Journal of Economic Perspectives—Volume 10, Number 1—Winter 1996—Pages 141–157

Dynamic Revenue Estimation

Alan J. Auerbach

On January 10, 1995, the newly constituted, Republican-led Budget Committees of the Senate and House of Representatives held an unusual joint hearing to evaluate the methods that government agencies use for estimating expected tax revenue. The heads of the Congressional Budget Office (CBO) and the Joint Committee on Taxation (JCT) appeared to describe their current procedures, and several prominent academic and nonacademic economists gave their views about the state of this process.

A central question motivating the hearing was whether these and other government agencies should adopt "dynamic" estimation methods. When forming estimates of the impact of a particular policy on revenues (or expenditures, for that matter), dynamic methods would incorporate the predicted macroeconomic effects of such policies. Motivating this question was the view that existing methods, which omit such macroeconomic "feedback effects," bias decisions against "supply-side" or "growth-oriented" policies by counting only the tax revenues lost from providing such incentives, not the salutary effect they would have on spurring economic growth and in this way increasing tax revenues. The hearing did not result in any immediate changes in practice. However, it signalled that legislators have become sharply aware that in these times of large budget deficits, the revenue estimation process can shape and sometimes constrain their policy decisions.

This paper will begin by describing the current revenue estimation process and indicating how these estimates are used—subject to a variety of budget rules—in the formation of policy. It will then evaluate this process as an exercise in

■ *Alan J. Auerbach is Robert D. Burch Professor of Economics and Law, University of California, Berkeley, California, and Research Associate, National Bureau of Economic Research, Cambridge, Massachusetts.*

forecasting, considering not only its methodology but also its performance. This evaluation will illustrate some of the difficulties that would accompany attempts to use dynamic methods. Finally, I will consider how the choice among different revenue estimation procedures relates to the political process and the use to which the revenue estimates are put. An approach that might be optimal in a narrow statistical sense may not be socially optimal for the revenue estimating process, given the incentives and information available to various agents.

My perspective is that of a researcher who has studied the behavioral effects of taxation and the impact of budget rules, but also of an academic who spent a brief period working for the Joint Committee on Taxation in the early '90s.

Revenue Estimation: Terminology and Description

Government agencies produce two types of revenue estimates: forecasts of the level of aggregate revenue and its components, and estimates of the impact of specific policies on revenues. The estimates for individual legislative changes are used in revising the estimates of overall revenue levels and in determining whether particular pieces of legislation are in accord with budget rules.

Table 1 illustrates the revenue forecasting process, using the period around the Omnibus Budget Reconciliation Act of 1990 (OBRA90) as an example. At least twice each year, the CBO and the Office of Management and Budget (OMB) independently publish estimates of aggregate federal revenues and their components, such as corporate income taxes, individual income taxes and excise taxes. These estimates cover not only the upcoming fiscal year, but also several future fiscal years. The estimates reflect policy assumptions and forecasts of macroeconomic conditions. The first column of Table 1 gives CBO's July 1990 forecast of total federal revenues for fiscal years 1991–95.

Each time that CBO and OMB update their projections, they attribute revisions to three exhaustive and mutually exclusive sources of error: policy changes, changes in macroeconomic forecasts (usually simply called economic changes) and changes conditional on policy and macroeconomic factors (commonly called technical changes). Estimated effects of policy changes are provided by the JCT and the Treasury's Office of Tax Analysis (OTA), respectively. (A similar forecasting process is carried out for the expenditure side of the federal budget by CBO and OMB themselves.) The second column of Table 1 shows CBO's estimates of the revenue effects of policies adopted between July 1990 and January 1991, mostly attributable to OBRA90. The third and fourth columns of the table give CBO's economic and technical revisions over the same period, and the final column gives CBO's January 1991 forecasts for fiscal years 1991–95, equal to the sum of the first four columns. The last column also provides the initial forecast for fiscal year 1996, so one can imagine this process of revisions and estimates rolling forward over time. The large negative economic revisions are due to the recognition of the recession of

Table 1
Revenue Forecasts and OBRA90
(billions of dollars)

		Revisions			
Fiscal Year	*CBO, 7/90*	*Policy (from JCT)*	*Economic*	*Technical*	*CBO, 1/91*
1991	1,123	18	−40	−7	1,094
1992	1,188	33	−46	−5	1,170
1993	1,260	32	−38	−3	1,251
1994	1,337	37	−39	−2	1,332
1995	1,417	39	−43	3	1,416
1996	—	—	—	—	1,496

Source: CBO (1991).

1990–91, which began in July 1990: a downward revision in projected levels of GDP translated into a drop in expected tax collections.

All revenue estimates take nominal GDP and other aggregates as fixed; to put it another way, the estimated effects of a given policy on tax revenue exclude any macroeconomic feedback effects. However, this does not mean that the macroeconomic effects of all policies are assumed to be zero, only that any such effects are not attributed to particular policies. Forecasting agencies will sometimes try to incorporate feedback effects in their contemporaneous macroeconomic forecasts or in subsequent forecast revisions. The policy effects given in Table 1 come from JCT estimates that take as given the macroeconomic aggregates underlying the July 1990 CBO macroeconomic projections. In turn, the January 1991 CBO budget forecasts and revised macroeconomic projections take into account the policy changes that occurred after July 1990; however, the revenue effects of the macroeconomic forecast revision are reported separately, with none traced back and attributed to policy changes.

Thus, the macroeconomic effects of policy changes ultimately must be accounted for—either in revised macroeconomic forecasts or observed aggregates. However, individual policy changes are treated as having no indirect revenue effects through changes in GDP, interest rates, inflation and so on. It is this lack of connection between individual provisions and their macroeconomic effects, rather than the macroeconomic forecasts themselves, that has been highlighted in the recent debate.

From a practical point of view, it would be difficult for current estimation procedures to draw a connection between most legislative changes and macroeconomic outcomes, because the macroeconomic forecasts are not done at the level of detail that would be needed to trace changes in macroeconomic variables to specific legislative changes. Incorporating macroeconomic effects on a provision-by-provision basis would require an integration of the separate forecasting

functions, along with far more detail in the macroeconomic forecasting model. Whether such a change is advisable, given the resources required and the built-in limits on forecasting abilities, is a central question.

The failure to incorporate macroeconomic feedback effects in an explicit manner has led many critics to refer to existing revenue estimates as "static," mistakenly implying that the estimates ignore all behavioral responses. The accusation is often simply wrong. It is common for revenue estimates to incorporate projections of behavioral responses in particular affected markets or transactions. For example, the revenue estimate of a change in the capital gains tax rate does incorporate predicted changes in capital gains realizations and shifts in the composition of income, but does not explicitly include changes in aggregate output or investment that might also result from the tax change. In other words, the existing methodology directly accounts for some "micro" behavioral responses to taxation, but no "macro" responses.[1]

Of course, there is no assurance that the "micro" behavioral responses that current procedures permit will be predicted without error; any prediction errors will show up eventually as technical revisions, the revisions that the forecasting agency cannot directly attribute to new policy changes or changes in macroeconomic conditions. Thus, the actual impact of a policy on tax revenue in a particular fiscal year equals the impact initially estimated plus some unknown (positive or negative) portion of the changes in revenue that eventually are attributed to economic (macro) and technical (micro) factors as that year draws closer and its actual revenues are estimated more accurately and, ultimately, observed. The first of these excluded components is left out by design (the revenue estimates ignore macroeconomic feedback effects), and the second arises as a residual of the forecasting process.

The recent debate over revenue estimation has highlighted only some of the potential macroeconomic responses to policy. Even though much of the criticism has been of the exclusion of one class of macroeconomic behavioral responses—supply-side effects of revenue changes—the possible macroeconomic effects of policies also include the demand-side effects associated with increased utilization of resources and the supply-side and demand-side effects of expenditure policies as well. For example, Keynesian economists would insist that when the economy is in a situation with underemployed resources, then expansionary fiscal policies will have an offsetting effect of increasing tax revenues. In their view, this expan-

[1] This distinction between macroeconomic and microeconomic behavioral responses is somewhat arbitrary, based on what it means to hold macroeconomic aggregates fixed. For example, suppose that a reduction in marginal income tax rates were expected to increase taxable income by reducing both the share of total compensation received as nontaxed fringe benefits and the share of labor devoted to household production. These both are shifts of "full" labor income away from nontaxed forms in response to a reduction in the tax on market income. But only the first of these responses could be accounted for under current practice, because total compensation, rather than "full" income, is the aggregate being held fixed.

sion might readily come from higher government spending, rather than from tax cuts.

The debate also has focused at times on microeconomic responses as well. One illustration of the distinction comes from the ongoing debate about the effects of capital gains taxes. While some of the debate has been on the macro side, considering whether changes in the capital gains tax rate have a significant effect on aggregate output, much of the controversy has been on the micro side, involving whether forecasters adequately incorporate the responsiveness of capital gains realizations to changes in the capital gains tax rate.[2] Another recent example involves the responsiveness of taxable income to changes in marginal tax rates, with much of the emphasis now being placed not on the macroeconomic labor supply effects, but on the shifting of income and changes in deductions (Feldstein, 1995). In this example, existing forecasts have been criticized not simply for omitting macroeconomic behavioral responses, but also for understating microeconomic behavioral responses.

Though revenue (and expenditure) forecasts are continually being updated, often substantially, they are never reported with standard errors or confidence intervals. There are several possible explanations for this omission. A simple reason is that confidence bounds are very difficult to compute for such estimates. Forecasters rarely have the time to construct estimates "from scratch," collecting data and running regressions. Rather, they must cobble together predictions using existing evidence from different sources.

Consider, for example, an estimate of the revenue effects of a change in certain excise taxes, such as a tax on alcoholic beverages. An economist producing the revenue estimate probably would begin by distilling the empirical literature. At best, the literature might report elasticities of demand and supply for the relevant commodities that could be incorporated directly by the revenue estimator. The need for educated guesses frequently arises when the literature is incomplete. Often, the tax changes involve the introduction of new tax provisions, for which there may be no direct evidence at all.[3] All too often, during legislative markup sessions, conference committee meetings, or floor debate, revenue estimates required for official use must be produced or updated in a matter of minutes. For estimates formed under constraints of time pressure and incomplete data, how is one to calculate a confidence interval? These difficulties also help to explain why estimators rarely offer a detailed description of the methodology used to construct estimates. Such explanations are typically reserved for cases of ongoing interest in which formal statistical evidence is available, as in the JCT (1990a) presentation of its methodology for estimating the revenue consequences of capital gains tax changes.

[2] For a discussion in this journal of how a cut in capital gains taxes would affect the rate of realizations, see Auten and Cordes (1991).

[3] A recent example is the "Medical Savings Account," which would extend the rules governing Individual Retirement Accounts to savings earmarked for future medical expenses. The controversy surrounding the JCT estimate of this provision was sufficient to merit a *Wall Street Journal* editorial (May 15, 1995) criticizing, by name, the JCT economist responsible for the estimate.

Presumably, forecasters could offer their own subjective confidence intervals for the estimates they produce, and this extra information ought to be helpful for policymakers. However, there is also the question of how well legislators without formal statistical training would grasp the notion of a confidence interval. As is discussed below, the budget rules that Congress has adopted are not well designed to accommodate forecast uncertainty.

Budget Rules and "Scoring"

Beginning with the Gramm-Rudman-Hollings Act of 1985, Congress has labored under a succession of budget-constraining rules of its own design. These budget rules have placed revenue and outlay estimates at the center of the political process. Rather than influencing legislation indirectly through the information they provide, the estimates now determine directly whether individual pieces of legislation can be enacted.

The current rules, introduced by the Budget Enforcement Act of 1990, differ from Gramm-Rudman-Hollings in placing limits on policy-induced *changes* in the deficit, rather than on the *level* of the deficit itself. These rules prohibit Congress (except in cases of national emergency) from adopting tax and mandatory spending policies that, taken together, would increase CBO-estimated deficits for the current or next five fiscal years, a period called the "budget window." However, there is no restriction at all concerning what actually happens to the deficit. This can be a problem if projections turn out to have been overly optimistic, as they have during the period. If, as time goes by, the deficit for a particular fiscal year seems likely to be well above what was estimated initially (that is, estimated five years earlier when that fiscal year entered the budget window), the rules simply require that, by estimate, subsequent actions not make that year's deficit even higher. The budget rule has no error-correction mechanism.

These forecasting errors have been large in recent years. For example, for fiscal years 1990–93, the cumulative technical revisions in the deficit forecast by CBO—those not attributed to economic or policy changes—averaged $132 billion per year, which is just over half the average deficit during these years (Auerbach, 1994). One might sensibly decide that there is no reason to offset all unexpected deficit increases, such as those resulting from a slowdown in economic activity. However, there does not seem any justification for simply ignoring all errors, even those unrelated to the overall state of the economy. This lack of any retrospective correction is quite relevant in considering whether to incorporate additional elements potentially subject to large forecast errors, such a macroeconomic feedback effects, in a forecast.

To implement the current budget rules, the House and Senate have formulated their own operating procedures with respect to individual pieces of legislation, generally not allowing tax legislation that, according to JCT estimates of the particular provisions, would increase estimated deficits within the budget window. At-

tempting to control deficit changes requires a meaningful measure of what policy would be in the absence of such changes, referred to as a *baseline.* Determining the appropriate baseline, and using revenue estimates to assess changes in policy, is sometimes referred to as *scoring.*

A baseline revenue forecast generally incorporates current tax rules, including any scheduled changes in these rules that have been put in place by previous legislation. However, there are many anomalies and inconsistencies inherent in this baseline concept. If existing law calls for a particular tax to be eliminated, then the elimination of the tax becomes part of the baseline, and simply preserving that tax is scored as a revenue increase. For example, the Omnibus Budget Reconciliation Act of 1993 (OBRA93) pushed through Congress by President Clinton is well known to have included an increase in marginal tax rates on high-income people. In addition to these actual increases in marginal tax rates, the act also extended two existing marginal tax rate surcharges on high-income individuals that were scheduled to disappear in 1996 and 1997, respectively. By extending these surcharges, the act scored additional revenue during the latter years of the budget window.[4] It is only one small step further to structure provisions actually to *lower* revenues in years beyond the current budget window. Then, simply failing to sustain such a tax cut would eventually count as a tax "increase." On the other hand, extensions of certain taxes set to expire, such as gasoline taxes earmarked for the Highway Trust Fund, are specifically excluded from being scored as additional revenue.

Another difficulty arises when it is unclear what current law implies, as in cases where revenues are influenced by administrative actions or court decisions. For example, when a stimulus-minded President Bush reduced the default levels of federal income tax withholding in 1992 to delay some of that year's tax collections into 1993, Congress could have offset the change by legislation (in this case, by increasing withholding back to the original level) and scored the resulting increase in 1992 revenues. Similarly, Congress can score revenue by offsetting a new Supreme Court interpretation of existing law that is favorable to taxpayers.

Defining the baseline represents one of the potential problems with dynamic revenue estimation and scoring. To know the macroeconomic impact of a particular policy, and hence the change in revenue resulting from that macroeconomic effect, we must have a macroeconomic baseline to compare macroeconomic conditions with and without the tax policy. But it is difficult to isolate the macroeconomic effects of a single tax change from those of other changes in fiscal and monetary policy.

[4] These provisions—structured as phase-outs of itemized deductions and personal exemptions for high-income individuals—originally were introduced as temporary measures by OBRA90. Their being temporary had no impact on OBRA90's initial scoring recorded in Table 1, because the budget window then ended in 1995. Indeed, the reason why one of the extended provisions was scheduled to disappear in 1997, rather than 1996, is that it already had been extended by a year by the Unemployment Compensation Amendments of 1992, in order to help "pay for" the extension of unemployment benefits.

For example, what behavior should such a baseline assume on the part of the Federal Reserve? As a tax cut is introduced, will the Fed seek to hold constant the money supply, the federal funds rate, or perhaps nominal GDP? What additional fiscal policy changes will follow a reduction in tax revenues that Congress enacts today? A tax cut, by itself, upsets the government's long-run fiscal balance and requires offsetting policy changes, but how agents expect this balance to be restored can influence their reaction to the initial tax cut (Auerbach and Kotlikoff, 1987). As these questions suggest, holding other fiscal and monetary policies "fixed" is not a meaningful option for revenue estimation. The option currently taken—to assume that GDP remains fixed—is at least well defined, if arbitrary and problematic in other respects.

Dynamic Scoring in Theory and Practice: Two Case Studies

Proponents of dynamic scoring would require official revenue estimates used for scoring to include the impact of proposed legislation on macroeconomic aggregates, and thus on revenues. Two recent examples illustrate the potential impact and the difficulties of including macroeconomic feedback effects in official revenue estimates.

The Tale of the Luxury Taxes

During the fall 1990 budget summit that led to the enactment of OBRA90, Congress and President Bush sought methods of raising the tax burden on high-income individuals that met the president's wish to limit explicit increases in marginal income tax rates. One measure adopted was an excise tax on luxury goods: more specifically, a 10 percent excise tax on individual purchases of personal-use aircraft in excess of $250,000, boats in excess of $100,000, automobiles in excess of $30,000, and furs and jewelry in excess of $10,000. According to estimates published by the JCT (1990b) at the time, these luxury excise taxes would raise net revenues by a total of $25 million in fiscal year 1991 and $65 million in fiscal year 1992—tiny sums compared to the budget agreement's total net revenue increases in those two years, estimated by the JCT to be $17.6 billion and $30.2 billion, respectively.

The problems facing revenue estimators in this case were fairly typical. When the option of introducing luxury taxes arose during the budget summit, estimates were needed promptly. According to the JCT (1995, p. 17), this set of estimates had to be produced although "little information was available from academic literature or from the affected industries on the elasticity of demand for cars, boats, jewelry and furs, and personal-use aircraft with values in excess of the proposed excise tax thresholds. Based on the available information, the Joint Committee assumed a significant change in consumption patterns stemming from the implementation of the tax, i.e., it assumed a significant decline in purchases of taxed items." Along with estimates for the rest of the legislation, these estimates were included as policy revisions in the next round of CBO budget forecasts, reported above in Table 1. As already noted, the tax increase was more than offset by the contemporaneous downward economic and technical revisions.

Table 2
Luxury Tax Receipts, Fiscal Years 1991–92
(millions of dollars)

Tax	*1991*	*1992*
Airplanes		
estimate	1.0	4.0
actual	0.1	0.4
Boats		
estimate	4.0	9.0
actual	7.3	12.4
Autos		
estimate	27.0	69.0
actual	151.5	296.5
Furs		
estimate	<1.0	1.0
actual	0.3	0.7
Jewelry		
estimate	1.0	3.0
actual	9.2	15.8
Total		
estimate	33.0	87.0
actual	168.4	325.8

Source: JCT (1995).

Proponents of a revised revenue estimation and scoring process argue that, to whatever extent these revisions in overall revenues were attributable to the luxury taxes, they should have been included in the initial policy estimates provided by the JCT, in the same way that the JCT attempted to account for the effects of reduced purchases of the taxed commodities, holding GDP constant. In this instance, those arguing for a change in methodology also believed that this correction would be substantial. A paper issued by Republican members of the Joint Economic Committee (1991) soon after the taxes took effect indicated that for the tax on boats, which the JCT had estimated would raise $4 million in fiscal year 1991, the "total cost to the Treasury of the tax-related job loss is estimated at $18.2 million in 1991." It suggested that the tax on aircraft would cost the Treasury over $5 million, far outweighing the JCT-estimated revenue gain of $1 million. Dramatic press accounts of the devastating impact of the taxes on affected industries reinforced such conclusions. With bipartisan support for a change, OBRA93 (for which no Republicans voted) indexed the tax threshold for autos and repealed the other luxury taxes.

Now that fiscal years 1991 and 1992 have come and gone, we can look back and see how much revenue the luxury taxes actually raised. The initial JCT estimates, along with actual IRS receipts, are reported in Table 2. These estimates are for changes in excise tax revenues. Notice that the figures reported at the bottom of Table 2 for growth in excise tax revenues are higher than the estimates given

earlier for net growth in revenue ($33 million and $87 million cited in the table, versus $25 million and $65 million referred to in the text earlier for 1991 and 1992, respectively). The net figures account for a predicted loss in income taxes equal to 25 percent of the predicted increase in excise tax revenues. This estimated reduction in income taxes is known as an *offset.* Offsets are a logical and necessary consequence of the assumption that nominal output doesn't change. If indirect business taxes rise and nominal GDP remains constant, then, by the national income identity, factor incomes must fall by the same amount. Thus, so should corporate and personal income taxes on these factor incomes.

Comparing the actual and estimated luxury tax collections in Table 2, one is struck immediately by the extent to which luxury tax collections *exceeded* predictions. Most of this excess is accounted for by automobiles, which in turn accounted for most of the luxury tax revenues. However, taken together, the taxes on airplanes, boats, furs and jewelry also raised more than was predicted.

The enormous difference between the actual and estimated tax collections raises the general issue of forecast accuracy (although, in this instance, not one of overoptimism). For example, how could the fiscal year 1991 proceeds from the tax on luxury autos have been underestimated by 82 percent *at the beginning of that same fiscal year?* One can identify three potential causes for these forecast inaccuracies, corresponding to the categories used in CBO forecast revisions. First, the behavioral elasticities may have been overestimated, leading to greater predicted declines in demand than actually occurred. As the luxury taxes applied only over high thresholds, relatively small percentage errors in predicted expenditures per item would have translated into large percentage errors in predicted revenues. Second, there may have been unexplained shifts in demand for these luxury items that defy simple explanation,[5] or simply a higher level of compliance than had been expected. Finally, the estimates were conditioned on the CBO macroeconomic forecast of summer 1990. Unexpected changes in aggregate demand would have brought changes in the demand for luxury goods and affected excise tax collections. We cannot separate this component precisely, because the CBO does not allocate the economic revision in its revenue forecast to specific taxes at this level of disaggregation. However, given that CBO's economic revision during this period was sharply *downward,* it is difficult to imagine how macroeconomic factors could have produced unexpectedly high demand in these income-sensitive industries.

This episode helps illustrate the difficulty of using empirical observation to form predictions of the macroeconomic effects of small policies. In this case, a small luxury tax is associated in time with a substantial macroeconomic downturn—but the luxury tax itself is surely far too small to have precipitated an economy-wide recession. With all the other changes in policy and economic conditions that oc-

[5] For example, Zimmerman (1992) reports that sales of luxury boats declined by 36 percent from 1989 to 1990, just *before* the imposition of the tax, dwarfing "any absolute or percentage change in unit sales since 1970." By comparison, Zimmerman estimates that the price elasticity of demand should have reduced unit sales in 1991 by 4.2 percent.

curred simultaneously, there is little hope of identifying empirically the marginal effects on income tax collections of such small provisions.

With no hope of adducing empirical evidence regarding the magnitude of feedback effects, researchers would be forced to rely heavily on the predictions of economic theory. However, this reliance would not be rewarded with a clear picture of macroeconomic effects. Textbook Keynesian analysis supports the argument that increased consumption taxes will increase unemployment, but also that such reductions in aggregate demand can be largely offset by accommodating monetary policy. More recent "new classical" theories of employment fluctuations might dismiss the importance of aggregate demand effects entirely, ruling out the significance of employment effects regardless of monetary policy. Real business cycle models might emphasize the importance of whether the consumption taxes are permanent or temporary, with larger current output effects occurring through intertemporal substitution in the latter case. Which theoretical modelling assumptions are appropriate for the point estimate forecast to be reported?

Some of the problems of producing and evaluating revenue estimates should now be clear. In producing them, economists must use judgment in place of hard statistical evidence, because the provisions are too specific to have been anticipated in previous empirical work. In evaluating our predictions, we have little hope of explaining why our forecasts of direct revenue effects erred, even if these errors were enormous. Were we attempting to incorporate macroeconomic feedback effects in our estimates, there would be no way to estimate these effects empirically or to measure our accuracy in doing so. If the typical revenue effect resulting from induced macroeconomic changes is small relative to those effects for which we already account, then the large errors likely to occur in attempting to incorporate such macroeconomic "feedback" effects should give us pause.

Dynamic Scoring Circa 1995

Literally thousands of individual revenue estimates are produced each year, and most are done at a level of detail far greater than exists in macroeconomic models. Whatever the potential benefits of incorporating macroeconomic feedback effects in revenue estimates, doing so for policies as minor as the 1990 luxury taxes seems impractical. But this need not preclude the incorporation of macroeconomic effects when large legislative changes are proposed, as when significant overall changes in the level and composition of taxes and expenditures are considered simultaneously. Such a change was proposed during the budget debates of the House and Senate during the spring of 1995, when both chambers passed budget resolutions aimed at reducing the federal budget deficit to zero by fiscal year 2002.

The House and Senate budget plans offered different policies to achieve a zero deficit by fiscal 2002, based on the revenue estimates provided by JCT and the expenditure forecasts of CBO, extended beyond the normal budget window for this purpose. They also used different scoring methods. While the Senate plan achieved budget balance using the standard approach, leaving out macroeconomic feedback effects, the House plan incorporated such feedback effects. Where did the House

Table 3
House Balanced Budget Plan for Fiscal Year 2002
(billions of dollars)

Budget Component	*CBO Baseline*	*House Plan Changes*
Discretionary Spending	636	−113
Medicare and Medicaid	523	−143
Other Mandatory Spending	720	−57
Interest	350	−59
feedback		−42
less		
Revenues	1880	−72
feedback		8
equals		
Deficit	349	−350

Source: CBO (1995) plus unpublished House Budget Committee documents.

estimates come from, given that neither CBO nor JCT incorporates estimated macroeconomic feedback effects in their own estimates? In a recent publication, CBO (1995) had provided, in an appendix, estimates of the macroeconomic effects of a generic policy of gradual budget balance by fiscal year 2002, including the implied feedback effects on revenues (through higher economic growth) and debt service expenditures (through lower interest rates). Although the CBO estimates were for a generic plan to balance the budget, rather than for the specifics of the House plan, the House chose to include these additional components in declaring that they had arrived at a balanced-budget plan.

Table 3 provides details of the scoring of the House budget plan, indicating changes relative to the CBO baseline for revenues, expenditures and the deficit in fiscal year 2002. In addition to its direct reductions in spending and revenues, the plan included an estimated reduction in debt service costs resulting from a smaller accumulated national debt. To get to the desired deficit reduction of at least $349 billion in 2002, the plan also included macroeconomic feedback effects on debt service (through lower interest rates) and revenues (through higher GDP growth) of $50 billion in that year, based on the CBO forecast of the macroeconomic effects of deficit reduction.

Table 4 presents key elements of this CBO macroeconomic forecast, indicating the initial baseline values and the effects of deficit reduction. The initial baseline shows the GDP growth rate converging to 2.3 percent per year, the unemployment rate rising to a natural rate of 6.0 percent, the inflation rate stable at 3.4 percent, and the rates on Treasury bills and notes falling to long-run values of 5.1 percent and 6.7 percent, respectively. Consistent with the feedback effects given in Table 3, the forecast impact of deficit reduction is a large drop in interest rates and an increase of 0.1 percentage points in the annual GDP growth rate. This higher growth rate is attributable to an increase in domestic investment due to reduced crowding out by government borrowing.

Table 4

The Macroeconomic Effects of Budget Balance in Fiscal Year 2002

(by calendar year, in percent)

	1996	*1997*	*1998*	*1999*	*2000*	*2001*	*2002*
Real GDP Growth Rate							
Baseline	1.8	2.4	2.3	2.3	2.3	2.3	2.3
Change	0.0	0.1	0.1	0.1	0.1	0.1	0.1
CPI Growth Rate							
Baseline	3.4	3.4	3.4	3.4	3.4	3.4	3.4
Change	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Unemployment Rate							
Baseline	5.7	5.8	5.9	6.0	6.0	6.0	6.0
Change	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3-Month Treasury Bill Rate							
Baseline	5.7	5.3	5.2	5.1	5.1	5.1	5.1
Change	−0.2	−0.4	−0.7	−0.9	−1.1	−1.1	−1.1
10-Year Treasury Note Rate							
Baseline	7.0	6.7	6.7	6.7	6.7	6.7	6.7
Change	−0.2	−0.5	−0.8	−1.1	−1.4	−1.6	−1.6

Source: CBO (1995)

Despite a reduction in the deficit of 3.5 percent of GDP and the fiscal contraction this implies, the policy is predicted neither to raise the unemployment rate nor to lower the inflation rate, even temporarily. The CBO (1995, p. 51) justification for this projection is straightforward: "[T]he fiscal restraint implied by the effort to balance the budget need not weaken the economy substantially as long as the Federal Reserve acts to offset that restraint." Were alternative assumptions made concerning monetary policy, the result might well be different. With little or no accommodation, most traditional models would predict a smaller decline in interest rates and slower GDP growth, making the feedback effects on the deficit smaller, or even negative, at least during the period considered.

The CBO report urges caution in the use of these projections, because of uncertainty about the path and efficacy of monetary policy. However, it offers no alternative estimates, nor does it consider other channels through which a significant fiscal contraction might increase unemployment, notably the large sectoral shifts implied by large reductions in discretionary spending and entitlements.[6]

While no macroeconomic projections are available to estimate the feedback effects of the luxury taxes considered in the previous example, it is worth considering what consistent application of these CBO assumptions might have implied in

[6] For example, based on cross-state empirical estimates, Hooker and Knetter (1994) suggest that the current reductions in defense spending may add as much as 0.2 to 0.4 percentage points to the unemployment rate, temporarily.

that instance: no decline in aggregate demand, thanks to Fed accommodation; lower interest rates and an increase in private investment, due to reduced crowding out by government borrowing; and no increase in unemployment resulting from sectoral demand shifts. Thus, through the same channels of higher growth and lower interest rates, the macroeconomic feedback effects of the luxury taxes implied by current CBO assumptions would have *reinforced* the estimated deficit reduction. Whether these assumptions would have been appropriate in 1990 or are appropriate now is hard to know.

This recent use of dynamic scoring by Congress has a certain irony. Although the Contract with America included in the House budget had several provisions promoted for their incentive effects, none of these effects play an apparent role in the dynamic scoring accorded the budget package. Instead, the macroeconomic feedback effects come through the lower interest rates and reduced crowding out resulting from the standard demand-side analysis of the combined impact of tighter fiscal policy and easier monetary policy. The additional output attributable to added investment provides just $8 billion in revenue feedback in fiscal year 2002—hardly enough to justify a prolonged and heated debate about scoring methods. The CBO forecast does not encompass supply-side changes in the incentive to save or work or the effects on these changes. Indeed, the CBO forecast does not even flesh out the policy changes that lead to reduced deficits, but simply assumes sufficient reductions in discretionary and mandatory spending. Thus, this exercise provides little evidence of the impact of including supply-side considerations in revenue forecasts.

The CBO macroeconomic forecast might have been somewhat different had it been based on the specific House budget package, rather than being performed earlier as a generic deficit reduction exercise. During the recent debate, various private forecasters did attempt to extend their models to permit an analysis of the effects of the Contract with America on labor supply and savings. However, these attempts at identifying the macroeconomic effects of more specific provisions also illustrate the difficulty of doing so.

For example, Brinner, Wyss and Latta (1995) provide simulations of the DRI model, one of the models used by CBO in forming its own projections. These simulations, like those of CBO, generate considerable feedback effects from predicted interest rate reductions. But they go further to incorporate projections of the impact of welfare reform and marginal tax rate reductions on labor supply and the effect of a capital gains tax reduction on share prices and GDP growth. The labor supply effects are motivated by results in the empirical literature, but they are not based on statistical estimates by the modelers themselves and represent but one choice from a considerable range of plausible scenarios that research to date have provided.[7] The estimates of the effects of capital gains tax reductions on share prices

[7] See Gravelle (1995) for further discussion of the issue of incorporating labor supply and savings effects in revenue forecasts.

and GDP growth are based on simple theoretical calculations of the impact of a lower capital gains tax on the cost of capital. While these calculations may be defensible, they are not based on empirical evidence; nor do they appear to account for the fact that these large wealth effects could well *reduce* capital formation and GDP growth through increased consumption and crowding out (Auerbach, 1992). Such simulations offer little help in predicting how important a contribution supply-side factors are likely to make in reducing the deficit, because we have little information about the confidence with which these particular assumptions can be made.

At least for major policy changes, CBO and other government forecasters could produce an estimate of the effects of revenue and expenditure changes, taking account of both demand-side and supply-side macroeconomic effects of policy. But the magnitude of many of these effects would represent little more than guesses. Current methods for estimating the effects of policies are already hampered by the problem of inadequate information. The addition of major new elements of uncertainty about the macroeconomic effects of particular policies would only compound this uncertainty.

Should Dynamic Scoring Be Used?

In principle, the concept of dynamic revenue estimation and scoring is attractive: why not account for all revenue effects rather than just some? The same argument would apply, of course, to the expenditure side of the budget as well, and to all macroeconomic effects, not just supply-side effects.

As a pure problem in forecasting, it makes sense to use all available information, at least to the extent that the benefits of improved forecasting outweigh the costs of doing so. This condition, in itself, poses an obstacle to the general use of dynamic scoring. The costs of expanded forecasts, in time and money, would not be insubstantial. It would be necessary to integrate the separate activities of macroeconomic forecasting and tax revenue forecasting, and to map out the detailed mechanisms by which revenue changes influence macroeconomic behavior. For budgetary reasons, it would probably be necessary to reserve such complete analysis for major proposals and continue to analyze many smaller proposals in the same way they are being done now.

But both evidence and logic suggest that revenue estimation is not a pure exercise in forecasting. Estimators are under considerable pressure to provide favorable estimates—that is, estimates that understate program costs and overstate revenues. This pressure isn't always easy to see (although the anecdote above in footnote 3 offers an example). But it is visible in the fact that recent CBO and OMB errors in revenue forecasting have not been unbiased and serially uncorrelated, as one would expect from the errors from optimal forecasts. Rather, at least since the mid-1980s, revenues have been consistently overestimated, and expenditures underestimated, by very large amounts (Auerbach, 1994, 1995). Of course, part of the reason may be that tax increases were predicted to raise too much money, because

the depressing behavioral consequences on economic behavior were not taken into account. But whatever the technical framework, the pressure under the present budget rules imparts a bias toward predicting a lower budget deficit.

This pressure works in subtle ways. No government revenue estimator in my acquaintance would consciously provide a biased estimate. But in many instances, the uncertainty is so great that one honestly could report a number either twice or half the size of the estimate actually reported. Facing the threat of job loss and public criticism by members of Congress and editorial writers each time an unfavorable estimate is reported, do we really expect estimators to flip a coin when they're unsure which number is more accurate? What government revenue estimator will attempt to hold back the tide of legislative zeal (or editorial opprobrium) by insisting on a pessimistic view of the likely outcome?

One could construct a model of the incentives for such a forecaster, incorporating the conflicting goals of accuracy and accommodation, to show that the poorer the information an estimator has, the more he or she is likely to report a biased policy estimate. In such a situation, requiring estimates in cases where information is scant may actually make the resulting estimates worse. It may be that, even if we have *some* information about the macroeconomic effects of policy proposals, reported estimates will actually be poorer if we insist that this information be incorporated (without standard errors) in estimates used for budget scoring.

There are intermediate steps that could be taken. One possibility is to encourage agencies to expand their practice of providing "unofficial" analyses of the macroeconomic effects of particular policies, to help guide the legislative process without providing numbers for scoring purposes.[8] However, the use by Congress in spring 1995 of the unofficial CBO estimate of the feedback effects of deficit reduction illustrates the difficulty of maintaining this distinction. A second step might be to review the present specification of "micro" behavioral effects. While current practice incorporates such effects in many instances, there may be compelling cases in which these effects are inaccurate or inadequate.

Conclusions

Revenue (and expenditure) estimates already are dynamic in taking account of behavioral effects of some policies, but they do hold macroeconomic aggregates fixed. Adding macroeconomic feedback effects is technically feasible, at least for major policy changes. Given the assumptions of the CBO macroeconomic forecasting model, the major feedback effects of deficit-reducing policies come through a reduction in interest rates. But one could also incorporate estimates based on supply-side or incentive effects envisioned by many champions of dynamic scoring,

[8] For example, CBO (1990) considers the potential effects of capital gains tax reductions on economic growth. The study illustrates the range of estimates available in the literature.

demand-side effects of moving the economy toward full employment as envisioned by many Keynesian economists, intertemporal effects of the sort envisaged by real business cycle economists, and more. Any such estimates will rely on assumptions about offsetting monetary, fiscal, behavioral and even legal and social effects.

In a technical sense, the primary argument against incorporating feedback effects in estimation and scoring is the cost of doing it, in relation to the limited information it is likely to convey. But in the intense political environment in which estimates are produced and used, and given the unavoidable uncertainties of compiling such forecasts, it also seems likely that fully dynamic forecasts may end up being even more biased and inaccurate than the forecasts of the last few years.

■ *I am grateful to Henry Aaron, Tom Barthold, David Bradford, Jane Gravelle, Glenn Hubbard, Bob Kilpatrick, Joyce Manchester, John Quigley, Carl Shapiro, Timothy Taylor, Jerry Tempalski and Randy Weiss for comments on an earlier draft.*

References

Auerbach, Alan J., "On the Design and Reform of Capital Gains Taxation," *American Economic Review,* 1992, *82*:2, 263–67.

Auerbach, Alan J., "The U.S. Fiscal Problem: Where We Are, How We Got Here, and Where We're Going." In Fischer, S., and J. Rotemberg, eds., *NBER Macroeconomics Annual,* Cambridge, Mass.: MIT Press, 1994, pp. 141–75.

Auerbach, Alan J., "Tax Projections and the Budget: Lessons from the 1980s," *American Economic Review,* 1995, *85*:2, 165–69.

Auerbach, Alan J., and Laurence J. Kotlikoff, *Dynamic Fiscal Policy.* Cambridge: Cambridge University Press, 1987.

Auten, Gerald, and Joseph Cordes, "Policy Watch: Cutting Capital Gains Taxes," *Journal of Economic Perspectives,* 1991, *5*:1, 181–92.

Brinner, Roger E., David A. Wyss, and Cynthia M. Latta, "Growth and Budget Repercussions of the Republican Contract with America," *DRI/McGraw-Hill U.S. Review,* February 1995, 27–43.

Congressional Budget Office, "Effects of Lower Capital Gains Taxes on Economic Growth," August 1990.

Congressional Budget Office, "The Economic and Budget Outlook: Fiscal Years 1992–1996," January 1991.

Congressional Budget Office, "An Analysis of the President's Budgetary Proposals for Fiscal Year 1996," April 1995.

Feldstein, Martin, "The Effect of Marginal Tax Rates on Taxable Income: A Panel Study of the 1986 Tax Reform Act," *Journal of Political Economy,* 1995, *103*:3, 551–72.

Gravelle, Jane G., "Behavioral Feedback Effects and the Revenue Estimating Process," *National Tax Journal,* 1995, *48*: 3, 463–77.

Hooker, Mark, and Michael Knetter, "Unemployment Effects of Military Spending: Evidence from a Panel of States." NBER Working Paper No. 4889, October 1994.

Joint Committee on Taxation, "Explanation of Methodology Used to Estimate Proposals Affecting the Taxation of Income from Capital Gains (JCS-12-90)," March 27, 1990a.

Joint Committee on Taxation, "Budget Reconciliation (H.R. 5835)—Revenue Provisions (JCX-45-90)," October 26, 1990b.

Joint Committee on Taxation, "Discussion of Revenue Estimation Methodology and Process (JCS-14-92)," August 13, 1992.

Joint Committee on Taxation, "Methodology and Issues in the Revenue Estimating Process (JCX-2-95)," January 23, 1995.

Joint Economic Committee (Republican Staff), "The Cost of Tax-Related Job Loss versus Projected Revenue Gain from Luxury Taxes in Fiscal 1991," July 1991.

Zimmerman, Dennis, "The Effect of the Luxury Excise Tax on the Sale of Luxury Boats," Congressional Research Service Report, February 10, 1992.