Brookings Papers
ON ECONOMIC ACTIVITY

BPEA Conference Draft, September 26-27, 2024

Dynamic Scoring: A Progress Report on Why, When, and How

Douglas Elmendorf (Harvard University)

Glenn Hubbard (Columbia University)

Heidi Williams (Dartmouth College)

Conflict of Interest Disclosure: Heidi Williams is a visiting scholar at the Congressional Budget Office. The authors did not receive financial support from any firm or person for this article or, other than the aforementioned, from any firm or person with a financial or political interest in this article. The authors are not currently an officer, director, or board member of any organization with a financial or political interest in this article.

Dynamic Scoring: A Progress Report on Why, When, and How

Douglas Elmendorf
Harvard University and NBER

Glenn Hubbard
Columbia University and NBER

Heidi Williams
Dartmouth College and NBER

September 10, 2024

This paper was prepared for the September 2024 conference of the *Brookings Papers on Economic Activity*. Please send comments to the authors at doug.elmendorf@hks.harvard.edu, rgh1@gsb.columbia.edu, and heidi.lie.williams@dartmouth.edu. Williams is also a visiting scholar at the Congressional Budget Office. The views expressed here are those of the authors and should not be attributed to any of the individuals acknowledged or any institution with which the authors are currently or formerly affiliated.

Abstract

By design, official budget estimates for legislative proposals generally exclude the proposals' likely effects on levels of labor, capital, productivity, and other economic outcomes, as well as any feedback effects from changes in those outcomes to the federal budget. Policymakers would benefit from knowing the expected sizes of those economic effects, and advances in research and in the estimating agencies' tools and experience have made providing estimates of those effects more feasible. If Congress requested that those effects be included more often in budget estimates—so-called "dynamic scoring" of legislation—the advantages and disadvantages of doing so would vary across policy areas. For some areas, the budgetary impacts of the currently excluded effects have been estimated to be significantly different from the impacts of the included effects. But producing dynamic estimates would be substantially more time-consuming than producing conventional estimates, and in some areas, the research base needed to inform modeling of the relevant economic effects is insufficient for credible estimation.

Economists who study the growth of productivity and output often emphasize the role of public policies that affect labor supply, savings, and federal investments in infrastructure and research. Yet, the official budget estimates provided to Congressional policymakers who consider changes in those policies generally exclude—by design—their impact on labor, capital, productivity, and other economic outcomes, as well as any feedback effects from changes in those outcomes to the federal budget.

The exclusion of those economic impacts is a longstanding convention of budget estimates produced by the Congressional Budget Office (CBO) and the staff of the Joint Committee on Taxation (JCT). CBO and JCT have demonstrated a growing capacity to analyze such economic impacts and their feedback to the budget, and the agencies have provided more analyses that incorporate these features over time. But those analyses have been limited in number and generally have been provided separately from the official estimates on which Congressional and public attention is often focused during policy debates. Those official budgetary effects of proposed legislation continue to be estimated by the agencies almost entirely through "conventional scoring" that excludes those economic impacts rather than so-called "dynamic scoring" that includes them.

This paper is not intended to make the case for dynamic scoring, but instead to analyze the economic and institutional issues presented by the choice between conventional and dynamic scoring. Thoughtful observers of the budget process, including former staff members of CBO and JCT, have a range of views about the best approach, and the agencies appropriately defer to Congress on this choice.

The current estimating convention has been supported by various lines of reasoning. One is that producing dynamic estimates takes too much time and work to be feasible given the large number of proposals for which Congress requests estimates, the many iterations through which some proposals are developed, and the short timeframes typically available for preparing estimates. We take the practical constraints faced by the estimating agencies very seriously. However, CBO and JCT have invested—to their credit—in the tools and experience needed to produce dynamic estimates, and those investments have reduced to some extent the additional time and work needed to generate such estimates. In our assessment, further targeted investments could make more progress along these lines, but producing dynamic estimates would nonetheless be substantially more time-consuming than producing conventional estimates in almost all cases.

Another line of reasoning is that the effects of policy changes on outcomes such as labor and productivity may be inherently more uncertain than other effects—and perhaps relatedly, estimates of those effects might be more subject to external pressure. In our assessment, uncertainty about the excluded effects is not necessarily greater than uncertainty about effects that are included. Indeed, for at least some important excluded effects, CBO and JCT can draw on substantial bodies of evidence and receive professional feedback on their use of the evidence. Moreover, including these additional effects could reduce the risk of bias in estimates.

A further line of reasoning in support of the current estimating convention is that the excluded effects are small for most proposals. The size of excluded effects can be evaluated in two ways. One is the impact on the aggregate economy, which is indeed small for most proposals, but that information itself can be useful for policymakers. The other is the impact on the overall budget estimate for a proposal, which is large in some cases. We document an especially stark example: For a potential change in the authorized number of high-skilled immigrants, the budgetary impact of the change in population that is excluded from conventional estimates is much larger than the impact of the included factors and has the opposite sign.

Yet another line of reasoning is that CBO's and JCT's credibility might be damaged by changing their estimating methodology, and in particular by a loss of transparency due to the complexity of including additional effects. Maintaining the agencies' credibility is indeed crucial for the policymaking process. But in our assessment, the challenges of change and complexity that would arise in doing more dynamic scoring are similar to those that arise in conventional scoring—for which the agencies have made many methodological advances over time, usually involving greater complexity and bolstered by substantial explanations. In addition, in our assessment, excluding from budget estimates behavioral responses that people expect to occur also can hamper the agencies' credibility.

We want to underscore that budget estimates from CBO and JCT are not the be-all and end-all of policymaking, even though estimates are sometimes blamed for blocking or inducing certain outcomes. The agencies provide considerable information beyond budget estimates, to Congress as additional material in published estimates and through informal exchanges, and to Congress and others through published reports on policy issues. Other government agencies, analysts outside government, and other private actors also communicate with policymakers and

the public. Still, official budget estimates are often consequential in the legislative process, so the conventions underlying those estimates warrant careful and ongoing attention.

In addition, the exclusion of certain economic impacts from budget estimates is not the only aspect of the estimating process that deserves scrutiny and thoughtful consideration of alternatives. So-called "scorekeeping guidelines" that have been developed by CBO, the Office of Management and Budget, and the House and Senate Budget Committees to help align the estimating process with legislative structures and procedures can result in the exclusion from official estimates of other effects of proposed policies beyond the ones we address in this paper (CBO, 2021a). For example, estimates of proposed activities to reduce fraud often exclude the impact of those activities on program spending or tax collections; Hall (2024) explains why this occurs and proposes changes to the scorekeeping guidelines to address the issue.

In the sections of the paper that follow, we elaborate on the advantages and disadvantages of dynamic scoring, describe potential processes and criteria for Congress to make decisions on the subject, and explain how dynamic scoring might be feasibly undertaken. Because discussions of dynamic scoring over the past three decades generally have focused on tax policies, and because we have greater familiarity with the procedures used by CBO, we focus on spending policies and CBO's methodology, but we include substantial references to JCT's important work on these topics as well. Many of these issues apply also to the Office of Management and Budget, the Treasury Department's Office of Tax Analysis, and unofficial budget estimators outside the government.

We then present three case studies that illustrate some of the issues raised in the first part of the paper. Our first case study is immigration policy, for which economic research provides a substantial evidence base, CBO and JCT have undertaken considerable modeling, and conventional scoring can be misleading. The second case study is federal investment in research and development (R&D), where a reasonable evidence base and modeling infrastructure for dynamic scoring are present. The third case study is federal permitting of investment in infrastructure, where the modeling infrastructure is present but the evidence base in the research literature is currently too slight for dynamic scoring to be feasible.

I. Why?

We begin with the crucial question of "why" or "why not" to do dynamic scoring; the subsequent sections turn to the following questions of "if so, when" and "if so, how."

I.A. Budget Estimates

CBO and JCT provide to Congress the official estimates of the effects of legislative proposals on the federal budget. CBO produces public estimates for bills after they have been approved by Congressional committees; the process of producing those "cost estimates" is often described as "scoring." For bills that would alter the tax code, CBO's cost estimates use changes in estimated revenues provided by JCT, which publishes its revenue estimates directly as well. In this paper we use the term "budget estimates" to emphasize that the estimates include changes in both spending and revenues when relevant. In addition, the agencies give private estimates to Members of Congress and their staffs for proposals that are being developed and have not been released publicly. In a typical year, CBO publishes roughly 700 public estimates, and it and JCT provide thousands of private estimates. For more information on this process, see CBO (2023c) and JCT (2023).

CBO and JCT also publish analyses of actual or potential legislative proposals beyond budget estimates, and some of those analyses include effects on labor, capital, productivity, and other economic measures. For example, in 2013, CBO released a cost estimate prepared by the agencies for an immigration bill being debated in the Senate (CBO, 2013a), and CBO simultaneously released a report with "estimates of the overall economic impact of the legislation and of the incremental federal budgetary effects of changes in the economy that the cost estimate does not reflect" (CBO, 2013b). Also, for some bills, JCT publishes what it terms "macroeconomic analyses," which include effects on broad economic measures and are examples of what we refer to in this paper as "dynamic analyses." We use the term "dynamic analysis" for work that addresses broad economic changes outside of official budget estimates and the term "dynamic scoring" for official budget estimates that include those economic changes. The distinction is important because official budget estimates are often the focus of Congressional and public attention during policy debates.

In addition, CBO provides projections of budgetary and economic outcomes under current law that are known as "baseline projections." These projections are generally updated

two times per year. Cost estimates are designed to show effects relative to these baseline projections.

Budget estimates are used by Congress to evaluate alternative proposals, to iterate in policy development, and to enforce budget plans. CBO has been providing cost estimates since its founding roughly half a century ago, and JCT has been in place for nearly a century.

Both baseline projections and cost estimates typically focus on nominal cash flows for the current fiscal year and each of the 10 subsequent years, a period that is often called the "budget window." Baseline projections are extended in less-detailed form beyond the budget window, and cost estimates are sometimes extended as well when Congress is especially interested in long-term effects and when CBO expects that a proposal's long-term effects would be notably different than its effects over the coming decade.

The cost estimates and baseline projections depend on assessments made by analysts at CBO and JCT and on procedures that have been codified in law or agreed to formally or informally by the relevant Congressional committees. The agencies provide objective, nonpartisan analysis at a high professional standard. As a result, their estimates are important not only to Congress but also to press coverage of policy deliberations and to the broader public.

I.B. Behavioral Responses

CBO's and JCT's estimates generally include the impact of behavioral responses to the proposed changes in law. For example, estimates for changes in benefit programs include shifts in take-up rates, and estimates for changes in income tax rates include shifts in the use of tax deductions. The agencies try to account for the behavior of households, businesses, implementers of legislation in the executive branch, and state and local governments; they do not attempt to predict future changes in federal law.

If budget estimates of proposed policies include all the foreseeable behavioral responses, they are more likely to be closely aligned with actual effects. Also, policymakers receive through budget estimates information about behavioral responses that may be independently relevant to their decision-making. For example, policymakers may want to understand how a potential change in a benefit program would affect take-up for its own sake—as, for example, with subsidies to health insurance—in addition to the impact of take-up on the budget.

However, CBO and JCT generally do not include in official budget estimates the budgetary impacts of behavioral responses that would affect the overall economy, which includes impacts on labor supply, on saving and investment, on productivity, and on aggregate demand. As CBO (2023c) explained: "According to long-standing practice, CBO's conventional cost estimates reflect the expectation that nominal gross domestic product [GDP] … would not change."[1]

For economists, the exclusion of those behavioral responses can seem quite odd. A change in a benefit program might induce an individual to stop or start taking up those benefits *and* to change their work effort; from an economic perspective, there is no categorical difference between those responses, and they may well be directly related. Similarly, a change in the tax code might induce an individual to adjust their tax deductions *and* their saving behavior, and there is no fundamental economic difference between those responses. Moreover, policymakers might be especially interested in receiving information about shifts in labor, capital, and productivity—and the follow-on effects on output, inflation, interest rates, and other economic measures—that would be induced by policy changes they are considering.

Why, then, are those behavioral responses excluded from conventional budget estimates? In concrete terms, why are some people who have produced and used budget estimates, and who are committed to improving their quality, skeptical about dynamic scoring? Four concerns have been most central to this long-running debate—the feasibility of dynamic scoring given time and resource constraints, the uncertainty of dynamic estimates, the magnitude of excluded effects, and the credibility of estimates. We take up these concerns in turn.

I.C. Feasibility of Producing Dynamic Cost Estimates Given Time and Resource Constraints

Including additional behavioral responses in budget estimates is challenging because of the large number of proposals to be analyzed, the many iterations that most proposals go through, and the short timeframes for producing estimates that Congress allows. These practical constraints are real, and they should be central to discussions about dynamic cost estimates.

A decade ago, one of us (Elmendorf) wrote: "All of the estimates of macroeconomic effects described above involved significant conceptual and practical challenges and required a

[1] JCT's cost estimates reflect the expectation that nominal GNP, rather than nominal GDP, would not change. This difference and other complications that arise in applying the general principle are not crucial for our analysis.

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. … For example, when CBO [2015a] examined three ways of reducing spending for the Supplemental Nutrition Assistance Program ([SNAP], 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" (Elmendorf, 2015, pages 110-113).

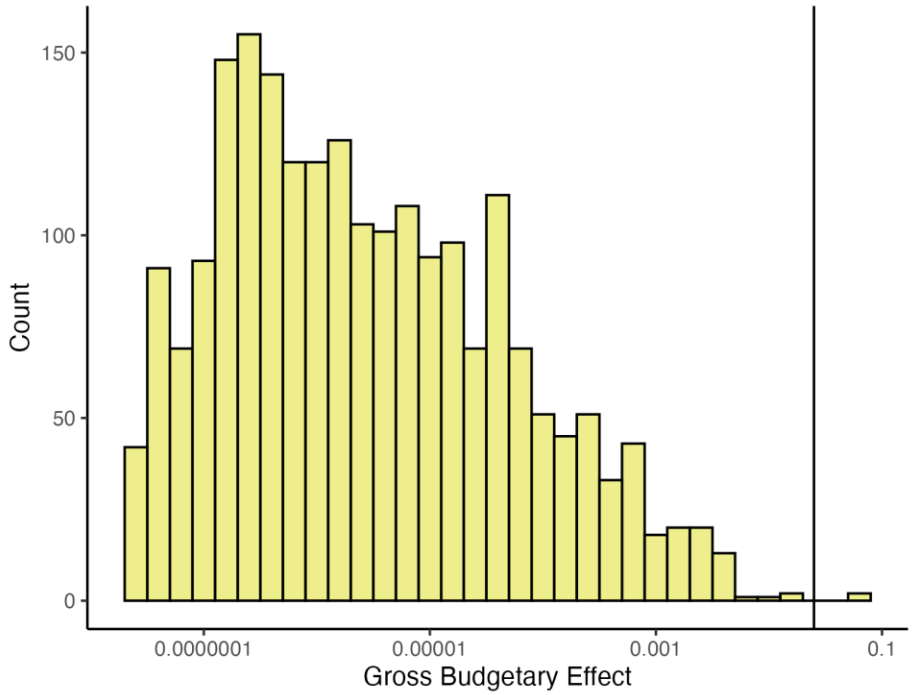
He argued therefore that dynamic scoring should be applied only to major proposals—specifically, proposals for which conventionally estimated changes in revenues, spending, or deficits exceeded one-quarter percent of projected output over the 10-year budget window (or about $900 billion currently). Indeed, the House of Representatives currently requires CBO and JCT to provide dynamic estimates, if practicable, for all bills that exceed that threshold in any year of the budget window (which ranges from about $70 billion to about $100 billion currently) or are deemed "major" by the chair of the House Budget Committee or of the Ways and Means Committee (Congressional Research Service, 2023). Some version of this rule has been in effect for much of the past decade.

As a practical matter, dynamic scoring has occurred very rarely. Among the roughly 2700 of CBO's published budget estimates between 2019 and January 2024 for which we could readily code gross budgetary effects, only one had effects large enough to warrant a dynamic estimate, shown on the far right of Figure 1. For this proposal, CBO and JCT found that the excluded effects were negligible, offering an example of the distinction between a bill's overall size and the size of excluded effects. JCT has published roughly a dozen dynamic estimates over the past decade, some of which we discuss below.

Whatever the shortcomings of conventional scoring, dynamic scoring unavoidably requires additional calculations and review relative to conventional estimates, and when Congress is impatient to have a policy debate, it wants estimates quickly. Thus, doing dynamic scoring on a regular basis would necessarily involve standardizing and simplifying some processes. Standardizing and simplifying present risks. CBO emphasizes that estimates are produced by analysts rather than models, because analysts build models, translate legislative language into model inputs, and review model outputs to ensure that the modeling is suitable for the proposals at hand. The agencies' investment in the tools of dynamic analysis and the experience they have gained in doing such analysis have reduced somewhat the work and time

required for future dynamic estimates and suggest that further advances are possible. However, even with further investment along these lines, dynamic estimates could not generally be produced as quickly as conventional estimates—so shifting to dynamic scoring would require Congress to accept longer cycle times in almost all cases.

Figure 1. Gross budgetary effects and the major legislation rule



Notes: This figure plots the distribution of estimated gross budgetary effects for the universe of CBO's cost estimates as drawn from the XML files of estimates from the 118th, 117th, and 116th Congresses through January 2024; see https://www.cbo.gov/cost-estimates/xml. The x-axis shows gross budgetary effects relative to GDP and uses a log scale for readability. CBO generally computes gross budgetary effects by summing the absolute value of the budgetary effects of provisions and their interactions (see Edelberg, 2016); a "provision" is a line in a table summarizing a CBO estimate. The gross budgetary effects presented here are approximations, because they ignore some uncommon complications that are difficult to calculate based on publicly available data. The sample shown here excludes two sorts of bills that are complicated to hand-code: reconciliation bills and the National Defense Authorization Act; the graph is visually quite similar if those observations are included. The vertical line denotes the threshold for the House rule in the 118th Congress.

To illustrate the work that CBO and JCT have done in recent years and the progress they have made in dynamic analysis, we point to reports from JCT on the economic effects of four major pieces of tax legislation that we return to shortly (JCT, 2015b, 2015c, 2017, and 2018b) and reports from CBO on repealing the Affordable Care Act, increasing federal investment, responding to the pandemic, funding more federal infrastructure, and raising the minimum wage

(CBO, 2015b, 2016, 2020, 2021c, and 2023f). In addition, CBO has introduced more dynamic complexities into its baseline projections; a striking example is the agency's painstaking collation of data on recent immigration and careful modeling of the likely economic and budgetary effects of that immigration over the coming decade (CBO, 2024f). Moreover, as we discuss later, the agencies have published reports describing models that have been created or improved and the ways those models have been used in specific analyses.

The agencies' new and refined tools and greater experience have made future dynamic estimates somewhat more tractable because the models and learning can be used repeatedly, just as the models and learning underlying conventional analyses are used repeatedly. Moreover, as the agencies have applied their technical frameworks to more policy issues, they have refined some of their analytic approaches to be more standardized, analogous to the way that conventional estimates can become standardized with practice. For example, CBO's 2016 and 2021c reports cover similar ground in describing how the agency models the economic effects of federally funded infrastructure investments, but the later report presents a practical framework that can be deployed in future analyses. Further development of that framework could give analysts standardized algorithms for capturing effects that are now excluded from cost estimates. Similarly, CBO has encapsulated its expertise in estimating the budgetary impact of economic changes into a "Budgetary Feedback Model" (Frentz et al, 2020), which has been further boiled down to an online workbook (CBO, 2023b) and back-of-the-envelope formulas (CBO, 2024b).

Still, producing dynamic estimates still will be substantially more time-consuming than producing conventional estimates in almost all cases. Even when standardized and simplified approaches have been created, analysts still need to assess whether they are applicable to the proposal at hand or require adjustment. We describe later the additional steps in the estimating process that generally would be needed to incorporate the full economic effects of a proposed policy.

I.D. Uncertainty of Dynamic Budget Estimates

Budget estimates are unavoidably very uncertain despite efforts to reduce that uncertainty. CBO and JCT derive expected behavioral responses to policy proposals using a broad array of evidence, including economic research, analyses done in-house, and institutional knowledge and anecdotal information from consultations with government agencies and others.

To ensure that CBO is drawing on the best economic research and applying it well, the agency also has standing panels of outside advisers and has outside experts review some work products.

Many proposals would change policies in ways that have never happened before or have happened only under different circumstances, would change factors that matter for the effects of policy changes but have never been examined closely in the research literature, or that propose relying on data that have not previously been collected.[2] Therefore, analysts at CBO and JCT often need to produce estimates with limited evidence. Meanwhile, the procedural rules of Congress rely on point estimates rather than ranges, so the agencies need to focus on point estimates. CBO is attentive to both the "centeredness" and "spread" of its estimates (see CBO, 2015c), and in some analyses, the agency measures accuracy primarily through the root mean square error (RMSE), which combines bias and variance (see CBO, 2023d).

If shifts in labor supply, saving, productivity, and aggregate demand due to policy changes are sufficiently uncertain, including those shifts in estimates would reduce the estimates' accuracy. To illustrate this concern, consider a simple example in which an outcome of potential interest has equal probabilities of taking the values -1, 0, 1, or 2, and an estimator is trying to predict the outcome. If the outcome is excluded from the estimating process by rule, the estimator is effectively predicting 0 all the time, and the RMSE of the estimates is 1.2, as shown in the first row of Table 1. If the estimator correctly predicts the average outcome but has no ability to distinguish between cases where the outcome will be low or high, the estimator predicts 0.5 all the time, and the RMSE dips to 1.1(as in the second row). If the estimator incorrectly predicts the average outcome to be zero but correctly orders the outcomes and limits the variability of the estimates, the estimator might predict (in order of the outcomes above) -1, -0.5, 0.5, and 1, and the RMSE falls to 0.6 (as in the third row); if the estimator also correctly predicts the sign of the average outcome, the RMSE falls further. But if the estimator incorrectly predicts the average outcome, correctly orders the different outcomes, and generates estimates that are more variable than the outcomes, the estimator might predict (in the order above) -3, -1, 2, and 4, and the RMSE rises to 1.6.

These different possibilities show, in the words of CBO (2015c), that including additional effects will probably improve the accuracy of estimates "if the estimates are for policies that are

[2] As just one example, consider the uncertainty in estimating the budgetary impact of a proposal to make eligibility for subsidized school lunches depend on personal information that is not currently being collected.

likely to have significant feedback effects; if the estimates of those feedback effects are centered around the actual effects; and if the spread of those estimates is small." That is, including additional effects will probably improve accuracy if their mean is notably different from zero, if the estimator knows the sign of the mean, and if the variance of the estimates is held down.

Table 1. Errors from different illustrative estimation processes

Estimation process	Outcome				RMSE
	-1	0	1	2	
Excluding the variable by rule					
Estimate	0	0	0	0	
Error of estimate	1	0	-1	-2	1.2
Predicting the average outcome but not the variation					
Estimate	0.5	0.5	0.5	0.5	
Error of estimate	1.5	0.5	-0.5	-1.5	1.1
Not predicting the average outcome but ordering the outcomes correctly and limiting the variance of estimates					
Estimate	-1	-0.5	0.5	1	
Error of estimate	0	-0.5	-0.5	-1	0.6
Predicting the average outcome and the order of outcomes but having a high variance of estimates					
Estimate	-3	-1	2	4	
Error of estimate	-2	-1	1	2	1.6

Notes: The root mean square error (RMSE) for the estimation process of excluding the variable is calculated as the square root of 0.25*{[0-(-1)]^2+[0-0]^2+[1-0]^2+[2-0]^2}); the other RMSEs are calculated similarly.

How do these considerations apply to dynamic scoring? In many cases, budgetary impacts of factors excluded from conventional estimates are significant relative to the budgetary impacts of the included factors and have signs that are widely understood: Examples include

changes in income tax rates, shifts in immigration policy that change the U.S. population, fiscal stimulus during recessions, and changes in federal investment. In addition, economists have paid more attention to at least some important responses that are excluded from conventional estimates than many responses that are included, so CBO and JCT can draw on larger bodies of evidence and receive more professional feedback on their use of the evidence.

At the same time, parameter values useful for estimates of budgetary effects cannot be extracted in a simple manner from the research literature on most topics: Estimates of parameters can vary widely, deciding which estimates to put more weight on can be difficult, and some observers use motivated reasoning in interpreting evidence and could try to pressure the agencies to adopt their interpretation.

In addition, the time dimension of dynamic analysis raises further issues. Many estimates in the research literature are for behavioral responses at a point in time, whereas dynamic analysis requires time paths for parameters. Linking cross-sectional models (such as microsimulation models) to dynamic models is complicated. And uncertainty in modeling economic effects over time goes beyond specific parameters to the choice of model itself; reflecting that uncertainty, JCT deploys three models in its macroeconomic analyses and varies the weights it attaches to each model depending on the nature of the legislation being analyzed (see JCT, 2018a).

However, it is important to recognize that many of the considerations about uncertainty that arise for dynamic estimates are also relevant for key behavioral responses that are included in conventional estimates, such as the effects of increasing the minimum wage on employment and earnings or the effects of various changes in health policy on insurance take-up, as well as in baseline projections.[3] CBO and JCT have responded to those challenges by reviewing evidence, discussing strong and weak points of evidence internally, reporting their evaluation, soliciting professional feedback, and defending their modeling choices to critics. The agencies have done the same for some behavioral responses included in dynamic estimates already, and in our assessment, Congress and the public would benefit from the agencies continuing to do so in the future.

[3] For example, how CBO and JCT might formally take on board the forecasting literature's lesson that accuracy can be improved by "shrinking" estimates is a question that applies to conventional as well as dynamic scoring. See CBO (2019, pages 30 to 35) for one application of CBO's approach to analyzing uncertainty in economic and budget projections.

For example, a dozen years ago CBO undertook a comprehensive review of evidence on the response of labor supply to changes in fiscal policy. In three linked papers (CBO, 2012c, McClelland and Mok, 2012, and Reichling and Whalen, 2012b), the agency documented the wide range of estimated elasticities in economic studies and explained the elasticities it chose to use in its analyses (including lower, central, and higher estimates of substitution and income elasticities for primary and secondary earners at different wage levels). Making choices was unavoidable, because CBO publishes dynamic analyses apart from cost estimates that use those elasticities. Writing about the elasticities provided a basis for those analyses. Similarly, CBO has reviewed research on the response of output to fiscal stimulus under different economic conditions and explained how it applies that research (for example, see Reichling and Whalen, 2012a, Whalen and Reichling, 2015, Seliski et al, 2020, and Demirel and Wilson, 2023), and CBO recently drew on a growing body of evidence related to immigration as it modeled the economic and budgetary effects of the recent immigration surge (see CBO, 2024f).

Note also on the matter of bias that CBO's evaluation of labor supply research did not result in substitution or income elasticities that equal zero (although they might be exactly offsetting for some proposed policy changes) nor fiscal-policy multipliers that equal zero. Deliberately excluding behavioral responses that are not zero greatly increases the chance that the resulting estimates will not be in the middle of the distribution of possible outcomes, which is the target the agency typically aims for.

Finally, evidence regarding the economic effects of changes in various public policies has expanded considerably over time and continues to do so. For the policy areas we discuss later in the paper, we highlight important new research from recent years.

I.E. Magnitude of Policy Effects on Labor, Capital, Productivity, and Other Economic Measures

For almost all proposals, their effects on labor supply, saving and investment, productivity, and other economic measures such as output, inflation, and interest rates would be very small relative to the economy's overall amounts of those objects. However, other considerations related to magnitude are important.

First, even if expected changes in labor, capital, productivity, output, and other measures are very small in macroeconomic terms, their budgetary effects could be large relative to other budgetary effects of the proposal and thus could significantly affect the total budget estimate. For

this reason, we do not refer to the excluded effects as "macroeconomic"; they are indeed excluded because of the assumption that proposals do not affect the size of the macroeconomy, but a key rationale for including them is to capture more accurately the budgetary impact of proposals rather than to improve forecasts of macroeconomic conditions.

For example, as we explain at greater length below, proposals to expand immigration generally would increase both benefit payments and tax revenue. Conventional estimates include the former effect but not the latter, leading to estimated increases in deficits; dynamic estimates include the additional revenue and tend to show estimated decreases in deficits. The greater accuracy of such dynamic estimates is unrelated to the number of people affected by a proposal.

Second, a finding that expected changes in labor, capital, productivity, output, and other economic measures are small—in macroeconomic terms or relative to some other benchmark—can be important information for policymakers regardless of the budgetary implications. Advocates for policies tend to expect large effects on desirable outcomes and small effects on undesirable ones, while opponents tend to expect the opposite. Objective analysis by CBO and JCT can help policymakers and observers sort out the competing claims.

For example, changes in the tax code have effects on labor supply, saving, and investment that are often asserted to be large by some observers and small by others, and that disagreement in turn generates disagreement about the budgetary impact of tax changes. Objective estimates of those effects, even recognizing the unavoidable uncertainty, are valuable to policymakers. In this vein, a comparison of JCT's conventional and dynamic estimates for four significant tax bills shows that the macroeconomic effects were expected to offset between 5 and 26 percent of the conventionally estimated revenue loss.[4] Both the larger and smaller offsets—shown in Table 2—provide information to policymakers.

[4] Table A-1 of Congressional Research Service (2023) lists five significant bills for which JCT has published what it terms "macroeconomic analyses"; JCT has published macroeconomic analyses for other bills that we do not address here. Of the five bills highlighted in that report, we exclude the Restoring Americans' Healthcare Freedom Reconciliation Act of 2015 because that bill—unlike the bills we show—had substantial spending changes that offset some of the tax changes, so comparing the deficit changes with and without the dynamic aspects is not useful for our purpose. In that bill, the dynamic effects reduce the estimated change in tax revenue by 21 percent, which is consistent with the data in Table 2.

Table 2. Comparing conventional and dynamic estimates for tax legislation

Legislative proposal	Budget window	Increase in deficit		
		Conventional estimate	Dynamic estimate	Difference
Tax Relief Extension Act of 2015 (S. 1946)	2015-2025	$97 billion	$87 billion	10%
A Bill to Amend the Internal Revenue Code of 1986 to Modify and Make Permanent Bonus Depreciation from 2015 (H.R. 2510)	2016-2025	$281 billion	$267 billion	5%
Act to Provide for Reconciliation Pursuant to Titles II and V of the Concurrent Resolution on the Budget for Fiscal Year 2018 (often referred to as the Tax Cut and Jobs Act)	2018-2027	$1456 billion	$1071 billion	26%
Protecting Family and Small Business Tax Cuts Act of 2018 (H.R. 6760)	2019-2028	$631 billion	$545 billion	14%

Notes: The increases in the deficit are in nominal dollars and are drawn from the official cost estimates published by CBO; those published documents explain that the estimates were prepared by JCT. The "Difference" column equals the conventional estimate less the dynamic estimate, divided by the conventional estimate.

I.F. Credibility of Estimates

The credibility of CBO's and JCT's estimates is crucial to the agencies' service to Congress and the country. Without agreement about "a common set of numbers" for the budget, policy discussions would be much less coherent and the budget process much less effective, as emphasized to us by our discussant Deborah Lucas. The current level of agreement about the numbers produced by CBO and JCT depends fundamentally on the credibility of the agencies' objectivity and methodologies for producing estimates, which is why the agencies confer

regularly with Congress regarding their activities and procedures. Moreover, maintaining CBO's and JCT's credibility may be especially challenging in the current era of amplified polarization and heightened suspicion about expertise. Therefore, any potential changes in their methodologies should be evaluated in part by whether such changes might diminish or enhance that credibility as well as by Congressional interest and support in such changes more broadly.

All else equal, the agencies' credibility is fostered by consistency over time in scoring methodology and by simplicity of scoring methodology. Members of Congress and their staffs need to be able to understand the key features of estimates, and they cannot be expected to become economic experts given that they are appropriately focused on the spending programs or tax-code items that are their primary interests as well as the many other aspects of their jobs. Including additional behavioral responses in budget estimates would represent a change in estimating approach and would make those estimates more complex than otherwise.

However, "all else" is not generally equal. *Excluding* from budget estimates behavioral responses that Members of Congress and their staffs expect to occur can reduce the agencies' credibility. Accordingly, when the agencies can draw on empirical evidence to include additional behavioral responses (of the conventional sort or the dynamic sort), that inclusion can strengthen the agencies' credibility.

Indeed, the history of CBO's and JCT's analysis is *not* one of stasis in methodology or particular focus on simplicity. Instead, the agencies have made ongoing advances in their methodologies for producing conventional estimates, usually involving greater complexity, in order to improve the accuracy of the estimates they provide to Congress and expand the information presented. And Congress does not appear to be asking to exclude some behavioral responses from conventional estimates to make the estimates simpler; instead, complexity in estimates seems to be expected in order to capture a complex economy and people's complex behavior. Moreover, the agencies have fostered transparency of their evolving analytic approaches by devoting considerable attention to explaining what they do.

Examples of increasing complexity in the methodology of conventional estimates include the following changes by CBO:

- In the 2000s, CBO improved its modeling and experience in analyzing policies regarding health care and health insurance, including requesting and receiving additional funding from Congress to bolster the agency's staff in this area. As a result, the methodology of

CBO and JCT's estimates during the debate about the Affordable Care Act was notably different and more complicated than the methodology underlying estimates of the Clinton health reform plan fifteen years earlier.

- In the 2010s, CBO expanded its use of so-called "fair-value" accounting for the government's financial activities as a complement to the accounting approach specified by the Federal Credit Reform Act of 1990; for example, see CBO (2012a and 2018b). Deborah Lucas played a key role in this effort as an employee and then consultant of the agency.
- CBO's estimates of the effects of increasing the minimum wage have become notably more sophisticated and complicated over time (even leaving aside the dynamic aspects that have been added).

In addition, CBO's construction of its baseline projections has become more complex when developments in the world have warranted it: The agency's most recent projections incorporate much more primary data on immigration and rigorous analysis of immigration's effects on the economy and budget than the agency's previous projections.

As CBO and JCT have adopted increasingly sophisticated approaches to producing both baseline projections and conventional budget estimates, they have developed effective approaches to explaining their work. Official estimates sometimes include substantial verbal descriptions, the agencies' leaders and staffs spend considerable time talking with Members and their staffs, and the agencies publish many separate reports documenting key methodologies. CBO's and JCT's openness about their methodologies through a stream of reports in recent years provides transparency that is important for professional feedback and for showing that estimating approaches are grounded in evidence and are not "arbitrary or politically motivated" (CBO, 1995).

Similar issues would arise with the additional behavioral responses that would be included in dynamic scoring. The agencies have published numerous descriptions of their approaches to modeling labor, capital, productivity, output, and other economic measures, as well as of the application of those approaches to specific policy changes. For CBO, the key models include a Solow-type growth model, a Keynesian multipliers model, and a small-scale policy model that combines the two preceding models; those models are described in CBO (2014b), Shackleton (2018), CBO (2021b), and Lasky (2022), among other references. For JCT,

the key models include a macroeconomic equilibrium growth model, an overlapping generations model, and a dynamic stochastic general equilibrium model; those models are described in JCT (2015), Auerbach et al (2017), JCT (2018a), JCT (2020), Moore and Pecoraro (2020), and Moore and Pecoraro (2023), among other references.

For future explanations of specific dynamic estimates, we can envision three types:

- Sometimes the additional responses would make little difference to the estimated budgetary impact and would not be of independent interest to policymakers. As a result, little explanation of the additional responses would be needed.
- In other cases, the additional responses would matter consequentially for the budgetary impact or would matter to policymakers for other reasons, and those responses would follow some standard patterns—for example, an increase in deficits would raise interest rates and thus debt service on existing debt. These responses could be explained primarily by reference to published material rather than through new descriptions.
- And in still other cases, the additional responses would matter and be somewhat idiosyncratic. In those cases, a bespoke and perhaps complex explanation might be required. The time need to explain and understand those expected effects would be among the tradeoffs to consider when evaluating whether to include them.

II. When?

In this section we address questions related to "when" dynamic scoring—or intermediate alternatives to conventional scoring—might be deployed. The current House rule, which specifies dynamic scoring for bills with conventionally estimated budget effects of a certain size (or when requested by the chair of a key committee), has advantages and disadvantages relative to alternatives such as designating certain policy topics for dynamic scoring or adopting an intermediate alternative. Special considerations arise when considering appropriation bills. In addition, investments in modeling and experience in dynamic scoring would have different value under different conditions. Moreover, all the foregoing considerations might evolve over time, which highlights the value of regularly reassessing these issues.

II.A. How Do Alternatives for Specifying the Deployment of Dynamic Scoring Compare?

Using dynamic scoring for proposals whose budget effects exceed a certain threshold has the advantage of focusing CBO and JCT's efforts on proposals that generally would have larger effects on the economy as well as the budget. Also, the trigger for deploying dynamic scoring is transparent, and lower or higher thresholds could be chosen to trigger dynamic scoring more or less often. On the other hand, when dynamic scoring is required for a bill with many disparate components, such as some budget reconciliation bills, evaluating the currently excluded effects of all components on a comparable basis may be infeasible owing to a lack of prior investment by the agencies (because the policy areas that might arise in proposals exceeding thresholds are difficult to predict).

Focusing dynamic scoring on certain policy areas would have the advantage of comparability of estimates for alternative policies that Congress might consider to meet a desired goal. Such comparability would be reinforced by applying dynamic scoring to both spending policies and tax policies in those chosen areas, such as symmetric treatment of federally funded R&D and tax provisions affecting the after-tax price of private R&D. Moreover, on the one hand, CBO and JCT could focus their capacity-building investments on the chosen areas, which could enhance the quality and timeliness of subsequent estimates. On the other hand, applying dynamic scoring to certain policy areas would be complicated to implement and to explain when legislation combines proposals from different areas. In addition, focusing on certain policy areas would have the disadvantage that deciding whether a proposal falls within a selected area would not be straightforward in all cases; in other contexts in which classification issues arise (such as decisions about whether to apply the scoring rules of the Federal Credit Reform Act), the agencies and Congressional staffers have used written guidelines, frequent communication, and experience in making classifications.

A further alternative in some cases would be to deploy approaches that are intermediate between conventional and dynamic scoring. The following examples are instructive.

One example that we discuss later in the paper is a "population-change" approach to scoring that CBO has used in a few estimates of proposed changes in immigration policy. That approach adds to a conventional estimate the direct budget effects of changing the number of people in the country (through the effects on compensation and tax revenue) but not all the effects that likely would occur. The resulting estimates are less comprehensive than dynamic

estimates, but they are more tractable and easier to explain because they do not include further macroeconomic complexities.

Another example is including specific types of additional behavioral responses but not all the responses that likely would occur. We noted earlier that changes in benefit programs might induce changes in both take-up and labor supply and that only the former are captured in conventional estimates. What if budget estimates included expected labor-supply responses but not other responses that are currently excluded? For proposals that would have direct effects on labor and either no direct effects on capital or direct effects on capital with the same sign as the direct effects on labor, adding only labor supply effects to budget estimates would tend to improve accuracy relative to conventional estimates. Changes in benefit programs would seem to meet this criterion. However, for proposals that would have direct effects on labor and direct effects on capital of the opposite sign, then adding only labor supply effects to budget estimates could reduce accuracy relative to conventional estimates. Shifts in the tax burden between labor and capital would seem to fit this criterion. Therefore, adopting an intermediate alternative of this sort would need to be done very carefully.

If Congress was interested in more dynamic scoring through any of these mechanisms, CBO and JCT might not be able to fulfill every request.[5] Many developments in the world and in Congress cannot be foreseen, and the current Congressional rule directing dynamic scoring for major legislation sensibly includes the phrase "if practicable."

II.B. What Issues Arise with Appropriation Bills?

CBO produces budget estimates for appropriation bills—which determine annual funding and traditionally have covered most federal spending on investments like R&D and infrastructure—that focus on how quickly available funding would be spent (see CBO, 2023c, page 9). Those estimates do not include behavioral responses of the sort included in other budget

[5] In writing "if Congress was interested in more dynamic scoring," we are skipping over the notable complication that Congress rarely speaks with one voice. For CBO and JCT, the committees that are most focused on the agencies' methodologies and allocation of effort are the House and Senate Budget Committees and the House Ways and Means Committee and Senate Finance Committee (whose chairs rotate as chairs of the Joint Committee on Taxation). But other committees that rely on CBO's and JCT's estimates may have their own views; resource levels for the agencies depend on the legislative branch subcommittees of the House and Senate Appropriations Committees; and the leadership of the House and Senate could be involved as well. In addition, of course, there can be partisan disagreements within each of these groups.

estimates. In addition, certain sorts of feedback from appropriations to other parts of the federal budget are excluded from estimates by scorekeeping guidelines (as noted earlier) and would complicate existing Congressional processes.

Gullo et al (forthcoming) comment on some of these issues in the context of R&D spending, and our discussant Bill Hoagland plans to tackle these issues in his remarks. Here we will only note that many similar procedural issues have been addressed over time. For example, CBO sometimes has reported as a memo item but not added into the basic tabulations the estimated rise in tax collections from increased appropriations for enforcement, and scorekeeping guidelines have been changed on occasion to make the information provided in estimates more useful to Congress. Analogously, the application of dynamic scoring to appropriation bills might take the form of memo items attached to basic tabulations that continue to exclude such effects, or it might engender changes in the scorekeeping guidelines.

II.C. What Factors Affect the Value of Investments in Dynamic Scoring Capacity?

In our assessment, the value of investment in modeling and practice for dynamic scoring depends principally on the degree of interest from policymakers in considering policy changes; the existence of a substantial research base about the impact of policies on labor, capital, and productivity; and the straightforwardness of potential approaches to modeling the economic effects of policy alternatives.

Focusing on topics of broad and enduring interest from policymakers generates value because there are fixed costs of developing the tools and experience to conduct dynamic analysis on a policy topic. Given the limitations on CBO's and JCT's resources, paying those costs is most worthwhile in cases where the long-term payoff is greatest. Past investments in modeling health proposals and the effects of fiscal policy, for example, are in areas of sustained interest.

A substantial research base regarding effects that are excluded in conventional scoring is important because the credibility of estimates is enhanced when the agencies can point to clear evidence underlying their analytic approaches.

Straightforwardness of modeling policy alternatives is valuable because the work and time required for modeling varies greatly across topics. Dynamic estimates depend on both general macroeconomic modeling (for example, how changes in labor supply affect capital investment and interest rates) and modeling of specific policy provisions (for example, how

various policy levers affect labor supply). CBO's and JCT's general macroeconomic modeling approaches are complex but, once developed, can be applied consistently to many different types of policies. By contrast, the agencies' modeling of specific policy provisions may be bespoke and can vary greatly in complexity.

Therefore, the feasibility of dynamic scoring for a policy topic depends crucially on the intricacy of the leading policy alternatives. For topics where alternatives are denominated in standardized units such as dollars or numbers of people and scale roughly linearly over the relevant range, estimates that have been made for some proposals might be scalable through simplified spreadsheets and rules of thumb to produce credible estimates for other proposals. Examples include federally funded research investments and immigration of higher-skilled people, and we return to those topics later.

For other topics where the policy space is more multidimensional because alternatives tend to involve different structures of policy, building simplified but still credible models is much more challenging. For example, in health care and health insurance, the diversity and complexity of proposals makes conventional scoring quite difficult in many cases and dynamic scoring even more difficult. CBO and JCT have produced dynamic analyses of some health proposals—including an increase in the tax on cigarettes (CBO, 2012b) and an expansion of Medicaid eligibility (Ash et al, 2023)—and provided a dynamic cost estimate for repeal of the Affordable Care Act (CBO, 2015b). However, the existence of those analyses does not imply that the agencies could provide comparable analyses for many other health proposals; indeed, the analysis of ACA repeal quantified some influences on labor supply and saving but did not quantify others. Over time, as the evidence base in the health area continues to expand and if Congress supports further improvement in the agencies' toolkits, CBO and JCT may be able to estimate the effects of more policies on improving people's health and enabling a more productive workforce.

Another challenge in building a simplified and credible model involves capturing effects of policies on labor, capital, productivity, output, and other economic measures that occur far in the future. In select analyses, CBO has provided numerical estimates or qualitative information on long-term effects (beyond the 10 years captured in the usual budget window) and their feedback to the federal budget. For example, see the analyses of comprehensive immigration reform (CBO, 2013a and 2013b) and potential changes to Medicaid (Ash et al, 2023).

II.D. How Might the Foregoing Considerations Evolve over Time?

Advances in economic research and in CBO's and JCT's capabilities might affect the advantages and disadvantages of the different approaches discussed in the preceding sub-sections. To improve both conventional and dynamic estimates, CBO and JCT could regularly publish "calls for research" on topics where a scarcity of evidence particularly hinders the agencies' work. One example is see CBO's (2024c) article in the *Journal of Economic Perspectives*, which was solicited by one of the authors of the present paper (Williams) in her role as *JEP* editor. In addition, there probably will be considerable "learning by doing" on the part of the agencies, with positive spillovers among investments in different aspects of dynamic analysis. Therefore, it could be useful for the strengths and weaknesses of different approaches for budget estimates to be reviewed regularly, and decisions about the best feasible approach might change over time.

Whether dynamic scoring is adopted more widely or not, we expect that CBO and JCT will continue to provide dynamic analyses apart from official budget estimates on at least some occasions. That information about the effects of policies on labor, capital, productivity, output, and other economic measures is useful for policymakers even if it is not included in official estimates. Moreover, once an agency has completed a dynamic analysis in a policy area, the benefits and costs of including such analysis in budget estimates can be ascertained more readily.

III. How?

In this section, we draw on CBO's past dynamic analyses to illustrate a process for dynamic scoring. JCT's past dynamic analyses have been constructed somewhat differently; we do not attempt to describe JCT's process fully (see the references cited earlier), but we discuss one particularly important difference from CBO's process below.

III.A. Estimating direct effects of a proposal on labor, capital, and productivity

The starting point for a dynamic cost estimate is to assess the direct effects of a proposal on labor supply, saving and investment, and total factor productivity, leaving aside for the moment the financing of the policy and any indirect effects of that financing. These effects can be disaggregated into categories of labor with different marginal products and categories of

capital with different gross marginal products and depreciation rates. The effects depend on responses by households, businesses, implementers of legislation in the executive branch, and state and local governments; for example, an increase in federal highway funding could spur reinforcing or offsetting changes in state highway funding. In addition, the effects are estimated with a time dimension: Conventional cost estimates recognize that federal funding affects outlays with a lag, and dynamic estimates should recognize also that federal outlays may have lagged effects on labor, capital, and productivity. Capturing all the conceivable effects of a proposal on these economic factors is not possible, any more than capturing all the conceivable effects of a proposal is possible in a conventional estimate; the challenge for analysts remains to identify and include the effects that seem most important and for which there is some evidence of likely sign and magnitude.

As discussed earlier, investments in modeling and practice in a policy area lower somewhat the marginal cost of future estimates because models and experience can be used repeatedly, in both conventional estimates and dynamic ones. But even when a model has been built and tested, an analyst still needs to translate a proposal's legislative language into model inputs and then closely review model outputs—and sometimes undertake this process many times as Congress iterates on alternative versions of proposals. Therefore, producing dynamic estimates still will be substantially more time-consuming than producing conventional estimates in almost all cases.

III.B. Estimating changes in output and other economic factors

With the estimated effects of a proposal on labor, capital, and productivity in hand, an analyst can estimate the follow-on effects of the proposal on potential output, actual output, interest rates, inflation, and other relevant economic factors. We address four issues: effects on aggregate demand, the impact of different methods of financing, the models used, and the challenge of aggregation across different elements of a proposal.

First, a proposal can cause actual output to deviate from potential output through its effects on aggregate demand. Some proposals might change aggregate demand roughly in line with potential output, and in cases when that does not occur, the Federal Reserve could be expected to adjust monetary policy to bring aggregate demand into balance with potential output. However, changes in government spending or revenue can affect output before monetary policy

offsets those changes, and shifts in monetary policy affect interest rates, which in turn affect government debt service. Moreover, some proposals are *designed* to shift aggregate demand relative to potential output, such as fiscal stimulus when economic activity is weak, and changes in aggregate demand relative to potential output can affect inflation. Dynamic estimates include shifts in aggregate demand to capture these economic impacts.

Second, including the effects of financing is crucial to generating a useful assessment of a policy's economic impact. For example, consider an increase in government investment financed by an increase in borrowing. Taking account of the positive impact on growth of the additional government investment without taking account of the negative impact of crowding out some private investment would be misleading. And including the positive impact on the budget of faster growth without including the negative impact of higher interest rates on debt service also would be misleading—especially because the current high level of debt means that changes in interest rates can have very large effects on the government's fiscal position.[6]

For an illustration of both these points, consider CBO's (2021c) analysis of a potential increase in infrastructure spending. In the first scenario considered, additional infrastructure spending is offset by a reduction in other government spending. CBO estimates little imprint on output in the short run, a growing positive effect on output over time, and little effect on interest rates; overall, the economic effects of the policy reduce the budgetary cost by about one-third. In the second scenario, additional infrastructure spending is financed by borrowing. CBO estimates a boost in output for a while but little effect in the long run, because additional borrowing crowds out some private investment, and higher interest rates in both the short and long runs; overall, the economic effects of the policy *increase* the budgetary cost by about one-fourth. These patterns—varying between financing methods and over time for either method—are important for policymakers to understand.

[6] Presumably the change in borrowing would be assumed to occur with a maturity structure that matched that of outstanding government debt. Some economic models imply that government budget deficits can be self-financing in certain circumstances so that the additional debt imposes no burden. But then the whole purpose of budget estimates for policy proposals should be reconsidered, and that effort is well beyond the scope of this paper; our maintained hypothesis is that deficits are not self-financing. Separately, in models in which self-financing does not occur, long-term analyses of government borrowing face the question of how long debt can increase before spending needs to be cut or taxes increased to avoid economic instability—and assuming any such future adjustment runs against the prohibition against CBO and JCT predicting policy changes. However, even with the current high level of government debt, this question can be avoided or answered in a fairly neutral manner over the next few decades for most proposals.

Regarding the financing of proposed policies, we note that, by longstanding convention, CBO's cost estimates do not include changes in debt service from changes in government borrowing: A proposal that would raise spending by $100 billion in the first year of the budget window and leave spending unchanged in subsequent years and taxes unchanged in every year would be reported to increase deficits over the budget window by a total of $100 billion, not $100 billion plus the interest costs that would be incurred later.[7] However, CBO's dynamic cost estimates *would* include the additional debt service that arises from changes in interest rates applied to the existing debt, which is what we are describing here; for example, see CBO (2016, footnote 19) and CBO (2021c, footnote 5). Thus, a dynamic estimate of the hypothesized proposal would include the additional debt service stemming from higher interest rates applied to currently outstanding debt, but it would not include debt service on the $100 billion itself.

Third, CBO and JCT have constructed a variety of models for estimating the effects of proposed policies on output and other economic factors, as described earlier. These models are aligned with CBO's baseline economic projections and are updated at least annually as those projections evolve. Economic conditions affect fiscal-policy multipliers (in part by influencing the Federal Reserve's reactions to fiscal-policy changes) and other impacts of legislative proposals, so reliable estimates depend on updating the models regularly and analyzing proposals in light of expected conditions. But the need for regular model updates and sensitivity to evolving external circumstances are hardly distinctive to dynamic estimates: All the agencies' models are updated regularly to reflect CBO's baseline projections and other new information, and budget estimates take those factors on board.

Fourth, aggregating different elements of a legislative proposal represents a significant challenge in estimating the broader economic effects. Bills often are constructed in pieces, and multiple analysts at CBO and JCT sometimes work on the estimate for a single bill.

For the spending bills on which CBO focuses, the pieces usually do not interact in significant ways, in which case piece-by-piece estimates can be constructed in a modular way and then added together for an overall conventional estimate. Then, suppose that analysts estimated for each piece not only the conventional effects but also the direct effects on labor, capital, and productivity, the resulting change in government borrowing, and then the follow-on effects on other economic measures. This approach can be applied to each element of a bill and

[7] CBO (2024d) addressed the possibility of changing this convention.

would aggregate correctly: For example, if a bill paid for extra spending for a chosen purpose by cutting other spending, then the sum across the bill's provisions of all the financing impacts would be zero, and the sum of the impacts of those provisions on labor, capital, productivity, and other measures would match the effects of estimating the impact of the bill as a package.

For the revenue bills on which JCT focuses, the interactions that often occur between provisions means that piece-by-piece estimation generally is not possible for either conventional or dynamic estimates (and sometimes revenue provisions interact with spending provisions). Alan Auerbach helpfully reminded us that the "stacking order" for tax provisions can affect the estimated impact of each provision, in both conventional and dynamic estimates. For example, if a proposed reduction in tax rates is combined with a strengthened minimum tax, the rate cut will cost more revenue and have larger dynamic effects if it is evaluated first than if it is evaluated second. As a result, conventional estimates of bills that change the tax code usually need to be produced holistically, and dynamic estimates are the same in that regard.

III.C. Estimating changes in the budget

Given the estimated effects of a proposed policy on labor, capital, output, and interest rates, the impact on the government budget can be estimated using the approaches noted earlier.

III.D. Reporting the estimates

The first report of a budget estimate is often a private conversation between CBO or JCT analysts and staff members for a Congressional committee that is developing a proposal. Because that estimate may influence the committee's further work, committee staff members want to receive a close approximation to what the agencies would ultimately publish if the proposal was approved by the committee.

Those initial conversations pose an additional challenge to dynamic scoring. If those conversations do not include the dynamic effects of a proposal, and those effects turn out to matter consequentially for the ultimate budget estimate, then the iterative process of developing the proposal will have been based on inadequate information. But to include consequential dynamic effects in initial conversations would require analysts at the estimating agencies to have a sense of those effects—which could be difficult to develop on the requisite timetable. This observation again highlights the value of simplified spreadsheets and rules of thumb in cases

where such simplifications are tractable and credible, as such approximations would allow more information on dynamic effects to be provided earlier in these conversations.

In terms of published estimates, the estimating agencies have—for transparency—shown the dynamic effects and conventional effects separately and then together. That approach has been taken by JCT for some time, and it was used in CBO's (2021) dynamic analysis of infrastructure spending (see Table 3 on page 6).

Because Congress's budget process and procedural rules rely on point estimates, CBO and JCT focus on point estimates. In addition, because estimates are not usually derived from probabilistic models, prediction intervals do not flow naturally out of the estimating process. But the agencies do develop and report intervals on occasion, especially for dynamic analyses, and that information can be useful for policymakers; for example, see CBO's (2013b) report on the economic effects of comprehensive immigration reform.

IV. An Illustrative Example: Potential Changes in Immigration Policy[8]

Immigration policy occupies an important place in many discussions of economic policy, ranging from specific rules about green cards for STEM (science, technology, engineering, and math) workers to the general issue of whether and how immigration policy should respond to slower growth of the domestic-born workforce. Under current practice, budget estimates of proposals to change immigration policy generally capture only the expected effects on federal spending and not the expected effects on federal revenue. In this section, we contrast that conventional scoring to dynamic scoring and to an intermediate alternative of "population-change scoring."

IV.A. Conventional Scoring

Consider a proposal that would increase the number of immigrants. In the conventional estimating approach of holding nominal GDP fixed, the additional immigrants cannot be modeled as earning income unless some existing residents are modeled as earning comparably less income than otherwise; because such a drop in income for existing residents is not consistent with the evidence, the estimating agencies simply do not include any additional income or

[8] This section draws in part on work by Elmendorf and Williams (2024), which was done in collaboration with Alex Arnon and Kent Smetters from the Penn Wharton Budget Model.

payroll tax revenue. However, the additional immigrants can be modeled as receiving federal benefits for which they are legally eligible, such as federal subsidies for health insurance under the Affordable Care Act; therefore, the estimating agencies do include additional federal spending. This asymmetry in handling expected changes in revenue and spending means that essentially any proposal that would increase the number of immigrants would be estimated to increase federal deficits.

The conventional approach has been applied to most immigration proposals being considered by Congress, and the resulting estimates have not provided a balanced picture of the likely impact of those proposals on the federal budget. Consider Section 80303 of H.R. 4521, which was a legislative provision in the America COMPETES Act that would have increased the availability of green cards for holders of advanced STEM degrees. Economic research has shown that the federal income and payroll tax revenue generated by such immigrants would be much greater than the spending for federal benefits they would receive. For example, a 2017 *National Academies* report concluded that, on average, high-skilled immigrants contribute hundreds of thousands of dollars more in federal tax revenue than they receive in federal benefits over their working lives. Yet, CBO (2022a) reported that this legislative provision would *cost* the federal government $3.1 billion over a decade.

Note that conventional scoring of immigration proposals relies on detailed modeling of how those proposals would affect the number and characteristics of people entering the country as well as the number and characteristics of domestic residents who would change their legal status, which could in turn affect emigration, eligibility for federal benefits, and perhaps a broader set of outcomes like job changes and earnings.[9] The results of this population modeling can be visualized as a table showing changes in the numbers of people with different combinations of attributes. Constructing a table of this sort is complicated and time-consuming—and the table offers much of the information needed to go beyond conventional scoring, as we explain shortly.

[9] As an example, one of us (Williams) worked with Matthew Esche and Jeremy Neufeld (2024) on population modeling for a representative policy in the spirit of Section 80303 and other recent proposals such as S. 2384 (the Keep STEM Talent Act of 2023). That representative policy would exempt from statutory limits employment-based green cards (EB-1, EB-2, and EB-3) for applicants who have earned a doctoral or master's degree in STEM at a U.S. research institution or foreign equivalent, as well as for an accompanying spouse and minor children. The code and documentation for that population modeling are publicly available here.

IV.B. Population-Change Scoring

The estimating agencies do not always use conventional scoring for immigration proposals. CBO (2013b, page 2) explained why and how the agency adopted a different approach in estimating the effects of S. 744, a proposal for comprehensive immigration reform: "[F]ollowing the standard convention of assuming that employment would remain unchanged relative to current law would have implied 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, which primarily affected the amount of additional tax revenues that would have resulted from enacting the bill."

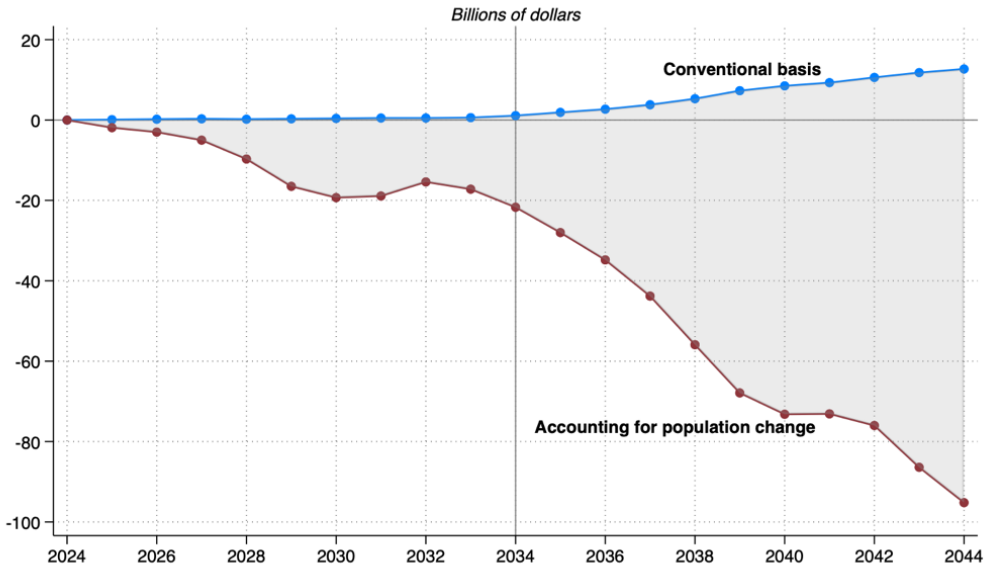
This population-change approach incorporates all the elements of a conventional estimate as well as the direct budgetary effects of changing the number of people in the country—in particular, the effects on taxable compensation and therefore on income and payroll tax revenue. Population-change estimates do not incorporate all the effects that CBO and JCT expect would occur, as would be captured by a fully dynamic approach; we return to this issue in the next sub-section. To our knowledge, CBO has applied the population-change approach to four legislative proposals over time: S. 2611 (CBO, 2006), Senate Amendment 1150 to S. 1348 (CBO, 2007), S. 744 (CBO, 2013a), and H.R. 2131 (CBO, 2014a). The first three of these proposals represented comprehensive immigration reform, and the fourth aimed to increase the number of visas available to high-skilled noncitizens.

Two examples illustrate how population-change estimates for immigration proposals can differ markedly from conventional estimates. First, consider that both H.R. 2131 and Section 80303 of H.R. 4521 would have increased the availability of green cards to holders of advanced STEM degrees. CBO's conventional estimate for Section 80303 showed an increase in budget deficits of about $3 billion from 2022 to 2031, whereas CBO's population-change estimate for H.R. 2131 (which included the expected increase in population, employment, and taxable compensation) showed a decrease in budget deficits of about $110 billion from 2014 to 2024.

Second, Elmendorf and Williams (2024) compared conventional and population-change approaches for a Section 80303-style proposal. Figure 2 summarizes their estimates of the effects

on budget deficits over the next 20 years. The conventional estimate (the blue line) shows a slight increase in deficits, while the population-change approach (the red line) shows a decrease in deficits in the first decade and a much larger decrease in the second decade as lawful permanent residents naturalize and sponsor family members (with an assumed nine-year lag, following Carr and Tienda, 2013).

Figure 2. Estimated effects on deficits of an increase in green cards to holders of advanced STEM degrees, conventional and population-change approaches



Note: Reprinted from Figure 1 in Elmendorf and Williams (2024), reflecting estimates produced by the Penn Wharton Budget Model.

A natural follow-up question is how much conventional estimates and population-change estimates would differ for immigration proposals that focused not on high-skilled immigrants but instead on less-skilled immigrants, whose taxable compensation probably would be lower on average and whose take-up of federal benefits may be higher on average.

CBO's recent update to its baseline economic and budget projections provides a partial answer. CBO (2024f) documented in detail the impact on the budget of the surge in immigration during 2022 and 2023 and the expected partial continuation of that surge for a few more years. The agency explained that the surge has comprised people who are probably considerably less educated, on average, than the people who would have been affected by a STEM-focused

legislative proposal.[10] CBO expected that about half of this group will receive authorization to work; some of the other half may work as well, but immigrants who work without authorization are less likely to pay taxes. Even so, CBO estimated that, over the next decade, the surge will increase federal revenue by roughly $1.2 trillion (mostly from additional income and payroll taxes) and increase federal spending by roughly $0.3 trillion (mostly from additional federal benefits), for a net reduction in deficits of roughly $0.9 trillion. These figures suggest that a legislative proposal that changed the number of less-educated immigrants might receive starkly different estimates under population-change scoring than under conventional scoring.

This brings us to a further question about the feasibility of producing population-change estimates—of estimating the effects of immigration proposals on revenue as well as spending—in a reasonable timeframe. One possibility is for the agencies to apply a standard microsimulation model to the estimated changes in numbers of people with different attributes. Because the attributes needed to predict employment and compensation are similar to the attributes needed to predict eligibility for benefits, the population modeling that is used to estimate spending changes would facilitate estimates of revenue changes. Still, this process would be time-consuming.

An alternative way to estimate revenue effects is to calibrate a microsimulation model each time that baseline projections are updated, apply the results to create a "calculator" with estimated changes in revenue per person for each combination of attributes in the population modeling, and then use that calculator for all legislative proposals until the baseline projections are updated again. Two of us (Elmendorf and Williams) are exploring this method in joint work with Theresa Gullo. This method would allow for much shorter turnaround times on Congressional requests relative to running a microsimulation model for each request.

IV.C. Dynamic Scoring

Population-change estimates capture more effects of immigration proposals than do conventional estimates, but they do not capture all the effects included in dynamic estimates. The dynamic approach mirrors CBO's handling of baseline projections by incorporating the

[10] CBO attributed the surge in immigration to people who are entering the country without inspection, who were allowed to enter the country lawfully through parole authority and are awaiting proceedings in immigration court, or who previously resided in the country under a temporary legal status and remained after that status expired.

budgetary impact of potential changes in capital investment, interest rates, productivity, and more. CBO has produced dynamic analyses of two legislative proposals on immigration—S. 2611 (CBO, 2006, appendix) and S.744 (CBO, 2013b)—although those dynamic analyses were separate from the official budget estimates for those bills. CBO (2024f) used an analogous approach to update the agency's baseline projections for the recent immigration surge.

Space constraints do not allow us to discuss all the economic effects that might arise in dynamic estimates of immigration proposals, so for illustrative purposes, we comment on just one effect featured in CBO's (2024f) new baseline projections—the impact of immigration on productivity.

As background, note that a number of papers present compelling evidence that when high-skilled immigrants move to the United States, their research becomes more productive (Prato, 2022), they generate valuable innovations at disproportionately high rates (Bernstein et al, 2023), and they start new firms that lead them to act more as "job creators" than "job takers" of domestic residents (Azoulay et al, 2022). Few papers link these innovation-related outcomes directly to productivity, although Peri (2012) and Prato (2022) are two exceptions.

CBO estimated that two percent of people in the U.S. immigration surge are high-skilled STEM workers. That is a small percentage but, given the size of the surge, a large absolute number—several hundred thousand people, which exceeds the statutory annual cap on H-1B visas. CBO estimated that these individuals will contribute to innovation-related activities that will raise total factor productivity roughly 0.2 percent by 2034 relative to what it would have been otherwise, boosting output and income notably in the following decades.

Constructing a dynamic estimate of an immigration proposal that incorporates all the potential economic and budgetary effects requires considerable modeling. Hence, a natural question is what share of the overall budgetary impact of a proposal can be captured by a simpler population-change estimate. The answer would vary across legislative proposals, but one datapoint is available: CBO estimated that the recent immigration surge will reduce deficits over the next decade by roughly $600 billion through only the change in population and by an additional nearly $300 billion when incorporating the full set of dynamic effects.

V. Illustrative Examples: Potential Changes in Federal Investment and Federal Permitting of Investment

Federal investments in infrastructure, education, and R&D are made with the intention of spurring economic growth by increasing physical capital, human capital, and productivity. CBO has analyzed federal investment on several occasions but has never included the economic impact of such investment in official cost estimates. In this section, we examine dynamic scoring for federal investment and federal permitting of investment.

V.A. CBO's 2021 Report on Physical Infrastructure Spending

CBO has analyzed the benefits and costs of federal investments for some time (for example, see CBO, 2013c and 2016). The agency made a substantial further advance in its 2021 report on "Effects of Physical Infrastructure Spending on the Economy and the Budget Under Two Illustrative Scenarios." The report was prepared in response to a request from then-Senator Rob Portman, who wanted to bring a more comprehensive analysis to bear in developing the bipartisan Infrastructure Investment and Jobs Act (Bolton, 2021).

The 2021 report essentially describes how CBO could produce a dynamic cost estimate of a proposal to change federal infrastructure spending if Congress requested it. The report is exemplary in its rigorous analysis and methodological transparency. Moreover, the modeling is presented in a flexible and modular fashion, so it can be applied readily to a variety of specific proposals for changing federal investment. Specifically, the report focuses on five key factors:

- *"How state and local governments respond to additional federal funding":* For the infrastructure policy in the report, additional federal spending is expected to generate a partially offsetting reduction in spending by state and local governments. For other federal investments, the expected response of state and local governments to a federal policy change could be different; moreover, other federal investments might spur an offsetting or reinforcing response by private actors.
- *"How quickly funding leads to outlays":* Federal spending generally lags the authorization to undertake spending, to a degree that varies across types of spending. This factor is a standard aspect of CBO's conventional estimates, so including it in dynamic estimates would be straightforward.
- *"How quickly outlays increase productivity":* CBO expects that additional federal infrastructure spending would increase productivity with a lag, with 40 percent of the effect occurring in the first year, 80 percent occurring by the second year, and 100

percent occurring by the seventh year. For other federal investments or for specific types of infrastructure spending, different lagged patterns might be appropriate and could be drawn from the research literature.

- *"How much outlays increase productivity":* CBO expects that federal infrastructure has a gross annual return of 12.4 percent and depreciates by 3.2 percent annually, implying a net annual return of 9.2 percent. Again, different figures could be appropriate for different federal investments.
- *"How outlays are financed":* As we described earlier, CBO examines two scenarios—one in which the additional infrastructure spending is financed by cuts in other spending, and one in which the additional spending is debt-financed. With these two scenarios modeled, CBO could presumably model any intermediate financing approach as well.

In sum, the modeling approach used and described by CBO (2021) can be applied consistently to any federal investment for which there is a reasonable basis for determining the five factors that CBO highlighted. Among the factors, the most challenging to ascertain are the third and fourth regarding the effects of federal investment on productivity, and we turn to those issues in the context of R&D spending in the next section. More generally, the evidence base on the economic effects of federal investments will improve over time, and as it does, the feasibility and value of dynamic estimates will improve as well.

V.B. Application to Potential Changes in Federal Spending for Research and Development

For federal spending on R&D, the research literature seems sufficiently developed that CBO's 2021 framework for analyzing infrastructure spending could be applied in a relatively straightforward way. One of us (Williams) is working with Theresa Gullo, Benjamin Page, and David Weiner (Gullo et al, forthcoming) to illustrate how that type of analysis could be done, so we refer interested readers to that paper and summarize the main points here.

For decades, economists have explored the effect of investments in research and innovation on economic growth; for example, see Romer (1990), Aghion and Howitt (1992), and Grossman and Helpman (1991). However, empirical evidence on the relationship between R&D and productivity growth has come slowly. Jones and Summers (2022) applied a somewhat indirect macroeconomic approach to infer that the marginal social return to R&D tends to be high (see their Table 7). Azoulay et al (2019) and Myers and Lanahan (2022) rigorously

estimated the spillovers from R&D funded by the National Institutes of Health (NIH) and Department of Energy, respectively, and found that the spillovers are substantial within a decade; however, neither of those papers quantified the effects of R&D on productivity and growth.

Fieldhouse and Mertens (2023) and Dyevre (2023) offered a key step forward, essentially stitching together the microeconomic evidence on returns to R&D with macroeconomic aggregates. Using quite different empirical approaches, those papers arrived at broadly similar quantitative estimates of the contribution of federal R&D to productivity. For example, Fieldhouse and Mertens used a narrative approach to classify postwar changes in appropriations for R&D at five major federal agencies and then applied those data to structurally estimate economic returns to those appropriations. They concluded that the returns are between 150 and 300 percent, which implies that federal R&D is responsible for around one-quarter of total postwar growth of total factor productivity in the United States.[11]

Gullo et al (forthcoming) discuss how this type of evidence can be applied in CBO's 2021 framework to analyze the economic and budgetary impact of changes in federal spending for R&D. For example, referring to some of the key factors identified by CBO: The research literature generally suggests that federally funded R&D crowds in private R&D, Fieldhouse and Mertens' estimated impulse response functions (their Figure 6) provide a basis for the time lags between R&D and productivity, and the Bureau of Economic Analysis has done extensive work on depreciation rates for R&D.

Much of federal R&D investment can be denominated readily in dollars and probably has effects that scale roughly linearly over relevant ranges. Therefore, modeling the effects of changing federal R&D in simplified ways could produce informative estimates. For example, one could imagine a spreadsheet with estimated effects on the economy and federal budget of each dollar of additional funding through different agencies (say, the National Institutes of Health versus the National Aeronautics and Space Administration) and for different purposes (say, basic research versus applied research). A tool like this could be updated with each baseline and provide a tractable way to produce dynamic budget estimates on compressed timelines.

[11] The research literature distinguishes between defense and nondefense R&D. CBO (2018a) argued that most defense R&D does not provide significant spillovers to the private sector, although there are exceptions, and therefore defense R&D does not substantially influence the agency's economic analysis. Recent research generally supports that position. More broadly, CBO (2014c) explored the impact of federal policies on innovation.

Investments in piloting, validating, and improving approaches of this sort could also provide a template that could be applied over time to other policy areas.

V.C. Application to Potential Changes in Federal Permitting for Physical Infrastructure

CBO's modeling capacity for potential changes in federal permitting for physical infrastructure seems ready, building on the same 2021 framework. However, in our view, the available research base for informing the key parameters in that framework is too limited for credible dynamic analysis to be feasible now. The description by CBO (2024c) of this disconnect between the agency's modeling capability and the available research base may spur additional research. More generally, as the research literature in this area develops over time, dynamic scoring will become more feasible.

The National Environmental Policy Act (NEPA) requires that "major" infrastructure projects that are federally funded or federally subsidized receive licenses, or "permits," before construction can occur. As discussed by Liscow (2024), the goal of this provision is to force review of how proposed projects would affect the "environment," including both nature and people. Potter, Datta, and Stapp (2022) argued that the NEPA permitting process is hindering other national priorities such as the clean energy transition and therefore should be restructured.

Consider, then, a proposal to accelerate the NEPA permitting process. To produce a conventional cost estimate for that proposal, CBO would need to quantify the additional spending on infrastructure projects that would occur within the budget window and the budgetary impact of that additional spending. For example, CBO would need to assess the extent to which organizations that opposed projects could shift their attention to procedures run by state, local, or tribal governments and prevent an acceleration of permitting in that way; unfortunately, we are aware of no empirical literature to inform such an assessment. In addition, CBO would need to assess the elasticity of projects with respect to changes in time to build, but again, there is little relevant evidence; in the words of CBO's (2023a) cost estimate for H.R. 1, The Lower Energy Costs Act: "CBO has insufficient information to determine the number of projects … that could be generated."[12]

[12] The issue is straightforward on a theoretical level: Shortening permitting time is analogous to shortening commercialization lags for new pharmaceutical drugs, which one of us (Williams) estimated with co-authors to have increased private investment in drug development (Budish et al, 2015).

Producing a dynamic estimate for that proposal would be even more challenging. CBO would need to build on a conventional estimate by also considering the effects of additional infrastructure projects on productivity and other economic measures while accounting for the relevant sources of financing. NEPA governs both federally funded and federally subsidized investments, and different types of investment could have different effects on productivity. For example, the productivity effects of accelerating construction of semiconductor fabrication facilities could be very different from the effects of accelerating construction of highways, which has been the focus of most research to date.

VI. Conclusion

In recent years Congress has asked its official estimators—the Congressional Budget Office and the staff of the Joint Committee on Taxation—to provide more information about the economic effects of policy changes. The agencies have responded by drawing on an expanding body of research to produce innovative analyses, improve their modeling tools, gain experience, and publish reports regarding their methodologies and results. This work has reinforced the value of the agencies providing unbiased and evidence-based analyses of the effects of potential policies on labor, capital, productivity, and other economic outcomes.

Decisions by Congress about the extent to which such analysis should be included in official budget estimates—through dynamic scoring or intermediate alternatives to conventional scoring—involve tradeoffs regarding time, resources, accuracy, and comparability. With labor force growth slowing and federal debt rising, interest in the intersection of public policy, economic outcomes, and budget impacts shows no signs of diminishing.

Acknowledgements

For valuable conversations and comments, we are grateful to Alexander Arnon, Alan Auerbach, Theresa Gullo, William Hoagland, Douglas Holtz-Eakin, Deborah Lucas, Donald Marron, Dylan Matthews, Alexander Mechanick, Benjamin Page, Kent Smetters, Jón Steinsson, Caleb Watney, David Weiner, David Wessel, and staff of the Congressional Budget Office and the Joint Committee on Taxation. The case studies we discuss draw on collaborative work with Arnon, Gullo, Page, Smetters, Weiner, Matthew Esche, Zachary Liscow, and Jeremy Neufeld. We are solely responsible for the contents of the paper and that the views expressed here should

not be attributed to any of the individuals acknowledged or any institution with which the authors are currently or formerly affiliated.

References

Aghion, Philippe, and Peter Howitt. 1992. "A Model of Growth Through Creative Destruction." *Econometrica* 60, no. 2: 323-351. https://doi.org/10.2307/2951599.

Ash, Elizabeth, William Carrington, Rebecca Heller, and Grace Hwang. 2023. "Exploring the Effects of Medicaid During Childhood on the Economy and the Budget." Working paper, Congressional Budget Office 2023-07, November.

Auerbach, Alan J., Itai Grinberg, Thomas Barthold, Nicholas Bull, W. Gavin Elkins, Pamela Moomau, Rachel Moore, Benjamin Page, Brandon Pecoraro, and Kyle Pomerleau. 2017. "Macroeconomic Modeling of Tax Policy: A Comparison of Current Methodologies." *National Tax Journal* 70, no. 4: 819-836. https://doi.org/10.17310/ntj.2017.4.06.

Azoulay, Pierre, Joshua S. Graff Zivin, Danielle Li, and Bhaven N. Sampat. 2019. "Public R&D Investments and Private-Sector Patenting: Evidence from NIH Funding Rules." *Review of Economic Studies* 86, no. 1: 117–52, https://doi.org/10.1093/restud/rdy034.

Azoulay, Pierre, Benjamin F. Jones, J. Daniel Kim, and Javier Miranda. 2022. "Immigration and Entrepreneurship in the United States." *American Economic Review: Insights* 4, no. 1: 71-88. https://doi.org/10.1257/aeri.20200588.

Bernstein, Shai, Rebecca Diamond, Abhisit Jirananphawiboon, Timothy McQuade, and Beatriz Pousada. 2022. "The Contribution of High-Skilled Immigrants to Innovation in the United States." Working paper, National Bureau of Economic Research 30797, December. https://doi.org/10.3386/w30797.

Bolton, Alexander. 2021. "CBO says bipartisan infrastructure bill would add $256B to deficit over 10 years." *The Hill*, August 5.

Budish, Eric, Benjamin Roin, and Heidi Williams. 2015. "Do Firms Underinvest in Long-Term Research? Evidence from Cancer Clinical Trials." *American Economic Review* 105, no. 7: 2044-85. https://doi.org/10.1257/aer.20131176.

Carr, Stacie and Marta Tienda. 2013. "Family Sponsorship and Late-Age Immigration in Aging America: Revised and Expanded Estimates of Chained Migration." *Population Research Policy Review* 32, no. 6. https://doi.org/10.1007/s11113-013-9300-y.

Congressional Budget Office. 1995. “Budget Estimates: Current Practices and Alternative Approaches.” January.

_____. 2006. “Cost Estimate for S. 2611, Comprehensive Immigration Reform Act of 2006.” May.

_____. 2007. “Cost Estimate for Senate Amendment to S. 1348, the Comprehensive Immigration Reform Act of 2007.” June.

_____. 2009. “The Role of the 25 Percent Revenue Offset in Estimating the Budgetary Effects of Legislation.” January.

_____. 2012a. “Fair-Value Accounting for Federal Credit Programs.” March.

_____. 2012b. “Raising the Excise Tax on Cigarettes: Effects on Health and the Federal Budget.” June.

_____. 2012c. “How the Supply of Labor Responds to Changes in Fiscal Policy.” October.

_____. 2013a. “Cost Estimate for S. 744, the Border Security, Economic Opportunity, and Immigration Modernization Act.” June.

_____. 2013b. “The Economic Impact of S. 744, the Border Security, Economic Opportunity, and Immigration Modernization Act.” June.

_____. 2013c. “Federal Investment.” December.

_____. 2014a. “Cost Estimate for H.R. 2131, Supplying Knowledge-based Immigrants and Lifting Levels of STEM Visas Act (SKILLS Visa Act).” March.

_____. 2014b. “How CBO Analyzes the Effects of Changes in Federal Fiscal Policies on the Economy.” November.

_____. 2014c. “Federal Policies and Innovation.” November.

_____. 2015a. “The Effects of Potential Cuts in SNAP Spending on Households With Different Amounts of Income.” March.

_____. 2015b. “Budgetary and Economic Effects of Repealing the Affordable Care Act.” June.

_____. 2015c. “Answers to Questions About Dynamic Analysis.” July.

_____. 2015d. “Cost Estimate for Tax Relief Extension Act of 2015.” August.

_____. 2015e. “Cost Estimate for A Bill to Amend the Internal Revenue Code of 1986 to Modify and Make Permanent Bonus Depreciation.” October.

_____. 2016. “The Macroeconomic and Budgetary Effects of Federal Investment.” June.

_____. 2018a. “Estimating the Long-Term Effects of Federal R&D Spending: CBO’s Current Approach and Research Needs.” June.

_____. 2018b. “How CBO Produces Fair-Value Estimates of the Cost of Federal Credit Programs: A Primer.” July.

_____. 2019. “The 2019 Long-Term Budget Outlook.” June.

_____. 2020. “The Effects of Pandemic-Related Legislation on Output.” September.

_____. 2021a. “CBO Explains Budget Scorekeeping Guidelines. January.

_____. 2021b. “CBO’s Policy Growth Model.” April.

_____. 2021c. “Effects of Physical Infrastructure Spending on the Economy Under Two Illustrative Scenarios.” August.

_____. 2022a. “Cost Estimate for H.R. 4521, the America COMPETES Act of 2022.” March.

_____. 2023a. “Estimated Budgetary Effects of Divisions A and B of H.R. 1, the Lower Energy Costs Act.” March.

_____. 2023b. “Workbook for How Changes in Economic Conditions Might Affect the Federal Budget: 2023 to 2033.” March.

_____. 2023c. “CBO Describes Its Cost-Estimating Process.” April.

_____. 2023d. “CBO’s Economic Forecasting Record: 2023 Update.” June.

_____. 2023e. “Cost Estimate for S. 2226, the National Defense Authorization Act for Fiscal Year 2024.” October.

_____. 2023f. “The Budgetary and Economic Effects of S. 2488, the Raise the Wage Act of 2023.” December.

_____. 2024b. “How Changes in Economic Conditions Might Affect the Federal Budget: 2024 to 2034.” April.

_____. 2024c. “How Economists Could Help Inform Economic and Budget Analysis Used by the US Congress.” *Journal of Economic Perspectives* 38, no. 2: 3-24. https://doi.org/10.1257/jep.38.2.3.

_____. 2024d. “Adding Debt-Service Effects to CBO’s Cost Estimates.” June.

_____. 2024e. “An Update to the Budget and Economic Outlook: 2024 to 2034.” June.

_____. 2024f. “Effects of the Immigration Surge on the Federal Budget and the Economy.” July.

Congressional Research Service. 2023. “Dynamic Scoring in the Congressional Budget Process.” March.

De Lipsis, Vincenzo, Matteo Deleidi, Mariana Mazzucato, and Paolo Agnolucci. 2023. Macroeconomic Effects of Public R&D. Working Paper, UCL Institute for Innovation and Public Purpose 2023-02.

Demirel, U. Devrim and Matthew Wilson, “Effects of Fiscal Policy on Inflation: Implications of Supply Disruptions and Economic Slack.” Working paper, Congressional Budget Office 2023-05. April.

Dyevre, Arnaud. 2023. “Public R&D Spillovers and Productivity Growth.” https://www.arnauddyevre.com/files/arnaud_dyevre_jmp.pdf.

Edelberg, Wendy. 2016. “Dynamic Analysis at CBO.” Presentation at the University of Chicago Booth School of Business. Congressional Budget Office.

Elmendorf, Douglas. 2015. “‘Dynamic Scoring’: Why and How to Include Macroeconomic Effects in Budget Estimates for Legislative Proposals.” *Brookings Papers on Economic Activity* 2015, no. 2: 91-149. https://dx.doi.org/10.1353/eca.2015.0005.

Elmendorf, Douglas and Heidi Williams. 2024. “How Does Accounting for Population Change Affect Estimates of the Effect of Immigration Policies on the Federal Budget?” Penn Wharton Budget Model. January.

Esche, Matthew, Jeremy Neufeld, and Heidi Williams, 2024. “Estimating Counterfactual Population and Status.” January.

Fieldhouse, Andrew J., and Karel Mertens. 2023. “The Returns to Government R&D: Evidence from U.S. Appropriations Shocks.” Working paper, Federal Reserve Bank of Dallas no 2305. December. https://doi.org/10.24149/wp2305r1.

Frentz, Nathaniel, Jaeger Nelson, Dan Ready, and John Seliski. 2020. “A Simplified Model of How Macroeconomic Changes Affect the Federal Budget.” Working paper, Congressional Budget Office Working Paper 2020-1. January.

Grossman, Gene M., and Elhanan Helpman. 1991. “Quality Ladders in the Theory of Growth.” *The Review of Economic Studies* 58, no. 1: 43-61. https://doi.org/10.2307/2298044.

Gullo, Theresa, Benjamin Page, David Weiner, and Heidi Williams. Forthcoming. "Estimating the Budgetary and Economic Effects of Research Investments." *Tax Policy and the Economy*. National Bureau of Economic Research. https://www.nber.org/conferences/39th-annual-nber-tax-policy-and-economy-conference-2024

Hall, Keith. 2024. "Beyond Dynamic Scoring: Enhancing CBO's Evaluation of Program Integrity Spending." Mercatus Center. June. http://dx.doi.org/10.2139/ssrn.4898038.

Hubbard, Glenn. 2015. "Comment on 'Dynamic Scoring': Why and How to Include Macroeconomic Effects in Budget Estimates for Legislative Proposals." *Brookings Papers on Economic Activity*, 2015, no. 2: 91-149. https://doi.org/10.1353/eca.2015.0015.

Joint Committee on Taxation. 2015a. "Macroeconomic Analysis at the Joint Committee on Taxation and the Mechanics of Its Implementation (JCX-3-15)." January.

_____. 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 (JCX-107-15)." August.

_____. 2015c. "A Report to the Congressional Budget Office of H.R. 2510, 'Macroeconomic Effects of the Bonus Depreciation Modified and Made Permanent,' as Ordered to be Reported by the House Committee on Ways and Means (JCX-134-15)." October.

_____. 2017. "Macroeconomic Analysis of the Conference Agreement for H.R. 1, the 'Tax Cuts and Jobs Act' (JCX-69-17)." December.

_____. 2018a. "Overview of Joint Committee Macroeconomic Modeling (JCX-33-18)." April.

_____. 2018b. "Macroeconomic Analysis Of H.R. 6760, The 'Protecting Family And Small Business Tax Cuts Act Of 2018' As Reported By The Committee On Ways And Means (JCX-79-18)." September.

_____. 2020. "An Overview of a New Overlapping Generations Model With an Example Application in Policy Analysis (JCX-22R-20)." October.

_____. 2023. "The Joint Committee on Taxation Revenue Estimating Process." Presentation. January.

Jones, Benjamin F. and Lawrence H. Summers. 2022. "A Calculation of the Social Returns to Innovation." In *Innovation and Public Policy*, edited by Austan Goolsbee and Benjamin F. Jones. University of Chicago Press.

Lasky, Mark. 2022. "The Congressional Budget Office's Small-Scale Policy Model." Working paper, Congressional Budget Office 2022-08. September.

Liscow, Zachary. Forthcoming. "Getting Infrastructure Built: The Law and Economics of Permitting." *Journal of Economic Perspectives*.

McClelland, Robert, and Shannon Mok. 2012. "A Review of Recent Research on Labor Supply Elasticities." Working paper, Congressional Budget Office Working Paper 2012-12. October.

Moore, Rachel and Brandon Pecoraro. 2020. "Macroeconomic Implications of Modeling the Internal Revenue Code in a Heterogeneous-Agent Framework." *Economic Modelling* 87: 72-91. April. https://doi.org/10.1016/j.econmod.2019.07.007.

Moore, Rachel and Brandon Pecoraro. 2023. "Quantitative Analysis of a Wealth Tax for the United States: Exclusions and Expenditures." *Journal of Macroeconomics* 78. December. https://doi.org/10.1016/j.jmacro.2023.103559.

Myers, Kyle R., and Lauren Lanahan. 2022. "Estimating Spillovers from Publicly Funded R&D: Evidence from the US Department of Energy." *American Economic Review* 11, no. 7: 2393-2423. https://doi.org/10.1257/aer.20210678.

National Academies of Sciences, Engineering, and Medicine. 2017. *The Economic and Fiscal Consequences of Immigration*, Washington, DC: The National Academies Press. https://doi.org/10.17226/23550.

Peri, Giovanni. 2012. "The Effect of Immigration on Productivity: Evidence from U.S. States." *The Review of Economics and Statistics* 94, no. 1: 348-358.

Potter, Brian, Arnab Datta, and Alec Stapp. 2022. "How to Stop Environmental Review from Harming the Environment." Institute for Progress. September.

Prato, Marta. 2022. "The Global Race for Talent: Brain Drain, Knowledge Transfer, and Economic Growth." Conditionally accepted at the *Quarterly Journal of Economics*.

Reichling, Felix, and Charles Whalen. 2012a. "Assessing the Short-Term Effects on Output of Changes in Federal Fiscal Policies." Working paper, Congressional Budget Office 2012-08. May.

Reichling, Felix, and Charles Whalen, 2012b. “Review of Estimates of the Frisch Elasticity of Labor Supply.” Working paper, Congressional Budget Office 2012-13. October.

Romer, Paul. 1990. "Endogenous Technological Change.” *Journal of Political Economy* 98, no. 5: 71-102.

Seliski, John, Aaron Betz, Yiqun Gloria Chen, U. Devrim Demirel, Junghoon Lee, and Jaeger Nelson. “Key Methods that CBO Used to Estimate the Effects of Pandemic-Related Legislation on Output.” Working paper, Congressional Budget Office 2020-07. October.

Shackleton, Robert. “Estimating and Projecting Potential Output Using CBO’s Forecasting Growth Model.” Working paper, Congressional Budget Office 2018-03. February.

West, Darrell. 2022. “R&D for the Public Good: Ways to Strengthen Societal Innovation in the United States.” Brookings Institution.

Whalen, Charles J., and Felix Reichling, “The Fiscal Multiplier and Economic Policy Analysis in the United States.” Working paper, Congressional Budget Office 2015-02. February.