Congressional Budget Office

Background Paper

Improving CBO's Methodology for Projecting Individual Income Tax Revenues

February 2011

THE CONGRESS OF THE UNITED STATES

Pub. No. 4038

Improving CBO's Methodology for Projecting Individual Income Tax Revenues

February 2011

The Congress of the United States ■ Congressional Budget Office

Preface

In preparing its annual report on the budget outlook, and updates to that report during the course of the year, the Congressional Budget Office (CBO) projects revenues from the federal individual income tax. This background paper discusses two possible ways to use information about tax collections to improve CBO's projections of tax receipts: Explicitly using the information provided by recent tax collections to adjust the projections, and basing the projections on multiple years of information from tax returns rather than relying on just the most recently available year. The results of CBO's analysis suggest that combining the two approaches—that is, using the information from recent tax collections and relying on multiple years of tax return information—can modestly improve the near-term projection of individual income tax revenues. Those findings have led CBO to use that combination of approaches in estimating revenues in recent years.

Kurt Seibert, Mark Booth, and David Weiner of CBO's Tax Analysis Division wrote the paper with assistance from Regis Barnichon, formerly of CBO, under the guidance of Frank Sammartino. Robert Arnold and Janet Holtzblatt of CBO and George Plesko of the University of Connecticut's School of Business provided helpful comments. (The assistance of an external reviewer implies no responsibility for the final product, which rests solely with CBO.)

Leah Mazade edited the paper, Chris Howlett proofread it, and Jeanine Rees prepared the report for publication. Monte Ruffin produced the print copies, Linda Schimmel coordinated the print distribution, and Simone Thomas prepared the electronic version for CBO's Web site (www.cbo.gov).

Douglas W. Elmendorf
Director

February 2011

Contents

Tables

Figures

Improving CBO's Methodology for Projecting Individual Income Tax Revenues

Summary and Introduction

Each January in *The Budget and Economic Outlook*, the Congressional Budget Office (CBO) publishes its projections of federal revenues spanning 11 years: the current fiscal year (that is, the fiscal year already under way at that point), the "budget year" (that is, the first full year after the year in which the projections are produced), and the nine fiscal years that follow. Those projections incorporate the assumption that the laws affecting revenues remain unchanged.[1] Since 1997, receipts of individual income taxes have differed from CBO's January projections for the current year by an average of about 6 percent—or by an average of 0.4 percent of gross domestic product (GDP). On average, receipts of individual income taxes for the budget year have differed from the projections made in January of the previous year by about 11 percent (1.0 percent of GDP). All of those differences have been calculated so that they exclude the effects of legislation enacted after the projections were prepared. They have also been calculated without regard to whether the discrepancies were positive or negative—that is, whether CBO overestimated or underestimated actual receipts—and are known as mean absolute errors.[2]

An analysis of the errors in CBO's past projections of individual income tax revenues suggests that, over time, the agency's estimates are about as likely to be too high as too low. For the years since 1997, positive and negative errors have largely offset each

1. CBO produces projections of federal revenues but does not estimate the revenue effects of individual pieces of legislation. That latter task is carried out by the staff of the Joint Committee on Taxation. CBO usually updates its baseline projections of revenues and spending in March and August; this analysis focuses on the projections that CBO releases in January.

2. For example, an overestimate of 5 percent for one year and either an overestimate or underestimate of 3 percent for another year would produce a mean absolute error of 4 percent. If the calculation for the current year had included projections for years before fiscal year 1997, the mean absolute error would be smaller than 0.4 percent of GDP.

other; so, for example, the average *net* error over that time in CBO's projections for the current fiscal year is 0.6 percent, and the average *net* error for the budget year is 1.3 percent (less than 0.1 percent of GDP in both cases).[3] Nevertheless, CBO's analysis also suggests that errors in the projections of revenues for the current and budget years tend to be correlated over time; that is, if the projections made in one January for the current year and the budget year were too low, the projections made the next January would probably also be lower than actual revenues in those years.

This paper analyzes two approaches for improving the results from the model that CBO uses to project revenues from individual income taxes.[4] The results of this analysis have been incorporated in CBO's revenue forecasting.

In the first approach, CBO sought to determine the best way to use the most recent information about tax collections to minimize errors in its forecasts. When actual collections differ from projections, forecasters must judge whether the discrepancy they observe is a temporary deviation or the beginning of a longer-term trend and thus whether and by how much to adjust the model's projections of future revenues to reflect that discrepancy. Such determinations are complicated, though, because detailed tax return data that could identify the sources of the difference are not available for some time. In the absence of such data, CBO examined whether the accuracy of its projections would improve if it simply assumed that actual revenues would differ from the model's projection in future years by the same percentage that they differed in the most recent year for which actual revenues were known.

In the second approach, CBO analyzed the improvement to be gained from using additional years of tax return data as the basis for its projections. CBO's individual income tax model starts with the most recent year of detailed tax return data available. But those so-called base-year data may not be representative of longer-term trends. Therefore, CBO investigated whether the model might achieve more-accurate results by producing multiple projections using tax return data from different years (rather than data from only a single year) as the starting point and averaging the projections.

Using a simplified version of its individual income tax model and examining projections for the 1984–2005 period, CBO evaluated various versions of those two approaches as well as combinations of the two. CBO's findings included the following:

---

3. Unlike the calculation of mean absolute errors, calculations of net errors incorporate the offsetting effects of positive and negative differences. Thus, a 5 percent overestimate for one year and a 3 percent underestimate for another year would produce an average net error of 1 percent.

4. CBO's individual income tax model is a stylized representation of the economy that relies on various simplifying assumptions; as such, it cannot capture all of the determinants of the taxes that people owe. CBO's projections of individual income tax revenues thus combine information from the individual income tax model with the most recent available data on tax collections.

- Projections of tax liabilities (what taxpayers owe) were more accurate—but only in the very short term—when CBO adjusted the model's results by the percentage error in those results in the most recent year for which actual revenues were known (called the "alignment year" in this paper).

- The use of multiple years of detailed tax return data reduced the size of the error in CBO's projections compared with the use of only the most recent year of such data. Using four years of tax return data to produce four separate projections, and then averaging them, yielded the best overall results.

- A combination of the two approaches was better than either one by itself. Of the options that CBO examined, the combination that was most effective in reducing errors was to average projections based on the four most recent years of tax return data and to adjust the estimated liabilities for the first year by 50 percent of the alignment-year error.

Those improvements have limitations, however. The reduction in the size of the errors is small relative to the errors seen before making those adjustments. The reason is that the use of the most recent information can do little to reduce the uncertainty that surrounds the future course of the economy and the rate of growth of income in the tax base. Also, the degree of improvement in the projections' accuracy declines rapidly over the 10-year forecast horizon.

Analyzing CBO's Estimates of Individual Income Tax Revenues

Studies that have examined CBO's historical "track record" in forecasting revenues have concluded that although errors in its projections are equally likely to be too high as too low, revisions to the projections and the projections' deviations from actual outcomes have been correlated over time.[5] For example, one study (by Alan J. Auerbach) examined 10-year revenue forecasts produced from 1986 to 1999 and found that when CBO increased (or lowered) its revenue projections from one forecast to the next, it tended to raise (or lower) them again in the next forecast. Another study (by Rudolph G. Penner) looked at CBO's projections over the 1984–2005 period and found that underestimates (or overestimates) of revenues for a year tended to be followed by underestimates (or overestimates) of revenues in the following year's forecast.

5. See Alan J. Auerbach, "On the Performance and Use of Government Revenue Forecasts," *National Tax Journal*, vol. 52, no. 4 (December 1999), pp. 767–782; and Rudolph G. Penner, "Federal Revenue Forecasting," in Penner, *Handbook of Government Budget Forecasting* (New York: Taylor & Francis, March 2008). See also George A. Plesko, "The Accuracy of Government Forecasts and Budget Projections," *National Tax Journal,* vol. 41, no. 2 (December 1988), pp. 483–501. All three studies considered projections of total revenues rather than projections of individual income tax revenues. The Plesko study found that, after legislative changes were taken into account, CBO's revenue forecasts were equally likely to be too high as too low, but it did not examine the correlation of errors.

Both findings suggest that CBO was not incorporating new information quickly enough into its forecasting methodology.

An analysis of CBO's forecasting record could indicate some ways to improve the accuracy of the agency's projections, but several factors complicate such an assessment.

- Forecasting methods change over time; the models and methods in use today would have produced different results than the results produced years ago with different methods. For example, adjustments to CBO's individual income tax model to reflect the most recent information from tax collections have changed over the years, making it difficult to isolate the effects of different methods of incorporating that information.

- Past errors in forecasting may have occurred for reasons unrelated to the two issues that are the focus of this analysis. For instance, CBO may have projected tax liabilities correctly but misestimated the fiscal year in which payments were made.

- The number of forecasts that CBO has produced over the years is relatively small for the purposes of statistical analysis and the development of reliable conclusions.

- Numerous changes in the tax code make it difficult to separate the effects of those changes from the effects of other factors. Because CBO's forecasts by design incorporate the assumption that for the 10-year projection period, current tax laws remain unaltered by future legislation, any assessment of CBO's track record in revenue forecasting must take into account the effects of subsequent legislation.

Those complicating factors would affect any analysis that relied on identifying and assessing errors in CBO's historical forecasting record. As a result, for this analysis, CBO instead recreated past forecasts by using a simplified version of its individual income tax model. This paper addresses the relationship between the error in CBO's estimate for the last known liability year and the errors in future years in CBO's model for projecting individual income tax liabilities.[6] It does not address other elements and characteristics of errors in the estimates of individual income tax revenues, nor does it address errors in projections of revenues from other sources.

The Timing of Tax Payments and the Availability of Tax Data

Two factors are critical in understanding the relationship between CBO's individual income tax model and its projections of tax receipts: the timing of those tax collections and the lag in the availability of detailed data from tax returns.

6. The liability year is the calendar year in which the tax liability is incurred, regardless of when payment is made.

CBO's individual income tax model produces estimates of tax liabilities for a particular tax year—generally, a calendar year. Those projections of liabilities are then converted into projections of collections by incorporating expectations about when taxpayers will pay what they owe. Some portion of those liabilities will be collected throughout the calendar year in the form of withheld and estimated tax payments; most of the rest will be collected the next year, after tax returns are filed.

In January 2010, for example, CBO released projections of federal revenues from the individual income tax for fiscal years 2010 through 2020. Those projections were based on estimates of liabilities generated by CBO's individual income tax model for 2009 through 2020. The projection of receipts for fiscal year 2010 (the "current year," according to the terminology in this paper) reflected payments for liabilities that the model projected for both calendar year 2009 and calendar year 2010 (see the bottom of Figure 1). The projection for the next year, fiscal year 2011 (the "budget year," in this paper's lexicon), reflected payments for liabilities that the model projected for calendar years 2010 and 2011.

The tax return data used in CBO's individual income tax model lag several years behind the year in which the projections are made, for multiple reasons. Although most people file their returns for one year by April of the next year, many taxpayers receive an extension for filing until October. It then takes the Internal Revenue Service (IRS) several months to compile the information; typically, details from the tax returns, in the form of a sample of returns that the IRS publishes as the Statistics of Income (SOI) Individual Income Tax File for that tax year, are available the following summer. CBO must then analyze and incorporate those data into its forecasting model.

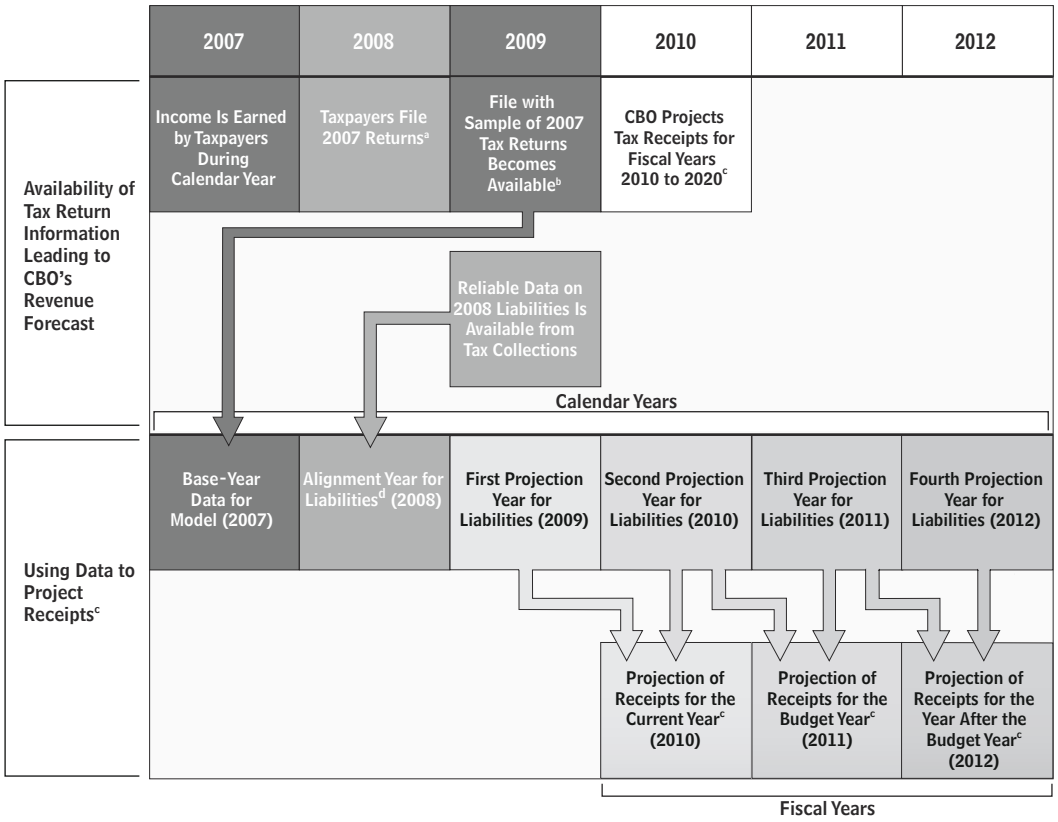
For example, for the projections released in January 2010, the base year of data for CBO's tax model was 2007, the latest year for which detailed tax return data were available. (Tax returns for that year were filed in 2008, and the detailed data became available in 2009.) CBO used those data to estimate what the tax returns in the 2007 SOI file (roughly 335,000 records) would look like for liability years 2008 through 2020 by adjusting the returns on the basis of actual and projected growth in the population and in income. Although tax returns for 2008 were not yet available at the time of the January 2010 forecast, CBO had an estimate of tax liabilities for 2008 (based on actual tax collections in fiscal years 2008 and 2009) that it could use to adjust the results from the model and improve its projections for subsequent years.

Forecasting Tax Receipts

CBO's individual income tax model relies on a microsimulation approach, using a sample of tax returns that reflects the diversity of households in the U.S. economy. The model is designed to capture certain characteristics of the population (for example, the distribution of tax filers by age) and how those characteristics interact with

Figure 1.

Timeline for Information Used in CBO's Revenue Projections Released in January 2010



Source: Congressional Budget Office.

Note: CBO's January 2010 revenue forecast included projections of receipts and liabilities through 2020. (The years 2013 to 2020 are not shown above.)

a. The initial deadline for filing taxes for a year is April 15 of the subsequent year, but taxpayers can automatically receive a six-month filing extension.

b. That is, the Internal Revenue Service's Statistics of Income Individual Income Tax File.

c. Receipts are based on liabilities from the current year and the previous year because some payments in the current year are for liabilities incurred in the previous year.

d. The alignment year is a year for which total tax receipts are known but detailed tax statistics are not yet available. The actual results for that year can be used to create an adjustment for the rest of the projection horizon by comparing the results from CBO's model for that year with actual individual income tax liabilities and then adjusting the model's projections for the remaining years of the horizon by that first-year difference.

provisions of the tax code.[7] Different tax provisions apply to different types of income and families. Because both the tax code and the characteristics of the population are changing over time, the microsimulation approach allows the model to reflect interactions between changes in the law and in the filing population's characteristics that affect revenues.

Once the sample of tax returns for a particular year has become available, CBO "ages" the sample from the base year to subsequent years by using actual data on the growth of the population and of income for the initial years and its projections of such growth for later years. CBO estimates changes in the different types of income reported on tax returns (such as earnings, interest, and dividends) on the basis of two factors: the actual history of comparable measures of those types of income available in the national income and product accounts (NIPAs) and CBO's projections of those measures for future years.[8] The NIPA measures of income do not exactly match the corresponding tax return measures because of differences in what each covers. In addition, certain kinds of income reported on tax returns, such as capital gains realizations and retirement income, have no analog in the NIPAs, and CBO must estimate those components of income separately. CBO also incorporates population growth into the model, using detailed estimates developed by the Social Security Administration, which are disaggregated by age, sex, and marital status.

CBO calculates the tax liability of each taxpayer in the sample by applying those estimates of changes in the distribution of income and the population as well as the effects of scheduled changes in the tax code. To generate total individual income tax liability for the entire population, CBO applies the weights from the sample—that is, the adjustments made to reflect the number of returns that each individual return in the sample represents. CBO then converts the estimate of total individual income tax liability from the model into projections of tax receipts, which measure how and when those liabilities will be collected in each fiscal year—specifically, how much will be collected through withholding, estimated payments, and the final payments taxpayers make when they file their returns.

The last step in forecasting receipts is to use the latest information on tax collections, which is available a year before the IRS's detailed tax return sample is released, to adjust the results of the model's projections of liabilities. For example, for CBO's January 2010 revenue forecast, the latest detailed tax return data covered the 2007 tax lia-

7. CBO supplements the data from tax returns with information from the Census Bureau's Current Population Survey to account for families that do not file tax returns. For a more detailed discussion of CBO's methods for projecting receipts from individual income taxes, see Congressional Budget Office, *Description of CBO's Models and Methods for Projecting Federal Revenues* (May 2001).

8. The NIPAs, which are maintained by the Department of Commerce's Bureau of Economic Analysis, track the level and composition of historical U.S. gross domestic product, the prices of its components, and the way in which the costs of production are distributed as income.

bility year. However, CBO used information about tax collections to estimate liabilities for the 2008 liability year, calculating overall taxes owed by adding together all of the tax payments made for that year: those for taxes withheld during calendar year 2008; the estimated tax payments made in 2008 (in April, June, and September) and in January 2009; and the final payments made and refunds issued for returns filed in 2009.[9] Once CBO had developed that estimate, it compared the projection of liabilities for 2008 from the model with its estimate of liabilities based on payments to determine whether and how to adjust the model's projections for other years of the forecast horizon.[10]

Why the Results of CBO's Model Differ from Actual Outcomes

When fuller information from tax returns becomes available, the discrepancies between actual tax receipts and CBO's projections can be traced, in part, to misestimates of the inputs for CBO's projection model. (For example, the actual rate of growth of incomes might have been different from what CBO projected.)[11] Other potential sources of errors include factors that affect liabilities but that the individual income tax model does not properly reflect. In particular:

- The most recent NIPA historical measures of wages and other types of income may contain inaccuracies that will be revised later, when better data become available.
- The relationship between categories of income in the NIPAs and the corresponding categories on tax returns may have changed. For example, wages recorded on tax returns may not grow at the same rate as do wages measured in the NIPAs.
- The income distribution may have changed in unexpected ways, causing misestimates of the average tax rates that taxpayers face.[12]
- The effects of recent legislative changes may be reflected incorrectly.

---

9. At this point in the forecasting process, CBO also has available preliminary data from 2009 tax returns for selected components of income, deductions, and tax liabilities, but those data generally do not supply much additional information beyond that available from the tax collections.

10. CBO's revenue projections also incorporated information from recent tax collections beyond that used to derive liabilities for 2008. For example, the revenue projections that CBO released in 2010 incorporated data about tax collections in 2009, which provided partial information about tax liabilities for that year. However, assessing the value of using that additional information from recent collections to improve CBO's projections was beyond the scope of this analysis.

11. Twice a year, in January and in August, CBO publishes a macroeconomic forecast (including projections for such variables as GDP, components of national income, interest rates, and inflation) that it uses to project federal revenues and spending.

12. The average tax rate is the ratio, for all taxpayers, of taxes paid to income.

- Other assumptions used in CBO's model—for example, about deductions, credits, and other components that are critical for determining tax liabilities but that are not part of CBO's macroeconomic projections—may be incorrect.

In certain cases, the divergence of CBO's revenue projections from actual outcomes might reflect some temporary factor. For example, income from the exercising of stock options was substantially higher in 1999 and 2000 than in previous years. That activity boosted total income and the share of income attributable to high-income filers—but only for those two years (the increase did not continue in the future). In other cases, an effect might persist. For instance, if CBO's estimate of the "cost" of a new tax credit was too low for the first year—because more people claimed the credit than CBO had expected—more people would probably claim it in future years than CBO had initially projected.

Until specific information on the sources of a divergence becomes available from tax returns, CBO has only limited data from which to determine whether the likely causes of a misestimate will prevail throughout the forecast period and whether the agency should adjust its projections accordingly. In those circumstances, CBO could follow one of two approaches:

- Assume that the misestimate resulted entirely from temporary phenomena and make no adjustments to the model's projections for the forecast period, or

- Assume that some or all of the discrepancy will continue into the forecast period and carry forward as an adjustment all or a portion of the amount by which the projection was too high or too low.

CBO generally follows that second approach, which is examined in the remainder of this paper, along with other methods for reflecting historical tax data in the agency's projections.

Measuring Forecasting Errors Using CBO's Simplified Tax Model

In analyzing the effectiveness of changes in its projection methods, CBO used a simplified tax model (STM) to more easily examine patterns of errors among different forecasts. The STM generates forecasts similar to those from the microsimulation model that CBO uses for its official forecasts. The differences between the forecasts generated by the STM and actual liabilities, expressed as a percentage of actual liabilities, were used to measure the errors in CBO's revenue forecasting. (For more information about CBO's simplified tax model, see Appendix A.)

For this analysis, CBO used the STM to project individual income tax liabilities for as many as eight years beyond each of the base years 1984 through 2005 (for which detailed tax return data are available). CBO then compared those projections with

actual liabilities to measure differences between the two for various projection years. As an example, in 1984, individual income tax liabilities totaled $309 billion. Starting with that amount, CBO used the most recent of its macroeconomic forecasts that would have been available—that is, the one issued in January 1987—and combined that forecast with information on income, tax rates, and tax rate elasticities derived from the tax return data to estimate total liabilities for future years.[13]

For example, CBO projected tax liabilities for 1985, 1988, and 1992 (for the purposes of this paper, "base year + 1," "base year + 4," and "base year + 8") of $334 billion, $444 billion, and $492 billion, respectively (see Figure 2). The projections diverged from the actual results for those years, overestimating liabilities by $3 billion for 1985, underestimating them by $27 billion for 1988, and overestimating them by $69 billion for 1992. Because the absolute size of such differences would tend to grow as the value of income tax liabilities grew, using that measure would make it difficult to compare the discrepancies over time. As a result, CBO "normalized" the errors by measuring them as a percentage of actual liabilities. For the previous example, actual liabilities differed from CBO's projections for 1985, 1988, and 1992 by -1.0 percent, 6.1 percent, and -13.9 percent, respectively (see Table 1).

Because actual data on income tax liabilities are available for the year after the base year, CBO can make certain adjustments to its estimates in that year. Thus, the year after the base year (base year + 1) is termed the "alignment year" in this analysis, the second year after the base year (base year + 2) is termed the first projection year, the third year after the base year is termed the second projection year, and so forth.

The years referenced here are liability years in the model, not the fiscal years of CBO's revenue projections. In the previous example, the first fiscal year covered in the January 1987 projections was what was then the current fiscal year—1987; the official revenue projection for that year was based on the (in this case, microsimulation) model's estimates of liabilities for 1986 and 1987. The projection for fiscal year 1988, the budget year, was based on the model's results for 1987 and 1988.

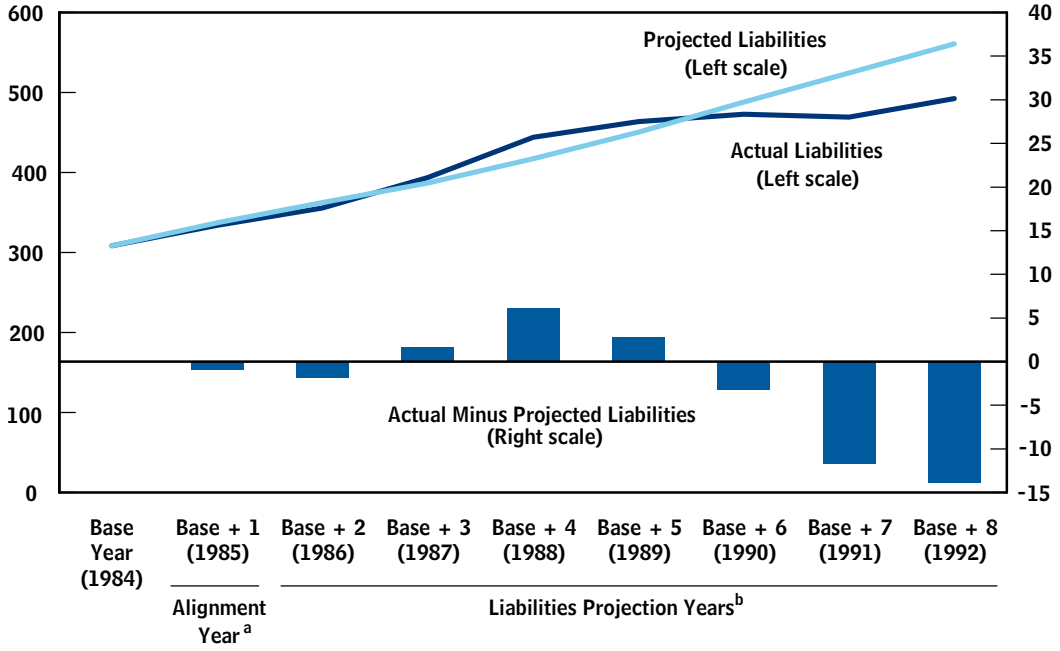
Estimating the Potential Improvement in the Accuracy of CBO's Revenue Forecasting

The accuracy of a forecast is the degree to which its projected values are narrowly dispersed around the actual outcomes. The root mean square error (RMSE) is a metric that can be used with different forecasting methodologies to compare their effects and to suggest the best method—at least over time—of using information from the error

13. An elasticity is a measure of the relationship between two variables—specifically, the ratio of the percentage change in one variable to the percentage change in another.

Figure 2.

Errors in Projecting Total Individual Income Tax Liabilities, as a Percentage of Actual Liabilities, Using Tax Return Data for 1984



Source: Congressional Budget Office.

Note: The figure is based on results from CBO's simplified tax model, which uses economywide totals and a small number of inputs to capture the essential features of CBO's more detailed microsimulation method.

a. The alignment year is a year for which total tax receipts are known but detailed tax statistics are not yet available. The actual results for that year can be used to create an adjustment for the rest of the projection horizon by comparing the results from CBO's model for that year with actual individual income tax liabilities and then adjusting the model's projections for the remaining years of the forecast by that first-year difference.

b. Projection years in the model representing the years in which tax liabilities are incurred, regardless of when payment is made.

in the projection for the first year of a forecast period to improve the forecast's accuracy in subsequent years. The RMSE is calculated by first squaring each of the errors in the projections for a forecast period and then taking the square root of the arithmetic average of the squared errors. (For example, if the errors in three projections were 2 percent, -3 percent, and 4 percent, the RMSE would equal 3.1 percent, the square root of the average of the three squared errors [(4 + 9 + 16)/3 = 9.7)]. Such a metric more clearly reflects the accuracy of the forecast than does the mean error because in the RMSE, overestimates and underestimates do not offset one another. The RMSE

Table 1.

Errors in Projecting Liabilities, as a Percentage of Actual Liabilities

	Liability Year for Which the Projection Was Made[b]							
Base Year [a]	Base Year + 1	Base Year + 2	Base Year + 3	Base Year + 4	Base Year + 5	Base Year + 6	Base Year + 7	Base Year + 8
1984	-1.0	-1.9	1.7	6.1	2.8	-3.2	-11.7	-13.9
1985	-1.8	0.6	6.1	3.8	-1.5	-9.0	-10.3	-14.3
1986	1.3	5.1	1.1	-3.9	-11.4	-13.1	-18.0	-18.7
1987	4.3	-2.5	-7.0	-15.1	-16.8	-21.6	-22.3	-19.6
1988	-2.4	-4.8	-9.8	-12.1	-18.0	-19.4	-17.2	-12.3
1989	-2.4	-3.4	-2.2	-7.9	-8.8	-6.4	-1.9	3.3
1990	-0.1	3.0	*	-0.6	1.7	6.8	12.9	18.8
1991	-0.3	-2.9	-4.0	-1.8	3.3	9.1	14.5	19.6
1992	-1.9	-3.7	-2.0	4.0	9.9	15.1	20.0	25.0
1993	1.1	1.4	6.8	12.9	18.2	23.0	27.9	17.3
1994	0.7	2.9	7.5	13.2	18.6	24.0	13.0	4.0
1995	0.8	2.0	8.2	14.4	20.8	10.2	1.9	-0.6
1996	1.8	4.2	10.1	16.8	7.0	-0.7	-2.5	5.8
1997	1.4	1.3	7.5	-3.9	-12.7	-14.4	-4.8	3.2
1998	1.3	2.2	-10.4	-21.1	-24.1	-14.1	-5.7	
1999	0.7	-8.9	-15.8	-20.4	-11.6	-3.6		
2000	-3.1	-6.0	-9.3	-2.7	2.9			
2001	-2.5	-2.8	4.0	8.8				
2002	-1.6	4.2	8.8					
2003	3.6	7.6						
2004	3.5							

Source: Congressional Budget Office.

Notes: The table is based on results from CBO's simplified tax model, which uses economywide totals and a small number of inputs to capture the essential features of CBO's more detailed microsimulation method.

Errors are actual liabilities minus projected liabilities, adjusted to remove the effects of legislation enacted after the base year and represented as a percentage of actual liabilities.

Missing values in the table represent years for which actual liabilities are not yet available or for which actual liabilities have not been calculated because of difficulties in making adjustments for enacted legislation.

* = between -0.05 percent and 0.05 percent.

a. Data on actual liabilities were taken from the samples of tax returns in the Internal Revenue Service's Statistics of Income Individual Income Tax Files published for these years.

b. The liability year is the calendar year in which the tax liability is incurred, regardless of when payment is made.

not only indicates the size of the overall discrepancy without regard to its direction, but it also gives greater weight to larger errors than does the mean error.

In its analysis, CBO first established a base case against which to assess the effectiveness of alternative forecasting methodologies. Using the simplified model, CBO identified the RMSEs for each liabilities projection year, making no adjustment to account for the error in the alignment year. The RMSE in the alignment year in that base case was 2.1 percent of the estimate of liabilities from the model; the RMSE was 4.1 percent in the first projection year (see Table B-1 on page 26). The error grew steadily over the first few years before leveling off at 14.7 percent of liabilities in the final (seventh) projection year.

CBO then measured RMSEs under three general approaches: carrying forward the alignment-year error into the estimates for subsequent years, using multiple years of tax return data from the IRS (rather than only the most recent year of available Statistics of Income data), and combining those two methods. Approaches that reduced RMSEs would be judged as increasing the accuracy of CBO's projections.

Approach 1: Adjusting the Simplified Tax Model's Results for the Alignment-Year Error

Despite the lack of detailed tax return data available beyond the base year, CBO has a good estimate of actual total income tax liabilities for the alignment year based on data from overall tax collections for that year. One way to use that information to improve projections of individual income tax receipts would be to adjust the model's projections for later years by the percentage error in the model's results for the alignment year. For example, the revenue projections that CBO produced in early 1987 were based on the IRS's Statistics of Income file for 1984. CBO recreated that revenue forecast for this analysis, using the STM and the macroeconomic forecast CBO produced in 1987, and found that actual tax liabilities in 1985 were 1 percent less than the projection from the STM.

In January 1987, when CBO was constructing its revenue forecast, analysts would have known that the model had overestimated liabilities by 1 percent in 1985 (the alignment year), but they would not have had detailed data on the actual liabilities for 1985 nor any data on actual liabilities for 1986. With those data now available, it can be seen that the STM was also overestimating liabilities for the first projection year (1986); the error for that year was 1.9 percent. CBO could have reduced that error by adjusting the model's result for 1986 by the amount of the 1985 error, which would have led to an error for 1986 of just less than 1 percent and improved the accuracy of the forecast in that instance.

One simple method for evaluating the effectiveness of adjusting future-year forecasts by the alignment-year error is to consider the number of years in which doing so would improve the forecasting of liabilities. Among the 20 forecasts that CBO considered, making the alignment-year adjustment for each of the seven years in a forecast's

projection window reduced the error in forecasting tax liabilities for the first projection year in 15 of those 20 forecasts. That is, it reduced the error in forecasting liabilities for the calendar year immediately preceding the January in which the forecast was being prepared—in the previous example, 1986. Carrying that same adjustment into the second projection year (1987 in the previous example) reduced the forecasting error for that year in 11 of 19 forecasts. However, carrying forward the alignment-year adjustment into years beyond the second year in the liabilities projection window reduced the error in fewer than half of those cases.

For all of the forecasts examined in this analysis, CBO used changes in the RMSE to measure the effect of making the full alignment-year adjustment to estimates for all years in the projection window. CBO found that:

- Using the full alignment-year adjustment for the first year of the projection window *reduced* the RMSE in that year by 12.7 percent (0.5 percentage points; see Figure 3 and Table B-1 on page 26);

- Using the entire alignment-year adjustment for the second year of the projection window *increased* the RMSE by 1.6 percent (0.1 percentage point); and

- Using the alignment-year adjustment for subsequent years in the projection window *increased* the RMSE in those years by about 3 percent (0.4 percentage points).

In sum, adjusting the model's results by the percentage error in the alignment year would have improved CBO's forecast of tax liabilities only for the first year of the projection period. A variant of that approach would be to adjust the model's results by a percentage that was smaller or larger than the alignment-year error. CBO's analysis of such adjustments revealed the following:

- Adjusting future-year results by only 50 percent of the alignment-year error *reduced* the RMSE in the first year of the projection window by 9.4 percent (0.4 percentage points), had no impact in the second year, and *increased* the RMSE by about 1 percent (0.2 percentage points) in subsequent years.

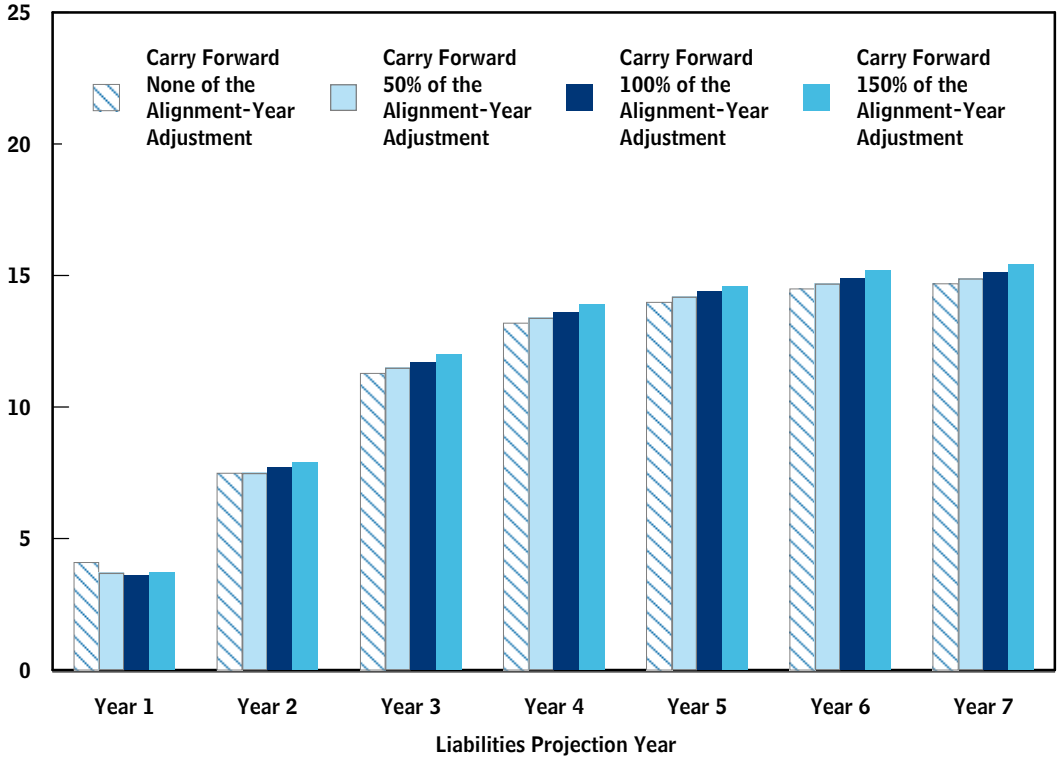
- Adjusting future-year results by 150 percent of the alignment-year error had an effect in the first year similar to that of adjusting the results by 50 percent of the error: The RMSE for the first projection year was *reduced* by 9.3 percent (0.4 percentage points). Using that approach for years beyond the first year of the projection window *increased* errors in those years by amounts ranging from 4.5 percent (0.6 percentage points) to 6.0 percent (0.7 percentage points).

Thus, CBO's analysis indicates that the best way to improve the accuracy of CBO's projections by adjusting the model's results for the alignment-year error is to use an adjustment equal to that error—but only for liabilities in the first year of the projection period.

Figure 3.

Effects of Carrying the Alignment-Year Adjustment into All Liabilities Projection Years

(Root mean square error)[a]



Source: Congressional Budget Office.

Notes: The figure is based on results from CBO's simplified tax model using one (base) year of tax return data. The simplified model uses economywide totals and a small number of inputs to capture the essential features of CBO's more detailed microsimulation method.

The first year after the base year is referred to as the alignment year; for that year, total tax receipts are known but detailed tax statistics are not yet available. The actual results for that year can be used to create an adjustment for the rest of the projection horizon by comparing the results from CBO's model for that year with actual individual income tax liabilities and then adjusting the model's projections for the remaining years of the forecast by that first-year difference.

a. The root mean square error is calculated by first squaring each of the errors in CBO's projections for a particular forecast year (for example, for the base year + 2, or the first liabilities projection year, as described in the text) and then taking the square root of the arithmetic average of the squared errors. In general, the smaller the error, the more accurate the projection will be.

Intuitively, it is reasonable to expect that carrying forward the alignment-year adjustment into the first year of the projection window would generally, although not always, improve the accuracy of CBO's revenue forecast. In a number of forecast periods, many of the sources of error in CBO's revenue projections (such as changes in the distribution of income or in the relationship between income as measured in the NIPAs and income as reported on tax returns) would persist for at least another year.

According to CBO's analysis, however, continuing the alignment-year adjustment (or some portion of it) beyond the first projection year does not generally improve the forecast's accuracy. One reason may be that CBO's projections, like those of most other forecasters, have been most accurate in periods in which the economy and tax liabilities were growing at a steady pace and least accurate when the projections were made at a turning point in the business cycle—that is, at the beginning or end of a recession or economic expansion. Thus, carrying forward the alignment-year adjustment at the beginning of a recession might worsen the forecast's accuracy, not because the source of the initial error was purely random but because the error was the result of an unexpected drop in income as the economy entered a downturn.

Approach 2: Basing Projections on Tax Return Data for Multiple Years

Errors in CBO's individual income tax projections could arise from temporary fluctuations in factors—such as the difference between wages and salaries reported on tax returns and wages and salaries reported in the NIPAs—that can vary from year to year. Using tax return data from years before the base year in addition to the base-year data might smooth those variations and improve the forecast's accuracy. For example, instead of projecting income tax receipts for 2010 solely on the basis of the tax returns filed in 2007, CBO could produce separate estimates—using the same macroeconomic forecast and tax return data from 2006 and 2007 (or even more years)—and then average the results.

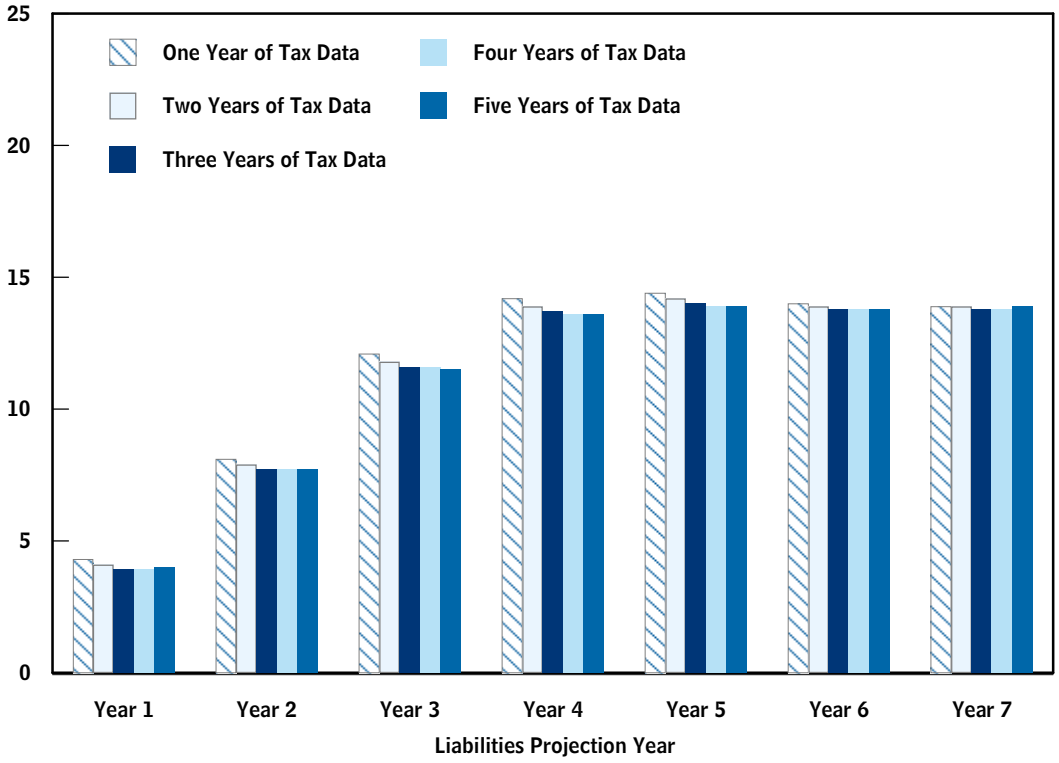
CBO therefore analyzed the impact of using multiple years of tax return data, weighted equally, instead of relying solely on data from the most recent year. It found that using tax return data from multiple previous years minimized the impact of minor annual fluctuations in those data and thereby improved the accuracy of its projections. Using four years of tax return data yielded the best overall improvement, reducing the forecasting error in all seven years of the projection period. That approach reduced the RMSE in the first projection year by 9.3 percent (0.4 percentage points) and the RMSE in the second year by 4.9 percent (0.4 percentage points; see Figure 4 and Table B-2 on page 28).[14] The reduction in the RMSE continued to

14. The results in this section differ slightly from the results in the previous section on the impact of the alignment-year adjustment because the use of multiple years of data means that fewer forecasts can be analyzed. Thus, for example, the RMSE for the first liabilities projection year shown in Figure 3 and Table B-1 is 4.1 percent and was constructed from 20 forecasts. The RMSE for the same projection year shown in Figure 4 and Table B-2 (the case using one year of tax return data) is 4.3 percent and was constructed from 16 forecasts.

Figure 4.

Effects of Using Multiple Years of Tax Return Data

(Root mean square error)[a]



Source: Congressional Budget Office.

Notes: The figure is based on results from CBO's simplified tax model, which uses economywide totals and a small number of inputs to capture the essential features of CBO's more detailed microsimulation method.

The tax return data come from the Internal Revenue Service's yearly Statistics of Income Individual Income Tax Files, which provide detailed information from a sample of tax returns.

a. The root mean square error is calculated by first squaring each of the errors in CBO's projections for a particular forecast year (for example, for the base year + 2, or the first liabilities projection year, as described in the text) and then taking the square root of the arithmetic average of the squared errors. In general, the smaller the error, the more accurate the projection will be.

diminish in subsequent years, dropping to 0.3 percent (less than 0.1 percentage point) by the seventh year in the projection window.

CBO found less improvement when it averaged multiple projections that were based on more or fewer than four years of data:

- Moving from a single year of base data to an average of two years of data reduced the RMSE in all projection years except the seventh.

- Compared with using projections based on two years of data, using the average of projections based on three years of data further reduced the RMSE in all projection years—but the reduction from using four years of data was greater.

- Using five years of data did not yield a noticeable reduction in the RMSE in the first projection year compared with the reduction from averaging data over four years, and it marginally increased the error in a number of the later projection years (see Figure 4 and Table B-2 on page 28).

Approach 3: Combining the Methods to Maximize Accuracy

Carrying forward the adjustment for the alignment year or averaging projections based on using several years of tax return information for the base year would each reduce the forecasting error in one or more years of the liabilities projection window. The final approach that CBO examined in this analysis was a combination of the two adjustments, to determine whether that strategy would yield a larger reduction in the forecasting error than the reduction achieved by either method alone. To test that approach, CBO analyzed several different combinations of adjusting for the error in the alignment year and using the average of multiple projections produced from different base years.

Of the options that CBO examined, the combination that was most effective in reducing the RMSE in the estimates of liabilities produced by the STM was to average projections based on the four most recent years of tax return data, carry about 50 percent of the alignment-year adjustment into the first year of the projection window, and make no change to the estimates for subsequent years.[15] That method reduced the RMSE for the first year of the projection window by 17.7 percent (0.7 percentage points; see Figure 5 and Table B-3 on page 30). The method reduced the RMSE for the second year by 5.5 percent (0.4 percentage points) and continued to reduce the error through the seventh year—by 2.1 percent for that year (0.3 percentage points).

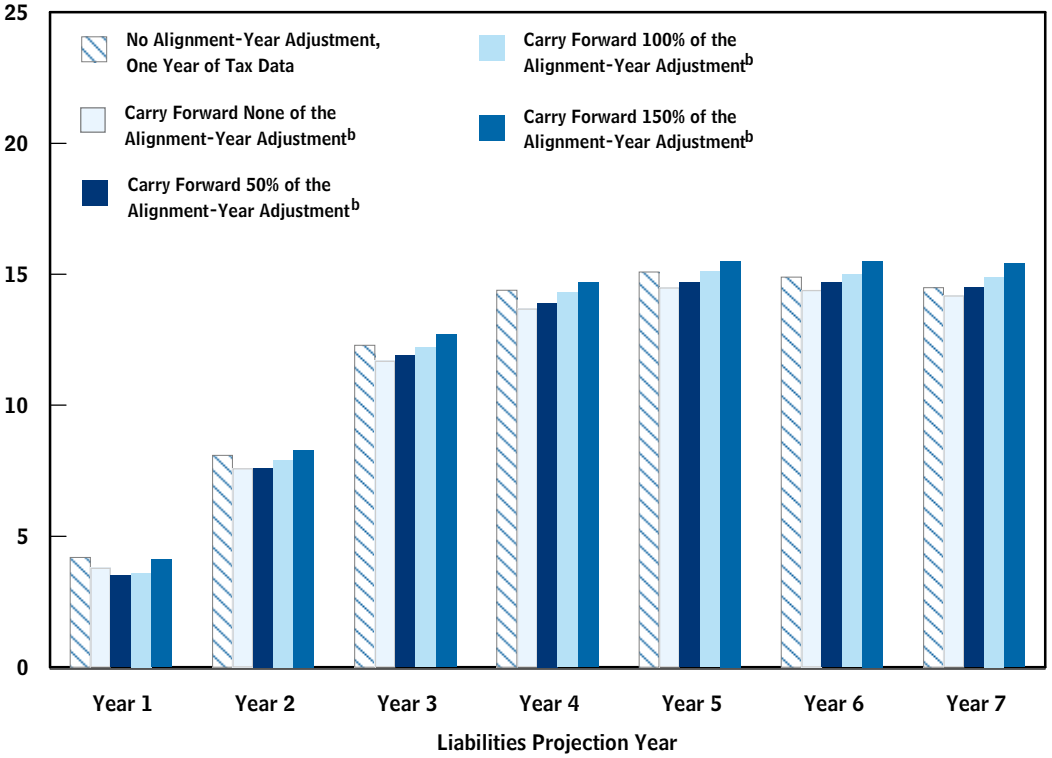
Carrying forward 100 percent or 150 percent of the alignment-year adjustment, in combination with averaging projections based on the four most recent years of tax data, reduced the RMSE in estimates of liabilities for the first year of the projection window but by less than did carrying forward only half of the adjustment. In addition, carrying forward all or more than all of the adjustment did not reduce the RMSE uniformly after the first projection year: In combination with averaging projections based on four years of tax data, carrying forward 150 percent of the alignment-year adjustment *increased* the RMSE for all years beyond the first

15. A more detailed analysis for the first projection year, which checked each integer between 50 percent and 100 percent, showed that carrying forward 62 percent of the alignment-year adjustment led to the smallest RMSE (3.47 percent). However, that error did not differ significantly from the RMSE generated by carrying forward 50 percent of the alignment-year error to the first projection year (3.48 percent).

Figure 5.

Effects of Carrying the Alignment-Year Adjustment into All Liabilities Projection Years and Using Four Years of Tax Return Data

(Root mean square error)[a]



Source: Congressional Budget Office.

Notes: The figure is based on results from CBO's simplified tax model, which uses economywide totals and a small number of inputs to capture the essential features of CBO's more detailed microsimulation method.

The tax return data come from the Internal Revenue Service's yearly Statistics of Income Individual Income Tax Files, which provide detailed information from a sample of individual income tax returns.

The first year after the base year is referred to as the alignment year; for that year, total tax receipts are known but detailed tax statistics are not yet available. The actual results for that year can be used to create an adjustment for the rest of the projection horizon by comparing the results from CBO's model for that year with actual individual income tax liabilities and then adjusting the model's projections for the remaining years of the forecast by that first-year difference.

a. The root mean square error is calculated by first squaring each of the errors in CBO's projections for a particular forecast year (for example, for the base year + 2, or the first liabilities projection year, as described in the text) and then taking the square root of the arithmetic average of the squared errors. In general, the smaller the error, the more accurate the projection will be.

b. In these cases, CBO used the four most recent years of tax return data available.

projection year, and carrying forward 100 percent of the adjustment *increased* the RMSE for the final two years of the projection window.

The findings presented here have limitations, however, in large part because they are based on results from the STM over a relatively small number of years. As a result, they may reflect the events that occurred during that period but not be representative of the range of outcomes likely to occur in the future. For example, the particularly severe recession of 2007 to 2009 was outside the period covered by the analysis, and the errors in CBO's revenue projections during the recession were much larger than those during the period that this analysis included. Nevertheless, the analysis period included several events—such as the 1990–1991 and 2001 recessions and the economic boom of the late 1990s—that might well be representative of the kinds of ups and downs most likely to lead to large forecasting errors in the future.

Appendix A: CBO's Simplified Tax Model

The simplified tax model (STM) that the Congressional Budget Office (CBO) developed for this analysis uses economywide totals and a small number of inputs to reflect the essential features of the microsimulation method that CBO uses for its official revenue forecasts.

Key Features

The STM comprises two modules, one for projecting liabilities from realizations of capital gains and another for projecting liabilities from all other sources of income. In the STM, estimates of tax liabilities other than from capital gains realizations are a function of several key inputs (see Table A-1); those inputs include the following:

- *Income.* Tax liabilities in the STM vary with the key components of income—wages, interest, dividends, and proprietors' income (capital gains are handled separately). CBO derived the totals for those components for the base year from the tax return data for that year; it derived those totals for subsequent years by increasing the base-year amounts by the estimated rate of growth of taxable personal income.[1]

- *Average Tax Rate for the Base Year.* The average tax rate is the ratio of taxes paid to income for all taxpayers in the base year. As with income, the calculation of the average tax rate excludes capital gains realizations and the taxes collected on those gains.

- *Tax Elasticities with Respect to Real Income Growth and Inflation.* In the STM, projected changes in the average tax rate from the rate in the base year are calculated on the basis of estimated changes in real (inflation-adjusted) income growth and inflation. The major parameters of the income tax—such as the brackets, the personal exemptions, and the standard deduction—are indexed for inflation to prevent average tax rates from rising when income grows solely as a result of a general rise in prices. However, if income grows faster than the rate of inflation, more

1. CBO's analysis measured income as the sum of the components of taxable personal income in the national income and product accounts—wages and salaries, proprietors' income, interest, and dividends—which were weighted to reflect their shares of taxable income as reported on tax returns.

Table A-1.

Key Inputs for CBO's Simplified Tax Model

Base Year[a]	Adjusted Gross Income[b] (Billions of dollars)	Average Tax Rate[c] (Percent)	Tax Elasticity with Respect to[d]	
			Real Growth per Capita[e]	Inflation
1984	2,081	13.5	1.71	1.05
1985	2,234	13.4	1.72	1.06
1986	2,345	13.4	1.70	1.04
1987	2,625	12.8	1.67	1.09
1988	2,920	12.8	1.56	1.07
1989	3,102	12.8	1.56	1.03
1990	3,283	12.8	1.55	1.00
1991	3,353	12.6	1.57	1.00
1992	3,502	12.8	1.55	1.05
1993	3,571	13.1	1.57	1.03
1994	3,755	13.3	1.56	1.03
1995	4,009	13.6	1.56	1.04
1996	4,275	13.8	1.54	1.01
1997	4,605	14.2	1.53	1.05
1998	4,961	14.1	1.54	1.06
1999	5,303	14.4	1.52	1.07
2000	5,721	14.9	1.50	1.08
2001	5,821	13.5	1.57	1.11
2002	5,765	13.0	1.61	1.17
2003	5,884	11.6	1.63	1.22
2004	6,292	12.1	1.59	1.22
2005	6,737	12.4	1.56	1.23

Source: Congressional Budget Office.

Note: CBO's simplified tax model uses economywide totals and a small number of inputs to capture the essential features of CBO's more detailed microsimulation method.

a. The table is based on detailed data from the samples of tax returns in the Internal Revenue Service's Statistics of Income Individual Income Tax Files published for these years.

b. Adjusted gross income—in this case, excluding capital gains—is calculated by subtracting deductions and other adjustments from a person's gross income.

c. Calculated as taxes on income divided by income, both excluding capital gains and losses.

d. An elasticity is a measure of the relationship between two variables—specifically, the ratio of the percentage change in one variable to the percentage change in the other.

e. Growth in inflation-adjusted income per person.

income will be taxed at higher rates, which will cause the average tax rate to rise over time. Similarly, if income grows more slowly than the rate of inflation, less income will be taxed at higher rates, which will cause the average tax rate to fall over time. In addition, not all parameters of the tax system are indexed for inflation; therefore, a general uptick in prices can also cause average tax rates to rise.[2]

The relationship between the percentage change in the average tax rate and the percentage change in other variables, such as real income per capita and the level of prices, is summarized in terms of elasticities—that is, the ratio of those changes. CBO simulated the impact of increases in real and nominal income for each base year using the tax return data that underlie its more detailed microsimulation model to determine how increases in real income and the price level affect tax liabilities. It then used those year-specific elasticities, which change over time as a result of alterations in the tax code and in the income distribution, as inputs for the STM.

All of the above inputs are combined to produce a forecast of tax liabilities other than those from capital gains realizations. The latter are excluded because liabilities resulting from capital gains realizations are determined by a different schedule of tax rates than that for most other forms of income. Consequently, in this analysis, CBO used a separate module to forecast such liabilities under the assumption that the amount of capital gains is known for the two years following the base year. For the first few years of each forecast horizon during the 1984–2005 period, CBO estimated capital gains realizations through a regression analysis based primarily on changes in the amount of businesses' fixed investment (for example, in structures and software) and the recent performance of the stock market relative to potential GDP, as well as changes in the rate at which gains are taxed.[3] For the latter years of a forecast horizon, CBO projected that capital gains realizations (and hence the resulting tax liabilities) would revert to a level consistent with their long-term relationship to GDP, which varies with the tax rates in effect.

Comparison with CBO's Microsimulation Model

In addition to its use of aggregated data, the STM differs in other respects from the microsimulation model used in CBO's official forecasting process. For example, the usual forecasting method uses detailed projections of the population, disaggregated by age, sex, and marital status, to account for the growth of various demographic groups represented in the base year of tax return data. In contrast, the STM implicitly incorporates the assumption that the population growth rate is the same for all such groups (and thus all returns). Although the resulting projections from the STM will

2. For example, the thresholds for taxing Social Security benefits are not indexed for inflation. When income rises with inflation, a greater share of benefits is subject to tax, and the tax rate on total income (including benefits) rises as well. The parameters of the alternative minimum tax also are not indexed for inflation, which results in a similar effect.

3. Potential GDP refers to the level of GDP that corresponds to a high level of resource (labor and capital) use and that neither adds to nor subtracts from inflationary pressures. It is never observed directly and must thus be estimated.

differ from the estimates produced using the microsimulation model, the difference introduced by the simplifying assumption of one growth rate is small, especially over the forecast horizons that CBO analyzes here.

Another difference between the two models is the way they treat the growth of certain types of taxable income, such as pension income and distributions from individual retirement accounts. The full microsimulation model projects the growth of each type of income separately, whereas the STM simply uses the rate of growth of overall taxable personal income (which includes such components as wages, interest, and dividends), excluding capital gains. The methodology used to project capital gains closely approximates the methodology that CBO ordinarily uses for such estimates.

Unlike projections from CBO's microsimulation model, the projections from the STM do not reflect changes in the tax code. As a result, CBO adjusted the STM's results for the effects of legislation enacted between the base year and a projection year before comparing those estimates with actual tax liabilities.[4]

Like the projections generated by CBO's microsimulation model, those from the STM were based on the macroeconomic forecast that CBO produced when it made its revenue projections. For example, in evaluating the projections that were based on actual tax return data for liabilities in 1984, CBO used its macroeconomic estimates from January 1987. The STM projections do not incorporate currently available data for the macroeconomic variables. As a result, CBO's analysis takes into account the errors that result from inaccuracies in the underlying macroeconomic forecast.

Despite several differences between the two approaches, the STM produces errors similar to those produced using the microsimulation model, thus permitting analysts to assess the effects of using different methods for incorporating available information from the recent past.[5] The STM also provides a tool for analyzing the impact of some of the assumptions CBO uses in forecasting. By comparison with the microsimulation model, however, the STM can more easily be used for each base year without additional adjustments.

---

4. A significant amount of uncertainty surrounds the effects of tax legislation on tax liabilities. In its analysis, CBO relied on published projections of the effects of changes in the tax code at the time the various laws were enacted and did not adjust those estimates to reflect subsequent information about a law's actual effects. (The estimates also reflected the impact of legislation that was enacted before the base year but that had provisions whose effects were phased in over time.) However, those published estimates were on a fiscal year basis, and in some cases, the periods covered by the estimates did not correspond to CBO's forecasting periods. As a result, CBO shifted the estimates to a liability year basis and, when necessary, extended them by producing estimates for additional years.

5. CBO compared the results of the STM with its projections of revenues from the microsimulation model, which exclude the effects of legislative changes occurring after the base year of data. Over each forecast horizon that CBO examined, the correlation between the forecasting errors for projections of liabilities produced by the STM and the forecasting errors for projections produced using the microsimulation methodology was roughly 0.9, suggesting that the STM successfully replicated the results of that more complex methodology.

Appendix B:
Detailed Data from the Analysis

This appendix comprises Tables B-1, B-2, and B-3, which provide the underlying data for Figures 3, 4, and 5 in the main text.

Table B-1.

Effects on Root Mean Square Errors in CBO's Simplified Tax Model from Carrying Forward the Alignment-Year Adjustment

Liabilities Projection Horizon	Number of Revenue Forecasts Providing Data[a]	RMSE with No Alignment-Year Adjustment[b]	With Alignment-Year Adjustment[b]		
			RMSE	Percentage-Point Change	Percentage Change
			Carry Forward 50 Percent of the Adjustment		
Alignment Year	21	2.1	2.1	n.a.	n.a.
Projection Year 1	20	4.1	3.7	-0.4	-9.4
Projection Year 2	19	7.5	7.5	*	*
Projection Year 3	18	11.3	11.5	0.2	1.3
Projection Year 4	17	13.2	13.4	0.2	1.2
Projection Year 5	16	14.0	14.2	0.2	1.1
Projection Year 6	15	14.5	14.7	0.2	1.3
Projection Year 7	14	14.7	14.9	0.2	1.2
			Carry Forward 100 Percent of the Adjustment		
Alignment Year	21	2.1	2.1	n.a.	n.a.
Projection Year 1	20	4.1	3.6	-0.5	-12.7
Projection Year 2	19	7.5	7.7	0.1	1.6
Projection Year 3	18	11.3	11.7	0.4	3.4
Projection Year 4	17	13.2	13.6	0.4	2.9
Projection Year 5	16	14.0	14.4	0.4	2.7
Projection Year 6	15	14.5	14.9	0.4	2.9
Projection Year 7	14	14.7	15.1	0.4	2.9

Continued

Effects on Root Mean Square Errors in CBO's Simplified Tax Model from Carrying Forward the Alignment-Year Adjustment

Liabilities Projection Horizon	Number of Revenue Forecasts Providing Data[a]	RMSE with No Alignment-Year Adjustment[b]	With Alignment-Year Adjustment[b]		
			RMSE	Percentage-Point Change	Percentage Change
			Carry Forward 150 Percent of the Adjustment		
Alignment Year	21	2.1	2.1	n.a.	n.a.
Projection Year 1	20	4.1	3.7	-0.4	-9.3
Projection Year 2	19	7.5	7.9	0.4	4.8
Projection Year 3	18	11.3	12.0	0.7	6.0
Projection Year 4	17	13.2	13.9	0.7	5.0
Projection Year 5	16	14.0	14.6	0.6	4.5
Projection Year 6	15	14.5	15.2	0.7	4.9
Projection Year 7	14	14.7	15.4	0.7	4.8

Source: Congressional Budget Office.

Notes: CBO's simplified tax model uses economywide totals and a small number of inputs to capture the essential features of CBO's more detailed microsimulation method.

The first year after the base year of actual data from the Internal Revenue Service is referred to as the alignment year. For that year, total tax receipts are known, but detailed tax statistics are not yet available. The actual results for that year can be used to create an adjustment for the rest of the projection horizon by comparing the results from CBO's model for that year with actual individual income tax liabilities and then adjusting the model's projections for the remaining years of the forecast by that first-year difference.

The root mean square error is calculated by first squaring each of the errors in CBO's projections for a particular forecast year (for example, for the base year + 2, or the first liabilities projection year, as described in the text) and then taking the square root of the arithmetic average of the squared errors. In general, the smaller the error, the more accurate the projection will be.

RSME = root mean square error; n.a. = not applicable; * = between -0.05 percent and 0.05 percent.

a. CBO's 10-year revenue forecasts for 1984 through 2005.

b. Results are based on one year of actual tax return data.

Table B-2.

Effects on Root Mean Square Errors in CBO's Simplified Tax Model from Using Multiple Years of Tax Return Data

Liabilities Projection Horizon	Number of Revenue Forecasts Providing Data[a]	RMSE with One Year of Tax Return Data	With Multiple Years of Tax Return Data		
			RMSE	Percentage-Point Change	Percentage Change
			Use Two Years of Tax Data		
Projection Year 1	16	4.3	4.1	-0.2	-5.1
Projection Year 2	15	8.1	7.9	-0.2	-2.5
Projection Year 3	14	12.1	11.8	-0.3	-2.5
Projection Year 4	13	14.2	13.9	-0.3	-2.1
Projection Year 5	12	14.4	14.2	-0.2	-1.5
Projection Year 6	11	14.0	13.9	-0.1	-0.8
Projection Year 7	10	13.9	13.9	*	*
			Use Three Years of Tax Data		
Projection Year 1	16	4.3	3.9	-0.4	-8.7
Projection Year 2	15	8.1	7.7	-0.4	-4.6
Projection Year 3	14	12.1	11.6	-0.5	-4.1
Projection Year 4	13	14.2	13.7	-0.5	-3.5
Projection Year 5	12	14.4	14.0	-0.4	-2.7
Projection Year 6	11	14.0	13.8	-0.2	-1.4
Projection Year 7	10	13.9	13.8	*	-0.3
			Use Four Years of Tax Data		
Projection Year 1	16	4.3	3.9	-0.4	-9.3
Projection Year 2	15	8.1	7.7	-0.4	-4.9
Projection Year 3	14	12.1	11.6	-0.6	-4.8
Projection Year 4	13	14.2	13.6	-0.6	-4.2
Projection Year 5	12	14.4	13.9	-0.4	-3.0
Projection Year 6	11	14.0	13.8	-0.2	-1.5
Projection Year 7	10	13.9	13.8	*	-0.3

Continued

Effects on Root Mean Square Errors in CBO's Simplified Tax Model from Using Multiple Years of Tax Return Data

Liabilities Projection Horizon	Number of Revenue Forecasts Providing Data[a]	RMSE with One Year of Tax Return Data	With Multiple Years of Tax Return Data		
			RMSE	Percentage-Point Change	Percentage Change
			Use Five Years of Tax Data		
Projection Year 1	16	4.3	4.0	-0.4	-8.2
Projection Year 2	15	8.1	7.7	-0.4	-4.7
Projection Year 3	14	12.1	11.5	-0.6	-5.2
Projection Year 4	13	14.2	13.6	-0.6	-4.5
Projection Year 5	12	14.4	13.9	-0.5	-3.3
Projection Year 6	11	14.0	13.8	-0.2	-1.7
Projection Year 7	10	13.9	13.9	*	-0.1

Source: Congressional Budget Office.

Notes: CBO's simplified tax model uses economywide totals and a small number of inputs to capture the essential features of CBO's more detailed microsimulation method.

The root mean square error is calculated by first squaring each of the errors in CBO's projections for a particular forecast year (for example, for the base year + 2, or the first liabilities projection year, as described in the text) and then taking the square root of the arithmetic average of the squared errors. In general, the smaller the error, the more accurate the projection will be.

RSME = root mean square error; * = between -0.05 percent and 0.05 percent.

a. CBO's 10-year revenue forecasts for 1984 through 2005.

Table B-3.

Effects on Root Mean Square Errors in CBO's Simplified Tax Model from Carrying the Alignment-Year Adjustment into All Liabilities Projection Years and Using Four Years of Tax Return Data

Liabilities Projection Horizon	Number of Revenue Forecasts Providing Data[a]	RMSE with No Alignment-Year Adjustment[b]	With Alignment-Year Adjustment[c]		
			RMSE	Percentage-Point Change	Percentage Change
			Carry Forward None of the Adjustment		
Alignment Year	18	2.6	2.6	n.a.	n.a.
Projection Year 1	17	4.2	3.8	-0.4	-9.8
Projection Year 2	**16**	**8.1**	**7.6**	**-0.4**	**-5.5**
Projection Year 3	**15**	**12.3**	**11.7**	**-0.7**	**-5.4**
Projection Year 4	**14**	**14.4**	**13.7**	**-0.7**	**-4.9**
Projection Year 5	**13**	**15.1**	**14.5**	**-0.6**	**-4.2**
Projection Year 6	**12**	**14.9**	**14.4**	**-0.4**	**-3.0**
Projection Year 7	**11**	**14.5**	**14.2**	**-0.3**	**-2.1**
			Carry Forward 50 Percent of the Adjustment		
Alignment Year	18	2.6	2.6	n.a.	n.a.
Projection Year 1	**17**	**4.2**	**3.5**	**-0.7**	**-17.7**
Projection Year 2	16	8.1	7.6	-0.4	-5.4
Projection Year 3	15	12.3	11.9	-0.5	-3.7
Projection Year 4	14	14.4	13.9	-0.5	-3.4
Projection Year 5	13	15.1	14.7	-0.4	-2.7
Projection Year 6	12	14.9	14.7	-0.2	-1.3
Projection Year 7	11	14.5	14.5	*	-0.2
			Carry Forward 100 Percent of the Adjustment		
Alignment Year	18	2.6	2.6	n.a.	n.a.
Projection Year 1	17	4.2	3.6	-0.6	-14.8
Projection Year 2	16	8.1	7.9	-0.2	-2.5
Projection Year 3	15	12.3	12.2	-0.1	-0.8
Projection Year 4	14	14.4	14.3	-0.2	-1.1
Projection Year 5	13	15.1	15.1	-0.1	-0.3
Projection Year 6	12	14.9	15.0	0.2	1.2
Projection Year 7	11	14.5	14.9	0.3	2.4

Continued

Effects on Root Mean Square Errors in CBO's Simplified Tax Model from Carrying the Alignment-Year Adjustment into All Liabilities Projection Years and Using Four Years of Tax Return Data

Liabilities Projection Horizon	Number of Revenue Forecasts Providing Data[a]	RMSE with No Alignment-Year Adjustment[b]	With Alignment-Year Adjustment[c]		
			RMSE	Percentage-Point Change	Percentage Change
			Carry Forward 150 Percent of the Adjustment		
Alignment Year	18	2.6	2.6	n.a.	n.a.
Projection Year 1	17	4.2	4.1	-0.1	-2.2
Projection Year 2	16	8.1	8.3	0.2	3.0
Projection Year 3	15	12.3	12.7	0.4	3.2
Projection Year 4	14	14.4	14.7	0.3	2.0
Projection Year 5	13	15.1	15.5	0.4	2.8
Projection Year 6	12	14.9	15.5	0.7	4.5
Projection Year 7	11	14.5	15.4	0.8	5.7

Source: Congressional Budget Office.

Notes: CBO's simplified tax model uses economywide totals and a small number of inputs to capture the essential features of CBO's more detailed microsimulation method.

The first year after the base year of actual data from the Internal Revenue Service is referred to as the alignment year. For that year, total tax receipts are known, but detailed tax statistics are not yet available. The actual results for that year can be used to create an adjustment for the rest of the projection horizon by comparing the results from CBO's model for that year with actual individual income tax liabilities and then adjusting the model's projections for the remaining years of the forecast by that first-year difference.

The root mean square error is calculated by first squaring each of the errors in CBO's projections for a particular forecast year (for example, for the base year + 2, or the first liabilities projection year, as described in the text) and then taking the square root of the arithmetic average of the squared errors. In general, the smaller the error, the more accurate the projection will be.

Data in the table that appear in boldface type reflect the combination of methods that was most effective in reducing the RMSE in the estimates of liabilities produced by CBO's simplified tax model.

RSME = root mean square error; n.a. = not applicable; * = between -0.05 percent and 0.05 percent.

a. CBO's 10-year revenue forecasts for 1984 through 2005.

b. Results are based on one year of actual tax return data.

c. Results are based on the four most recent years of tax return data available.