingency Operations (mainly funding for activities in Afghanistan and similar missions); disaster relief and activities related to wildfire suppression; and program integrity initiatives. In theory, the latter exception was intended to give Congress an incentive to adequately fund program integrity, and about \$14 billion on such programs that would have been stopped by the caps was allowed.<sup>13</sup> The Fiscal Responsibility Act (FRA) of 2023 included provisions that again establish discretionary spending caps for fiscal years 2024

---

11. See, for example, the extensive news coverage of the scoring for the 2017 legislation that attempted to amend the Affordable Care Act and for the Tax Cuts and Jobs Act of 2018.

12. The categories were later changed to just defense and nondefense.

13. Megan S. Lynch, “Were the Discretionary Spending Caps Effective?” (CRS Insight IN12093, Congressional Research Service, Washington, DC, February 9, 2023).



and 2025. As with the earlier spending caps, the FRA designates some spending as exempt. In particular, spending as an emergency requirement and for certain purposes, such as disaster funding, program integrity initiatives, and reemployment services, is not subject to the current caps.<sup>14</sup> The problem with both of these cap exceptions is that, while direct spending on program integrity is allowed, there is still no accurate scoring of the full effects of the spending on the government cost of waste, fraud, and abuse.

Congress has recognized that the limitations imposed by scorekeeping guidelines 3 and 14 will create a bias against reducing waste and fraud when spending caps are in place. It has not, however, fixed the problems with the guidelines directly but instead has tried to offer exemptions for program integrity spending (as seen in the two acts previously discussed) in specific legislation setting up spending caps. Congress has done this many times, but the limitations have not been very effective.<sup>15</sup> Table 1 reproduces a table from Scott Levy’s article “Spending Money to Make Money” that summarizes the cap exemptions used by Congress. Nevertheless, they are incomplete fixes for the cap problem created by the guidelines, require a fair amount of extra legislative work, aren’t used as much as needed, and do nothing for the basic lack of information problem created by the guideline holes.

These cap exemptions have been incomplete fixes to the problem of scoring program integrity. They, of course, help out with the cap problem, but not with the other issues with the biased scoring of program integrity. While the exemptions allow spending on program integrity to exceed a cap level, there is no “budget credit” for the reduction in waste and fraud. Plus, the cap exemptions help create more spending but do nothing to prevent a decrease in spending for program integrity; that is, decreasing spending gives Congress a biased reduction in scored spending because the resulting rise in the cost of waste and fraud is ignored. In addition, the cap exemptions have offered only a prespecified level of program integrity funding to an agency and are legislatively costly—after a cap adjustment for program integrity is set, sufficient funds must be appropriated for the program, and then the adjustment must be used. This involves two separate bills, and either one could fail to pass. The process is therefore costly and can easily fail to get enough support.

---

14. Grant A. Driessen, “Discretionary Spending Caps in the Fiscal Responsibility Act of 2023” (CRS Insight IN12168, Congressional Research Service, Washington, DC, June 28, 2023).

15. For a more detailed discussion of the use of cap exemptions for program integrity, see Levy, “Spending Money to Make Money” (965–74).

TABLE 1. Historical cap adjustments created for four programs by year

Year	Internal Revenue Service	Unemployment Insurance	Health Care Fraud and Abuse Control Program	Continuing Disability Reviews
1991	Y			
1992	Y			
1993	Y			
1994	Y			
1995	Y			
1996				Y
1997				Y
1998	Y			Y
1999	Y			Y
2000	Y			Y
2001	Y			Y
2002	Y			Y
2003				
2004				
2005				
2006	Y	Y	Y	Y
2007				
2008	Y	Y	Y	Y
2009	Y	Y	Y	Y
2010	Y	Y	Y	Y
2011				
2012			Y	Y
2013			Y	Y
2014			Y	Y
2015			Y	Y
2016			Y	Y
2017			Y	Y
2018			Y	Y
2019			Y	Y
2020			Y	Y
2021			Y	Y

Source: Scott Levy, "Spending Money to Make Money: CBO Scoring of Secondary Effects," *Yale Law Journal* 127, no. 4 (2018): 936-1008.

Accurate scorekeeping is also important when applying PAYGO rules. Such rules are designed to force Congress and the president to prioritize spending items and exercise fiscal restraint. There have been PAYGO rules off and on by statute (starting with the BEA of 1990) and in congressional rules with both the House and the Senate. They generally require that any legislation estimated to increase mandatory spending or reduce revenues must be offset so that it does not increase the deficit. Instead, additional offsetting provisions must be added to the legislation to fix the problem by lowering mandatory spending, raising revenue, or both. The current active statutory PAYGO started in 2010. Statutory PAYGO violations are determined by the executive branch scorekeeper, OMB.<sup>16</sup>

### Limitations on Scorekeeping Estimates

Scoring requires CBO to assess the extent to which proposed legislation changes or creates policies that would affect individuals, firms, and government entities and how the individuals, firms, and government entities, in turn, would change their behavior in ways that would affect federal revenues or spending. These are the so-called behavioral responses to the legislation. CBO staff members generally use their technical expertise to estimate the budgetary impact of behavioral responses as best they can. They must decide when the budgetary impact is too small or too uncertain to estimate. There are two types of behavioral responses, however, for which they are not allowed to do this. The first is dynamic analysis, in which legislation will change the GDP, and the second is the effect of program integrity spending designed to lower the level of waste, fraud, and abuse in federal programs.

The first constraint on scoring does not come from the scorekeeping guidelines but from the budget committees themselves. It is the absence of estimating and including the macroeconomic effects of new legislation as part of the behavioral responses in official scores. Including the GDP effect is usually referred to as *dynamic scoring*. Dynamic scoring has received a great deal of attention and would permit CBO and the Joint Committee on Taxation (JCT) to estimate the effects of large pieces of legislation on the GDP. Changes in the GDP would generally affect the federal budget, so as with other behavioral responses, the addition of the GDP effect might create a secondary effect of the legislation on the budget. For much of the time since 1997, congressional rules

---

16. Megan Carroll and Justin Humphrey, “The Statutory Pay-As-You Go Act and the Role of the Congress” (Congressional Budget Office, Washington, DC, August 18, 2020).

have asked for dynamic scoring under certain circumstances but for informational purposes only. In 2023, the House reinstated the dynamic scoring requirement for CBO and JCT. As before, it's only for "major legislation" and for informational purposes. Not using a dynamic score may have an impact on both the unbiasedness and the accuracy of estimates of budget effects. Opponents of dynamic scoring generally argue that these estimates are too uncertain to be relied on for budget enforcement. Their use, the critics argue, would lower the accuracy of scorekeeping and make comparisons between dynamically scored and nondynamically scored legislation difficult. Proponents of dynamic scoring argue that its official use in scorekeeping would increase the accuracy of legislative changes that are likely to affect economic productivity in the private sector. Further, they point out that not using dynamic analysis creates a bias *against* policy proposals designed to raise productivity in the economy and a bias *for* proposals that would lower productivity.<sup>17</sup>

A second constraint that is placed on official scorekeeping comes from two guidelines that limit CBO's ability to estimate the effect of legislation on mandatory spending. These are guidelines 3 and 14. It turns out that either one can stand in the way of accurate and unbiased estimates of the impact of program integrity on waste, fraud, and abuse. The concern is that Congress may be underinformed on the full effects of spending on program integrity and may miss opportunities to counter a rise in waste, fraud, and abuse in a mandatory program. Normally, appropriations bills control discretionary spending, while mandatory spending is controlled in other laws. When changes in mandatory spending are included in an appropriations act, however, guideline 3 dictates the scoring. It reads as follows:

**Guideline 3: Direct Spending Programs**

Revenues, entitlements and other mandatory programs (including offsetting receipts) will be scored at current law levels, as defined in section 257 [of the Balanced Budget and Emergency Deficit Control Act of 1985, also called the Deficit Control Act], unless congressional action modifies the authorizing legislation. Substantive changes to or restrictions on entitlement law or other mandatory spending law in appropriations laws will be scored against the Appropriations Committee's section 302(b) allocations in the House and the Senate. For the purpose of

---

17. See Lynch and Gravelle, "Dynamic Scoring in the Congressional Budget Process," for a discussion of the issue and the current congressional requirements for the use of dynamic scoring.

[scoring under the Congressional Budget Act of 1974], direct spending savings that are included in both an appropriations bill and a reconciliation bill will be scored to the reconciliation bill and not to the appropriations bill. For scoring under sections 251 or 252 of [the Deficit Control Act], such provisions will be scored to the first bill enacted.<sup>18</sup>

Scorekeeping guideline 3 is nicknamed the “fingerprint rule.” It applies when changes in mandatory spending are included in an appropriations act. If the changes are substantive, then they are treated as discretionary spending for purposes of budget enforcement. These provisions are referred to as CHIMPs. They are commonly used as an offset for discretionary spending.<sup>19</sup> However, if the legislative language changes only the level of discretionary spending for an existing program without substantive changes or restrictions on the mandatory program, then the effects on mandatory spending are not scored at all. This leaves a “hole” in the official score in which significant budgetary effects may be ignored. This hole can have a particular effect on the scoring of legislation designed to increase or decrease spending on program integrity. Program integrity activities are designed to prevent fraud, waste, and abuse of government resources. As such, higher or lower spending on the former may be expected to lower or raise spending on the latter. Because of its nature, program integrity activity is often funded in annual agency budgets by the appropriations process but is designed to control fraud, waste, and abuse in mandatory programs. Thus, it crosses that discretionary/mandatory line described in guideline 3.

To illustrate the distortion created by the scorekeeping guidelines, CBO used the concept of return on investment (ROI) to discuss how it would have to score an increase in additional antifraud funds in the Health Care Fraud and Abuse Control (HCFAC) program. Through the program, about \$1.4 billion was spent in 2014 to reduce fraud and other improper payments. Over the prior few years, lawmakers had been adding an average of \$300 million in additional funding to this program through annual appropriations bills. As CBO said in its report from October 2014,<sup>20</sup> the office expected that “increased funding for antifraud activities through the HCFAC program would result in reductions in [mandatory] direct spending for health care that exceed the amount invested; that [posi-

---

18. Congressional Budget Office, “CBO Explains Budgetary Scorekeeping Guidelines” (report, Washington, DC, January 2021), 2.

19. Aherne et al., “Congressional Rules Pertaining to Changes in Mandatory Program Spending.”

20. Congressional Budget Office, “How Initiatives to Reduce Fraud in Federal Health Care Programs Affect the Budget,” 10.

TABLE 2. Projected changes in spending and revenues from increased IRS funding in the 2022 Reconciliation Act

	By fiscal year, millions of dollars											
	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2022-2026	2022-2031
Increases in direct spending												
Subtitle A. Deficit reduction												
Part 3. Funding the Internal Revenue Service and improving taxpayer compliance												
Sec. 10301. Enhancement of Internal Revenue Service resources												
Budget authority	79,622	0	0	0	0	0	0	0	0	0	79,622	79,622
Estimated outlays	0	3,823	3,380	4,970	6,248	7,996	10,106	12,617	15,072	15,388	18,421	79,600
Memorandum:												
Nonscorable increases in revenue												
Enhancement of Internal Revenue Service resources	0	2,012	5,106	11,125	16,116	21,716	26,314	31,218	34,877	31,904	34,359	180,388

Source: Congressional Budget Office, “Estimated Budgetary Effects of Public Law 117-169, to Provide for Reconciliation Pursuant to Title II of S. Con. Res. 14,” September 7, 2022.

tive] return on investment would generally be realized over several years.” In fact, CBO believed that such spending would have an ROI factor of about 1.5 to 1. That is, every dollar increase in spending on the HCFAC would lower fraud and other improper payments by \$1.50. However, the authors explained, those reductions cannot be counted as decreases in mandatory spending in official scores because of scorekeeping guideline 3. So, if legislation increased funding for the HCFAC program by \$100 million and CBO was asked to provide an unofficial estimate of the benefits of the new spending, the impact might look like what is shown in table 2.<sup>21</sup> Here, the CBO example supposes that it would take three years to spend the added funds, with most (\$80 million) spent in the first year and yielding an estimated savings of \$150 million, which would accrue over six years. Those savings would not count for the purpose of enforcing budget rules. For example, if discretionary spending caps did not have an exemption for program integrity spending, Congress would also not be able to spend the additional \$100 million without finding some offsetting spending reduction or revenue increase.

21. Congressional Budget Office cost estimate for P.L. 117-169 as enacted on August 16, 2022, “Estimated Budgetary Effects of Public Law 117-169, to Provide for Reconciliation Pursuant to Title II of S. Con. Res. 14,” September 7, 2022.

The second problem with how the scorekeeping guidelines treat program integrity spending comes from guideline 14, which relates to increases in mandatory spending or declines in revenue that involve spending on program management or administration.

**Guideline 14: Scoring of Receipt Increases or Mandatory Spending Reductions for Additional Administrative Program Management Expenses**

No increase in receipts or decrease in direct spending will be scored as a result of provisions of a law that provides direct [mandatory] spending for administrative or program management activities.<sup>22</sup>

Guideline 14 says that if there is a proposed rise in mandatory spending through changes in authorizing legislation on administrative or program management activities, then any resulting reduction in spending or rise in revenues may not be included in official scoring estimates. Note that this is an asymmetric rule in that it does not limit the full scoring of behavioral effects when legislation decreases mandatory spending or raises revenue for administrative or program management activities. A recent example of this was from the 2022 Inflation Reduction Act (P.L. 117-169). An increase in mandatory spending for Internal Revenue Service (IRS) enforcement activity was included in the bill, but the official score included it as just a rise in spending of \$79.6 billion.<sup>23</sup> However, CBO believed that the increase in enforcement activity would, through its behavioral effects, raise tax revenue by considerably more than the cost of additional enforcement funding. Guideline 14 prevented its formal inclusion for budget enforcement purposes. As is sometimes the case, CBO added an “unofficial” estimate of the nonscorable increase in revenues of \$180 billion. This is an ROI in IRS enforcement of 2.3 to 1.<sup>24</sup> For budgetary enforcement purposes, it was treated as a \$79.6 billion rise in the deficit, but the actual projection was to lower it by \$100 billion.<sup>25</sup> Another example cited by CBO is the SGR Repeal and Medicare Beneficiary Improvement Act of 2013.<sup>26</sup> That act

---

22. Congressional Budget Office, “CBO Explains Budgetary Scorekeeping Guidelines,” 3.

23. See the CBO estimate for P.L. 117-169 from September 7, 2022.

24. A useful concept for discussing the benefits and cost of program integrity activity is its ROI. Several agencies actively estimate the ROI for their program integrity activities to justify the cost—that is, they estimate the reduction in government spending from these activities and compare that with the cost.

25. For more on this issue, see the presentation by Janet Holtzblatt, “Estimating the Revenue Effects of Proposals to Increase Funding for Tax Enforcement” (CBO presentation, June 23, 2016).

26. Congressional Budget Office, “S. 1871, SGR Repeal and Medicare Beneficiary Improvement Act of 2013” (CBO cost estimate, January 24, 2014).

would have provided mandatory funding to reduce improper payments in Medicare and Medicaid and reduced spending for benefits. However, guideline 14 did not permit the overall savings in spending to be part of the official score for budgetary rule enforcement.

### How Program Integrity Spending Could Be Analyzed

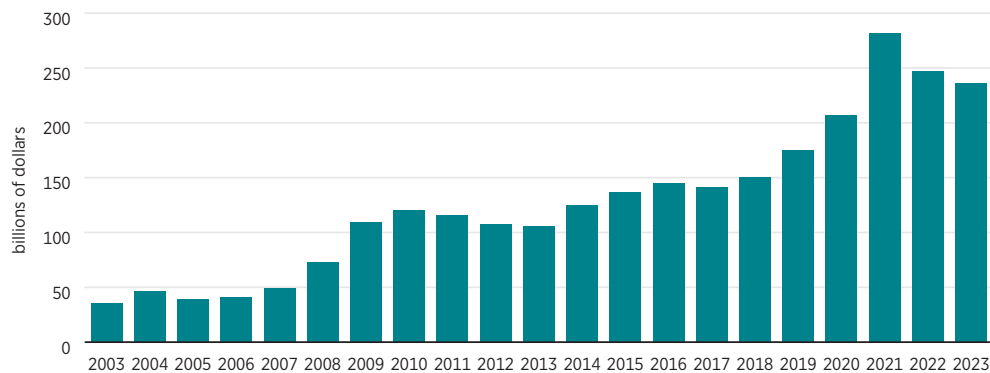
There are many federal programs for which estimating the behavioral responses to the budgetary legislation is difficult, and CBO must make its best effort to offer an accurate, unbiased estimate of the budgetary effects. Just as an example, note the CBO scoring of types of federal spending on disease prevention.<sup>27</sup> Government spending on preventive medical care includes services that can prevent diseases from occurring and or detect them before symptoms appear. For example, the preventive services could be vaccinations or medical screenings. Just as with other legislation, when legislative proposals affect such services, the CBO role is to project the 10-year federal budgetary effects of the legislation. And much like spending on program integrity, the costs of providing preventive medical services are part of the scorekeeping estimates. In this case, delivering such services results in costs for each person using the service. So, the direct cost of vaccinations and medical screenings are estimated. This results in improved health outcomes where disease is avoided or treated earlier than would have happened without screening. The cost of future healthcare is therefore reduced, so federal spending on disease treatment is likely to decline as a result. Unlike with program integrity spending, the full net effects of this spending can be taken into account in the official score. The overall rise in federal spending will be less than just the cost of preventive medical services. In fact, the decline in future healthcare spending might entirely wipe out the cost so that it results in overall budgetary savings. However, CBO notes that, historically, about 80 percent of preventive medical services have been found to still lead to higher healthcare spending overall. If the full budgetary effects of the behavioral response to the spending on better health outcomes were not counted, there would be a scorekeeping bias against spending on preventive medicine. The reverse would also be true. A reduction in spending on preventive health services would appear to reduce more federal spending if the rise in bad health outcomes were not scored.

---

27. See, for example, Congressional Budget Office, “How CBO Analyzes Approaches to Improve Health through Disease Prevention” (report, Washington, DC, June 2020).



FIGURE 1. Improper payments



Source: Figure 1 in Government Accountability Office, “Improper Payments: Information on Agencies’ Fiscal Year 2023 Estimates” (GAO-24-106927, Washington, DC, March 26, 2024).

### Dealing with the Surge in Waste, Fraud, and Abuse

Government programs need to have the ability to allocate resources to program integrity activities as the level of fraud, waste, and abuse rises. This effort includes focusing on activities that have the highest ROI in producing government savings. The United States has had an unprecedented surge in waste, fraud, and abuse, and thus spending on program integrity—activities designed to ensure that government programs deliver on their mission while reducing fraud, waste, and abuse—needs to be able to rise significantly to combat this surge. According to GAO,<sup>28</sup> “[w]hen the COVID-19 pandemic began, federal agencies acted swiftly to establish emergency programs and deliver financial relief to the American people. Many agencies were able to distribute funds quickly, but the tradeoff was that they did not have systems in place to prevent and identify payment errors and fraud.” The level of improper payments has now cost the government about \$970 billion over the past four years (see figure 1). While all of this has been a loss to the government, fraud was a particular source of additional loss.<sup>29</sup> In a recent report, GAO concluded that the pandemic, recession, and resulting surge

28. Government Accountability Office, “Emergency Relief Funds: Significant Improvements Are Needed to Ensure Transparency and Accountability for COVID-19 and Beyond” (GAO-22-105715, Washington, DC, March 17, 2022).

29. Improper payments are any payments that should not have been made or that were made in an incorrect amount. Fraud involves obtaining a thing of value through willful misrepresentation. While all fraudulent payments are considered improper, not all are due to fraud. Similarly, there are types of fraudulent activity that do not result in improper payments. See Government Accountability Office, “Improper Payments and Fraud: How They Are Related but Different” (GAO-24-106608, Washington, DC, December 7, 2023).

in related federal spending were responsible for a surge in fraudulent claims on federal programs. In the study, they surveyed agencies on efforts to reduce fraud and found that availability of sufficient staffing and funding was a challenge for 90 percent of them.<sup>30</sup>

Program integrity activities are designed to ensure that government programs deliver on their mission while reducing fraud, waste, and abuse. Improper payments are ones that should not have been made or were made in an incorrect amount. According to GAO, improper government payments alone have totaled \$2.6 trillion over the past 20 years.<sup>31</sup> An improper payment, whether it involves fraud or not, represents a waste of federal spending. Since 1990, GAO has kept and monitored its High Risk List of federal programs and operations vulnerable to waste, fraud, abuse, or mismanagement or needing broad reform, and the agency has encouraged programs to make sure that they have sufficient capacity—including staffing and funding—to reduce those levels. The problem has never been worse than it is right now. Currently, a record 37 federal programs are on the High Risk List, and there has been a surge in improper payments starting with the federal government’s \$4.7 trillion pandemic response in 2020.

As OMB states,<sup>32</sup>

[T]here is compelling evidence that investments in administrative resources can significantly decrease the rate of improper payments and recoup many times their initial investment for certain programs. In such programs, using adjustments to base discretionary funding for program integrity activities allows for the expansion of oversight and enforcement activities in the largest benefit programs, including Social Security, Unemployment Insurance, Medicare, and Medicaid. In such cases, where return on investment using discretionary dollars is proven, adjustments to base discretionary funding are a useful budgeting tool. Formerly, when statutory spending caps on the discretionary budget were in place under the BCA, the law allowed the caps to be adjusted upward to account for additional discretionary

---

30. Government Accountability Office, “Fraud Risk Management: Agencies Should Continue Efforts to Implement Leading Practices” (GAO-24-106565, Washington, DC, November 1, 2023).

31. Government Accountability Office, “Improper Payments: Fiscal Year 2022 Estimates and Opportunities for Improvement” (GAO-23-106285, Washington, DC, March 29, 2023). See also Matthew Dickerson, “Improper Payments Exceed Funding for the U.S. Army,” Economic Policy Innovation Center (blog), November 30, 2023.

32. Office of Management and Budget, *Analytical Perspectives, Budget of the U.S. Government, Fiscal Year 2024* (Washington, DC: Office of the President of the United States, March 13, 2023), 32.

funding that supported savings in these mandatory programs. These adjustments continue in congressional budget enforcement under the Congressional Budget Act and are called allocation adjustments. Such adjustments are needed because budget scoring rules do not allow the mandatory savings from these initiatives to be credited for budget enforcement purposes.

## Unemployment Insurance Programs

There is a particular need to increase program integrity spending in the unemployment insurance (UI) program. According to the Congressional Research Service (CRS),<sup>33</sup> waste, fraud, and abuse in unemployment insurance programs have long been a serious concern. Estimates of the rate of improper payments have been above 10 percent for most of the past two decades. The new UI programs related to the pandemic have caused this rate to surge—hitting 17.9 percent (\$74 billion) in 2021, and a total loss to the government from fraudulent payments has been between \$100 billion and \$135 billion over three years, according to GAO. This amount does not include another \$50.5 billion in nonfraudulent improper payments.<sup>34</sup> OMB still designates UI as a high-priority program, with estimated improper payments of more than \$100 million a year. The most recent budget request by the Department of Labor has argued that new and expanded tools and controls addressing fraud in the UI system would have a net savings of \$2 billion over the 10-year budget window.<sup>35</sup>

According to CRS, the pre-pandemic struggle with improper payments stemmed from two issues. The first was a decrease in administrative funding and staffing problems over time, along with an increased use of outdated automated systems. GAO estimated that funding available for state UI administration declined 32 percent in real terms from 2010 through 2019.<sup>36</sup> This was made worse by the continual push to prioritize the timeliness of payments during periods of

---

33. Congressional Research Service, “Unemployment Insurance Program Integrity: Recent Developments” (CRS In Focus IF12243, Washington, DC, November 2, 2022).

34. Government Accountability Office, “Unemployment Insurance: Estimated Amount of Fraud during Pandemic Likely between \$100 Billion and \$135 Billion” (GAO-23-106696, Washington, DC, September 12, 2023).

35. “OIG Oversight of the Unemployment Insurance Program,” Office of the Inspector General for the Department of Labor (website), updated December 15, 2023, <https://www.oig.dol.gov/doloiguioversightwork.htm>.

36. Government Accountability Office, “Unemployment Insurance: Transformation Needed to Address Program Design, Infrastructure, and Integrity Risks” (GAO-22-105162, Washington, DC, June 2022).

high unemployment. The second was that the added administrative burden came from the automatic expansion of permanent UI benefits during recessions and the new temporary benefits often added in response to higher unemployment rates. States faced difficulties in increasing staffing during these periods.

The new UI programs enacted during the pandemic meant that states struggled with the identity verification of claimants and the significant magnitude of benefits. For example, the Pandemic Unemployment Assistance (PUA) program provided unemployment benefits to workers not typically eligible to receive benefits, such as self-employed and gig workers. The program required applicants to self-certify their own eligibility for benefits without proof of income or identity. In addition to the challenges of identity verification in the PUA program, the Federal Pandemic Unemployment Compensation program benefits were \$300 or \$600 per week compared with the more common recession boost in benefits of \$25 per week (as in 2009 and 2010).

## Conclusion

Fully scoring spending on program integrity—whether through discretionary funding or mandatory funding—has some clear benefits. Two changes to the guidelines need to be made to achieve this goal. Some of guideline 3 applies to the existence of CHIMPs—substantive changes in mandatory programs that originate from an appropriations bill. The “hole” comes from instances in which the changes in mandatory spending are on an existing program. These should be treated the same as substantive changes in mandatory programs. In other words, they should simply become CHIMPs and be fully scored as changes in discretionary spending. The other change is to eliminate guideline 14.

## Implications

1. The accuracy of official scorekeeping is important for taxpayers, Congress, the president, and federal agencies receiving funding. Having a bias against spending on program integrity may lead to unnecessarily high levels of waste, fraud, and abuse. CBO scores continue to routinely leave out the full budgetary effects of changes in program integrity spending.<sup>37</sup>

---

37. For example, see the “Status of Discretionary Appropriations Report” for both the House and the Senate (Congressional Budget Office, Washington, DC, April 4, 2024), in which the CBO again points out that the estimated budgetary effects from appropriation legislation leave out the effects on mandatory programs.

2. The opportunities to lower waste, fraud, and abuse, with an increase in spending on program integrity, certainly exist. As this report has cited, there have been a number of instances in the past when CBO has estimated a high ROI for an unscorable increase in program integrity. While an ROI above 0 means that there is a scoring bias in the CBO estimate, an ROI above 1 means that there is an opportunity to actually lower the deficit.
3. Rising levels of waste, fraud, and abuse from new or expanded federal programs, seen during the past three years, will likely raise the expected ROI of higher levels of spending on program integrity. Allowing Congress to simply raise or lower funding for program integrity with full scorable credit will make it easier to moderate the effects of such abuses.
4. There are many programs for which the ROI of program integrity efforts is not known. Very significant opportunities may exist to raise spending on program integrity activities and provide overall budgetary savings because of significant reductions in waste, fraud, and abuse. Similarly, there may be activities with a low ROI, and federal spending would be better directed toward other things—including, possibly, more-effective program integrity activities.
5. Little incentive exists for federal agencies to establish an ROI for their program integrity activities because scorekeeping guidelines prevent Congress from giving those agencies “budgetary credit” for their success (or failure). This also applies to the incentive to improve the effectiveness of these activities to raise their ROI.
6. GAO spends considerable time and effort finding where significant waste, fraud, and abuse exist in government programs. When a program joins GAO’s High Risk List, it should be encouraged through its budget oversight to justify and improve its program integrity efforts. Fully scorable changes in funding would help this.

## About the Author

Keith Hall is a distinguished visiting fellow at the Mercatus Center at George Mason University. He has more than 25 years of public service, notably serving as the director of the Congressional Budget Office and commissioner of the Bureau of Labor Statistics.

In addition, he served as the chief economist and director of economics at the US International Trade Commission (USITC), chief economist for the White House's Council of Economic Advisers, chief economist for the Department of Commerce, and senior international economist for the USITC. He also previously worked as a senior research fellow at Mercatus. He has been a full-time faculty member at the Georgetown University McCourt School of Public Policy and in the economics departments at the Universities of Arkansas and Missouri.

## About the Mercatus Center at George Mason University

The Mercatus Center at George Mason University is the world's premier university source for market-oriented ideas—bridging the gap between academic ideas and real-world problems.

As a university-based research center, the Mercatus Center trains students, conducts research of consequence, and persuasively communicates economic ideas to solve society's most pressing problems and advance knowledge about how markets work to improve people's lives.

Our mission is to generate knowledge and understanding of the institutions that affect the freedom to prosper and to find sustainable solutions that overcome the barriers preventing individuals from living free, prosperous, and peaceful lives.

Since 1980, the Mercatus Center has been a part of George Mason University, located on the Arlington and Fairfax campuses.