Forecasting U.S. Inflation by Bayesian Model Averaging

Jonathan H. Wright*

First Version: September 2003
This version: May 2004

Abstract: Recent empirical work has considered the prediction of inflation by combining the information in a large number of time series. One such method that has been found to give consistently good results consists of simple equal weighted averaging of the forecasts over a large number of different models, each of which is a linear regression model that relates inflation to a single predictor and a lagged dependent variable. In this paper, I consider using Bayesian Model Averaging for pseudo out-of-sample prediction of US inflation, and find that it gives more accurate forecasts than simple equal weighted averaging. This superior performance is consistent across subsamples and inflation measures. Meanwhile, both methods substantially outperform a naive time series benchmark of predicting inflation by an autoregression.

Keywords: Shrinkage, Phillips curve, model uncertainty, forecasting, inflation.
JEL Classification: C32, C53, E31, E37.

* International Finance Division, Board of Governors of the Federal Reserve System, Washington DC 20551. I am grateful to Ben Bernanke, David Bowman, Jon Faust, Lutz Kilian, Mike McCracken, Matt Pritsker and Pietro Veronesi for helpful comments, to Jeremy Rudd for providing me with data used in this paper, and to Sergey Chernenko for excellent research assistance.

1. Introduction.

Forecasting inflation is clearly of critical importance to the conduct of monetary policy, regardless of whether or not the central bank has a numerical inflation target. A simple Phillips curve, which uses a single measure of economic slack such as unemployment to predict future inflation, is probably the most common basis of inflation forecasting. The usefulness of the Phillips curve as a means of predicting inflation has however been questioned by several authors. For example, Atkeson and Ohanian (2001) found that Phillips curve based forecasts of inflation give larger out-of-sample prediction errors than a simple random walk forecast of inflation, although this specific result is very sensitive to the sample period and to the choice of inflation measure (Sims (2002)). Cecchetti, Chu and Steindel (2000) consider inflation prediction with individual indicators, including unemployment, and argue that none of these gives reliable inflation forecasts. Stock and Watson (2001, 2002a) consider prediction of inflation in each of the G7 countries using a large number of possible models. Each model has a single predictor (plus lagged inflation). They find that most of the models they consider give larger out-of-sample root mean square prediction error than a simple naive time series forecast based on fitting an autoregression to inflation. When a model *does* have predictive power relative to the naive time series forecast, this tends to be unstable. That is, the model that has good predictive power in one subperiod has little or no propensity to have good predictive power in another subperiod.

In recent years, researchers have however made substantial progress in forecasting inflation using large datasets (i.e. a large number of predictive variables), but where the information in these different variables is combined in a judicious way that avoids the estimation of a large number of unrestricted parameters. Bayesian VARs have been found to be useful in forecasting: these often use many time series, but impose a prior that many of the coefficients in the VAR are close to zero. Approaches in which the researcher estimates a small number of factors from a large dataset and forecasts using these estimated factors have also been shown to be capable of superior predictive performance (Stock and Watson (1999, 2002b) and Bernanke and Boivin (2003)). Stock and Watson (2001, 2002a) however argue that the best predictive performance is obtained by constructing forecasts from a very large number of models and simply averaging these forecasts. Stock and Watson report that this gives the best predictive performance of international inflation (and also output growth), and that this is remarkably consistent across subperiods and across countries. Although the basic idea that forecast combination outperforms any individual forecast is part of the folklore of economic forecasting, going back to Bates and Granger (1969), Stock and Watson underscore how consistent this is across time periods and variables being forecast. It is of course crucial to the result that the researcher just average the forecasts (or take a median or trimmed mean). It is in particular tempting to run a forecast evaluation regression in which the weights on the different forecasts are estimated as free parameters. While this leads to a better in-sample fit, it gives less good out-of-sample prediction.[1]

Stock and Watson (2001, 2002a) do not offer a definitive explanation for *why* simple averaging of forecasts does so well, but the finding is sufficiently strong and

[1] Better out-of-sample predictive power is obtained if the weights in the forecast evaluation equation are instead estimated by ridge regression (Chan, Stock and Watson (1998), Stock and Watson (1999)). Ridge regression is a shrinkage technique, so this is another example of how methods that avoid the estimation of a large number of unrestricted parameters give better forecasts.

general that forecasters ought to pay attention to this result, even without necessarily understanding exactly what is so effective about this particular form of shrinkage.

The result that equal weighted averaging gives the best forecasts is however odd. Indeed it cannot be correct as a general principal that equal weighted averaging of forecasts is always optimal. For example, one can always just average the first two forecasts, call that a new forecast, and throw that back in the set of forecasts being considered. If equal weighting is always best, then this new forecast will get an equal weight. But this changes the weights on the original forecasts, which are no longer all equal. While it is easy to say this, it is much harder to come up with a concrete alternative forecasting strategy that actually does better than simple averaging in terms of out-of-sample prediction of inflation. That is the goal of this paper.

This paper considers the prediction of US inflation by Bayesian Model Averaging, a technique which was not considered by Stock and Watson (2001, 2002a). Bayesian Model Averaging has been developed mainly, but not exclusively, by statisticians as opposed to econometricians. The idea is to consider prediction when the researcher does not know the true model, but has several candidate models. A forecast can be constructed putting weights on the predictions from each model. If these weights are all equal, then this is simple forecast averaging. The researcher can however start from the prior that all the models are equally good, but then estimate the posterior probabilities of the models, which can be used as weights instead.

The contribution of this paper is to argue that Bayesian Model Averaging generally does better than simple equal weighted model averaging for predicting US inflation. The result is remarkably consistent across measures of inflation, and across different time periods. Equal weighted model averaging substantially outperforms a naive time series forecast, but Bayesian Model Averaging does better again. Since equal weighted model averaging has been found in the literature to give good forecasting performance, and is found by Stock and Watson (2001, 2002a) to give the best inflation forecasts of all the methods that they consider, I conclude that Bayesian Model Averaging goes straight to the top of the class, or at least should be taken very seriously in the toolkit of inflation forecasters.

One does not have to be a subjectivist Bayesian to believe in the usefulness of Bayesian Model Averaging, or of Bayesian shrinkage techniques more generally. A frequentist econometrician can interpret these methods as pragmatic smoothing devices that can be useful for out-of-sample forecasting.

The plan for the remainder of the paper is as follows. In section 2, I shall describe the idea of Bayesian Model Averaging. The out-of-sample inflation prediction exercise is described in section 3. Section 4 concludes.

2. Bayesian Model Averaging

The idea of Bayesian Model Averaging was set out by Leamer (1978), but has recently received a lot of attention in the statistics literature, including in particular Raftery, Madigan and Hoeting (1997), Hoeting, Madigan, Raftery and Volinsky (1999) and Chipman, George and McCulloch (2001). It has also been used in a number of econometric applications, including output growth forecasting (Min and Zellner (1993), Koop and Potter (2003)), cross-country growth regressions (Doppelhofer, Miller and Sala-i-Martin (2000) and Fernandez, Ley and Steel (2001)) and stock return prediction

(Avramov (2002) and Cremers (2002)). Avarmov and Cremers both report improved pseudo-out-of-sample predictive performance from Bayesian model averaging.

Consider a set of n models $M_1,...M_n$. The ith model is indexed by a parameter vector θ - this is a different parameter vector for each model, but for compactness of notation I do not explicitly subscript θ by i. The researcher knows that one of these models is the true model, but does not know which one.[2] The researcher has prior beliefs about the probability that the ith model is the true model which we write as $P(M_i)$, observes data D, and updates her beliefs to compute the posterior probability that the ith model is the true model:

$$P(M_i \mid D) = \frac{P(D \mid M_i)P(M_i)}{\Sigma_{j=1}^{n} P(D \mid M_j)P(M_j)} \tag{1}$$

where

$$P(D \mid M_i) = \int P(D \mid \theta, M_i)P(\theta \mid M_i)d\theta$$

is the marginal likelihood of the ith model, $P(\theta \mid M_i)$ is the prior density of the parameter vector in this model and $P(D \mid \theta, M_i)$ is the likelihood. Each model implies a forecast density $f_1,...f_n$. If we knew which model was the true model, we would pick the associated forecast density. In the presence of model uncertainty, our forecast density is

$$f^* = \Sigma_{i=1}^{n} P(M_i \mid D) f_i$$

Likewise, each model implies a point forecast. In the presence of model uncertainty, our point forecast weights each of these forecasts by the posterior for the model.[3] This is all

[2] The assumption that one of the models is true is of course unrealistic, though it may be a useful fiction for getting good forecasting results. Recent theoretical work has considered Bayesian Model Averaging when none of the models is in fact true (see Bernardo and Smith (1994) and Key, Perrichi and Smith (1998)).

[3] This is the point forecast that minimizes mean square error. Likewise, the density forecast f^* is the best forecast evaluated by the logarithmic scoring rule (Raftery, Madigan and Hoeting (1997)).

there is to Bayesian Model Averaging. The researcher needs only specify the set of models, the model priors, $P(M_i)$, and the parameter priors, $P(\theta | M_i)$. The rest is just computation.

The models do not have to be linear regression models, but I shall henceforth assume that they are. The ith model then specifies that

$$y = X\beta + \varepsilon$$

where y is a time series that the researcher is trying to forecast (such as inflation), X is a matrix of predictors, β is a px1 parameter vector, $\varepsilon = (\varepsilon_1, ...\varepsilon_T)'$ is the disturbance vector and T is the sample size and the disturbances are iid with mean zero and variance σ^2. I shall define the models in the context of the empirical application below.

For the model priors, I shall assume that all models are equally likely, so that $P(M_i) = \frac{1}{n}$. For the parameter priors, I shall take the natural conjugate g-prior specification for β (Zellner (1986)), so that the prior for β conditional on σ^2 is $N(\overline{\beta}, \phi\sigma^2 (X'X)^{-1})$ where $\overline{\beta}$ is some prior mean. For σ^2, I assume the improper prior that is proportional to $1/\sigma^2$. This is a standard choice of the prior for the error variance, that was made by Fernandez, Ley and Steel (2001) and many others. As shown in the Technical Appendix, one can then calculate the required likelihood of the model

$$P(D | M_i) = \frac{1}{\sqrt{2}} \frac{\Gamma(T/2)}{\pi^{T/2}} (1+\phi)^{-p/2} S^{-(T+1)}$$

where

$$S^2 = Y'Y - Y'X(X'X)^{-1} X'Y \frac{\phi}{1+\phi}$$

Combining the model priors and parameter priors, equation (1) can then be evaluated for each model giving the posterior probability of each model and hence the weights to be assigned to forecast.

The prior for β is centered around zero and so within each model the parameter is shrunken towards zero, which corresponds to no predictability. The extent of this shrinkage is governed by ϕ. A smaller value of ϕ means more shrinkage, and makes the prior more informative, but this may help in out-of-sample forecasting. Researchers often try to make the prior as uninformative as possible (corresponding to a high value of ϕ), but at least in the inflation forecasting problem considered in this paper, a more informative prior turns out to give better predictive performance.

One way of thinking about the role of ϕ is that it controls the relative weight of the data and our prior beliefs in computing the posterior probabilities of different models. If ϕ=0, then $P(D|M_i)$ is equal for all models and so the posterior probability of each model being true is equal to the prior probability. The larger is ϕ, the more we are willing to move away from the model priors in response to what we observe in the data.

3. Application to U.S. Inflation Forecasting

The application I consider is to forecasting U.S. inflation. Following Stock and Watson (2001, 2002a), each model that will be used for forecasting is of the form

$$\pi_{t,t+h} = \alpha + \gamma Z_t + \rho \pi_{t-h,t} + \varepsilon_t \tag{2}$$

where $\pi_{t,t+h}$ denotes the inflation rate from time t to time t+h, h is the forecasting horizon, Z_t is a scalar predictor, and ε_t is the error term. Each model has a different scalar predictor and different parameters but I do not explicitly subscript these to denote their dependence on the model. The forecasts will be compared with the naive time series model in which inflation is a simple autoregression

$$\pi_{t,t+h} = \alpha + \rho \pi_{t-h,t} + \varepsilon_t \tag{3}$$

The data I consider are quarterly from 1960Q1-2003Q2. Four measures of inflation are used: CPI, Core CPI, the GDP deflator and the PCE deflator. For the predictor Z_t in equation (2), I consider a total of 93 possible variables, as listed in the Appendix. All of the variables are available for the whole sample period, yielding a balanced panel. This gives a large number of alternative measures of economic slack and several asset prices. The predictors used are similar to those considered by Stock and Watson (1999, 2001, 2002a).

I consider pseudo out-of-sample prediction of inflation using equal weighted averaging across the models defined by the different predictors in equation (2). Concretely, the equal weighted averaging h-step ahead forecast is given by

$$\frac{1}{n}\Sigma_{i=1}^{n}\hat{\alpha}_t + \hat{\gamma}_t Z_t + \hat{\rho}_t \pi_{t-h,t} \tag{4}$$

where $\hat{\alpha}_t$, $\hat{\gamma}_t$ and $\hat{\rho}_t$ are the OLS estimates of the parameters of equation (2) obtained using only data from date t and earlier. Meanwhile, the Bayesian Model Averaging forecast is given by

$$\Sigma_{i=1}^{n} P_t(M_i \mid D)(\hat{\alpha}_t + \hat{\gamma}_t Z_t + \hat{\rho}_t \pi_{t-h,t}) \tag{5}$$

where $P_t(M_i \mid D)$ is the posterior probability that model i is the true model, computed using only data from date t and earlier. Notice that in equation (5), the Bayesian approach is only used to determine the posterior probabilities for each model. Within each individual model, the forecast is constructed using OLS estimates of α, γ and ρ, not the posterior means. This means that the Bayesian Model Averaging forecast in (5) reduces to the equal-weighted averaging forecast in (4) when the posterior model probabilities are all equal.

These posterior model probabilities are calculated as described in section 2 above except for one issue: the forecasts are overlapping h-step ahead forecasts and so forecast errors less than h periods apart are likely to be serially correlated. Meanwhile my derivation of the model likelihood only applies with serially uncorrelated errors. I circumvent this problem by the simple device of using only every hth observation in computing the posterior model probabilities $P(M_i \mid D)$, although all of the observations are used in calculating the OLS estimates of the parameters α, γ and ρ in each model.

The only parameters I need to specify are ϕ and the prior mean $\bar{\beta}$. I set the prior mean of γ to zero meaning that my prior is that Z_t is not useful for prediction. On the other hand zero does not seem like a good choice for the prior mean for α and ρ. So I use the OLS estimates from estimating equation (3) on CPI data over the pre-sample period 1913Q1-1959Q4 as the prior mean for α and ρ.

As a benchmark, I also consider the naive time series forecast

$$\tilde{\alpha}_t + \tilde{\rho}_t \pi_{t-h,t} \tag{6}$$

where $\tilde{\alpha}_t$ and $\tilde{\rho}_t$ are the OLS estimates of the parameters of equation (3) obtained using only data from date t and earlier.

For each quarter from 1971Q1 on, I computed the out-of-sample mean square prediction error of the equal weighted averaging forecast in equation (4) and the Bayesian Model Averaging forecast in equation (5), both relative to the out-of-sample mean square prediction error from the naive forecast in (6). For the Bayesian Model Averaging, the forecasts are computed as described in the previous section, with the prior probabilities for all models being equal, for $\phi = 100, 5, 2, 1, 0.5$. A relative mean square prediction error less than one means that the forecast outperforms the naive time series forecast.

To investigate the possibility of forecast instability (forecasts working well in one subperiod but not another), I also computed these relative out-of-sample mean square prediction errors in two subsamples: for 1971Q1-1986Q4 and for 1987Q1-2003Q2.

The results are reported in Tables 1, 2, 3 and 4, for prediction of CPI inflation, CPI core inflation, GDP deflator inflation and PCE deflator inflation, respectively. These results use all the predictors enumerated in the appendix for a total of 93 models. It is also possible to compute the relative out-of-sample mean square prediction errors using only asset prices as predictors, which is useful because these are available in real-time and are not subject to revision, although there are only 23 such models. The results, using asset prices only are reported in Tables 5, 6, 7 and 8 for prediction of CPI inflation, CPI core inflation, GDP deflator inflation and PCE deflator inflation, respectively. The key results from these Tables are as follows:

1. Both the equal weighted forecasts and the Bayesian Model Averaging forecasts nearly always have a relative mean square prediction error below one, indicating that both forecasts are consistently outperforming the naive time series benchmark. The improvement relative to the benchmark is typically substantial.

2. Bayesian Model Averaging nearly always outperforms equal weighted model averaging (as well as the naive time series benchmark) unless the value of φ is large (low shrinkage). This is a consistent result across for all four inflation measures, for both the 1971Q1-1986Q4 and 1987Q1-2003Q2 subperiods[4]. The margin by which Bayesian Model Averaging outperforms equal weighted model is often substantial. As an example, forecasting CPI inflation at a 4-quarter ahead horizon, simple equal weighted model averaging gives a 20% reduction in mean square prediction error relative to the naive time series benchmark. but Bayesian Model Averaging with φ=5 gives a 25.5% reduction relative to this benchmark.

3. Bayesian Model Averaging is very similar to equal weighted model averaging for small values of φ (high shrinkage). Raising φ usually improves the forecasting power of Bayesian Model Averaging up to a point. But a sufficiently large value of φ usually eventually leads to less good forecasting performance.

[4] Bayesian Model Averaging outperforms equal weighted averaging and the naive time series benchmark in both subperiods, but the margin of improvement is greater in the second subperiod, which could be the result of having a longer span of data on which to base the forecasts.

4. The margin by which both equal weighted averaging and Bayesian Model Averaging outperform the naive time series benchmark grows with the horizon, at least up to a horizon of about 4-6 quarters.

5. The margin by which both equal weighted averaging and Bayesian Model Averaging outperform the naive time series benchmark is greater when using all predictors than when using asset prices only.

6. These results are not qualitatively different across the four alternative inflation measures.

A natural question to ask is what fraction of the time the Bayesian Model Averaging gives a forecast that turns out to be better than that from equal weighted model averaging, in the sense of being closer to the subsequent realized inflation rate. In Table 9, I report the proportion of times that Bayesian Model Averaging gives the more accurate out-of-sample forecast, over the whole period 1971Q1-2003Q2 for all four inflation measures when using all possible predictors. In Table 10, I report the proportion of times that Bayesian Model Averaging gives the more accurate out-of-sample forecast using asset prices alone. The elements of Tables 9 and 10 are mostly above 0.5, substantially so at longer horizons, indicating that Bayesian Model Averaging is usually more accurate than equal weighted averaging. This result is consistent across different inflation measures, and applies both when using all predictors and when using asset prices alone.

It is in fact easy to conduct a statistical test of the null hypothesis that Bayesian Model Averaging and equal weighted averaging are equally likely to give forecasts that are closer to the actual inflation rate (i.e. that the population probability being estimated in Tables 9 and 10 is in fact 0.5). Under this null hypothesis, the proportion of times that Bayesian Model Averaging gives a more accurate forecast is approximately $N(0.5, \frac{h}{4T^*})$ where T^* denotes the number of out-of-sample observations. Using this approximation, I do a two-tailed 10% test of the null hypothesis and denote the cases for which the hypothesis is rejected with an asterisk in Tables 9 and 10. The null is generally rejected for all the inflation series, at least at the horizons of 3 quarters and above and with φ of 5 or less.

This test may not be very powerful, but since it rejects the null, lack of power is not relevant. Meanwhile, it affords a very simple way of testing the significance of the improvement which Bayesian Model Averaging offers over simple equal weighted averaging. Significance testing by a standard bootstrap might be quite tricky. One could proceed by fitting a vector autoregression to inflation and all of the predictors and use this to generate bootstrap samples of inflation and the different predictors. But the number of predictors is so large that this is unlikely to work well.

Each of the models I have considered consists of a single predictor (plus a constant and a lagged dependent variable). This keeps the exercise close to that considered by Stock and Watson (2001, 2002a). However, it is nonstandard in Bayesian Model Averaging methodology. A more standard Bayesian Model Averaging approach would use all possible permutations of predictors, generating a large number of candidate models. The computational burden of such an approach is considerable. If λ is the number of predictor variables, there will be a total of 2^{λ} models. Given my dataset of inflation predictors, it is impossible to evaluate the posterior probability for all of these models. Madigan and York (1995) and Geweke (1996) discuss simulation based methods for implementing Bayesian Model Averaging that are practical with an extremely large number of models. I do not however implement this.

In a different application of Bayesian Model Averaging to exchange rate forecasting (Wright (2003)), I consider exchange rate prediction using both (i) all possible single-predictor models and (ii) models consisting of all possible permutations of these predictors. I find that considering all possible permutations of predictors, although computationally harder to deal with, gives forecasts that are no better and perhaps even a little worse.

4. Conclusion and Future Research

A theme of much recent empirical work on inflation forecasting is that the judicious pooling of information from a large number of indicators provides the best approach to predicting inflation. One method that has been particularly promising is to simply average the forecasts from a large number of models, each of which has a single predictor variable. In this paper, I have considered instead using Bayesian Model Averaging for U.S. inflation forecasting and found that it fairly consistently outperforms this equal weighted forecast averaging. This result is consistent across different subperiods and across different inflation measures.

Equal weighted forecast averaging is a benchmark that has been found to provide good forecasts of inflation (and of many other variables). Stock and Watson (2001, 2002a) indeed argue that it is the best method for predicting inflation in the US and other G-7 countries among a wide range of forecasting methods that they consider. So, since Bayesian Model Averaging does better than equal weighted averaging in predicting US inflation, it should be taken very seriously as a method for forecasting inflation. I do not mean to claim by this that Bayesian Model Averaging as I have implemented it in this paper is necessarily the best thing that a researcher could ever do. It may be possible to get still better forecasts by incorporating nonlinear models in the exercise, by incorporating Greenbook and private sector survey forecasts of inflation, by considering models with more than one predictor, or by using different shrinkage techniques.

The researcher using Bayesian Model Averaging has to select some prior hyperparameters, and the promising results obtain for values of these hyperparameters that imply considerable shrinkage. One approach would be to select prior hyperparameters at each point in time that maximize the historical pseudo-out-of-sample forecasting performance.[5] This kind of adaptive estimation strategy seems appropriate if one views Bayesian Model Averaging simply as a pragmatic forecasting device, as I do. A purist Bayesian would however reject this approach because it gets the conditioning wrong by allowing the prior to depend on the data.

Inflation forecasting is of direct interest to those who set monetary policy, but also of interest to researchers in empirical macroeconomics. For example, researchers are often interested in estimating a forward-looking Taylor rule of the form

---

[5] This is similar in spirit to the empirical Bayes methodology, considered in the context of model selection by George and Foster (2000). The empirical Bayes approach selects prior hyperparameters so as to maximise the marginal likelihood of these hyperparameters.

$$r_t = \mu_0 + \mu_1 r_{t-1} + \mu_\pi E_t \pi_{t,t+4} + \mu_g g_t + v_t \quad (7)$$

where r_t is the short-run interest rate, E_t denotes the expectations operator at time t and g_t is the output gap (see, for example, Clarida, Gali and Gertler (2000)). The expectation of future inflation is not observable. The standard empirical strategy is to rewrite (7) as

$$r_t = \mu_0 + \mu_1 r_{t-1} + \mu_\pi \pi_{t,t+4} + \mu_g g_t + v_t^*$$

where $v_t^* = v_t - \mu_\pi(\pi_{t,t+4} - E_t \pi_{t,t+4})$ and then to estimate the coefficients from this regression. The future realized inflation in this regression will however clearly be endogenous and so researchers look for instruments that will be correlated with $\pi_{t,t+4}$, but uncorrelated with v_t^*. They are looking for instruments that have predictive power for future inflation, but that are in the information set at time t. The more predictive power for future inflation a variable has, the stronger an instrument it is. The inflation forecast provided by Bayesian Model Averaging is therefore likely to be an excellent instrument allowing equation (7) or similar related equations to be estimated reliably.

One caveat with the results in this paper is that although I have shown that Bayesian Model Averaging would have worked well in out-of-sample prediction over the last 30 years, I cannot be confident that it will continue to work well in the future. Inflation is now at a lower level than it has been for the last 30 years and there is at least a possibility that there is some substantive nonlinearity leading inflation to behave differently at very low levels.[6] This caveat of course applies to equal weighted averaging and to other inflation forecasting methods too. Indeed one of the strengths of the

[6] It would be possible to use Bayesian Model Averaging to obtain a density forecast for inflation, and to compute a probability of deflation arising over a given horizon. While such an exercise would clearly be of contemporary policy relevance, the concern that inflation and the variance of inflation might behave differently at very low levels would perhaps be especially pertinent here.

Bayesian Model Averaging method is that I have found its superior performance relative to both a naive time series model and to equal weighted forecast averaging to be stable in the sense that it applies over at least two different subperiods.

Appendix: List of Predictors

Predictor	Transformation(s)
Industrial Production Total	FDL
Industrial Production Total Products	FDL
Industrial Production Final Products	FDL
Industrial Production Consumer Goods	FDL
Industrial Production Consumer Durables	FDL
Industrial Production Consumer Nondurables	FDL
Industrial Production Business Equipment	FDL
Industrial Production Intermediate Materials	FDL
Industrial Production Nondurable Goods Materials	FDL
Industrial Production Manufacturing	FDL
Purchasing Managers' Index	Level
Capacity Utilization	Level
NAPM Production Index	Level
Personal Income Less Transfers	FDL
Help Wanted Index	FDL
Ratio of Help Wanted Index to Number Unemployed	Log
Number of Employed	FDL
Number of Employed, Nonagricultural	FDL
Unemployment Rate, 16+	Level
Unemployment Rate,	Level
Unemployed less than 5 weeks	Level
Unemployed 5-14 weeks	Level
Unemployed more that 15 weeks	Level
Unemployed 15-26 weeks	Level
Nonfarm Payroll Employment	FDL
Private Nonfarm Payroll Employment	FDL
Nonfarm Payrolls, Goods Producing	FDL
Nonfarm Payrolls, Goods Producing Production Workers	FDL
Nonfarm Payrolls, Construction	FDL
Nonfarm Payrolls, Manufacturing	FDL
Nonfarm Payrolls, Durable Goods Manufacturing	FDL
Nonfarm Payrolls, Nondurable Goods Manufacturing	FDL
Nonfarm Payrolls, Service Producing	FDL
Nonfarm Payrolls, Wholesale and Retail Trade	FDL
Nonfarm Payrolls, Finance, Insurance & Real Estate	FDL
Nonfarm Payrolls, Services	FDL
Nonfarm Payrolls, Government	FDL
Average Weekly Hours	Level
Average Weekly Hours, Overtime	Level
NAPM Employment Index	Level
Aggregate Hours, Manufacturing	FDL
Aggregate Hours, Mining	FDL
Aggregate Hours, Construction	FDL
Real Personal Consumption, Total	FDL
Real Personal Consumption, Durables	FDL
Real Personal Consumption, Nondurables	FDL
Real Personal Consumption, Services	FDL
Real Personal Consumption, New Cars	FDL
Real Personal Consumption, Retail	FDL
Housing Starts, Total	Log

Housing Starts, Northeast	Log
Housing Starts, Midwest	Log
Housing Starts, South	Log
Housing Starts, West	Log
Mobile Homes, Manufacturers’ Shipments	Log
NAPM Inventories Index	Level
NAPM New Orders Index	Level
NAPM Supplier Deliveries Index	Level
New Orders, Consumer Goods and Materials	FDL
New Orders, Nondefense Capital Goods	FDL
Money Stock, M1	SDL
Money Stock, M2	SDL
Money Stock, M3	SDL
Money Stock, Real M2	SDL
Monetary Base	SDL
Total Reserves	SDL
NAPM Prices Paid	Level
Average Hourly Earnings, Construction	SDL
Average Hourly Earnings, Manufacturing	SDL
Michigan Index of Consumer Sentiment	Level
Asset Prices	
Prime Rate	FD, Spread over Fed Funds
3 month T bill rate	FD, Spread over Fed Funds
6 month T bill rate	FD, Spread over Fed Funds
1 year Treasury constant maturity yield	FD, Spread over Fed Funds
5 year Treasury constant maturity yield	FD, Spread over Fed Funds
10 year Treasury constant maturity yield	FD, Spread over Fed Funds
Moody’s AAA Corporate Yield	FD, Spread over Fed Funds
Moody’s BAA Corporate Yield	FD, Spread over Fed Funds
Fed Funds Rate	FD
NYSE Composite Index	FDL
S&P Composite Index	FDL
S&P Dividend Yield	Level
S&P Price Earnings Ratio	Level
Oil Price	FDL
Gold Price	FDL

Notes: FD means first differences, FDL means first differences of the logs, SDL means second differences of the logs.

Technical Appendix: Derivation of Model Likelihood

Setting $\overline{\beta}$ to zero without loss of generality, the prior for β in model M_i conditional on σ is $N(0,\sigma^2\phi(X'X)^{-1})$, i.e.

$$p(\beta\,|\,\sigma)=\frac{1}{|2\pi\sigma^2\phi(X'X)^{-1}|^{1/2}}\exp(-\frac{1}{2}\beta'(\phi\sigma^2(X'X)^{-1})^{-1}\beta)=\frac{|X'X|^{1/2}}{(2\pi)^{p/2}\phi^{p/2}\sigma^p}\exp(-\frac{1}{2\phi\sigma^2}\beta'X'X\beta)$$

and the prior for σ is $p(\sigma)=\sigma^{-2}$. The Gaussian likelihood function is

$$p(D\,|\,\theta,M_i)=\frac{1}{(2\pi)^{T/2}\sigma^T}\exp(-\frac{1}{2\sigma^2}(y-X\beta)'(y-X\beta))$$

$$=\frac{1}{(2\pi)^{T/2}\sigma^T}\exp(-\frac{1}{2\sigma^2}[Q+(\beta-\hat{\beta})'X'X(\beta-\hat{\beta})])$$

where $Q=(y-X\hat{\beta})'(y-X\hat{\beta})$ and $\hat{\beta}=(X'X)^{-1}X'y$, using the fact that $(y-X\hat{\beta})'X=0$. Integrating this over β and σ,

$$p(D\,|\,M_i)=\iint p(D\,|\,\theta,M_i)\,p(\beta\,|\,\sigma)\,p(\sigma)\,d\beta d\sigma$$

$$=\frac{1}{(2\pi)^{T/2}}\frac{|X'X|^{1/2}}{(2\pi)^{p/2}\phi^{p/2}}\iint\frac{1}{\sigma^T}\frac{1}{\sigma^p}\frac{1}{\sigma^2}\exp(-\frac{1}{2\sigma^2}[Q+(\beta-\hat{\beta})'X'X(\beta-\hat{\beta})])\exp(-\frac{1}{2\phi\sigma^2}\beta'X'X\beta)d\beta d\sigma$$

$$=\frac{|X'X|^{1/2}}{(2\pi)^{(T+p)/2}\phi^{p/2}}\iint\frac{1}{\sigma^{T+p+2}}\exp(-\frac{1}{2\sigma^2}[Q+(\beta-\hat{\beta})X'X(\beta-\hat{\beta})+\frac{1}{\phi}\beta'X'X\beta])d\beta d\sigma$$

But ,

$$(\beta-\hat{\beta})'X'X(\beta-\hat{\beta})+\frac{1}{\phi}\beta'X'X\beta=(1+\frac{1}{\phi})\beta'X'X\beta+\hat{\beta}'X'X\hat{\beta}-2\hat{\beta}'X'X\beta$$

$$=(\frac{\phi+1}{\phi})\beta'X'X\beta-2\hat{\beta}'X'X\beta+\frac{\phi}{\phi+1}\hat{\beta}'X'X\hat{\beta}+\frac{1}{\phi+1}\hat{\beta}'X'X\hat{\beta}$$

$$=(\frac{\phi+1}{\phi})(\beta-\frac{\phi\hat{\beta}}{\phi+1})'X'X(\beta-\frac{\phi\hat{\beta}}{\phi+1})+\frac{1}{\phi+1}\hat{\beta}'X'X\hat{\beta}$$

So letting $S^2=Q+\frac{1}{\phi+1}\hat{\beta}'X'X\hat{\beta}$, $p(D\,|\,M_i)$ can be written as

$$\frac{|X'X|^{1/2}}{(2\pi)^{(T+p)/2}\phi^{p/2}}\iint\frac{1}{\sigma^{T+p+2}}\exp(-\frac{S^2}{2\sigma^2})\exp(-\frac{1}{2\sigma^2}[(\frac{\phi+1}{\phi})(\beta-\frac{\phi\hat{\beta}}{\phi+1})'X'X(\beta-\frac{\phi\hat{\beta}}{\phi+1})])d\beta d\sigma$$

$$=\frac{|X'X|^{1/2}}{(2\pi)^{(T+p)/2}\phi^{p/2}}\int\frac{1}{\sigma^{T+p+2}}\exp(-\frac{S^2}{2\sigma^2})\int\exp(-\frac{1}{2\sigma^2}[(\frac{\phi+1}{\phi})(\beta-\frac{\phi\hat{\beta}}{\phi+1})'X'X(\beta-\frac{\phi\hat{\beta}}{\phi+1})])d\beta d\sigma$$

Using the properties of the multivariate normal pdf,

$$\int\exp(-\frac{1}{2\sigma^2}[(\frac{\phi+1}{\phi})(\beta-\frac{\phi\hat{\beta}}{\phi+1})'X'X(\beta-\frac{\phi\hat{\beta}}{\phi+1})])d\beta=\int\exp(-\frac{1}{2\sigma^2}\frac{\phi+1}{\phi}\beta'X'X\beta)d\beta$$

$$=(2\pi)^{p/2}\sigma^p(\frac{\phi}{\phi+1})^{p/2}|X'X|^{-1/2}$$

So

$$p(D\,|\,M_i) = \frac{|X'X|^{1/2}}{(2\pi)^{(T+p)/2}} \frac{(2\pi)^{p/2}}{|X'X|^{1/2}\,\phi^{p/2}} (\frac{\phi}{\phi+1})^{p/2} \int \frac{\sigma^p}{\sigma^{T+p+2}} \exp(-\frac{S^2}{2\sigma^2}) d\sigma$$

$$= \frac{1}{(2\pi)^{T/2}} (\frac{1}{\phi+1})^{p/2} \int \frac{1}{\sigma^{T+2}} \exp(-\frac{S^2}{2\sigma^2}) d\sigma$$

For $a, b > 0$,

$\int_0^\infty \sigma^{-a} \exp(-2b\sigma^{-2}) d\sigma = 2^{(a-3)/2} b^{-(a-1)/2} \int_0^\infty \theta^{(a-3)/2} e^{-\theta} d\theta = 2^{(a-3)/2} b^{-(a-1)/2} \Gamma(\frac{a-1}{2})$, integrating using the change of variable $\theta = 2b\sigma^{-2}$. So

$$p(D\,|\,M_i) = \frac{1}{(2\pi)^{T/2}} (\frac{1}{\phi+1})^{p/2} 2^{(T-1)/2} S^{-(T+1)} \Gamma(\frac{T+1}{2}) = \frac{1}{\sqrt{2}} \frac{\Gamma(T/2)}{\pi^{T/2}} (1+\phi)^{-p/2} S^{-(T+1)}$$

where

$$S^2 = (Y - X\hat{\beta})'(Y - X\hat{\beta}) + \frac{1}{\phi+1} \hat{\beta}' X' X \hat{\beta} = Y'Y - \hat{\beta}' X' X \hat{\beta} + \frac{1}{\phi+1} \hat{\beta}' X' X \hat{\beta}$$

$$= Y'Y - \frac{\phi}{\phi+1} Y' X' (X'X)^{-1} X' Y$$

References

Atkeson, A. and L.E. Ohanian (2001): Are Phillips Curves Useful for Forecasting Inflation?, *Federal Reserve Bank of Minneapolis Quarterly Review,* 25, pp.2-11.

Avramov, D. (2002): Stock Return Predictability and Model Uncertainty, *Journal of Financial Economics*, 64, pp.423-458.

Bates, J.M. and C.W.J. Granger (1969): The Combination of Forecasts, *Operations Research Quarterly*, 20, pp.451-468.

Bernardo, J.M. and A.F.M. Smith (1994): *Bayesian Theory*, Wiley, New York.

Bernanke, B.S. and J. Boivin (2003): Monetary Policy in a Data-Rich Environment, *Journal of Monetary Economics*, 50, pp.525-546.

Clarida, R., J. Gali and M. Gertler (2000): Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory, *Quarterly Journal of Economics*, 115, pp.147-180.

Cecchetti, S., R. Chu and C. Steindel (2000): The Unreliability of Inflation Indicators, *Federal Reserve Bank of New York Current Issues in Economics and Finance*, 6, pp.1-6.

Chan, Y.L., J.H. Stock and M.W. Watson (1999): A Dynamic Factor Model for Forecast Combination, *Spanish Economic Review*, 1, pp.91-122.

Chipman, H., E.I. George and R.E. McCulloch (2001): The Practical Implementation of Bayesian Model Selection, mimeo.

Cremers, K.J.M. (2002): Stock Return Predictability: A Bayesian Model Selection Perspective*, Review of Financial Studies,* 15, pp.1223-1249.

Doppelhofer, G., R.I. Miller and X. Sala-i-Martin (2000): Determinants of Long-Term Growth: A Bayesian Averaging of Classical Estimates (BACE) Approach, NBER Working Paper 7750.

Fernandez, C., E. Ley and M.F.J. Steel (2001): Model Uncertainty in Cross-Country Growth Regressions, *Journal of Applied Econometrics,* 16, pp.563-576.

George. E.I. and D.P. Foster (2000): Calibration and Empirical Bayes Variable Selection, *Biometrika*, 87, pp.731-747.

Geweke, J. (1996): "Variable Selection and Model Comparison in Regression," in J.M. Bernardo, J.O. Berger, A.P. Dawid and A.F.M. Smith (eds.), *Bayesian Statistics Volume 5*, Oxford University Press, Oxford.

Hoeting, J.A., D. Madigan, A.E. Raftery and C.T. Volinsky (1999): Bayesian Model Averaging: A Tutorial, *Statistical Science*, 14, pp.382-417.

Key, J.T., L.R. Pericchi and A.F.M. Smith (1998): “Choosing Among Models when None of Them are True”, in W. Racugno (ed.), *Proceedings of the Workshop on Model Selection, Special Issue of Rassegna di Metodi Statistici ed Applicazioni*, Pitagore Editrice, Bologna.

Koop, G. and S. Potter (2003): Forecasting in Large Macroeconomic Panels Using Bayesian Model Averaging, Federal Reserve Bank of New York Staff Report 163.

Leamer, E.E. (1978): *Specification Searches*, Wiley, New York.

Madigan, D. and J. York (1995): Bayesian Graphical Models for Discrete Data, *International Statistical Review*, 63, pp.215-232.

Min, C. and A. Zellner (1993): Bayesian and Non-Bayesian Methods for Combining Models and Forecasts with Applications to Forecasting International Growth Rates, *Journal of Econometrics*, 56, pp.89-118.

Raftery, A.E., D. Madigan and J.A. Hoeting (1997): Bayesian Model Averaging for Linear Regression Models, *Journal of the American Statistical Association*, 92, pp.179-191.

Sims, C.A. (2002): The Role of Models and Probabilities in the Monetary Policy Process, *Brookings Papers on Economic Activity*, 2, pp.1-40.

Stock, J.H. and M.W. Watson (1999): Forecasting Inflation, *Journal of Monetary Economics*, 44, pp.293-335.

Stock, J.H. and M.W. Watson (2001): Forecasting Output and Inflation: The Role of Asset Prices, mimeo.

Stock, J.H. and M.W. Watson (2002a): Combination Forecasts of Output Growth in a Seven Country Dataset, mimeo.

Stock, J.H. and M.W. Watson (2002b): Forecasting Using Principal Components from a Large Number of Predictors, *Journal of the American Statistical Association*, 97, pp.1167-1179.

Wright, J.H. (2003): Bayesian Model Averaging and Exchange Rate Forecasts, International Finance Discussion Paper, 779.

Zellner, A. (1971): *An Introduction to Bayesian Inference in Econometrics*, Wiley, New York.

Zellner, A. (1986): "On Assessing Prior Distributions and Bayesian Regression Analysis with g-Prior Distributions," in P.K. Goel and A. Zellner (eds.), *Bayesian Inference and Decision Techniques: Essays in Honour of Bruno de Finetti*, North Holland, Amsterdam.

Table 1: Out-of-Sample Mean Square Error of Averaged Forecasts of CPI (relative to naive time series benchmark)

Horizon	Bayesian Model Averaging					Simple Avg.
(quarters)	φ=100	φ=5	φ=2	φ=1	φ=0.5	
			1971Q1-2003Q2			
1	**0.957**	**0.947**	**0.940**	**0.937**	**0.939**	**0.947**
2	**0.883**	**0.857**	**0.853**	**0.859**	**0.868**	**0.882**
3	**0.748**	**0.735**	**0.744**	**0.763**	**0.784**	**0.822**
4	**0.870**	**0.767**	**0.740**	**0.747**	**0.765**	**0.802**
5	**0.899**	**0.782**	**0.763**	**0.768**	**0.780**	**0.806**
6	**0.828**	**0.769**	**0.771**	**0.784**	**0.797**	**0.818**
7	**0.858**	**0.820**	**0.822**	**0.826**	**0.829**	**0.833**
8	**0.849**	**0.828**	**0.835**	**0.841**	**0.844**	**0.848**
			1971Q1-1986Q4			
1	**0.954**	**0.942**	**0.933**	**0.929**	**0.930**	**0.936**
2	**0.863**	**0.833**	**0.830**	**0.838**	**0.848**	**0.865**
3	**0.683**	**0.683**	**0.705**	**0.734**	**0.761**	**0.803**
4	**0.856**	**0.748**	**0.726**	**0.737**	**0.757**	**0.796**
5	**0.902**	**0.783**	**0.767**	**0.774**	**0.786**	**0.809**
6	**0.845**	**0.791**	**0.792**	**0.801**	**0.810**	**0.827**
7	**0.875**	**0.841**	**0.841**	**0.843**	**0.844**	**0.845**
8	**0.860**	**0.843**	**0.850**	**0.855**	**0.858**	**0.860**
			1987Q1-2003Q2			
1	**0.965**	**0.963**	**0.960**	**0.959**	**0.962**	**0.974**
2	**0.955**	**0.946**	**0.941**	**0.941**	**0.943**	**0.948**
3	**1.076**	**0.995**	**0.941**	**0.913**	**0.903**	**0.917**
4	**0.973**	**0.898**	**0.844**	**0.819**	**0.819**	**0.850**
5	**0.871**	**0.772**	**0.723**	**0.715**	**0.729**	**0.780**
6	**0.667**	**0.558**	**0.569**	**0.619**	**0.665**	**0.727**
7	**0.693**	**0.629**	**0.645**	**0.672**	**0.694**	**0.722**
8	**0.742**	**0.686**	**0.693**	**0.707**	**0.719**	**0.736**

Notes: Both Bayesian Model Averaging and simple equal-weighted averaging use 93 models corresponding to all predictors listed in the appendix. Elements in bold correspond to cases where Bayesian Model Averaging outperforms simple equal-weighted averaging for some ϕ. Elements in italics correspond to cases where simple equal-weighted averaging does better.

Table 2: Out-of-Sample Mean Square Error of Averaged Forecasts of Core CPI (relative to naive time series benchmark)

Horizon	Bayesian Model Averaging					Simple Avg.
(quarters)	φ=100	φ=5	φ=2	φ=1	φ=0.5	
			1971Q1-2003Q2			
1	*1.126*	*1.085*	*1.036*	*0.993*	*0.969*	*0.959*
2	**0.898**	**0.872**	**0.862**	**0.865**	**0.872**	**0.883**
3	**0.907**	**0.821**	**0.758**	**0.759**	**0.780**	**0.824**
4	**0.677**	**0.693**	**0.725**	**0.755**	**0.779**	**0.813**
5	**0.812**	**0.772**	**0.772**	**0.787**	**0.802**	**0.824**
6	**0.922**	**0.893**	**0.833**	**0.814**	**0.817**	**0.834**
7	**0.966**	**0.871**	**0.850**	**0.846**	**0.846**	**0.848**
8	**0.856**	**0.809**	**0.825**	**0.840**	**0.850**	**0.862**
			1971Q1-1986Q4			
1	*1.138*	*1.097*	*1.047*	*1.002*	*0.975*	*0.962*
2	**0.898**	**0.878**	**0.872**	**0.877**	**0.883**	**0.893**
3	**0.907**	**0.815**	**0.751**	**0.759**	**0.789**	**0.835**
4	**0.676**	**0.693**	**0.735**	**0.773**	**0.798**	**0.827**
5	**0.832**	**0.794**	**0.798**	**0.812**	**0.824**	**0.840**
6	**0.965**	**0.936**	**0.868**	**0.841**	**0.839**	**0.850**
7	*1.006*	*0.902*	*0.874*	*0.866*	*0.864*	*0.863*
8	**0.889**	**0.834**	**0.845**	**0.857**	**0.866**	**0.876**
			1987Q1-2003Q2			
1	**0.989**	**0.957**	**0.919**	**0.899**	**0.901**	**0.922**
2	**0.900**	**0.759**	**0.671**	**0.648**	**0.660**	**0.701**
3	**0.908**	**0.904**	**0.865**	**0.753**	**0.643**	**0.654**
4	**0.700**	**0.700**	**0.598**	**0.501**	**0.525**	**0.619**
5	**0.560**	**0.494**	**0.448**	**0.466**	**0.523**	**0.623**
6	**0.399**	**0.372**	**0.414**	**0.485**	**0.551**	**0.636**
7	**0.489**	**0.510**	**0.562**	**0.606**	**0.635**	**0.670**
8	**0.486**	**0.537**	**0.604**	**0.648**	**0.674**	**0.702**

Notes: See the notes to Table 1.

Table 3: Out-of-Sample Mean Square Error of Averaged Forecasts of GDP Deflator (relative to naive time series benchmark)

Horizon (quarters)	Bayesian Model Averaging φ=100	φ=5	φ=2	φ=1	φ=0.5	Simple Avg.
			1971Q1-2003Q2			
1	*1.026*	*1.007*	*0.993*	*0.985*	*0.982*	*0.982*
2	**0.970**	**0.947**	**0.940**	**0.942**	**0.944**	**0.948**
3	**1.047**	**0.950**	**0.880**	**0.866**	**0.869**	**0.878**
4	**0.714**	**0.724**	**0.753**	**0.781**	**0.803**	**0.832**
5	**0.711**	**0.754**	**0.769**	**0.782**	**0.793**	**0.812**
6	**0.921**	**0.794**	**0.759**	**0.771**	**0.786**	**0.805**
7	**0.867**	**0.775**	**0.784**	**0.793**	**0.800**	**0.807**
8	**0.721**	**0.746**	**0.776**	**0.793**	**0.802**	**0.811**
			1971Q1-1986Q4			
1	*1.054*	*1.027*	*1.005*	*0.991*	*0.985*	*0.983*
2	**0.951**	**0.948**	**0.950**	**0.951**	**0.951**	**0.951**
3	**1.045**	**0.961**	**0.895**	**0.880**	**0.880**	**0.884**
4	**0.640**	**0.694**	**0.760**	**0.800**	**0.822**	**0.842**
5	**0.715**	**0.775**	**0.795**	**0.807**	**0.815**	**0.826**
6	**0.979**	**0.854**	**0.801**	**0.801**	**0.809**	**0.820**
7	**0.942**	**0.828**	**0.822**	**0.821**	**0.821**	**0.822**
8	**0.762**	**0.789**	**0.809**	**0.818**	**0.822**	**0.826**
			1987Q1-2003Q2			
1	**0.951**	**0.955**	**0.963**	**0.971**	**0.976**	**0.980**
2	**1.049**	**0.941**	**0.896**	**0.902**	**0.917**	**0.937**
3	**1.058**	**0.877**	**0.777**	**0.774**	**0.800**	**0.839**
4	**1.358**	**0.979**	**0.693**	**0.612**	**0.643**	**0.737**
5	**0.680**	**0.578**	**0.538**	**0.558**	**0.602**	**0.690**
6	**0.423**	**0.285**	**0.400**	**0.514**	**0.592**	**0.679**
7	**0.269**	**0.356**	**0.484**	**0.574**	**0.629**	**0.686**
8	**0.413**	**0.420**	**0.531**	**0.607**	**0.652**	**0.698**

Notes: See the notes to Table 1.

Table 4: Out-of-Sample Mean Square Error of Averaged Forecasts of PCE Deflator (relative to naive time series benchmark)

Horizon (quarters)	Bayesian Model Averaging φ=100	φ=5	φ=2	φ=1	φ=0.5	Simple Avg.
			1971Q1-2003Q2			
1	*1.008*	*0.993*	*0.982*	*0.975*	*0.971*	*0.968*
2	**0.927**	**0.910**	**0.904**	**0.906**	**0.909**	**0.917**
3	**0.940**	**0.863**	**0.844**	**0.846**	**0.854**	**0.869**
4	**0.842**	**0.773**	**0.768**	**0.781**	**0.798**	**0.830**
5	**0.758**	**0.763**	**0.768**	**0.778**	**0.790**	**0.811**
6	**0.901**	**0.801**	**0.776**	**0.783**	**0.793**	**0.809**
7	**0.850**	**0.796**	**0.799**	**0.806**	**0.810**	**0.816**
8	**0.924**	**0.802**	**0.809**	**0.816**	**0.821**	**0.827**
			1971Q1-1986Q4			
1	*1.020*	*0.999*	*0.982*	*0.972*	*0.966*	*0.962*
2	**0.897**	**0.882**	**0.881**	**0.886**	**0.892**	**0.902**
3	**0.936**	**0.848**	**0.828**	**0.831**	**0.840**	**0.857**
4	**0.827**	**0.755**	**0.755**	**0.773**	**0.793**	**0.825**
5	**0.751**	**0.763**	**0.772**	**0.784**	**0.795**	**0.813**
6	**0.957**	**0.846**	**0.804**	**0.801**	**0.806**	**0.816**
7	**0.890**	**0.826**	**0.822**	**0.823**	**0.825**	**0.826**
8	**0.963**	**0.826**	**0.829**	**0.833**	**0.836**	**0.839**
			1987Q1-2003Q2			
1	**0.985**	**0.982**	**0.981**	**0.980**	**0.980**	**0.981**
2	*1.025*	*0.998*	*0.979*	*0.969*	*0.965*	*0.962*
3	**0.962**	**0.937**	**0.926**	**0.923**	**0.925**	**0.933**
4	**0.955**	**0.909**	**0.866**	**0.842**	**0.840**	**0.865**
5	**0.823**	**0.763**	**0.731**	**0.730**	**0.746**	**0.789**
6	**0.393**	**0.397**	**0.521**	**0.621**	**0.681**	**0.740**
7	**0.512**	**0.537**	**0.604**	**0.656**	**0.689**	**0.728**
8	**0.611**	**0.604**	**0.645**	**0.677**	**0.700**	**0.729**

Notes: See the notes to Table 1.

Table 5: Out-of-Sample Mean Square Error of Averaged Forecasts of CPI using Asset Prices Only
(relative to naive time series benchmark)

Horizon	Bayesian Model Averaging					Simple Avg.
(quarters)	φ=100	φ=5	φ=2	φ=1	φ=0.5	
			1971Q1-2003Q2			
1	**0.923**	**0.923**	**0.924**	**0.923**	**0.920**	**0.926**
2	**0.865**	**0.848**	**0.845**	**0.856**	**0.872**	**0.900**
3	**0.808**	**0.826**	**0.842**	**0.854**	**0.868**	**0.905**
4	**0.911**	**0.899**	**0.894**	**0.891**	**0.894**	**0.911**
5	*1.175*	*1.149*	*1.037*	*0.957*	*0.929*	*0.914*
6	**0.923**	**0.905**	**0.901**	**0.903**	**0.908**	**0.919**
7	**0.957**	**0.922**	**0.919**	**0.922**	**0.925**	**0.929**
8	**0.968**	**0.941**	**0.939**	**0.939**	**0.940**	**0.942**
			1971Q1-1986Q4			
1	**0.907**	**0.908**	**0.910**	**0.912**	**0.911**	**0.918**
2	**0.850**	**0.831**	**0.829**	**0.841**	**0.857**	**0.884**
3	**0.780**	**0.802**	**0.823**	**0.839**	**0.858**	**0.893**
4	**0.935**	**0.922**	**0.912**	**0.901**	**0.899**	**0.911**
5	*1.199*	*1.173*	*1.055*	*0.971*	*0.939*	*0.919*
6	*0.962*	*0.950*	*0.940*	*0.932*	*0.929*	*0.928*
7	*1.010*	*0.966*	*0.952*	*0.947*	*0.944*	*0.942*
8	**0.992**	**0.961**	**0.956**	**0.956**	**0.956**	**0.957**
			1987Q1-2003Q2			
1	**0.967**	**0.965**	**0.960**	**0.952**	**0.944**	**0.946**
2	**0.920**	**0.911**	**0.908**	**0.913**	**0.927**	**0.958**
3	**0.948**	**0.946**	**0.939**	**0.927**	**0.921**	**0.960**
4	**0.744**	**0.737**	**0.769**	**0.815**	**0.859**	**0.917**
5	**0.964**	**0.941**	**0.884**	**0.839**	**0.837**	**0.871**
6	**0.555**	**0.478**	**0.529**	**0.626**	**0.715**	**0.826**
7	**0.459**	**0.502**	**0.606**	**0.692**	**0.749**	**0.809**
8	**0.742**	**0.758**	**0.774**	**0.786**	**0.796**	**0.808**

Notes: Both Bayesian Model Averaging and simple equal-weighted averaging use 23 models corresponding to asset price predictors only. Elements in bold correspond to cases where Bayesian Model Averaging outperforms simple equal-weighted averaging for some ϕ. Elements in italics correspond to cases where simple equal-weighted averaging does better.

Table 6: Out-of-Sample Mean Square Error of Averaged Forecasts of Core CPI using Asset Prices Only (relative to naive time series benchmark)

Horizon (quarters)	Bayesian Model Averaging φ=100	φ=5	φ=2	φ=1	φ=0.5	Simple Avg.
			1971Q1-2003Q2			
1	*1.092*	*1.054*	*1.019*	*0.987*	*0.957*	*0.919*
2	**0.874**	**0.861**	**0.845**	**0.838**	**0.838**	**0.843**
3	**0.894**	**0.841**	**0.799**	**0.789**	**0.796**	**0.823**
4	**0.885**	**0.857**	**0.840**	**0.833**	**0.833**	**0.842**
5	*1.328*	*1.114*	*0.978*	*0.922*	*0.895*	*0.871*
6	*0.921*	*0.924*	*0.908*	*0.895*	*0.889*	*0.887*
7	*1.045*	*0.969*	*0.932*	*0.916*	*0.909*	*0.905*
8	**0.989**	**0.904**	**0.905**	**0.913**	**0.920**	**0.929**
			1971Q1-1986Q4			
1	*1.101*	*1.062*	*1.026*	*0.994*	*0.964*	*0.923*
2	**0.869**	**0.864**	**0.854**	**0.849**	**0.848**	**0.850**
3	**0.893**	**0.837**	**0.793**	**0.786**	**0.800**	**0.833**
4	**0.853**	**0.844**	**0.846**	**0.849**	**0.851**	**0.859**
5	*1.366*	*1.146*	*1.010*	*0.952*	*0.922*	*0.893*
6	*0.964*	*0.969*	*0.950*	*0.932*	*0.920*	*0.912*
7	*1.089*	*1.016*	*0.975*	*0.954*	*0.942*	*0.932*
8	**1.018**	**0.936**	**0.937**	**0.944**	**0.950**	**0.958**
			1987Q1-2003Q2			
1	*0.997*	*0.973*	*0.942*	*0.909*	*0.885*	*0.881*
2	**0.974**	**0.804**	**0.683**	**0.634**	**0.641**	**0.704**
3	*0.908*	*0.906*	*0.892*	*0.835*	*0.738*	*0.672*
4	**1.322**	**1.032**	**0.754**	**0.618**	**0.587**	**0.616**
5	**0.838**	**0.714**	**0.579**	**0.540**	**0.552**	**0.593**
6	**0.395**	**0.367**	**0.392**	**0.443**	**0.500**	**0.581**
7	**0.530**	**0.419**	**0.432**	**0.481**	**0.526**	**0.588**
8	**0.656**	**0.551**	**0.541**	**0.557**	**0.575**	**0.605**

Notes: See the notes to Table 5.

Table 7: Out-of-Sample Mean Square Error of Averaged Forecasts of GDP Deflator Using Asset Prices Only (relative to naive time series benchmark)

Horizon (quarters)	Bayesian Model Averaging φ=100	φ=5	φ=2	φ=1	φ=0.5	Simple Avg.
			1971Q1-2003Q2			
1	*0.993*	*0.978*	*0.968*	*0.962*	*0.957*	*0.953*
2	**1.048**	**0.992**	**0.961**	**0.950**	**0.948**	**0.951**
3	*1.084*	*1.057*	*1.015*	*0.976*	*0.954*	*0.940*
4	**0.833**	**0.839**	**0.849**	**0.864**	**0.882**	**0.910**
5	**1.136**	**1.026**	**0.910**	**0.885**	**0.885**	**0.885**
6	**0.938**	**0.875**	**0.851**	**0.853**	**0.865**	**0.883**
7	**0.921**	**0.840**	**0.846**	**0.866**	**0.880**	**0.893**
8	**0.886**	**0.866**	**0.877**	**0.887**	**0.894**	**0.902**
			1971Q1-1986Q4			
1	*1.018*	*0.995*	*0.978*	*0.965*	*0.957*	*0.948*
2	*1.090*	*1.024*	*0.984*	*0.964*	*0.955*	*0.949*
3	*1.119*	*1.091*	*1.042*	*0.993*	*0.964*	*0.941*
4	**0.828**	**0.846**	**0.866**	**0.884**	**0.897**	**0.913**
5	*1.166*	*1.052*	*0.934*	*0.907*	*0.901*	*0.891*
6	**0.978**	**0.935**	**0.907**	**0.893**	**0.890**	**0.892**
7	**0.973**	**0.906**	**0.901**	**0.902**	**0.903**	**0.904**
8	**0.909**	**0.897**	**0.904**	**0.909**	**0.912**	**0.915**
			1987Q1-2003Q2			
1	**0.926**	**0.934**	**0.944**	**0.952**	**0.958**	**0.964**
2	**0.869**	**0.854**	**0.863**	**0.890**	**0.919**	**0.958**
3	**0.852**	**0.832**	**0.839**	**0.861**	**0.886**	**0.929**
4	**0.879**	**0.780**	**0.702**	**0.688**	**0.749**	**0.882**
5	**0.868**	**0.793**	**0.696**	**0.693**	**0.745**	**0.834**
6	**0.601**	**0.362**	**0.374**	**0.517**	**0.653**	**0.813**
7	**0.513**	**0.322**	**0.415**	**0.589**	**0.702**	**0.804**
8	**0.714**	**0.630**	**0.676**	**0.725**	**0.760**	**0.800**

Notes: See the notes to Table 5.

Table 8: Out-of-Sample Mean Square Error of Averaged Forecasts of PCE Deflator Using Asset Prices Only (relative to naive time series benchmark)

Horizon (quarters)	Bayesian Model Averaging φ=100	φ=5	φ=2	φ=1	φ=0.5	Simple Avg.
			1971Q1-2003Q2			
1	*0.995*	*0.990*	*0.987*	*0.985*	*0.983*	*0.980*
2	**1.062**	**0.995**	**0.961**	**0.953**	**0.954**	**0.961**
3	**0.977**	**0.953**	**0.932**	**0.923**	**0.925**	**0.940**
4	**0.875**	**0.880**	**0.882**	**0.887**	**0.898**	**0.920**
5	*1.063*	*1.052*	*0.951*	*0.909*	*0.906*	*0.904*
6	**0.933**	**0.904**	**0.886**	**0.885**	**0.893**	**0.907**
7	**0.863**	**0.862**	**0.882**	**0.898**	**0.907**	**0.916**
8	**0.910**	**0.900**	**0.909**	**0.916**	**0.921**	**0.927**
			1971Q1-1986Q4			
1	*1.019*	*1.008*	*1.001*	*0.996*	*0.991*	*0.984*
2	**1.110**	**1.023**	**0.973**	**0.955**	**0.952**	**0.955**
3	**0.994**	**0.963**	**0.936**	**0.921**	**0.920**	**0.932**
4	**0.891**	**0.906**	**0.907**	**0.903**	**0.904**	**0.918**
5	*1.071*	*1.067*	*0.962*	*0.917*	*0.912*	*0.906*
6	*0.975*	*0.957*	*0.935*	*0.919*	*0.913*	*0.913*
7	**0.917**	**0.914**	**0.919**	**0.922**	**0.923**	**0.923**
8	**0.939**	**0.920**	**0.925**	**0.929**	**0.932**	**0.936**
			1987Q1-2003Q2			
1	**0.949**	**0.954**	**0.959**	**0.964**	**0.968**	**0.973**
2	**0.907**	**0.908**	**0.925**	**0.946**	**0.963**	**0.981**
3	**0.891**	**0.900**	**0.914**	**0.932**	**0.949**	**0.977**
4	**0.756**	**0.690**	**0.689**	**0.765**	**0.846**	**0.938**
5	**0.987**	**0.923**	**0.854**	**0.837**	**0.852**	**0.890**
6	**0.555**	**0.429**	**0.436**	**0.583**	**0.718**	**0.858**
7	**0.401**	**0.418**	**0.563**	**0.693**	**0.774**	**0.850**
8	**0.684**	**0.740**	**0.783**	**0.809**	**0.827**	**0.847**

Notes: See the notes to Table 5.

Table 9: Fraction of Quarters when Bayesian Model Averaging gives Closer Prediction than Simple Equal Weighted Model Averaging
All Predictors

Horizon (qtrs)	φ=100	φ=5	φ=2	φ=1	φ=0.5
			CPI		
1	0.48	0.48	0.46	0.47	0.49
2	0.51	0.52	0.53	0.54	0.55
3	0.50	0.53	0.55	0.59	0.58
4	0.53	0.55	0.57	0.58	0.59
5	0.53	0.52	0.54	0.57	0.59
6	0.54	0.65	0.67	0.72*	0.71*
7	0.45	0.64	0.70*	0.71*	0.70*
8	0.53	0.62	0.65	0.64	0.65
			CPI Core		
1	0.47	0.50	0.53	0.55	0.58*
2	0.47	0.52	0.58	0.62*	0.66*
3	0.40	0.42	0.42	0.52	0.58
4	0.65*	0.67*	0.65*	0.65*	0.68*
5	0.56	0.57	0.61	0.63	0.64
6	0.59	0.63	0.69*	0.71*	0.71*
7	0.55	0.58	0.62	0.64	0.65
8	0.60	0.68	0.68	0.70	0.71*
			GDP Deflator		
1	0.52	0.52	0.53	0.55	0.56
2	0.48	0.50	0.52	0.53	0.55
3	0.52	0.53	0.56	0.57	0.58
4	0.58	0.58	0.60	0.61	0.63
5	0.56	0.63	0.67*	0.68*	0.67*
6	0.60	0.68*	0.71*	0.72*	0.74*
7	0.55	0.60	0.62	0.64	0.62
8	0.62	0.75*	0.77*	0.78*	0.78*
			PCE Deflator		
1	0.45	0.45	0.44	0.45	0.45
2	0.47	0.48	0.47	0.50	0.52
3	0.49	0.53	0.56	0.55	0.55
4	0.59	0.57	0.60	0.65*	0.65*
5	0.62	0.57	0.58	0.60	0.61
6	0.62	0.65	0.68*	0.68*	0.68*
7	0.55	0.60	0.65	0.66	0.66
8	0.46	0.62	0.65	0.69	0.69

Notes: An asterisk denotes that the proportion is significantly different from 0.5 using a 10% two-tailed test as described in the text.

Table 10: Fraction of Quarters when Bayesian Model Averaging gives Closer Prediction than Simple Equal Weighted Model Averaging
Asset Prices Only

Horizon (qtrs)	φ=100	φ=5	φ=2	φ=1	φ=0.5
			CPI		
1	0.44	0.45	0.44	0.45	0.45
2	0.55	0.52	0.52	0.55	0.57
3	0.49	0.51	0.52	0.52	0.53
4	0.56	0.57	0.61	0.63	0.65*
5	0.45	0.47	0.49	0.55	0.59
6	0.58	0.66	0.68*	0.68*	0.67
7	0.57	0.60	0.64	0.65	0.66
8	0.53	0.57	0.58	0.58	0.58
			CPI Core		
1	0.39*	0.40*	0.42*	0.43	0.46
2	0.45	0.51	0.51	0.54	0.56
3	0.48	0.50	0.48	0.48	0.45
4	0.42	0.42	0.47	0.53	0.56
5	0.39	0.40	0.48	0.52	0.51
6	0.62	0.67	0.65	0.66	0.67
7	0.42	0.49	0.55	0.58	0.61
8	0.47	0.55	0.57	0.59	0.62
			GDP Deflator		
1	0.50	0.51	0.50	0.50	0.48
2	0.52	0.52	0.55	0.55	0.55
3	0.52	0.56	0.56	0.55	0.56
4	0.59	0.60	0.64	0.70*	0.72*
5	0.48	0.52	0.60	0.57	0.57
6	0.62	0.65	0.67	0.65	0.67
7	0.44	0.58	0.62	0.62	0.65
8	0.55	0.65	0.69	0.74*	0.75*
			PCE Deflator		
1	0.52	0.50	0.51	0.51	0.51
2	0.54	0.54	0.55	0.55	0.55
3	0.54	0.55	0.55	0.56	0.59
4	0.65*	0.66*	0.68*	0.69*	0.69*
5	0.46	0.49	0.59	0.61	0.62
6	0.61	0.67	0.69*	0.68*	0.66
7	0.58	0.63	0.68	0.68	0.69
8	0.62	0.67	0.68	0.68	0.69

Notes: An asterisk denotes that the proportion is significantly different from 0.5 using a 10% two-tailed test as described in the text.