DOUGLAS W. ELMENDORF
Brookings Institution[1]

"Dynamic Scoring": Why and How to Include Macroeconomic Effects in Budget Estimates for Legislative Proposals

ABSTRACT Official estimates of the budgetary effects of legislative proposals generally include anticipated behavioral responses except for those that would alter overall output or employment. Based on my experience as director of the Congressional Budget Office and on the analysis in this paper, I conclude that such macroeconomic effects of legislative proposals should be included in budget estimates—that is, so-called dynamic scoring should be used—for major (but not minor) proposals and for proposals affecting federal spending as well as revenues. However, such macroeconomic effects should not be included when the estimating agencies do not have the tools or time needed to do a careful analysis of those effects. Current rules governing the official estimating process do not fully meet those conditions.

When legislation is being developed in the U.S. Congress, the Congressional Budget Office (CBO) and the staff of the Joint Committee on Taxation (JCT) prepare estimates of the effects of that legislation on the federal budget. Those estimates often play a critical role in congressional deliberations and public discussion.

The estimates produced by CBO and JCT generally incorporate the effects of anticipated behavioral responses to the proposed changes in federal tax or spending policies. For example, estimates for changes in benefit programs include shifts in take-up rates among eligible people, and estimates for changes in income tax rates include shifts in the use of tax deductions. However, by long-standing convention, the estimates

1. The author was working at Brookings when he produced this paper; in January 2016 he became dean of Harvard's John F. Kennedy School of Government.

have excluded behavioral responses that would have macroeconomic effects, in the sense of altering overall output, employment, or similar variables. For example, CBO and JCT's original estimate of the budgetary impact of the Affordable Care Act (ACA) included the effects of employers altering the mix of taxable and nontaxable compensation provided to their employees but not the effects of employees altering their supply of labor.

The convention of excluding macroeconomic effects may seem odd from an economics perspective. Estimates for legislative proposals include behavioral responses in order to improve the accuracy of the predicted budgetary effects and to illuminate noteworthy nonbudgetary effects, and that rationale appears to apply equally to behavioral responses that affect overall output and those that do not. Indeed, some analysts and policymakers have argued for years that the estimates produced by CBO and JCT should include macroeconomic effects—an approach that has become known as "dynamic scoring." However, other analysts and policymakers have argued in response that including macroeconomic effects would degrade the quality and usefulness of CBO's and JCT's estimates.[2] That debate has achieved greater prominence recently because a rule adopted by the House of Representatives and the budget resolution approved by the House and the Senate both call for dynamic scoring in certain circumstances.

Based on my experience as the director of CBO from January 2009 through March 2015, I believe the principal concerns expressed about estimating the macroeconomic effects of proposals apply with equal force to other aspects of budget estimates or can be addressed by CBO and JCT. In my view, including macroeconomic effects in budget estimates for certain legislative proposals would improve the accuracy of those estimates and would provide important information about the proposals' economic effects. Moreover, if certain key conditions were satisfied, those estimates would meet the general goals of the estimating process, namely: that estimates be understandable and resistant to misinterpretation, that they be based on a consistent and credible methodology, that they be produced quickly enough

2. The advantages and disadvantages of dynamic scoring have been considered by numerous authors, including Auerbach (1996, 2005), Burman (2006), Committee for a Responsible Federal Budget (2012), CBO (1995, 2002), Furman (2006), Gale (2002), Hassett (2002), Holtz-Eakin and Mandel (2015), Ip (2015), Orszag (2002), and Van de Water and Huang (2014).

to serve the legislative process, and that they be prepared using the resources available to CBO and JCT.

Therefore, I conclude that the macroeconomic effects of legislative proposals should be included in budget estimates—that is, that dynamic scoring should be used—under the following conditions:

—Macroeconomic effects should be included in estimates only for *major* proposals, defined as those that would have a large estimated budgetary impact excluding macroeconomic effects, and when estimates of such effects are requested by the chair or ranking member of the House or Senate Budget Committee. CBO and JCT do not have sufficient staff or time to carefully analyze macroeconomic effects for every proposal under consideration, and using rules of thumb in place of careful analysis risks the credibility of the estimates.

—Macroeconomic effects should be included in estimates for major proposals affecting federal *spending* as well as revenues. Changes in either spending or revenues can have notable macroeconomic effects, and the estimating process should treat proposals affecting the two sides of the budget as comparably as possible subject to other constraints.

—Macroeconomic effects should *not* be included in estimates when CBO and JCT find that they do not have the tools or time needed to do a careful analysis of those effects. That situation will arise most often for proposals that are being developed and amended quickly and for proposals regarding certain types of regulatory policy in which the estimators do not have significant expertise.

Those conditions, and others discussed in this paper, can be readily satisfied. However, the current House rule and congressional budget resolution do not fully meet the specified conditions. The current requirements for dynamic scoring explicitly exclude appropriations bills (which cover about one-third of federal noninterest spending) and give only the chairs but not the ranking members of key committees the right to request the incorporation of macroeconomic effects in certain estimates. In addition, the threshold budgetary impact for presumptively including macroeconomic effects in estimates is lower than ideal, from my perspective.

There are advantages to an alternative approach in which CBO's and JCT's estimates of macroeconomic effects and their budgetary feedback would be provided in supplementary reports rather than being included in official budget estimates. In my judgment, though, the advantages of that alternative approach, as compared with the agencies' current plans for dynamic scoring, are limited and are outweighed by significant disadvantages.

I. The Basics of Budget Estimates for Legislative Proposals

CBO and JCT provide the official estimates used by Congress of the effects of legislative proposals on the federal budget. The estimates are based on procedures that have been developed over time and on the professional judgment of the two agencies. The analysts at CBO and JCT stay in their jobs regardless of political shifts in the control of Congress (although the director of CBO and the chief of staff for JCT are chosen by the congressional leadership), and the organizations have strong reputations for providing objective, nonpartisan analysis.

I.A. The Mechanics of Estimates

CBO, which began work in 1975, produces public estimates for bills after they have been approved by congressional committees or before they are voted on by the full House or Senate. For bills that would alter the tax code, CBO is required by its founding statute to use revenue estimates provided by JCT, which was created in 1926; for bills that would alter spending policies, CBO uses its own estimates; and for bills that would make changes in both tax and spending policies, the agencies prepare estimates together. The estimating process is sometimes referred to as "scoring," and the estimates are called "cost estimates." In addition to those public estimates, the agencies provide private estimates to members of Congress and their staffers for proposals that are being developed and have not been released publicly. In a typical year, CBO publishes between 500 and 600 public estimates, and it and JCT give committees thousands of private estimates for legislation under development.

Each estimate shows effects relative to the "baseline," which is CBO's projection of what would occur in the absence of the proposal. The baseline generally reflects current law, although Congress has specified certain exceptions.

The estimates present changes in nominal cash flows for the current fiscal year and each of the 10 subsequent years, a period that is often called the "budget window."[3] The use of cash flows and a limited time period

3. The principal exception to this statement is estimates for federal credit programs, which are based on the accrual of financial commitments by the federal government (CBO 2012a). All cost estimates exclude changes in federal interest payments that would result from changes in federal borrowing. However, CBO includes changes in interest payments when it provides estimates for overall budget packages, as in its annual analysis of the President's budget proposals.

mean that estimates do not always measure a proposal's full budgetary effect. Indeed, some proposals deliberately delay cash costs beyond the budget window or accelerate cash receipts into the budget window in order to lower the apparent budgetary impact.[4] However, when Congress is especially interested in a proposal's long-term budgetary effects, the agency tries to provide information about those effects. And when CBO expects that a major proposal would have notably different budgetary effects beyond the coming decade than during the decade, the agency can provide information about those effects without a specific request from the Congress.

The estimates are point estimates that are intended to show what is colloquially described as "the middle of the distribution of possible outcomes," but is specifically the mean outcome as judged by the agencies (CBO 1999). Although CBO and JCT are acutely aware of the uncertainty of estimates, the agencies focus on point estimates because the budget process and the procedural rules of the House and Senate rely on point estimates and because measuring the uncertainty of estimates is often especially difficult.

I.B. Behavioral Responses

CBO's and JCT's estimates generally include the impact of behavioral responses to the proposed changes in law—that is, the estimates are not based on an assumption that the economy is static. For example, estimates of changes in benefit programs include shifts in take-up rates for those benefits among eligible people, and estimates for changes in income tax rates include shifts in the use of tax deductions. More generally, CBO and JCT try to account for the behavior of households, businesses, federal regulators, and state, local, and foreign governments. However, the agencies do not attempt to predict future changes in federal law. CBO and JCT estimate the magnitude of behavioral responses using a broad range of evidence, including formal statistical analyses done by the agencies themselves and by other researchers as well as anecdotal information from consultations with government agencies and private businesses (CBO 2011; JCT 2011a).

The scope of the included behavioral responses varies greatly across estimates. Some potential responses are omitted because the available evidence does not indicate the order of magnitude or even the sign of a response.

4. For example, in a policy change known as "pension smoothing," companies are allowed to defer required payments into pension funds, thereby increasing their reported profits and thus tax payments in the budget window while reducing them later.

Other potential responses are omitted because CBO or JCT do not have the time or resources to collect the available evidence and build and apply an appropriate model. Still other potential responses are omitted from estimates because including the responses would greatly complicate the legislative process; for example, the effects that certain changes in spending would have on taxable incomes and thus on revenues are omitted from estimates because including them would generate jurisdictional conflicts between committees.

In addition, some potential behavioral responses are excluded from estimates because the responses would affect overall output, and overall output has been held fixed in cost estimates by long-standing convention. Therefore, CBO's and JCT's estimates have not included the budgetary effects of changes in labor supply, consumption, saving, productivity, and other aggregate variables. Under dynamic scoring, this convention of "fixed output" would be dropped. The principal exception to this convention before 2015 was the production of estimates for comprehensive immigration legislation in 2006, 2007, and 2013. In CBO's view, "assuming that those bills would have had no effect on overall output would have ignored one of the primary effects of the bills and distorted those estimates too severely" (CBO 2015j).

I.C. Current Status of Dynamic Scoring

In early 2003, the House adopted rule XIII.3.(h)(2), which required JCT to provide an analysis of the macroeconomic impact of all tax legislation approved by the Ways and Means Committee. That rule was adopted again by subsequent Congresses and remained in effect in the House through 2014. JCT (2015a, p. 12) summarized its response to this rule as follows:

> For most tax bills, the expected effects were so small that a brief statement to that effect was all that was required. Short qualitative analyses were provided for legislation that JCT macro models were not configured to model. For major tax legislation, JCT staff has provided detailed quantitative analysis of a possible range of effects on GDP, employment, investment, and revenues, based on the results of multiple models using multiple parameter assumptions.

In early 2015, the House modified that rule to require dynamic scoring by CBO and JCT for "major" legislation, defined as legislation that would have significant estimated budgetary effects or was designated as major by the chair of the Budget Committee or the Ways and Means Committee. The rule excludes appropriations bills, requires a qualitative assessment of

budgetary impact including macroeconomic effects for 20 years beyond the 10-year budget window, and includes the caveat that the requirements be met "to the extent practicable."

In the spring of 2015, the House and Senate approved a budget resolution that included requirements for dynamic scoring similar to those in the House rule. Under the budget resolution, CBO and JCT will, to the greatest extent practicable, incorporate the budgetary effects of changes in macroeconomic variables resulting from legislation that has a "gross" budgetary effect of a quarter of a percent of output in any year over the next 10 years, or is selected for such analysis by the chair of the House or Senate Budget Committee. That threshold equals about $45 billion in 2015 and about $70 billion in 2025 based on projected output (CBO 2015a); I address the interpretation of "gross" budgetary effects later in the paper. The resolution excludes appropriations bills and requires a qualitative assessment regarding the two decades following the budget window.

CBO (2015i) summarized the agency's plans for meeting the requirements of this budget resolution. In June 2015, in response to a request from the Senate Budget Committee, CBO and JCT applied dynamic scoring to a proposal to repeal the Affordable Care Act (CBO 2015h). And in August 2015, pursuant to the resolution, JCT applied dynamic scoring to a bill approved by the Senate Finance Committee that would extend for 2 years a number of tax credits, deductions, and exclusions that primarily affect businesses (JCT 2015b; CBO 2015k).

II. CBO's and JCT's Past Estimates of Macroeconomic Effects and Their Budgetary Feedback

The arguments for and against dynamic scoring can be understood best after briefly examining CBO's and JCT's past analysis of proposals' macroeconomic effects and their budgetary feedback. Most of that analysis has appeared not in cost estimates—given the conventional exclusion of macroeconomic effects from such estimates—but rather in supplemental reports.

II.A. Analytic Approach

CBO (2014d) provided an overview of its methodology for estimating the macroeconomic effects of legislative proposals in a November 2014 report, "How CBO Analyzes the Effects of Changes in Fiscal Policies on the Economy"; that report referred to a set of other reports describing specific aspects of the agency's methods (CBO 2001, 2012d, 2012f, 2012g, 2012h, 2013e, and 2014b). JCT (2003b, 2005, 2006, 2011a, 2011b, and

2015a) has summarized its methodology as well. In addition, in each report that presents such a macroeconomic analysis, CBO and JCT explain the key factors affecting the estimates.[5]

The approaches used by the two agencies differ in various specifics but are quite similar in their overall structure. For estimating the *short-term* effects of changes in fiscal policies, the agencies focus on effects on the demand for goods and services (and also include effects from changes in labor supply). Reductions in taxes and increases in federal spending boost demand directly, while the opposite changes diminish it; those direct effects propagate through the economy to an extent that depends on the response of monetary policy and other factors.[6] Changes in demand are estimated to lead to changes in output relative to potential output.

For estimating the *longer-term* effects of changes in fiscal policies, the agencies examine effects on potential output. Both agencies use a Solow-type growth model and a life cycle (overlapping generations) growth model.[7] In its Solow-type model, CBO focuses on the effects of changes in federal borrowing, marginal and average tax rates (through income and substitution effects), transfer payments (through income effects and, in some cases, substitution effects), and federal investment in physical infrastructure, education and training, and research and development. For example, an increase in the marginal tax rate on labor income is estimated to reduce the supply of labor, which in turn reduces capital accumulation. Similarly, JCT's Solow-type model captures responses to changes in federal borrowing, marginal and average tax rates, and other factors. In their life cycle models, CBO and

5. There are recurring calls for CBO and JCT to be more transparent regarding many aspects of their analyses. However, achieving greater transparency would require the agencies to allocate more of their resources to explaining existing estimates rather than producing new ones, and Congress has been reluctant to accept that trade-off.

6. CBO projects actions by the Federal Reserve as part of its baseline economic projections, and the agency has explained its method for estimating the Federal Reserve's reaction to changes in fiscal policies. For example, CBO (2015g) expects that the negative short-term effects of deficit reduction on output (stemming from a decrease in demand) will be "stronger when short-term interest rates are near zero . . . because under those conditions the Federal Reserve is unlikely to adjust short-term interest rates to try to offset the effects of changes in federal spending and taxes" (p. 88).

7. JCT (2011a, 2011b) also sometimes uses a growth model with infinitely lived agents. Separately, CBO (2014d, p. 12) discusses the possibility that changes in demand in the short term could affect potential output in the long term and concludes that the significance of the channels through which that might occur are "unclear," and thus "CBO does not currently incorporate such channels in its analyses, although the agency continues to investigate the issue."

JCT include many of the same channels, although expectations also matter explicitly—which presents a challenge for dynamic scoring that I return to below.

CBO generally reports both a central estimate and a range of estimates for the macroeconomic effects of proposals. The range is based on the values for the key parameters found in the research literature; the range for each variable "is intended to cover roughly the middle two-thirds of the likely values for the variable" (CBO 2015g, p. 73). The central estimate is intended to represent the middle of the distribution of possible outcomes (and can give weight to estimates from both the Solow-type model and the life cycle model).

To estimate the feedback from economic changes to the federal budget, CBO accounts for the impact of changes in income on tax revenues and benefits (with the latter much less affected than the former), as well as other factors. A one-dollar increase in overall output reduces the budget deficit by roughly 20 to 25 cents, holding all else equal.[8] The estimated budgetary effects in a given year influence estimated economic developments in subsequent years.

A key challenge for CBO and JCT is assessing the changes that proposals would generate to effective marginal tax rates on labor and capital, the income of people with different propensities to consume, differences in tax rates across types of capital, and changes in federal investment. CBO and JCT also modify their models as needed to capture the features of specific proposals. That process—as undertaken, for example, in the analyses of immigration reform and tax reform discussed below—sometimes requires a great deal of time and effort. In addition, the agencies adjust parameter values over time in response to new evidence; for example, see CBO's (2012g) paper "A Review of Recent Research on Labor Supply Elasticities."

CBO's and JCT's analyses of the macroeconomic effects of proposals are generally produced on much longer timetables than their budget estimates. That difference arises both because estimating macroeconomic effects can take considerable time and because the estimated budgetary impact (excluding macroeconomic effects) is one of the inputs into estimating a proposal's macroeconomic effects.

8. CBO (2015a, p. 133) provides a rule of thumb for the budgetary impact of lower output growth.

II.B. Comprehensive Immigration Legislation

In 2013, the Senate passed a bill to substantially increase the number of people who could enter the country legally and to create a process through which many people who are currently present in the country on an illegal basis could gain legal status. CBO and JCT provided estimates for the legislation in two separate documents released simultaneously: a cost estimate that included some but not all of the expected macroeconomic effects of the bill (CBO 2013b), and a supplemental analysis of the bill's total macroeconomic effects and the incremental budgetary impact of the economic changes not included in the cost estimate (CBO 2013c). CBO (2013c, p. 2) explained the analysis this way:

> [Since the legislation] would significantly increase the size of the U.S. labor force, assuming that total employment was unchanged would imply that any employment of the additional immigrants would be offset one-for-one by lower employment elsewhere in the population. Because that outcome would be highly implausible, CBO and JCT relaxed the assumption of fixed GDP and employment and incorporated into the cost estimate their projections of the legislation's direct effects on the U.S. population, employment, and taxable compensation. Nevertheless, to remain as consistent as possible with the estimating rules CBO and JCT follow for almost all other legislation, the cost estimate . . . does not incorporate the budgetary impact of every economic consequence of the bill. The [supplemental] analysis . . . includes some additional budgetary effects stemming from changes in the productivity of labor and capital, the income earned by capital, the rate of return on capital (and therefore the interest rates on government debt), and the differences in wages for workers with different skills.

That is, the cost estimate excluded macroeconomic changes that could be excluded without making the estimate nonsensical, and the supplemental analysis included all of the macroeconomic changes that CBO was able to estimate.

In the cost estimate, CBO and JCT estimated that the bill would reduce cumulative budget deficits by about $200 billion during the first decade after enactment and about $700 billion during the following decade. In the supplemental report, CBO estimated that the bill would raise output by roughly 3 percent by the end of the first decade; additionally, it estimated that economic effects not included in the cost estimate would have no further net effect on the cumulative deficit in the first decade but would further reduce the cumulative deficit in the second decade by about $300 billion.

II.C. Affordable Care Act

When CBO and JCT estimated the budgetary effects of the ACA and its precursors in 2009 and 2010, they incorporated the impact of many changes

in the behavior of individuals, employers, health insurers, and health care providers. However, the agencies did not include the impact of certain other changes in behavior because of the long-standing convention for cost estimates that overall output would be unaffected.

By contrast, given the recent congressional push for dynamic scoring, CBO and JCT's estimate in mid-2015 of the effects of *repealing* the ACA included macroeconomic changes. The estimate incorporated, among other factors: short-term effects on aggregate demand of changes in federal spending and taxes; effects on labor supply of changes in tax rates for certain higher-income people and of changes in subsidies for health insurance for certain lower-income people; and effects on capital investment of changes in tax rates, labor supply, and federal borrowing. The largest macroeconomic impact of repealing the ACA was estimated to stem from repealing the subsidies for health insurance.[9] The agencies concluded:

> Repealing the ACA would increase federal budget deficits by $137 billion over the 2016–2025 period . . . , [which incorporates] the net effects of two components: Excluding the effects of macroeconomic feedback . . . , federal deficits would increase by $353 billion over the 2016–2025 period if the ACA was repealed. Repeal of the ACA would raise economic output, mainly by boosting the supply of labor; the resulting increase in GDP is projected to average about 0.7 percent over the 2021–2025 period. Alone, those effects would reduce federal deficits by $216 billion over the 2016–2025 period. (CBO 2015h, p. 1)

II.D. Congressman Camp's Tax Reform Proposal

In early 2014, Congressman Dave Camp, then the chair of the Ways and Means Committee, put forward a comprehensive proposal for broadening the bases of the individual and corporate income taxes, adjusting tax rates,

9. The estimate of the effects on labor supply drew heavily on CBO's published analysis during the preceding several years. CBO (2009c) examined various channels through which changes to the health insurance system could affect labor markets; however, specific proposals were still in formative stages at the time, so the report did not provide quantitative estimates. In late 2009 and early 2010, congressional interest in the evolving health care legislation focused on its effects on the federal budget, health insurance coverage, insurance premiums, and existing federal programs, so CBO and JCT's analysis focused on those issues. After the ACA was enacted, CBO needed to incorporate the law's economic effects into the baseline economic projections. As part of that process, CBO (2010b, pp. 48–49) reported that it expected the ACA to reduce aggregate labor supply by an amount that would reduce labor compensation by roughly one-half percent after it was fully phased in. A few years later, during a careful review of its labor-market projections, CBO (2014a) updated that estimate to roughly one percent, with the revision arising because the agency "incorporated into its analysis additional channels through which the ACA will affect labor supply, reviewed new research about those effects, and revised upward its estimates of the responsiveness of labor supply to changes in tax rates" (p. 118).

and making other changes to those tax systems. JCT published a set of reports analyzing that proposal, including estimates of the proposal's distributional consequences, revenue effects (excluding any impact on the overall economy), and overall economic impact. Based on the convention of fixed output, JCT (2014b) estimated that the proposal would be effectively revenue-neutral, raising federal revenues by $3 billion over the next decade. In its macroeconomic analysis, JCT (2014c, p. 21) wrote:

> Broadening of the individual and corporate income tax bases through elimination of many preferences in the form of deductions, exemptions, and tax credits allows for a reduction in average and effective marginal tax rates for most individual taxpayers, which provides both an incentive for increased labor effort, and an increase in demand for goods and services. These changes also reduce the after-tax return to investment under many modeling assumptions, providing an incentive for a reduction in the U.S. domestic capital stock. On net, these changes are expected to result in an increase in economic output relative to present law.

JCT estimated that the proposal would raise the level of output by between 0.1 percent and 1.6 percent, on average, during the 2014–23 period. That additional output was estimated to reduce cumulative deficits by between $50 billion and $700 billion during the 2014–23 period.

II.E. Other Illustrative Analyses

In the past several years, CBO and JCT have provided estimates of the budgetary feedback from the macroeconomic effects of other proposals as well.

Each year, the agencies publish detailed estimates of the president's budget proposals based on the conventional assumption that the overall economy would be unaffected, and CBO separately (and somewhat later) publishes an analysis of the economic effects of the proposals and the feedback to the federal budget. As an example, CBO (2012b, 2012c) estimated that, excluding macroeconomic effects, the cumulative deficits under the president's proposals would be $3.2 trillion during the 2013–17 period and another $3.2 trillion during the 2018–20 period—and that including macroeconomic effects, the cumulative deficits would be $3.0 trillion to $3.2 trillion during the first half-decade and $3.3 trillion to $3.6 trillion during the second half-decade.

In addition, CBO's annual analysis of the long-term budget outlook includes estimates of economic and budgetary outcomes under alternative policies, with the budgetary effects taking into account the economic effects and vice versa (CBO 2015e). Moreover, in some years, CBO has published

estimates of the economic effects and budgetary feedback of deficit paths specified by the chair of the House or Senate Budget Committee (for example, CBO 2015e). JCT has also released macroeconomic analyses of the Jobs and Growth Reconciliation Tax Act of 2003 (JCT 2003a), the American Recovery and Reinvestment Tax Act of 2009 as reported by the Ways and Means Committee (JCT 2009, pp. 234–39), and other proposals.

III. The Case for Including Macroeconomic Effects in Budget Estimates

A natural starting point for evaluating dynamic scoring is to consider the objectives of the budget estimating process. After providing that context, I explain why including macroeconomic effects in estimates for certain legislative proposals would both improve the accuracy of budget estimates for those proposals and provide important information about their economic effects. However, attempting to include macroeconomic effects in all budget estimates would not be appropriate because of limited staffing and time, and because it would raise another set of issues that would have to be addressed as well.

III.A. What Should Be the Objectives of the Budget Estimating Process?

In my judgment, *CBO and JCT should provide estimates for legislative proposals that measure the full budgetary effects as accurately as possible and illuminate notable nonbudgetary effects, subject to several significant constraints.*

To "measure the full budgetary effects as accurately as possible," estimates would ideally equal the expected present value of proposals' budgetary effects over a long time horizon. A long horizon is appropriate, because the effects of proposals could last into the indefinite future, and CBO and JCT should analyze proposals as specified and not presume the enactment of any future legislation to modify or undo them. In particular, it is untenable for agencies working for Congress to make specific predictions about the future decisions of current members or of members who will be elected over time.[10] Present value is appropriate because future events are

10. However, some approaches to estimating the macroeconomic effects of legislation do require limited predictions about future policies; this issue is addressed in section IV.

discounted in other contexts. The expected outcome is appropriate because it minimizes the mean squared error of estimates.[11]

To "illuminate notable nonbudgetary effects"—such as effects on the overall economy—estimates would ideally provide quantitative, or at least qualitative, information about such effects. That objective is appropriate because reliable and timely information about the nonbudgetary effects of proposals is important for policymakers to receive and is not readily available from sources other than CBO and JCT. In particular, advocates and opponents of proposals often generate overly optimistic or pessimistic estimates of their effects, while independent analysts often are not familiar with the details of proposals and do not possess the models needed to estimate their effects, so they have difficulty producing reliable estimates quickly.

However, those ideal approaches cannot be fully put into practice, due to four significant constraints, as follows:

First, estimates should be easily understandable by members of Congress, their staffs, and outside observers, and they should be resistant to misinterpretation. Many members, staff, and observers have little training in quantitative analysis or budgeting, and most have limited time for reviewing budget estimates. Also, advocates and opponents of proposals often try to cite estimates in ways that support their own positions, so it is important that estimates be clear and difficult to use in misleading ways.

Second, estimates should be based on methodologies that are applied consistently across related proposals and are credible to members of Congress, their staffs, and outside analysts. Using consistent methodologies is crucial to ensuring that proposals can be compared meaningfully. Using methodologies that are credible is crucial to maintaining Congress's confidence in the estimates, to ensuring that the estimates reflect the consensus of informed professional thinking, and to protecting CBO and JCT from political pressure. By contrast, using methodologies that seem arbitrary or can be easily manipulated by lawmakers' construction of proposals in particular ways undermines confidence in the agencies' estimates for those proposals and for other proposals as well.

Third, estimates should be produced quickly enough to serve the legislative process and structured in ways that fit the process. Thus, estimates should include the information sought by congressional leaders or committees as

11. For certain financial activities of the government where risk is apparent and can be readily assessed, I think the estimated budgetary effect should not equal the expected present value of the activities but instead should incorporate an adjustment for the cost of the risk. That issue lies beyond the scope of this paper.

they develop proposals and by members of Congress as they vote on proposals. And when the legislative process moves swiftly, estimates should be prepared and updated rapidly as proposals are modified.

Fourth, estimates need to be prepared using the resources available to CBO and JCT. Although the number of congressional requests for estimates has increased considerably in recent years, the funding provided to the agencies has left their staffing little changed, on balance.

Many aspects of the budget estimating process represent compromises between those constraints and the idealized estimates described above. As one important example, official budget estimates generally apply to the coming decade rather than a longer time period, because the estimating methodology needed for a longer period would require additional resources to develop, would usually be less credible, and would lead to estimates that were more prone to misinterpretation. However, certain proposals would have longer-term effects that are quite different from their effects in the coming decade, and in those cases CBO and JCT try to provide some information on longer-term effects. The precision of that information and the time period for which it is provided vary across proposals, depending on congressional interest, on the agencies' assessment of the resources required to generate the information, on the credibility of the methodology used, and on the risk of leaving results open to misinterpretation. Thus, CBO has analyzed certain proposals to change Social Security over 75 years (CBO 2010a), but it generally does not analyze proposals to change federal health care programs beyond 25 years because of the especially large uncertainty involved in predicting the evolution of the health care delivery and financing systems.

As another example, official budget estimates generally show nominal cash flows rather than inflation-adjusted or present-value cash flows because nominal flows are more straightforward than the alternatives and because the distortion relative to showing present values is fairly small over a decade. The principal exception is estimates for federal credit programs, for which nominal cash flows over a decade are often a gross misrepresentation of the full budgetary effects over a long horizon; for these estimates, accrued costs are therefore used instead.

As a final example, official budget estimates sometimes exclude factors that might affect the budgetary impact of proposals but whose sign or magnitude are especially uncertain. That exclusion may seem inconsistent with the objective of measuring budgetary impact as accurately as possible: The mean squared error of a budget estimate reflects the underlying uncertainty of all relevant factors, even if estimates of some of the factors are set to

zero; therefore, if CBO or JCT can generate informed estimates of those factors, including them in the overall budget estimate would probably increase the accuracy of the estimate. However, that point cannot be proven as a general rule. Although in-sample prediction errors from linear regression models are minimized by including all relevant factors, including additional factors does not necessarily minimize out-of-sample prediction errors from nonlinear models (which is what CBO and JCT often use).[12]

Moreover, including factors that are especially uncertain can diminish the credibility of the estimating process because when the likelihood function for a factor is particularly flat, the agencies' choice of a specific value often seems arbitrary. Therefore, for factors whose budgetary impacts are probably small and are especially uncertain, the probable improvement in accuracy from including them in budget estimates may be outweighed by the risks of inadvertently diminishing accuracy and weakening the agencies' credibility. That condition is particularly likely to be satisfied when the net budgetary impact of a set of excluded factors might be either positive or negative.[13]

III.B. Similarity between Macroeconomic and Nonmacroeconomic Effects

Changes in federal tax and spending policies can affect people's behavior in many ways, and those behavioral responses can affect the federal budget. Some of those responses affect the composition of output or distribution of income but not total output and income, while other responses affect total output and income as well as their composition and distribution. A natural presumption is that measuring the full budgetary effects of legislative proposals as accurately as possible requires including the impact of all of those behavioral responses.

For example, if marginal income tax rates were increased, a number of responses would ensue. The share of people's income devoted to activities whose costs can be deducted from income, such as mortgage interest

12. See CBO (2015f, pp. 21–22) for a related discussion.

13. For example, CBO (2014a, p. 123) discussed some ways in which the ACA might affect productivity and concluded: "Whether any of those changes would have a noticeable influence on overall economic productivity, however, is not clear. Moreover, those changes are difficult to quantify and they influence labor productivity in opposing directions. As a result, their effects are not incorporated into CBO's estimates of the effects of the ACA on the labor market."

payments and charitable contributions, would increase. The share of compensation received in nontaxable forms, such as employers' contributions to pensions and health insurance, would increase. The amount of labor supplied would decrease (if the substitution effect outweighed the income effect), and the amount of saving would decrease (again, if the substitution effect outweighed the income effect). The first two responses are typically included in conventional estimates, and the latter two are not. However, when the responses are described in this manner, there is no clear conceptual reason to treat the latter two responses differently from the first two.

Of course, one difference between those two sets of responses is that reductions in labor supply and saving would affect total output whereas shifts in the uses of income and types of compensation would not. Thus, the reductions in labor supply and saving can be labeled macroeconomic effects, while the other shifts are purely microeconomic. Still, because all of the responses stem from actions by people and firms, there is no clear rationale for including some in budget estimates and excluding others.

The reductions in labor supply and saving that are spurred directly by the increase in marginal tax rates could generate further economic changes, such as shifts in pretax wages and in the pretax return to capital, which would have further effects on labor supply and saving. One might argue that those additional effects should be excluded from budget estimates because of their indirectness. However, indirect effects can be quantitatively important. Consider an example from a legislative proposal that did not change tax rates: The immigration legislation approved by the Senate in 2013 would have significantly increased the supply of labor, which would have induced additional capital investment. Ignoring the increase in labor supply would have substantially understated the impact of the legislation on output; including that increase in labor supply without including the induced growth of the capital stock would still have understated the impact of the legislation on output and would also have overstated its impact on wages.

Changes in federal policies can affect total output and income in many other ways as well. Changes in tax rules can affect investment in human capital and the allocation of physical capital, changes in federal benefits can affect labor supply and saving, and changes in federal spending for infrastructure, education and training, and research and development can affect labor supply, saving, and productivity. As with the effects of changes in marginal tax rates, there is no clear conceptual basis for including in budget estimates the effects of such policy changes on specific parts of the economy but not the effects on aggregate economic variables.

III.C. Advantages of Including Macroeconomic Effects in Certain Budget Estimates

Based on CBO's and JCT's past analysis of the macroeconomic effects of legislative proposals and their budgetary feedback, I conclude that *using dynamic scoring in budget estimates for certain proposals would improve the accuracy of those estimates, provide important information about the economic effects of those proposals, and (under certain conditions) satisfy the significant practical constraints for budget estimates listed earlier.*

Some proposals' estimated macroeconomic effects would have significant budgetary consequences. For example, the estimated macroeconomic effects of the Senate's 2013 immigration bill, the ACA, and Congressman Camp's tax plan (based on the midpoint of the reported estimates) all have budgetary impacts equal to hundreds of billions of dollars over a decade. Moreover, if dynamic scoring had been applied to the economic stimulus legislation of 2009 (the American Recovery and Reinvestment Act), its estimated budgetary effect would have been reduced by hundreds of billions of dollars: CBO (2009b) estimated that the legislation would raise output by more than $800 billion over the following decade, and that additional income would have been estimated to reduce budget deficits by about $200 billion compared with an estimated budgetary cost of the bill of roughly $800 billion (CBO 2009a).[14]

To be sure, accounting for the estimated macroeconomic effects of those proposals would have improved the accuracy of the official budget estimates only if the estimates of the macroeconomic effects had been somewhat accurate.[15] Unfortunately, assessing the accuracy of CBO's and JCT's estimates is quite difficult. Many proposals that the agencies examined were not enacted, and the proposals that were enacted were just a few of many factors affecting the economy and the budget, so isolating their impact is hard even in retrospect (CBO 2013a, 2015b). In my judgment, however, both agencies' methodology for conducting macroeconomic analysis reflects the consensus of informed professional thinking, and that consensus provides a useful, albeit imperfect, basis for predicting the macroeconomic effects of legislative proposals. Moreover, there is no reason to believe that their

14. The central estimate in JCT (2009) was that the tax provisions in the bill (as approved by the Ways and Means Committee) would increase output by about one-half percent in the short run, leading to a reduction in the cost of those provisions of about one-seventh of the conventional estimate.

15. The further step of estimating the budgetary feedback from estimated macroeconomic effects is fairly straightforward and can be done reasonably accurately.

estimates of macroeconomic effects are generally less accurate than their estimates of other effects of complex proposals, although certainly the agencies should continue to strive to improve their analyses.

The effects of some legislative proposals on the overall economy are very important for policymakers to understand. For example, while the macroeconomic effects of immigration reform and tax reform are among the most touted reasons for pursuing those policy changes, different approaches to immigration reform—such as increasing the numbers of high-skilled and low-skilled immigrants by different amounts—could lead to very different macroeconomic effects.[16] Similarly, different approaches to tax reform, such as using revenues raised by broadening tax bases to reduce marginal tax rates or to make targeted inframarginal tax reductions, could lead to very different macroeconomic effects.

As another example, major changes to benefits for lower-income people could have notable effects on the economy by altering labor supply, and those effects could be an important criterion in evaluating such changes. To use Arthur Okun's famous metaphor, we should understand the leakiness of different buckets for transferring resources to lower-income people. And as a further example, policy changes that reduced federal deficits to different degrees and at different speeds would generally have different macroeconomic effects in the next few years and in the longer run.

Estimates of macroeconomic effects can be valuable even when those effects appear small to some observers. For example, CBO (2015e) found that this year's budget resolution—which calls for a reduction in cumulative deficits over the next decade of about $5 trillion excluding interest savings and macroeconomic effects—would raise the level of real output in 2025 by 1½ percent, which amounts to an increase in the average annual growth rate over the coming decade of 0.15 percentage point. If that effect is surprisingly small to some people, the value of the estimate is increased, not diminished.

In addition, objective and timely information about the macroeconomic effects of legislative proposals is not readily available from sources other than CBO and JCT. Advocates and opponents of particular policies usually find ways to have their perspectives well represented in the congressional and public debates. However, independent, reliable analysts generally

16. Changes in overall output do not necessarily correspond to changes in economic well-being and should not be interpreted as such. For example, CBO (2013c) distinguished carefully between the effects of the-Senate's 2013 immigration legislation on total output and on output per resident.

have more difficulty than the agencies do in completing their analyses on a timely basis (because doing careful analysis is hard and because those analysts often are not close to the policy development process) and having their analyses heard. Including macroeconomic effects in budget estimates would ensure that CBO and JCT devote resources to analyzing those effects in a timely way. Also, because of the attention paid to official budget estimates in the legislative process, including macroeconomic effects in those estimates would ensure that the effects received attention.

The usefulness of such attention depends in part on the clarity of CBO's and JCT's descriptions of estimated macroeconomic effects. Describing such effects can be challenging, but the agencies now have considerable practice doing so in their supplementary analyses and are quite able to do so in official budget estimates as well. For example, in the agencies' report on repealing the ACA, the logic and magnitude of the macroeconomic effects are laid out clearly and in a manner that nicely parallels the discussion of the repeal's nonmacroeconomic effects.

Lastly, under certain conditions CBO's and JCT's estimates of the macroeconomic effects of legislative proposals can satisfy the key constraints described earlier: being understandable and resistant to misinterpretation, based on a consistent and credible methodology, produced quickly enough to serve the legislative process, and prepared using the resources available to the agencies. Those issues are addressed in the remainder of the paper.

III.D. Limiting Macroeconomic Effects to Budget Estimates for Major Proposals

Despite the advantages of including macroeconomic effects in budget estimates for legislative proposals, I conclude that *such effects should be incorporated only in estimates for major proposals.* Specifically, apart from proposals for which dynamic scoring is requested by the chair or ranking member of the House or Senate Budget Committee, I think that dynamic scoring should be applied only to proposals whose estimated nonmacroeconomic effects on revenues, spending, or deficits, relative to the baseline, exceed a given threshold.

CONSIDERATIONS IN SUPPORT OF LIMITED DYNAMIC SCORING That recommendation is based on three considerations (although there are legitimate counterarguments that are discussed below). First, CBO and JCT have the resources to conduct careful macroeconomic analyses for only a limited number of legislative proposals each year. All of the estimates of macroeconomic effects described above involved significant conceptual and practical challenges and required a great deal of analysts' time to complete.

The agencies can devote that much time to only a very small share of the thousands of proposals they examine each year.

Second, applying rules of thumb to produce estimates of the macroeconomic effects of other proposals would generally violate the important constraint that estimates be based on methodologies that are credible and cannot be easily manipulated. The macroeconomic effects of proposals can be complex and can vary considerably with the specifics of the proposed policy changes, the state of the economy, and the time horizon being examined. For example, when CBO (2015c) examined three ways of reducing spending for the Supplemental Nutrition Assistance Program (commonly known as "food stamps"), the agency found that even the sign of some policies' net effect on labor supply was unclear without detailed analysis. Similarly, when JCT (2005) examined three approaches to reducing taxes by $500 billion, it found that their effects on the economy differed greatly.

In addition, the agencies' estimates of short-term macroeconomic effects depend importantly on the posture of monetary policy. Rules of thumb would not capture those kinds of crucial nuances. Moreover, developers of proposals might exploit rules of thumb by structuring their proposals or labeling aspects of their proposals in ways that would generate more favorable estimated macroeconomic effects and thus lower estimated budgetary costs. CBO and JCT currently minimize such gaming by basing their nonmacroeconomic estimates on a careful understanding of the substance of proposals rather than the application of arbitrary rules. Even if using rules of thumb to estimate macroeconomic effects could improve the accuracy of budget estimates on average, doing so would endanger the credibility of the estimating process.

Third, the proposals for which CBO's and JCT's estimates of macroeconomic effects would generally be most valuable are the ones with the largest estimated budgetary impacts apart from such effects—because those proposals are likely to produce significant macroeconomic effects. To be sure, *some* proposals that would not cause large changes in revenues or spending would also produce significant macroeconomic effects, but identifying them with a mechanical algorithm would be difficult. For example, Congressman Camp's comprehensive tax plan had a very small estimated effect on revenues apart from macroeconomic effects—because the large estimated effects of some individual provisions of the plan were largely offsetting—but significant estimated macroeconomic effects. A threshold for dynamic scoring based on the *gross* budgetary effects of a proposal's individual provisions would have identified his plan. However, a criterion based on so-called gross effects would not be very robust, because the

method of constructing an estimate and the level of detail shown in a table describing the estimate can alter the magnitude of the reported increases and decreases.[17]

Thus, the best way to choose which proposals with small estimated budgetary impacts relative to the baseline should be scored dynamically is to allow for requests from key congressional leaders. In principle, at least, the budget committees are responsible for the budget process, and the official budget estimates are designed to support that process, so it makes sense to allow for requests from those committees. One might also allow for requests from the House Ways and Means Committee, the Senate Finance Committee, and the House and Senate Appropriations Committees, which are the other principal consumers of CBO's and JCT's estimates. However, if all of those committees could request dynamic estimates, the number of requests might become unmanageably large, so I slightly prefer to vest the authority only in the budget committees. What is more important is to ensure that such requests are not skewed in the direction of any one party's political interests. Therefore, the ranking members as well as the chairs of the committees should be allowed to make requests. By contrast, the new House rule and this year's budget resolution grant that power only to the budget committee chairs.

In my view, a sensible threshold for automatically including macroeconomic effects in budget estimates would be estimated changes in revenues, spending, or deficits (excluding any macroeconomic impact) relative to the baseline exceeding one-quarter of one percent of projected output over the 10-year budget window. That threshold equals about $575 billion currently (based on CBO [2015a]) and probably would lead to dynamic scoring for only a few proposals each year, which would be a manageable increase in CBO's and JCT's workloads. Neither the Senate's 2013 immigration proposal nor Congressman Camp's tax plan would have met that threshold, though clearly at least one leader of a budget committee would have requested dynamic scoring for each.

The new House rule and this year's budget resolution use a quarter-point threshold but apply it to any single year in the budget window rather than

17. For example, the ACA included significant changes to the drug benefit in Medicare, some of which increased federal spending and others of which reduced it. One might view the estimated net effect on spending of those changes to be the combination of a gross estimated increase and a gross estimated decrease. However, the changes interacted with each other in significant ways, so CBO estimated their effects as a package and never identified elements of the estimate separately.

the 10-year period as a whole. That approach generates thresholds of about $45 billion in 2015 and about $70 billion in 2025 based on projected output, and naturally it will cause more bills to be scored dynamically than my preferred approach would. In particular, short-term extensions of expiring tax or spending provisions are more likely to receive dynamic scores, as in JCT (2015b). However, the estimated macroeconomic effects of such extensions are not always illuminating: Many people expect such provisions to be extended, even retroactively, but the baseline reflects the expiration of the provisions and therefore the assumption that people will gradually recognize that extensions are not occurring; as a result, the macroeconomic effects of extending the provisions are primarily the effects of people not being surprised. Such effects can be complicated to estimate and difficult to explain. In addition, extensions of that sort are often negotiated and voted on under tight timetables, which further complicates doing macroeconomic analysis.

TWO COUNTERARGUMENTS There are two noteworthy counterarguments to limiting dynamic scoring to major proposals. One is that even if the macroeconomic effects of a proposal with limited budgetary impact are small relative to the overall economy, their feedback effect on the federal budget could still be large relative to the nonmacroeconomic budgetary impact of the proposal. In those circumstances, careful dynamic scoring would significantly improve the accuracy of the budget estimate. However, CBO and JCT cannot do careful analyses of the macroeconomic effects of all proposals, and, as described above, using rules of thumb in place of careful analyses could reduce the accuracy of those estimates and diminish the credibility of CBO's and JCT's estimates more generally. In my judgment, those costs outweigh the benefits.

The other counterargument is that focusing dynamic scoring on major proposals would create an incentive for certain proposals to be bundled together or separated into pieces in order to lower their estimated budgetary cost. However, sufficiently few proposals have budgetary impacts close to a quarter of a percent of output that this distortion would probably not be significant in practice. Moreover, allowing key congressional leaders to request dynamic scoring for less-significant proposals should ameliorate this problem.

Given the inability of CBO and JCT to apply dynamic scoring to all proposals, one might wonder whether avoiding dynamic scoring altogether is the best feasible approach because it would make the estimating methodology more consistent across proposals. However, policymakers do not usually compare major proposals to less-significant proposals; major proposals

are usually compared to other major proposals addressing similar issues, or else they are compared to the status quo, and both types of comparisons would be improved by dynamic scoring. In addition, as just noted, key congressional leaders could request dynamic scoring for less-significant proposals whose comparison to major proposals they considered especially valuable.

III.E. Other Important Issues

Five other issues concerning the inclusion of macroeconomic effects in budget estimates deserve comment.

TIME HORIZONS First, estimates of macroeconomic effects of proposals should include both short-term effects stemming from shifts in aggregate demand and longer-term effects stemming from shifts in potential output. Longer-term economic effects may be better guides to proposals' effects beyond the 10-year budget window, and since that window is a compromise between the ideal of an even longer horizon and the practical advantages of a shorter horizon, a focus on longer-term effects may seem preferable. However, the severe recession and slow recovery of the past several years are a stark reminder that shortfalls in the demand for goods and services can have large and persistent effects on the economy and the federal budget, so the effects of policy changes on aggregate demand can be quite important. Increasing aggregate demand was the principal objective of some proposals considered by Congress in the past several years, such as the economic stimulus legislation of 2009.

In addition, policy changes are sometimes reversed or modified in subsequent years, so the short-term effects of changes are the effects most likely to occur, and policymakers may therefore give estimates of those effects greater weight in their decisions. Further, different policy changes that Congress sometimes compares—such as different time paths for reducing budget deficits by a given amount—would have different effects on aggregate demand, and illuminating those differences would be an important benefit of dynamic scoring.

RESOURCE CONSTRAINTS Second, when CBO and JCT conclude that they do not have the tools or time needed to do a careful analysis of a proposal's macroeconomic effects, they should state as much and not include such effects in the official budget estimate. Estimating macroeconomic effects carefully often requires a great deal of analysis, and legislation is sometimes developed and amended quickly. Doing dynamic scoring without sufficient tools and time would endanger the credibility of the estimating process, as discussed above. That situation is especially likely to arise for

changes in regulatory policy that have small effects on the federal budget apart from any macroeconomic effects. For example, substantial changes in federal regulation of the financial system or the environment could have significant macroeconomic effects that would be important for policymakers to understand and that could feed back to the federal budget in notable ways. However, CBO does not have much expertise in estimating the macroeconomic effects of such regulatory changes, and acquiring enough expertise to do so quickly during the legislative process would require a significant increase in CBO's resources and would distract the agency from its core responsibility of informing budget policy.

SHARED RESPONSIBILITY Third, CBO and JCT should share responsibility for estimating the macroeconomic effects of proposals along the lines of their sharing responsibility for estimating the budgetary impact of proposals excluding macroeconomic effects. Specifically, JCT should produce estimates of the macroeconomic effects of major proposals to change the federal tax code, CBO should do the same for major proposals to change federal spending, and the two agencies should collaborate in estimating the macroeconomic effects of major proposals that would change both tax and spending policies (as they did in their estimate for repealing the ACA, released in June 2015 [CBO 2015h]). One challenge is ensuring that proposals that are similar in their substance receive similar estimates of their macroeconomic effects regardless of whether they are structured as changes in tax policy or spending policy. For example, similar subsidies for similar activities should be estimated to have similar macroeconomic effects, whether those subsidies take the form of tax credits or explicit federal spending. But that same challenge arises currently in estimating the budgetary impact of policies excluding their macroeconomic effects, and the challenge is met by ongoing interaction and coordination between CBO and JCT.

CONGRESSIONAL PROCEDURES Fourth, some congressional procedures may need to be modified to accommodate dynamic scoring. For example, dynamic scoring may generate jurisdictional conflicts between some congressional committees. Because changes in spending that affected total output and income would affect revenues, the House and the Senate would need to develop procedures for assigning budget targets to committees that allowed for such interactions. That problem would be somewhat ameliorated by the limited number of bills that would be scored dynamically. Other congressional procedures would adapt naturally to dynamic scoring. For example, official budget estimates are used to assess the applicability of certain parliamentary "points of order"; if a proposal were scored dynamically, the

estimated budgetary feedback from the macroeconomic effects of the proposal would be incorporated in that assessment.

SUPPLEMENTARY REPORTS Fifth, there are some advantages to taking an alternative approach, in which CBO's and JCT's estimates of macroeconomic effects and their budgetary feedback would be provided in supplementary reports rather than being included in official budget estimates, although I think those advantages are outweighed by the approach's disadvantages. The advantages of that alternative approach include the following: It would avoid delaying the publication of budget estimates until macroeconomic analyses could be completed, which might be a considerable period of time in some cases. It would maintain a consistent basis of fixed output for all official budget estimates, which might make it easier for CBO and JCT to exclude macroeconomic effects when they did not have a solid analytic basis for assessing them. And it would avoid the jurisdictional problems that can arise between committees just discussed.

One might view as a further advantage of the alternative approach that estimated macroeconomic effects and their budgetary feedback could be reported as ranges of possible outcomes rather than as the point estimates that the congressional budget process requires. This would also demonstrate the uncertainty of such macroeconomic analyses. Indeed, CBO and JCT should quantify that uncertainty by reporting ranges of estimates whenever feasible, as I discuss in greater detail later. However, CBO (2014f, p. 12) explains that "providing ranges sometimes muddies, rather than enhances, general understanding of our analysis because people tend to cite the part of a range they prefer," so CBO already tries to clarify the agency's findings by reporting point estimates as well as ranges in the agency's macroeconomic analyses. Using those point estimates in official budget estimates would not distort the analyses or their presentation.

One might view as a different advantage of the alternative approach that it would enable the House and Senate to disagree about whether to include macroeconomic effects in official estimates and would enable members of Congress, their staffs, and outside observers to evaluate the estimated macroeconomic effects separately from the other estimated effects. However, this would already be the case without the alternative approach. CBO (2015d, p. 23) has explained that "cost estimates [with dynamic scoring] will include all of the information that typically would be included if macroeconomic effects were not incorporated in the analysis, as well as additional information related to the macroeconomic effects." Accordingly, the recent estimates for repealing the ACA and extending certain expiring tax provisions reported three projections: the estimated budgetary impacts

excluding macroeconomic effects, the estimated budgetary impacts of macroeconomic effects alone, and the estimated total budgetary impacts including macroeconomic effects. Including estimated macroeconomic effects in official budget estimates in this manner does not preclude Congress and others from evaluating and using that information as they see fit.

Yet another seeming advantage of the alternative approach might be to facilitate a procedural transition between the historical exclusion of macroeconomic effects and their possible future inclusion. Such a transition would give CBO and JCT an opportunity to experiment with different methods of analysis and presentation, and it would give Congress and others an opportunity to learn about the agencies' macroeconomic analysis and develop procedures for using that information. In fact, however, that transition has effectively been under way for some time. As discussed above, both CBO and JCT have published many analyses of the macroeconomic effects of legislative proposals as well as reports on their methodology for such analyses. Moreover, the recently published estimates for repealing the ACA and extending expiring tax provisions did not reveal any problems that would suggest dynamic scoring is "not ready for prime time."

In addition, the alternative approach would have some significant disadvantages. Leaving macroeconomic effects out of budget estimates would reduce the attention those effects receive, even though the effects can be important for policymakers to understand. Also, separating the two types of budgetary effects of legislative proposals would make CBO's and JCT's estimates more difficult to understand for members of Congress, their staffs, and outside observers.

In my judgment, all those disadvantages of the alternative approach outweigh its limited advantages.

IV. Concerns about Including Macroeconomic Effects in Budget Estimates

A number of observers have expressed the view that including the macroeconomic effects of legislative proposals in official budget estimates would worsen rather than improve the information those estimates provide. In this section I examine six important concerns noted by those observers that were not addressed, or were addressed only briefly, in the preceding section. In my view, the first two of these concerns apply with roughly equal force to macroeconomic and nonmacroeconomic effects of proposals and thus are not compelling reasons for treating the former differently,

the next three concerns represent true challenges in estimating macroeconomic effects but can be addressed adequately by CBO and JCT, and the final concern is premised on an incorrect view of the agencies' role in the budget process.

IV.A. Potential Political Pressure on CBO and JCT

One concern is that including macroeconomic effects of proposals in official budget estimates would increase political pressure on CBO and JCT to adopt modeling approaches or elasticities of behavioral responses that would overstate the positive effects of certain types of policies. I welcome vigilance to guard against the risk that the agencies will be subjected to political pressure to modify their analysis in any regard. However, political pressure has not altered CBO's or JCT's estimates for legislation in past decades, and I doubt that dynamic scoring would substantially increase such pressure.

In my six years as CBO director, members of Congress, their staffs, and other observers were not hesitant to speak up when they disagreed with an estimate from CBO. That is not surprising, nor is it objectionable, because CBO's and JCT's estimates are not above reproach. The appropriate response by the agencies to such criticism is to collect any information that can be provided by those who disagree with the estimate, reconsider whether the agencies' analysis was correct, and revise the estimate if, and only if, the reconsideration shows that a different figure would be more accurate. That process occurred a number of times while I was at CBO. However, at no point in those six years did anyone in a position of authority in Congress attempt to dictate a change in a CBO estimate or in a CBO modeling assumption. Congressional leaders appear to understand the long-term value of maintaining CBO's and JCT's analytic independence.

Similar disagreements with CBO's and JCT's macroeconomic analyses have occurred in the past and would occur in the future if dynamic scoring were adopted. However, the risk of political pressure does not seem greater for estimates of macroeconomic effects than for estimates of nonmacroeconomic effects. On the one hand, there may be more observers who have strongly held views about key parameters underlying macroeconomic estimates—such as the elasticity of labor supply—than have strongly held views about the less well-known parameters underlying nonmacroeconomic estimates. On the other hand, because macroeconomic estimates depend heavily on a small number of parameters and other modeling choices, CBO and JCT have publicly documented those choices more thoroughly than they have publicly documented the analytic underpinnings of some

nonmacroeconomic estimates. It is important, then, that CBO and JCT report publicly on any changes in their macroeconomic modeling (as CBO did during the past few years when it updated a number of aspects of that modeling).

IV.B. Uncertainty of Macroeconomic Effects

Another concern is that the macroeconomic effects of proposals are so uncertain that including them in official budget estimates would degrade the quality of the estimates. However, as noted above, the accuracy of the budget estimates would probably be improved by including all of the factors for which CBO or JCT can generate informed estimates, and macroeconomic effects meet that criterion when the agencies have the time and tools to do a careful analysis. Moreover, many nonmacroeconomic effects of proposals are very uncertain as well, and in many cases there is less evidence to use in quantifying those effects than in quantifying the macroeconomic effects of proposals.

As one important example, the extent of uncertainty regarding the macroeconomic effects of changes in tax rates can be gleaned from CBO's recent review of evidence of the elasticity of labor supply. In a table summarizing estimates of the substitution elasticity for men and single women, CBO (2012g) showed values ranging from 0.04 to 0.84; in a corresponding table for married women, CBO showed values ranging from 0.03 to 0.70. Similar uncertainty exists about other behavioral responses and other aspects of CBO's and JCT's models of the macroeconomic effects of fiscal policies. However, uncertainty about those parameters is so apparent in part because they have been the subject of substantial research, which at least provides evidence for CBO and JCT to draw on.

For many of the behavioral responses underlying the nonmacroeconomic effects of fiscal policies, there is much less evidence for CBO and JCT to use, but that does not imply that uncertainty about the responses is smaller. For example, the agencies' estimates of the nonmacroeconomic effects of the Senate's 2013 immigration legislation, the ACA, and Congressman Camp's tax reform plan all relied heavily on assessments of behavioral responses for which there is little evidence and experts are highly uncertain. In addition, as noted above, the agencies' public documentation of their methodologies for estimating nonmacroeconomic effects is less expansive in some cases than their public documentation of their methodology for estimating macroeconomic effects. Thus, there is no good reason to view the agencies' estimates of macroeconomic effects as less credible than their estimates of nonmacroeconomic effects.

Note also that *excluding* macroeconomic effects from budget estimates for proposals that might have significant macroeconomic effects—and doing so only because of a historical convention that many consider arbitrary—itself diminishes the credibility of budget estimates. In sum, I think that including macroeconomic effects (under the conditions described in this paper) would enhance the credibility of the official budget estimates.

Although the budget process focuses on point estimates, quantifying the uncertainty of estimated macroeconomic effects may be useful to members of Congress, their staffs, and outside analysts, and I think that CBO and JCT should do that whenever feasible.[18] For example, the range of estimates that JCT published for Congressman Camp's tax plan provided a measure of uncertainty, and, as noted above, CBO typically presents its estimates of macroeconomic effects with a range as well as a central estimate.

IV.C. Potential Bias toward Tax Cuts Relative to Spending Increases

A further concern is that applying dynamic scoring to proposals affecting federal taxes but not proposals affecting federal spending would distort policymakers' decisions in favor of tax cuts relative to spending increases. For example, if lower tax rates raise output by increasing labor supply, and greater infrastructure spending raises output by increasing the capital stock, then including the former effect in official budget estimates but excluding the latter effect would inappropriately encourage tax rate cuts relative to infrastructure spending increases. That concern is, in some ways, the opposite of a concern about the conventional approach to cost estimates, namely that excluding effects on labor supply tends to overstate the budgetary cost of tax cuts and understate the budgetary cost of benefit increases, thereby encouraging policymakers to increase taxes and benefits. In any event, the concern can be addressed by applying dynamic scoring to proposals that change spending as well as those that change revenues.

Indeed, as described earlier, CBO's analyses of macroeconomic effects include the effects of federal spending on the demand for goods and services, the effects of federal benefits on labor supply, and the effects on the economy of federal investments in infrastructure, education and training,

18. Regarding uncertainty in CBO's estimates and the appropriate response by policymakers, see Manski (2011), CBO (2014f), and CBO (2015g, pp. 108–9).

and research and development.[19] For example, CBO's annual economic analysis of the President's budget incorporates the effects of proposed changes in both spending and taxes. In addition, the estimated effects of the ACA on labor supply stem partly from changes in the tax code and partly from changes in spending for Medicaid (as well as some other aspects of the law), and those effects are treated in a completely parallel manner in the estimates.

Nonetheless, there are two obstacles to the goal of applying dynamic scoring equally to federal spending changes and tax changes. One obstacle is that the congressional budget process treats certain types of spending differently from other types of spending and revenues. Roughly a third of noninterest federal spending arises from annual appropriations by Congress (sometimes called "discretionary spending"), with the remaining roughly two-thirds reflecting payments for ongoing benefit programs (sometimes called "mandatory spending"). Appropriations are currently split about equally between defense and nondefense purposes, and about half of nondefense appropriations go to investments in infrastructure, education and training, and research and development. The Congressional Budget Act of 1974—which established CBO, the House and Senate Budget Committees, and many of the ground rules that govern the budget process—specified that CBO should not produce estimates for appropriations bills that are comparable to those it produces for other legislation, but instead should tally the amounts specified in those bills and provide those tallies to the appropriations committees. Potential implications of appropriations for future tax revenues or benefit payments are not considered in that tallying process or in the subsequent legislative process. Perhaps because of that different procedural treatment, appropriations bills are excluded from the requirements for dynamic scoring in the new House rule and in the congressional budget resolution.

Moreover, if dynamic scoring were applied only to proposals with a significant budgetary impact (excluding macroeconomic effects) relative to the

19. As CBO has noted, its analysis of those effects would benefit from further methodological advances. CBO (2014d, p. 9) explained that the agency is "developing the capability to apply substitution elasticities as well [as income elasticities], but it does not currently have that capability for all transfer payments (although the agency has incorporated substitution elasticities in some specific analyses)." Similarly, the agency should enhance its capability to estimate the effects of federal investments in disaggregated categories and to estimate the long-term effects of federal benefits—although its analysis in those areas is limited mostly by a paucity of available research.

baseline, as recommended above, few appropriations bills would meet that criterion. The baseline for appropriations equals the previous year's appropriations adjusted for inflation, or the statutory cap if one exists. Actual appropriations in a single year rarely differ from the baseline by an amount that exceeds a quarter of a percent of output over the coming decade, with the most recent exception being the one-time burst of appropriations under the 2009 stimulus bill.

However, if the chairs and ranking members of the budget committees were allowed to request dynamic scoring for some bills with small estimated budgetary effects, also as recommended above, then they should be allowed to make those requests for appropriations bills. It would be useful, in my view, for CBO to provide estimates of the macroeconomic effects and resulting budgetary feedback of consequential changes in appropriations.[20] Unfortunately, the new House rule and congressional budget resolution explicitly exclude appropriations bills. I also think that CBO should publish a report with estimates of the macroeconomic effects of alternative multiyear paths for federal investment and the budgetary feedback from those macroeconomic effects.

A second obstacle to applying dynamic scoring equally to spending changes and tax changes is that the macroeconomic effects of certain sorts of federal spending are not fully felt within the 10-year budget window. Of course, that same issue arises for certain sorts of federal tax changes. For example, reductions in the marginal tax rate on capital income encourage additional private investment, and the resulting increase in the capital stock (and thus output) occurs gradually. However, the problem may be especially acute for spending changes that involve investments. For example, most of the increment to output from a new bridge or improved highway (allowing for construction time and subsequent depreciation) occurs beyond the budget window. And most of the increment to output and income that may arise from improved health care, preschool education, or housing for low-income children occurs after those children have entered the labor force. Therefore, the extent to which the budget window distorts the estimated macroeconomic effects of a change in spending (or tax) policy varies considerably depending on the characteristics of the policy change and the private or public investment that is increased or decreased.

This problem can be at least partly addressed by having CBO and JCT provide information about the effects of proposals beyond the 10-year budget

20. For a recent estimate of the macroeconomic effects of a change in appropriations—but not the budgetary feedback from those macroeconomic effects—see CBO (2015l).

window. As noted above, when Congress is especially interested in a proposal's long-term effects or when CBO or JCT expect that a proposal's long-term effects would be very different from its effects within the budget window, the agencies can provide that information. For example, CBO and JCT provided estimates for the second decade after enactment for the Senate's 2013 immigration legislation and for repeal of the ACA, and in some cases CBO has estimated the effects of policy changes over even longer horizons (CBO 2010a, 2012e). In addition, the new House rule and congressional budget resolution explicitly require that qualitative information be provided about budgetary effects (including macroeconomic effects) beyond the 10-year budget window. Unfortunately, providing such information beyond the coming decade is challenging because, as noted earlier, the estimating methodology needed for a longer period would require additional resources to develop, would usually be less credible, and would lead to estimates that were more prone to misinterpretation in certain ways. In particular, the evidence base that could be used to construct estimates of the long-term effects of benefit changes for low-income children remains limited.

Given the limitations on applying dynamic scoring to changes in federal spending, one might wonder whether avoiding dynamic scoring for changes in federal taxes is the best feasible approach to official budget estimates because it increases the comparability of certain policies in a second-best way. Indeed, one might wonder whether dynamic scoring for changes in taxes always provides useful information if some of the macroeconomic effects of those changes occur beyond the 10-year budget window. However, the macroeconomic effects of policy changes within the budget window can be important, even if later effects are somewhat different. As noted above, the severe recession and slow recovery of the past several years, and the fact that policy changes are sometimes reversed or modified over time, may make policymakers especially interested in short-term effects.[21]

In addition, the limitations discussed here do not create significant distortions in comparisons among many types of tax changes, comparisons between tax changes and the continuation of current tax law, comparisons between tax changes and spending changes that do not have notable effects on investment, or comparisons between spending changes apart from investment and the continuation of current law. Even for comparisons

21. The particular problems that might arise in providing estimated macroeconomic effects for unsustainable policy changes are discussed in a later section.

between tax changes and spending changes that affect investment, estimates for the decade-long budget window or estimates that extend into a second decade may still be more comparable if they include macroeconomic effects. At the least, including such effects in budget estimates, even when imperfectly measured, would bring additional attention to those effects.

IV.D. Difficulty in Capturing Expectations of Future Fiscal Policy

Yet another concern is that accurately projecting the macroeconomic effects of proposals might require modeling expectations of future fiscal policy, which would raise significant conceptual and practical difficulties. In particular, a reduction in current tax rates may have different effects on labor supply and saving if people's expectations of future tax rates increase, decrease, or remain the same—and modeling those expectations is challenging, especially if the proposal at hand represents an unsustainable change in policy. However, expectations of future fiscal policy are frequently left aside in other economic analyses, and CBO and JCT have developed methods for handling this issue in their macroeconomic analyses.

Consider proposals that reduce marginal tax rates on labor income. When CBO or JCT analyze such a proposal using their Solow-type growth models, they apply labor supply elasticities drawn from the large empirical literature to the change in tax rates between the baseline and the proposal. Many papers in that literature do not explicitly measure expected tax rates but simply examine changes in labor supply that have resulted from given changes in contemporaneous tax rates. Therefore, the estimated elasticities can be interpreted as the effect of a given change in current tax rates with expected future tax rates adjusting in whatever way people expected them to adjust, on average, in the past—which may appropriately reflect the fuzziness of people's expectations about future tax rates. The Solow model does not explicitly include expectations, so this sort of estimated elasticity fits logically, although it will generate a less accurate estimate in circumstances when anticipatory effects are important.

In contrast, people's behavior in life cycle growth models depends explicitly on their expectations. If forward-looking people expected that federal debt would rise relative to output without limit, they would not hold federal bonds, so the models can be used only to analyze sustainable changes in policies. Therefore, when CBO or JCT use their life cycle models to analyze a proposal that would increase deficits indefinitely, the agencies incorporate future policy changes not specified in the proposal to offset the deficit increases. That situation is awkward, because a key principle of budget

estimates is that the agencies take proposals as written and do not predict future legislation.

To minimize the influence of the assumed future policy changes on their estimates, CBO and JCT generally report results for multiple alternative changes—for example, one estimate under the assumption that future revenues are increased and another under the assumption that future spending is reduced. In fact, the results often do not differ very much under alternative assumptions (CBO 2015i). The agencies also defer the assumed changes as long as possible while still being able to solve the models (although congressional interest in receiving information about the effects of proposals beyond the 10-year budget window increases the difficulty of deferring the assumed changes long enough that they would not affect the reported estimates). Even so, if a proposal would increase deficits indefinitely, which would have harmful economic effects, the inclusion in a budget estimate of additional policy changes not specified in the proposal might make the proposal look better than it really is.

Given those issues, I think that CBO and JCT should give less weight to estimates based on their life cycle models than to estimates using their Solow-type models, except in circumstances where the anticipatory effects of proposals might be especially important.[22] Indeed, the agencies have used their life cycle models less often than their Solow-type models in their macroeconomic analyses, perhaps because of the issue described here or else because the life cycle models can be more cumbersome in other respects.

It bears emphasis that the dependence of behavior on expectations of future fiscal policy also arises with estimates of certain nonmacroeconomic effects of proposals, although it is rarely discussed in that context. For example, a reduction in current tax rates may have different effects on mortgage interest deductions or on employers' payments for health insurance, depending on expectations of future tax rates. However, the empirical literature regarding such behavior does not explicitly measure expected tax rates, so CBO's and JCT's estimates can be interpreted as the effects of a given change in current tax rates with expected future tax rates adjusting in whatever way people had expected them to adjust, on average, in the past. That approach is somewhat unsatisfying, especially if the policy change at hand is not sustainable and therefore future changes will be needed, but there is no feasible alternative.

22. See CBO (2014c, pp. 14–15) for an example of how the agency combines results from its two models.

IV.E. Potential Distortion to Estimates of Unsustainable Policy Changes Stemming From the 10-Year Budget Window

A further concern is that estimates of the macroeconomic effects of unsustainable policy changes would be distorted by the 10-year budget window. For example, suppose that a reduction in tax rates was estimated to raise output over the next decade but also to generate revenue losses that were unsustainable (because the increase in output was not large enough for the tax reduction to pay for itself). If those same rates were later raised to satisfy the government budget constraint, output would be estimated to be lower in the long run. In that scenario, dynamic scoring over the budget window would credit the proposal with raising output even though the ultimate effect on output would be negative.

That scenario might occur, but it is not likely. First, a reduction in tax rates without an offsetting broadening of the tax base or reduction in spending might well be estimated to *lower* output within the 10-year budget window, depending on the specifics of the tax reduction. For example, JCT (2003a) estimated that the tax cuts enacted in 2003 would increase output during the first five years after enactment but decrease output later in the decade, in part because the harmful effects of greater federal debt were estimated to outweigh the favorable effects of lower tax rates. Similarly, CBO (2010c) estimated that extending the tax cuts originally enacted in 2001 and 2003 would raise output in the following few years but lower output later in the decade.[23]

Second, the harmful effects of greater federal debt increase over time as debt compounds, while the favorable effects of lower tax rates generally do not, so any tax-rate reductions that were estimated to raise output throughout the first decade after enactment would be less likely to be estimated to do so in the second decade. Indeed, CBO (2010c) estimated that the negative effects on output of extending the tax cuts would be much larger after 30 years than after 10 years. Those long-term effects would be reported by CBO and JCT because, as noted above, the agencies try to provide information about long-term effects when they expect them to be very different from effects within the budget window. In addition, the

23. That result may seem surprising in light of CBO's (2005) estimate that a 10-percent reduction in federal tax rates on individual income would probably increase output in the second half of the decade after enactment. The difference between CBO's 2005 and 2010 estimates stems from several factors, including: the inclusion of inframarginal cuts in taxes under the 2001 and 2003 legislation; an increase in outstanding federal debt, which means that the rise in interest rates resulting from greater federal borrowing has a more significant effect on future deficits; and various improvements in CBO's modeling.

current rules regarding dynamic scoring require CBO and JCT to provide information about the long-term effects of proposals.

Despite those points, suppose that a reduction in tax rates was, in fact, estimated to raise output over the next few decades but also to generate revenue losses that were unsustainable. The unsustainable nature of the rate reduction would be shown by the estimated effect of the proposal on federal debt, so the need to make further policy changes to offset the budgetary losses would be quite apparent. In addition, the increase in deficits might be offset later *not* by reversing the tax-rate reduction but by making some other policy change—and because that other change would have an effect on output that was not simply the opposite of the effect of the tax-rate reduction, it is unclear whether the estimated effects of the rate reduction on output in the first few decades would truly be misleading. Moreover, the estimated macroeconomic effects of the rate reduction would not be the only aspect of the budget estimate that could be misleading; the estimated nonmacroeconomic effects on the budget could be misleading as well, as would any distributional analysis or other analysis based on the rate reduction. In any event, it is not tenable for CBO and JCT to ignore the policy changes included in a legislative proposal even if those changes are not, by themselves, sustainable.

IV.F. Potential Benefits of Estimates that Err on the Side of Overstating Budgetary Costs

One other concern about dynamic scoring arises from the view that policymakers tend to give insufficient weight to budgetary costs when developing and voting on legislative proposals, so CBO and JCT should provide budget estimates that tend to err in the direction of overstating those costs. As summarized (but not necessarily endorsed) by the Committee for a Responsible Federal Budget (2012, p. 1), the argument is then made that excluding the macroeconomic effects of proposals from official budget estimates would make "dynamic gains a 'bonus' to help further reduce the deficit."

However, dynamic scoring does not consistently reduce the estimated budgetary cost of proposals relative to nondynamic scoring. As noted earlier, CBO (2010c) estimated that extending the broad tax cuts originally enacted in 2001 and 2003 would *reduce* output by the latter part of the decade after enactment, so dynamic scoring would have shown a more negative impact on the budget than nondynamic scoring. Also, CBO (2014a) estimated that the expansion of federal subsidies for health insurance under the ACA was reducing labor supply and thereby federal revenues, so dynamic scoring of

the ACA would have shown a more negative impact on the budget than nondynamic scoring.

More fundamentally, I do not think it is appropriate for CBO and JCT to try to nudge policymakers toward smaller budget deficits by providing estimates that tend to overstate the budgetary costs of proposals. Instead, the agencies should provide estimates that are in the middle of the distribution of possible outcomes and leave policymakers to make decisions based on their own views of desirable outcomes and acceptable risks.

V. Conclusion

Including macroeconomic effects in budget estimates for major legislative proposals—except when CBO and JCT do not have the tools or time to do a careful macroeconomic analysis—would improve the accuracy of those budget estimates and would provide important information about the economic effects of those proposals. Therefore, I conclude that dynamic scoring should be used for major proposals when the agencies have the tools and time to do a careful macroeconomic analysis.

To complement that greater commitment to analyzing the macroeconomic effects of legislative proposals, CBO and JCT should also make a greater commitment to analyzing proposals' distributional effects. JCT currently provides estimates of the distributional consequences of certain changes in federal taxes, as it did when it analyzed Congressman Camp's comprehensive tax reform proposal (JCT 2014a). However, CBO does not provide corresponding estimates for changes in federal spending. CBO's recent reports on the distribution of federal taxes under current law (CBO 2014e) and the distribution of federal taxes and spending under current law (CBO 2013d) described a number of conceptual complications and data limitations that arise in estimating the distributional impact of existing taxes and spending, and those problems are more acute when estimating the distributional impact of *changes* in taxes and spending. For example, distributional analysis of proposals ideally would incorporate macroeconomic analysis to capture the partial shift in the burden of capital taxes from capital owners to workers stemming from changes in the amount of capital. However, CBO and JCT have not developed the models needed to conduct such "dynamic distributional analysis."

In my view, CBO and JCT should continue to enhance their capabilities in this area. Policy choices can have significant effects on the distribution of income as well as on total income, and outcomes that are quantified often receive greater attention in policy discussions than outcomes that are

not. Therefore, it would be very valuable for CBO and JCT to quantify the effects of legislative proposals on both total income and its distribution.

ACKNOWLEDGMENTS I am grateful to Thomas Barthold, Wendy Edelberg, Glenn Hubbard, Jeffrey Kling, Donald Marron, Pam Moomau, Ben Page, and the editors for helpful comments on earlier drafts of this paper and to many former colleagues at the Congressional Budget Office and on the staff of the Joint Committee on Taxation for discussions that helped to form my views on this topic. The views expressed here are my own and should not be attributed to any of those people or to any institution with which I was or am currently affiliated.

References

Auerbach, Alan J. 1996. "Dynamic Revenue Estimation." *Journal of Economic Perspectives* 10, no. 1: 141–57.

———. 2005. "Dynamic Scoring: An Introduction to the Issues." *American Economic Review* 95, no. 2: 421–25.

Burman, Leonard E. 2006. "Dynamic Analysis and Scoring." Statement before the House Committee on the Budget, September 13. Washington. http://www.taxpolicycenter.org/UploadedPDF/900999_dynamicanalysis.pdf

CBO (Congressional Budget Office). 1995. "Budget Estimates: Current Practices and Alternative Approaches." Washington. https://www.cbo.gov/publication/13599

———. 1999. "Estimating the Costs of One-Sided Bets: How CBO Analyzes Proposals with Asymmetric Uncertainties." Washington. https://www.cbo.gov/publication/11843

———. 2001. "CBO's Method for Estimating Potential Output: An Update." Washington. https://www.cbo.gov/publication/13250

———. 2002. "Federal Budget Estimating." Statement of Dan L. Crippen before the Subcommittee on Legislative and Budget Process, House Rules Committee, May 9. Washington. https://www.cbo.gov/publication/13648

———. 2005. "Analyzing the Economic and Budgetary Effects of a 10 Percent Cut in Income Tax Rates." Washington. https://www.cbo.gov/publication/17507

———. 2009a. "H.R. 1: American Recovery and Reinvestment Act of 2009." Cost Estimate. Washington. https://www.cbo.gov/publication/41759

———. 2009b. "Estimated Macroeconomic Impacts of H.R. 1 as Passed by the House and by the Senate." Letter to the Honorable Judd Gregg, Ranking Member, Committee on the Budget, United States Senate, February 11. Washington. https://www.cbo.gov/publication/20474

———. 2009c. "Effects of Changes to the Health Insurance System on Labor Markets." Washington. https://www.cbo.gov/publication/20910

———. 2010a. "Social Security Policy Options." Washington. https://www.cbo.gov/publication/21547

———. 2010b. "The Budget and Economic Outlook: An Update." Washington. https://www.cbo.gov/publication/21670

———. 2010c. "The Economic Outlook and Fiscal Policy Choices." Statement of Douglas W. Elmendorf before the Committee on the Budget, United States Senate, September 28. Washington. https://www.cbo.gov/publication/21836

———. 2011. "CBO's Use of Evidence in Analysis of Budget and Economic Policies." PowerPoint Presentation at the Annual Fall Research Conference of the Association for Public Policy Analysis & Management, November 3. https://www.cbo.gov/publication/42722

———. 2012a. "Fair-Value Accounting for Federal Credit Programs." Washington. https://www.cbo.gov/publication/43027

———. 2012b. "An Analysis of the President's 2013 Budget." Washington. https://www.cbo.gov/publication/43083

———. 2012c. "The Economic Impact of the President's 2013 Budget." Washington. https://www.cbo.gov/publication/42972

———. 2012d. "Assessing the Short-Term Effects on Output of Changes in Federal Fiscal Policies." Working Paper no. 2012-08. Washington. https://www.cbo.gov/publication/43278

———. 2012e. "Raising the Excise Tax on Cigarettes: Effects on Health and the Federal Budget." Washington. https://www.cbo.gov/publication/43319

———. 2012f. "How the Supply of Labor Responds to Changes in Fiscal Policy." Washington. https://www.cbo.gov/publication/43674

———. 2012g. "A Review of Recent Research on Labor Supply Elasticities." Working Paper no. 2012-12. Washington. https://www.cbo.gov/publication/43675

———. 2012h. "Review of Estimates of the Frisch Elasticity of Labor Supply." Working Paper no. 2012-13. Washington. https://www.cbo.gov/publication/43676

———. 2013a. "The Accuracy of CBO's Budget Projections." CBO blog post by Doug Elmendorf, March 25. Washington. https://www.cbo.gov/publication/44017

———. 2013b. "S. 744: Border Security, Economic Opportunity, and Immigration Modernization Act." Cost Estimate. Washington. https://www.cbo.gov/publication/44225

———. 2013c. "The Economic Impact of S. 744, the Border Security, Economic Opportunity, and Immigration Modernization Act." Washington. https://www.cbo.gov/publication/44346

———. 2013d. "The Distribution of Federal Spending and Taxes in 2006." Washington. https://www.cbo.gov/publication/44698

———. 2013e. "Fiscal Policy Effects in a Heterogeneous-Agent Overlapping-Generations Economy with an Aging Population." Working Paper no. 2013-07. Washington. https://www.cbo.gov/publication/44941

———. 2014a. "The Budget and Economic Outlook: 2014 to 2024." Washington. https://www.cbo.gov/publication/45010

———. 2014b. "The Long-Run Effects of Federal Budget Deficits on National Saving and Private Domestic Investment." Working Paper no. 2014-02. Washington. https://www.cbo.gov/publication/45140

———. 2014c. "The Economic Effects of the President's 2015 Budget." Washington. https://www.cbo.gov/publication/45540

———. 2014d. "How CBO Analyzes the Effects of Changes in Federal Fiscal Policies on the Economy." Washington. https://www.cbo.gov/publication/49494

———. 2014e. "The Distribution of Household Income and Federal Taxes, 2011." Washington. https://www.cbo.gov/publication/49440

———. 2014f. "Communicating Uncertainty in Budgetary and Economic Estimates." Presentation by Douglas W. Elmendorf at the Hutchins Center on Fiscal

and Monetary Policy at the Brookings Institution, December 15. Washington. https://www.cbo.gov/publication/49862

———. 2015a. "The Budget and Economic Outlook: 2015 to 2025." Washington. https://www.cbo.gov/publication/49892

———. 2015b. "CBO's Economic Forecasting Record: 2015 Update." Washington. https://www.cbo.gov/publication/49891

———. 2015c. "The Effects of Potential Cuts in SNAP Spending on Households with Different Amounts of Income." Washington. https://www.cbo.gov/publication/49978

———. 2015d. "Answers to Questions for the Record Following a Hearing on the Budget and Economic Outlook for 2015 to 2025 Conducted by the Senate Committee on the Budget." Washington. https://www.cbo.gov/publication/49974

———. 2015e. "Budgetary and Economic Outcomes under Paths for Federal Revenues and Noninterest Spending Specified in the Conference Report on the 2016 Budget Resolution." Washington. https://www.cbo.gov/publication/50115

———. 2015f. "Answers to Questions for the Record Following a Hearing by the Senate Committee on the Budget on the Work of the Congressional Budget Office." Washington. https://www.cbo.gov/publication/50279

———. 2015g. "The 2015 Long-Term Budget Outlook." Washington. https://www.cbo.gov/publication/50250

———. 2015h. "Budgetary and Economic Effects of Repealing the Affordable Care Act." Washington. https://www.cbo.gov/publication/50252

———. 2015i. "Dynamic Scoring at CBO." Presentation at the NABE Foundation's 12th Annual Economic Measurement Seminar, July 21. Washington. https://www.cbo.gov/publication/50409

———. 2015j. "About CBO: Processes." Web page. https://www.cbo.gov/about/processes

———. 2015k. "Tax Relief Extension Act of 2015." Cost Estimate. Washington. https://www.cbo.gov/publication/50704

———. 2015l. "The Macroeconomic Effects of Eliminating Automatic Reductions to Discretionary Spending Caps." Letter to the Honorable Bernard Sanders, Ranking Member, Committee on the Budget, United States Senate, August 11. Washington. https://www.cbo.gov/publication/50725

Committee for a Responsible Federal Budget. 2012. "Understanding Dynamic Scoring." Washington.

Furman, Jason. 2006. "A Short Guide to Dynamic Scoring." Washington: Center on Budget and Policy Priorities.

Gale, William G. 2002. "Notes on Taxes, Growth, and Dynamic Analysis of New Legislation." *Tax Notes, 30th Anniversary Edition*: 30–43.

Hassett, Kevin A. 2002. "Dynamic Scoring." Testimony before the House Committee on the Budget, May 2. Washington. https://www.aei.org/publication/dynamic-scoring/

Holtz-Eakin, Douglas, and Michael Mandel. 2015. "Dynamic Scoring and Infrastructure Spending." New York: McGraw Hill Financial Global Institute.

Ip, Greg. 2015. "Q&A: Dynamic Scoring Is Neither Salvation Nor Scourge." Real Time Economics blog, *Wall Street Journal,* March 30.

JCT (Joint Committee on Taxation). 2003a. "Macroeconomic Analysis of H.R. 2, the 'Jobs and Growth Reconciliation Tax Act of 2003.'" *Congressional Record* 149, no. 68: H3829–H3832.

———. 2003b. "Overview of Work of the Staff of the Joint Committee on Taxation to Model the Macroeconomic Effects of Proposed Tax Legislation to Comply with House Rule XIII.3.(h)(2)." Report no. JCX-105-03. Washington.

———. 2005. "Macroeconomic Analysis of Various Proposals to Provide $500 Billion in Tax Relief." Report no. JCX-4-05. Washington.

———. 2006. "Exploring Issues in the Development of Macroeconomic Models for Use in Tax Policy Analysis." Report no. JCX-19-06. Washington.

———. 2009. "Providing for a Portion of the Economic Recovery Package Relating to Revenue Measures, Unemployment, and Health." House of Representatives Report no. 111-8, Part 1. Washington.

———. 2011a. "Summary of Economic Models and Estimating Practices of the Staff of the Joint Committee on Taxation." Report no. JCX-46-11. Washington.

———. 2011b. "Testimony of the Staff of the Joint Committee on Taxation before the House Committee on Ways and Means regarding Economic Modeling." Report no. JCX-48-11. Washington.

———. 2014a. "Distributional Effects of the 'Tax Reform Act of 2014.'" Report no. JCX-21-14. Washington.

———. 2014b. "Estimated Revenue Effects of the 'Tax Reform Act of 2014.'" Report no. JCX-20-14. Washington.

———. 2014c. "Macroeconomic Analysis of the 'Tax Reform Act of 2014.'" Report no. JCX-22-14. Washington.

———. 2015a. "Macroeconomic Analysis at the Joint Committee on Taxation and the Mechanics of Its Implementation." Presentation at the Brookings Institution, January 26. Report no. JCX-3-15. Washington.

———. 2015b. "A Report to the Congressional Budget Office of the Macroeconomic Effects of the 'Tax Relief Extension Act of 2015,' as Ordered to Be Reported by the Senate Committee on Finance." Report no. JCX-107-15. Washington.

Manski, Charles F. 2011. "Policy Analysis with Incredible Certitude." *Economic Journal* 121, no. 554: F261–F289.

Orszag, Peter R. 2002. "Macroeconomic Implications of Federal Budget Proposals and the Scoring Process." Testimony before the Subcommittee on Legislative and Budget Process, House Rules Committee, May 2. Washington. http://www.brookings.edu/~/media/Research/Files/Testimony/2002/5/02federalbudget%20orszag/20020502.pdf

Van de Water, Paul N., and Chye-Ching Huang. 2014. "Budget and Tax Plans Should Not Rely on 'Dynamic Scoring.'" Washington: Center on Budget and Policy Priorities.

Comments and Discussion

COMMENT BY

GLENN HUBBARD "Dynamic scoring" is surely an in-the-budget-weeds topic for most people, even most economists. Nevertheless, it is an important topic for study and reflection, particularly given the major tax and expenditure proposals likely to surface in the context of the 2016 presidential campaign. Answering the question of whether dynamic scoring can and should be done is straightforward, but I can attest from both government and academic experience that the ratio of heat to light in such a discussion is often high. Enter Douglas Elmendorf's paper. Elmendorf is the right author, with background both as a scholar in related research and as a distinguished former director of the Congressional Budget Office (CBO). Wisely, his paper quickly shifts the analytical discussion from "yes or no" to "how and when."

Responsible budget analysis largely incorporates estimates of the revenue or expenditure consequences of budget proposals—"scoring" them. Consider, for example, a proposal to reduce marginal tax rates across the board on individual incomes. A conventional revenue estimate would calculate lost revenue from the existing tax base and analyze the behavioral effects of the policy (for example, working harder as a result of the lower marginal tax rate on work). Such static scoring is not as naïve as the term *static* suggests; analysis of behavioral effects can be complex and rich, using a variety of elasticity estimates and microsimulation models.

While potentially rich, such work is necessarily incomplete for major proposals, in that a static analysis holds GDP constant. That is, while estimates consider compositional effects, they abstract from macroeconomic impacts. For example, suppose the Ways and Means Committee were considering a major tax reform proposal to scrap the present federal income tax on corporations and individuals and replace it with a broad-based con-

sumption tax. Almost surely, much of the motivation for such a proposal would be the goal of raising GDP and household incomes, and higher output and incomes, all else equal, will raise revenue. So conventional static scoring cannot provide the answer on *Jeopardy!* to the question, "What is the revenue impact of the tax reform proposal?"

PROS AND CONS OF DYNAMIC SCORING A consideration of the macroeconomic impacts of certain proposals is obviously the right answer to the last question. By "certain proposals" I mean to include not only tax reform but also immigration reform (with its effects on labor supply), health care reform (with effects on labor supply), and large expenditure programs (such as major infrastructure expansion initiatives). Importantly, dynamic scoring sheds light on aggregate effects that motivate congressional and White House interests in the first place. And dynamic scoring leans against the concern that current estimation procedures are stacked against policies that would advance economic growth or higher overall incomes. For example, the revenue consequences of a major tax reform would be affected according to whether the reform raised aggregate economic activity or incomes and, hence, expected revenue. In that sense, ignoring aggregate feedback effects, which are the basic element of dynamic scoring, makes the budget cost of expansionary tax and spending policies appear too expensive.

The foregoing comments notwithstanding, dynamic scoring has remained controversial in some policy circles. One objection is that formal incorporation of dynamic scoring within budget estimates introduces a bias for tax cuts—or spending increases—eroding budget disciplines. A second fear is that uncertainty in forecasts makes dynamic scoring unreliable. Finally, a technical concern is often expressed that the inclusion of macroeconomic feedback effects complicates the estimation process so much that dynamic scoring, however meritorious in theory, is simply too difficult to implement in practice. As Elmendorf observes in the paper, such concerns are off the mark.

A CONSUMPTION TAX EXAMPLE Many economists have estimated large gains on output and incomes from a shift to a broad-based consumption tax (because of reduced capital taxation and inter-asset tax distribution), including studies by academic researchers, the Treasury Department's Office of Tax Analysis, and the Joint Committee on Taxation (JCT). A study by staff economists of the JCT (1997) drew on a range of models to estimate that a shift to a consumption tax would raise GDP in steady state by 5 percent. Even the more modest reform plan in 2014 of then-House Ways and Means Committee chairman Dave Camp was estimated

by John Diamond and George Zodrow (2014) to raise GDP by as much as 3.1 percent in the long run.

With reasonable estimates of the marginal revenue effect of the change, the revenue impact would be about 1 percent of GDP in steady state, a very large adjustment. As work by Greg Mankiw and Matthew Weinzierl (2006) shows, static revenue estimates are considerably off the mark. Using a Ramsey growth model, Mankiw and Weinzierl (2006) estimate that the dynamic revenue effect of a cut in the capital income tax rate is about half the static estimate. (The difference would be modestly attenuated with a finite-horizon case and accentuated in a model with imperfect competition and markups.)

Again, as Elmendorf notes, there are legitimate concerns that macroeconomic modeling is not an exact science—but the same can be said for microsimulation models underlying conventional analysis. In addition, some fears about dynamic scoring are really concerns about the distributional consequences of proposals such as tax reform—of course, information on both budget and distributional effects should be presented to policymakers. Finally, the question arises as to whether dynamic scoring should be incorporated in budget rules—but surely such information is the best answer for major proposals.

Returning to the consumption tax example: failure to provide dynamic scoring denies policymakers information on the economic gains from tax reform. That lack of information effectively denies policymakers the ability to understand and manage trade-offs between the distributional and economic efficiency consequences of tax reform.

Finally, the uncertainty argument against dynamic scoring of tax reform proposals is not compelling. The idea that the method's uncertainty suggests the desirability of adopting a static-scoring answer that is *known with certainty to be incorrect* is not logical. And while the Federal Reserve must formulate monetary policy in the face of macro uncertainty, its own reliance on economic models is uncontroversial.

NEXT STEPS I agree with Elmendorf's basic point that the task ahead is to figure out how and when to do dynamic scoring, ending the existential debate. I think of this next step as incorporating analysis, process, and politics.

Analysis A key first step is to define candidate policies for dynamic scoring. Such candidates include policies with a material impact on aggregate demand, productivity, and/or hours worked. The second step is to encourage the staff economists of the JCT, the CBO, and the Office of Tax Analysis to refine models, including open-economy features and realistic

heterogeneity on the household side, with exercises incorporating the contribution and judgment of outside experts. Third, toward this end, organizations like the National Bureau of Economic Research and think tanks can develop forums to investigate the professional consensus regarding aggregate economic effects of major tax and spending proposals.

Process Moving to dynamic scoring requires a shift in both resources and best practices. That is, integrating microeconomic and macroeconomic models would require substantial incremental financial resources for the staffing of official scorekeepers. And a dynamic score should provide information as to why an estimated macroeconomic effect differs from a consensus estimate by economists. Assessing research on the best process would be enhanced by the use of a panel of outside experts, by official scorekeepers, and by the existence of a nonofficial, "open source" alternative that could be employed by outside researchers or policymakers seeking advice.[1]

Politics The political concern that dynamic scoring will inappropriately soften attention to the budget deficit must be addressed head on. The key, as Elmendorf notes, is to focus dynamic scoring on major proposals and to use consensus estimates of the macroeconomic effects of policy changes. In this regard, analysis of tax policy changes is more straightforward than analysis of spending changes, since it is easier to estimate the aggregate effects of tax policy over short- and medium-term horizons. More research on the spending side will help frame potential dynamic scoring of spending programs for infrastructure, education, and training support. Finally, political questions about the applicability of dynamic scoring to budget rules must be addressed. The answer here should be simple: To the extent that dynamic scoring is conceptually correct (it is), and implemented rigorously (it can be), the dynamic score should be the official score under budget rules. Including it only in an impact statement presented to decisionmakers is too limiting.

Elmendorf's thoughtful and careful paper makes a strong case for dynamic scoring and identifies next steps for implementation. Read one

1. The Open Source Policy Center, recently inaugurated by the American Enterprise Institute, is a welcome development in this regard. Incorporating dynamic scoring involves exporting output from microsimulation models of policy changes to a dynamic macroeconomic model that models a substantial portion of aggregate economic activity. In this model-bridging exercise, the macroeconomic model output from one time period can be fed into the microsimulation model for the next period. Openness of both models exposes assumptions and can allow policy analysts to pinpoint sources of disagreement.

way, the paper is so sensible and straightforward it seems "apple pie." But I think it reflects how far this policy debate has come. A generation ago, while working as a tax official at the Treasury Department, I occasionally felt like a referee on dynamic scoring between proponents who thought it was the savior of economic policy and opponents who thought it represented the end of western civilization. That we are now having a reasonable discussion about *how* to implement dynamic scoring is surely a good sign.

REFERENCES FOR THE HUBBARD COMMENT

Diamond, John W., and George R. Zodrow. 2014. "Dynamic Macroeconomic Estimates of the Effects of Chairman Camp's 2014 Tax Reform Discussion Draft." Houston: Tax Advisers LLC. http://businessroundtable.org/sites/default/files/reports/Diamond-Zodrow%20Analysis%20for%20Business%20Roundtable_Final%20for%20Release.pdf

JCT (Joint Committee on Taxation). 1997. "Joint Committee on Taxation Tax Modeling Project and 1997 Tax Symposium Papers." Report no. JCS-21-97. Washington.

Mankiw, N. Gregory, and Matthew Weinzierl. 2006. "Dynamic Scoring: A Back-of-the-Envelope Guide." *Journal of Public Economics* 90, nos. 8–9: 1415–33.

COMMENT BY

DONALD B. MARRON Douglas Elmendorf lays out a compelling case that the Congressional Budget Office (CBO) and the staff of the Joint Committee on Taxation (JCT) should account for macroeconomic feedback when scoring proposed legislation. I agree. The two agencies have been developing and refining their macroeconomic modeling techniques for more than a decade. They have successfully applied those techniques in a host of analyses. The next step in this gradual evolution is to incorporate them in official scoring.

Concerns about dynamic scoring are understandable in light of highly politicized fiscal policy debates and the way claims about dynamic effects often align with ideological views. But CBO and JCT have a strong track record of navigating such shoals and delivering nonpartisan analysis to Congress. I expect the same will be true with dynamic scoring, which will not fully live up to the hopes of its proponents nor "live down" to the fears of its detractors. Instead, it will modestly improve the budget estimates that inform policymakers and the public.

Seven points in Elmendorf's paper deserve particular emphasis:

First, including macroeconomic effects *can improve budget estimates.* The budget process requires year-by-year estimates of the revenue and spending implications of proposed legislation. CBO's and JCT's missions are to make those estimates as accurate as possible, given the constraints under which they operate. Fiscal policies can change how people work, save, invest, and spend and thus can raise or lower macroeconomic activity. Including such effects in official scores will improve budget deliberations as long as there is a sufficient evidentiary base for estimating them and as long as doing so is consonant with other constraints the agencies face (such as timeliness, resource limits, and transparency requirements).

Second, including macroeconomic effects *can improve policy comparisons.* Policy debates often distill macroeconomic effects to the level of tweets and bumper stickers: "Tax cuts boost growth"; "Spending stimulates the economy." The extent to which such claims are true, however, depends on policy specifics. Marginal tax cuts likely do more for long-run economic growth than do inframarginal ones, for example, and spending in recessions likely boosts the economy more than at times of full employment (CBO 2014).

CBO and JCT have traditionally quantified those differences in supplementary analyses undertaken outside of official budget scoring. Including them in official scores would make them more politically salient. Tax cuts that weaken long-run growth, for example, will get worse budget scores than under conventional estimating, while cuts that encourage growth will get better scores. Differing scores will give lawmakers more reason to consider seriously the macroeconomic effects of competing policy proposals.

Third, dynamic scoring *should apply to both taxes and spending.* Most of the public debate about dynamic scoring has focused on tax policies, but spending programs have the same potential for macroeconomic effects. Investments in infrastructure and education can boost long-run economic potential; so can private investment induced by tax policy changes. The phase-out of benefits in social insurance programs can discourage labor supply; so can taxes on wages. Spending can soften recessions; so can tax cuts. Spending financed by deficits can crowd out private investment; so can tax cuts financed by deficits. Treating spending and taxes equally in dynamic scoring thus makes perfect sense. This is straightforward for taxes and mandatory spending, but it is more challenging for the discretionary spending that Congress handles through its annual appropriation process. Whether and how to apply dynamic scoring to discretionary spending therefore deserves further attention in budget process discussions.

Fourth, dynamic scoring *should be applied only to major pieces of legislation.* Dynamic scoring is logistically challenging. It takes time and talented staff. As a result, there are practical limits to how many dynamic scores CBO and JCT can produce. It makes sense to focus solely on the largest bills, while allowing legislators to require dynamic scoring in other cases where macroeconomic effects may be important. If CBO and JCT develop ways to make macroeconomic analysis easier in the future, lawmakers can expand the scope of required dynamic scoring.

Fifth, *the authority to require dynamic scoring should be shared by the two parties.* Under current congressional procedures, only the chairs of the House and Senate budget committees can request dynamic scoring of particular bills. As Elmendorf recommends, however, that power should also be given to the ranking members of those committees. Sharing this power would reduce the risk of the majority's using dynamic scoring strategically and would maintain more consistency in the scoring's application.

Sixth, the adoption of dynamic scoring at the start of 2015, as important as it is, *is not as big a break with past practice as it first appears.* CBO and JCT have been publishing dynamic analyses for more than a decade, including analyses of major tax reforms (JCT 2014a, 2014b), stimulus proposals (CBO 2009), and presidential budgets (CBO 2015). The techniques used in those analyses have been refined through experience and external review.

Immigration reform proposals in 2006, 2007, and 2013 provided special opportunity for the agencies and Congress to prepare and consider scores that included some macroeconomic effects. When CBO and JCT did this in 2006—at a time when I served as CBO's acting director—we expected controversy, but there was none; lawmakers and outside analysts understood that it made sense for CBO and JCT to consider the effects of an increased labor force when evaluating immigration reform despite the convention of not including any macroeconomic effects for other bills (Marron 2013).

In addition, the agencies incorporate the macroeconomic effects of fiscal policy when constructing their twice-yearly budget baselines. In August 2012, for example, CBO had to project the budget outlook in the face of the then-looming "fiscal cliff," a panoply of scheduled tax increases and spending cuts. Those provisions would have amounted to $500 billion in fiscal tightening in 2013. As a result, CBO (2012) projected that the economy would be pushed into recession. That projection provided important context for fiscal cliff deliberations.

Dynamic analysis, immigration reform, and baseline projections have thus paved the way for the agencies to include macroeconomic feedbacks in official scores.

Seventh, *the strongest concerns about dynamic scoring—uncertainty and the risk of bias—are understandable but manageable.* Macroeconomic impacts are uncertain, and experts disagree on how best to model them. How myopic or forward-looking are individuals when making work, saving, and consumption decisions? How will the Federal Reserve respond to changes in fiscal policy? What are the feedbacks between the U.S. economy and the rest of the world? How much do deficits crowd out private investment? Can fiscal policy reduce hysteresis effects in the aftermath of a deep recession?

In principle, such uncertainties could create opportunities for the agencies to put a thumb on the scale to favor results preferred by their political masters. In reality, the cultures and staffing of CBO and JCT are fundamentally nonpartisan. The two agencies' directors are chosen by the congressional majority, and they certainly hear from congressional leaders about important bills, but they have done an admirable job maintaining their nonpartisan credibility.

Moreover, concerns about uncertainty and potential bias apply equally to many estimates the agencies have traditionally produced. Future economic conditions and behavioral responses are highly uncertain. What will be the take-up rate of a never-before-seen benefit program? How will state governors and legislatures respond to new flexibility in a federal program? What are the probability and magnitudes of potential terrorist attacks in the United States? What will oil prices be in 2025? What will electromagnetic spectrum sell for in 2022? What new medicines will come to market in the next decade? At what price and usage? How will the Supreme Court respond to a possibly unconstitutional piece of legislation?

The strong nonpartisan culture of the agencies has allowed them to make objective calls on such questions in the past and will allow them to do so with dynamic scoring in the future. That culture also implies that the effects of dynamic scoring will be less rosy than some proponents claim and less dire than some opponents fear. President Harry S. Truman once asked for a one-armed economist. He would not find any at CBO and JCT. Instead, the agencies will provide classic many-handed analyses that include offsetting effects.

In analyzing macroeconomic feedback from tax cuts, for example, the agencies will consider how they might encourage working, saving, and

investing. But they will also consider how tax cuts may increase after-tax income and thus reduce work and how any resulting deficits may eventually reduce private investment. To use the jargon, CBO and JCT will consider the income effects and the crowding-out effects of tax cuts, not just the substitution effects that proponents emphasize. That three-handed approach tempers the potential macroeconomic effects of tax cuts. Indeed, it reveals that some tax cuts reduce economic growth and thus have a larger budget cost than conventionally estimated (CBO 2010).

The same is true of the dynamic scoring of spending provisions, stimulus efforts, and other policies that often have the opposite political valence. Proponents of stimulus often emphasize the potential boost from putting money in peoples' pockets and the multiplier effects that this may set in motion. But CBO and JCT also consider whether and how much Federal Reserve policy may offset such effects and what long-term drag will result from accompanying deficits and accumulated debt (CBO 2014).

The reality of dynamic scoring is thus unlikely to live up to the hype. Instead, dynamic scoring will modestly improve the budget projections that inform fiscal policy deliberations.

REFERENCES FOR THE MARRON COMMENT

CBO (Congressional Budget Office). 2009. "The State of the Economy and Issues in Developing an Effective Policy Response." Statement of Douglas W. Elmendorf before the Committee on the Budget, U.S. House of Representatives, January 27. Washington. https://www.cbo.gov/publication/41755

———. 2010. "The Economic Outlook and Fiscal Policy Choices." Statement of Douglas W. Elmendorf before the Senate Committee on the Budget, September 28. Washington. https://www.cbo.gov/publication/21836

———. 2012. "An Update to the Budget and Economic Outlook: Fiscal Years 2012 to 2022." Washington. https://www.cbo.gov/publication/43539

———. 2014. "How CBO Analyzes the Effects of Changes in Federal Fiscal Policies on the Economy." Washington. https://www.cbo.gov/publication/49494

———. 2015. "A Macroeconomic Analysis of the President's 2016 Budget." Washington. https://www.cbo.gov/publication/50734

JCT (Joint Committee on Taxation). 2014a. "Estimated Revenue Effects of the 'Tax Reform Act of 2014.'" Report no. JCX-20-14. Washington.

———. 2014b. "Macroeconomic Analysis of the 'Tax Reform Act of 2014.'" Report no. JCX-22-14. Washington.

Marron, Donald. 2013. "Immigration, Dynamic Scoring, and CBO." Blog post, May 3. Washington: Urban-Brookings Tax Policy Center, TaxVox. http://taxvox.taxpolicycenter.org/2013/05/03/immigration-dynamic-scoring-and-cbo/

GENERAL DISCUSSION Peter Orszag spoke first to say that the arguments against dynamic scoring have never been theoretical but, instead, have been pragmatic. A political economy argument can be made on both sides. Moreover, the Congressional Budget Office (CBO) has been providing estimates of macroeconomic feedback effects for some time in the form of dynamic budget analysis. The question is whether they should be directly incorporated into the budget score. Echoing discussant Donald Marron—who worked as acting director of the Congressional Budget Office (CBO) in 2006—Orszag admitted that a decade ago he would have been more concerned about dynamic scoring. Nowadays he is less concerned, and he agreed with Douglas Elmendorf that it is marginally better to incorporate dynamic scoring directly into the budget score than not to. In his view, the debate over the 2009 stimulus package would have been much better informed if this had been done at the time.

Orszag offered two notes of caution. First, he stressed the importance of applying dynamic scoring to both spending *and* tax proposals, noting that this may not necessarily be consistent with the current political environment. Second, he warned about potential abuses of the process, and offered two suggestions to protect against it. The appointment of the CBO director has traditionally been somewhat bipartisan, with informal cooperation between the chairs and ranking members, but that arrangement can easily fall apart, especially in times like the present where polarization reigns. The answer is to make the appointment formally bipartisan. Orszag also would encourage outside entities—such as the Tax Policy Center—to play a greater role than they do now in providing a check on the reality of what the CBO is doing in the dynamic scoring.

Alan Blinder agreed with Orszag that it is important to think about how to increase the independence of the CBO from political meddling given the new scope for difficult judgments that dynamic scoring would open. He pointed to the well-known fact that long-term projections, which often go "out of sample," can magnify standard errors, which makes it all the more important to protect the neutrality of the estimating process. Blinder suggested that the greater independence from political meddling enjoyed by the Federal Reserve Board is a standard that should be aimed for, hard as that may be.

William Gale voiced similar concern over the political economy implications. He argued that one could conclude from Elmendorf's paper that while dynamic scoring should be used, the way the House currently narrows its application to tax cuts and disallows the minority from requesting it renders it unbalanced. Gale felt that as a budget rule, dynamic scoring

could be used to prop up political agendas as much as it could be used for illumination.

Alan Viard also acknowledged that the risks of political bias could be heightened by requiring dynamic scoring. However, he added, estimates of macroeconomic effects by a respected agency whose directors have come from both political parties could help combat exaggerated claims presented by outside advocates, such as claims that tax cuts are likely to fully pay for themselves.

Jason Furman pointed to the value of the second half of the paper, which discussed the CBO's role in providing other information to policymakers that may be relevant to them, in addition to estimating budget impacts. This role deserves more careful thought, since some policymakers can have limited attention spans and often ignore important details of interpretation. For example, fiscal policy analysis that just shows growth or jobs effect but gives short shrift to welfare effects reflects an elementary cost-benefit mistake we would not make in the regulatory arena. Furman stressed that growth effects should be reported in tandem with welfare effects, since policymakers will not make the right choices if CBO only reports growth or job metrics without embedding them into a broader context of distribution and welfare. He argued for the creation of some type of simple summary statistic—not just in footnotes or caveats—to capture such total effects.

Gregory Mankiw characterized Elmendorf's paper as eminently sensible and then also raised two points of concern. He noted that in the federal government, a lot of attention is paid to the 10-year budget window, and the CBO projections are no different. From a president's standpoint, 10 years may seem like forever, but for the economy it is a relatively short time frame. Mankiw's worry was that the focus on a 10-year frame might overemphasize short-run Keynesian demand effects and underemphasize longer-run classical growth effects. Dynamic scoring—which currently projects only 10 years into the future—could also lead to shortsightedness. He recommended that a balance be struck in the estimating between short-run and long-run effects.

Along similar lines, Mankiw also stressed that any well-specified general equilibrium model must close budget gaps in the long run, otherwise the proposal will be incomplete. Proposals that cut taxes without stating how the budget gap will be closed in the future are incomplete, and so are those that propose to increase spending without providing a way to pay for it. Outcomes will depend crucially on how the policies are closed. To take two current examples, one might propose to pay for a stimulus package by cutting Social Security in the long run; or by increasing capital taxes

in the long run. In fact, those two approaches would have very different steady-state effects. On this point, he strongly disagreed with Elmendorf, who seemed to say that if one had a well specified model, how this loop was closed would not matter for the long-run effects. In short, Mankiw felt, dynamic scoring needs to take into account the long-run effects of closing policy proposals.

Gale echoed the importance of fiscal closure rules. He agreed with Mankiw on the principle that tax cuts financed by future tax increases have different long-term consequences than future spending cuts. Budget constraints are real, and policymakers must specify how a policy is going to be paid for.

Martin Feldstein praised Elmendorf's paper for its advocacy of dynamic scoring as well as its discussion of the technical problems associated with it. However, he expressed worry in how dynamic scoring is actually implemented. He reminisced about the "bad old days" of budget scoring, before the mid-1970s, when revenue estimation was done with the assumption that changes in tax rates had no effect at all on taxpayer behavior. This assumption was even applied to taxation of capital gains, which Feldstein noted can actually have a very large effect. He recalled how when he explained this assumption in testimony before the Senate Finance Committee in the late 1970s, Senator Russell Long's shocked reaction shamed the estimators into changing their approach. Since then the joint tax committee and the Treasury staff have applied the convention that changes in tax rates do change behavior, though not in the way Elmendorf has emphasized, with an impact on GDP.

In thinking about the response to changes in the personal income tax, Feldstein distinguished between three kinds of effects, which he believed were not adequately captured by the current dynamic scoring methods. The first kind are the short-run aggregate demand effects, including how the Federal Reserve responds to offset the fiscal impact. The second are the longer-run growth effects with their impacts on savings, investment, and human capital. The third kind—which in Feldstein's view are the most overlooked—are the permanent revenue effects of behavioral changes in response to changes in marginal tax rates. To illustrate, he named three ways behavior changes in response to a reduction in the marginal tax rate: by increasing the labor supply, broadly defined, including occupational choice; by influencing the form of compensation, since workers will prefer taxable cash over benefits when rates are lowered; and through changes in spending on tax-favored consumption, such as mortgage interest and charitable contributions. He said the response of labor supply elasticities,

hours, and participation to tax rate changes is small in comparison with the response of taxable income, according to available microdata. Any dynamic model needs to look much more closely at the latter.

Jeffrey Kling, representing the CBO, assured Feldstein that the CBO has long attempted to incorporate some of the issues he described regarding compensation and the labor supply. Its staff continues to study, including impacts on taxable income, to try to synthesize the findings and apply them to proposals that would affect forms of compensation. At the same time, he noted that while the CBO could probably do more in this area, it actually falls under the jurisdiction of the Joint Committee on Taxation.

Alice Rivlin noted that most of the discussion thus far had been about tax changes, with little attention given to government spending and its macroeconomic effects. She acknowledged that spending is much harder to model, both because the evidence base is not nearly as well developed as for taxation and because many public investments, such as those in early childhood education, can only improve productivity over a very long term and in the aggregate. However, Rivlin stressed, it is as important to consider the very long-range effects of spending as it is to examine the very long-run effects of taxation.

Caroline Hoxby noted that to a microeconomist, taking the central tendency of estimates seems quite unnatural. Microeconomists would prefer to take the best econometrically identified estimate, the one that is closest to what one believes would occur were a policy to go into effect, ideally based on using a randomized trial. Applying a midpoint or consensus estimate essentially assumes that mistakes made on one side are offset by those made on the other or that modeling decisions somehow are distributed in some normal way.

Kling—again speaking for the CBO—clarified that a "central estimate" is informed by a judgment of the reading of the literature, which is not necessarily arraying all of the point estimates and taking the midpoint. As an example, if analysis showed that a policy had a 60 percent chance of having no effect and a 40 percent chance of having a positive effect, the CBO would synthesize that as some positive number.

Justin Wolfers remarked that there seemed to be an emerging consensus in the room that there need be no more debate about the principle of dynamic scoring. If so, he disagreed. One reason to continue the argument over its merits stems, he said, from a simple statistical principle of shrinkage estimators. When one has a raw, unbiased but noisy statistical estimate, a forecast is improved in a mean-squared-error sense by shrinking it back to some prior, yet for many of the policies under discussion the prior

was simply that "anything could happen." If that is the case, according to Wolfers, the right thing to do would be to shrink toward zero. This led him to agree with Marron, who proposed that the midpoint of macro thinking indeed adds value. He concluded that fighting about dynamic scoring is sort of a second-best way of actually getting shrinkage estimators and more efficient estimates.

Wolfers also returned to the earlier points made about political bias. Will dynamic scoring yield a pro–tax cut bias? He proposed thinking about the social and political leanings of people within the economics profession itself. One could caricature (in good humor) the profession as consisting of people who hate inflation becoming monetary economists, those who love global trade becoming international economists, those who like workers becoming labor economists, and those who like capital and rich people becoming financial economists. Wolfers conjectured that those in the field of public finance—and particularly those who analyze tax cuts—contain a higher share of Republicans than any other fields in the profession, and if that really is so it could add a pro–tax cut bias to their dynamic scoring.

Brad DeLong noted that when he was a Treasury political appointee, one of the Treasury career staff economists lectured him about dynamic scoring thus: "Brad, you people come in with your exaggerated belief in the productivity benefits of public investment. And so you command us to score your policies as having a very favorable impact on the deficit. They come in with their exaggerated belief in the benefits of tax cuts. They command us to score their policies as having a very favorable impact. We cannot say we disagree with our bosses' analytic judgments. But by holding the line and stating that we do not consider any macroeconomic effects of policies, we can at least prevent being whipsawed by this partisan rosy-scenario ratchet."

He noted that being whipsawed by the partisan rosy-scenario ratchet is a serious danger, as evidenced most recently by the recent semi-score of the Jeb Bush tax plan by Feldstein and others.[1] There would be an upside if appropriate real technocratic dynamic-scoring corrections were significant. But, he concluded, they mostly likely are not.

Phillip Swagel raised the issue of Congress's budget process itself and its influence on policies and the economy. Economists too often overlook the significance of budget rules, he said, including Senate procedures, which

1. John Cogan, Martin Feldstein, Glenn Hubbard, and Kevin Warsh, "Fundamental Tax Reform: An Essential Pillar of Economic Growth; An Assessment of Governor Jeb Bush's 'Reform and Growth Act of 2017'" (New York: Center for Global Enterprise, 2015).

have a real effect on economic outcomes. Swagel has been impressed by the CBO's and the Joint Committee on Taxation's willingness to push back against budget gimmicks, citing as an example their analysis of proposals for repatriation of foreign earnings.

Viard voiced support for dynamic scoring, noting that its most important role was not to favor or disfavor tax cuts, as Wolfers had suggested. Rather, he suggested that a more important purpose was to help illuminate which kinds of tax cuts are better for economic growth than others. Similarly, the extension of dynamic scoring to spending proposals could help illuminate which kinds of spending are better for growth.

George Perry voiced a concern that no one else had raised, namely that while it had been discovered fifty years earlier that fiscal policy has a special role to play in a depressed economy as a stabilization tool, today this gets overlooked in a scramble to tease out what he considers third-order effects. Supply-side effects, like hysteresis and the use of unemployment insurance are, in the bigger picture, not nearly as important in an underemployed economy as the first-order effects of fiscal policy.

Elmendorf spoke in response to all the comments, first by agreeing with Furman about the importance of distinguishing between the effects on GDP and the effects on people's welfare. He noted that in its analysis of immigration reform, the CBO was careful to distinguish between total GDP and GDP per capita, which is a better measure of welfare. He noted that it is always a challenge when publishing analyses to explain what the numbers mean, since so many policymakers will go just to the numbers without reading the words around them.

Elmendorf disagreed with Mankiw's view that too much attention to the 10-year window made it harder to properly understand the steady state. He believes the CBO's analyses of short-run effects of policy changes remain vital, as they can be very large effects. And because policies are rarely permanent, whether the long-run steady-state effects of a policy can ever be realized is unclear. Moreover, when the CBO does think the longer-term effects of a policy will differ from the effects in the first decade, it generally explains those longer-term effects as well, its work on the Affordable Care Act and immigration reform being two salient examples.

To the question about the value of picking a midpoint estimate in dynamic modeling, he answered that in his first few years at the CBO, the practice was to publish only the low end and the high end of the range. That left policymakers free to pick their favorite number at the extreme, however, and offered them no other simple choice. The CBO

improved on this a few years ago by adding in the central number and releasing that along with the full range.

Addressing the political economy aspect of dynamic scoring and the independence of the CBO and the staff of the Joint Committee on Taxation, Elmendorf assured everyone that there is an ongoing understanding and respect among congressional leaders for that independence. He pointed out that he had originally been appointed director of the CBO by two Democrats and was reappointed by a Democrat and a Republican. Likewise, Thomas Barthold was originally appointed as the staff director for the Joint Committee on Taxation when Democrats controlled both the Finance and Ways and Means committees, and was subsequently reappointed when Republicans controlled both. In response to Blinder's comparison with the greater independence of the Federal Reserve Board, he noted that many Members of Congress are in fact not very happy with having given so much latitude to the Federal Reserve, and surely they would be much less happy doing the same for agencies working on fiscal policy.