

HOW SHOULD THE FED REPORT UNCERTAINTY?

By

Ray C. Fair

June 2012

COWLES FOUNDATION DISCUSSION PAPER NO. 1864



**COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281**

<http://cowles.econ.yale.edu/>

How Should the Fed Report Uncertainty?

Ray C. Fair*

June 2012

Abstract

In January 2012 the Fed began reporting ranges of its economic forecasts. The ranges, however, measure differences of opinion, not variances of economic forecasts. This paper discusses what the Fed could report in a world in which it used a single macroeconomic model to make its forecasts and guide its policies. Suggestions are then made as to what might be feasible for the Fed to report given that it is unlikely to be willing to commit to a single model.

1 Introduction

In January 2012 the Fed began reporting ranges of its economic forecasts along with central tendencies. Projections are collected from each member of the Board of Governors and each president of a Federal Reserve bank, up to 19 in all. The ranges are taken from these projections. The ranges thus measure differences of opinions among the participants. These ranges, while interesting, are not measures of uncertainty usually associated with economic forecasts. They are not, say, estimated standard errors of forecasts. There could be, for example, broad consensus about the mean paths of future economic variable values, making the ranges

*Cowles Foundation and International Center for Finance, Yale University, New Haven, CT 06520-8281. Phone: 203-432-3715; e-mail: ray.fair@yale.edu; website: fairmodel.econ.yale.edu.

small, and yet considerable uncertainty around the mean paths—say because of uncertainty about Europe or Asia or U.S. fiscal policy.

This paper discusses what the Fed could report in a world in which it used a single macroeconomic model to make its forecasts and guide its policies. Suggestions are then made as to what might be feasible for the Fed to report given that it is unlikely to be willing to commit to a single model. A particular model—the MC model in Fair (2004)—is used as an illustration.

The general model is presented in Section 2; estimation is discussed in Section 3; and solution is discussed in Section 4. Stochastic simulation, a tool that can be used to estimate uncertainty, is discussed in Section 5. The solution of optimal control problems in the deterministic case is discussed in Section 6, and stochastic simulation and optimal control are discussed in Section 7. The example is presented in Section 8. Section 9 then provides suggestions for what the Fed might be able to do in practice.

A key procedure in this paper is the use of certainty equivalence. When a nonlinear model is solved setting the errors to their expected values (assumed to be zero throughout this paper), the solution value of a variable is not its expected value. The assumption in this paper is that the difference between the two values is small enough to be ignored. Expected values can be estimated using stochastic simulation, and when this is done it is usually the case for macroeconomic models that the estimated expected value is close to the solution value with the errors set to zero.¹ The use of stochastic simulation in this paper is not to improve upon estimates of expected values, but to estimate uncertainty.

¹See, for example, Fair (2004, Table 9.4, p. 127).

2 A General Model

The general model considered in this paper is dynamic, nonlinear, simultaneous, and may have rational (model consistent) expectations:

$$f_i(y_t, y_{t-1}, \dots, y_{t-p}, E_{t-1}y_t, E_{t-1}y_{t+1}, \dots, E_{t-1}y_{t+h}, x_t, \alpha_i) = u_{it} \quad (1)$$

$$i = 1, \dots, n, \quad t = 1, \dots, T,$$

where y_t is an n -dimensional vector of endogenous variables, E_{t-1} is the conditional expectations operator based on the model and on information through period $t - 1$, x_t is a vector of exogenous variables, and α_i is a vector of coefficients. The first m equations are assumed to be stochastic, with the remaining equations identities. The vector of errors, $u_t = (u_{1t}, \dots, u_{mt})'$, is assumed to be *iid*. The function f_i may be nonlinear in variables, coefficients, and expectations. u_i will be used to denote the T -dimensional vector $(u_{i1}, \dots, u_{iT})'$. α will be used to denote the vector of all the coefficients in the model. The vector x_t can include lagged values of exogenous variables as well as values for period t .

This specification is fairly general. It includes as a special case the VAR model. It also incorporates autoregressive errors. If there are no expectation variables on the right hand side in an equation i and if the original error term in the equation follows a r th order autoregressive process, say $w_{it} = \rho_{1i}w_{it-1} + \dots + \rho_{ri}w_{it-r} + u_{it}$, then the equation can be assumed to have been transformed into one with u_{it} on the right hand side. The autoregressive coefficients $\rho_{1i}, \dots, \rho_{ri}$ are incorporated into the α_i coefficient vector, and additional lagged variable values are introduced. This transformation makes the equation nonlinear in coefficients if it were not otherwise, but this adds no further complications because the model is already allowed to be

nonlinear. The coefficient vector α thus includes any autoregressive coefficients. The assumption that u_t is *iid* is thus not as restrictive as it would be if the model were required to be linear in coefficients. If there are expectation variables on the right hand side of an equation, the treatment of autoregressive errors is more complicated because there is now more than one viewpoint date. The treatment in this case is discussed in Fair and Taylor (1983, 1990). In what follows the model in (1) will be called a “RE” model if there is at least one right-hand-side expectation variable in at least one of the equations.

3 Estimation

One method for estimating the model in (1) is two-stage least squares (2SLS) equation by equation. For an equation i with no right-hand-side expectation variables, the 2SLS estimate of α_i is obtained by minimizing

$$S_i = u_i' Z_i (Z_i' Z_i)^{-1} Z_i' u_i \quad (2)$$

with respect to α_i , where Z_i is a $T \times K_i$ matrix of first stage regressors. The first stage regressors can be a subset of the predetermined variables in the model, where the predetermined variables are assumed to be correlated with the right-hand-side endogenous variables in the equation but not with the error term.

If there are right-hand-side expectation variables, Hansen’s (1982) method can be used to estimate α_i . For example, say that $E_{t-1}y_{2t+1}$ and $E_{t-1}y_{2t+2}$ are postulated to be explanatory variables in equation i . If it is assumed that variables in a matrix Z_i are used in part by agents in forming their (rational) expectations, then

Hansen's method in this context is simply 2SLS with adjustment for the moving average process of the error term. The expectations variables are replaced by the actual values y_{2t+1} and y_{2t+2} , and the first stage regressors are the variables in Z_i . Consistent estimation does not require that Z_i include all the variables used by agents in forming their expectations. The requirement for consistency is that Z_i be uncorrelated with the expectation errors, which is true if expectations are rational and Z_i is at least a subset of the variables used by the agents.²

Another estimation method is full information maximum likelihood (FIML). Under the assumption that u_t is independently and identically distributed as multivariate normal $N(0, S)$, FIML estimates of α are obtained by maximizing

$$L = -\frac{T}{2} \log |S| + \sum_{t=1}^T \log |J_t| \quad (3)$$

with respect to α , where S is the $m \times m$ covariance matrix of the errors and J_t is the $n \times n$ Jacobian matrix for period t . The ij element of S is $(1/T) \sum_{t=1}^T u_{it}u_{jt}$.

For non RE models estimates of $u_t, t = 1, \dots, T$, can be obtained from (1) given the data and a value of α . The elements of the Jacobian can also be computed. For a given value of α , L can thus be computed, and so the FIML maximization problem can be turned over to a nonlinear maximization algorithm. My experience is that the Parke (1982) algorithm works well for large models.

For RE models estimates of u_t cannot be obtained until the expectations have been computed. The first step given the data and a value of α is to compute the expectations. This can be done using the EP method discussed in the next section. Given the expectations, estimates of u_t and of the Jacobian elements can be

²For more details, including the case in which u_{it} is serially correlated, see Fair (1993) or Fair (1994), pp. 65-70.

obtained. Since the expectations have viewpoint date $t - 1$, they are predetermined from the point of view of taking derivatives for the Jacobian. For RE models L can thus be computed for a given value of α , and the problem can be turned over to a maximization algorithm. The extra work for RE models is solving for a given value of α the expectations for each of the T viewpoint dates. Various tricks to lessen computational time are discussed in Fair and Taylor (1990). FIML estimates have the advantage of incorporating all the nonlinear restrictions on the reduced form coefficients. For RE models, for example, all the nonlinear restrictions are used in computing the expectations.

4 Solution

For non RE models the solution of (1) is trivial. Consider the solution for period t . Given an estimate of α , given values of the lagged endogenous variables (values for periods $t - 1$ and back), given values of the variables in x_t , and given values of the errors u_t , values for y_t can be computed using a technique like Gauss-Seidel. A solution for period $t + 1$ is dynamic if the solution values of the endogenous variables for period t are used in computing the solution values for period $t + 1$, where values of x_{t+1} and u_{t+1} are also needed. A static solution for period $t + 1$ is one in which the actual values of y_t are used. The values of the errors are usually set to their expected values, which, as mentioned in Section 1, are taken to be zero in this paper.

For RE models the solution for period t requires that the expectations for periods $t, t + 1, \dots, t + h$ be computed first. Agents are assumed to solve the model to

compute these expectations given values they choose before solution of the current and future values of the exogenous variables and errors. It does not have to be the case that the values they use for x_t are the values used in the solution of the model for period t , but for simplicity this will be assumed here. It will also be assumed that the agents always use zero for the errors.

Consider solving the model for period t . A popular method is the extended path (EP) method in Fair and Taylor (1983). The method iterates over solution *paths*. Values of the expectations for period t through period $t+h+k+h$ are first guessed, where h is the maximum lead in the model and k is chosen as discussed below. Given these guesses, the model can be solved for periods t through $t+h+k$ in the usual ways (usually period by period using the Gauss-Seidel technique). This solution provides new values for the expectations through period $t+h+k$, namely the solution values. Given these new values, the model can be solved again for periods t through $t+h+k$, which provides new values for the expectations, and so on. Convergence is reached when the predicted values for periods t through $t+h$ from one iteration to the next are within a prescribed tolerance level of each other. (There is no guarantee of convergence, but in most applications convergence is not a problem.)

In this process the guessed values of the expectations for periods $t+h+k+1$ through $t+h+k+h$ (the h periods beyond the last period solved) have not been changed. If the solution values for periods t through $t+h$ depend in a nontrivial way on these guesses, then overall convergence has not been achieved. To check for this, the entire process can be repeated for k one larger. If increasing k by one has a trivial effect (based on a tolerance criterion) on the solution values for

t through $t + h$, then overall convergence has been achieved; otherwise k must continue to be increased until the criterion is met. In practice what is usually done is to experiment to find the value of k that is large enough to make it unlikely that further increases are necessary for any experiment that might be run and then do no further checking using larger values of k . Since for a given k agents solve the model through period $t + h + k$, values of the exogenous variables and errors are needed through this period. Agents must have expectations of these values before solving the model.

After the expectations are computed the model can be solved for period t . In fact, if the agents use the same value of x_t as is used for the solution of the model, which is assumed here, the model has already been solved for t , namely the computed expectation for t . Given the solution for t , one can move on to the solution for $t + 1$. If the solution is dynamic and if agents use the same exogenous variable values as are used for the solution of the model, the solution for $t + 1$ has already been computed, namely the computed expectation for $t + 1$.

For FIML, T uses of the EP method are required per evaluation of L since the solution is static. Note regarding certainty equivalence that this procedure for computing FIML estimates for RE models is based on the assumption that the solution values are the expected values. Note also that values of the exogenous variables and errors are required beyond the last observation T because expectations beyond T have to be computed. For example, for the solution for the last period T , values are needed for $T + 1$ through $T + h + k$, where k is defined in the previous section.

5 Stochastic Simulation or Bootstrapping

The solutions described above are deterministic in that the errors have been set to fixed values and the model solved once per period. Estimates of uncertainty can be made by using stochastic simulation. Stochastic simulation has a long history in macroeconomics. The seminal paper in this area is Adelman and Adelman (1959), which introduced the idea of drawing errors to analyze the properties of econometric models. Closely related to stochastic simulation is the bootstrap procedure, introduced in statistics in 1979 by Efron (1979). In Fair (2003 and 2004, Chapter 9) I discuss the relationship between these two literatures and integrate for a model like (1) the bootstrap approach to evaluating estimators and the stochastic simulation approach to evaluating models' properties.

When solving a model there are two sources of uncertainty of the predicted values. One is from the fact that the coefficients are only estimated rather than known, and the other is from the additive errors. Much of the stochastic-simulation literature has taken the coefficients to be fixed and focused on the additive errors. The bootstrap literature has been concerned with evaluating estimators. In this paper uncertainty from the coefficient estimates is not taken into account. It is straightforward to do this—see Fair (2004, Chapter 9)—but beyond what a central bank is likely to do.

The aim in this paper, as should be the aim of the Fed, is to estimate as precisely as possible standard errors of the forecasts. The complication regarding the Fed is that it controls a key variable in the model, namely the federal funds rate. For present purposes there are at least three assumptions that can be made about Fed

behavior when computing standard errors. One is that the Fed simply sets a path of the funds rate and never deviates from this path. This is, of course, an unrealistic assumption since the Fed does respond to surprise changes in the economy. The second is that the Fed at the beginning of each period solves an optimal control problem in choosing the interest rate path. The third is that the Fed uses an interest rate rule, which then simply makes the interest rate an endogenous variable in the model. This section will assume that an interest rate rule is being used, and the next section discusses optimal control.

First assume that the Fed has made at the beginning of period $T + 1$ a “baseline” forecast for $T + 1$ and beyond. This forecast would be based on a set of coefficient estimates, lagged endogenous variables, and choices of current and future values of the exogenous variables and errors. As noted above, for RE models values of the exogenous variables and errors are needed beyond the end of the forecast horizon.

For a model estimated for periods 1 through T , estimates of u_t , $t = 1, \dots, T$, are available. These are the estimated residuals. One way of doing stochastic simulation is to draw from these errors. A vector of errors, u_t^j , can be drawn with replacement from the T vectors of estimated errors, with probability $1/T$ of drawing any particular vector. Another way of doing stochastic simulation is to assume that u_t is distributed $N(0, \hat{S})$, where \hat{S} is the estimated $m \times m$ covariance matrix. u_t^j would be a random draw from this distribution. The first procedure has the advantages that it is distribution free and uses the actual historical residuals. The second procedure has the advantage that it is consistent with the assumptions behind the estimation of the model. For present purposes all that needs to be assumed is that u_t^j can be drawn in some way.

Consider the solution for period $T + 1$ for a non RE model. For a given draw w_{T+1}^j the model can be solved, producing solution values y_{T+1}^j . Doing this J times produces a distribution of solution values, from which measures of central tendency and dispersion can be computed. The mean for endogenous variable i is

$$\bar{y}_{iT+1} = \frac{1}{J} \sum_{j=1}^J y_{iT+1}^j \quad (4)$$

and the variance is

$$\sigma_{iT+1}^2 = \frac{1}{J} \sum_{j=1}^J (y_{iT+1}^j - \bar{y}_{iT+1})^2 \quad (5)$$

Ignoring simulation error from the fact that J is finite, \bar{y}_{iT+1} is the expected value of y_{iT+1} . Again, however, the main reason for doing stochastic simulation is not to compute more accurate expected values but measures of dispersion.

For a dynamic simulation of, say, two periods, errors would be drawn for periods $T+1$ and $T+2$ and the model solved for the two periods using these draws. Solution values for the two periods would be computed for each endogenous variable, which is one trial. J trials would be done, producing J values of the predictions for each of the two periods. As many periods ahead can be done as desired. If the Fed were doing this, the interest rate rule would presumably be deterministic, and so errors would not be drawn for it.

For RE models if agents do not observe the error draw and continue to form expectations using zero values for the errors for periods $T + 1$ and beyond, the story for period $T + 1$ is the same as for non RE models. The expectations are predetermined regarding the draws. For a dynamic simulation the story for period $T + 2$ is different. Using the error draw for $T + 1$, the final solution values of the endogenous variables for $T + 1$ are different from what the agents expected them

to be, unlike in the deterministic case. At the beginning of $T + 2$ they would use the observed (solution) values of the endogenous variables for $T + 1$ in forming their expectations for $T + 2$ and beyond. The EP method must thus be used for each period solved, not just for period $T + 1$ as in the deterministic case. For a horizon of r periods, r error vectors would be drawn and the EP method used r times. This is one trial. After J trials there would be J solution values of each endogenous variable for each period, as in the non RE case.

6 Optimal Control: Deterministic Case

Now consider the case in which the Fed sets the federal funds rate by solving an optimal control problem. For a linear non RE model and a quadratic objective function, analytic, closed-form solutions are available—see, for example, Chow (1975). In this case, if the Fed reported the feedback equation, the coefficients in the model, and the covariance matrix of the errors, users would have all the information they need to compute means and variances.

In practice models are not linear, objective functions are generally not quadratic, and there may be rational expectations. The following is a more general problem. The deterministic case is considered in this section and the stochastic case in the next. Assume that the horizon is $t = T + 1, \dots, T + r$ and that the objective is to maximize the expected value of W , where W is

$$W = g(y_{T+1}, \dots, y_{T+r}, x_{T+1}, \dots, x_{T+r}) \quad (6)$$

In most applications the objective function is assumed to be additive across time,

which means that (6) can be written

$$W = \sum_{t=T+1}^{T+r} g_t(y_t, x_t) \quad (7)$$

Assume that the control variable of the Fed, the federal funds rate, is variable x_{1t} , and let $z = (x_{1T+1}, \dots, x_{1T+r})$. If all the errors are set to zero, then for each value of z one can compute a value of W by first solving the model for y_{T+1}, \dots, y_{T+r} and then using these values along with the values for x_{T+1}, \dots, x_{T+r} to compute W in (6) or (7). Stated this way, the optimal control problem is choosing values (the elements of z) to maximize an *unconstrained* nonlinear function. By substitution, the constrained maximization problem is transformed into the problem of maximizing an unconstrained function of the control variables:

$$W = \Phi(z) \quad (8)$$

where Φ stands for the mapping $z \longrightarrow y_{T+1}, \dots, y_{T+r}, x_{T+1}, \dots, x_{T+r} \longrightarrow W$. Given this setup, the problem can be turned over to a nonlinear maximization algorithm like DFP. For each iteration of the algorithm, the derivatives of Φ with respect to the elements of z , which are needed by the algorithm, can be computed numerically. An algorithm like DFP is generally quite good at finding the optimum for a typical control problem.³

Regarding the choice of r , the end of the horizon, it is sometimes the case that unusual results are obtained near the end of the horizon because there is no tomorrow. In practice the end of the horizon should be taken to be large enough so that end-of-horizon effects have small effects on the earlier control values of

³See Fair (1974) for various applications of this procedure.

interest. It will be assumed that this has been done in the following discussion. r should thus be thought of as much larger than the actual horizon of interest.

In practice the Fed would solve its control problem at the beginning of $T + 1$ and implement x_{1T+1}^* , the optimal value of the funds rate for period $T + 1$. It could also announce its plans for periods after that: $x_{1T+2}^*, \dots, x_{1T+r}^*$. In the deterministic case this would be it. The Fed would simply implement the optimal values for $T + 2$ and beyond as the time came. To look ahead to the next section, however, the world is not deterministic, and in practice the economy in $T + 1$ would not be what the Fed expected it to be when it solved its control problem. After period $T + 1$ is over, the Fed at the beginning of $T + 2$ could reoptimize. The optimal value for $T + 2$ would no longer be the value it computed at the beginning of $T + 1$ because the errors for $T + 1$ would not in general be zero, which was what the Fed was assuming at the beginning of $T + 1$. The Fed could thus behave by solving a series of open-loop optimal control problems, one at the beginning of each new period.

Turn now to the RE case and assume that the agents know what the Fed is doing. It is still the case that one can compute a value of W given a value of z . The extra work in the RE case is that the EP method must be used in the solution of the model. For a given z the expectations would be computed first and then the model solved. The Fed would assume zero current and future errors when solving its control problem, as would the agents in computing their expectations. The DFP algorithm could still be used to find the optimal value of z , and this value would be consistent with the expectations of the agents.

7 Stochastic Simulation and Optimal Control

Consider non RE models first. One trial is a draw of $u_{T+1}^j, \dots, u_{T+r}^j$. At the beginning of $T + 1$ the computed optimal value of z is not affected by the draws because the Fed assumes zero current and future errors. x_{1T+1}^* is implemented at the beginning of $T + 1$. The solution of the model for $T + 1$ uses this value and the error draw for $T + 1$. Then at the beginning of $T + 2$ the process is repeated, where the Fed uses the solution values of the endogenous variables for $T + 1$ and the assumption of zero errors for $T + 2$ and beyond. x_{1T+2}^* is implemented at the beginning of $T + 2$, and when the model is solved the error draw for $T + 2$ is used. The optimal value of x_{1T+2} is different from what the Fed computed at the beginning of $T + 1$ because the actual error draws for $T + 1$ are not in general zero. This process is repeated r times. This is one trial, so each trial requires the solution of r optimal control problems. After J trials are performed, the J values of each endogenous variable for each period can be used to compute variances. These variances would incorporate the optimal behavior of the Fed.

Now consider the RE case. Since agents also use zero current and future errors in computing their expectations, the expectations solved at the beginning of $T + 1$ are the same as in the deterministic case since the optimal value of the control variable for $T + 1$ is the same. Again, for each value of z tried by the algorithm, the EP method must be used. At the beginning of $T + 2$ the process is repeated, just as in the non RE case, where the agents, along with the Fed, use the solution values of the endogenous variables for $T + 1$, which are affected by the error draw. The process is done r times, which is then one trial.

The RE case is thus no different from the non RE case except that the EP method must be used each time W is computed. This is expensive, and various tricks would probably be needed in practice to lessen computational time.

8 An Example

The multicountry (MC) macroeconometric model in Fair (2004) is used for this example. The model has been updated for purposes of this paper—version dated April 27, 2012. The model is of the form (1) without any expectation variables, and it is estimated by 2SLS. The estimation periods begin in 1954 for the United States and 1962 for the other countries and go through the latest data at the time of the update. The model is quarterly, but some of the equations are estimated using annual data. There are 59 countries in the model.

There are 1,643 estimated equations in the MC model, of which 1,333 are bilateral trade share equations. For each estimated equation there are estimated residuals over the estimation period. Let \hat{u}_t denote the 1643-dimension vector of the estimated residuals for quarter t .⁴ Most of the estimation periods have the 1972:1–2009:4 period—152 quarters—in common, and these 152 observations on \hat{u}_t are used for the draws in the procedure discussed below.⁵

A baseline forecast was made for the 2012:1–2022:4 period using the MC model—44 quarters. This forecast is based on a particular set of exogenous variable

⁴For equations estimated using annual data, the error is put in the first quarter of the year with zeros in the other three quarters (which are never used).

⁵If an estimation period does not include all of the 1972:1–2009:4 period, zero errors are used for the missing quarters.

values and zero error values. This yields a predicted value of each endogenous variable for each quarter. The interest here is estimating the uncertainty around each predicted value. Each trial of the stochastic simulation procedure is as follows. First, 44 error vectors are drawn with replacement from the 152 vectors of residuals, each with probability $1/152$. (Each vector consists of 1,643 residuals.) Using these error vectors, one per quarter, the model is solved dynamically for the 2012:1–2022:4 period. The same set of exogenous variable values is used as was used for the baseline forecast. The solution values are recorded. This is one trial. The procedure is then repeated, say, J times. This gives J values of each endogenous variable for each quarter, from which measures of dispersion can be computed.

Results are presented in Table 1. They are based on 500 trials. (There were no solution failures on any of the trials.) The measure of dispersion used in the table is as follows. Rank the J values of a given variable for a given quarter by size. Let m_r denote the value below which r percent of the values lie. The measure of dispersion is $(m_{.8413} - m_{.1587})/2$. For a normal distribution this is one standard error.

In the MC model there is an estimated interest rate rule for each major country, including the United States. Monetary policy is thus endogenous in the model. For the results in Table 1 errors were drawn for the interest rate rules except for the rule for the United States. For the United States the estimated rule without any errors was taken to be the exact rule that the Fed uses. When for a particular draw the rule called for a negative interest rate, zero was used instead.

Remember that the results in Table 1 are based on historical residuals between 1979 and 2009, i.e., historically observed uncertainty. The residuals are assumed

Table 1
Estimated Standard Errors of Forecasts
Values are in Percentage Points

qtr	Y	UR	P	D	R
2012:1	0.57	0.23	0.31	0.42	0.15
2012:2	1.05	0.38	0.44	0.80	0.32
2012:3	1.51	0.60	0.58	1.22	0.41
2012:4	2.01	0.87	0.82	1.74	0.60
2013:1	2.20	1.12	1.02	2.10	0.78
2013:2	2.37	1.27	1.33	2.44	0.92
2013:3	2.52	1.38	1.45	2.76	1.01
2013:4	2.78	1.49	1.71	3.12	1.15
2014:4	2.95	1.66	2.56	3.92	1.46
2015:4	3.15	1.73	3.10	4.44	1.64
2016:4	3.31	1.80	3.55	4.83	1.73
2022:4	3.78	1.91	4.60	7.45	1.93

Y = real GDP

UR = unemployment rate

P = GDP deflator

D = nominal federal debt/nominal GDP

R = three-month Treasury bill rate

to be *iid* since, as discussed in Section 2, serial correlation has been removed when necessary by estimating autoregressive error coefficients.

The estimated standard errors in Table 1 increase with the length of the horizon, as expected. After four quarters the standard errors are 2.01 for real GDP, 0.87 for the unemployment rate, 0.82 for the GDP deflator, 1.74 for the debt/GDP ratio, and 0.60 for the interest rate. After eight quarters the respective standard errors are 2.78, 1.49, 1.77, 3.12, and 1.15. Note that even though no errors are drawn for the Fed interest rate rule, there is still considerable variation in the interest rate as the Fed reacts to the shocks.

The results in Table 1 are thus illustrative of what the Fed could report. Even though the MC model is large, the time taken per solution is modest, and the 500 trials for Table 1 took about 2 hours on a high-end laptop. The time would, of course, be larger if the model were a RE model or if optimal control problems were solved for each trial.

9 What Could the Fed Do In Practice?

An interest rate rule, as used in the MC model in the previous section, is likely too simple an explanation of Fed behavior. The solution of formal optimal control problems is likely too complicated. The following is something in between that the Fed might do that would allow uncertainty estimates to be computed.

First, the Fed must begin with a baseline forecast using a model. The model could, however, be subjectively adjusted by using non zero values of various current and future errors. Error values could be chosen to have the forecast from the model be what the Fed thinks is most likely. These error values would then be taken to be exogenous. This baseline forecast would include the Fed's chosen funds rate path.

Second, the Fed must have a way of drawing values of the errors in the model. This could be from a set of historically estimated residuals or from an assumed probability distribution with a computed covariance matrix.

The procedure could then be as follows. At the beginning of $T + 1$ draw error vectors for $T + 1$ through $T + r$ and solve the model using these errors and the baseline path of the funds rate. The error values added would be on top of any of

the exogenous error values discussed above. If the model is an RE model, the EP or similar method would have to be used in the solution. This solution path will obviously differ from the baseline path, and it is likely that if this path actually occurred, the Fed would not pick the baseline funds-rate path. The next step is thus for the Fed to change the funds-rate path to be the one it would pick if the particular error draw occurred. This could possibly be done by hand, with a few iterations needed to find the desired path. Or possibly this could be turned over to an algorithm. At any rate, at the end the funds-rate path would be consistent (in the eyes of the Fed) with the particular error draws. This is one trial. Now do this J times and compute measures of uncertainty for the J values of each endogenous variable for each period and for the funds rate.

The Fed could also do this for more than one model and report more than one set of results. This is likely to be more informative than to try to average uncertainty estimates across models and report the averages.

A key advantage of a procedure like this is that there is a different funds-rate path for each set of error draws, namely a path appropriate to that particular set. The path could be chosen by a rule, by solving optimal control problems, or in the in between manner discussed above. It is obviously not appropriate to keep the baseline funds-rate path the same for each set of draws, and a key part of any reporting should be estimates of the variances of the future funds rates.

The Bank of Norway—Aastveit, Gerdrup, and Jore (2011a) and Gerdrup and Nicolaisen (2011b)—reports forecast densities. It has 221 models for forecasting GDP and 171 models for forecasting the CPI. Each model produces density forecasts, and these densities are then combined into one for each of the two variables.

The combined forecasts are then fed into a DSGE model (NEMO) for policy analysis. The policy analysis in NEMO is conditional on these forecasts and judgment. The density forecasts from the individual models are not affected by the policy rate, and so the procedure used by the Bank of Norway is not the same as the one recommended above. No single model is used in the analysis. The final density forecasts from NEMO are in part based on the individual model's density forecasts and in part on judgment. An alternative procedure would be to use a single model for the entire analysis, as outlined in this paper, but then report results from many different models. There would then be consistency within each model.

10 Conclusion

Using certainty equivalence, it is feasible to compute measures of dispersion using stochastic simulation. This can be done for large nonlinear simultaneous equations models, including those with rational expectations. It is also possible to incorporate optimal control behavior into the analysis. This framework, or an approximation to it, could be used by monetary authorities in reporting uncertainty estimates. The key ingredients needed are 1) a model, possibly subjectively adjusted, 2) a set of historically estimated errors for drawing or an estimated probability distribution of errors, and 3) a way of changing the optimal path of the interest rate when errors are drawn, either an interest rate rule, optimal control, or something in between. The Fed's current reporting is only of differences of opinions, which are not measures of dispersion of endogenous variable forecasts.

References

- [1] Aastveit, Knut Are, Karsten R. Gerdrup, and Anne Sofie Jore, 2011b, “Short-Term Forecasting of GDP and Inflation in Real-Time: Norges Bank’s System for Averaging Models,” Norges Bank, Staff Memo, No. 09.
- [2] Adelman, I., and F.L. Adelman, 1959, “The Dynamic Properties of the Klein-Goldberger Model,” *Econometrica*, 27, 596–625.
- [3] Chow, Gregory C., 1975, *Analysis and Control of Dynamic Economic Systems*. New York: John Wiley & Sons.
- [4] Efron, B., 1979, “Bootstrap Methods: Another Look at the Jackknife,” *The Annals of Statistics*, 7, 1–26.
- [5] Fair, Ray C., 1974, “On the Solution of Optimal Control Problems as Maximization Problems,” *Annals of Economic and Social Measurement*, 3, 135–154.
- [6] Fair, Ray C., 1993, “Testing the Rational Expectations Hypothesis in Macroeconometric Models,” *Oxford Economic Papers*, 45, 169–190.
- [7] Fair, Ray C., 1994, *Testing Macroeconometric Models*, Cambridge, MA: Harvard University Press.
- [8] Fair, Ray C., 2003, “Optimal Control and Stochastic Simulation of Large Nonlinear Models with Rational Expectations,” *Computational Economics*, 245–256.
- [9] Fair, Ray C., 2003, “Bootstrapping Macroeconometric Models,” *Studies in Nonlinear Dynamics & Econometrics*, Vol. 7, No. 4, Article 1. <http://www.bepress.com/snede/vol7/iss4/art1>.
- [10] Fair, Ray C., 2004, *Estimating How the Macroeconomy Works*. Cambridge, MA: Harvard University Press.
- [11] Fair, Ray C., and John B. Taylor, 1983, “Solution and Maximum Likelihood Estimation of Dynamic Rational Expectations Models,” *Econometrica*, 51, 1169–1185.

- [12] Fair, Ray C., and John B. Taylor, 1990, “Full information Estimation and Stochastic Simulation of Models with Rational Expectations,” *Journal of Applied Econometrics*, 5, 381–392.
- [13] Gerdrup, Karsten R., and Jon Nicolaisen, 2011a, “On the Purpose of Models—The Norges Bank Experience,” Norges Bank, Staff Memo, No. 06.
- [14] Hansen, Lars Peter, 1982, “Large Sample Properties of Generalized Method of Moments Estimators,” *Econometrica*, 50, 1029–1054.
- [15] Parke, William R., 1982, “An Algorithm for FIML and 3SLS Estimation of Large Nonlinear Models,” *Econometrica*, 50, 81-96.