

Statement of Leonard E. Burman¹

Senior Fellow, the Urban Institute
Director, the Tax Policy Center
www.taxpolicycenter.org

Before the House Committee on the Budget

Dynamic Analysis and Scoring

September 13, 2006

Chairman Nussle, Ranking Member Spratt, and members of the committee. Thank you for inviting me to present my views on dynamic analysis and scoring.

With three economists on this panel, I think it is safe to say that we would all like more attention paid to the economic effects of public policies. The big question is whether dynamic scoring or dynamic analysis is the best way to bring such analysis to bear on public programs. (To define terms, dynamic scoring involves adding a point estimate of macroeconomic feedback effects into official revenue estimates. Dynamic analysis is a supplemental analysis of plausible macroeconomic responses under a range of models and parameter assumptions.)

A related question, which I will also touch on, is how the policy-making process itself could be made more conducive to meaningful economic analysis.

In short, my conclusions are these:

- Many behavioral responses are already included in official revenue estimates of tax changes; that is, they are not static. The estimates could be improved in several ways. However, given the current state of economic knowledge, including macroeconomic feedback effects (dynamic scoring) is not one of them.
- There are three key problems in analyzing the effects of tax policy proposals: for deficit-financed tax and budget proposals, the long-term economic effects depend critically on how the deficit is financed (that is, on who ultimately pays for the tax cuts or new spending), and that is inherently unknowable by any estimator; there is tremendous uncertainty about key parameters that reflect how people make decisions about working and saving that can have large effects on estimates; and the limits of

¹ Views expressed are mine alone and do not necessarily reflect the views of any organization with which I am affiliated.

data, computing power, and economists' ingenuity mean that our models have little relationship to the way real people make real decisions.

- That said, most economists would agree that certain kinds of tax and spending policies are better for economic growth than others, so we could produce a rough ranking regardless of the financing mechanism or long-term economic model. For that reason, economic analysis of specific provisions as well as entire packages is useful. While dynamic analysis typically has been equated with macroeconomic modeling, for many specific provisions, an analysis of the microeconomic effects is all that is feasible at present. Such analysis would be a useful complement to policymaking.

Introduction

Federal tax and spending policies have an effect on the economy and citizens' well-being. Obviously, we should measure those effects as accurately as we can simply as a matter of responsible budgeting. What's more, the effects of policies on the economy clearly should be considered as a factor in assessing their desirability. All else equal, pro-growth policies are better, although there is often a trade-off between economic efficiency and other goals, such as fairness; growth is only one factor to consider.

Micro-Dynamic Revenue Estimates

By longstanding practice, official revenue estimates are dynamic in a microeconomic sense. They account for all the measurable behavioral responses that can be anticipated consistent with an assumption that macroeconomic aggregates—including labor supply, saving, and gross domestic product—are held constant. Thus, the official estimates of the income tax rate cuts enacted in 2001 assumed that at lower tax rates, taxpayers would report more taxable income because, for example, they would earn a smaller fraction of compensation in the form of untaxed fringe benefits and perhaps be less prone to cheating. However, the estimates accounted for neither a boost in hours worked or saving, which might have increased growth, nor a drop in investment or demand for homes and other consumer durables as a result of swelling public debt and higher interest rates, which would have retarded growth.

There are ways to improve revenue estimates, but dynamic scoring is not, at present, one of them. A problem that could be rectified is that estimators must provide a single point estimate that assumes that a host of unknown factors are known with certainty. This can cause the cost of particular types of tax proposals to be consistently underestimated. For example, in 2004, Congress effectively created a price support program for certain low-yielding oil wells. If prices fell below a certain trigger price, a tax credit would offset the difference between the actual price and the trigger price. Since the trigger was set below then-prevailing oil prices, the provision was scored as having no revenue effect, even though under some scenarios it could have been very costly to the Treasury. A better rule would be to estimate the *expected* revenue loss—that is, the average across all the

plausible price scenarios—to get an idea of what the price guarantee would cost the government (and be worth to recipients).²

More fundamentally, the legislative process itself may introduce biases into revenue estimates in a subtle way. The reason is that revenue estimates are subject to error. Under the best of circumstances, the errors will average out to zero. However, if overall budget targets are binding, then tax cuts and spending programs that appear to cost less will be favored over those that appear to cost more. (Indeed, the principal argument for dynamic scoring is that advocates believe that tax cuts would be more feasible if official estimators predicted that they would cost less in terms of lost revenues.) That means that tax cuts that are underestimated (and tax increases that are overestimated) will be more likely to be enacted than those that err in the opposite direction. As a result, despite the best efforts of estimators, the errors in policies that are actually adopted will tend to go in the same direction—they will not average out to zero. Revenue estimates will be consistently over-optimistic and deficits larger than predicted (or surpluses smaller). This might argue for a deliberate offsetting conservative bias in revenue estimating to make estimates more accurate on average.

Berkeley economist, Alan Auerbach looked at the accuracy of baseline receipts forecasts over many years and did not find evidence of consistent bias one way or the other.³ However, Auerbach found that receipts projections tend to be inefficient in the sense that aggregate errors tend to repeat from year to year. Building on Auerbach's work to adjust baseline outlay and receipts forecasts could make budgets more accurate although, as he notes, that is easier said than done.

Pitfalls of Incorporating Macroeconomic Effects in Estimates

What about macroeconomic effects? Most economists would agree that a major tax reform in which loopholes were eliminated and tax rates lowered, holding overall revenues constant, would increase economic growth, although there would be a wide range of estimates of how much.

Unfortunately, the vast majority of tax proposals considered by Congress would not fit in this no-brainer category of growth enhancers. While everyone likes lower tax rates, base broadening is a lot more popular with economists than it is with the people who pay higher taxes as a result. Tax cuts enacted since 2001, for example, have lowered marginal tax rates, but they also narrowed the tax base by creating a slew of new targeted tax breaks—including that one for unproductive oil wells I mentioned earlier—that are likely to hurt the economy rather than help it. This makes assessing the net effect problematic.

The biggest problem, though, is that recent tax bills have produced significant revenue losses with no indication of how those losses will be offset. Without knowing that, it is

² The CBO now applies “probabilistic scoring” to measure the cost of similar spending programs.

³ Alan J. Auerbach, “On the Performance and Use of Government Revenue Forecasts,” *National Tax Journal*. Vol. 52 no. 4 (December 1999) pp. 765-782

impossible to assess the economic effects, or even to measure whether the economy will be stronger or weaker in the long run.

Depending on how the deficits are closed, there could be dramatically different economic results. The best-case scenario for economic growth is for deficits to be financed by cuts in transfer programs or increases in lump-sum taxes (fixed per capita taxes not related to ability to pay). That deficits might force spending constraint appears to be the logic behind the “starve the beast” rationale for deficit-financed tax cuts, but there is no evidence that this tack actually works. It is not clear why spending cuts would be easier in the future than they are now. Will it be easier to cut Social Security and Medicare 20 years from now when all the baby boomers are retired (and AARP’s membership has exploded)?

The worst-case scenario for economic growth is this: years from now, our profligate budgetary policies lead to dramatically higher interest rates and a massive recession, if not a depression. Taxpayers blame this on the tax cuts for the rich and decide to deal with budget problems by raising tax rates on high-income folks. (And they leave in place all the middle-class tax cuts like the child credit, higher standard deduction, and 10-percent bracket.) I think it is safe to say that in JCT’s, CBO’s, and Treasury’s models, such a tax increase would prove most damaging to growth. The net effect would be a much smaller economy than would exist had the tax cuts not been enacted.

To be clear, this long-term risk also means that deficit-financed spending would also be more costly than would appear in either a balanced-budget scenario or one assuming less damaging deficit offsets in the future.

Since it is impossible for official scorers to predict how the deficits will be closed, you cannot expect them to produce a dynamic score for the long-term effect of deficit-financed tax cuts. For related reasons, it is a challenge to predict the short-term effects as well. In the standard Keynesian macroeconomic model, short-term fiscal stimulus (a spending increase or tax cut) boosts the economy during downturns by spurring households to spend and businesses to invest, creating more demand and thus more jobs. When the economy is at full employment, deficit-financed tax cuts can hurt by creating inflationary pressure. If the economy is running at capacity, companies will respond to higher demand by bidding up wages to try to keep or retain workers, which translates into higher product prices and inflation.

The wild card is the Federal Reserve, which tries to stimulate the economy when it is underperforming and slow it down when inflationary pressures arise. Fed policymakers are likely to respond to tax cuts by tightening up monetary policy to prevent inflation. Since monetary policy affects the economy more slowly than fiscal policy, short-term deficits that are larger than the Fed had expected can still have an immediate effect, but the effect beyond that is complicated by the Fed’s response. While this is probably more predictable than how future Congresses will deal with the national debt, it significantly complicates forecasting the effects of fiscal policy beyond a year or so.

Value and Limitations of Long-term Macroeconomic Models

A model of the economy can be a very useful tool in assessing tax reform options. A well-designed model incorporates individuals' and firms' decision processes and their interaction with each other and with government policy. Although the point predictions of such models are of questionable value, since they depend on parameters that are highly uncertain, the models do demonstrate the channels through which tax policy can affect the economy. They also allow for consistent comparisons of different policy options.

There are four basic kinds of models used for macroeconomic analysis of tax policy (with almost infinite variations): neoclassical growth models; disequilibrium (or Keynesian) models; infinite horizon models; and overlapping generations (OLG) models.⁴ The first two types of models represent a very stylized version of the economy. Individuals and firms do not make optimizing decisions. Instead, the results of those optimization decisions are reflected in numerical measures of the response of saving, labor supply, and factor substitution (firm's ability to substitute capital and labor for each other in the production process), which are known as elasticities. Taxes can affect the economy by altering the return to saving and working for individuals or the costs of labor and capital for firms. Such models are often enhanced by disaggregating different economic sectors (such as manufacturing, agriculture, services, etc.) on the assumption that they have different production technologies (i.e., use capital and labor differently) and often include different classes of workers who have different skill levels.

The neoclassical growth model assumes full employment: markets for all goods and services always clear instantaneously so unemployment, which is a disequilibrium between the supply and demand for labor, is not possible. Disequilibrium or Keynesian models assume that such disconnects are the norm, but that they can be affected by government policy and the business cycle. Thus, during a recession, tax cuts can reduce unemployment and increase GDP because lower taxes spur individuals to spend more or companies to invest more, which increases the demand for goods and services, translating into more jobs.

Modern disequilibrium models are typically combined with standard growth models. In such models, tax changes can affect labor supply and saving, which affects output and the demand for capital in the next period. The changes alter the return to capital and labor, which adjust again affecting output and the demand for productive inputs. The process continues until the demand and supply for capital return to equilibrium, from which point on the economy grows at a constant steady-state rate.

Such models have been used for decades and are well understood. They are attractive because their results are fairly easy to explain and intuitive, but they have some limitations. First, Keynesian models, being inherently short-term in focus, do not tell

⁴ For a good survey, see Jane G. Gravelle. 2003. "Issues in Dynamic Revenue Estimating," CRS Report for Congress RL31949, U.S. Congressional Research Service; or the Appendix of U.S. Congressional Budget Office. 2003. "How CBO Analyzed the Macroeconomic Effects of the President's Budget." *CBO Paper*. Available at: <http://www.cbo.gov/ftpdocs/44xx/doc4454/07-28-PresidentsBudget.pdf>.

policymakers about the long-term effects of tax policy, when, presumably, the level of equilibrium is of most interest. Indeed, they may not provide the answer tax cut advocates want in the short term. Such models typically predict that spending increases or cuts will have a larger effect than tax changes because government spending immediately generates additional demand for goods and services, whereas tax cuts affect demand only to the extent that the recipients choose to spend them rather than save. In such models, tax cuts are good, but spending is better and deficits are a plus in the short-term.

A more fundamental problem with both the disequilibrium and the growth models are that they are too aggregated. They assume, for example, that labor supply and saving decisions of individuals (or groups of individuals) depend only on the average tax rate on labor and capital income. Thus, replacing a progressive income tax with a flat rate tax that raises the same amount of tax revenue would be expected to have no effect on work or saving decisions since the average tax rate remains the same. But reducing high tax rates is likely to produce a larger positive effect than the negative effect of increasing tax rates at the bottom or broadening the tax base. The individual decisions do not average out to zero. Similarly, policies that affect individuals' and firms' expectations about the future can have big effects on their behavior now, but the neoclassical and disequilibrium models are not forward-looking.

Infinite horizon (or Ramsey) models and OLG models are more modern representations of the economy based on the decisions of individuals and firms. Individuals maximize utility subject to an intertemporal budget constraint (that is, typically, they cannot die in debt). Firms maximize profits. The government must balance its budget over the long term (although not necessarily over any finite interval). In some such models, people have perfect foresight: they can predict the future accurately. In other more realistic models, the future is uncertain so results depend on how people are assumed to form expectations. Rational expectations models assume that people have a very good macro model inside their heads so that their forecasts are correct on average.⁵ Other models assume myopia—people assume that the present will continue—or adaptive expectations.

In infinite horizon models, people (and firms) live forever. Obviously, this is an unrealistic assumption, but advocates of such models argue that they are a good and relatively simple representation of a world where people care about their children as much as they care about themselves. Therefore, the preferences of children enter their parents' utility functions and are represented in a motive to leave bequests. Since children will also care about their children, and so on, the very long time horizon may be warranted.

The OLG models represent the very long term by assuming that individuals live for a fixed number of years, but are replaced by children and grandchildren with similar preferences who are young when the parents age. By solving for the decision process of

⁵ Ironically, a consequence of this assumption is that it is impossible for economists to develop an accurate aggregate model of the economy because individuals' behavior changes the parameters of the economists' models, invalidating them.

each generation and connecting them (primarily through interest rates), there is, again, a very long horizon in such models.

An advantage of this class of models is that it is possible to build in great detail on the tax structures facing individuals and firms. A disadvantage is that the models depend on parameters about which little is known. In particular, the models depend on the parameters of individuals' utility functions: their trade-off between consumption and leisure (and thus labor) in the current period, and their willingness to trade future consumption and leisure for current consumption and leisure. For example, if people expect taxes to increase in 20 years, will they work harder and spend more now, and if so, by how much? These intra- and intertemporal elasticities (and, to a lesser extent, factor substitution elasticities of firms) are critical to the predictions of such models, but very little is known about the proper values. Critics have also pointed out that there is a considerable amount of evidence that individual decisions deviate in important ways from the predictions of the life cycle model, which underpins both of these frameworks.

These models are also very sensitive to their exact structure. The CBO found larger growth effects in the infinite horizon model than in the OLG model; the smallest long-term effects arose in the neoclassical growth model. Different forms of uncertainty and assumptions about individuals' attitudes towards uncertainty can also produce markedly different predictions about the effects of a given policy.

A key implication is that a single model will not be adequate for evaluating the long-term effects of public policies since the results may be very sensitive to the choice of model. When employing any model, parameter assumptions should be subjected to extensive sensitivity analysis. That is, different values for key parameters, such as labor supply and saving elasticities, in the case of the disequilibrium and growth models, and intra- and inter-temporal elasticities of substitution, in the case of the infinite horizon and OLG models, should be tried to see how sensitive the results are to the parameter assumptions.

Put differently, there is no basis for producing a single point estimate for the macroeconomic effect of tax or spending policies, even when they are not deficit-financed. While the kind of dynamic analysis that CBO and JCT have done using a range of models can be an informative input to public policy, it is of virtually no value in improving estimates of the short- or long-term effects on revenues. Indeed, since none of these models has been validated in actual practice, choosing a particular model, a set of parameters, and assumptions about the way deficits will be offset and the Federal Reserve will react would almost surely add new biases and significantly increase the variance of revenue estimates. This problem also raises the risk that policymakers will gravitate towards policies whose macroeconomic feedback effects are most overstated in a particular model chosen by estimators, even though these policies may not necessarily be the best ones for the economy in the long run. It certainly raises the risks that forecast accuracy would be significantly worse with this approach.

The Potential Usefulness of Dynamic Analysis

Despite its limitations, dynamic analysis is potentially a useful complement to policy making, although existing models are quite limited in what they can simulate. All else equal, it would be nice to discriminate in favor of pro-growth policies, especially if the growth benefits are widely shared rather than concentrated at the top.

Note, however, that this is not necessarily an argument in favor of tax cuts. First, as noted, when financing is considered, almost any tax cut could turn out to be counterproductive over the long run. Second, some tax cuts would tend to reduce economic growth no matter how they are financed and some spending increases would tend to enhance growth.

For example, a horse-and-buggy tax credit would certainly create jobs in the horse and buggy industry, potentially reversing a century-long downturn, but almost nobody would argue that this would be good for the economy overall. The resources that were diverted into horses and buggies could surely be better used in any of thousands of goods and services that consumers value more.

You might think that Congress would never enact such a thing, but the manufacturers' tax deduction and many of the other targeted tax breaks, enacted as part of the American Jobs Creation Act of 2004, are not much different. They distort market prices and interfere with the efficient allocation of scarce economic resources. While there may be a role for targeted taxes or subsidies in markets when they are not working—for example, when there is pollution—many if not most tax breaks cannot be justified on those grounds.

On the other side, some kinds of government spending may produce economic benefits over and above their direct value to beneficiaries. Some examples include investments in infrastructure, education, information, and research and experimentation. Not all such projects produce benefits in excess of their costs (as the recent debate about the bridge to nowhere in Alaska illustrates), but well chosen public investments can produce substantial payoffs.

Unfortunately, the kinds of models designed to do dynamic analysis are not well suited to discriminating among good and bad kinds of targeted tax incentives or spending programs. Typically, such models represent different sectors at a highly aggregated level and have only a rudimentary representation of the tax system and no detail at all about spending. Indeed, in some macroeconomic models, government spending is tantamount to throwing the money (and the real resources it represents) into the ocean.

However, government analysts can do a microeconomic analysis of the efficiency effects of different programs and, indeed, the CBO, GAO, and CRS do when Congress asks. . Often, the analysis can be informed by empirical estimates—as in the case of investments in infrastructure, education, and research—although the research findings can vary wildly.

At a minimum, as part of the summary analysis of each piece of proposed legislation, it would be nice to include a qualitative analysis of the likely efficiency effects of each provision.

Other Advantages of Specifying Financing for Tax Cuts

As noted, the way tax cuts (or spending increases) are financed can fundamentally alter the assessment of their long-term economic effects. It also affects how we assess the distribution of tax benefits. The standard distribution table ignores the question of how deficits will be financed. As a result, tax cuts can look like good news for almost everyone. Advocates can argue that everyone is a winner and may appear to be right. Anyone arguing against the tax cuts is just a selfish demagogue practicing the “politics of envy.”

If we were explicit about financing, however, the picture would change. A deficit-neutral tax change has to make some people worse off and those people often object. The people who will pay for the government’s current generosity appear to be our children, and maybe that works politically because children and those not yet born don’t vote. However, if one accepts the key assumption of the infinite horizon model that parents care as much about their children and grandchildren as they do about themselves, then being explicit about who will bear the burden of current tax cuts would create political fallout.

How the tax cuts are financed then becomes very important. Under the scenario where a future Congress decides to close the deficit by soaking the rich, the President’s tax cuts become a lot more progressive than they appeared before financing was considered. If, instead, we follow the growth-maximizing path and slash spending, then many, if not most, taxpayers will find that they lose much more in future benefits than they gain in short-term tax cuts.

Thus, the best case for economic growth (all households take an equal share of resulting future debt service) produces the bleakest case for progressivity. For example, under any of the major tax cuts passed since 2001, the vast majority of households would be worse off under this financing option unless the economic benefits turned out to be implausibly large.⁶ And the best case for progressivity (high-income taxpayers face income tax rate hikes to offset the debt) is the worst case for long-term economic growth.

If Congress has in mind financing options that involve less draconian trade-offs, it should be explicit about them.

Conclusion

⁶ Jason Furman, “A Short Guide to Dynamic Scoring,” Center on Budget and Policy Priorities, August 24, 2006. Available at <http://www.cbpp.org/7-12-06bud2.pdf>.

Dynamic scoring is not feasible because of lack of knowledge about how deficits will be offset, uncertainty about key parameters in economic models, and inherent limitations in those models themselves. Dynamic analysis, however, is useful, but it should be applied to spending as well as taxes. What's more, the economic analysis of tax and spending provisions should be done on a provision-by-provision basis, not just overall packages.

Finally, both dynamic analysis and the assessment of the distribution of winners and losers from tax changes could be made much more accurate if Congress specified a financing mechanism for each major piece of tax legislation.