Board of Governors of the Federal Reserve System

International Finance Discussion Papers
Number 779
September 2003

Bayesian Model Averaging and Exchange Rate Forecasts

Jonathan H. Wright

Bayesian Model Averaging and Exchange Rate Forecasts

Jonathan H. Wright*

Abstract: Exchange rate forecasting is hard and the seminal result of Meese and Rogoff (1983) that the exchange rate is well approximated by a driftless random walk, at least for prediction purposes, has never really been overturned despite much effort at constructing other forecasting models. However, in several other macro and financial forecasting applications, researchers in recent years have considered methods for forecasting that combine the information in a large number of time series. One method that has been found to be remarkably useful for out-of-sample prediction is simple averaging of the forecasts of different models. This often seems to work better than the forecasts from any one model. Bayesian Model Averaging is a closely related method that has also been found to be useful for out-of-sample prediction. This starts out with many possible models and prior beliefs about the probability that each model is the true one. It then involves computing the posterior probability that each model is the true one, and averages the forecasts from the different models, weighting them by these posterior probabilities. This is effectively a shrinkage methodology, but with shrinkage over models not just over parameters. I apply this Bayesian Model Averaging approach to pseudo-out-of-sample exchange rate forecasting over the last ten years. I find that it compares quite favorably to a driftless random walk forecast. Depending on the currency-horizon pair, the Bayesian Model Averaging forecasts sometimes do quite a bit better than the random walk benchmark (in terms of mean square prediction error), while they never do much worse. The forecasts generated by this model averaging methodology are however very close to (but not identical to) those from the random walk forecast.

Keywords: Shrinkage, model uncertainty, forecasting, exchange rates, bootstrap.
JEL Classification: C32, C53, F31.

* International Finance Division, Board of Governors of the Federal Reserve System, Washington DC 20551. I am grateful to Ben Bernanke, David Bowman, Jon Faust, Matt Pritsker and Pietro Veronesi for helpful comments, and to Sergey Chernenko for excellent research assistance. The views in this paper are solely the responsibility of the author and should not be interpreted as reflecting the views of the Board of Governors of the Federal Reserve System or of any person associated with the Federal Reserve System.

1. Introduction.

Out-of-sample forecasting of exchange rates is hard. Meese and Rogoff (1983) argued that all exchange rate models do less well in out-of-sample forecasting exercises than a simple driftless random walk. Although this finding was heresy to many at the time that Meese and Rogoff wrote their seminal paper, it has now become the conventional wisdom. Mark (1995) claimed that a monetary fundamentals model can generate better out-of-sample forecasting performance at long horizons, but that result has been found to be very sensitive to the sample period (Groen (1999), Faust, Rogers and Wright (2003)). Claims that a particular variable has predictive power for exchange rates crop up frequently, but these results typically apply just to a particular exchange rate and a particular subsample. As such, they are by now met with justifiable skepticism and are thought of by many as the result of data-mining exercises.

Exchange rates are not the only data that are hard to predict. Atkeson and Ohanian (2001) showed that Phillips-curve based forecasts of inflation give larger out-of-sample prediction errors than a simple random walk forecast of inflation. Stock and Watson (2001, 2002a) consider prediction of inflation and output growth in each of the G7 countries using a large number of possible models. 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 or output growth. 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. The models that Stock and Watson consider are simple: each model consists of a regression of inflation/output growth on a single leading indicator and a lagged dependent variable. Heavily parameterized models, with large numbers of variables or capricious nonlinear specifications can provide extraordinarily good fits in sample, but generally make matters worse in terms of out-of-sample prediction.

However, in the context of inflation and output growth prediction, researchers have recently made substantial progress in forecasting 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 (2002b) and Bernanke and Boivin (2003) are among the many possible cites). 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 output growth and inflation, 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.

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 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.

In this paper, I plan to use forecast combination methods that have been found to be useful in other contexts, but to apply them to the problem of out-of-sample exchange rate prediction. I shall pool forecasts from a large number of different models, to see whether this idea that has been so successful in the context of output growth and inflation forecasting makes any dent in the context of exchange rate forecasting. But I shall also try to apply the closely related idea of Bayesian Model Averaging (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 may be useful for out-of-sample forecasting of exchange rates in the 1990s. It seems to work better than simple equal-weighted model averaging, in this particular context at least.

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 exchange rate prediction exercise is described in section 3. Using a large number of models, combined using Bayesian Model Averaging methods, gives promising results for out-of-sample exchange rate forecasting. 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.[1] 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.[2] This is all there is to Bayesian Model Averaging. The researcher needs only specify the set of

---

[1] 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)).

[2] 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)).

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 exchange rate returns), 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. Motivated by the possibility of overlapping data in my subsequent application, I assume that the error term is an MA(h-1) process with variance σ^2 such that

$$Cov(\varepsilon_t, \varepsilon_{t-j}) = \sigma^2 \frac{h-j}{h}, j \leq h-1$$

I shall define the models and the model priors in the context of the empirical application below. 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 $N(0, \phi\sigma^2 (X'X)^{-1})$. 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. Routine integration (Zellner (1971)) then yields the required likelihood of the model

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

where

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

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 exchange rate 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 Exchange Rate Forecasting

The application I consider is to forecasting exchange rate returns. Specifically, each model is of the form

$$e_{t+h} - e_t = \beta' X_t + \varepsilon_t$$

where e_t denotes the log exchange rate, h is the forecasting horizon, X_t is a vector of regressors, and ε_t is the error term, assumed to satisfy the restrictions above (so that it is a moving average process of order h-1). A possible model is one with no predictors at all which is simply the driftless random walk

$$e_{t+h} - e_t = \varepsilon_t$$

I shall consider prediction of the bilateral exchange value of the Canadian dollar, pound, yen and mark/euro, relative to the US dollar.

3.1 *Monthly Financial Dataset*

I first consider pseudo-out-of-sample exchange rate prediction using a dataset of financial variables as the possible predictors. These data are available at a monthly frequency, are available in real-time and are never revised. Data vintage issues can substantially affect the results of exchange-rate forecasting exercises, as noted by Faust, Rogers and Wright (2003). The predictors are (i) the relative stock prices (foreign-US) (logs and log differences), (ii) the relative dividend yield, (iii) relative long term interest rates, (iv) relative short term interest rates, (v) the relative term spread, (vi) oil prices (logs and log differences) (vii) exchange rate returns over the previous month, and (viii) the sign of exchange rate returns over the previous month. Data sources are given in Appendix A. The data cover the months 1973:01 to 2002:12. The models I consider are the driftless random walk model plus all linear regression models in which the exchange rate return is predicted by any one of these variables (plus a constant). This gives a total of 11 candidate models. Note that it is also possible to consider prediction using all possible permutations of these predictors rather than just using them one at a time. I shall implement this as well at the end of section 3.

The pseudo-out-of-sample prediction exercise involves forecasting the exchange rate for 1993:01 to 2002:12 as of h months previously, for $h = 3,6,9,12$. For example, the first 3-month ahead forecast is the prediction of the exchange rate in 1993:01 that was made in 1992:10. Of course, this forecast is constructed only using data from 1992:10 and earlier.

3.2 *Quarterly Macro and Financial Dataset*

Although the monthly financial dataset has some advantages, it is missing a great many of the variables that researchers claim have predictive power for exchange rates. To include these, I switch to quarterly data, and give up on the "real-time" feature of the monthly asset price dataset.

This larger dataset contains all the same variables as the monthly data, aggregated to quarterly frequency. In addition it includes (i) relative GDP (foreign-US) (logs and log differences), (ii) relative money supply (logs and log differences), (iii) the relative price level (logs and log differences), (iv) the relative ratio of current account to GDP (level and cumulated) and (v) the monetary fundamentals as defined by Mark (1995).

The data cover the quarters 1973:1 to 2002:4. The models I consider are the driftless random walk model and all linear regression models in which the exchange rate return is predicted by any one of these variables (plus a constant). This gives a total of 20 models. The pseudo-out-of-sample prediction exercise involves forecasting the exchange rate for 1993:1 to 2002:4 as of h quarters previously, for $h = 1, 2, 3, 4$.

3.3 *Results for Equal Weighted Model Averaging*

I first considered the out-of-sample mean square prediction error of the forecast obtained by averaging the predictions across all the different models, giving all models equal weight, relative to the out-of-sample mean square prediction error for the forecast assuming that the exchange rate is a driftless random walk. Table 1 shows this equal weighted relative out-of-sample root mean square prediction error (RMSPE) in both the monthly and quarterly datasets. A number greater than 1 means that equal-weighted model-averaging is forecasting less well than a random walk. Except for the Canadian dollar, most entries in these tables are greater than 1.[3] Simple equal-weighted model averaging, that is such an effective strategy in many forecasting contexts, does not seem to buy us very much in exchange rate forecasting, at least not with these models.

3.4 *Results for Bayesian Model Averaging*

I now turn to Bayesian Model Averaging, which weights the forecasts from different models by their posterior probabilities. Table 2 shows the out-of-sample RMSPE for Bayesian Model Averaging. In the monthly dataset, for sterling, the out-of-sample RMSPE is uniformly slightly above 1 indicating that the random walk gives better forecasts. But for the other three currencies, the RMSPE is nearly uniformly below 1 in the monthly dataset, indicating that Bayesian Model Averaging gives better forecasts.

Similar results are obtained in the quarterly dataset, also shown in Table 2. The addition of the macro variables in the quarterly dataset does little on net to either improve (or worsen) predictive performance.

Although Bayesian Model Averaging can give good results for some currency-horizon pairs with a large value of ϕ, overall the best results are obtained with a smaller value of ϕ (e.g. $\phi = 1$). In other words, a fairly informative prior with substantial shrinkage improves the forecasting performance of model averaging. In this sense, it does not pay to try to make the prior as uninformative as possible.

For small ϕ it is fair to say that Bayesian Model Averaging can help quite a bit, but cannot hurt much. For example if $\phi = 1$, it can lower mean square prediction error by

[3] I do not show results for the out-of-sample RMSPE for the individual models but, not surprisingly, although the RMSPE is below 1 for some models and currency-horizon pairs, there is no model for which it is below 1 on average across all currency-horizon pairs, in either the monthly or quarterly datasets.

up to 12%, while the worst case is that it raises mean square prediction error by 2%, relative to the random walk benchmark.

Bootstrap p-values of the hypothesis that the out-of-sample RMSPE is one are reported in Table 3. In each bootstrap sample an artificial dataset is generated in which the exchange rate is by construction a driftless random walk, using the bootstrap methodology described in Appendix B. The p-values in Table 3 represent the proportion of bootstrap samples for which the RMSPE is smaller than that which was actually observed in the data. These are therefore one-sided p-values, testing the null of equal predictability against the alternative that Bayesian Model Averaging gives a significant improvement over the driftless random walk. The null is rejected for several currency-horizon pairs in both the monthly and quarterly datasets, at conventional significance levels.

Researchers are rightly suspicious of significant p-values in a test of the hypothesis that a particular model forecasts the exchange rate better than a random walk. The key reason is that these p-values ignore the data mining that was implicit in choosing the particular model to use. Researchers publish the results of these tests only if they find a model which forecasts the exchange rate significantly better than a random walk, and thus "significant" results can be expected to crop up from time to time even if the exchange rate is totally unpredictable. But to the extent that the Bayesian Model Averaging approach is starting out with a set of models that spans the space of all models researchers would ever want to consider, the results and specifically the p-values in the forecast comparison test are then immune to any such data-mining critique.

Bayesian Model Averaging forecasts are not necessarily very different from random walk forecasts. The driftless random walk forecast is of course for no change in the exchange rate. The root mean square forecast exchange rate change in the Bayesian Model Averaging gives a metric for how different this forecast is from a random walk. I report this root mean square forecast exchange rate change in Table 4. In general, the higher is ϕ and the longer is the forecast horizon, the larger is the magnitude of the forecast exchange rate changes. But the key thing to note is that generally the Bayesian Model Averaging procedure is not forecasting large exchange rate fluctuations. For $\phi = 1$, where the Bayesian Model Averaging procedure outperforms the random walk for most currency-horizon pairs, the root mean square forecast exchange rate change at a one-year horizon is at most 1.54 percent.

Since most economists believe that the exchange rate is very well approximated by a random walk, this is a reassuring feature of the Bayesian Model Averaging procedure.

Bayesian Model Averaging predicts small exchange rate changes. One question of some interest is whether it predicts the *sign* of the exchange rate change correctly or not. Among other things, this metric is robust to the possibility of outliers that artificially enhance/inhibit predictive performance. Table 5 reports the proportion of times that it does predict the correct sign of the exchange rate change. Bayesian Model Averaging predicts the correct sign more than half the time for the Canadian dollar and mark/euro, at all horizons and for all choices of φ. Results for the yen and pound are mixed. For the case $\varphi=1$, the proportion of times that Bayesian Model Averaging predicts the sign of the exchange rate change correctly is at least one standard deviation above a coin toss at all horizons.

Cheung, Chinn and Pascual (2002) also considered forecasting exchange rates over the 1990s, but not using Bayesian methods. They found that they were able to predict the direction of exchange rate changes more than half the time, but typically underperformed the random walk in mean square error. Meanwhile, Bayesian Model Averaging methods are typically outperform the random walk both in mean square error and in the sign of the exchange rate change. These results are all very much consistent with the idea that model and parameter uncertainty are the stumbling blocks to exchange rate forecasting (given that the exchange rate is so close to being a random walk), and that the researcher who wants to get good out-of-sample prediction, rather than in-sample fit, should use shrinkage methods.

3.5 *How Stable Is the Predictive Power of Bayesian Model Averaging?*

In many forecasting applications, 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 no propensity to have good predictive power in another subperiod.

To give a little evidence on whether Bayesian Model Averaging prediction of exchange rates suffers from this problem, I computed the out-of-sample RMSPE in two subsamples: 1993-1997 and 1998-2002 for all 16 currency-horizon pairs in both datasets for the case $\phi = 1$. Figure 1, using a graphical device adapted from Stock and Watson (2001, 2002a), plots this out-of-sample RMSPE with the value in the 1998-2002 subperiod on the vertical axis and the value in the 1993-1997 subperiod on the horizontal axis. The figures are split into 4 quadrants: a currency-horizon pair in the lower left quadrant is forecast by Bayesian Model Averaging better than a random walk in both subperiods, a currency-horizon pair in the upper right quadrant is forecast better by the random walk in both subperiods, and currency-horizon pairs in the other quadrants are forecast better by the random walk in one subperiod but not in the other.

If the predictive power of Bayesian Model Averaging were highly unstable, we would expect to see few observations in the bottom left quadrant, but many observations in the top left and bottom right quadrants. There is some such forecast instability, but it is not too severe. In the quarterly dataset, 9 out of the 16 currency-horizon pairs are in the bottom left quadrant (consistently better predicted by Bayesian Model Averaging). In the monthly dataset, 8 out of 16 are in this quadrant (and 2 more are very close). None of the pairs are in the top right quadrant (consistently better predicted by the random walk) in either dataset. While Bayesian Model Averaging is clearly not guaranteed to outperform the random walk exchange rate forecast for all currency-horizon pairs in all subsamples, these graphs still look very different from their counterparts shown by Stock and Watson (2001, 2002a) for some individual forecasting models for inflation and output growth where most entries were in the top right quadrant and very few were in the bottom left quadrant.

3.6 *Forecasting at Longer Horizons*

In this paper, I have reported exchange rate forecasting at short to medium horizons, up to one year. These are the horizons considered by Meese and Rogoff (1983), but more recent work has forecast exchange rates at longer horizons. I have also experimented with using Bayesian Model Averaging at longer horizons, but found that it predicts consistently less well than a random walk at horizons of two years or more.

3.7 *Bayesian Model Averaging with All Possible Permutations of Regressors*

Each of the models I have considered (with the exception of the driftless random walk) consists of a single predictor, plus a constant. The idea of averaging over models with a single predictor in each model has been used successfully in other contexts (Stock and Watson (2001, 2002a)), albeit weighting the forecasts from the different models equally instead of using posterior probabilities as weights. It seems to work well in the context of Bayesian Model Averaging for exchange rate prediction. However, it is nonstandard in Bayesian Model Averaging methodology.

A more standard Bayesian Model Averaging approach would use all possible permutations of predictors (including none/all of the predictors), generating a large number of candidate models. One has to however think carefully about the priors for such models. Assigning equal prior probability to each model means that models with a small number of predictors may receive too little prior weight. If λ is the number of predictor variables and each model consist of some permutation of these predictors for a total of 2^{λ} models, then a standard approach is to specify that the prior probability for a model with κ predictors is

$$P(M_i) = \rho^{\kappa}(1-\rho)^{\lambda-\kappa}$$

This is implemented by Cremers (2002) and Koop and Potter (2003), among others. If $\rho = 0.5$, then all the models get equal weight - a smaller value of this hyperparameter favors smaller models. The probability that the true model has no predictors is $(1-\rho)^{\lambda}$. The expected number of predictors is $\rho\lambda$.

I considered Bayesian Model Averaging using all possible permutations of predictors with this prior in the monthly and quarterly datasets. In the monthly dataset, there are 9 predictors[4] and so a total of 2^9=512 models. I augment each of these models with a constant, except for the model with no predictors which is just a driftless random walk.

In the quarterly dataset, there are 17 predictors[5] and so a total of 2^{17}=131072 models. Again, I include an intercept in each of these models except for the model with no predictors which is just a driftless random walk.

Although the number of models is large, especially in the quarterly dataset, it is still computationally possible to evaluate the posterior probabilities of all of the models by simply applying the formula in equation (1)[6]. If there were many more models, it would be necessary to use simulation based methods instead.[7]

Table 6 shows the out-of-sample RMSPE for Bayesian Model Averaging using all of these models and this prior with ρ=0.2, 0.1, 0.05 and setting φ=1. To compare these results with those from Bayesian Model Averaging over single-predictor models, compare the elements of Table 6 with the elements of the φ=1 column in Table 2. The

[4] The predictors are the relative stock prices (logs and log differences), the relative dividend yield, relative short-term interest rates, relative term spreads, oil prices (logs and log differences) exchange rate returns over the previous month and the sign of exchange rate returns over the previous month. The relative long-term interest rate was dropped from the list of predictors, because there would otherwise be a perfect multicollinearity problem when considering all possible permutations of regressors.

[5] These are the same predictors as for the monthly dataset (aggregated to quarterly frequency) plus relative GDP (logs and log differences), relative money supply (logs and log differences), the relative price level (logs and log differences) and the relative ratio of current account to GDP (level and cumulated). The Mark (1995) monetary fundamentals were dropped from the list of predictors, because there would otherwise be a perfect multicollinearity problem when considering all possible permutations of regressors.

[6] For any one currency-horizon pair and any one choice of φ, the computation in the quarterly dataset takes about 15 minutes on a 2.5 Ghz computer.

[7] 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.

results are quite similar overall, though on balance the more complicated procedure actually seems to work a little less well.

4. Conclusion and Future Research

In this paper I have considered a specific approach to pooling the forecasts from different models, namely Bayesian Model Averaging, and argued that it gives promising results for out-of-sample exchange rate prediction. With suitable shrinkage, Bayesian Model Averaging can help quite a bit and cannot hurt much. That is, depending on the currency-horizon pair, the Bayesian Model Averaging forecasts sometimes do quite a bit better than the random walk benchmark in terms of mean square prediction error, while they never do much worse. The finding that this method for pooling model forecasts works well (while the individual forecasts do not) is related to the folklore in the forecasting literature that averaging the forecasts from many different models gives superior predictive performance. However, simple equal-weighted averaging does not work very well when applied to the hard problem of exchange rate forecasting. Bayesian Model Averaging which averages the forecasts from different models but chooses the weights based on a combination of prior beliefs and the observed data, gives better results for exchange rate prediction.

I have considered both Bayesian Model Averaging using simple models with a single predictor in each model, and using all possible permutations of regressors. Interestingly, while both methods give good forecasting results, the latter method does not work much better and can indeed work less well. And it is computationally harder.

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.[8] 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.

It would be possible and interesting to include Federal Reserve Greenbook exchange rate forecasts in the model averaging exercise. This involves certain complications. Firstly, the forecasts are made on an FOMC schedule, rather than at a monthly or quarterly frequency. Some months have no FOMC meeting and so no Greenbook forecast. Secondly, the forecasts are not public (and cannot be used in any non-internal Fed work) until 5 years after they are made, substantially restricting the length of the sample.

It would also be possible to include nonlinear models, notably Markov switching and threshold models in the model averaging exercise. Such models have been considered by Engel and Hamilton (1990) and Kilian and Taylor (2003) among others. These models may contain information that could make them useful as elements of a forecast pooling exercise such as Bayesian Model Averaging. Stock and Watson (1999) discuss the potential for using nonlinear models as components of a macroeconomic forecast combination exercise.

[8] 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.

Although I have reported promising results on the performance of Bayesian Model Averaging relative to the driftless random walk benchmark, I do not make any claim that this “overturns” the result of Meese and Rogoff (1983). The Bayesian Model Averaging forecasts are very close to random walk forecasts. I find this to be a reassuring feature. Many researchers are skeptical of any model of exchange rates that does not give a flatline or near-flatline forecast of future exchange rates, and I agree that they are right to be skeptical. The exchange rate is at least very close to being a driftless random walk. Methods using large datasets, such as Bayesian Model Averaging, may be able judiciously to pool information from a large number of indicators so as to do a little better than a flatline forecast. I claim that there is evidence for this and that it is useful, but at the same time it remains true that the exchange rate is very close to being a driftless random walk.

Appendix A: Data Sources

Exchange Rates: OECD Main Economic Indicators (MEI).

Long Term Interest Rates: 10 year rates from MEI and IFS.

Short Term Interest Rates: 3-month rates from MEI.

Oil Prices: Index of spot oil prices from IFS.

Prices*: CPI from MEI. Not seasonally adjusted.

GDP*: MEI (real, sa).

Money Stock*: M1 from MEI (sa).

Mark fundamentals*: Computed from money stock, GDP and prices.

Stock Prices: MSCI price indices (local currency).

Dividend Yields: Computed from MSCI price and total return indices.

Current Account*: MEI (sa).

Current Account/GDP ratio*: Computed from MEI (nominal and sa in both numerator and denominator).

*: included in quarterly dataset only.

Appendix B: Construction of Bootstrap Samples

To construct bootstrap samples in the monthly dataset, I fitted a VAR(12) to the exchange rate, log relative stock prices, the relative dividend yield, relative short term interest rates, the relative term spread, and log oil prices. I estimated this 6-variable VAR but anchored the bootstrap at the bias-adjusted estimates of Kilian (1998), not the OLS estimates. I also imposed that all the coefficients in the exchange rate equation were equal to zero, except for the coefficient on the first lag of the exchange rate which I set to one. So the exchange rate is a driftless random walk by construction (the null hypothesis of no forecastability is imposed). I then generated 500 bootstrap samples of all of the variables in this VAR. All of the predictors in the monthly dataset can be constructed from these 6 variables. So in this way I get a bootstrap sample of the exchange rate and all of the predictors in which the exchange rate is a random walk (but may affect future values of the predictors and so is not strictly exogenous).

In the quarterly dataset, I fitted a VAR(4) to the exchange rate, log relative stock prices, the relative dividend yield, relative short term interest rates, the relative term spread, and log oil prices, relative log GDP, relative log money supply, relative log prices and the relative ratio of current account to GDP, and then proceeded in exactly the same way as for the monthly dataset.

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.

Cheung, Y-W, M.D. Chinn and A.G. Pascual (2002): Empirical Exchange Rate Models of the Nineties: Are Any Fit to Survive?, NBER Working Paper 9393.

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.

Engel, C. and J.D. Hamilton (1990): Long Swings in the Dollar: Are They in the Data and Do Markets Know It?, *American Economic Review*, 80, pp.689-713.

Faust, J., J.H. Rogers and J.H. Wright (2003): Exchange Rate Forecasting: The Errors We've Really Made, *Journal of International Economics*, 60, pp.35-59.

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.

Groen, J.J. (1999): Long Horizon Predictability of Exchange Rates: Is it for Real?, *Empirical Economics*, 24, pp.451-469.

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.

Kilian, L. (1998): Small-Sample Confidence Intervals for Impulse Response Functions, *Review of Economics and Statistics*, 80, pp.218-230.

Kilian, L. and M.P. Taylor (2003): Why is it so Difficult to Beat the Random Walk Forecast of Exchange Rates?, *Journal of International Economics*, 60, pp.85-107.

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.

Mark, N. (1995): Exchange Rates and Fundamentals: Evidence on Long-Horizon Predictability, *American Economic Review*, 85, pp.201-218.

Meese, R. and K. Rogoff (1983): Empirical Exchange Rate Models of the Seventies: Do They Fit Out of Sample?, *Journal of International Economics*, 14, pp.3-24.

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.

Stock, J.H. and M.W. Watson (1999): "A Comparison of Linear and Nonlinear Univariate Models for Forecasting Macroeconomic Time Series", in R.F. Engle and H. White (eds.), *Cointegration, Causality and Forecasting: A Festschrift in Honor of Clive W.J. Granger*, Oxford University Press, Oxford.

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.

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 RMSPE for Equal-Weighted Averaged Forecasts

Horizon	Canadian $	Mark/Euro	Yen	Pound
		Monthly Financial Data		
3 months	0.949	1.001	1.006	1.022
6 months	0.883	1.011	1.019	1.036
9 months	0.834	1.016	1.031	1.044
12 months	0.799	1.026	1.055	1.030
		Quarterly Data		
1 quarter	0.917	0.992	1.000	0.990
2 quarter	0.877	1.013	1.012	1.044
3 quarter	0.842	1.019	1.020	1.048
4 quarter	0.837	1.037	1.056	1.049

Notes: This Table reports the out-of-sample mean square prediction error from the forecasts taken by simple equal-weighted averaging all of the different model forecasts, relative to the mean square prediction error from a driftless random walk forecast. A number less than 1 means that averaging the model forecasts predicts better than the random walk benchmark. The models considered are described in the text – each has a constant and one other predictor, except for the driftless random walk.

Table 2: Out-of-sample RMSPE for Bayesian Model Averaging

Currency	Horizon	φ=20	φ=5	φ=2	φ=1	φ=0.5
				Monthly Financial Data		
Canadian $	3 months	0.970	0.961	0.960	0.965	0.973
	6 months	0.927	0.909	0.911	0.924	0.944
	9 months	0.897	0.872	0.877	0.896	0.924
	12 months	0.873	0.844	0.853	0.877	0.910
Mark/Euro	3 months	1.016	0.981	0.952	0.953	0.975
	6 months	0.897	0.869	0.870	0.914	0.964
	9 months	0.853	0.845	0.876	0.929	0.973
	12 months	0.836	0.838	0.890	0.944	0.981
Yen	3 months	0.993	0.991	0.992	0.994	0.996
	6 months	0.977	0.978	0.983	0.989	0.994
	9 months	0.973	0.974	0.979	0.985	0.991
	12 months	0.998	0.997	0.996	0.995	0.995
Pound	3 months	1.025	1.027	1.017	1.008	1.003
	6 months	1.031	1.037	1.027	1.016	1.008
	9 months	1.020	1.029	1.021	1.012	1.005
	12 months	1.007	1.015	1.009	1.002	0.998
			Quarterly Data			
Canadian $	1 quarter	0.968	0.951	0.940	0.943	0.956
	2 quarters	0.933	0.919	0.917	0.925	0.942
	3 quarters	0.917	0.900	0.895	0.904	0.926
	4 quarters	0.976	0.947	0.919	0.914	0.928
Mark/Euro	1 quarter	1.095	1.034	0.977	0.959	0.974
	2 quarters	0.966	0.898	0.867	0.907	0.964
	3 quarters	0.871	0.847	0.873	0.931	0.975
	4 quarters	0.850	0.864	0.914	0.960	0.988
Yen	1 quarter	1.021	1.007	0.994	0.989	0.990
	2 quarters	0.980	0.975	0.974	0.979	0.986
	3 quarters	0.958	0.956	0.961	0.970	0.980
	4 quarters	1.007	1.001	0.993	0.990	0.990
Pound	1 quarter	0.977	0.976	0.979	0.984	0.990
	2 quarters	1.046	1.048	1.035	1.023	1.013
	3 quarters	1.037	1.042	1.032	1.022	1.013
	4 quarters	1.033	1.039	1.031	1.021	1.012

Notes: This Table reports the out-of-sample mean square prediction error from the forecasts taken by Bayesian Model Averaging, relative to the mean square prediction error from a driftless random walk forecast. A number less than 1 means that Bayesian Model Averaging predicts better than the random walk benchmark. The models used in the Bayesian Model Averaging procedure are described in the text – each has a constant and one other predictor, except for the driftless random walk.

Table 3: Test That Bayesian Averaging & Random Walk Have Out-of-Sample RMSPE of 1 (p-values)

Currency	Horizon	φ=20	φ=5	φ=2	φ=1	φ=0.5
			Monthly Financial Data			
Canadian $	1 quarter	0.02	0.03	0.04	0.05	0.06
	2 quarters	0.02	0.02	0.04	0.05	0.06
	3 quarters	0.02	0.03	0.04	0.05	0.06
	4 quarters	0.05	0.05	0.05	0.06	0.09
Mark/Euro	1 quarter	0.49	0.08	0.02	0.02	0.03
	2 quarters	0.01	0.01	0.01	0.02	0.07
	3 quarters	0.02	0.02	0.02	0.06	0.14
	4 quarters	0.02	0.02	0.05	0.12	0.21
Yen	1 quarter	0.10	0.12	0.13	0.14	0.16
	2 quarters	0.07	0.09	0.11	0.14	0.18
	3 quarters	0.08	0.11	0.14	0.16	0.19
	4 quarters	0.21	0.22	0.24	0.26	0.27
Pound	1 quarter	0.73	0.74	0.62	0.52	0.45
	2 quarters	0.58	0.57	0.56	0.53	0.50
	3 quarters	0.37	0.41	0.41	0.42	0.42
	4 quarters	0.27	0.31	0.33	0.34	0.35
			Quarterly Data			
Canadian $	1 quarter	0.07	0.04	0.02	0.02	0.01
	2 quarters	0.07	0.06	0.05	0.03	0.03
	3 quarters	0.08	0.07	0.06	0.05	0.03
	4 quarters	0.17	0.14	0.11	0.09	0.07
Mark/Euro	1 quarter	0.97	0.68	0.06	0.02	0.03
	2 quarters	0.06	0.02	0.00	0.02	0.04
	3 quarters	0.03	0.02	0.03	0.04	0.12
	4 quarters	0.03	0.04	0.06	0.11	0.22
Yen	1 quarter	0.43	0.25	0.18	0.14	0.13
	2 quarters	0.11	0.11	0.13	0.15	0.17
	3 quarters	0.11	0.11	0.13	0.15	0.18
	4 quarters	0.18	0.18	0.21	0.23	0.26
Pound	1 quarter	0.06	0.07	0.08	0.11	0.14
	2 quarters	0.58	0.56	0.56	0.55	0.55
	3 quarters	0.33	0.36	0.40	0.43	0.43
	4 quarters	0.28	0.31	0.35	0.39	0.41

Notes: This Table reports the bootstrap p-values for a one-sided test of the hypothesis that the driftless random walk and Bayesian Model Averaging forecasts have equal out-of-sample mean square prediction error. Specifically the entries are the fraction of bootstrap samples in which the RMSPE is below the sample value as reported in Table 2. The bootstrap methodology is described in Appendix B.

Table 4: Root Mean Square Forecast of Exchange Rate Change with Bayesian Model Averaging

Currency	Horizon	φ=20	φ=5	φ=2	φ=1	φ=0.5
			Monthly Financial Data			
Canadian $	3 months	0.36	0.37	0.30	0.22	0.14
	6 months	0.61	0.66	0.54	0.41	0.27
	9 months	0.80	0.90	0.75	0.57	0.38
	12 months	0.95	1.13	0.97	0.74	0.49
Mark/Euro	3 months	2.35	1.96	1.36	0.80	0.28
	6 months	3.68	3.10	1.94	0.95	0.29
	9 months	3.62	2.94	1.63	0.72	0.23
	12 months	2.54	2.13	1.10	0.42	0.23
Yen	3 months	0.64	0.61	0.48	0.35	0.23
	6 months	1.14	1.12	0.93	0.69	0.46
	9 months	1.79	1.82	1.48	1.08	0.72
	12 months	2.26	2.41	2.02	1.54	1.01
Pound	3 months	0.39	0.42	0.35	0.26	0.17
	6 months	0.58	0.72	0.62	0.47	0.31
	9 months	0.83	1.06	0.91	0.69	0.46
	12 months	1.01	1.31	1.14	0.86	0.57
			Quarterly Data			
Canadian $	1 quarter	0.49	0.43	0.34	0.24	0.16
	2 quarters	0.68	0.64	0.52	0.40	0.27
	3 quarters	0.97	0.93	0.75	0.57	0.38
	4 quarters	1.24	1.22	1.01	0.76	0.50
Mark/Euro	1 quarter	2.14	1.77	1.18	0.64	0.25
	2 quarters	4.18	3.45	2.28	1.04	0.30
	3 quarters	4.43	3.26	1.96	0.74	0.26
	4 quarters	2.99	2.16	1.01	0.41	0.33
Yen	1 quarter	0.94	0.79	0.59	0.39	0.24
	2 quarters	1.15	0.99	0.70	0.48	0.32
	3 quarters	1.50	1.32	0.93	0.70	0.46
	4 quarters	2.02	1.87	1.46	1.05	0.68
Pound	1 quarter	0.36	0.34	0.25	0.17	0.10
	2 quarters	0.44	0.48	0.40	0.30	0.19
	3 quarters	0.60	0.68	0.56	0.42	0.27
	4 quarters	0.76	0.85	0.71	0.54	0.36

Notes: This Table reports the root mean square forecast of exchange rate changes from Bayesian Model Averaging. The exchange rate was transformed by taking logs and then multiplying by 100, so the elements in this table can be interpreted as approximate percentage point forecast changes.

Table 5: Proportion of Times Bayesian Model Averaging Predicts Correct Sign of Exchange Rate Change

Currency	Horizon	φ=20	φ=5	φ=2	φ=1	φ=0.5
		Monthly Financial Data				
Canadian $	3 months	0.58	0.58	0.58	0.58	0.58
	6 months	0.68	0.68	0.68	0.68	0.68
	9 months	0.68	0.68	0.68	0.68	0.68
	12 months	0.74	0.74	0.74	0.74	0.74
Mark/Euro	3 months	0.58	0.58	0.58	0.60	0.63
	6 months	0.63	0.63	0.64	0.67	0.73
	9 months	0.67	0.68	0.68	0.73	0.73
	12 months	0.66	0.68	0.71	0.74	0.54
Yen	3 months	0.44	0.43	0.44	0.41	0.41
	6 months	0.55	0.53	0.48	0.48	0.48
	9 months	0.52	0.49	0.49	0.49	0.49
	12 months	0.49	0.49	0.49	0.49	0.49
Pound	3 months	0.52	0.52	0.53	0.53	0.53
	6 months	0.48	0.48	0.48	0.48	0.48
	9 months	0.50	0.50	0.50	0.50	0.50
	12 months	0.51	0.51	0.51	0.51	0.53
		Quarterly Data				
Canadian $	1 quarter	0.63	0.58	0.60	0.63	0.63
	2 quarters	0.68	0.68	0.70	0.70	0.70
	3 quarters	0.63	0.63	0.63	0.65	0.65
	4 quarters	0.58	0.63	0.63	0.63	0.63
Mark/Euro	1 quarter	0.55	0.55	0.55	0.55	0.60
	2 quarters	0.63	0.63	0.63	0.65	0.78
	3 quarters	0.63	0.63	0.65	0.73	0.68
	4 quarters	0.68	0.70	0.75	0.63	0.48
Yen	1 quarter	0.63	0.63	0.63	0.58	0.55
	2 quarters	0.55	0.53	0.53	0.50	0.48
	3 quarters	0.53	0.53	0.53	0.50	0.50
	4 quarters	0.48	0.48	0.48	0.45	0.40
Pound	1 quarter	0.55	0.55	0.55	0.58	0.53
	2 quarters	0.48	0.48	0.48	0.48	0.48
	3 quarters	0.50	0.50	0.50	0.50	0.50
	4 quarters	0.50	0.50	0.50	0.50	0.50

Notes: Asymptotic standard errors can be obtained from the formula for the variance of a binomial distribution, adjusting for the overlapping forecasts. The standard errors so obtained vary, but are approximately 0.08, 0.11, 0.14 and 0.16 at horizons of 1, 2, 3 and 4 quarters ahead, respectively.

Table 6: Relative Out-of-sample MSPE for Bayesian Model Averaging with All Permutations of Regressors as Models

Currency	Horizon	ρ=0.2	ρ=0.1	ρ=0.05
		Monthly Financial Data		
Canadian $	3 months	0.962	0.974	0.984
	6 months	0.913	0.945	0.967
	9 months	0.881	0.925	0.955
	12 months	0.858	0.912	0.948
Mark/Euro	3 months	0.938	0.945	0.956
	6 months	0.859	0.898	0.935
	9 months	0.857	0.915	0.955
	12 months	0.860	0.928	0.966
Yen	3 months	0.990	0.993	0.995
	6 months	0.983	0.988	0.993
	9 months	0.975	0.984	0.991
	12 months	0.986	0.991	0.995
Pound	3 months	1.013	1.004	1.001
	6 months	1.043	1.017	1.006
	9 months	1.040	1.013	1.004
	12 months	1.029	1.006	1.000
		Quarterly Data		
Canadian $	1 quarter	0.946	0.942	0.950
	2 quarters	0.923	0.934	0.949
	3 quarters	0.895	0.914	0.935
	4 quarters	0.927	0.933	0.946
Mark/Euro	1 quarter	0.936	0.937	0.944
	2 quarters	0.810	0.837	0.880
	3 quarters	0.785	0.849	0.912
	4 quarters	0.809	0.888	0.946
Yen	1 quarter	0.983	0.984	0.988
	2 quarters	0.954	0.965	0.978
	3 quarters	0.925	0.950	0.972
	4 quarters	0.949	0.969	0.985
Pound	1 quarter	0.947	0.965	0.981
	2 quarters	1.095	1.046	1.020
	3 quarters	1.106	1.045	1.018
	4 quarters	1.115	1.045	1.017

Notes: This Table reports the out-of-sample mean square prediction error from the forecasts taken by Bayesian Model Averaging over all possible permutations of regressors (as described in the text) relative to the mean square prediction error from a driftless random walk forecast.

Fig. 1: Out-of-sample Bayesian Model Averaging RMSPE for all currency-horizon pairs in two subsamples