Technical Paper Series
Congressional Budget Office
Washington, DC

USING TIME-SERIES MODELS TO PROJECT OUTPUT OVER THE MEDIUM TERM

Ufuk Demiroglu
Congressional Budget Office
Washington, DC
(Email: ufukd@cbo.gov)

Matthew Salomon
Joint Economic Committee
U.S. Senate
Washington, DC
(Email: Matt_Salomon@jec1.senate.gov)

September 2002

2002-1

Using Time-Series Models to Project Output Over the Medium Term

This paper examines how multivariate time-series models might be used to project output over the medium term—that is, over a 10-year span. Fairly simple time-series models are known to yield short-term forecasts comparable in accuracy with those of large-scale macroeconometric models. Could the information embodied in projections based on time-series models be informative to the Congressional Budget Office (CBO) in constructing its own medium-term projections of output?[1]

Based on our examination of several time-series models, we conclude that such models can help inform the medium-term projections in several respects. Because time-series models generally place fewer a priori restrictions on the dynamic properties of the projections than do CBO's models, an independent time-series representation might be a useful check on the efficacy of CBO's standard, more restrictive, modeling approach. Additionally, prospective forecast uncertainty is readily computable for each time-series model but not for CBO's standard projections. Finally, implicit within the dynamic structure of a time-series model is a transition path along which the economy moves toward its trend rate of growth. That transition path can be compared with the transition path in CBO's medium-term projections, and the latter might be reassessed if the time-series predictions so warrant. The models examined here allow ready calculation of measures of trend or potential output, facilitating such a comparison.

This paper is only an exploratory effort to understand whether and how such

---

1. CBO's current procedures for making projections over the medium term are described in Congressional Budget Office, *CBO's Method for Estimating Potential Output: An Update*, CBO Paper (August 2001).

models could be useful for CBO's purposes. Accordingly, the focus is limited to some basic time-series specifications, which are taken from articles that have been published in peer-reviewed journals.

The standard errors for the alternative projections of output are estimated to range from about 4 percent to 8 percent of CBO's winter 2000 baseline projection for 2010.[2] The uncertainty bands around the projection of real output imply a narrower range of uncertainty for the federal surplus than was recently estimated by CBO in its examination of errors in its surplus projections.[3] However, uncertainty about real output growth is only one source of uncertainty in the projection of the federal surplus, which is also affected by uncertainty in the levels of revenues and spending for given levels of real output.

CBO's Procedures for Projecting Output Over the Medium Term

In constructing its projections of output over the medium term, CBO assumes that actual output eventually approximates potential output. Accordingly, CBO's projection of actual output growth is built upon two measures: CBO's historical estimate and projection of potential output, as well as an assumption as to the rate at which and the degree to which the gap between actual and potential output narrows over time.

CBO's Measure of Potential Output. CBO's estimates and projections of potential

---

2. All the analyses in this paper make use of the data and CBO baseline available as of February 2000.

3. Congressional Budget Office, *Uncertainties in Projecting Budget Surpluses: A Discussion of Data and Methods*, CBO Report (February 2002).

output are based on a model that assumes that growth in potential output equals the weighted sum of the trend growth in labor and capital services plus the trend growth in total factor productivity (TFP).[4] To measure the trend component of potential output (and its constituent productive factor services and TFP) in isolation from cyclical influences, CBO relies on a variant of the well-known and widely used relationship known as Okun's law. According to that relationship, actual output exceeds its potential level when the rate of unemployment is below its "trend" level. Conversely, when the unemployment rate exceeds its trend level, output falls short of potential. The unemployment gap—the difference between the trend and actual rates of unemployment—is the pivotal indicator of the business cycle in models built upon Okun's law.

CBO uses as the trend level of unemployment a measure known as the nonaccelerating inflation rate of unemployment, or NAIRU. The NAIRU is a conceptual measure that builds upon the view that, over the short term, the unemployment rate varies inversely with changes in inflation. In that view, periods of high unemployment are associated with falling inflation; conversely, periods of low unemployment are associated with rising inflation. The NAIRU is defined as that rate of unemployment that is consistent with a constant rate of inflation.[5]

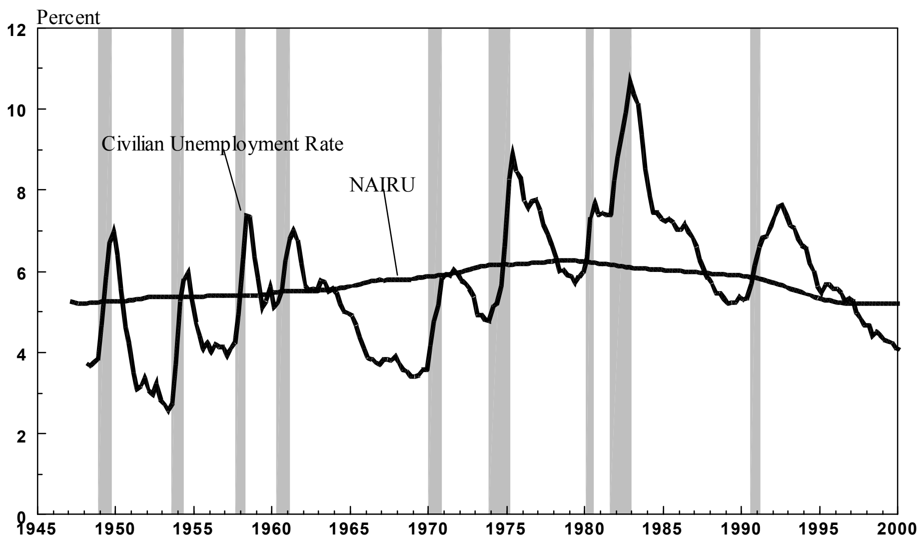
---

4. Total factor productivity (TFP) is defined as that part of nonfarm business output not explained by factor service flows. CBO assumes that the production function for nonfarm business output depends on the flows of labor and capital services. Measures of TFP depend critically on the precise functional form assumed for the production function as well as the many additional assumptions necessary to obtain an estimate of the flow of capital services. Changes in any of those assumptions would alter the estimate of TFP.

5. Whether or not there is a stable inverse relationship between the unemployment rate and changes in inflation—a relationship inspired by the work of A. W. Phillips—has always been a matter of controversy among economists. In recent years, the controversy has focused on the statistical precision with which the NAIRU is measured. Some economists have argued that estimates of the NAIRU have been too imprecise for the NAIRU to be useful to policymakers; see D. Staiger, J. H. Stock, and M. W. Watson, "The NAIRU, Unemployment, and Monetary Policy," *Journal of Economic Perspectives*, vol. 11, no. 1 (Winter 1997), pp. 33-49. Other studies have suggested that the imprecision of NAIRU estimates has been exaggerated; see Flint Brayton, John M. Roberts, and John C. Williams, *What's Happened to the Phillips Curve?* Finance and Economics Discussion Series Paper No. 1999-49 (Washington, D.C.: Federal Reserve

CBO estimates the NAIRU by first estimating a NAIRU for married males. CBO starts with married males because it has confirmed that demographic shifts, such as changes in labor force participation and household formation, have distorted

Figure 1. Civilian Unemployment Rate and CBO's Estimate of the NAIRU



SOURCES: Congressional Budget Office; Department of Labor, Bureau of Labor Statistics.

the relationship between the change in inflation and the overall civilian unemployment rate.[6] The unemployment rate for married males, however, is relatively free of such distorting demographic shifts. Accordingly, CBO initially estimates the NAIRU using the unemployment rate for married males. Then, CBO constructs its estimate of the NAIRU for the entire labor force by taking into account the demographic changes with respect to age, sex, and race. The resulting estimate

Board, September 1999). Ongoing independent analyses within CBO have yielded results that are consistent with the latter findings.

6. The procedures are presented in greater detail in Congressional Budget Office, *The Economic and Budget Outlook: An Update* (August 1994), pp. 59-63; and *CBO's Method for Estimating Potential Output: An Update* (August 2001).

of the overall NAIRU varies slowly over time because of variations in the composition of the labor force (see Figure 1).[7]

CBO uses the unemployment gap as a proxy for the business cycle in regression equations to estimate the potential levels of the labor input and TFP. Each regression equation breaks down the natural logarithm of the respective independent variable (labor input or TFP) into a deterministic time trend, a cyclical component proportional to the unemployment gap, and a random residual. Once the equations have been estimated over the historical period, CBO identifies the estimated deterministic trends as the so-called potential levels of labor input or TFP.

Those estimated potential levels of labor input and TFP are combined with CBO's estimates of the service flow of capital to yield an estimate of the level of potential output for the nonfarm business sector.[8] CBO assumes that the production function for the nonfarm business sector has a Cobb-Douglas form, with the elasticity of output with respect to changes in labor input equal to 0.7. Under those assumptions, a proportionate change in the level of potential output equals the proportionate change in potential TFP plus 70 percent of the proportionate change in labor input flows plus 30 percent of the proportionate change in the service flow from capital.

By construction, the proportionate gap between actual and potential output is

7. In recent years, CBO has lowered its estimate of the NAIRU below that implied by this procedure to accommodate what has become the consensus view—namely, that the NAIRU has fallen at least temporarily for reasons other than changes in the composition of the labor force with respect to age, sex, and race. CBO estimates the overall NAIRU to have been about 5.2 percent in recent years.

8. In its estimate of potential output, CBO assumes that the service flow from the aggregate stock of nonfarm business capital is always at its potential level. This assumption allows CBO's estimate of the output gap—though not the projected *level* of potential output—to be independent of the service flow of capital.

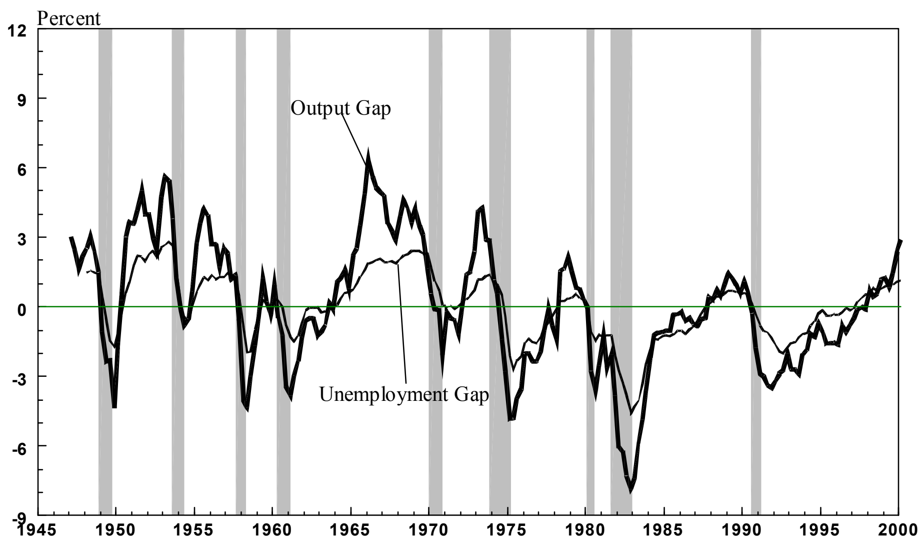
equal to the TFP gap plus 70 percent of the labor input gap. Because the TFP and labor gaps vary directly with the unemployment gap, the output gap does as well (see Figure 2).

Transition to the Medium Term. Given the levels of historical and projected potential output, CBO computes a projected path for actual nonfarm business output by "closing" the output gap at some point during the projection period. Once the gap has been completely closed, the projected level of output grows at the same rate as the potential level of output.

Both the degree to which the gap is closed and the pace at which it closes are determined through assumptions made by CBO. Those assumptions have varied throughout CBO's history. Until recently, CBO assumed that the gap would eventually reach and remain at its historical average of between 0.2 percentage points and 0.4 percentage points of output, after which output was projected to grow at the same rate as potential. Currently, however, the projection period is extended to 10 years, and CBO assumes that the gap goes to zero at least two years before the end of the projection period and stays there. Of course, the growth rate of actual output depends on the most recent historical estimate of the gap—if actual GDP exceeds potential GDP, the growth rate of actual output must be less than the projected growth rate of potential output for the gap to close.

Some Limitations of CBO's Procedures. CBO's approach to measuring and projecting potential output is a version of a widely used structural model that allows the influence of exogenous factors (such as demographic shifts) to be easily incorporated in estimates of potential output, directly via changes in the estimate of labor input and indirectly via changes in the NAIRU. However, the advantages of CBO's approach come at the cost of some limitations.

Figure 2. CBO's Estimate of the Output and Unemployment Gaps



SOURCES: Congressional Budget Office; Department of Labor, Bureau of Labor Statistics; Department of Commerce, Bureau of Economic Analysis.

NOTE: The output gap is measured as the difference between actual and potential GDP as a percentage of potential GDP. The unemployment gap is measured as the difference between the estimated NAIRU and the civilian unemployment rate.

First, the difference between the unemployment rate and the NAIRU may not be the best method of identifying cycles in output. CBO's standard approach to estimating potential GDP assumes the Phillips-Okun structure. Thus, it could fail if that structure turns out not to be a good description of reality.

Moreover, CBO's approach to estimating the output gap does not shed any light on how quickly or slowly the gap tends to return to its long-run level over the medium term. As a result, CBO has to make assumptions regarding the pace at which the gap closes over the projection horizon. Although those assumptions are informed by factors such as the most recent historical estimates of the gap and the trend reversion patterns exhibited in the past, the determination of the gap remains exogenous to the statistical models that are used to estimate and project potential output.

Finally, CBO's method depends on possibly imprecise estimates of the capital stock. Although the use of measures of the capital stock to estimate potential output certainly has a strong theoretical basis, such an approach comes at the cost of possible measurement error. Capital stocks must be estimated and then linked to CBO's projections of business investment, requiring additional steps that introduce the possibility of error into the process of measuring potential output for the nonfarm business sector. For instance, conventional measures of the capital stock appear to overstate short-run movements in the capital input.

Specification and Estimation of the Multivariate Statistical Models

In principle, multivariate time-series models offer a way around some of the limitations of CBO's standard procedures for projecting output over the medium term. This section describes the three statistical models chosen for this study, the estimation issues that arose in fitting the models to post-World War II data, and the techniques used to isolate trend components from cyclical components in each model's projections.

Three Models. The three alternatives examined here are variants of models that have appeared in leading economics journals over the past decade or so. Two of the models are bivariate vector autoregressions (VARs), and the other is a bivariate vector error-correction model (VECM). A VAR is a linear multivariate stochastic model in which each variable is affected by the past and contemporaneous values of all of the other variables in the model as well as its own past, in addition to random influences. VECMs are an extension of VARs in which long-run relationships among variables are explicitly built in.

- The first model is based on a seminal empirical study by Blanchard and Quah (BQ).[9] It is a dynamic version of the Okun relationship, which CBO's method also relies on. This model may be viewed as a VAR benchmark to CBO's current procedures.

- The second model is taken from a recent study by Gali.[10] Unlike the neoclassical growth-accounting framework that CBO's current methodology uses, this model combines TFP and the capital input into a labor productivity variable, circumventing the need to estimate the capital stock.

- The third model, which is based on work by Cochrane, hinges on a completely different economic theory, namely that the level of consumption is an indicator of expected future income.[11]

The BQ model estimates a bivariate VAR relationship for the growth rate of real

---

9. Olivier Blanchard and Danny Quah, "The Dynamic Effects of Aggregate Demand and Supply Disturbances," *American Economic Review*, vol. 79, no. 4 (September 1989), pp. 655-673.

10. Jordi Gali, "Technology, Employment, and the Business Cycle: Do Technology Shocks Explain Aggregate Fluctuations?" *American Economic Review*, vol. 89, no. 1 (March 1999), pp. 249-271.

11. John H. Cochrane, "Permanent and Transitory Components of GNP and Stock Prices," *Quarterly Journal of Economics*, vol. 109, no. 1 (February 1994), pp. 241-265.

GDP and the unemployment rate for married males.[12] As Blanchard and Quah originally pointed out in their 1989 study, this specification is a generalized variant of the basic Okun relationship, which allows, potentially, for a richer range of dynamics than does the contemporaneous Okun relationship.

Gali's bivariate VAR uses growth in labor productivity for the nonfarm business sector along with growth in labor hours for that sector.[13] This model also allows for the comovement of the labor input and labor productivity as indicated by the Okun relationship.

Cochrane's VECM imposes a cointegration relationship between the logarithms of consumption and real GDP. The model is motivated by the permanent income hypothesis, which postulates that households determine their consumption in relation to their expectations of average lifetime income. One implication of that hypothesis is that changes in income that are not accompanied by changes in consumption must be temporary. This suggests that a given change in income may be identified as temporary or permanent by determining the degree to which consumption also changes.

Specification Issues. There were two major specification issues with the models: how the data would be transformed to make them stationary and the number of lags to use on the right-hand side. The assumptions regarding stationarity and lag length

12. In their original study, Blanchard and Quah used real GNP along with the unemployment rate for males 25 years and older. At the expense of a shorter sample, we've chosen to use the unemployment rate for married males because it is the unemployment measure used in CBO's estimates of potential GDP. Additionally, we used BEA's chained-Fisher indexes of GDP, which were not available to Blanchard and Quah in 1989.

13. The sum of the two growth rates equals growth in real GDP for the nonfarm business sector.

bear directly on the ways in which the trend, or permanent, component of output is isolated. Although a trend-cycle decomposition is not necessary to obtain forecasts, some defensible measure of the trend is necessary to compare the implications of the alternative models for the estimated output gap. For the baseline specifications in this study, we adopted the stationarity transformations and lag lengths of the original studies.

For each of the alternative models, we estimated a measure of trend output using a procedure first developed by Beveridge and Nelson (BN).[14] Put simply, the BN decomposition identifies the trend component of output at any point by subtracting from output the part that the model estimates to be transient.

In the case of our estimate of the Blanchard-Quah model, we additionally computed another measure of trend output using the same long-run identifying restrictions as were introduced in the original BQ study. The basic intuition underlying those restrictions is that short-term or transitory shocks should have no permanent effects on output growth. (Implementation of those restrictions and other technical details on the decomposition methods are discussed in Appendix A.)

Finding a Stationary Series. The decomposition methods used here require output to be a difference-stationary time series. Determining whether this is truly the case, however, is a dicey matter in practice: numerous analysts over the past two decades have examined the question for real GDP and no decisive conclusion is at hand. Part

14. S. Beveridge and C. R. Nelson, "A New Approach to Decomposition of Economic Time Series Into Permanent and Transitory Components with Particular Attention to Measurement of the 'Business Cycle'," *Journal of Monetary Economics*, vol. 7, no. 2 (March 1981), pp. 151-174. A multivariate version of the Beveridge-Nelson decomposition is presented in G. Evans and L. Reichlin, "Information, Forecasts, and Measurement of the Business Cycle," *Journal of Monetary Economics*, vol. 33, no. 2 (April 1994), pp. 233-254.

of the reason for this uncertainty is that, using the available finite data on real GDP, it is effectively impossible to distinguish the hypothesis of a difference-stationary series from one that has an autoregressive root very close to unity. This observational equivalence may be of little consequence in short-run forecasting—indeed, it appears that these alternatives tend to produce comparable short-run forecasts—but the implications for the very long term are radically different.[15]

These difficulties are illustrated with the macroeconomic time series to be used in the study. A first criterion for stationarity is whether the sample autocorrelation function approaches zero sufficiently quickly for sufficiently long lags. Table 1 presents the sample autocorrelations for a number of time series, each of which is subjected to several alternative transformations.

Except for the unemployment rate for married males, the autocorrelations for the untransformed series show very little dampening—all of the autocorrelations for the logarithms or levels of the series appear to be significantly different from zero for the first 10 quarters. When the series are differenced, however, the sample autocorrelations for all of the series dampen very quickly—in most cases, they approach zero after a couple of quarters.

The autocorrelations also tend to dampen when their time trends are removed from the series. If, in addition to a break in the linear trend, a break is allowed in the trend rate of real GDP growth, the autocorrelations dampen even more quickly. But

15. See M. W. Watson, "Univariate Detrending with Stochastic Trends," *Journal of Monetary Economics*, vol. 18, no. 1 (July 1986), pp. 49-75.

Table 1. The Autocorrelations of Selected Macroeconomic Time Series Under Alternative Transformations

	Autocorrelations by Number of Quarters Lagged									
	1	2	3	4	5	6	7	8	9	10
Real GDP										
Level (natural logs)	0.98	0.97	0.95	0.94	0.93	0.91	0.90	0.88	0.87	0.85
Differenced	0.33	0.18	-0.01	-0.12	-0.18	-0.11	-0.10	-0.05	0.03	0.07
Deterministic Trend	0.96	0.90	0.84	0.77	0.71	0.66	0.61	0.57	0.53	0.49
Deterministic Trend with Break	0.93	0.82	0.68	0.55	0.43	0.34	0.26	0.20	0.15	0.09
Real GDP, Nonfarm Business										
Level (natural logs)	0.98	0.97	0.95	0.94	0.92	0.91	0.89	0.88	0.86	0.85
Differenced	0.24	0.16	-0.03	-0.11	-0.22	-0.11	-0.09	-0.09	0.08	0.07
Deterministic Trend	0.94	0.85	0.75	0.65	0.57	0.50	0.45	0.41	0.37	0.33
Deterministic Trend with Break	0.91	0.79	0.64	0.49	0.37	0.28	0.21	0.15	0.11	0.06
Unemployment Rate, Married Males										
Level (percent)	0.95	0.86	0.75	0.65	0.56	0.48	0.41	0.35	0.30	0.26
Differenced	0.57	0.19	-0.05	-0.21	-0.14	-0.05	-0.12	-0.15	-0.08	0.00
Deterministic Trend	0.95	0.85	0.74	0.63	0.53	0.45	0.38	0.31	0.26	0.22
Deterministic Trend with Break	0.95	0.85	0.74	0.63	0.53	0.45	0.38	0.32	0.27	0.22
Labor Productivity, Nonfarm Business										
Level (natural logs)	0.98	0.97	0.95	0.94	0.92	0.91	0.89	0.88	0.86	0.85
Differenced	-0.14	0.14	-0.06	0.01	-0.10	-0.00	-0.07	0.04	0.15	-0.01
Deterministic Trend	0.97	0.94	0.92	0.89	0.86	0.84	0.82	0.79	0.77	0.75
Deterministic Trend with Break	0.85	0.74	0.61	0.50	0.39	0.32	0.26	0.22	0.19	0.13
Labor Services, Nonfarm Business										
Level (natural logs)	0.99	0.97	0.95	0.94	0.92	0.91	0.89	0.88	0.86	0.84
Differenced	0.59	0.24	0.04	-0.15	-0.26	-0.20	-0.15	-0.18	-0.09	0.03
Deterministic Trend	0.95	0.86	0.75	0.63	0.53	0.44	0.37	0.33	0.30	0.28
Deterministic Trend with Break	0.93	0.79	0.62	0.44	0.28	0.15	0.06	-0.01	-0.05	-0.08
Consumption of Real Nondurables										
Level (natural logs)	0.99	0.97	0.96	0.95	0.93	0.92	0.90	0.89	0.88	0.86
Differenced	0.13	0.11	0.14	-0.02	-0.05	0.00	0.10	-0.15	0.02	0.01
Deterministic Trend	0.98	0.96	0.94	0.91	0.88	0.85	0.82	0.79	0.76	0.74
Deterministic Trend with Break	0.92	0.82	0.74	0.64	0.55	0.47	0.39	0.31	0.25	0.20

SOURCE: Authors' calculations using data from the Department of Commerce, Bureau of Economic Analysis; and the Department of Labor, Bureau of Labor Statistics.

the rate of dampening is still much slower than in the case of differenced output.

Running a battery of augmented Dickey-Fuller tests on the data suggests that, with the exception of the unemployment rate, most of the series can adequately be characterized by unit roots. However, those tests are not powerful enough to distinguish unit roots from mean-reversion, especially in the presence of breaks in trend.[16] Additional testing for unit roots after allowing for a break in trend in 1974 did result in mild rejections of the null hypothesis of a unit root in several cases. But those results are not completely reliable, either, because of size distortion problems that bias the results toward rejecting the null.[17] Those problems make it difficult to identify what the right form of stationarity transformations should be.

For the purposes of this study, we adopted as baseline specifications the same stationarity transformations as were used by the authors of the original studies. We assumed that all variables, except for the unemployment rate, were stationary in first differences of their natural logarithms (that is, growth rates). As a check on the robustness of the basic specification, however, we also examined alternative stationarity transformations.[18]

Choosing a Lag Length. The issue of lag length is critical. The efficacies of both of the trend-cycle decompositions used in this study (BN or BQ) rely on the quality of the empirical approximation to the long-run properties of the time-series model. These, in turn, are sensitive to the chosen lag length. Ideally, one would like to choose a lag length that is long enough to ensure the accuracy of the approximation but not so long as to chew up too many degrees of freedom.

Established criteria are available to determine what the lag length should be on

16. Pierre Perron, "The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis," *Econometrica*, vol. 57, no. 6 (November 1989), pp. 1361-1401.

17. N. Gregory Mankiw and Matthew D. Shapiro, "Do We Reject Too Often? Small Sample Properties of Tests of Rational Expectations Models," *Economics Letters*, vol. 20 (1986), pp. 139-145.

18. All computations of the output gaps in this report, however, reflect the basic specifications used in the respective published works.

the basis of those considerations. It turns out that the lag lengths of the original specifications of the three models exceeded those selected by the various information criteria. However, a recent study has shown that the lag orders favored by those criteria tend to be too low, resulting in inconsistent estimates of a model's long-run coefficients.[19] That inconsistency points to the difficulty associated with determining the appropriate lag length. On balance, we did not find a sufficiently strong case in favor of lag orders different from the ones used in the original studies.[20]

Analysis of the Model Forecasts

The models are estimated using quarterly postwar data through the end of 1999 (as published in February 2000). Those estimated models are then used to forecast real GDP through 2010. The comparisons are made with CBO's winter 2000 estimates and forecasts. The range of forecasts for real GDP is substantial, rising to nearly 8 percent of the midpoint forecast by 2010.

The analysis below breaks down the forecast differences into three components: differences in the estimated trend rate of growth, the most recent historical value of the output gap, and the speed of decay of the gaps through the projection period.[21]

19. See Alain DeSerres and Alain Guay, *Selection of the Truncation Lag in Structural VARs (or VECMs) with Long-Run Restrictions*, Working Paper No. 95-9 (Ottawa, Canada: Bank of Canada, October 1995).

20. A lag length of eight quarters is reported for the Blanchard-Quah and Gali VARs. Cochrane does not report a lag length in his published study; we assumed a lag length of two quarters for the basic specification.

21. When VARs are used for forecasting, it is not actually necessary to identify the gap or make a separate assumption about how quickly it closes. Estimating the model and simulating the reduced form regression are sufficient to obtain forecasts. Nevertheless, the models' estimates for the trend rate of growth and the present deviation from the trend are worth looking into because they provide a convenient framework with which to understand and compare forecasts of different models. They also provide a parallel to the CBO method and therefore facilitate an understanding of how a comparison can be made with the various components that form the CBO forecast. As discussed in Section II, CBO has historically used estimates of the rate of growth of potential output and the output gap to guide its projections of growth in real GDP over the medium term, by assuming that the output gap closes at a given rate over the projection period.

The differences in the estimates of the output gap's level at the end of the sample period account for about half of the differences in output forecasts for 2010. As the estimated gap widens, the part of current output taken to be cyclical and temporary increases and the predictions of future growth decrease. The models vary in their estimates of the output gap mainly because they rely on different economic principles and choose different economic variables. The other half of the differences in 2010 forecasts is due to differences in the estimated trend rate of growth of GDP, which, in turn, critically depend on whether or not a break in the trend in 1974 is assumed. The speed with which the gap closes affects the short-term forecast but has essentially no effect on 2010 forecasts.

Forecast Results. The Blanchard-Quah VAR projected real GDP to grow at an average annual rate of 2.9 percent between 2000 and 2010. The Gali VAR projected a 3.6 percent per year advance over the period, and the Cochrane VECM projected growth of 3.3 percent per year, on average. The range of forecasts is substantial, rising from just under 1 percent of the midpoint forecast in 2000 to nearly 8 percent of the midpoint forecast by 2010. In that year, the difference between the largest and smallest projected level of real GDP is a trillion chained-1996 dollars—those two forecasts would yield projections of federal revenues that differed by over $450 billion in 2010, more than CBO's current projection of the federal surplus in that year.

Accounting for Differences in the Forecasts. More than half of the 8.3 percent difference between the Gali and Blanchard-Quah forecasts stems from the difference in their estimates of the gap at the end of the sample period; the remainder arises from the difference in their estimates of the trend rate of growth (see Table 2).[22] The gap at the end of the sample period in the BQ VAR is 4.5 percent higher than in the Gali VAR. Because the gap disappears over time, that difference reduces the BQ VAR's projection of output growth by 4.5 percent between 1999 and 2010. The remaining difference between the Gali VAR and the BQ VAR in the 2010 forecasts is due to the faster (by 0.3 percent per year) trend growth rate estimated by the BQ VAR.[23]

Table 2. Reasons for Differences Among the Output Forecasts for 2010 (Percentage of 2010 GDP as estimated by the lower forecast)

Vector Autoregressions	2010 Forecast	Components of the Total Difference Due to	
		Estimates of the Output Gap in 1999:4	Estimates of the Trend Growth Rate
Gali VAR vs. BQ VAR	8.3	4.5	3.6
Cochrane VECM vs. BQ VAR	5.0	2.4	2.6
Gali VAR vs. Cochrane VECM	3.1	2.0	1.0

SOURCE: Authors' calculations using data from the Department of Commerce, Bureau of Economic Analysis; and the Department of Labor, Bureau of Labor Statistics.

The 2010 forecast difference between the BQ VAR and the Cochrane VECM is smaller. It is nearly equally accounted for by the difference in estimates of the output gap at the end of the sample period and estimates of the trend rate of growth. As for the forecast differences of the third pair (the Cochrane VECM and the Gali

22. This ignores the fact that the estimate of the output gap is not independent of the estimate of the trend rate of growth.

23. This accumulates to a difference of approximately 3.6 percent of the level of real GDP by 2010 (see Table 1). 1.036*1.045 = 1.0826, so the 4.5 and 3.6 percentage points fully account for the 8.3 percent difference. Note that all percentages are rounded to the nearest first decimal in the tables.

VAR), the forecast difference is small and arises mostly from the difference in their estimates of the output gap.

Trend Rate of Growth. The trend rates of growth range from 3.2 percent to 3.5 percent. That range is quite large considering the effects of such differences in average growth carried out over 10 years. The range reflects the different assumptions about a trend break in 1974. If a trend break is not assumed, as in the Gali and Cochrane models, the estimate of the trend rate of growth is approximately equal to the average growth rate of output in the estimation sample, which is 3.4 percent. If a trend break in 1974 is assumed, as in the Blanchard-Quah VAR, the average real output growth for the post-1973 sample (as opposed to the full sample) determines the estimate of the trend rate of growth. Because the growth rate of output in the years after 1973 is substantially lower (3.2 percent) than it was prior to 1974, the trend break assumption results in a lower estimate of the trend rate of growth. Of the three alternatives, the estimate of the trend component in the Blanchard-Quah VAR was closest to CBO's estimate, which is not surprising given the fact that CBO's approach also assumes trend breaks.

Output Gap. The Beveridge-Nelson decomposition was used to estimate the output gap for all three models. In the Blanchard-Quah VAR, the BQ decomposition was used as well; the estimate of the output gap produced under that decomposition was similar to the gap produced by the BN decomposition for that model. The estimated gap shows some striking differences across models for recent years. The three models estimate the gap at the end of the sample period (the last quarter of 1999) as follows:

BQ VAR:	4.2 percent[24]	(the largest boom since the 1960s)
Gali VAR:	-0.2 percent	(economy below potential)
Cochrane VECM:	1.8 percent	(a mild boom)

By contrast, CBO's estimate of the output gap for that quarter at the time was approximately 3 percent, which lies midway between the BQ VAR and the Cochrane VECM.

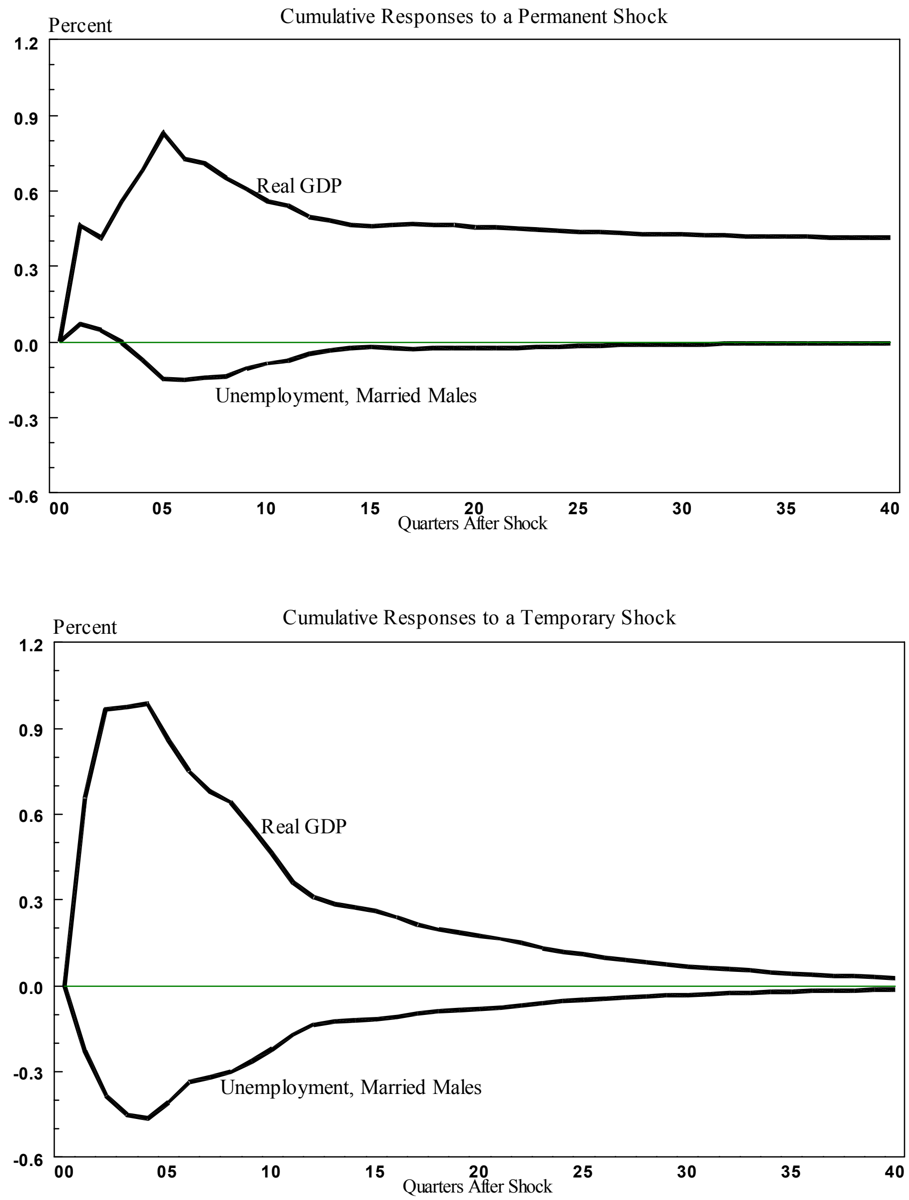
The differences in the gap estimates of the Blanchard-Quah VAR and the Cochrane VECM arise from the fact that they are based on very different economic principles. Because the BQ VAR perceives a concurrent decrease in unemployment and increase in output as a temporary shock, it interprets the activity in the last years of the sample period as a long and large cyclical boom.[25] The cyclical increase in the VECM gap model is not as large as that in the BQ VAR because the VECM model estimates potential output as approximately proportional to consumption. The increases in output in the boom of the late 1990s have at least partly been accompanied by increases in consumption; thus, the VECM model infers that part of the increase in output is due to an increase in potential output.

Not surprisingly, the BQ VAR's estimate of the output gap at the end of 1999 is closest to CBO's estimates. After 1960, the output gap implicit in the BQ VAR follows the cycle turning points of the National Bureau of Economic Research (NBER) fairly closely; the exception is the recession of 1970 (see Figure 4). In most

24. The 4.2 percent is the estimate obtained by using the BN decomposition. When long-run restrictions are used to identify the trend, the estimate becomes 3.8 percent.

25. This can be seen from the impulse response functions for this model (Figure 3, lower panel).

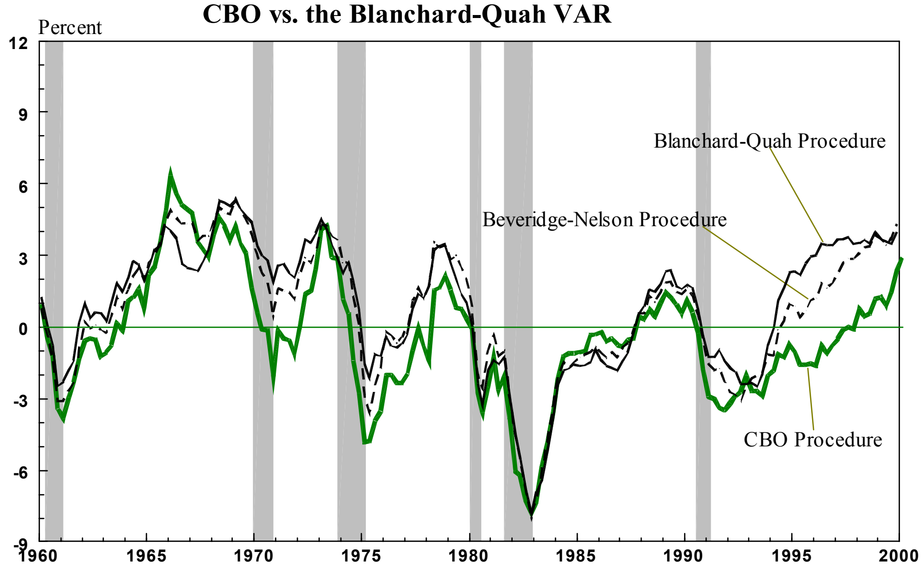
Figure 3. Cumulative Impulse Responses in the BQ VAR Under Blanchard-Quah Restrictions



SOURCE: Authors' calculations using data from the Department of Commerce, Bureau of Economic Analysis; and the Department of Labor, Bureau of Labor Statistics.

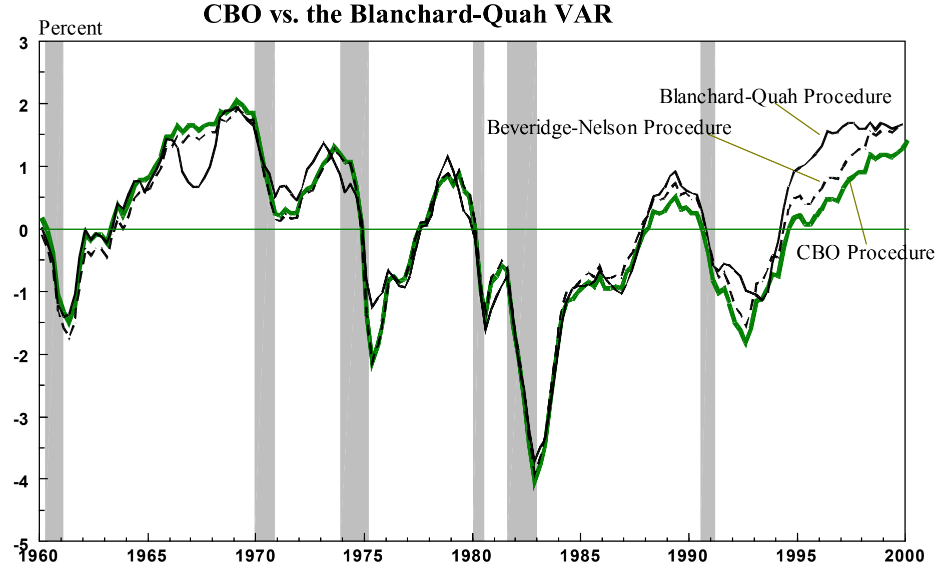
respects, until the early 1990s, the measure of the output gap from the BQ VAR mirrors CBO's fairly consistently. After 1994, however, the BQ VAR estimates the

Figure 4. Comparison of Output Gap Estimates: CBO vs. the Blanchard-Quah VAR



SOURCE: Authors' calculations using data from the Department of Commerce, Bureau of Economic Analysis; and the Department of Labor, Bureau of Labor Statistics.

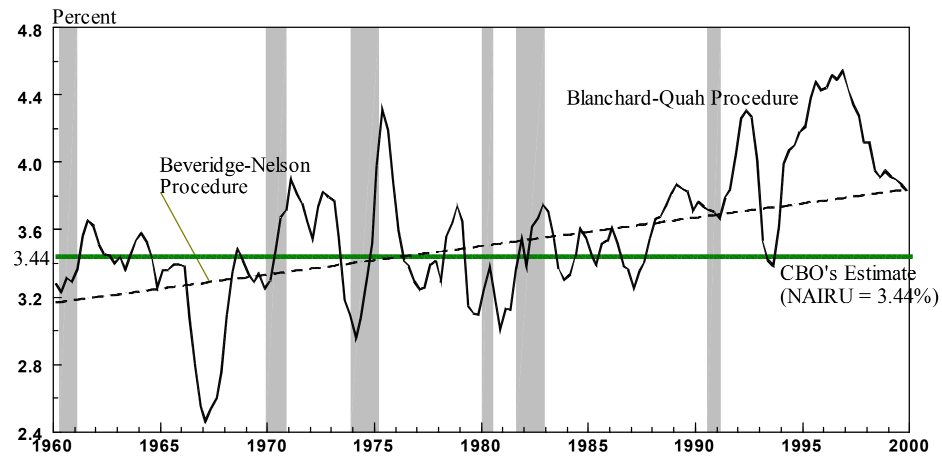
Figure 5. Comparison of Unemployment Gap Estimates: CBO vs. the Blanchard-Quah VAR



SOURCE: Authors' calculations using data from the Department of Commerce, Bureau of Economic Analysis; and the Department of Labor, Bureau of Labor Statistics.

NOTE: The married-male unemployment gap is measured as the difference between the estimated trend and the actual unemployment rate for married males.

Figure 6. Comparison of Estimates of the Trend Rate of Unemployment for Married Males: CBO vs. the Blanchard-Quah VAR



SOURCE: Authors' calculations using data from the Department of Commerce, Bureau of Economic Analysis; and the Department of Labor, Bureau of Labor Statistics.

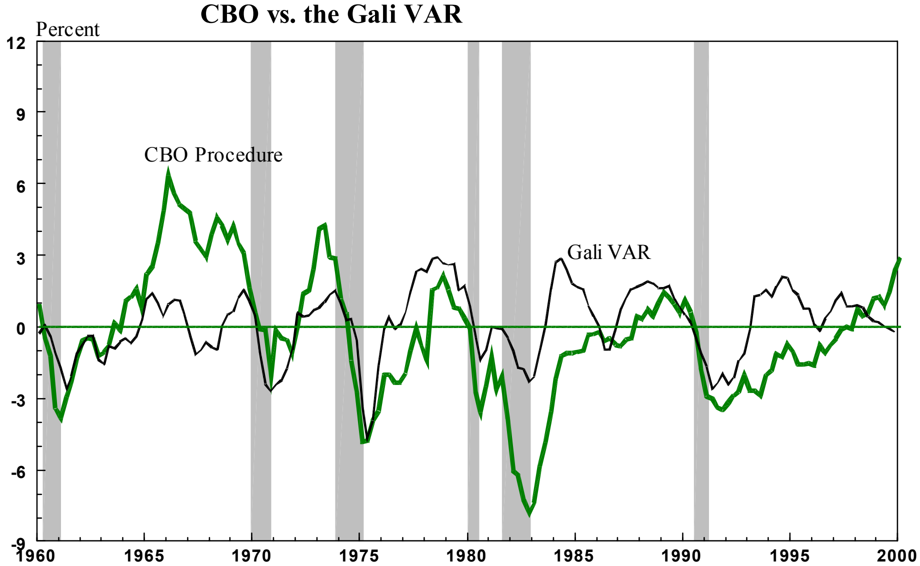
trend growth rate somewhat below the growth rate of potential estimated by CBO. As a result, the BQ VAR indicates a substantially larger gap than CBO does, especially between 1995 and 1998, although the gap narrows somewhat after that. A related observation is that the BQ VAR estimates a wider gap between the trend and actual rates of unemployment for married males (see Figure 5).

Part of the reason for the divergence of the output and unemployment gaps between the BQ VAR and the CBO estimates is that the VAR has built into it a rising trend of unemployment for married males, in contrast with CBO's assumption of a constant married-male NAIRU (see Figure 6).[26] Another reason is that CBO's

26. If the VAR is estimated without removing a linear trend from unemployment, the gap is estimated to be 0.8 percent smaller. Removing the trend, therefore, can account for a substantial part of the 1.2 percent difference between the gap estimates of CBO and Blanchard-Quah VAR for the end of the sample period. (In fact, if the long-run restrictions are used to identify the trend in BQ VAR, the difference is 0.8 percent to start with—CBO's estimate of 3.0 percent versus the BQ VAR's estimate of 3.8 percent—and in that case the treatment of the unemployment trend can explain all the difference.)

method, unlike the time-series methods, incorporates information on demographic shifts on top of the information contained in past trends.

Figure 7. Comparison of Output Gap Estimates: CBO vs. the Gali VAR



SOURCE: Authors' calculations using data from the Department of Commerce, Bureau of Economic Analysis; and the Department of Labor, Bureau of Labor Statistics.

NOTE: The output gap is measured as the difference between actual and potential GDP as a percentage of potential GDP.

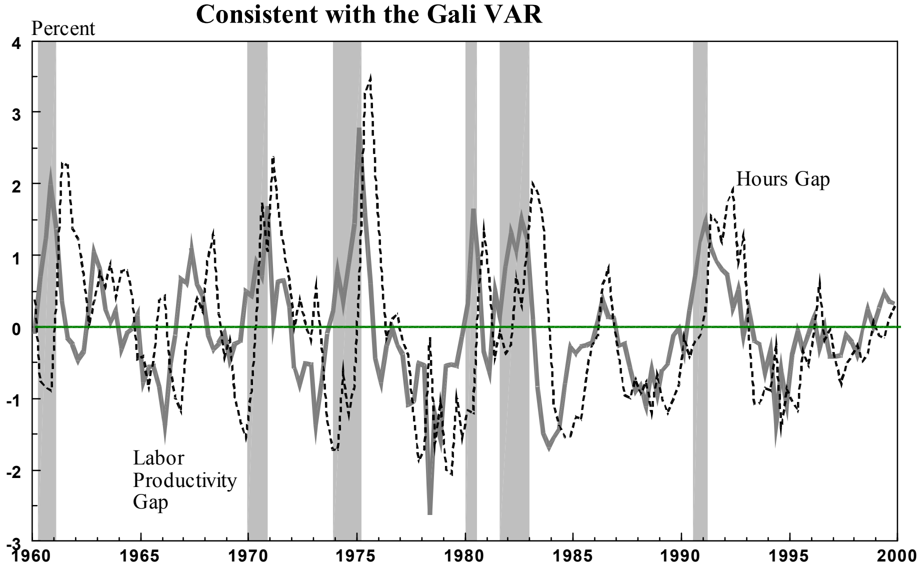
The Gali VAR produces a picture of the output gap that is very different from all of the other models (see Figure 7).[27] Although the estimated gap demonstrates some propensity to follow the cycle, it is unusual in other respects. The movements in this measure of the output gap indicate that the economy was operating below its potential at the end of 1999—an assessment with which most public and private

27. The measure of the output gap was estimated by using the Beveridge-Nelson decomposition to calculate the permanent component of the two variables of this VAR, which are the growth rates of labor productivity and hours in nonfarm business, and then adding together those two rates to estimate the growth rate of the permanent component of nonfarm business GDP. To compute the trend growth rate in overall GDP from the estimate in the nonfarm business sector, the ratio of the two was set to grow at the same rate as it does in CBO's projection. The underlying gaps for the two variables appear to move almost in lockstep throughout the historical period (see Figure 8).

forecasters would disagree. Moreover, the Gali VAR judges the 1990-1991 recession as more severe than the back-to-back recessions of 1980 and 1982.

The measure of the output gap based on the Cochrane VECM shows a more subdued cyclical tendency over history, though one that tracks the NBER's turning points fairly well (see Figure 9). This is not surprising given that the estimate of potential output tracks consumption closely in this model and consumption is cyclical. Consequently, the estimated potential output tends to be low when output is low and vice versa. In other words, this model's estimate of potential output is cyclical—not a desirable quality for a potential output estimate. The apparent lack

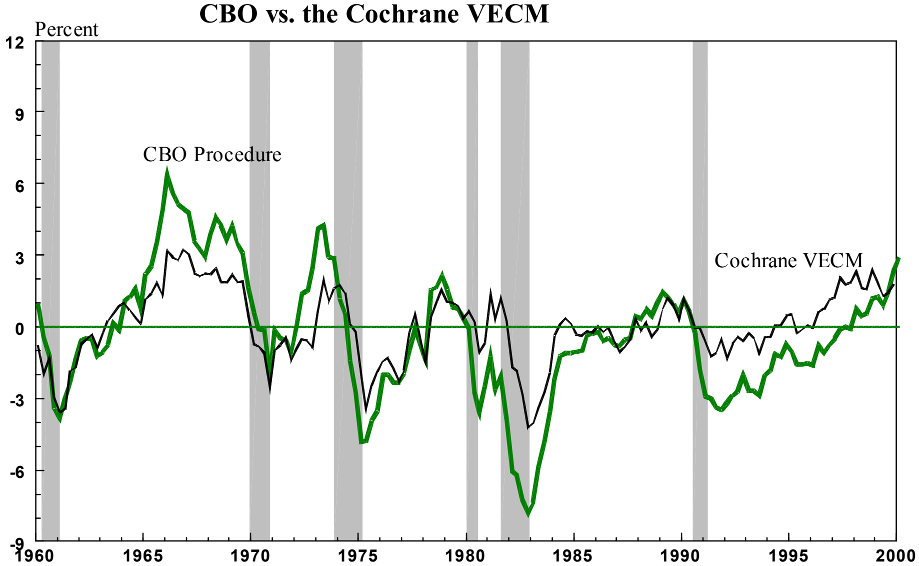
Figure 8. Measures of the Hours Gap and Labor Productivity Gap Consistent with the Gali VAR



SOURCE: Authors' calculations using data from the Department of Commerce, Bureau of Economic Analysis; and the Department of Labor, Bureau of Labor Statistics.

NOTE: The hours gap and productivity gap are measured as the difference between actual hours worked and productivity, respectively, as a percentage of their trend levels.

Figure 9. Comparison of Output Gap Estimates: CBO vs. the Cochrane VECM



SOURCE: Authors' calculations using data from the Department of Commerce, Bureau of Economic Analysis.

NOTE: The output gap is the difference between actual and potential GDP as a percentage of potential GDP.

of severity of the early 1980s downturns (measured by the size of the output gap) can also be explained in the same way. In those years, consumption fell along with output, which the model interprets as a decline in potential output, making the gap smaller than that suggested by the other models.

<u>Rates of Gap Closure</u>. The speed at which the gap disappears is important for the shorter-term forecasts, especially two- or three-year horizons, but not for the longer-term forecasts, when the gaps are closed. The "half-life" of the output gap—defined as the number of quarters it takes for the output gap to be half closed from its most recent historical value—varies between five and eight quarters in the three models (see Table 3). The lag order appears to be an important determinant of the speed of decay. (The lag length in the Cochrane VECM is two quarters as compared with

eight quarters in the other two models.) Although this does not have to be the case in general, the model that is specified with a shorter lag length tends to revert to its trend more rapidly.[28]

Table 3. Estimates of the Growth Rate of Potential Output and the Output Gap as Implied by Different Models

	Estimated Annual Growth Rate of Potential GDP in 1999:IV (In percent)	Estimated Output Gap in 1999:IV (As a percentage of potential GDP)	Decay Times of Output Gap (In quarters)			
			25% closed	50% closed	75% closed	90% closed
Blanchard-Quah VAR	3.2	4.2	5	8	15	26
Gali VAR[a]	3.5	-0.2	4	5	13	21
Cochrane VECM	3.4	1.8	4	6	9	12

SOURCE: Authors' calculations using data from the Department of Commerce, Bureau of Economic Analysis; and the Department of Labor, Bureau of Labor Statistics.

a. The output gap is close to zero at the end of the sample period, which may have inflated the reported decay times.

Forecast Performance. Table 4 compares the out-of-sample forecast performances of the three models with each other and with CBO's annual "real-time" projections.[29] Annual forecasts of real GDP beginning in 1976 were computed and compared with CBO's projections of the growth of real output made over the same period. For the first quarter of each year beginning in 1976, all variants of each of the models were estimated through the last quarter of the previous year and used to project real GDP

28. Note that the term "half-life" is typically used for processes with smooth exponential decay, but the decay implied by a VAR is generally not smooth. Table 3 therefore reports the number of quarters it takes for the output gap to lose 25 percent, 50 percent, 75 percent, and 90 percent of its most recent value.

29. Unlike the out-of-sample forecasts, real-time forecasts make use of the data available at the time the forecasts were made without the benefit of later revisions. CBO's projections of growth in real output are presented in Congressional Budget Office, *CBO's Economic Forecasting Record* (January 2002), available at www.cbo.gov.

for the next five years. Projected growth rates were calculated for two and five years ahead. Then, the process of estimation and forecasting was repeated for the next year, and so on.

The performance of the model-based forecasts is generally similar to that of forecasts produced by CBO over its history (see the rows showing the root mean square errors in Table 4). The main exceptions are that the Gali VAR performs substantially worse than CBO does in forecasts looking two years ahead, and the BQ VAR performs better than CBO does over the five-year horizon. Large errors (such as those occurring near turning points in the business cycle) appear in the VARs at about the same time as they do in CBO's real-time projections. Strikingly, the relative accuracy of the projections of the three models improves in the longer-term horizons.

However, the better performance of the BQ VAR over the five-year horizon (as compared with CBO's projections) must be discounted because the BQ VAR assumes a trend break in 1974. (The trend break was not yet established at the time, so CBO's real-time forecasts could not incorporate it.) The projections made before 1979 (the first four lines of the five-year projections in Table 4) are the ones in which CBO's forecasts have much larger errors than those of the BQ VAR.

Table 4. Comparison of Model Forecasts with CBO's Real-Time Forecasts of Real Output

	CBO	Blanchard-Quah VAR	Cochrane VECM	Gali VAR
Two-Year Growth				
1976 - 1977	1.1	-0.4	-2.3	-3.8
1977 - 1978	0.5	-2.1	-2.1	-5.3
1978 - 1979	0.5	-1.0	-1.5	-4.7
1979 - 1980	1.4	1.9	1.6	1.2
1980 - 1981	-0.3	0.9	1.4	1.3
1981 - 1982	2.2	3.2	2.7	4.8
1982 - 1983	1.3	2.5	0.2	1.0
1983 - 1984	-2.0	-0.8	-2.3	-3.8
1984 - 1985	-0.3	-0.5	-1.1	-2.8
1985 - 1986	0.3	0.1	-0.8	-2.3
1986 - 1987	0.3	0.3	-0.3	-0.5
1987 - 1988	-0.5	-0.4	-0.9	-0.8
1988 - 1989	-1.2	-0.3	-0.6	-1.1
1989 - 1990	0.2	0.4	0.6	0.5
1990 - 1991	1.9	1.6	1.7	2.5
1991 - 1992	0.8	0.6	-0.1	1.5
1992 - 1993	0.1	0.8	-0.3	1.1
1993 - 1994	0.0	1.1	0.2	0.9
1994 - 1995	-0.1	0.4	-0.2	-1.3
1995 - 1996	-0.4	-0.1	-0.3	-1.4
1996 - 1997	-1.7	-1.4	-1.4	-2.0
1997 - 1998	-1.8	-1.4	-1.8	-1.9
Mean Error	0.1	0.2	-0.3	-0.8
Mean Absolute Error	0.9	1.0	1.1	2.1
Root Mean Square Error	1.1	1.3	1.4	2.6
Five-Year Growth				
1976 - 1980	2.0	-0.0	-1.7	-2.7
1977 - 1981	2.3	0.1	-0.3	-1.9
1978 - 1982	3.2	1.3	1.0	0.1
1979 - 1983	2.5	1.2	1.6	1.7
1980 - 1984	0.4	0.2	0.6	0.3
1981 - 1985	0.0	0.2	-0.3	0.1
1982 - 1986	0.1	0.2	-1.2	-1.2
1983 - 1987	-0.3	0.2	-1.4	-2.6
1984 - 1988	0.0	-0.2	-1.0	-1.9
1985 - 1989	0.1	-0.1	-0.9	-1.4
1986 - 1990	0.5	0.2	-0.2	-0.2
1987 - 1991	0.8	0.9	0.5	0.9
1988 - 1992	0.5	0.7	0.6	0.7
1989 - 1993	0.6	0.8	1.0	1.1
1990 - 1994	0.6	0.5	0.6	0.7
1991 - 1995	0.4	0.5	0.1	0.3
1992 - 1996	-0.2	0.3	-0.4	-0.4
1993 - 1997	-0.3	0.2	-0.3	-0.5
1994 - 1998	-0.7	-0.3	-0.8	-1.6
Mean Error	0.7	0.4	-0.1	-0.4
Mean Absolute Error	0.8	0.4	0.8	1.1
Root Mean Square Error	1.2	0.6	0.9	1.3

SOURCE: Authors' calculations.

Forecast Uncertainty. The standard deviation of bootstrap forecasts is about 1.3 percent to 1.9 percent of the mean bootstrap forecast for 2000, rising to 4 percent to 8 percent by 2010 (see the last three rows of Table 5, and Figures 10-12). Not surprisingly, the forecasts' standard errors increase as the forecast horizon lengthens, but less than proportionally to the forecast horizon.[30]

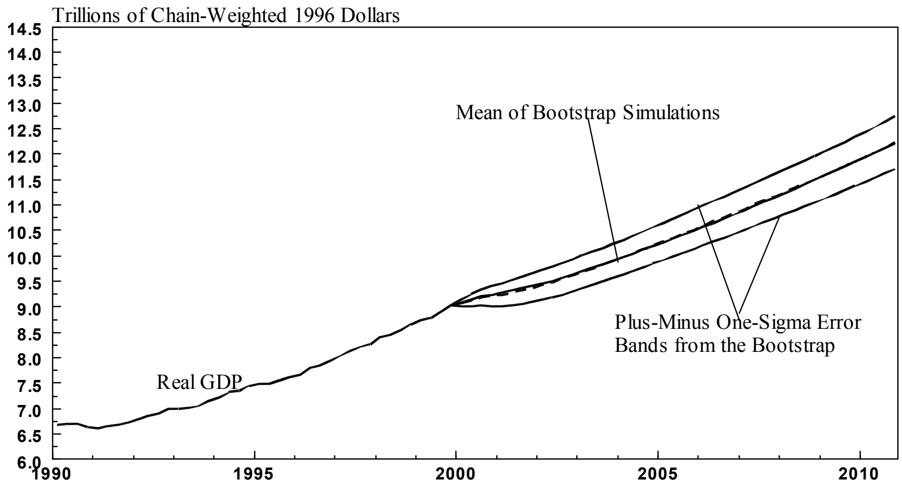
Table 5. Characteristics of the Time-Series Models' Real GDP Forecasts

	Year				
	2000	2001	2002	2005	2010
Forecasts of Real GDP (In billions of chained 1996 dollars)					
Blanchard-Quah VAR	9,170	9,340	9,540	10,380	12,080
Average Growth from 1999 (In percent)	3.4	2.7	2.5	2.7	2.9
Gali VAR	9,250	9,580	9,910	11,010	13,080
Average Growth from 1999 (In percent)	4.4	4.0	3.8	3.7	3.6
Cochrane VECM	9,200	9,430	9,710	10,720	12,690
Average Growth from 1999 (In percent)	3.8	3.2	3.1	3.2	3.3
The Range of the Forecasts (In billions of chained 1996 dollars)					
Size of range	80	240	370	630	1,000
Midpoint of range	9,210	9,460	9,725	10,695	12,500
The range (As a percentage of the midpoint)	0.9	2.5	3.8	5.9	7.9
Standard Deviation of Bootstrap Forecasts (As percentage of mean bootstrap forecast)					
Blanchard-Quah VAR	1.3	2.6	3.1	3.5	4.1
Gali VAR	1.9	3.4	4.1	5.7	8.0
Cochrane VECM	1.6	3.2	4.0	5.1	6.5

SOURCE: Authors' calculations using data from the Department of Commerce, Bureau of Economic Analysis; and the Department of Labor, Bureau of Labor Statistics.

30. Appendix B describes how the confidence bands are computed.

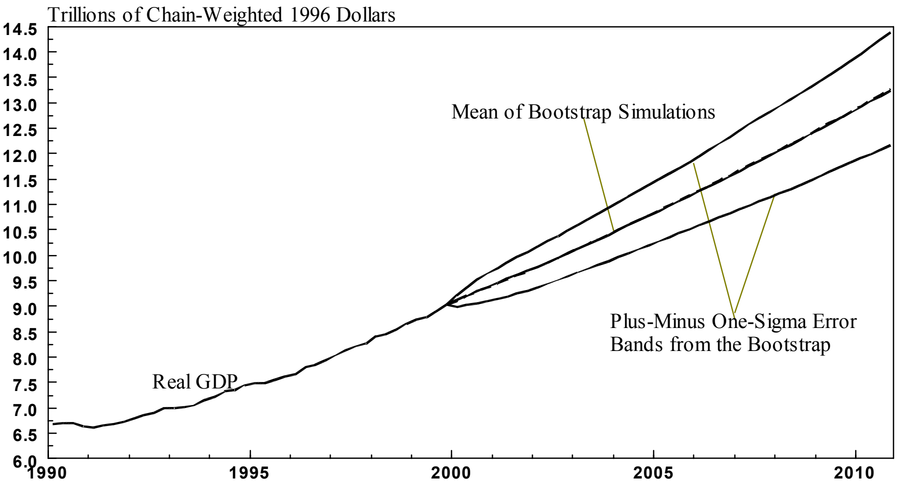
Figure 10. Medium-Term Projections of Real GDP with Bootstrap Error Bands: Blanchard-Quah VAR



SOURCE: Authors' calculations using data from the Department of Commerce, Bureau of Economic Analysis; and the Department of Labor, Bureau of Labor Statistics.

NOTE: The bootstrap is carried out by back-casting the series conditional on the last *p* observations of the data. The number of replications is 2,000. See Table 5 for error percentages.

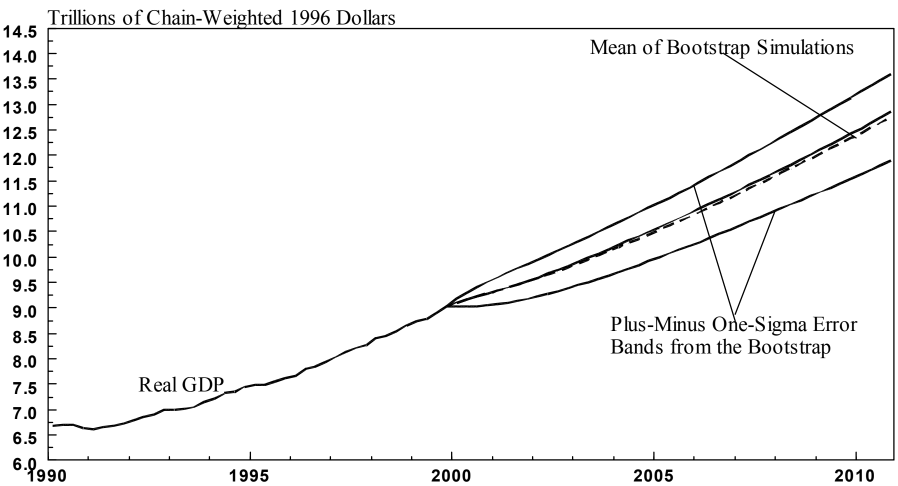
Figure 11. Medium-Term Projections of Real GDP with Bootstrap Error Bands: Gali VAR



SOURCE: Authors' calculations using data from the Department of Commerce, Bureau of Economic Analysis; and the Department of Labor, Bureau of Labor Statistics.

NOTE: The bootstrap is carried out by back-casting the series conditional on the last *p* observations of the data. The number of replications is 2,000. See Table 5 for error percentages.

Figure 12. Medium-Term Projections of Real GDP with Bootstrap Error Bands: Cochrane VECM



SOURCE: Authors' calculations using data from the Department of Commerce, Bureau of Economic Analysis; and the Department of Labor, Bureau of Labor Statistics.

NOTE: The bootstrap is carried out by back-casting the series conditional on the last p observations of the data. The number of replications is 2,000. See Table 5 for error percentages.

The most important source of uncertainty in the forecasts for real GDP in 2010 is uncertainty in the level of potential output in that year; in contrast, the uncertainty resulting from the cyclical position in 2010 is a minor source of error (see columns 3 and 2, respectively, of Table 6). The standard error in the estimate of potential output in 2010 is comparable to the error in the 2010 real GDP forecast in all three models (see columns 1 and 3).

Table 6. Uncertainty of the Forecasts and Forecast Components (Standard deviations of the models' projections for real GDP in 2010, as a percentage of their mean forecasts)

	Standard Deviation of the Forecast Components							
	(1) 2010 GDP Forecast	(2) 2010 GDP Gap Estimate	(3) 2010 Potential GDP Forecast	(4) Random Walk Component of Potential GDP in 2010	(5) Trend Line in 2010	(6) Accumulated Trend Growth	(7) Trend Line Intercept	(8) Correlation of Columns 6 and 7
BQ VAR	4.8	2.4	4.0	2.9	2.5	2.0	1.3	(0.13)
Gali VAR	8.3	1.7	8.0	7.2	3.4	3.3	0.5	(0.19)
Cochrane VECM	6.7	2.1	5.7	5.0	2.7	2.3	0.7	(0.46)

SOURCE: Authors' calculations using data from the Department of Commerce, Bureau of Economic Analysis; and the Department of Labor, Bureau of Labor Statistics.

NOTE: The reported values are the standard deviations as a percentage of real GDP in 2010, computed using the natural logs of variables multiplied by 100. The GDP forecasts for 2010 (the standard deviation of which is in column 1) equals the sum of 2010 GDP gap (whose standard deviation is in column 2) and potential GDP in 2010 (column 3). Potential GDP has both random walk and trend components (columns 4 and 5). The uncertainty in year 2010 position of the trend line, in turn, arises from the uncertainty in the trend growth rate, which accumulates between the current period and year 2010 (column 6) and the uncertainty in the intercept of the trend line (column 7). The variables in columns (4) and (5) are virtually uncorrelated. The correlation of columns (6) and (7) is reported in column (8). The trend line intercept (column 7) is determined by the estimate of the current output gap.

In turn, the forecast error in the projection of future potential GDP (whose standard deviation is in column 3) can arise either from an inaccurate estimate of the trend rate of growth (column 6) or from unexpected developments between the end of the sample period and 2010 that will make potential GDP grow faster or slower than the trend rate (column 4). The latter is the more important source of uncertainty in estimating future potential GDP. This is a pessimistic result for medium-term forecasts. It means that even if the gap and the average trend rate of growth were known precisely, the randomness in growth of potential output, which makes the growth of potential output in a given year differ from the average trend, would accumulate over time and create a large amount of uncertainty. The estimate of the intercept of the trend line (column 7), which is mainly determined by the estimate of the output gap at the end of the sample period, is responsible for a small portion of the error in the estimate of potential GDP in 2010.

The bootstrap error bands are conditional on the assumption that the models' specifications are correct. For that reason, the bands probably understate the uncertainty of the models' forecasts (they might overstate the uncertainty as well). Because the correct specifications are unknown, that uncertainty adds to the inherent uncertainty of the models' forecasts. Finally, as the large differences among the forecast standard errors of the three models illustrate, different models imply different levels of uncertainty—for example, the standard errors for the forecast of the Gali VAR are about twice those for the BQ VAR in 2010 (see Table 5). Similarly, none of those standard errors necessarily corresponds to the uncertainty of CBO's forecasts.

Sensitivity of Results to Alternative Specifications. We have examined the sensitivity of the models' forecasts to lag lengths and the assumptions regarding the transformations required to make the time series stationary. As discussed in the previous section, we adopted the lag lengths used in the original studies underlying those alternative time-series models. It turns out, in each case, that this lag length was at least as large as the lag length selected by a standard battery of information criteria (see Table 7). Using the shorter lag lengths would have had a minimal effect on the forecast for real GDP in 2010. Moreover, the out-of-sample root mean square errors seemed to be affected relatively little by lag length for two-year forecasts, and only small differences were manifest for the 10-year forecast horizon.

Although lag length may not have appeared to matter to a great degree, the choice of the stationarity transformation does matter. In the case of the BQ VAR, which seemed superior to the other VAR and the VECM in forecasting over the longer horizons, the exclusion of the trend break in 1974 substantially worsened the performance in forecasting 10 years ahead. Including the trend break did not consistently improve the forecasts of the other models, however.

Table 7. Effects of Alternative Lag Lengths and Transformations on the Forecast

		Selected Autoregressive Order by Information Criterion (Lag in quarters)					Root Mean Square Errors for Out-of-Sample Projections of Real GDP, Estimated over the Period 1970-1999 (Percentage of historical level)				Forecast of Average Annual Growth of Real GDP, 2000-2010 (In percent)	
							Two-Year Horizon		Ten-Year Horizon			
		AIC	SBIC	FPE	HQ	Order Used in Original Study	Smallest Lag Selected	Order Used in Original Study	Smallest Lag Selected	Order Used in Original Study	Smallest Lag Selected	Order Used in Original Study
Blanchard-Quah VAR (Q denotes output of the nonfarm business sector, U denotes unemployment rate for married males)												
No trend break in Q growth	No linear trend in U	3	2	2	3	8	3.3	3.5	6.2	6.4	3.4	3.5
No trend break in Q growth	Linear trend in U	3	2	2	2	8	3.3	3.4	6.1	6.3	3.3	3.4
Trend break in Q growth	No linear trend in U	6	2	2	2	8	2.9	3.0	2.9	3.2	2.9	2.9
Trend break in Q growth	Linear trend in U	6	2	2	2	8	2.9	3.0	2.8	3.1	2.9	2.9
Linear trend in log Q, with trend break	No linear trend in U	6	2	2	2	8	2.6	2.7	2.3	2.4	2.8	2.8
Linear trend in log Q, with trend break	Linear trend in U	3	2	2	2	8	2.6	2.9	2.3	2.6	2.8	2.9
Gali VAR (dY/L denotes the growth rate of nonfarm business productivity, dL denotes the growth rate of hours worked in the nonfarm business sector)												
No trend break in dY/L	No trend break in dL	3	2	2	2	8	4.7	4.6	7.9	6.6	3.7	3.6
No trend break in dY/L	Trend break in dL	3	2	2	2	8	5.4	4.8	12.4	7.9	4.2	3.9
Trend break in dY/L	No trend break in dL	3	2	2	2	8	5.2	5.3	12.2	11.7	3.0	3.0
Trend break in dY/L	Trend break in dL	3	2	2	2	8	5.8	5.4	15.1	11.7	3.3	3.3
Cochrane VECM (dQ denotes growth of real nonfarm business output, dC denotes growth of real consumption)												
No trend break in dQ	No trend break in dC	2	1	1	1	2	3.9	3.8	11.0	10.8	3.1	3.0
No trend break in dQ	Trend break in dC	2	1	1	1	2	3.1	3.0	5.9	6.1	3.3	3.0
Trend break in dQ	No trend break in dC	3	1	1	1	2	3.4	3.1	6.0	5.7	3.3	3.3
Trend break in dQ	Trend break in dC	1	1	1	1	2	3.2	3.4	8.4	9.4	3.1	3.1

SOURCE: Authors' calculations using data from the Department of Commerce, Bureau of Economic Analysis; and the Department of Labor, Bureau of Labor Statistics.

APPENDIX A: Alternative Approaches to Estimating the Output Gap

Because estimates of potential GDP and the output gap in the three vector autoregression (VAR) models rely on the vector-moving-average (VMA) representation of real GDP, we are including a brief explanation of those calculations.

Time-series analysis relies heavily on a result known as Wold's Decomposition Theorem. If $Z_t / [Z_{1t}, Z_{2t}, \ldots , Z_{nt}]$ is a stationary n-vector time series, then it can be decomposed into:

(i) a linearly deterministic component, the n-vector δ_t, that can be estimated arbitrarily well given the information set through period t, and

(ii) a random component, consisting of a moving average of a white noise process.

In practice, not much is lost if it is assumed that δ_t is independent of time ($\delta_t = \delta$). Therefore, Z can be written in the form:

$$Z_t = \delta + C(‹)\varepsilon_t \tag{A-1}$$

where C is an n-by-n matrix whose elements are infinite-order polynomials in the lag operator ‹ (that is, $‹^k Z_t = Z_{t-k}$ for any integer k), the poles of $C(‹)$ are outside the unit circle, and $C(0) / I_n$ so that the decomposition is unique.

Isolating the Trend Component

Two different approaches to decomposing the observed output into trend and cycle components are used. In cases in which both methods are applied, the estimated components turn out to be fairly similar. The first method is the Beveridge-Nelson (BN) decomposition. It defines the cycle component as the transitory component of output that is bound to disappear in the long run; the trend component is defined as the remaining part. The other decomposition method is long-run restrictions. In that method, the random shocks to the system are grouped on the basis of whether they have a permanent long-run impact. The system is simulated using only the shocks that have no permanent effect so as to identify the part of output that the transitory shocks are responsible for. That part is then taken as the cycle component, and the remaining part is the trend component.

Those two basic ways of extracting the trend component of output are examined in more detail below. The relationship (A-1) is central to each.[1]

The Beveridge-Nelson Decomposition

The first approach to extracting the trend component examined here comes from Beveridge and Nelson.[2] They note that if the stationary time series Z represents the first differences of an I(1) vector time series, say X, then the application of (A-1) to ΔX gives rise to an intuitively appealing definition of the trend, or permanent, component of X. Beveridge and Nelson define the permanent component of X_t to be the midpoint of the predictive distribution (at time t) for the future path of X excluding its trend, and the transitory component of X_t to be the expected momentum in X that is predictable at time t.[3] In the present notation, the BN decomposition

---

1. Some expansion of this notation would be required to accommodate the presence of cointegration relationships among the variables in Z. For the sake of expositional simplicity, the examples in this section assume no cointegration. The only model examined here that includes cointegrated terms is the estimate of the Cochrane model. In an appendix to his study, Cochrane outlines his method for extracting the permanent component of output assuming one cointegrating relationship. For a general exposition of the themes of this section, which does not preclude the possibility of cointegrating relationships among the variables in Z, see G. Evans and L. Reichlin, "Information, Forecasts, and Measurement of the Business Cycle," *Journal of Monetary Economics*, vol. 33, no. 2 (April 1994), pp. 233-254.

2. S. Beveridge and C. R. Nelson, "A New Approach to Decomposition of Economic Time Series Into Permanent and Transitory Components with Particular Attention to Measurement of the 'Business Cycle'," *Journal of Monetary Economics*, vol. 7, no. 2 (March 1981), pp. 151-174. The multivariate decomposition is presented in ibid., G. Evans and L. Reichlin.

3. Ibid., Beveridge and Nelson, pp. 153-154.

gives:

$$\Delta X^*_t = \delta + C(1)\varepsilon_t \qquad \text{(A-2a)}$$

$$\Delta X_t = \Delta X^*_t - [(1-\text{‹})\sum_{k=0}^{\infty}\sum_{j>k}^{\infty} C(j)\,\text{‹}^{k}]\ \varepsilon_t \qquad \text{(A-2b)}$$

in which X^*_t denotes the permanent component of X under the BN decomposition. From (A-2a), the permanent component of the n-vector X follows a random walk with drift δ. The period t innovation to the random walk, ε_t, has a permanent effect on the level of X* at time t. The matrix C(1), which is the sum of the VMA coefficients in (A-1), is the long-run impact multiplier. The product $C(1)\varepsilon_t$ cumulates the current and all future effects of ε_t on X and allocates them to X* at time t.

The key to estimating the BN decomposition is estimating C(1). Following the original studies, the assumption that C(‹) in (A-1) is an invertible matrix is imposed so that the resulting vector autoregressive form can be estimated:

$$[I_n\text{-}A(\text{‹})](Z_t - \delta) = \varepsilon_t \qquad \text{(A-3)}$$

where $A(\text{‹})$ / $I_n - C^{-1}(\text{‹})$. The coefficients of the (possibly) infinite-order VAR, A(‹), can be estimated by assuming a finite VAR order of p sufficiently large to approximate the order of the polynomials A(‹) and by applying ordinary least squares to the resulting specification. If those estimates of A(‹) are denoted by $A_{OLS}(\text{‹})$, then the entire sequence of VMA coefficients can be recovered as:

$$C(\text{‹}) – C_{OLS}(\text{‹}) \;/\; [I_n - A_{OLS}(\text{‹})]^{-1} \qquad \text{(A-4)}$$

in which the quality of the approximation depends on the lag length p used in estimating the VAR.[4]

Long-Run Identifying Restrictions

Although the BN decomposition's association of the permanent component of a series with its long-run forecast seems intuitive, the decomposition is by no means unique. In their study, Blanchard and Quah (BQ) adopted an alternative approach to extracting the permanent component of output by introducing identifying restrictions on the matrix C(1). Their decomposition is easiest to describe for the case of n=2.[5] In this case, the stochastic dynamics of the two-variable system are governed by two types of random shocks, one permanent and the other transitory. The permanent and transitory shocks are assumed to be statistically independent. If the two-dimensional vector of structural shocks is denoted by $\xi \equiv (\xi^{\pi}, \xi^{\tau})'$ for permanent and transitory, respectively, the Wold form for the BQ decomposition of the stationary vector Z can be written as

$$Z_t = \delta + B(L)\xi_t \tag{A-5}$$

where $E(\xi_t\xi_t')$ is a diagonal 2-by-2 matrix, and the matrix of long-run coefficients, B(1), is lower triangular. This lower triangularity means that ξ^{τ} (temporary shock) has no long-run effect on the first variable in Z (taken by BQ to be growth in real GDP).

The coefficients of the BQ structural form (A-5) are related to equation (A-1) as follows:

---

4. In the computations, a standard recursive formula was used to evaluate the VMA coefficients. See, for example, James D. Hamilton, *Time Series Analysis* (New Jersey: Princeton University Press, 1994), p. 260. Of course, the quality of the approximation depends even more critically on the invertibility assumption that is maintained throughout this study.

5. Higher-order variants of the Blanchard-Quah decomposition are easily implemented. For an interpretation of the BQ decomposition with n=3, see C. Dupasquier, A. Guay, and Pierre St-Amant, *A Comparison of Alternative Methodologies for Estimating Potential Output and the Output Gap*, Working Paper No. 97-5 (Ottawa, Canada: Bank of Canada, February 1997).

$$\xi_t \;/\; S^{-1} \varepsilon_t \qquad \text{(A-6a)}$$

$$B(\text{‹}) \;/\; C(\text{‹})S \qquad \text{(A-6b)}$$

with B(1) a lower triangular matrix. For the two-dimensional case, the nonsingular matrix S is exactly identified by the orthogonality requirement on the structural shocks and the requirement that the transitory shocks have no permanent effect on the first variable in the system.

Using this structural VAR approach for identifying a two-variable model, the permanent components of Z under the BQ decomposition are given by:

$$Z^*_{1t} = \delta_1 + B_{11}(\text{‹})\, \xi^{\pi}{}_t \qquad \text{(A-7a)}$$

$$Z^*_{2t} = \delta_2 + B_{21}(\text{‹})\, \xi^{\pi}{}_t. \qquad \text{(A-7b)}$$

By design, the permanent component of Z_1 is unaffected by the transitory shocks.

APPENDIX B: The Bootstrap Procedure

A bootstrap procedure is used to estimate confidence intervals for the alternative forecasts.[1] The steps are:

- The backward representation of each model is estimated.[2]

- The *L* most recent observations in the data are taken as given (where *L* is the longest lag length in the model). Holding these last *L* values fixed at their historical values, artificial data are generated for the remainder of the historical period by simulating the estimated backward representation and resampling the residuals estimated from the backward representation.

- The resulting artificial sample is then used, in the usual way, to generate a projection for the variables in the model.

The latter two steps were repeated 2,000 times. The resulting 2,000 alternative projections were used to generate the forecast confidence bands.

The advantage of using resampling techniques in general is that they generate nonparametric confidence intervals—confidence intervals that do not rely on particular assumptions about the distribution of disturbance terms. Although the residuals from the backward regression are not always serially independent, they are

---

1. The procedure is more fully described in Lori A. Thombs and William R. Schucany, "Bootstrap Prediction Intervals for Autoregression," *Journal of the American Statistical Association*, vol. 85, no. 410 (June 1990), pp. 486-492.

2. The backward representation of an autoregressive system expresses the current values of the endogenous variables in terms of their future values.

uncorrelated.[3] The serial independence, which is needed to justify resampling of residuals, is ensured, for instance, in the case of normally distributed disturbance terms.

3. For a discussion of this point, see Lutz Kilian and Ufuk Demiroglu, "Residual-Based Tests for Normality in Autoregressions: Asymptotic Theory and Simulation Evidence," *Journal of Business and Economic Statistics*, vol. 18, no. 1 (January 2000), p. 41.