SPRING 2018
NEW YORK UNIVERSITY
SCHOOL OF LAW

"Removing the free lunch from dynamic scores: Reconciling the scoring perspective with the optimal tax perspective"

Greg Leiserson
Washington Center for Equitable Growth

January 16, 2018
Vanderbilt Hall – 208
Time: 4:00 – 5:50 p.m.
Week 1

SCHEDULE FOR 2018 NYU TAX POLICY COLLOQUIUM

(All sessions meet from 4:00-5:50 pm in Vanderbilt 208, NYU Law School)

1. Tuesday, January 16 – Greg Leiserson.Washington Center for Equitable Growth. "Removing the Free Lunch from Dynamic Scores: Reconciling the Scoring Perspective with the Optimal Tax Perspective."

2. Tuesday, January 23 – Peter Dietsch, University of Montreal Philosophy Department. "Tax Competition and Global Background Justice."

3. Tuesday, January 30 – Andrew Hayashi, University of Virginia Law School. "Countercyclical Tax Bases."

4. Tuesday, February 6 – Gerald Auten, U.S. Treasury Department. "Income Inequality in the United States: Using Tax Data to Measure Long-Term Trends."

5. Tuesday, February 13 – Vanessa Williamson, Brookings Institution.

6. Tuesday, February 27 – Jacob Goldin, Stanford Law School.

7. Tuesday, March 6 – Lisa Phillips, Osgoode Hall Law School. "Gendering the Analysis of Tax Expenditures."

8. Tuesday, March 20 – Michelle Hanlon, MIT Sloan School of Management.

9. Tuesday, March 27 – Damon Jones, University of Chicago Harris School of Public Policy.

10. Tuesday, April 3 – Ajay Mehrotra, American Bar Foundation and Northwestern University School of Law. "T.S. Adams and the Beginning of the Value-Added Tax."

11. Tuesday, April 10 – Jason Furman, Harvard Kennedy School.

12. Tuesday, April 17 – Emily Satterthwaite, University of Toronto Law School. "Electing into a Value-Added Tax: Survey Evidence from Ontario Micro-Entrepreneurs."

13. Tuesday, April 24 – Wolfgang Schon, Max Planck Institute. "Taxation and Democracy."

14. Tuesday, May 1 – Mitchell Kane, NYU Law School.

Removing the free lunch from dynamic scores: Reconciling the scoring perspective with the optimal tax perspective

Greg Leiserson, Washington Center for Equitable Growth
January 3, 2018

Conventional estimates of the revenue effects of proposed tax legislation assume that the legislation would not change macroeconomic aggregates such as output, the capital stock, and employment. Dynamic estimates relax this assumption and—in the emerging consensus approach—replace it with two alternative assumptions. First, the macroeconomic analysis supporting dynamic estimates assumes future policy changes sufficient to address the fiscal imbalances that exist in CBO's current-law baseline. These changes are assumed to take effect after the period for which economic results are reported. Second, in many but not all cases, the analysis assumes additional future policy changes that offset any change in the government's present value fiscal position that the proposed legislation would cause, again taking effect after the period for which results are reported. This approach to scoring tax legislation severs the link between tax revenue and non-interest spending and thus ignores the primary benefit of raising revenues while highlighting the primary cost. By minimizing the benefits of tax revenues in this way, dynamic scores suggest that tax cuts offer an apparent free lunch. This paper outlines an alternative approach to presenting macroeconomic analyses and the dynamic scores they support that would offer policymakers a clearer illustration of the economic tradeoffs they face.

I. Introduction

The nonpartisan staff of the congressional Joint Committee on Taxation (JCT) prepares revenue estimates for tax legislation to inform congressional consideration of the legislation and for use in enforcing the rules that govern the congressional budget process. These revenue estimates show the impact that a legislative proposal would have on revenues over the next ten years relative to a baseline projection of federal revenues.[1]

Conventional estimates produced by JCT assume that the legislation would not change macroeconomic aggregates such as output, the capital stock, and employment. However, subject to this constraint, the analyses do assume that people will respond to a change in tax law in a manner consistent with current research. For example, an analysis of a proposal to restrict the home mortgage interest deduction might conclude that the proposal would reduce the value of mortgages and the value of the owner-occupied housing stock over the next decade. But the analysis might also assume an increase in the capital stock in other sectors such that the aggregate capital stock remains fixed even as the value of owner-occupied housing decreases.

This approach to preparing revenue estimates creates a sharp difference between margins that are allowed to respond and those that are not. This difference is particularly relevant for proposals that are, at least in part, justified on the basis of claims about economic growth that would result from their enactment.

In recent years, increasing attention has been devoted to so-called dynamic scores, which relax the assumption that aggregate economic activity does not change as a result of proposed legislation. In 2015, the House of Representatives adopted a rule that requires a dynamic score for certain legislation considered in that chamber, and the budget resolution subsequently adopted by both chambers had a similar requirement. While JCT produced a couple of dynamic estimates in 2015, JCT's first dynamic estimates of complex tax legislation under the new rules were released in 2017 during the consideration of the Tax Cuts and Jobs Act.[2,3] In addition to the official scores produced by JCT, independent organizations produce both conventional and dynamic scores as well.

In what appears to be an emerging consensus approach, dynamic scores replace the assumption that aggregate economic activity does not change as a result of proposed legislation with two alternative assumptions.[4] First, the macroeconomic analysis supporting dynamic estimates assumes future policy changes sufficient to address the fiscal imbalances that exist in CBO's current-law baseline. These changes are assumed to take effect after the period for which economic results are reported. Second, in many but not all cases, the analysis assumes additional future policy changes that offset any change in

[1] Revenue estimates in this sense also include effects classified as changes in spending through, for example, changes in refundable tax credits.

[2] JCT had produced macroeconomic analyses of complex legislation prior to the adoption of the rules requiring such analyses at the beginning of 2015. See, for example, JCT's analysis of the proposal for tax reform released by Representative Camp in 2014 (JCT 2014).

[3] For convenience, this article refers to the enacted legislation as the Tax Cuts and Jobs Act notwithstanding that this title was removed late in the legislative process.

[4] For examples of this approach see JCT (2017a, 2017b, 2017c), Page et al. (2017a, 2017b, 2017c), and Penn Wharton Budget Model (2017a, 2017b, 2017c). The Tax Foundation (2017a, 2017b, 2017c) takes a different approach that effectively assumes the government budget constraint never binds.

the government's present value fiscal position that the proposed legislation would cause, again taking effect after the period for which results are reported.

The primary benefit of raising revenues is the opportunity to spend those revenues on some set of programs, whether they be for national defense, social insurance, poverty alleviation, the legal or regulatory regime, or something else entirely.[5] Yet the assumptions adopted in preparing a dynamic score serve to largely sever the link between revenues raised and government spending. By assuming that policy actions necessary to address the fiscal imbalances in CBO's current law baseline occur after the period for which results are reported, higher revenues do not result in any increase in programmatic spending. By assuming that any policy actions necessary to address a reduction in revenues likewise occurs after the period for which results are reported, lower revenues cause no reduction in programmatic spending. Changes in revenues matter only for their implications for debt dynamics.

The primary cost of raising revenues is the impact those taxes have on economic behavior. Taxes can discourage labor supply, investment, or other productive economic behavior, and can encourage unproductive avoidance activities. These types of issues are the central focus of a dynamic analysis.[6]

Thus, the assumptions used in dynamic scores downplay the primary benefit of tax revenues while highlighting their primary cost. By minimizing the benefits of tax revenues, dynamic scores can suggest that deficit-financed tax cuts offer an apparent free lunch and deficit-reducing tax increases have few economic benefits.

More concretely, the emerging consensus approach to dynamic scoring suffers from several related problems. First, future policy changes can affect behavior today. Thus, the assumptions made about how policymakers will respond in the future can affect the results reported for the budget window. Second, economic impacts can differ sharply inside and outside the traditional ten-year budget window. Thus, results reported for the budget window can be misleading. Third, dynamic scores can improperly rank policies in terms of their true long run cost and economic effects. Fourth, critical attention is directed to the wrong set of assumptions. Namely, critical attention is focused on assumptions relating to the debt dynamics that occur before future fiscal adjustments are made, not the future fiscal adjustments themselves. Finally, macroeconomic analyses can be inconsistent with distribution analyses that provide information about how changes in tax policy affect economic well-being.

This paper suggests that the current approach to dynamic scoring is motivated by—and makes sense according to—a perspective termed the scoring perspective. The scoring perspective prioritizes the accuracy of the results inside the traditional ten-year budget window and apparent fidelity to the legislation under consideration. The emerging consensus approach, consistent with the scoring perspective, defers necessary fiscal adjustments until well after the analysis period and focuses on results for the ten-year window. As a result, the choice of future fiscal adjustment is relatively less important and analysts can assert that the results speak to the policy option under consideration.

[5] In the language of welfare economics, the effects of the revenue in isolation would be termed a transfer, not a benefit. This article adopts a colloquial definition of benefits which is more useful in the current context and avoids the need to make explicit assumptions about whether utility is transferrable.

[6] Conventional scores will also include the revenue loss attributable to avoidance activities that do not change macroeconomic aggregates.

If instead, the analysis is motivated by an alternative perspective—termed the optimal tax perspective—a different approach to the analysis would be more appropriate. The optimal tax perspective prioritizes an economically robust analysis of tradeoffs in tax and spending policy. Under this perspective, it makes little to sense to ask how a change in tax law that changes the level of revenues (and is not accompanied by a commensurate change in spending) affects the economy, because such a policy change most likely requires additional legislation in the future.

There is a striking disconnect between these two approaches. Under the optimal tax perspective, the revenue requirement is fundamental to any analysis and often stipulated at the outset. Under the scoring perspective, the change in revenues is precisely the object that analysis is intended to estimate. To the extent that a change in the level of revenues poses a practical problem for the analysis, this is a problem to be worked around.

The current approach to generating dynamic scores reflects, in part, an effort to minimize the analytical problems resulting from changes in the government's long-run fiscal position. This is a reasonable response to the task with which scoring organizations are faced, namely providing a meaningful analysis of a policy change that requires some future offsetting policy change that has not been specified by legislators. But, in doing so, analysts are in a sense exploiting the flaws of a fixed budget window and turning that to their advantage. If the budget window were extended, these strategies would not be viable.

However, this approach also reduces the quality of the information provided to policymakers. The analytical approach underlying the macroeconomic analyses that support existing dynamic scores operates as a timing gimmick.

Informed by the optimal tax perspective, this paper recommends an alternative approach to presenting the information in macroeconomic analyses of tax proposals—and the dynamic scores they support—to provide policymakers with more information in making their decisions.

First, analysts should make explicit assumptions about future policy offsets. As a default approach, analysts should evaluate two policy scenarios, one assuming a future fiscal adjustment resulting from a change in transfer spending and one assuming a future fiscal adjustment resulting from a change in taxes intended to proxy for a reversal of the proposed legislation. As under the emerging consensus approach, these fiscal adjustments should occur after a substantial lag.[7]

Second, scorekeepers should report the results of their analysis extending beyond the time the fiscal adjustment takes effect. In other words, the analysis should show the economic effects of the policy change for three distinct periods: the budget window, the period following the budget window and before the fiscal offset is enacted, and the period after the fiscal offset is enacted.

Third, a welfare analysis should be presented that is consistent with the macroeconomic results. A welfare analysis attempts to evaluate the impact of policy changes on economic well-being. As is well-known, GDP is not a measure of economic well-being. This is particularly true in scenarios—like those contemplated in dynamic scores of tax legislation—where a change in output is often generated by a change in investment or labor supply which results in a consequent cost or benefit for investors or

[7] Implicit in this recommendation is the recommendation that an explicit statement of future policies should be made even for analysis relying on models that do not require such an assumption to solve as a technical matter.

workers. Thus, a welfare analysis, standard in academic treatments of optimal tax analysis, becomes essential to understand the economic effects of proposals for fundamental tax reform.

Fourth, a supplemental distribution analyses should be presented in conjunction with the traditional distribution analysis. JCT's traditional distribution analysis reports select measures of the change in tax liabilities consistent with the organization's conventional revenue estimate. However, these distribution estimates often report gains and losses attributable to a change in the government's net fiscal position that are not sustainable. As a potentially more accessible complement to the welfare analysis, a supplemental analysis should report the impact on families assuming a fiscal offset necessary to achieve dynamic revenue balance consistent with each of the two financing adjustments imposed above but imposed beginning in the first year after the budget window.

The recommendations above apply to the macroeconomic analysis that supports the estimates of the budgetary effects of a piece of legislation. The score itself, however, reports only the impacts on the government budget. In this case, the key additional results are simply the budgetary impacts up to and beyond the time the fiscal adjustments take effect.

The changes above would likely have—and are intended to have—relatively little impact on the ten-year scores presented by the Joint Committee on Taxation. Ideally, the difference in the score under each of the assumptions about future fiscal adjustments would be modest during this period. However, it is to be expected that the two paths would diverge more substantially after the budget window in the case of unbalanced fiscal policies. The primary purpose of this proposal is to provide additional information that would more clearly illustrate the economic tradeoffs are more clearly illustrated involved in proposals to increase or decrease revenues.

If this recommendation is adopted, the analysis of proposals for deficit-financed tax cuts that generate growth today at the cost of future spending cuts, future tax increases, or a combination, would clearly indicate this tradeoff. Analysis of proposals for deficit-reducing tax increases that reduce growth today and allow for higher spending in the future, future tax cuts, or a combination, would likewise show this. By directly reflecting the link between tax revenues and spending, these analyses would replace the apparent free lunch of tax cuts with the real economic tradeoffs.

This recommendation would give the specification of future policy options a central role in the macroeconomic analysis of legislative proposals. It thus places the staff of nonpartisan organizations devoted to scoring proposals in a somewhat awkward position as they aim to provide a nonpartisan analysis of what will happen under a proposal, not make policy recommendations. Consistent with the role of these organizations, it would be appropriate for policymakers to formalize the high-level approach to analyzing unbalanced fiscal policies in rules like those that require dynamic scores in the first place such that scoring organizations themselves are able to evaluate policies without appearing to choose policy options.

Section two of this paper describes scores in more detail and their use in the congressional budget process. Section three describes current approaches to dynamic scoring and how they affect proposals for tax changes. Section four outlines several problems with the current approach. Section five describes the proposal for improving the presentation of dynamic scores. Section six provides a discussion of select issues implicated in the proposal. Section seven considers alternatives to the proposal here that would require more changes to the substance of the economic analysis. A final section concludes.

II. Scoring legislative proposals

Legislative proposals are scored by the Congressional Budget Office, the Joint Committee on Taxation, or the two organizations working together. The scores produced by these organizations are intended to inform legislators in making decisions about policies but they also have implications for the policymaking process.

II.1. The conventional revenue (and budget) estimate

The basic score produced by the Joint Committee on Taxation (JCT) measures the change in federal revenues and outlays that would result from enacting a proposal relative to the baseline projections of revenues and outlays.

The baseline is often described as a reflection of current law or what would happen if the proposal were not enacted, but in practice the baseline is specified in statute and can differ from what would happen under current law for a variety of reasons. As CBO states in its most recent budget projections, "[t]he baseline is not intended to be a forecast of budgetary outcomes; rather, it is meant to provide a neutral benchmark that policymakers can use to assess the potential effects of policy decisions" (CBO 2017b).

JCT's revenue estimates are typically presented for the ten fiscal years that make up the budget window. For example, scores for the Tax Cuts and Jobs Act were presented for fiscal years 2018-2027 (JCT 2017d). The score reports the change in nominal revenues attributable to each provision of a proposal for each year during the window.

In estimating the effect of a proposal, analysts must estimate the impact of the proposal on the behavior of many other actors in the economy. Depending on the provisions of a proposal, it could affect the behavior of workers and firms by changing the tax incentives they face, and it could lead to actions by other entities with policymaking power, such as state and local governments or foreign countries. In addition, a proposal may raise questions about the specification of federal policy itself, for example, if there is ambiguity in the ways in which executive branch agencies would interpret the legislation.

However, as noted above and as will be discussed in greater detail below, conventional scores assume that macroeconomic aggregates remain fixed. Macroeconomic aggregates include the capital stock, labor supply, gross domestic product, or other similar variables. Thus, while policies can shift employment into or out of certain industries or shift capital into or out of certain sectors, they will not change total employment or the aggregate capital stock. The exact way in which this assumption is implemented can vary across organizations. For example, JCT describes the organization's assumption as one of fixed income and thus fixed GNP (JCT 2017e).

Scoring organizations frequently produce information beyond the impact a proposal would have on the budget. For example, JCT frequently produces distributional analyses of legislative proposals. CBO often publishes a wider array of supplemental information about the effects of a proposal. For example, in recent analyses related to the Affordable Care Act and other proposals to change the federal government's policies on health insurance, CBO has typically included estimates of changes in insurance coverage. Estimates of insurance coverage necessarily underlie any estimate of the impact of proposals related to health care and health insurance on the government budget, but they can also be of substantial interest in their own right.

II.2. Interpreting revenue (and budget) estimates

A natural interpretation of the score is the impact of a proposal's enactment on the federal budget. Yet this simplified definition of scoring requires some additional nuance. Scores have purposes and uses beyond those of simply explaining what will happen that complicate interpretation. Scores provide the information necessary to enforce targets established by the budget process, to inform policymakers about the costs of their policies, and to inform policymakers about the impact of their policies more generally. Indeed, CBO places budget enforcement at the center of its description of cost estimates on its website:

> …cost estimates are intended to ensure that when the House and Senate consider legislation recommended by committees, Members have information about the budgetary consequences of enacting that legislation that can be used to enforce budgetary rules or targets (CBO 2017a).

Revenue estimates can differ from an assessment of the impact enacting a law would have for several reasons. First, as noted above, the baseline against which policies are measured is not an assessment of what will happen under current law, but rather an amalgamation of current law and explicit statutory rules. For example, the baseline departs from current law in assuming that adequate funding is available to pay for spending from trust funds, certain expiring excise taxes are assumed to continue at current rates, and the debt limit does not constrain federal borrowing. Thus, for example, a proposal to change Social Security disability insurance tax rates would not be scored as changing disability insurance outlays, even though there is a direct relationship between the two under current law.[8,9]

Second, the policy alternative examined is not the same as the policy regime that would result if the proposal were enacted. The deviations between current law and the baseline noted above would necessarily carry over into the policy alternative. Another example arises from the operation of the Statutory Pay-As-You-Go Act (statutory PAYGO). Statutory PAYGO imposes automatic spending cuts if the net effect of legislation affecting certain types of spending and revenues enacted by Congress would be to increase deficits. Thus, under statutory PAYGO, enacting policies that would increase the deficit (judged in isolation) would not do so under a complete depiction of current law. While recognizing statutory PAYGO in revenue estimates would make such estimates uninformative about the impacts of legislation and, as such, there are strong arguments against doing so, it highlights the extent to which scores do not measure only the impact a proposal would have. Instead, they compare one counterfactual policy regime to another counterfactual policy regime with the purpose of informing Congressional deliberations and enforcing budget process.

Third, when considering proposals that require certain behavior by executive branch agencies but do not provide additional funding for those agencies to undertake those tasks, CBO still assigns a cost to the proposal:

> When a law imposes a new requirement on an agency (such as preparing a plan or completing a study), complying with that new requirement will entail the use of resources, and the cost of

[8] See footnote 2 of CBO's cost estimate for H.R. 1314, available at https://www.cbo.gov/sites/default/files/114th-congress-2015-2016/costestimate/hr1314.pdf, for an example.

[9] There are legal ambiguities in precisely what would happen when one of the Social Security trust funds is exhausted. Beneficiaries would remain entitled to the full benefits but the Social Security Administration would be prohibited from paying those benefits. See Morton and Liou (2017) for more details.

> carrying out that requirement is the amount of resources used. In general, in bills that are being considered, such requirements would apply to future fiscal years, for which appropriations have not yet been determined—so the requirements could, in fact, influence the amount of budget authority available to the agency in the future. Even if future funding was not affected, the agency would have to spend appropriated resources on that new activity instead of spending them to carry out other responsibilities. The resources used to carry out the new activity would be a measure of the "opportunity cost" of not carrying out other responsibilities (CBO 2017a).

Finally, under the Fair Credit Reporting Act, budget estimates for loan programs are computed as the present value of the expected subsidy or profit on loans issued by the government using Treasury yields in the year the loans are disbursed. Such an estimate is clearly not an accurate statement of the cash flows of the government. Moreover, under the widely discussed fair value approach to measuring the costs of loan programs often suggested as an alternative, the cost would be measured using interest rates that assume a higher rate of return to account for the risk of non-payment. This approach still would not provide an accurate statement of the cash flows of the government. Proponents of the fair value approach typically justify it on the grounds that it would provide a more accurate statement of the social cost of the program.[10]

Thus, a simple interpretation of revenue or cost estimates as the impact a proposal would have on federal revenues or spending is incomplete.[11] Revenue estimates provide information about how a proposal will change revenues and spending under a complex set of assumptions about the policy environment. Conventional and dynamic scores should both be understood in this light. The laws, rules, and conventions governing the two types of scores reflect how a proposal will change revenues and spending under a particular set of assumptions about government policy and behavior, but they do not provide a direct measurement of the effect a policy will have on the government budget.

II.3. Conventional scores and dynamic scores

The prior section outlined several reasons that conventional scores differ from an estimate of the effect a proposal would have on the federal budget if it were enacted grounded in the assumptions about the policy regime and the measurement of budgetary effects. Another reason a conventional score could differ from an estimate of the effect a proposal would have on the budget if enacted is grounded in the assumption of fixed macroeconomic aggregates embedded in a conventional score. This section reviews different assumptions about behavior that could be made as part of a revenue estimate and classifies the assumptions in four types: fully static analysis, conventional analysis, dynamic analysis, and a fourth type referred to as comprehensive analysis.

A *fully static analysis* holds all behavior fixed and mechanically assesses the changes in revenues and spending that that would result if a different policy regime applied to the current behavior. For example, a fully static evaluation of a change in tax law would assume the same income levels, spending patterns, and reporting patterns as projected under current law and simply compute the level of tax liability that

[10] See CBO (2012), Kamin (2014), and Marron (2014) for further discussion of the issues involved in budgeting for federal lending programs.

[11] The discussion in this section is not a complete list of deviations. For example, Rules 3 and 14 of the scorekeeping guidelines formalized in the Balanced Budget Act of 1997 effectively require CBO and JCT to assume that funding provide to the IRS to improve compliance does not increase revenues in the official estimates. See Holtzblatt and McGuire (2016) for additional discussion.

would be paid under this alternative regime. A fully static analysis of an alternative health care system would assume that nobody would change their decisions about their choices about sources of insurance coverage.

As the health care example makes clear, it is difficult to even conceive of how this type of analysis should apply in many cases. If a government program is opt-in rather than opt-out, would a fully static analysis assume that nobody signs up? Similar questions could be raised about tax proposals. Would an increase in the employer share of payroll taxes and a decrease in the employee share of payroll taxes increase total compensation and the employer wage bill?

For this reason, no scoring entity engages in fully static analysis. It would not be a useful or informative exercise.[12]

A *conventional analysis* assumes that macroeconomic aggregates remains fixed. Examples of macroeconomic aggregates include the capital stock, total labor supply, gross domestic product, or other similar variables.[13]

However, conventional analyses assume that the behavior of the people and firms that make up the economy can change. In many cases these analyses will also necessarily assume that the behavior of non-federal government entities must change as well. In the case of the Affordable Care Act, for example, conventional scores included the wholesale creation of new markets, large changes in the decisions about insurance coverage made by families and employers, major changes in state health policy, and major changes in the design of health insurance products. Likewise, in essentially any score of proposals to change employment taxes with statutory incidence on employers, analysts assume meaningful changes in the compensation packages for employees.

The assumption that macroeconomic aggregates remain fixed can raise questions parallel to those for a fully static analysis above. For example, if a legislative proposal encourages additional savings in some population, how should savings elsewhere be adjusted such that the capital stock does not change? However, in the case of conventional analysis, these ambiguities are typically more modest in scope than they would be for a fully static analysis.

A *dynamic analysis* relaxes the assumption that macroeconomic aggregates remain fixed. Relaxing this assumption is most frequently highlighted in the cases of changes in tax policy, for which changes in the relevant tax rates can affect behavior and could increase or decrease the total output. However, the difference between conventional and dynamic analysis could also be important in proposals in a variety of different areas, including immigration reform and infrastructure development, among others.

In the emerging consensus approach to producing dynamic scores, two alternative assumptions are added to replace the assumption of fixed macroeconomic aggregates. First, the macroeconomic analysis supporting dynamic estimates assumes future policy changes sufficient to address the fiscal imbalances

[12] Certain entities produce (nearly) fully static distributional analyses of proposals, including the U.S. Treasury (2015) and the Tax Policy Center (see, e.g. Tax Policy Center Staff 2017). As I have argued elsewhere, fully static distributional analyses are informative about the impact of a tax proposal on welfare, particularly in the case of a revenue-neutral tax change (Leiserson 2017).

[13] As noted above, there is scope for difference in the precise implementation of an assumption of fixed macroeconomic aggregates in conventional scores across organizations.

that exist in CBO's current-law baseline. These changes are assumed to take effect after the period for which economic results are reported. Second, in many but not all cases, the analysis assumes additional future policy changes that offset any change in the government's present value fiscal position that the proposed legislation would cause, again taking effect after the period for which results are reported. While results are typically reported only for the period before the fiscal adjustment takes effect, some analysts further focus on effects in the budget window.

It is worth distinguishing conceptual issues in defining these types of scores with practical issues. Dynamic scores created by numerous entities suffer from ad hoc assumptions about behavior. For example, some have suggested—though it is not clear from publicly available materials—that the Joint Committee on Taxation assumes that foreign countries do not change their statutory corporate tax rates in response to a change in the U.S. statutory corporate tax rate, even as the analysts also assume that changes in relative tax rates drive meaningful international shifting of activity that would create pressure on foreign governments. Such an assumption would raise questions about the accuracy of a dynamic score, but it is not essential to the conceptual definition of a dynamic score. This paper focuses on the conceptual issues, but the practical issues in generating a dynamic score merit attention in their own right.

A *comprehensive analysis*, which is not a widely-used term and is introduced here for conceptual completeness, would seek to eliminate any ad hoc assumptions about future policy action. At least two alternative approaches could be imagined and possibly more. In the first, a model for fiscal crises could be adopted and the analysis could report the details of the change in the timing or impacts of fiscal crises that results from the policy. This approach would be closer to the conventions of scoring as it would assume no future policy action. However, it would likely be incorrect as a depiction of reality in the U.S. context, as people and firms would most likely assume an endogenous response by policymakers rather than an eventual crisis.

In the second, a political economy model of the feedback from budgetary outcomes to policies could be adopted that allows for an explicit estimation of the policies future Congresses would enact to resolve the fiscal gap. This model could be applied to the baseline to resolve questions of fiscal sustainability under current law as well as to the policy alternative. This approach would depart more radically from the traditional conventions of scoring, but has a stronger claim to reflecting a correct conceptual framework as—assuming the political economy model were correct—it would indicate the likely effects of the policy change.

This fourth conceptual approach makes clear that the assumptions underlying a dynamic score are incorrect in much the same sense that the assumptions underlying a conventional score are. Economic actors will make assessments about likely future policy changes and incorporate them in their behavior. Moreover, ignoring the likelihood of future policy action to partially reverse the economic impacts of a legislative proposal that causes a deterioration in the fiscal position likely means that not only will deficit-increasing policies see overstated effects as a result of the short-term focus induced by the budget window, but that they will also suggest overstated economic effects by failing to recognize an expectation of future policy changes.

One interpretation of the assumptions of conventional scoring is that they apply an ad hoc penalty that undervalues certain types of policies with growth-enhancing effects. Similarly, the assumptions of dynamic scoring apply an ad hoc bonus that overvalues certain types of growth-enhancing reforms. The

proposal made in this paper attempts to achieve a practical implementation of some of the insights provided by this fourth approach within the general framework of current dynamic scoring assumptions.

II.4. The implicit assumptions underlying dynamic scores

This paper argues that the macroeconomic analysis supporting dynamic scores is best understood as replacing the assumption of fixed macroeconomic activity with two alternative assumptions. First, such analysis assumes future policy changes sufficient to address the fiscal imbalances that exist in CBO's current-law baseline. These changes are assumed to take effect after the period for which economic results are reported. Second, in many but not all cases, the analysis assumes additional future policy changes that offset any change in the government's present value fiscal position that the proposed legislation would cause, again taking effect after the period for which results are reported. This section describes this argument in additional detail.

Models used for dynamic scoring can be split into two types: those that require an explicit specification of offsetting future fiscal policies to solve as a mechanical matter and those that do not. In the former case, it is clear that analysts are making an assumption about future fiscal policies. Of the three models used by the Joint Committee on Taxation in its macroeconomic analyses, two require explicit specification of future fiscal policies.[14] The model used by the Penn Wharton Budget model is also of this type (Page and Smetters 2017).

However, even in the case of models that do not require an explicit specification of future fiscal policies to solve, the underlying analysis is generally best understood as implicitly requiring such an assumption for the analysis to make sense in the longer run.

There is an inherent limit on the government's ability to borrow. In the most generous case, this limit would manifest when government debt becomes so large that the required interest payments exceed the maximum revenue yield of the economy assuming revenue-maximizing tax rates. Such a level of debt would be quite high relative to current levels.[15] In practice, concerns about the credit-worthiness of the government would likely arise much sooner.

If some limit on debt is recognized, even models that do not require an explicit assumption about how the government's fiscal position is modified in the future effectively make such an assumption if they assume that the riskiness of government debt does not change over time when fiscal policies are unsustainable. The assumption that the riskiness of government debt does not change over time necessarily assumes a government's ability to repay, and it thus also assumes that unsustainable policies are ultimately reversed.

[14] The three models JCT uses are the MEG model, the OLG model, and the DSGE model. The OLG model and the DSGE model require specification of offsetting fiscal policies to solve as a mechanical matter. See JCT (2017a) and cites therein for more on this point.

[15] Strictly speaking, this logic applies to a barter economy or a government that borrows in a currency other than its own. In the case of a government that borrows in a currency that it issues, default is always a choice because it is possible to simply print additional money to pay past debts. In the models used for dynamic scoring this possibility is typically ruled out by assumption. It is not considered further here.

In this sense, only a model that explicitly attempts to model a government fiscal crisis ultimately avoids an implicit assumption about future fiscal policies. Modeling a fiscal crisis directly effectively restores the link between revenues and spending commitments.

Thus, since most analysts do not incorporate increasing risk premia on government debt or fiscal crises in their macroeconomic analysis, they are implicitly assuming future changes in fiscal policy whenever policies are unsustainable. Moreover, since most analysts also follow the assumptions of CBO's baseline, which assumes that current policies are unsustainable, they necessarily assume future policy action in the baseline. If a policy causes a deterioration in the fiscal balance they assume an additional future policy response.

That said, models that do not require an explicit specification of future fiscal policies to solve as a mechanical matter are insensitive to the assumptions chosen as to how the required future fiscal adjustment is made. Moreover, if the increase in the riskiness of government debt occurs only slowly over time, the impact of abstracting from this consideration in a first decade score of a modestly sized bill could be negligible. Thus, the difference between models that do and do not require an explicit specification of future fiscal policies to address unsustainable policies is likely better understood as whether the assumption about the nature of the future fiscal adjustment matters for nearer-term economic outcomes, rather than whether the model requires a specification of the future fiscal adjustment or not. This perspective is further strengthened by recognizing that the existence of a constraint on government borrowing is a basic feature of the broad class of neoclassical models used for dynamic scoring, not a feature specific to certain modeling methodologies.

This argument suggests that analysts make some implicit assumptions about future fiscal policies whenever they evaluate policies under current fiscal conditions (as judged by CBO), and make stronger assumptions in the case of most deficit-increasing policies.[16] As was seen in the case of the Tax Cuts and Jobs Act, deficit-increasing tax policies are likely to be a particularly relevant scenario for the kinds of tax legislation actively considered in Congress for which a dynamic score achieves prominence.[17]

Nonetheless, this argument also leaves open other scenarios in which the evaluation of policies could be done without an explicit or implicit assumption of future fiscal adjustment. These are the scenarios in which the government starts from a position of fiscal sustainability and considers policy proposals that do not put the government in an unsustainable position.[18] In this case, an analysis that is based solely on changes in the government's debt levels can be indicative of future economic outcomes. The essence of such an analysis, however, is that the government budget constraint does not bind because the government raises more money than it spends and does not plan to spend this money now or in the future. Thus, while it would be feasible to analyze policies without assuming any future policy action in the case of fiscal surplus—as will be highlighted in the sections that follow—it is not particularly informative about the economic tradeoffs facing policymakers to do so.

[16] See CBO (2017c) for the organization's latest assessment of the federal government's longer-term fiscal outlook.
[17] The analysis of the Camp proposal (JCT 2014) is the exception that proves the rule as the macroeconomic analysis achieved some prominence but the proposal had no political prospects.
[18] There is a sense in which this is true even for models for which an assumption about future fiscal policy is required to solve as a mechanical matter, because such models generally could be solved using the assumption that the government raises excess revenue and devotes it to spending with no public or private value.

III. The impact of dynamic scoring on revenue estimates

As discussed above, there appears to be an emerging consensus approach to generating dynamic scores. Fiscal imbalances in CBO's baseline are addressed by a fiscal adjustment that takes effect after the period for which analytic results are reported. And, if confronted with a proposal that increases the present value of net government deficits, analysts assume a fiscal adjustment several decades in the future and then present analysis primarily for the period prior to the point at which this adjustment occurs. The proposed policy change is then evaluated relative to the baseline policy regime as defined by statute.

These assumptions serve to sever links between government revenues and non-interest spending in the baseline. The use of the statutory baseline, for example, means that spending from trust funds is assumed to continue even when trust funds are exhausted. Moreover, in assessing the policy change, certain other feedback, such as that which would result from statutory PAYGO, is likewise excluded. As a result, the channels by which changes in tax law affect the economy will thus be the direct effect of changes in tax policy, such as the changes in rates, deductions, credits, and so forth that make up the tax plan; the changes in debt that result from a change in revenues; and the change in future fiscal policy necessary to resolve any long-term change in net deficits.

This section reviews the qualitative economic effects of changes in each of these areas and then considers the timing of such effects. It also includes a summary of the macroeconomic analyses released by JCT for the Tax Cuts and Jobs Act.

III.1. The impact of changes in tax policy

The focus of dynamic analyses of tax policy is the impact of changes in tax policy on the supply side of the economy. In the long run, output is determined by the supply of labor, the supply of capital, and the technology that converts capital and labor into output. In this context, technology is a broadly defined concept that includes not just the state of scientific knowledge, but frequently is a reduced form representation of the institutions, arrangements, and relationships that make up society.

Dynamic scoring models typically assume that changes in policy have no effect on the long run rate of technological progress and instead operate by changes in the quantity of labor and capital supplied. The primary focus of these models are the marginal tax rates that apply to labor and the marginal investment. All else equal, a reduction in the marginal tax rate on labor encourages workers to supply more labor and a reduction in the marginal tax rate on firms encourages firms to invest more.

Models vary in their treatment of other factors. For example, some models suggest an important role for the differences in tax rates on business income across countries that may then play a role in the shifting of internationally mobile capital or profits by firms.

Thus, the basic effect of adopting dynamic scoring model as far as changes in tax policy are concerned is to indicate that a reduction in effective tax rates on labor or investment will increase output and an increase in tax rates will reduce output. The change in output corresponds to a change in income, and this change in incomes will affect the revenue impact of the proposal. Higher incomes will increase revenues and reduce the cost of a tax cut. Lower incomes will reduce revenues and decrease the revenue gain from a tax increase.

III.2. The impact of changes in debt

Proposed tax legislation will frequently result in a change in the government's net fiscal position. Deficit-financed tax cuts increase deficits and debt and deficit-reducing tax increases reduce deficits and debt. These changes in federal borrowing can affect the economy through two main channels.

First, increases in federal borrowing can compete with private uses of the funds and thus increase interest rates in both the private and public sectors. Mechanically, the higher interest rates will increase government interest payments. They will also discourage investment and thus reduce the capital stock and output, reducing incomes and tax payments.

Second, increases in federal borrowing can increase the risk premium on government. While this would not necessarily crowd out private sector investment, it would increase the interest payments required on government debt and thus cause a deterioration in the government's fiscal position. However, this channel is excluded from most dynamic scoring models as they treat the resolution of the government's fiscal position in the future as a guaranteed event.

III.3. The impact of changes in future fiscal policy

The emerging consensus approach to dynamic scores implicitly assumes future changes in fiscal policy to resolve unsustainable changes in the government's fiscal position. The typical approach, however, defers the enactment of these policies for a substantial period of time after the end of the budget window in an effort to minimize the direct impact of these policies on economic outcomes in the budget window.

Nonetheless, in certain models these future policies are likely to have indirect effects in the window through their impact on the savings decisions of forward-looking households. These households may choose to save more or less today—depending on whether the proposed legislation increases net deficits or decreases them—so that they have the funds available to offset the future change in fiscal policies and thus achieve a more stable level of consumption over time. These policies may also have some direct effects. For example, a change in the taxation of business income in 30 years could have a modest effect on investment decisions in the final year of the budget window.

III.4. The timing of economic effects in dynamic scores

JCT's official dynamic scores focus on the impacts that are realized within the budget window. Thus, the impact of changes that occur quickly and over a longer time has an important impact on the results.

In most dynamic models, changes in labor tax rates generate roughly contemporaneous changes in labor supply. Thus, economic effects resulting from these changes occur rapidly. Changes in economic activity resulting from shifting of real economic activity across countries likewise occurs rapidly. Changes in capital accumulation resulting from changes in tax rates on investment occur over time, but have meaningful effects within ten years. Thus, in general, the direct behavioral effects of tax changes will be realized to a substantial extent within the window, though the precise extent will depend on the nature of the tax changes.

Changes resulting from the accumulation of debt will occur over a longer horizon. While models vary as to the precise way in which these channels are modeled, the impact of a proposal on deficits and debt—whether positive or negative—tends to compound over time and thus lead to larger effects in the short-

run rather than the long run. Thus, these effects are typically very modest in the first few years, notable by the end of the window, and continue to grow outside the window.

Finally, as noted above, changes in future fiscal policy are typically specified such that they have little effect inside the window as a result of shifting of economic behavior across time. Thus, there is a sense in which they have little effect inside the window. However, as will be discussed in greater detail below, there is an important sense in which the existence of future offsetting policies affects behavior inside the window even when the precise contours of that policy does not in certain types of models.

Thus, the practical impact of adopting dynamic scoring with a focus on a ten-year budget window is to highlight the relatively short-run supply-side effects of changes in tax policy, downplay the impact of changes in debt accumulation, and largely obscure the impact of changes in revenues on government policy.

III.5. JCT's macroeconomic analyses of the Tax Cuts and Jobs Act

JCT published three macroeconomic analyses of the Tax Cuts and Jobs Act: one for the version reported by the Senate Finance Committee, one for the version passed by the House of Representatives, and one for the conference report (JCT 2017a, 2017b, 2017c).

The three analyses are broadly similar. The introduction of each analysis highlights the percentage change in GDP on average over the ten-year budget window, the conventional score for the proposal, the additional revenue raised from additional economic activity, and the increase in interest payments on federal debt resulting from higher interest rates. In addition, a qualitative statement about growth in the second decade is provided.

In the body of the analysis, the most detailed estimates reported are for the budget effects of the legislation. Year-by-year estimates of the revenue impact for the conventional proposal and the change in that estimate from the macroeconomic analysis are provided.

In addition, the analysis provides (with slight variations across the reports), the average percent change in output, the capital stock, employment, and consumption over the budget window. For each of these measures, as well as the impact of changes in estate taxation and changes in the treatment of foreign activity, a qualitative discussion of relevant economic issues is also provided. Finally, the analysis offers a qualitative discussion of economic and budgetary effects in the second and third decades.

The analyses also provide various technical information, including a summary of salient features of the legislation for the macroeconomic analysis, certain parameters in the models used to evaluate the analysis, and a description of the method for producing the analysis, which consists of a weighted average of the results from three different models.

Finally, for the two models for which an explicit specification of future fiscal policies is required, the analyses indicate the assumptions used to resolve the unsustainable nature of baseline federal policies and the policies proposed in the Tax Cuts and Jobs Act. In the OLG model, the baseline assumes that government purchases decline over time to maintain a stable debt to GDP ratio. For the additional debt resulting from the proposed legislation, the analysis assumes that the policy is as specified for 30 years and then adjusts transfer payments to stabilize the debt to GDP ratio.

In the DSGE model, actors in the model assume that there is a probability that the government will adjust policy to stabilize the debt to GDP ratio. After ten years, the debt-to-GDP ratio returns to a level consistent with a steady state. The policy mix to support the changes in the debt to GDP ratio in the DSGE model are not specified.

IV. Problems with the current approach to dynamic scoring

The emerging consensus approach to dynamic scoring downplays the value of revenues and highlights the costs of raising revenues. This section outlines several flaws that cause the current approach to dynamic scoring to generate misleading results. Specifically: future policy changes can affect behavior today; economic impacts differ sharply inside and outside the budget window; dynamic scores improperly rank policies; critical attention is directed to subsidiary assumptions rather than those that truly determine the effects of policy changes; and the macroeconomic analysis may be inconsistent with the distributional analysis separately reported.

These technical issues all underlie a common high-level theme: the apparent link between taxes and growth inside the window from dynamic scores can be misleading as an indicator of the economic tradeoffs policymakers face.

Take the case of a deficit-financed tax cut for all families. The distribution analysis would then report that everybody comes out ahead. The dynamic feedback inside the window would likely show that the reform would have a positive impact on growth. Yet the economic costs of the policy would be shunted outside the window, and the distribution analysis would be inconsistent with the welfare analysis to the extent families are assumed to save money to pay for future offsetting fiscal policies. The in-the-window analytic results would be quite different from a more rigorous economic analysis and could offer a misleading impression that deficit-financed tax cuts offer a free lunch.

IV.1. Future policy changes can affect behavior today

Government borrowing cannot increase too rapidly for too long or the ability of the government to repay the debt will ultimately be called into question and, eventually, the government could default. Thus, any policy change that increases deficits too much, assuming the government does not start from a position of positive assets, will require future policy changes.

The emerging consensus approach to dynamic scoring defers necessary fiscal adjustments for many years in an effort to minimize the impact of the policy change on behavior inside the window. Deferral can reduce the direct effect of changes in fiscal policy on behavior today. For example, with reasonable discounting, changes in policy taking effect in 30 years will likely have modest effects on the investment decisions made today. Similarly, shifting of activity over time to take advantage of changes in policy will likely be modest in the first decade if future policy changes do not take place for several decades.

However, while it is possible to reduce the direct impact of policy changes on behavior inside the window by deferring them, in models with forward-looking agents it is difficult to eliminate the impact of tax changes on behavior inside the window entirely. Moreover, even when tax changes and spending changes can have similar impacts on behavior inside the window, they do not necessarily have the same behavior as simply not paying for the policy through fiscal policy changes. One could conceive of this scenario as comparing the outcome of a scenario with a future fiscal offset to the scenario in which the government started from a position of surplus and was simply discarding the revenues. The latter would

be expected to result in a more substantial increase in consumption today with corresponding implications for labor supply and other economic variables. The impact on consumption would be highly consequential for the impact on well-being.

IV.2. Economic effects can differ sharply inside and outside the window

Delaying a required fiscal adjustment to offset an unbalanced policy creates the potential for important dynamics involving the accumulation of government debt. The accumulation of government debt can drive changes in interest rates in the broader economy, affect the risk premium on government debt, and change the appetite for risk throughout the broader economy, among other potential impacts. A reduction in government debt can have similar effects with the opposite sign.

Of these, perhaps the most immediately consequential channel is that accumulating government debt can drive higher interest rates in the broader economy. Higher interest rates in the broader economy discourage capital accumulation, reducing output and labor productivity. Because these debt dynamics emerge only gradually over time, they can make the economic results quite different inside and outside the budget window.

Proposed legislation may also have sharply different impacts inside and outside the window if major provisions are sunset as is often the case for legislation passed through the reconciliation process, as the Tax Cuts and Jobs Act was. While JCT provided qualitative assessments of the second decade impact of the proposal, the analysis was much less detailed. At the same time, the pattern of effects in the first decade is strongly suggestive of important differences in the impact of the legislation over time.

By emphasizing economic outcomes in the first decade, the current approach to dynamic scoring highlights economic impacts in the window and downplays what happens outside the window. As a result, it creates a bias in favor of policies with a short run payoff and long run cost and a bias against policies with a short run cost and a long-run payoff.

One manifestation of this bias will be that policies that raise revenues to improve the fiscal position of the government and policies that raise revenues to finance new spending—assuming that the proposed spending has few supply-side effects—will appear much more similar in their impact on the economy than would necessarily be the case. This is an unsurprising result given that the basic framework of dynamic scores is to sever any link between taxes and spending.

IV.3. Dynamic scores can improperly rank policies

One argument for dynamic scoring is that it ranks policies appropriately in terms of their costs and economic effects while conventional scores do not. For example, dynamic scoring would show that not all tax policies with the same cost have the same effect on the economy. But because the costs of tax cuts are hidden outside the budget window and the costs of existing taxes are highlighted inside the window, dynamic scores typically suggest that all tax cuts are good and bigger tax cuts are better than smaller tax cuts.

Moreover, in addition to suggesting that larger tax cuts are generally better for the economy than smaller tax cuts, by shunting the cost of tax cuts outside the window, scores would typically show larger gains for policies that have more negative effects on the economy in the long run.

To the extent that sunsets or other policies cause economic activity to be shifted into the budget window, dynamic scores would also suggest that policies are generating positive increases in economic behavior even when what they do is shift economic activity across time.

IV.4. Critical attention is directed to the wrong assumptions

Perhaps the most important byproduct of delaying the fiscal offset is that it places great attention on the debt dynamics in the intervening period as a determinant of the economic consequences of a change in tax law. Proponents of tax cuts frequently argue that the economic benefits of tax cuts are understated because analysts assume the impact of higher deficits and debt on interest rates is overstated.

But, assuming some limit on government borrowing, the fundamental trade-off facing policymakers is not between taxes and debt, it is between taxes and spending or between different types of taxes. By developing an approach to scoring changes in tax policy that downplays the need for fiscal offsets, the attention shifted to the impact of changes in debt.

IV.5. Macroeconomic analyses are inconsistent with distribution tables

Static distribution tables provide an approximation to the gains and losses resulting from a tax bill to families across the income distribution. However, as is well known, distribution tables include gains made possible by increased borrowing and thus overstate the gains from a tax cut and include costs resulting from reduced borrowing and the overstate the costs from a tax increase.

Macroeconomic analyses make implicit assumptions about the use of tax cuts, particularly with regard to consumption and savings decisions. Perhaps the most clear-cut scenario is for models with Ricardian equivalence, in which taxpayers save a large fraction of the tax cut to finance the future fiscal offset. The size of this savings response would depend in part on the nature of the future fiscal adjustment.

Thus, the macroeconomic analysis supporting a dynamic score may suggest that the distribution estimates for the same policy are misleading.

V. Improving the presentation of macroeconomic analyses

Scores serve several purposes: to provide the information necessary to enforce targets established by the budget process, to inform policymakers about the costs of their policies, and to inform policymakers about the impact of their policies more generally. The current approach to dynamic scoring could be improved along all of these dimensions.

In the legislative process leading to the enactment of the Tax Cuts and Jobs Act, the congressional budget resolution established an envelope for the conventional score based on a loose justification about an alternative policy baseline and the revenue loss that would be offered by a dynamic score. (Of course, policymakers can agree to the budget resolution for a variety of reasons; inframarginal members of the caucus likely did not care about the revenue impact of the proposal.) However, when the dynamic scores produced by the Joint Committee on Taxation failed to be consistent with the loose justification, there was no parliamentary procedure to enforce it. Thus, JCT's actual dynamic score was replaced by a hypothetical dynamic score for purposes of the budget process.

As a presentational matter, the dynamic scores produced during consideration of the Tax Cuts and Jobs Act focused on the first ten-years and implicitly obscured the costs of the accumulated debt, thus

providing an incomplete picture of the economic impact of the proposals. The practical impact of the dynamic score for many policymakers and the press was to reinforce the idea that tax cuts—including deficit-financed tax cuts—generate growth and that tax cuts are a simple and direct mechanism for increasing growth. However, the underlying economic tradeoffs confronting policymakers are between different types of taxes or between taxes and spending. A longer-term focus would thus show that unsustainable tax cuts deliver temporary growth unless enacted alongside spending cuts.

The problems with the current approach to dynamic scoring identified in the prior section result in large part from the presentation of information. Thus, dynamic scores could be substantially improved by presenting the information underlying existing analyses in a manner that highlights the true economic tradeoffs underlying policy decisions.

This section first reviews the basic approach of optimal tax theory, which provides the motivation for this paper, and then recommends an alternative presentation that would offer a clearer exposition of the tradeoffs facing policymakers. The additional information highlighted in this alternative presentation would help policymakers understand what existing estimates imply would happen and would reduce the risk of simplistic inferences based on partial information.

V.1. Optimal tax theory

A lengthy literature in economics considers questions of optimal taxation. In contrast to scores, which attempt to evaluate the question of what would be, optimal tax analyses consider the question of what should be. However, there is not as large a gap between these two types of analyses as may at first seem given the questions posed. Any attempt to evaluate what policy should be will necessarily embed an answer to the question of how the economy works and thus what will happen in response to a change in tax policies.

However, a key difference between scoring exercises and optimal tax analyses is that models of optimal taxation frequently impose constraints on the choice of policy itself. The most important of these is generally that the policy raise the level of funds needed to finance the government's activities.[19] If the problem did not pose a constraint on government revenues, most analyses would conclude that the optimal policy is to provide unlimited transfers to the public (i.e. taxes would be negative). In effect, in this scenario, the government can successfully run a Ponzi scheme. If that is in fact the case, it stands to reason that the government should exploit this ability to provide benefits to the public.

Thus, optimal tax theory does not evaluate, for example, the merits of raising or lowering the statutory corporate tax rate in isolation. It considers the merits of lowering the statutory corporate tax rate and offsetting the cost by increasing the tax rate on dividends. This focus on comparing feasible policy alternatives that is missing from current dynamic scores.

[19] In most models, the revenue requirement could be set to zero with little impact on the results and—unless there are critical relationships between the source of government funds and the uses of government funds, incorporating an analysis of what the optimal level of public spending should be would have relatively little impact on the results as well.

V.2. A proposal for presenting the macroeconomic analysis supporting dynamic scores

The key ingredient needed in an improved presentation of dynamic scores is a recognition of the key tradeoffs in designing tax policy, which are the tradeoffs between different types of taxes, different patterns of taxation across time, or a change in both the level of taxation and the level of spending.

From the scoring perspective, it is untenable to exclude unbalanced policies from the set that can be analyzed. But in the logic of an optimal policy model, it is not informative to consider unbalanced policy changes. Policy action today frequently requires additional policy action in the future.

The emerging consensus approach to dynamic scoring has largely dealt with this issue by assuming future policy actions but at the same time not clearly reporting the impact of these policies. This paper recommends making such assumptions explicit and reporting longer term results that incorporate the effects of these policies.

At its core, the proposal is intended to restore the link between taxes and spending in economic analyses of tax legislation that is severed in the basic approach to dynamic scoring pursued today.

Concretely, the macroeconomic analysis underlying dynamic scores should be modified as follows:

First, analysts should make explicit assumptions about future policy offsets. As a default approach, analysts should evaluate two policy scenarios, one assuming a future fiscal adjustment resulting from a change in transfer spending and one assuming a future fiscal adjustment resulting from a change in taxes intended to proxy for a reversal of the proposed legislation. If the proposed legislation focuses on reducing statutory rates, the reversal would increase statutory rates. If the proposed legislation broadens the tax base, the reversal would narrow the tax base. And if the legislation does a combination of these, the reversal would do a similar combination. As under the emerging consensus approach, these fiscal adjustments should occur after a substantial lag. Implicit in this recommendation is the recommendation that an explicit statement of future policies should be made even for analysis relying on models that do not require such an assumption to solve and the analysis should be extended to include this period. When the results for the two approaches differ substantially, they should be prominently described as conditional on these policy assumptions.

Second, scorekeepers should report the results of their analysis extending beyond the time the fiscal adjustment takes effect. In other words, the analysis should show the economic effects of the policy for three distinct periods: the budget window, the period following the budget window and before the fiscal offset is enacted, and the period after the fiscal offset is enacted.

Third, a welfare analysis should be presented that is consistent with the macroeconomic analysis. A welfare analysis attempts to evaluate the impact of policy changes on economic well-being. As is well-known, GDP and other measures of output are not measures of economic well-being. Moreover, inferences based on the level of GDP can have implications that are starkly different from those based on welfare in scenarios where the increase in output is often generated by an increase in investment or labor supply which imposes a direct cost on people. Such is often the case for dynamic scores of tax legislation. It can also be the case when GDP deviates from GNP as additional investment is financed by foreign investors. Thus, a welfare analysis, standard in academic treatments of optimal tax analysis, would help policymakers and the public understand the economic effects of proposals for fundamental tax reform.

The nature of the welfare analysis, and particularly any implicit redistributive objectives, will fundamentally be dependent on the nature of the model used for the underlying macroeconomic analysis. As an initial approach, it would be reasonable to conduct the welfare analysis relying on the predicted economic effects for aggregates and using a stylized representative agent. However, the sensitivity of the results to different approaches is an issue that requires more substantial consideration.

Fourth, supplemental distribution analyses should be presented in conjunction with the traditional distribution analysis. JCT's traditional distribution analysis reports the aggregate and percent change in federal taxes by income class, aggregate federal taxes and the percent of total taxes under current law by income class, aggregate federal taxes and the percent of total taxes under the proposal, and the average tax rate under current law and the proposal. These provide a range of perspectives on the allocation of the change in taxes that would result from a proposal.

However, embedded in these distribution estimates is a gain or loss attributable to a change in overall federal revenues. Such a gain or loss will require a fiscal offset in the future or allow for increased future spending. Under current approaches to dynamic analysis, such a reversal is assumed but the impact generally not reported. As a potentially more accessible complement to the welfare analysis, a supplemental analysis should report the impact on families assuming a fiscal offset necessary to achieve dynamic revenue balance consistent with each of the two financing adjustments imposed above but imposed beginning in the first year after the budget window.[20]

The recommendations above apply to the macroeconomic analysis that underlies the dynamic score. The score itself, however, reports only the impacts on the government budget. In this case, the key results are revenue estimates for the budget window, the period following the budget window before the policies required for a fiscal adjustment take effect, and the period after the fiscal adjustment takes effect.

In addition to the substance of the analysis, a key question is what the findings in the analysis are that are to be highlighted. As noted above, JCT's analyses of the Tax Cuts and Jobs Act highlight the percent change in GDP on average over the budget window and the impact of the macroeconomic feedback on the conventional revenue estimates.

The key results to be highlighted should be short and longer-term impacts on GDP and revenues as well as the results of the welfare analysis. Highlighting longer term results for economic output and revenue feedback would clarify the results when policies improve or worsen the government's net fiscal position. Policies that achieve a temporary increase in GDP by borrowing from the future to reduce taxes would be shown to ultimately require spending cuts that make people worse off or tax increases that reduce GDP. Policies that result in a reduction in GDP from raising taxes to improve the government's fiscal position would be shown to allow future spending increases that make people better off or tax cuts that increase GDP. In either case, the nature of the reallocation of resources across time would be made explicit rather than implicit. While there is substantial uncertainty about longer-term projections, when shorter-term results depend on longer-term outcomes either as a result of future policy changes or directly, it is appropriate to describe those results on an even footing.

[20] As these analyses would fundamentally be assuming a different policy specification, they would amount to doing an analysis of a third and fourth specification of the underlying policy.

VI. Discussion

The proposal above would allow for a more informed discussion of policy changes. The current approach to dynamic scoring obscures the actual tradeoffs inherent in economic policy and highlights an apparent free lunch for tax cuts obtained by borrowing from the future while understating the benefits of deficit reduction. The proposal would present additional information that serves to highlight the actual tradeoffs available to policy-makers: (i) swapping one tax instrument for another, (ii) changing spending to finance a corresponding change in taxes, or (iii) changing taxes today and reversing that change in the future, and combinations of these options.

This section elaborates on the issues with existing approaches to scoring and the advantages of the proposed approach.

VI.1. Recognizing the role of future fiscal policies

Future fiscal policies can affect behavior today through their impact on spending and savings decisions, particularly in the case of models that assume agents with substantial foresight. In addition, depending on the precise timing of the fiscal adjustment and the nature of the adjustment, it could also affect certain types of behavior inside the window.

A primary advantage of the approach described above is that it recognizes the role of future fiscal policies and gives them a more central role in the analysis. Under the scoring perspective, this is perhaps undesirable as it makes it clear that the reported results are sensitive to assumptions about future policy. However, under the optimal tax perspective, this is desirable as it more clearly elucidates the actual economic tradeoffs that policymakers face. Moreover, to the extent that future policy action affects behavior today, whether in ways that are sensitive to the particular fiscal adjustment chosen or merely the existence of some adjustment, it is appropriate that the analysis highlight the role of future policy changes.

The most substantial disadvantage of this approach is that it results in two scores, an issue discussed in more detail below, and it makes clear that the results are sensitive to choices about future fiscal policy. It is understandable that nonpartisan scorekeepers would want to minimize the role of policy choices they make in assessing the impact of policies. As a means of addressing this concern, it may be appropriate for the same congressional rules or budget resolutions that require dynamic scores to also explicitly state how that both tax and spending offsets should be used and the nature of the tax and spending offsets. This would allow the scorekeepers to continue their role of providing rigorous nonpartisan analysis subject to guidelines created by partisan policymakers and eliminate the awkward position such scorekeepers are currently placed in by virtue of the lack of clear guidance.

VI.2. Short-run and long-run results

This paper has argued that the implicit lack of attention to the link between taxes and spending that is a significant problem in the creation and interpretation of dynamic scores and that the results for the next ten years can differ from those in the longer run. In contrast, Elmendorf (2015) argues that this is not a major concern:

> A further concern is that estimates of the macroeconomic effects of unsustainable policy changes would be distorted by the 10-year budget window. For example, suppose that a reduction in tax rates was estimated to raise output over the next decade but also to generate

> revenue losses that were unsustainable (because the increase in output was not large enough for the tax reduction to pay for itself). If those same rates were later raised to satisfy the government budget constraint, output would be estimated to be lower in the long run. In that scenario, dynamic scoring over the budget window would credit the proposal with raising output even though the ultimate effect on output would be negative.
>
> That scenario might occur, but it is not likely. First, a reduction in tax rates without an offsetting broadening of the tax base or reduction in spending might well be estimated to lower output within the 10-year budget window, depending on the specifics of the tax reduction…Second, the harmful effects of greater federal debt increase over time as debt compounds, while the favorable effects of lower tax rates generally do not, so any tax-rate reductions that were estimated to raise output throughout the first decade after enactment would be less likely to be estimated to do so in the second decade…Despite these points, suppose that a reduction in tax rates was, in fact, estimated to raise output over the next few decades but also generate revenue losses that were unsustainable. The unsustainable nature of the rate reduction would be shown by the estimated effect of the proposal on federal debt, so the need to make further policy changes to offset the budgetary losses would be quite apparent…Moreover, the estimated macroeconomic effects of the rate reduction would not be the only aspect of the budget estimate that could be misleading; the estimated nonmacroeconomic effects on the budget could be misleading as well, as would any distributional analysis or other analysis based on the rate reduction. In any event, it is not tenable for CBO and JCT to ignore the policy changes included in a legislative proposal even if those changes are not, by themselves, sustainable.

The recent JCT macroeconomic analysis of the Tax Cuts and Jobs Act provides a useful benchmark against which to assess these claims. First, while the analysis would seem to imply that the proposal is unsustainable in the long-term, this conclusion is not emphasized or stated clearly. In addition, while there is no explicit statement of GDP in the final year of the window, it is likely (given the pattern of revenues) that JCT estimates the proposal will increase GDP in the final year of the window as well. Thus, it would appear to be the case of an unsustainable fiscal policy increasing GDP in the window.

Second, while it is the case that the effects of debt compound over time in JCT's analysis and JCT does express uncertainty as to the longer-term economic effects: "Because of competing incentive effects over time, it is uncertain whether GDP and possibly revenue feedback will continue to be higher than under present law by that time," this statement is caveated with substantially more uncertainty than the short-run analysis. Thus, JCT does not place the short-term results and the longer-term results on a comparable footing.

Thus, the reasons Elmendorf offers to suggest that an analysis might show an unsustainable change in fiscal policy in an unattractive light are not compelling in practice. Moreover, JCT's written presentation further focuses on the effects in the budget window. In the summary, the organization offers a qualitative statement of the second decade effects in the summary and ignores the long run effects entirely. This short run focus is emphasized even more in most public commentary and press coverage of the scores, suggesting that the analysis is not offering a balanced picture of the short and long-run consequences.

The JCT analyses of the Tax Cuts and Jobs Act thus offer reason to be concerned about the treatment of unsustainable policies in macroeconomic analyses. Yet there is a certain parallel between Elmendorf's defense of dynamic scoring in the excerpt above and the proposal made in this paper as both suggest that more careful attention paid to the presentation of information associated with the score could help protect against misleading results. At present, however, the score emphasizes the short run even when such outcomes could be quite different from the long run.

VI.3. Ranking policies appropriately

One prominent argument in favor of adopting dynamic scoring is that conventional scoring incorrectly ranks policies and that dynamic scoring would correctly rank policies. For example, Holtz-Eakin (2016) argues

> Scoring is not forecasting. Scoring is ranking proposals. And in the particular small world in which this is important, it's typically legislative proposals in the House and the Senate of the United States of America and the idea is to be able to give policymakers the ability to rank them and thus make decisions about which ones are the ones they would like to enact into law.
>
> For that reason, the most important thing about scoring is doing it the same for all the proposals. It's desirable if at all possible to have them be accurate, but it's more important to make sure they get ranked. And you'll notice that the scoring process has lots of things built into it that make sure that they're actually kind of inaccurate. For example, the CBO does a January economic projection and then in March that projection is locked in stone and combined with the current law baseline projections of revenue and spending and that becomes the quote scoring baseline. If there's a recession in July and you're passing legislation in august it's still the scoring baseline, you haven't updated it, you're not forecasting, and so…there's often a sort of tendency to veer off into discussions about how that's not right and that's going to be inaccurate…that may be true, but remember the top criteria should be…to get things ranked right.

In many ways, the focus on the comparison between proposals in this excerpt is consistent with the optimal tax perspective motivating the alternative proposal presented in this paper. Economists would typically rather compare between two alternative options than make a single representation of what one proposal would do. In part, this is because a comparison between two alternatives is often a better posed question.

However, any ranking of policies should necessarily be based on their substantive merits, and such a ranking requires a clear and accurate assessment of their impacts. And, as noted above, current approaches to dynamic scoring, in large part because of their short run focus, tend to emphasize the size of a reduction in taxes rather than design of the proposal. Indeed, it is difficult to see the JCT estimates for the Tax Cuts and Jobs Act as being widely used in a comparative manner. While it is true that some observers compared the GDP growth estimates for the three versions of the bill, the general focus was on the comparison of GDP under the proposal to GDP under current law.

The longer-run focus recommended in this analysis would likely improve the relative rankings of policies by focusing on the design of the tax system and the alignment of spending and revenues rather than the size of the tax cut.

VI.4. Coherent welfare analysis

Welfare analysis in the sense the term is used by economists—meaning the impact of policy changes on economic well-being—does not play a central role under that label in the analysis of potential tax changes. However, distribution analysis does play a central role, and distribution analyses are closely linked to welfare analyses. In the special case of revenue-neutral tax reform, static distribution analyses provide an approximation to the change in welfare that will result from a change in tax policy (Leiserson 2017).

There are several sources of potential tension between the distribution analysis and the macroeconomic analysis implicit in dynamic scores. Placing greater emphasis on welfare changes in macroeconomic analysis can help reconcile and clarify these issues.

First, most static distribution analyses use an approximation for the incidence of the corporate income tax (and, potentially, taxes on income from pass-through businesses). However, as the macroeconomic models and distribution analysis used by most scorekeeping organizations have evolved separately, the two are not necessarily consistent with each other.

Second, the macroeconomic analysis will generally imply particular assumptions about behavior that may be at odds with the assumptions underlying the static distribution table and the public's interpretation. For example, in models that feature Ricardian equivalence, or a near approximation thereof, individuals will be assumed to save a substantial fraction of the tax cut to offset the future fiscal adjustment. This increased savings will be implicit in analytic results for consumption and similar variables, but will not be captured in distribution analyses. These responses bear a certain relationship to distribution analyses that incorporate financing as they provide a partial measure of the future costs. Focusing on welfare thus offers a path to addressing the potential for distribution tables to provide a misleading depiction of the effects of unbalanced proposals.

Third, the results of macroeconomic analyses are prone to misinterpretation, particularly as they relate to changes in employment. In supply-determined models, an increase in employment is generally a cost borne by workers, not a benefit. By shifting focus from changes in GDP to changes in welfare, the analysis would be more informative about measures that are directly relevant to assessing the merits of a policy.

Thus, a presentation of macroeconomic analysis that places greater emphasis on welfare as recommended above offers several potential improvements in the quality of the information provided to policymakers.

VI.5. Recognizing the link between revenues and spending

The primary purpose of taxation is to raise revenues for the government. Indeed, in models used for dynamic scoring there are typically no efficiency or growth-enhancing reasons to have taxes. Taxes are inherently undesirable. As discussed above, the policy assumptions embedded in the emerging consensus approach to dynamic scoring serve to largely sever any link between taxes and spending. As a result, raising taxes offers few benefits and potentially high costs. It is difficult to have a sensible conversation about tax policy in a framework in which revenues have so little value. The alternative

presentation recommended above, implicitly restores the value of taxes as manifest in the potential for higher public spending.[21]

VI.6. How will policymakers deal with multiple scores?

Only one score can be produced for purposes of dealing with the congressional budget process, but multiple scores can be produced for any other purpose. The experience with the Tax Cuts and Jobs Act suggests that the integration of dynamic scores into the congressional budget process is far from complete and politicians may in fact prefer that it not be completed. Thus, the practical argument that only a single score can be produced—at least at present—does not seem to have much weight.

As a result, the value of showcasing multiple scores seems quite valuable as it would provide more relevant information at little cost. Policymakers would certainly be free to choose the score they prefer (and the associated policy) in talking about the impacts their policies would have on the economy, but they would not be able to exploit policy ambiguity as they are at present where the score assumes future policy actions but politicians do not acknowledge them.

If, as suggested above, the future policy actions scoring entities should assume were established by policymakers in the same rules that establish a requirement for dynamic scores in the first place, policymakers could likewise establish a rule for how multiples scores should be incorporated into the budget process. For example, budgeting could be conducted using a particular specified policy option or a weighted average of multiple options.

VI.7. Are policies scored twice?

One issue in assuming a future fiscal adjustment is that any policy enacted by Congress today necessarily involves future legislation as well. Thus, it raises the question when Congress considers future legislation as to whether the impacts of such legislation have already been analyzed.

For example, JCT's scores of the Tax Cuts and Jobs Act assume a future fiscal adjustment to offset the debt impact. The precise details of that adjustment vary across models within the score. Thus, if Congress were to consider a future fiscal adjustment, it would be natural to ponder whether the macroeconomic effects of the legislation have already been incorporated into the analysis. Of course, in practice, since there is no single policy adjustment in JCT's scores and the assumed policy response may not align with the precise contours of JCT's assumptions, the answer would be no. Moreover, incorporating future fiscal adjustments into CBO's baseline and then carrying them forward in time as time passes would lead to bizarre results and there is no reason to think CBO would do so nor that this would be good idea.

Yet it is not unreasonable for proponents of policies to reduce the fiscal gap in the future to argue that such policies were assumed in the score of the Tax Cuts and Jobs Act, and thus the policies to reduce the fiscal gap are being held to a higher standard when they are enacted. A similar argument could be made that if policymakers pursue deficit-reducing legislation in the short run, these effects are overstated

[21] Recognizing the link between revenues and spending does raise the question of how results would change in there is no limit on government's ability to borrow. If this were the case, the recommendations of optimal tax theory would change radically. Most canonical models would suggest that the government should provide infinite sums of money to every person yielding infinite utility. I set aside the consideration of how dynamic scoring should be conducted if there is truly no binding constraint on government spending or borrowing for purposes of this paper.

because they were implicitly required in the long run in the score of the Tax Cuts and Jobs Act, and the analysis simply proceeded under an incorrect assumption about timing.

This issue is inherent in any analysis of unbalanced policies and the proposal of this paper does not improve on this front.

VI.8. Comparison with academic approaches

The information presented in dynamic scores as produced by JCT and other independent analysts differs sharply from optimal tax analyses prepared by the academic community.

As would be expected given their role in the budget process, the most detailed results in JCT's dynamic scores of the Tax Cuts and Jobs Act conference report are those for the revenue feedback. JCT provides year-by-year estimates of the revenue loss according to the conventional score and the change resulting from the dynamic analysis, and finally the two combined. In addition to these estimates, JCT provides a percent change in average GDP during the budget window, some information about the change in GDP before and after the sunset of certain tax provisions, a percent change in the capital stock, a percent change in employment and consumption, and qualitative statements about second and third decade effects.

Based on the summary in the introduction, JCT considers the most important elements of the analysis the first decade GDP effect and the first decade revenue feedback. A qualitative statement about second decade effects appears in the introduction and no information about the third decade or later appears.

As a benchmark, Altig et al. (2001) is a highly cited academic paper simulating fundamental tax reforms. Consistent with most academic work, it considers reforms that maintain a fixed level of revenues, so the revenue analysis is not central. In contrast, the analysis focuses on additional information about the transition and the pace of transition. In addition, the academic analysis provides a more robust welfare analysis that illustrates heterogeneous effects on different birth cohorts and income groups. These elements of the macroeconomic analysis are completely absent from the JCT analysis.

The proposal of this paper would thus move the substance of JCT's analysis in the direction of the academic optimal tax analysis. It would place more weight on welfare analysis and focus on longer-term outcomes.

VII. Alternative approaches to conducting macroeconomic analyses

The proposed approach above is not the only alternative to the emerging consensus approach to dynamic scoring. It focuses on a readily implementable option that would leave current dynamic scores largely unchanged inside the window. This section outlines three additional alternative approaches and briefly considers their merits.

VII.1. Model fiscal crises that result from unsustainable policies

An alternative approach to dealing with the enactment of unbalanced fiscal policies would be to model the fiscal crisis resulting from unsustainable policies in a reduced form way. For example, the current baseline could be replaced with one in which trust fund payments stop when they are exhausted, discretionary spending is limited to revenues when debt reaches some critical threshold, and mandatory spending is reduced across the board in some ad hoc manner.

The major disadvantage of this approach is that the analysis of fiscal crises would require ad hoc assumptions about policy and be subject to substantial uncertainty in the economic modeling. While the trust fund approach has a clear motivation, the maximum debt threshold is arbitrary and the treatment of mandatory spending would be inconsistent with the law. Whether a future administration confronted with this reality would engage in this way is unclear and political considerations would come into play long before this outcome is reached. However, at the same time, the scoring perspective prioritizes the evaluation of policies under the assumption that they remain in place without future action and thus this approach is perhaps the most faithful to that perspective.

The advantage of this approach is that it would provide a direct link between revenues and spending by virtue of the spending cuts or increases that would result from an improvement or deterioration in the government's fiscal position. It would also appropriately highlight that the revenue consequences ultimately manifest in a long-run impact on government spending.

VII.2. Model a more realistic depiction of current law

As a compromise between modeling a fiscal crisis and the current approach, which severs the link between taxes and programmatic spending, macroeconomic analysis could provide a more direct modeling of select current law policies. For example, rather than using the baseline, under which spending from major trust funds continues even after they are exhausted, the analysis could assume that spending stops when trust funds are exhausted.

While this would establish only a partial link between spending and revenues, it has the potential to have a substantial effect on the result. Payroll taxes account for roughly one-third of federal revenues and thus would be expected to mediate a substantial portion of any revenue feedback in a dynamic score. Feedback from increased economic activity would thus directly lead to increased spending and obviate the impacts on the federal budget.

An advantage of this approach relative to the approach of attempting to model a fiscal crisis in a reduced form way is that the assumptions necessary to model the behavior of the trust fund are easier to define and implement than those governing a fiscal crisis.

However, this approach still suffers from several drawbacks. Perhaps most importantly, it would retain the problems of assuming away any relationship between general fund revenues and spending. Moreover, the practical effect of trust fund exhaustion is subject to substantial legal ambiguity and so assuming a particular meaning of exhaustion would not be without controversy.

VII.3. Rely on the conventional score

Given the problems embedded in current approaches to dynamic scoring and the difficulties in defining a superior approach, a third option would be to rely on the conventional score either in all cases or for the analysis of any policy that results in unbalanced fiscal changes, and return macroeconomic analyses to a supplemental role.

In the interest of brevity, this paper does not repeat the arguments in favor of or against dynamic scores in general to the extent that they are not implicit in the preceding discussion, other than to note this as an option.

VIII. Conclusion

Policymakers and the press are placing increasing attention on dynamic scores of tax legislation. Dynamic scores differ from conventional scores in that they allow changes in public policy to affect macroeconomic aggregates such as GDP, labor supply, and the capital stock. However, when policymakers enact policies that do not respect a balanced budget constraint, they create a challenge: how will the resulting improvement or deterioration of the government's fiscal position be addressed?

The emerging consensus approach to dynamic scoring implicitly adopts two new assumptions. First, the macroeconomic analysis supporting dynamic estimates assumes future policy changes sufficient to address the fiscal imbalances that exist in CBO's current-law baseline. These changes are assumed to take effect after the period for which economic results are reported. Second, in many but not all cases, the analysis assumes additional future policy changes that offset any change in the government's present value fiscal position that the proposed legislation would cause, again taking effect after the period for which results are reported.

While this approach offers obvious advantages in facilitating the creation of a ten-year score that appears faithful to the substance of a legislative proposal, it obscures the key economic tradeoffs. Indeed, in some ways it amounts to analyzing the impact of tax policy on the economy under the assumption that revenues are largely unnecessary and thus have little value. Thus, dynamic scores can offer an apparent free lunch, where tax cuts offer higher GDP growth with little economic cost and no difficult tradeoffs.

This paper outlines a proposal for an alternative presentation of dynamic scores that would leave the existing scores largely unaffected within the window, but would clarify the underlying economic tradeoffs. It proposes reporting results for two potential future policy responses to highlight the sensitivity of economic estimates to future policies and emphasizes short and long-term outcomes to provide a more complete picture of the findings as well as incorporating greater attention to economically grounded measures of well-being rather than output in isolation.

References

Altig, David, Alan J. Auerbach, Laurence J. Kotlikoff, Kent A. Smetters, and Jan Walliser. 2001. Simulating Fundamental Tax Reform in the United States. American Economic Review. 91 (3): 574-595.

CBO. 2012. Fair-Value Accounting for Federal Credit Programs. Available at https://www.cbo.gov/publication/43027.

CBO. 2017a. Frequently asked questions about CBO cost estimates. https://www.cbo.gov/about/products/ce-faq.

CBO. 2017b. An Update to the Budget and Economic Outlook: 2017 to 2027. Available at https://www.cbo.gov/publication/52801.

CBO. 2017c. The 2017 Long-Term Budget Outlook. Available at https://www.cbo.gov/publication/52480.

Elmendorf, Douglas W. 2015. "Dynamic Scoring": Why and How to Include Macroeconomic Effects in Budget Estimates for Legislative Proposals. Brookings Papers on Economic Activity. Fall 2015. 91-149.

Holtz-Eakin, Douglas. 2016. Remarks at "Dynamic Scoring of Tax Proposals". Tax Policy Center, September 16, 2016. Available at http://www.taxpolicycenter.org/event/dynamic-scoring-presidential-candidates-tax-proposals2.

Holtzblatt, Janet and Jamie McGuire. 2016. Factors Affecting Revenue Estimates of Tax Compliance Proposals. CBO Working Paper 2016-05 and JCX-90-16. Available at https://www.cbo.gov/publication/52199.

Joint Committee on Taxation. 2014. Macroeconomic Analysis of the Tax Reform Act of 2014.

Joint Committee on Taxation. 2017a. Macroeconomic Analysis of the Conference Agreement for H.R. 1, The "Tax Cuts and Jobs Act".

Joint Committee on Taxation. 2017b. Macroeconomic Analysis of the "Tax Cuts and Jobs Act" as Passed by the House of Representatives on November 16, 2017.

Joint Committee on Taxation. 2017c. Macroeconomic Analysis of the "Tax Cuts and Jobs Act" as Ordered Reported by the Senate Committee on Finance on November 16, 2017.

Joint Committee on Taxation. 2017d. Estimated Budget Effects of the Conference Agreement for H.R. 1, the "Tax Cuts and Jobs Act".

Joint Committee on Taxation. 2017e. Joint Committee Revenue Estimation Process. https://www.jct.gov/about-us/revenue-estimating.html.

Kamin, David. 2014. Accounting for Risk in the Federal Budget. Indiana Law Journal. 88. 723-772.

Leiserson, Greg. 2017. If U.S. tax reform delivers equitable growth, a distribution table will show it. Washington Center for Equitable Growth. Available at http://equitablegrowth.org/research-analysis/if-u-s-tax-reform-delivers-equitable-growth-a-distribution-table-will-show-it/.

Marron, Donald. 2014. A Better Way to Budget for Federal Lending Programs. Tax Policy Center. Available at http://www.taxpolicycenter.org/publications/better-way-budget-federal-lending-programs.

Morton, William R. and Wayne Liou. 2017. Social Security: What Would Happen if the Trust Funds Run Out? Congressional Research Service. RL33514.

Page, Benjamin R., Joseph Rosenberg, James R. Nunns, Jeffrey Rohaly, and Daniel Berger. 2017. Macroeconomic Analysis of the Tax Cuts and Jobs Act. Tax Policy Center. Available at http://www.taxpolicycenter.org/publications/macroeconomic-analysis-tax-cuts-and-jobs-act.

Page, Benjamin R., Joseph Rosenberg, James R. Nunns, Jeffrey Rohaly, and Daniel Berger. 2017. Macroeconomic Analysis of the Tax Cuts and Jobs Act as Passed by the Senate. Tax Policy Center. Available at http://www.taxpolicycenter.org/publications/macroeconomic-analysis-tax-cuts-and-jobs-act-passed-senate.

Page, Benjamin R., Joseph Rosenberg, James R. Nunns, Jeffrey Rohaly, and Daniel Berger. 2017. Macroeconomic Analysis of the Tax Cuts and Jobs Act as Passed by the House of Representatives. Tax Policy Center. Available at http://www.taxpolicycenter.org/publications/macroeconomic-analysis-tax-cuts-and-jobs-act-passed-house-representatives.

Page, Benjamin R. and Kent Smetters. 2017. Dynamic Analysis of Tax Plans: An Update. Available at http://www.taxpolicycenter.org/publications/dynamic-analysis-tax-plans-update.

Penn Wharton Budget Model. 2017. The Tax Cuts and Jobs Act, as Reported by Conference Committee (12/15/17): Static and Dynamic Effects on the Budget and the Economy. Available at http://budgetmodel.wharton.upenn.edu/issues/2017/12/18/the-tax-cuts-and-jobs-act-reported-by-conference-committee-121517-preliminary-static-and-dynamic-effects-on-the-budget-and-the-economy.

Penn Wharton Budget Model. 2017. The Senate Tax Cuts and Jobs Act, as Passed by Senate (12/2/17): Static and Dynamic Effects on the Budget and the Economy. Available at http://budgetmodel.wharton.upenn.edu/issues/2017/12/8/the-senate-tax-cuts-and-jobs-act-as-passed-by-senate-12217-static-and-dynamic-effects-on-the-budget-and-the-economy.

Penn Wharton Budget Model. 2017. The House Tax Cuts and Jobs Act, Amended (11/9/17): The Dynamic Effect on the Budget and the Economy. Available at http://budgetmodel.wharton.upenn.edu/issues/2017/11/13/the-house-tax-cuts-and-jobs-act-amended-11917-the-dynamic-effect-on-the-budget-and-the-economy.

Tax Foundation. 2017. Updated Details and Analysis of the 2017 House Tax Cuts and Jobs Act. Available at https://taxfoundation.org/2017-tax-cuts-jobs-act-analysis/.

Tax Foundation. 2017. Preliminary Details and Analysis of the Senate's 2017 Tax Cuts and Jobs Act. Available at https://taxfoundation.org/details-analysis-2017-senate-tax-cuts-and-jobs-act/.

Tax Foundation. 2017. Preliminary Details and Analysis of the Tax Cuts and Jobs Act. Available at https://taxfoundation.org/final-tax-cuts-and-jobs-act-details-analysis/.

Tax Policy Center Staff. 2017. Distributional Analysis of the Conference Agreement for the Tax Cuts and Jobs Act. Available at http://www.taxpolicycenter.org/publications/distributional-analysis-conference-agreement-tax-cuts-and-jobs-act.

U.S. Treasury. 2015. Treasury's Distribution Methodology and Results. Available at https://www.treasury.gov/resource-center/tax-policy/tax-analysis/Documents/Summary-of-Treasurys-Distribution-Analysis.pdf.