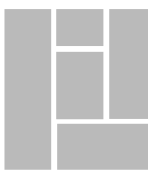
Errors IN Budget Forecasting

RUDOLPH G. PENNER

The Urban Institute
Washington, D.C.

Errors IN Budget Forecasting

RUDOLPH G. PENNER



The Urban Institute

The Urban Institute is a nonprofit, nonpartisan policy research and educational organization that examines the social, economic, and governance problems facing the nation. The views expressed are those of the authors and should not be attributed to the Urban Institute, its trustees, or its funders.

Contents

Errors IN Budget Forecasting

RUDOLPH G. PENNER

Since budgeting is about the future, budget decisions regarding the allocation of resources must be based on forecasts. But budget forecasts are always wrong, and often they are wrong by a lot (Aaron 2000). Indeed, changes in budget projections caused by alterations in economic forecasts and technical assumptions generally far exceed adjustments caused by changes in fiscal policies. A tax cut of $50 billion planned in the summer for the next fiscal year would represent a politically important change in policy, but an identical change in the surplus projected for the next fiscal year would be relatively minor compared to a typical forecasting error. Between April and July of 2000, for example, the Congressional Budget Office (CBO) adjusted the surplus estimated for 2001 upward by $95 billion for economic and technical reasons. The degree to which politicians are annoyed by such major adjustments is only dampened slightly by the fact that we are now enjoying the best budget outcomes of the last 70 years.

Forecasts are unlikely to become significantly more accurate in the future. They are compiled by skilled technicians who do their work as conscientiously as possible. The problem is partly due to the unscientific state of economic science. But economic and budget forecasting is also affected by forecasting in other fields. Forecasters are notoriously unreliable at predicting things like the next wave of technological change, the weather, and earthquakes, all of which can have significant economic and budget implications. Given that we cannot expect major improvements in the accuracy of forecasts, the question becomes how policymakers should use these highly inaccurate forecasts as effectively as possible and how journalists should report on them. But before tackling this difficult question, it is useful to examine the historical record.

Sins of the Past

Surplus Projections

The budget surplus is of great interest to policymakers and the press. It provides an immediate, if crude, indicator of how the nation is doing fiscally. Errors in projecting the surplus seem inordinately large because budget forecasters do not project the surpluses directly. They project two much larger numbers, receipts and outlays, and take the difference. As a result, relatively small percentage errors in projecting receipts or outlays translate into very large percentage errors in projecting the surplus. For example, receipts in fiscal 2000 were $2,025 billion and outlays were $1,788 billion, so the surplus was $237 billion. Between 1980 and 2000, the average absolute error in projections of receipts and outlays for the next year was a little more than 3 percent.[1] Such an error in the forecast of receipts for 2000 would have implied a change of 26 percent in the surplus forecast.

Chart 1 examines the economic and technical errors made in forecasting the budget balance between the promulgation of CBO's baseline budget, which is usually prepared in March or April, and the final number for the subsequent fiscal year.[2] For example, the baseline estimate for fiscal 1998 (prepared in March 1997) is compared with the actual outcome in 1998 after the effects of legislative action between the two dates are removed.

CHART 1 Economic and Technical Errors for One-Year Forecasts, as a Percentage of GDP, 1980–2000

Although CBO's first budget projection, generally produced in January, gets more publicity than the baseline forecast produced in the spring, it is the latter that serves as a foundation for congressional budget deliberations, so this analysis starts with the spring projection. Typically, the baseline differs only slightly from the January forecast, and only because CBO learns of administrative decisions in the president's budget that may affect receipts or expenditures. Sometimes the spring projections also correct errors found in the January projections. The economic assumptions are seldom changed between the two projections. Chart 1 shows the cumulative economic and technical errors between the baseline and the final outcome as a percentage of gross domestic product (GDP).

Two things about the chart stand out. First, the absolute errors during the three years from 1998 through 2000 were unusually large. The only other comparable errors occurred in 1982 and 1991, years affected by unpredicted recessions. Second, the earlier period, 1980 through 1991, is distinctly different from the later period, 1992 through 2000. In the earlier period, forecasts were overwhelmingly too optimistic, with only two years—1984 and 1989—that were slightly too pessimistic. In the latter period, forecasts have been overwhelmingly too pessimistic. If CBO is pessimistic (optimistic) in one year, it is highly probable that it will be pessimistic (optimistic) the next year as well. Statisticians refer to this as hav-

CHART 2 Frequency Distribution of One-Year Errors as a Percentage of GDP, 1980–2000

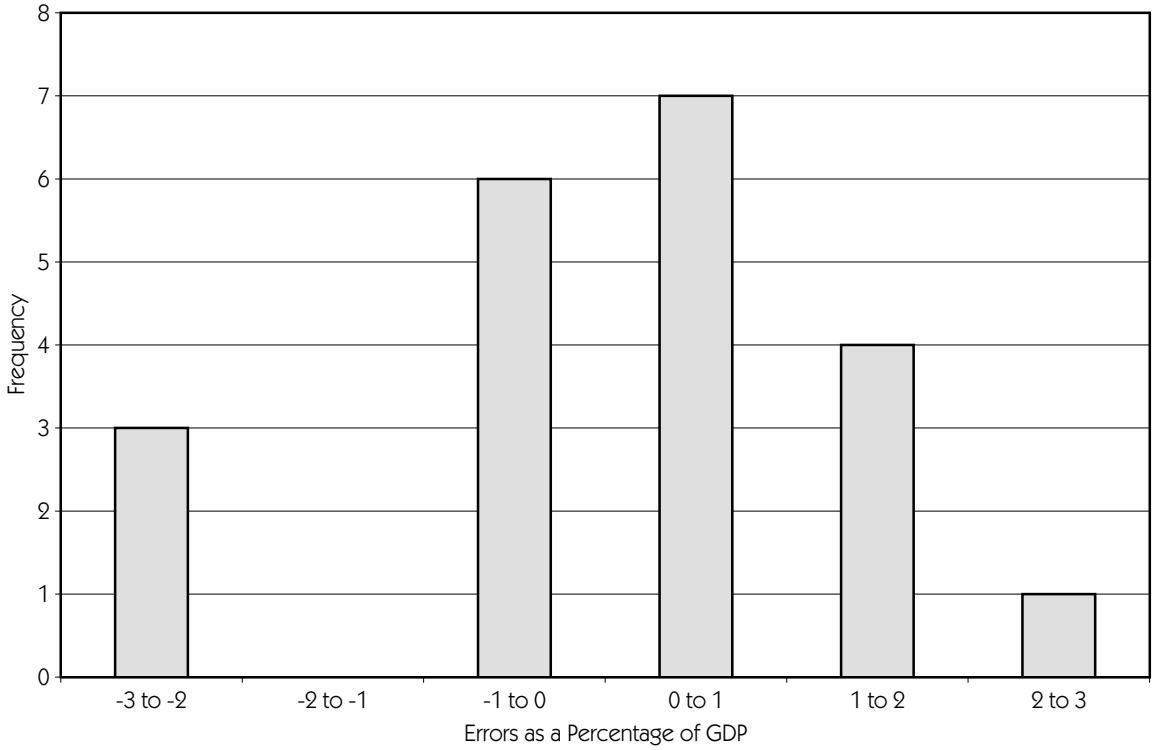
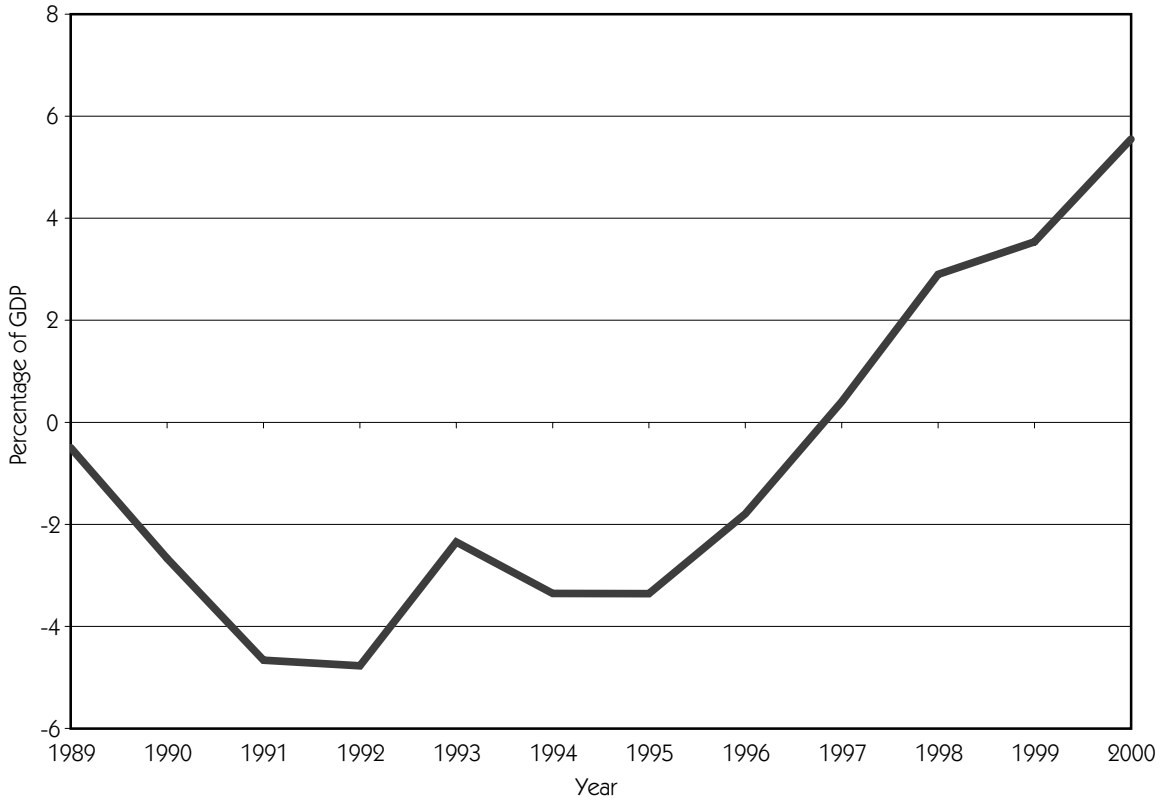


CHART 3 Economic and Technical Errors for Five-Year Forecasts, as a Percentage of GDP, 1980–2000



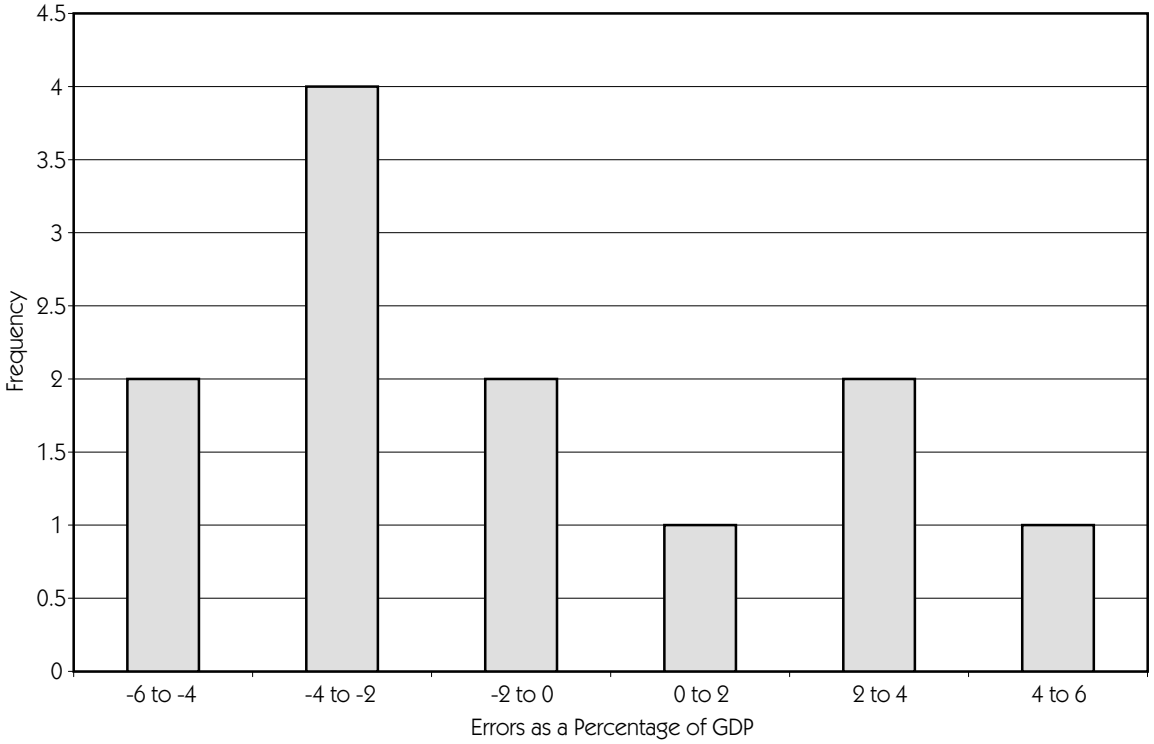
ing a high serial correlation among the errors. Possible reasons for the persistence of optimistic or pessimistic errors for long periods will be considered later.

Chart 2 shows that smaller errors are more common than larger errors. A statistician would say that the errors appear to be normally distributed. It must be emphasized that there are only 21 years of experience in which errors have been computed consistently; it is difficult to reach statistically sound conclusions with such a small sample.

The most important point about the errors is that they can be very large: In 4 of 21 cases, they exceeded 2 percent of GDP. At current levels of GDP, that would imply an error of over $215 billion in the spring 2001 forecast of the 2002 surplus. The average absolute error is 1.1 percent, implying an error of about $120 billion for 2002.[3]

The available sample for errors made in five-year projections is extremely small. CBO has only separated economic and technical forecasting errors from the effects of policy changes consistently since the projection made for 1989 in the spring of 1984. Of course, the errors in the spring baseline projections for the fiscal year five years out are huge relative to the errors for the following fiscal year. The errors for five-year projections are shown in chart 3.

CHART 4 Frequency Distribution of Five-Year Errors as a Percentage of GDP, 1980–2000



The most serious errors are for projections made in 1987 about fiscal 1992 and those made about fiscal 2000 in the spring of 1995. The errors exceeded 5 percent of GDP, equivalent to an error of over $660 billion in the budget balance forecast made for 2006 in the spring of 2001; the January 2001 projection of the 2006 surplus is only $505 billion. The average absolute error in the 12 observations in chart 3 is about 3 percent of GDP, equivalent to an error of almost $400 billion in the spring 2001 projection for 2006.

The frequency distribution of the errors in five-year projections is not very revealing, since the sample is so small. Nevertheless, it is somewhat disturbing that small errors have been less likely than large errors, as shown in chart 4.

Starting in the middle of 1996, CBO has made 10-year projections of the budget surplus. It is too early to test those projections against actual outcomes, but it is clear from adjustments made so far that those long-term projections will be extremely inaccurate. When the budget balance for 2007 was first projected, in the spring of 1997, the projection called for a deficit of $278 billion. By the summer of 2000, the projection had changed to a surplus of $523 billion, a difference of $801 billion. But policy changes used up some of the surplus in the interim, so the cumulative sum of economic and technical forecasting errors over the period was $844 billion, or over 6 percent of expected GDP. If errors continue to be seri-

ally correlated, the cumulative error will grow over time. If there is a shift toward pessimism, the cumulative error will shrink.

However, the change so far in the 2007 projection is huge relative to likely policy changes, and is worth far more than the tax cut proposed by President Bush. If the projections had not become so much more optimistic over time, it is unlikely that either presidential candidate would have promised substantial tax cuts during the campaign, and we probably would not be talking about policy initiatives such as a prescription drug benefit under Medicare. We can only hope that the projection does not become more pessimistic with equal rapidity. A comparable change over the next three years in our view of the 2011 budget balance would amount to over $1 trillion, an amount greater than the total 2011 surplus forecast in CBO's January 2001 report.

Projections of Receipts and Outlays

Errors in forecasting receipts are only slightly more important than errors in forecasting outlays in explaining the errors in one-year forecasts of the deficit. The absolute average error in forecasting receipts from 1980 to 2000 was 0.7 percent of GDP, while the average absolute error in forecasting outlays was 0.5 percent.[4] Simply adding up the positive and negative errors makes it clear that there is no statistically significant bias in forecasting outlays or receipts over the entire period, a proposition Auerbach's (1999) rigorous analysis confirms for receipts.

Errors in forecasting receipts and outlays can go in the same direction or in opposite directions, depending on the event that catches forecasters by surprise. For example, a surprising rise in inflation will raise both receipts and outlays, while a surprising recession will lower receipts but raise outlays on safety net programs. From 1980 to 2000, the sum of economic and technical errors on the receipts and outlay sides of the budget reinforced each other's effect on the budget balance in 15 of the 21 years.

In his careful analysis of errors in forecasting receipts, Auerbach documents the same serial correlation in these errors as discussed above for the budget balance. He divides the 1986–1999 period into pre-Clinton (1986–1993) and post-Clinton (1994–1999) eras, showing persistent optimism in the earlier period and persistent pessimism in the latter. The same type of persistence can be shown for errors in outlay forecasts, albeit with a somewhat different dividing line. Between 1980 and 1991, economic and technical errors caused actual outlays to exceed forecast outlays in all but 2 of the 12 years. Between 1992 and 2000, actual outlays fell short of the forecast in every year.

The examples described above imply that outlay and receipts forecasting errors reinforced each other in creating the serial correlation in the errors in the budget forecast. Although receipts and outlay errors were, on average, of similar relative importance between 1980 and 2000, very large errors in predicting receipts were the main reason for the surprising emergence of the surplus in

1998. Large errors occurred both because GDP growth was understated and because revenues rose unexpectedly as a share of GDP. The unpredicted rise in the stock market led to a surge in revenues from capital gains taxation, and a rapid rise in the share of income earned by the very rich raised the average individual tax rate. The corporate profit share of GDP also exceeded CBO's expectations and added to corporate tax collections.

The five-year projections of receipts and outlays show similar patterns, in that forecasts tend to be persistently optimistic early in the period and persistently pessimistic later. The errors, of course, grow as the forecast period lengthens. These longer-run forecasts will not be examined in detail here. The long-run errors in receipts and outlay projections manifest themselves in the very large errors discussed above for five-year projections of the budget balance.

The Forecasting Process

The procedures for forecasting budget totals are very similar at CBO and at the Office of Management and Budget (OMB), so it is not surprising that their past errors have also been very similar.[5] The process must begin with an economic forecast, because economic variables are crucial in projecting most types of outlays and receipts. At OMB, the economic forecast is produced by the "Troika," economists drawn from the Council of Economic Advisers, the Department of the Treasury, and OMB. In preparing their forecasts, economic forecasters in the administration and at CBO pay close attention to the forecasts made by business economists and private forecasting companies such as DRI and Macroeconomic Advisors. CBO, in particular, does not want to deviate far from the consensus, and receives input from a diverse set of academic and business advisers; to deviate significantly from the consensus would be considered to be partisan one way or the other.[6] During the Reagan and Bush administrations, OMB forecasts were often, but not always, somewhat more optimistic than the consensus forecasts, but during the Clinton administration, forecasts were much more conventional.

Theoretically, one would expect to see differences between CBO and OMB forecasts, since CBO bases its forecast on the fiscal stance implied by current policy, whereas the OMB forecast is based on the assumption that all presidential recommendations are adopted. Because most of the president's policy proposals are aimed at improving the state of the economy, one would expect the differences in policy assumptions to make the president's economic outlook more optimistic than CBO's. As a practical matter, however, other uncertainties in the forecasting process tend to overwhelm these differences in policy assumptions.

CBO carefully characterizes its estimates for the current and the next calendar year as a "forecast" and estimates for following years as projections. The projections assume that GDP follows a linear trend from the end of the forecast period to a point on the potential growth path of the economy 10 years hence. The potential or full-capacity growth path is deemed to be consistent with an

unemployment rate of 5.2 percent toward the end of the period. CBO explicitly notes that it does not mean to rule out business cycles that could cause GDP to vary above and below the potential growth path.

However, business cycles do not have a major long-run impact on the budget balance. When the economy recovers from a recession, the budget will as well. The only major lasting difference is an increase in the interest bill on the debt, because the reduced budget balance while the economy was depressed adds to the outstanding debt. A recession can have a longer-run indirect impact on the budget balance if it provokes a permanent tax cut or expansion of a safety net program. Policy uncertainty is always with us, however. For obvious political reasons, CBO and OMB avoid speculating on how Congress might change current law or react to presidential proposals.

The budget balance projected for the longer run is crucially dependent on the rate at which it is assumed the economy can grow without provoking inflation. If that rate falls from the currently assumed 3.3 percent to rates below the 2.5 percent commonly assumed in the 1980s, a very large proportion of the projected unified budget surplus disappears.

Dozens of experts have some input into the forecast; for example, energy experts comment on oil prices and agricultural experts suggest farm prices. In checking the forecast, forecasters make considerable use of accounting and other definitions. For example, accounting definitions are supposed to yield equal numbers for total domestic and net foreign investment and the total amount of domestic saving. Any forecast that caused the two numbers to deviate significantly would have been viewed with some suspicion in the past. Unfortunately, a large discrepancy between reported investment and saving has emerged in recent years. This indicates imperfections in the data-gathering system, because the data should show them to be equal. The imperfection means that we must forecast a statistical discrepancy, which is a bit like forecasting a random number. This is just one of many problems that make economic forecasting a challenging task.

Once economic forecasters have produced forecasts for variables like the consumer price index (CPI), unemployment, interest rates, and corporate profits, the analysts who compute program outlays and different types of receipts go to work. For example, specialists in the Medicaid program are responsible for projecting its costs. Medicaid is the third-largest entitlement program, and in the past errors in projecting its costs have been quite large. Medicaid costs depend on the number of people eligible for the program, a number that may be influenced by many factors, including the unemployment rate. Then the analysts must estimate the proportion of eligibles who will actually enroll in the program, Medicaid enrollees' per capita consumption of health services, and the cost of those services. Medical prices will, of course, be influenced by the overall inflation rate.

Errors are inevitable in each step of the process, which involves estimates for hundreds of programs and sources of revenue. A good forecast, which emerges if

the hundreds of errors offset each other, is largely a matter of luck. Highly competent analysts making hundreds of tiny errors that generally go in the same direction can make the forecasters of budget aggregates appear highly incompetent, whereas true incompetence may yield a satisfactory forecast, so long as individual analysts are incompetent in opposite directions.

Why Is the Direction of Forecast Errors So Persistent?

The analysis in this section is highly subjective. It is based on many years of watching budget analysts at work from the sanctity of think tanks and from my own participation in the forecasting process as chief economist at OMB from 1974 to 1977 and as director of CBO from 1983 to 1987. Based on this personal experience, I believe that there are a number of powerful forces that push forecasters in the same direction. As a result, an error in one direction in one year makes it highly probable that analysts will make an error in the same direction the next year.

There may be purely statistical reasons for this phenomenon. In some program and revenue estimating areas, analysts rely heavily on equations whose parameters are estimated on the basis of historical data. Typically, the equations are re-estimated periodically as new information comes in. However, adding one year's new data, even if it is highly unusual, does not change the estimates of the parameters very much. Consequently, if some important structural change has occurred in relationships among variables, it may not be recognized for many years, and forecasting errors in the interim are likely to all go in the same direction. Although such statistical phenomena may be of some importance, I suspect that the more informal judgments that analysts introduce into their forecasts are at least equally important. These judgments are often motivated by psychological and political factors that enhance the tendency for errors to go persistently in one direction or another.

In the Medicaid example used above, analysts attempt to model different components of the problem. They may modify the results of the models based on judgments about the effects of factors that are difficult to model. The important point is that they will consider all the information available to them and ultimately produce a forecast. Say, for example, that they decide that Medicaid outlays will rise 8 percent next year, but later find that the actual outcome is an increase of 4 percent. It would not be easy to identify the reason for the discrepancy quickly, because the requisite administrative and other data would only be available after a long time lag. Having used all the available information to produce the original forecast of 8 percent, there is a natural human tendency for forecasters to believe that the discrepancy must have been the result of an anomaly. If their techniques again yielded an estimate of 8 percent outlay growth for the following year, the analysts might think up some excuse for lowering the forecast to 7 percent just in case some unknown factor were working to depress Medicaid

costs, but unless they were certain that something important was changing, they would be very slow to adjust their estimates downward. They would not dream of lowering their forecast all the way to 4 percent.

This is not an irrational approach to forecasting. A whole branch of statistics, called Bayesian statistics, relies on such an approach. Analysts allow previously held views, presumably based on good information, to restrain statistical estimates of relationships among variables, thereby protecting the analyst against aberrations in the sample. If the restrained estimates do not work very well over time, prior views are adjusted, but the process of finding truth works slowly. A similar, formal approach to adjusting prior views is being tested in forecasting state and local government revenues. The approach applies so-called "neural networks." Assumed relationships among variables are gradually adjusted as the data come in, much as the brain processes information inconsistent with previously held views.

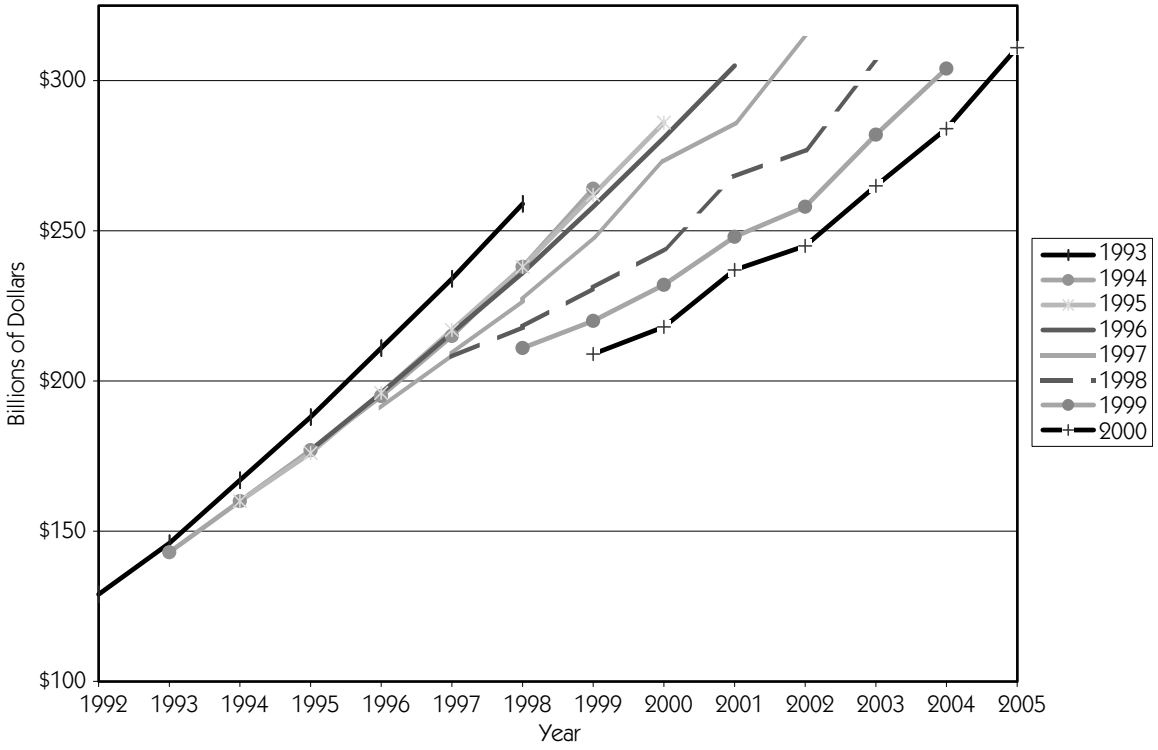
The gradual approach to adjusting a forecast is clearly indicated by Medicare cost projections from 1993 through 1999. Medicare costs are difficult to forecast, and Medicare projections generally contribute significantly to errors in projecting total outlays. The 1993 projection was clearly too high, but those between 1994 and 1996 were quite stable, 1996 being the first fiscal year for which 10-year projections were published in CBO documents. Between 1996 and 2000, Medicare outlay projections have constantly been adjusted downward as each year those outlays have fallen short of expectations. In the projection of 1995, Medicare outlays were forecast at $286 billion for 2000; the actual outcome was $218 billion, $68 billion (24 percent) lower. Legislation passed in 1997 explains part of the decline, but at the time, the cost-saving impact of the legislation was significantly understated (chart 5).

The above discussion implicitly assumes that significant forecasting errors occur mainly because the models used by analysts turn out to be wrong, either because the world changed and the analyst did not notice or because the model was poorly constructed. Errors caused by temporary statistical disturbances in a good model would not be serially correlated.

The models used by program analysts and revenue forecasters are often fairly crude and sometimes are little more than a collection of rules of thumb. Analysts seldom have time to construct highly sophisticated models. Moreover, it may not be worth trying to be highly sophisticated. Policy changes and changes in underlying economic relationships can cause the most sophisticated model to go astray. In other words, sophistication does not necessarily improve accuracy. Consequently, the art of forecasting may involve constantly adjusting crude models in a way that corrects the direction of errors very slowly.

There are other common forecasting procedures that tend to make errors persist. Where there is little basis for forecasting changes, as in financial markets, analysts tend to assume that values will return to some historical norm, regardless of where they begin. Forecasters often make that assumption, for example,

CHART 5 Medicare Spending Projections (in billions of dollars)



with regard to the relationship between short- and long-term interest rates and the ratio of realized capital gains to GDP. It was the assumption that relationships among financial variables would return to normal that got Long-Term Capital Management into so much trouble.

Assuming that things will return to normal, of course, still leaves the question of how fast the return to normalcy will proceed, but even beyond that the process is fraught with peril. Financial and other variables can remain at abnormal values for many years. If an analyst fails to predict a surge in the stock market and therefore underpredicts capital gains revenues for a particular year, it is extremely difficult to know what to predict for the next year. If the analyst assumes that a trend toward normalcy will begin and it does not, or it proceeds less rapidly than the analyst assumed, the forecast error will be in the same direction for a second year. An alternative would be to assume that the market is following a random walk in the short run and is about as likely to go up as down. It may then be appropriate to assume that the market will remain the same and to try to figure out what that implies for capital gains realizations. However, Congress would probably not be reassured to know that the revenue forecast was based in part on random-walk theory. Assuming a return to historical normalcy

seems prudent, but it often does not work very well and has been a significant contributor to the persistently over-pessimistic revenue forecasts in recent years.

A last psychological factor that intensifies the propensity for the errors' direction to persist afflicts everyone from technical analysts to the directors of CBO and OMB. If forecasting techniques and assumptions are changed significantly every time an error is made, long-run budget projections will jump all over the map from year to year, making the analysts appear incompetent and infuriating policymakers. Consequently, analysts or directors demand convincing proof that something fundamental has changed before significantly changing forecasting techniques or assumptions. Because it takes time to gather evidence of fundamental change, forecasting techniques and assumptions are likely to change more slowly than reality, and that lag may be the most important reason that errors persist for so long. The resulting slow change sounds much like the result of the Bayesian process described above. But the reluctance to make too many changes too quickly is an added factor that slows the adjustment of forecasts to reality. It might be called a political factor, but it is nonpartisan, and simply reflects a desire not to become too unpopular too fast.

The Usefulness of Flawed Forecasts

The damage done by forecast errors depends on how forecasts are used. For some purposes, inaccuracy is not much of a problem. In other contexts, an inaccurate forecast can badly mislead policymakers and media pundits.

Forecasts serve two important purposes in the budget process. They indicate the fiscal health of the nation and they provide assumptions that can be used to help evaluate tax and expenditure policies.

It has been shown that forecasts become rapidly less reliable as the forecast period is extended. Forecasts for periods beyond five years are not totally misleading, but it would be foolish to believe that they provide anything but the crudest indication of the nation's fiscal health. Given constant policies, the surplus for 2011 was projected in January 2001 to be $889 billion, or 5.3 percent of GDP. There is, however, a good chance that the final outcome could be twice that figure or zero, even without any policy changes. On the other hand, there is only the tiniest possibility that 2011 could see a deficit in the range experienced between 1982 and 1986, when the unified deficit varied between 4.0 and 6.1 percent of GDP. Although the uncertainty is enormous, the estimates indicate that the nation's budget is very likely to remain healthy or at least not become extremely unhealthy for the next 10 years.

Given that the long-run estimates provide some information, regardless of how crude it may be, they should continue to be produced. However, they need not be given the prominence that they now receive.

If 10-year forecasts are too inaccurate to be displayed prominently, what are we to make of the 75-year time horizon used to evaluate the financial well-being

of the Social Security system or the comparably long projections of budget aggregates periodically published by CBO, GAO, and OMB? It is important to repeat that despite severe inaccuracies, forecasts can be useful for some purposes even though they are useless for others.

It is also important to note that the 75-year projections made by the Social Security trustees are reported more sensibly than CBO's 10-year budget projections. No one would dream of adding up the trust fund surpluses and deficits for 75 years, nor is the cumulative dollar value of proposals reported. The present value of the deficit of the system is reported as a percent of the present value of payrolls over the entire period, and policy proposals estimate the extent to which they would raise or lower this deficit as a percent of payroll. Although these numbers are sensitive to economic and demographic assumptions made, they are reported in a way that makes their significance much easier to understand.

The press also pays a great deal of attention to each year's estimate of when the trust fund will be exhausted; this estimate is also highly sensitive to economic and demographic assumptions. Immediately after the 1983 Social Security reforms, it was thought that the trust fund would last until the 2050s. The date steadily fell to the late 2020s, and more recently it has risen again into the 2030s. But the main point is that the fund will eventually be exhausted under the so-called intermediate assumptions used by most observers. Not too much attention should be paid to estimates of the exact year in which this is supposed to occur.

The Social Security trustees' projections of the trust fund's deficit appear to be relatively stable from year to year. This is partly because the indexing of the system causes revenues and outlays to move in the same direction, as assumptions concerning economic growth and inflation are moved up and down. They are also stable because the trustees, when making 75-year projections, are even more reluctant than budget analysts making shorter forecasts to significantly alter very long range economic and demographic assumptions, regardless of how much those assumptions may be off from year to year. But stability should not be confused with reliability. The projections could be off by huge amounts. Remember that demographers failed to forecast either the postwar baby boom or the baby dearth that followed. The Social Security outlook would be very different today had these huge swings in the birth rate not occurred. The current projections are only valuable because they suggest that, under a wide range of assumptions, the system is not sustainable in its current form. However, precise estimates of the seriousness of the problem or of the exact year in which the trust fund is exhausted should be taken with a grain of salt.

The very long run budget projections by CBO, GAO, and OMB do not get much attention from the media. The main point of the projections is to show that current budget policy is not sustainable in the very long run because of the future burden to be imposed by Social Security, Medicare, and Medicaid as baby boomers retire and expected life lengthens. Eventually the national debt will explode toward infinity unless taxes are raised substantially or programs reformed. The

estimated date of the explosion can, however, be moved back and forth from the 2030s to the 2060s with very small changes in economic and demographic assumptions, which may be why the press does not take these projections very seriously. However, the projections' basic point cannot be denied: Current policy is unsustainable. That is a valuable conclusion, and given that the press and policymakers pay only cursory attention to the details of such projections relative to the attention paid to CBO's 10-year projections, the huge inaccuracies do little harm.

As noted above, budget estimates provide assumptions for evaluating policy changes. This is known as scorekeeping, a crucial part of the budget process. Policymakers must have estimates of the dollar value of policy changes. It is particularly important for them to know whether a change in an entitlement program will cost a great deal more in the long run than it does in the first few years, and, similarly, whether the long-run revenue effects of a change in tax law will dwarf its short-run effects. It is impossible to make such judgments without a long-run forecast.

CBO extended its forecast period from 5 to 10 years when the Senate extended the time horizon for applying its pay-as-you-go rule from 5 to 10 years. That rule insisted that tax cuts or entitlement increases be paid for by tax increases or entitlement cuts. When the pay-as-you-go rule was applied with a five-year horizon, it became common for legislators to propose phased-in tax cuts or spending increases that cost much less annually during the first five years than they did in the longer run. Ten-year scorekeeping was introduced to limit such gaming.[7] The estimate of the precise quantitative effect of the policy change in years 5 and 6 may be inaccurate, but even so an assessment of the qualitative nature of the pattern of effects is unlikely to be badly misleading. This is especially true if a policy has been designed to delay its effects until after the time horizon used for scorekeeping, because then the change in the time path of effects is very apparent.

If a 5-year scorekeeping horizon encourages gaming between years 5 and 6, it is natural to ask whether a 10-year horizon will encourage gaming between years 10 and 11. It probably will, but as the scorekeeping horizon lengthens, gaming of this type becomes less of a concern. A policy change enacted this year is likely to be altered many times over a 10-year period and may not even be recognizable by year 11. Moreover, if the largest effects of a policy change are long delayed, it will be much less attractive politically.

Although attempts at gaming and other reasons for large swings in the costs of a policy change are likely to be identified even if estimates are based on a bad forecast, a bad forecast may, of course, sometimes substantially misrepresent the time pattern of a policy change's budget effects. For example, the time pattern of the effects of a change in the taxation of capital gains will obviously depend on the time pattern of capital gains realizations, which are notoriously difficult to forecast. Our view of the cost path for a new entitlement may be badly distorted if we inaccurately forecast the participation rate of those who are eligible. But

in general, the problems bad forecasts pose for scorekeeping are minor relative to the value of having a rough idea of how the costs or savings associated with a policy change are likely to vary over the long run.

Recommendations for the Congressional Budget Office

- *CBO should de-emphasize forecasts produced for the second 5 years of the 10-year forecast period.*

CBO should show the estimates for years 6 through 10 of the projection period in an appendix of their reports, rather than in the main body of the reports, to indicate that they regard the later projections as being considerably less reliable than shorter-run forecasts. The division of the forecasts into two five-year periods is somewhat arbitrary, because the reliability of the forecast declines steadily as the period lengthens (i.e., the forecast for year 5 is only somewhat more accurate than the forecast for year 6). However, the tradition that a five-year projection is appropriate was established by the Budget and Impoundment Control Act of 1974, and five years is a reasonable projection period, even though it may contain very large errors.

- *CBO should not provide a cumulative value for the budget balance or policy changes.*

Even if the projection period that is given prominence is only five years, it is extremely misleading to add the surplus projection for year 5 to the surplus projection for year 1 to provide an estimate of the cumulative surplus. The process of adding up surpluses gives equal weight to longer and shorter projections, implying that the estimate for year 5 is as reliable as the estimate for year 1. In truth, however, longer-term projections are much more uncertain.

- *CBO should continue to publish and refine estimates of the degree of confidence that should be attached to estimates of the budget balance.*

In the past, CBO has found it difficult to describe the degree of uncertainty that should be attached to its estimates. In its report of January 2001, CBO introduced a new approach, providing a distribution of the probabilities of different outcomes, given the baseline estimates. Because CBO's history is not long enough to provide a large sample of past errors, any probability distribution based on the historical record is necessarily crude, but it is nonetheless extremely valuable. The CBO report also describes the budget implications of different economic scenarios, and it includes sensitivity tables showing the impact that errors in the economic forecast might have on the budget. For example, the tables show the effects of a 0.1 percent annual error in the assumed growth rate and of errors in estimating specific amounts of inflation. While such analyses are useful and should

be continued, their usefulness is limited because of the difficulty of assessing the likelihood of a particular economic scenario or of a forecast error involving only one economic variable.

Recommendations for Congress

- *Congress should not insert quantitative targets into laws governing the budget process.*

Forecasting is not accurate enough to support laws that set quantitative targets for budget totals. The Gramm-Rudman-Hollings Act of 1985 (GRH) required Congress to balance the unified budget within five years and specified a precise path of declining deficits on the way to the ultimate target. If targets were not met, spending would be cut according to a complex formula. The law was bound to fail, and it did. When the economy failed to perform as well as forecast, the size of the tax cuts or spending reductions necessary to meet the targets became totally implausible, as was doing so entirely on the spending side using cuts determined by an automatic formula.[8]

The same fate will await any of the "lock box" laws recently discussed in Congress. They would make it illegal to run deficits outside the Social Security system, presumably enforcing the goal with some type of automatic mechanism. It will be no easier to balance the budget under this definition than it was to balance the unified budget under GRH.

So how do states manage to create the balanced budgets mandated by their laws or constitutions? Sometimes they do it honestly, with real tax increases and spending cuts, but often they use bookkeeping changes and artificial changes in the timing of expenditures and revenues (GAO 1993).

At both the federal and the state level, difficulty in balancing the budget—however balance may be defined—is usually caused by an unanticipated weakening of the economy. It is not appropriate for the federal government to be raising taxes and cutting spending under these circumstances, and while it is not desirable for states and localities to be doing it either, the nation has learned to live with this perverse reaction.

- *Congress should be cautious about promising to hit precise budget targets.*

Although they create less of a problem than precise, legislated targets, more informal promises involving very precise budget goals are also likely to be violated. For example, reducing the national debt to below 20 percent of GDP is not a bad goal, but promising that it will be accomplished in 2005 is hazardous. Similarly, it may be appropriate to have the goal of saving the Social Security surplus—that is, balancing the budget outside of Social Security—but to imply that it is a mortal sin for the rest of government to borrow from Social Security is unwise. As

noted above, if Congress aims at the target of balancing the non–Social Security budget, forecasting errors will certainly cause it to miss this target from time to time—sometimes by very large amounts that imply very large deficits.

- *Congress should devote more effort to analyzing the risks of outcomes that may occur because the forecast is wrong.*

Because of the uncertainty that must be attached to projections, policymakers proposing a specific policy change should not only assess its desirability given the current set of budget and economic projections, they should also examine its implications if current projections turn out to be far too optimistic or pessimistic.

During the 1980s, such analysis would have yielded a clear message: Policies that would have increased the deficit could have been extremely harmful if budget forecasts turned out to be overly optimistic. Projected deficits were so large that they threatened to become dangerous, eventually causing the debt to rise so rapidly that interest costs would soar and the country would be unable to keep up by raising taxes or cutting taxes. Many countries have resorted to printing money and have experienced hyperinflation under those circumstances. If the projections of the day had turned out to be even more overly optimistic than they were, the danger of an exploding debt would have grown more rapidly and have been harder to control. It was clear that Congress should err on the side of caution and pursue a highly constrained fiscal policy. They essentially did so, although perhaps the policy should have been even more constrained, given the circumstances.[9]

As noted above, it now appears highly unlikely that we will return to the dangerous deficits of the 1980s even if current projections turn out to be overly optimistic. In contemplating tax cuts or spending increases, Congress faces a more symmetrical set of risks. If it avoids a major tax cut and the projections become consistently more optimistic, the nation will not have the increase in private consumption that might otherwise have been enjoyed. If Congress passes major tax cuts and the projections become more pessimistic, the nation may not save as much as it should, given the approaching retirement of the baby boomers. In particular, future surpluses may fall short of the surplus in the Social Security trust fund—a goal now enunciated by a majority of Congress.

There may be a slight asymmetry in the risks, in that it is probably somewhat easier to cut taxes and increase spending if projections become more optimistic than it is to do the reverse if projections become more pessimistic. However, Congress did manage to raise taxes and restrain civilian spending during the 1980s, so it is not impossible. But the bottom line is that the risks are now much more symmetrical than they were through much of the 1980s, and the way that they should be balanced is much less apparent.

Because of the currently superb fiscal situation, it should also be emphasized that the risks of being wrong are not very large. If, between now and 2011, $1 trillion less

in debt is retired than planned, interest payments will be $50 to $60 billion higher than anticipated, representing about 2.5 percent of outlays. Economic growth will be minutely lower, but if the shortfall is not the result of slower-than-anticipated growth, private consumption will be higher in the interim. The effect on the nation's well-being will be trivial relative to the $17 trillion GDP expected in 2011. If the shortfall is the result of lower-than-anticipated growth, it is appropriate to lower the target for the surplus in any case. The lower growth will cause a significant loss of welfare, but the fiscal policy mistake will not.

- *Congress should seek to improve the quality of economic statistics.*

A last suggestion for Congress is a favorite of economists, but has few other constituencies: The statistical agencies of the government need larger budgets. The agencies currently cannot do the research necessary to adjust their data-gathering practices to rapid changes in the structure of the economy or to otherwise improve the quality of data used in forecasting efforts. In addition, very small increases in the statistical budgets of other agencies could improve the quality and timeliness of administrative data generated by spending programs and the tax system.

Better data would probably result in improvements in the accuracy of forecasts in the long run, as better data improves our understanding of how the economy works. There would still be enormous gaps in our knowledge of economic relationships and severe problems in forecasting the exogenous events that have a large impact on the economy, so having better data would result in, at best, minor improvements in forecasting accuracy, but every little bit helps.

Recommendations for the Media

- *The media should change the language it uses to describe budget totals and policy changes.*

Although the language of the policy debate seems to be fairly set in the media lexicon, the media should consider the fact that the discussions of budget totals or the budget effects of tax and spending policies that are cumulated over long periods are misleading, and would not be very descriptive even if they were not misleading. As was noted previously, adding the projected surplus for 2002 to the surplus for 2006 makes little sense because the two numbers differ so much in their reliability. In other words, it is frequently reported that the projected surplus for the 2002–11 period is $5.6 trillion. If, next summer, projections change and $100 billion is subtracted from the 2002 surplus, but $100 billion is added to the 2011 surplus, the cumulated surplus will still be $5.6 billion, but the sum's reliability will have fallen considerably.

More generally, reporting on spending and tax issues in the last campaign left much to be desired, especially in terms of conveying a sense of how much uncertainty should be attached to estimates of the surplus and to the estimated value of the candidates' tax and spending initiatives. There was much discussion during the campaign as to whether the candidates' proposals more than used up the entire non–Social Security surplus, but the argument was more than a little absurd, because no one could possibly know within hundreds of billions of dollars what the surplus would be or the true costs of the candidates' proposals. Between January and July of 2000, CBO increased its cumulative surplus estimate by more than enough to finance the entire direct cost of Bush's promised tax cut, but the increase received little media coverage.

Quantitative comparisons of the size of various candidates' tax and spending proposals should not be avoided entirely, but they should be reported with much more caution. There are ways of reporting the size of tax or spending initiatives that are less susceptible to forecast error than is an absolute dollar sum. For example, instead of saying that a tax cut is worth $1.6 trillion—a figure that must mean little to most Americans—the cut could be reported as equal to slightly more than 5 percent of average revenues through 2011. If the reporter wanted to emphasize the change in income tax policy in the long run, the cut might be described as being about 12 percent of personal income tax revenues, when fully phased in.

Because most surprising economic events will push total revenues and the size of the tax cut in the same direction, a report of the percentage cut is likely to be less susceptible to forecast errors than a dollar figure. It would be even more revealing to report the percentage tax cut in 2002 and in 2006, to show whether the cut is growing quickly or whether the revenue loss is relatively stable. Similarly, the cost of a prescription drug plan could be reported as a percent of total outlays or as a percent of current Medicare outlays.

- *The media should place more emphasis on the qualitative characteristics of proposed policies.*

Given that numerical estimates are extremely unreliable, it would be instructive to place more emphasis on debates regarding the qualitative merits of policy proposals rather than discussing whose numbers are correct. The latter debate is unlikely to be easily resolved.

Issues of equity or economic efficiency are unlikely to be resolved either, but it is much easier for readers to understand the relevant arguments. For example, discussions of tax proposals should ask whether the tax cuts simplify or complicate the system. Are they efficient? Are they fair? If they are targeted, are they likely to be effective?

Conclusions

Because budget projections tend to be highly inaccurate and are unlikely to get better soon, policymakers, policy analysts, and the media must live with tremendous uncertainty. That uncertainty should be recognized more explicitly than it has been in the past, and the dialogue regarding policy decisions should pay more attention to the risks of being wrong. Almost everyone involved in the policymaking process recognizes this intellectually, but their time is scarce and it is difficult enough to analyze policies under the assumption that the future is certain. Nevertheless, relatively simple changes in language, more caution in discussing and legalizing precise quantitative targets, and a more detailed discussion of the risks associated with bad forecasts have the potential to significantly improve the policy dialogue.

Addendum on Technical Issues

Because of the size of the errors in past forecasts, there are sometimes calls for radical changes in the forecasting techniques used by CBO and OMB. As noted above, the budget forecast process begins with an economic forecast that usually relies heavily on results derived from complex structural models of the economy. Robert Lucas (Lucas 1976 and 1980, Lucas and Sargent 1978) won a Nobel Prize for critiquing such models. He argued that the structure of the economy is constantly changing as decisionmakers learn more about how the economy functions and as important changes continually occur in technologies and other relationships. In particular, as people learn more about how policies are formulated, their reactions to policy changes will be altered. For example, if people come to expect that government will provide a tax incentive for investment during recessions, they might delay investments in order to enjoy the tax advantage if a recession appears imminent. Such behavior would, of course, make a recession more likely.

Lucas concluded that forecasts derived from the assumption that people do not learn such things and that the structure of behavior is constant were bound to be wrong. The fact that practical forecasters constantly make judgmental changes to individual equations in structural models was taken as evidence that they are accommodating structural changes in the economy.

The alternative to using structural models is to use analyses of variables over time that are not restrained by specific structural assumptions. Past patterns in time series may be able to tell us something about future patterns. Using these new techniques, it might be possible to forecast things like personal income tax revenues and Medicare outlays directly without going through the step of first making an explicit, detailed economic forecast.

CBO experimented with such techniques in the mid-1980s—specifically with vector auto regression approaches—but they yielded totally implausible forecasts. By comparison, forecasts based on structural models did not look so bad.

When practical forecasters judgmentally adjust equations in a structural model in the hope of improving the model's forecast, it is unlikely that many believe that they are adjusting for changes in the basic structure of the economy. Even if the economy's structure is constant, it would be remarkable if it could be described with complete accuracy by a system of equations. Because of problems in describing the economy and fitting the equations to historical data, a blind run

of a complex model will probably provide implausible results for particular variables. If the forecast for, say, business equipment investment is not credible, analysts will adjust the relevant equations. The beauty of a structural model is that one can quickly assess the effects of an adjustment in one place on all other variables in the system. If correcting an absurdity in one area creates a problem in another area, the analyst quickly returns to the drawing board and makes more adjustments. Forecasting becomes an iterative process, and although the forecast may turn out to be wildly wrong, it is, at least, logical.

Most time-series analysis does not provide the same checks on the forecaster's judgment as a structural model does. Despite my earlier contention that the serial correlation of errors often occurs because analysts take a long time to recognize fundamental structural changes in economic relationships, I cannot imagine relying heavily on time-series analysis, given the current state of the art. Unfortunately, we are a very long way from having a solution for this problem where the cure is not worse than the disease.

I once suggested reforming the process in a radically different direction (Penner 1982). Given that forecasts were so inaccurate, I thought it might be preferable to rely on projections based on simple rules of thumb. My specific suggestion was a simple extrapolation of the last five years of history. The process would be simple, transparent, and immune from political influence and the biases of technical analysts. It would avoid creating the illusion that we could forecast with accuracy.

Unfortunately, the suggestion was made at a time that revealed the proposal's weakness. In the early 1980s, the rampant inflation of the late 1970s came to an end much more quickly than anyone thought possible. Basing the budget projections of the early 1980s on an extrapolation of the inflation rates of the late 1970s would have been more wildly misleading than the most inaccurate forecast imaginable.

A similar break in historical trends occurred in the late 1990s, when economists constantly raised the rate at which they assumed the economy could grow without provoking inflation. Again, budget projections based on historical economic trends would have been even worse than the forecasts that were actually used.

It seems evident that there is no easy way of improving on current procedures. CBO and OMB should continue to experiment with novel approaches, including time-series analyses and perhaps neural networks, but it is unlikely that we will see old-fashioned structural models of the economy abandoned any time soon. CBO, OMB, and private forecasters all rely on them to a considerable degree, and none of the proposed approaches is significantly more accurate than any of the others, either in the two-year forecasts so important to the budget process or in their five-year projections. If there were better ways to forecast, private forecasting companies would clearly have a very powerful incentive to adopt them.[10]

Notes

The author would like to thank Barry Anderson, Joseph Antos, Alan Auerbach, William Hoagland, Arlene Holen, Richard Kogan, and Carol Wait for extremely useful comments. He would also like to thank the Smith Richardson Foundation for financial support.

1. Since this analysis was completed, CBO has altered its techniques for estimating errors in outlay projections. All changes in projections of discretionary spending are now defined as legislative, reducing the measured forecast error on the outlay side of the budget. Whether this change is appropriate involves a number of complex issues that will not be discussed here. The basic point is that the division of changes in the projections into legislative, economic, and technical changes is somewhat arbitrary. Different approaches yield somewhat different results, but changes in definitions do not alter the conclusion that errors are very large.
2. There are gaps in CBO's records that prevent a completely accurate separation of the effects of policy changes on the budget deficit from the effects of economic forecasting and technical errors made in the early 1980s. However, the gaps are small, and it is unlikely any of the important conclusions of this analysis would be altered if discrepancies in the data were completely resolved.
3. Because of the change in the definition of forecast error reported in note 1, CBO reports a somewhat smaller possible error using a similar sample period. See *The Budget and Economic Outlook,* 2001.
4. CBO reports that outlay errors are considerably smaller relative to revenue because of the change in definitions reported in note 1.
5. The accuracy of the economic forecasts of the two agencies are compared in *The Budget and Economic Outlook: An Update*, July 2000, Appendix B.
6. Indeed, as director of CBO, I sometimes semi-seriously characterized our process as forecasting what the consensus forecast would be several months hence, when we would be defending our forecast before Congress.
7. Section 303 of the Congressional Budget Act provides a point of order against having any provisions, including phase-ins, effective beyond the time period covered by the budget resolution. However, the provision does not apply to revenue measures in the House and can be waived by a majority of senators. The provision has not been effective in the past. Ten-year scorekeeping can, at least identify the gaming that this provision tries to prevent.
8. For a more favorable account of the effects of GRH, see Gramlich (1990).
9. Carol Wait has suggested that when risks are asymmetrical, Congress should consciously bias the forecast in an optimistic or pessimistic direction. While this notion has some appeal, it would not be appropriate to push CBO to provide biased forecasts, since it is CBO's obligation to give Congress its best judgment. The budget committees could introduce such a bias,

but when they have deviated from the CBO forecast in the past, it has generally been in an optimistic direction.

10. Bachman (1996) provides an excellent description of how practical forecasters approach their task.

Bibliography

Aaron, Henry J. 2000. "Presidential Address—Seeing through the Fog: Policy-making with Uncertain Forecasts." *Journal of Policy Analysis and Management* 19 (2): 193–206.

Auerbach, Alan J. 1999. "On the Performance and Use of Government Revenue Forecasts." *National Tax Journal* 52 (4): 767–82.

Bachman, Daniel. 1996. "What Economic Forecasters Really Do." The WEFA Group, Bala Cynwyd.

Congressional Budget Office. 1987–2001a. *An Analysis of the President's Budgetary Proposals* Washington, D.C.: U.S. Government Printing Office.

———. 1987–2001b. *The Budget and Economic Outlook*. Washington, D.C.: U.S. Government Printing Office.

———. 1987–2001c. *The Economic and Budget Outlook: An Update*. Washington, D.C.: U.S. Government Printing Office.

Council of Economic Advisors. 2000. *Economic Report of the President: February 2000*. Washington, D.C.: U.S. Government Printing Office.

GAO. *See* U.S. General Accounting Office.

Gramlich, Edward M. 1990. "U.S. Federal Budget Deficits and Gramm-Rudman-Hollings." *American Economic Review* 80 (2): 75–80.

Lucas, Robert E. 1980. "Methods and Problems in Business Cycle Theory." *Journal of Money, Credit and Banking* 12 (2): 696–713.

———. 1976. "Econometric Policy Evaluation: A Critique." In *The Phillips Curve and Labor Markets*, edited by Karl Brunner and Alan Meltzer. Carnegie-Rochester Conference Series on Public Policy (1). Amsterdam: North Holland.

Lucas, Robert E., and Thomas Sargent. 1978. "After the Phillips Curve." In *Rational Expectations and Econometric Practice*, edited by Robert E. Lucas and Thomas Sargent. Minneapolis: University of Minnesota Press.

Penner, Rudolph. 1982. "Forecasting Budget Totals: Why Can't We Get It Right?" In *The Federal Budget: Economics and Politics*, edited by A. Wildavsky and M. Boskin (89–110). San Francisco: Institute for Contemporary Studies.

U.S. General Accounting Office. 1993. *Balanced Budget Requirements: State Experience and Implications for the Federal Government*. GAO/AFMD-93-58BR. Washington, D.C.: General Accounting Office.

About the Author

Rudolph G. Penner, an internationally known authority on government budgeting and public finance, holds the Arjay and Frances Miller Chair in Public Policy at the Urban Institute. He directed the U.S. Congressional Budget Office from 1983 to 1987 and has had a distinguished career as a senior government official at the Council of Economic Advisers and the U.S. Department of Housing and Urban Development. Dr. Penner has also served as chief economist at the Office of Management and Budget and has taught at several universities, including the University of Rochester and Princeton University.

The Urban Institute
2100 M Street, N.W.
Washington, D.C. 20037

Phone: 202.833.7200
Fax: 202.429.0687
E-Mail: paffairs@ui.urban.org
http://www.urban.org

Research Reports