



Despite

<The Dynamics of Scoring>

A Congressional Tale

by Rudolph G. Penner

MICHAEL SIPE

what you think,

the most important form of scoring in Congress has nothing to do with sex.

Every bill reported out of committee must be scored – that is, its potential impact on government revenues and outlays estimated. The Joint Committee on Taxation considers bills that affect major revenue sources, while the Congressional Budget Office scores bills' spending impact along with revenues from less important sources like import tariffs. And though the process seems dull on first consideration – I know, you're already wondering whether to flip to the next article – budget scoring can make a world of difference in how we are taxed and how much is spent.

Hence, the analysts who labor deep in the background to produce scores are extremely powerful. A prime example: the recent tax cut.

The Senate had decreed that changes in tax policies would not be allowed to increase the budget deficit by more than \$350 billion over the next 10 years. To meet the target, the bill was trimmed to fit, with phase-outs, phase-ins and sunsets that would make an Enron accountant blush.

Although the contortions in the tax law were particularly outrageous this year, it is common to rejigger policies to produce scores that meet Congressional spending or revenue targets, or that satisfy



SCORING

some budget rule. Indeed, if a policy initiative cannot be shoehorned into a preagreed numerical limit, it is likely to die.

As a result, those favoring a bill press the analysts to assign as low a cost as possible to it, while opponents like to see its costs exaggerated. Lobbyists watch the process carefully, hoping to find analytic errors if a score goes against them – or to persuade analysts to adopt assumptions favorable to their cause before the analysis begins. During my tenure as the CBO's director, criticism of our economic forecasts and estimates of budget aggregates was tempered. But more than once I had to hold the phone away from my ear when listening to a member of Congress who claimed that our scoring methods had destroyed some pet initiative that, but for our stupidity, would save the nation.



In spite of these pressures, the process proceeds with remarkable integrity. CBO and Joint Committee on Taxation analysts almost always work objectively, regardless of the political consequences.

Indeed, because scoring can be so controversial, the CBO has laid down explicit rules defining what's right and what's wrong. This ensures that analysts will score on a consistent basis – and that the office will be protected against charges of bias. But most of the rules do contain arbitrary elements. Some are prag-

matically designed to ease the task, so that scores can be produced quickly. And some are affected by in-house politics. For example, scoring rarely takes account of the effect of one committee's actions on the costs of programs in another committee's jurisdiction.

One rule particularly annoys supply-siders and other fervent tax cutters: scorers do not take account of some indirect macroeconomic effects of tax changes on revenues or spending. In particular, supply-siders argue that if the positive effects of tax cuts on economic growth were factored in, the revenue losses would be much lower than conventionally scored.

Many outsiders charge that the failure to account for indirect macroeconomic effects implies that scorers pay no attention to the impact of policy on human behavior. Critics refer to conventional scores as "static" – and

thus inherently inferior to dynamic estimates. Actually, scorers do take account of *microeconomic* behavioral responses to policy changes. For example, the revenue estimates associated with an increase in the gasoline tax would take account of the likely decline in gasoline consumption when the price goes up. Changes in the capital gains tax rate are assumed to affect the timing and volume of asset sales. And so forth.

However, analysts do stop at the water's edge; these micro responses are implausibly assumed to have no effect on total economic output, inflation, unemployment or other macroeconomic variables that feed back on government spending and revenue. If

RUDOLPH PENNER is a senior fellow at the Urban Institute and a former director of the Congressional Budget Office.

MICHAEL SIPE

I often had to hold the phone away from my ear when listening to a member of Congress who claimed that our scoring methods had destroyed some pet initiative that, but for our stupidity, would save the nation.

pressed, the analysts may say they do not bother because the Federal Reserve will offset relevant macro effects. But the real reason that macro responses are not considered is purely pragmatic: dynamic scoring is just too time-consuming and too hard to coordinate.

Analysts must routinely contend with complex provisions that affect both revenue and spending. Dozens of analysts working both at the Joint Committee on Taxation and the CBO may be involved in scoring one bill, each focusing on a subset of its provisions. What's more, they sometimes have only hours to complete their task.

Assumptions about GDP growth and inflation have a crucial impact on estimates of the revenue effects of tax changes. The macro assumptions are generally provided by CBO's annual economic forecast. This forecast is completely dynamic in that it considers the consequences of all behavioral responses to the policies in effect when the forecast is made.

Think, too, of the dimensions of the coordination problem. Ideally, any change in the macro forecast made by analyst A would affect the estimates of analysts working on most other provisions of the legislation. If, after A makes an estimate, and B decides that her provisions affect the macro economy, too, A should go back and redo his own estimate to maintain internal consistency. And that's only the beginning; changes in the macroeconomic environment should change cost and rev-

enue estimates throughout the budget – even for programs not directly affected by the policy change. In other words, the whole budget base line should be re-estimated every time a significant bill passes.

Even if that were practical it would drive Congress crazy, because ever-changing cost estimates would disrupt negotiations and force the revision of deals that had been made before the estimates changed. CBO used to provide a baseline in the spring and update it in the summer. But even that arrangement proved cumbersome. So, in 1985, Congress directed the agency to stick with its spring baseline all year for purposes of scoring, regardless of how inaccurate it became.

The complexities of dynamic scoring might still be manageable if Congress allowed analysts a couple of weeks to do their work. But often CBO must come up with estimates the day after they are requested, and that makes adequate dynamic scoring logistically impossible.

It is obviously absurd to argue that the Congress should therefore be shielded from knowledge of the effect of major policy changes on economic growth, inflation and unemployment. When requested, the CBO has done comprehensive dynamic analyses of important hypotheticals like capital gains tax changes,



SCORING

but the process takes months and is thus hard to link to specific legislation. Remarkably, Congress has seldom asked for such studies.

When the Republicans took control of both houses of Congress and supply-siders gained leadership roles, the pressure grew to do some sort of dynamic scoring of tax bills. The House first passed a rule saying that dynamic analyses could be requested for specific tax bills. But more recently, it ordered the Joint Committee on Taxation to estimate the effects of *all* tax legislation on economic growth, or to inform the House why such estimates were not practical. The first full-blown dynamic analysis by the Joint Committee on Taxation was performed for the House tax bill of 2003.

Note that the rule states dynamic analysis is for *informational purposes* only. That is, the analysis will not be used to enforce numerical targets created in the process of hammering out a budget. A purist would say, therefore, that it is not really dynamic scoring, but just analysis that complements scoring-as-usual. This distinction may seem a nitpick. However, it has important implications for how the analysis is done – a point I shall come back to.

To improve the quality of its dynamic analysis and to foster consensus, the Joint Committee created an advisory body of economists from both political parties and with diverse ideologies. But that is a daunting goal, because estimates of the effects of tax changes depend critically on assumptions about the

behavioral responses of both investors and consumers on the one hand, and the Federal Reserve on the other.

To deal with such thorny issues, the House rule gives the scorers a luxury that they do not usually enjoy. Because the analysis is not an official scoring exercise, analysts can offer a range of estimates rather than the single number demanded in traditional scoring. For example, the JCT provided different results for different assumptions regarding the Fed reaction to the tax-cut initiative and analyzed the policy change using different models of the economy. In one variant, the Fed raised interest rates, thus washing out a large portion of the demand-side stimulus associated with lower taxes. In others, the Fed created enough liquidity to accommodate differing portions of the demand effects.

The assumptions in the economic models used by the JCT differed in a number of respects, but one of the key factors was the extent to which households and businesses reformulated their own forecasts in response to changes in government fiscal policy and how they responded to those forecasts. Economic actors may be totally myopic, not worrying about the affect of federal tax changes on variables ranging from interest rates to inflation. On the other hand, people may recognize that a tax cut today probably means higher taxes or lower government spending tomorrow to counter the effect on the deficit.

If higher taxes are indeed expected down the road, people may work harder now to take

A whole branch of economic theory has been built on the idea that fiscal and monetary policy have little or no net effect because producers' and consumers' knowledge of the long-term consequences leads them to behave in ways that fully offset the impact.

advantage of temporarily low rates and enjoy more leisure later when rates are again raised. They may also save more now to prepare for paying future tax increases. Indeed, a whole branch of economic theory has been built on the idea that fiscal and monetary policy have little or no net effect because producers' and consumers' knowledge of the long-term consequences leads them to behave in ways that fully offset the impact.

In all the models used by the JCT, the tax-rate cuts had both positive and negative effects on economic growth. They increased productive capacity because lower tax rates increased incentives to work, save and invest. Moreover, this supply-side effect was complemented by a positive effect on demand that varied in importance, depending on the assumptions.

On the other hand, the increased deficit that resulted from tax cuts reduced national savings and capital formation. Tax cuts taking the form of tax credits also had a negative impact – the resulting income is like a no-strings-attached gift, and thus provides no extra incentive to work. In fact, research suggests that lump-sum income gains reduce work effort – especially among second (and third) earners in a household.

The Joint Committee on Taxation estimated that in the first five years the positive effects of the tax cut would dominate. In later years, the negative effects of the deficit be-

came relatively more important as the resulting slowdown in capital formation begins to bite. And after all the brouhaha over giving supply-siders a new peg to hang their hats, dynamic analysis made little difference, on balance. The House tax bill was projected to raise real GDP by something on the order of 0.2-0.9 percent between 2003 and 2008, and lower it by a maximum 0.2 percent between 2009 and 2013.

The reduction in the estimated loss of revenue resulting from dynamic analysis of the tax cut was similarly disappointing to proponents. For the 2003-2008 period, the budget savings ranged from 5.8 to 27.5 percent, depending on the model and assumptions used. Over the entire 2003-2013 period the savings ranged from 2.6 to 23.4 percent.

In March 2003, the CBO provided a dynamic analysis of the Administration's whole budget, including spending as well as tax proposals. Although the analysis also provided a wide range of results, the average differed little from zero. The estimated cumulative deficit from 2004 to 2008 ranged from \$1.24 trillion to \$1.04 trillion, compared to a traditionally computed estimate of \$1.16 trillion.

Being able to provide a range of estimates greatly eases the analyst's task. I am



It is something of a game – one that brings some discipline to the budget process when it is played well. In recent years, though, the game has not been played well.

sure they would like to provide a range for static scores as well, but Congressional rules require a single number. Thus, when confronted with a range of estimates for, say, the demand response to an increase in the gas tax, analysts must choose just one.

There are differences between what was done by the JCT and the CBO and a comprehensive dynamic analysis. The agencies did not go back and re-estimate the revenue effect of specific provisions of the bill, given different macro results. They also did not provide year-by-year estimates, and they made no effort to recompute the whole budget baseline for different macroeconomic responses. All that made the task much easier. It does mean, however, that the results are not quite internally consistent.

How can we expect tax-cut enthusiasts to respond to this less-than-meets-the-eye result? Their first line of attack will be to argue that the JCT and CBO analysts used the wrong models and made the wrong assumptions. They will argue for assumptions that enhance the positive supply effect of tax-rate cuts and minimize or delay the negative impact of deficit increases. However, if the scorers do not use assumptions consistent with the precepts of mainstream economics, they will lose credibility. And credibility is their only asset.

The center of the economics profession may, of course, move to accommodate the supply-siders. If dynamic analysis had been performed 30 years ago, the focus would have been on demand effects, while the supply side

would have largely been ignored. Now the worm has turned. Economists seem more confident of supply-side effects than they are of the impact on demand. This drift toward the right may well continue, but such changes generally proceed glacially.

I suspect that the Congress may tire of dynamic analysis long before its results become more favorable to tax cutters – if they ever do. And Congress will tire of the effort more quickly if the JCT pursues the task aggressively. After all, a large portion of the tax cuts considered by Congress are clearly hostile to economic growth. The cuts often favor one industry or group over another, thereby distorting the allocation of labor and capital, and making the economy less productive. If I'm right, the rule requiring dynamic analysis is likely to quietly disappear.

Although the contretemps over dynamic scoring has been educational, it has diverted attention from much more serious problems with the current scoring process. For example, important distortions result from using a 10-year time horizon that scores the impact for the 11th year as zilch. These distortions are particularly serious when considering changes in the taxation of capital income.

Take the case of individual retirement accounts that give owners some relief from taxation. Under reasonable assumptions, the present value of the revenue loss associated with traditional IRAs (where deposits are deducted from current taxable income), is the same as the loss from Roth IRAs, where deposits are not deductible but the income

generated over the life of the account is never taxed. But the costs are all backloaded; for the first 10 years, Roth IRAs cost the Treasury less – which explains why Congress created them. Conversely, an increase in the generosity of investment depreciation allowances, which defer the payment of taxes, reduces revenues more in the short run than in the long run. In all such cases, we should be willing to consider scores consisting of present values computed over very long time horizons.

The rules of scoring also produce very misleading results in some, albeit rare, odd cases. For example, scores do not take account of changes in the interest owed each year on the public debt. Thus, although a tax cut increases the deficit and so adds to the interest burden, the latter is not counted in the official score. Most of the time, this omission does little harm – but it obviously should be considered in comparing a temporary tax cut to a permanent one.

Rules can also have weird effects on scores for minor policy proposals. Consider the curious case of a proposal to replace the dollar bill with a coin. The official score in this case took account of the fact that since coins last much longer than bills, it would cost much less to manufacture the needed amount of currency. But it did not take account of the reality that it takes more than one coin to replace each bill because, among other things, people tend to throw coins on dressers at night and leave them there.

Nor did it account for the potential savings in interest payments on the federal debt. If deficits can be financed in small part by the distribution of new coins (as opposed to selling government bonds) the interest saving ought to be a significant part of the scoring calculation. But CBO is not allowed to count interest savings in its scores.

In this case, CBO estimated the interest

savings for informational purposes only, much as the new House rule provides information on the growth effects of tax cuts. But that did not placate organizations of the blind, vending machine operators, and others who support the transition to coins. They were rightly annoyed that an important component of the savings from their proposal was omitted from the official score.

I do not wish to leave the impression that bizarre results are common. But odd examples do reveal that a lot about scoring is close



to arbitrary. It is something of a game – one that brings some discipline to the budget process when it is played well. In recent years, though, the game has not been played well.

As noted earlier, a number of swing-vote senators forced the Republican leadership to limit the tax cut to \$350 billion in the 2004 budget. But that target was attained only with an array of phase-outs, phase-ins and sunsets that made a mockery of the \$350 billion limit and undermined the tax cut's potential to stimulate work, investment and, ultimately, economic growth. The distortions introduced by such shenanigans make aberrations in scoring rules and the absence of dynamic scoring seem a trivial problem. ■

Can China

Fix its

Banking

System?

Reformers at the Brink

by Marsha Vande Berg

China's new rulers are circling around what could prove the toughest challenge of their reign – reform of the state-run banking system and the crony Socialist system for allocating capital. They are taking steps to repair the balance sheets of major banks, which are groaning under the weight of billions of dollars in dud loans. And they must manage the process in ways that sustain the credibility of the Communist Party.

It didn't help, of course, that Beijing

