University of Innsbruck

Working Papers in Economics and Statistics

Forecasting euro exchange rates: How much does model averaging help?

Jesus Crespo Cuaresma

2007-24

Forecasting euro exchange rates: How much does model averaging help?*

Jesus Crespo Cuaresma†

Abstract

We analyze the performance of Bayesian model averaged exchange rate forecasts for euro/US dollar, euro/Japanese yen, euro/Swiss franc and euro/ British pound rates using weights based on the out-of-sample predictive likelihood. The paper also presents a simple stratified sampling procedure in the spirit of Sala i Martin *et alia* (2004) to obtain model weights based on predictive accuracy. Our results indicate that accounting explicitly for model uncertainty when constructing predictions of euro exchange rates leads to improvements in predictive accuracy as measured by the mean square forecast error. While the forecasting error of the combined forecast tends to be systematically smaller than that of the individual model that would have been chosen based on predictive accuracy in a test sample, random walk forecasts cannot be beaten significantly in terms of squared forecast errors. Direction of change statistics, on the other hand, are significantly improved by Bayesian model averaging.

Keywords: Forecasting, model averaging, Bayesian econometrics, exchange rates.

JEL Classifications: C11, C53, F31.

*The author is indebted to Helmut Berrer for suggestions that helped improve the paper.

†Department of Economics, University of Innsbruck, Universitätstrasse 15, 6020 Innsbruck, Austria. Email: `jesus.crespo-cuaresma@uibk.ac.at`.

1 Introduction

Since the influential paper by Meese and Rogoff (1983), which showed the predictive superiority of a random walk model over exchange rate determination models, a myriad of studies have been written evaluating the forecasting ability of multiple modelling techniques for exchange rates. Although modern developments in the specification of cointegration and error correction models have led to some progress in improving out-of-sample forecasts for relatively long horizons (see for example, MacDonald and Taylor, 1993, MacDonald and Taylor, 1994, or Mark, 1995), this evidence is not as robust as many authors seem to claim (see Faust *et alia*, 2003). In particular, models that perform well for a given currency pair and sample do not necessarily deliver good forecasts for different currencies, or do not predict well in other subsamples.

Most of the studies on forecasting models for exchange rates concentrate on a model or set of models, whose predictive ability is compared with that of simple univariate time series models, prominently the random walk model. This methodological approach ignores systematically a dimension of uncertainty, namely that related to the choice of variables which enters the model. Most econometric exchange rate models take the theoretical structure behind the specification (which gives rise to the choice of variables) for granted and therefore do not exploit the predictive improvement which may be caused by combining forecasts from different models.

Bates and Granger (1969) and Newbold and Granger (1974) initiated the literature on forecast combination, which is extensively surveyed in recent contributions by Hendry and Clements (2004) and Timmermann (2006). Bayesian model averaging (BMA) presents a systematic methodology which integrates in a statistically solid framework the determination of weights for such combinations. In this paper we analyze the forecasting ability of BMA of exchange rate forecasts based on the out-of-sample predictive likelihood for the exchange rate of the euro against the US dollar, British pound, Japanese yen and Swiss franc. This method weights the forecasts of different exchange rate models based on their predictive ability and not on in-sample fit, like standard BMA techniques do. Wright (2003) presents results based on "*classical*" Bayesian model averaging (based on in-sample fit) for a relatively small set of exchange rate models, with results which are supportive of the averaging technique but not too impressive. Recently, BMA methods based on predictive likelihood have been studied theoretically by Eklund and Karlsson (2007) and applied to forecasts of Swedish inflation by Jacobson and Karlsson (2004). While these studies apply Markov Chain Monte Carlo methods in order to exploit the model space efficiently, here we present a new method to obtain the model weights based on the sampling procedure put forward by Sala-i-Martin *et alia* (2004).

Our results for the EUR/USD, EUR/JPY, EUR/CHF and EUR/GBP exchange rates indicate that accounting explicitly for model uncertainty when constructing predictions of exchange rates tends to lead to improvement over the use of the single best forecasting model in terms of predictive accuracy, although the results of the averaged forecasts perform poorly compared to the traditional random walk benchmark. The forecast averaging technique also tends to improve over the single best model in terms of direction of change statistics, reaching values which are significantly over 50% in several cases.

The paper is structured as follows. Section 2 presents the methodology of BMA using the out-of-sample predictive likelihood and describes the sampling procedure used. Section 3 gives the results of the forecasting competition for euro exchange rates

2 Bayesian model averaging using the out-of-sample predictive likelihood

2.1 Bayesian model averaging and exchange rate models

The empirical literature on exchange rate forecasting tends to concentrate on a given theoretical framework that determines the nature of the variables to be used in the econometric specification (and most probably, also the functional relationship linking them). In the case of the monetary model of exchange rate determination (see for example Frenkel, 1976 or Dornbush, 1976, for the original formulations), the variables that should be included in the empirical model depends on the equilibrium conditions assumed in theory. If the uncovered interest rate parity (UIRP) is not assumed to hold, for instance, then interest rate differentials could play a role in the determination of exchange rates. On the other hand, if the UIRP is assumed to be fulfilled, the interest rate differential contains information on exchange rate expectations and the extra assumption of rational expectations would imply that the interest rate variable should not be included in the econometric specification (see for example the derivations in Groen, 2002). The same way, while macroeconomic models aimed at exchange rate prediction concentrate on monetary variables and do not tend to use data on financial markets, models using variables such as stock indices have shown good predictive power at similar forecasting horizons (Chu and Lu, 2006, is a recent example of this literature).

In this paper we propose averaging over different alternative models using Bayes factors so as to evaluate the relative importance of different variables as predictors of the exchange rate. In the situation where there are M competing models, $\{M_1, \ldots, M_M\}$ Bayesian inference about the quantity of interest, which in our case will be the predictive density at forecasting horizon h, $P(\mathbf{y}_h)$ is based on its posterior distribution (that is, the distribution given the data, $\mathbf{Y}$),

$$\mathrm{P}(\mathbf{y}_h|\mathbf{Y}) = \sum_{m=1}^{M} \mathrm{P}(\mathbf{y}_h|\mathbf{Y}, M_m)\mathrm{P}(M_m|\mathbf{Y}), \tag{1}$$

where $\mathrm{P}(M_k|\mathbf{Y})$ are the posterior model probabilities,

$$\mathrm{P}(M_k|\mathbf{Y}) = \frac{\mathrm{P}(\mathbf{Y}|M_k)\mathrm{P}(M_k)}{\sum_{m=1}^{M} \mathrm{P}(\mathbf{Y}|M_m)\mathrm{P}(M_m)}. \tag{2}$$

The posterior model probabilities can thus be obtained as the normalized product of the integrated likelihood for each model ($\mathrm{P}(\mathbf{Y}|M_k)$) and the prior probability of the model ($\mathrm{P}(M_k)$). Notice that for the simple case $m = 2$ the posterior odds for a model against the other can be readily written as the product of the Bayes factor and the prior odds. Further assuming equal priors across models, the posterior odds are equal to the Bayes factor ($\mathrm{P}(\mathbf{Y}|M_2)/\mathrm{P}(\mathbf{Y}|M_1)$).The Bayes factor, in turn, can be accurately approximated (see Leamer, 1978, and Schwarz, 1978) as

$$\frac{\mathrm{P}(\mathbf{Y}|M_2)}{\mathrm{P}(\mathbf{Y}|M_1)} = N^{(k_1-k_2)/2}\left(\frac{Lik_2}{Lik_1}\right), \tag{3}$$

where N is the number of observations, k_j and Lik_j are respectively the number of parameters and the likelihood of model j. This simple approximation allows us to compute (2) and the corresponding statistics based on (2).

Since our only interest in the exercise is prediction, we redefine the posterior model probabilities based on the predictive densities of the models being entertained for the forecasting horizon considered, h. This approach to model averaging based on out-of-sample predictive likelihoods instead of in-sample fit has been recently proposed by Kapetanios *et alia* (2006), for instance. In practice, this amounts to replacing the in-sample residuals by out-of-sample forecasting errors in (2) when computing the corresponding likelihood. The forecasting errors are obtained from a model estimated on a subsample of the available data, which is used in order to predict the remaining sample. The corresponding Bayes factor can thus be approximated by

$$\frac{\mathrm{P}(\mathbf{Y}|M_2)}{\mathrm{P}(\mathbf{Y}|M_1)} = T_F^{(k_1-k_2)/2}\left(\frac{MSFE_1}{MSFE_2}\right)^{\frac{T_F}{2}}, \tag{4}$$

where $MSFE_j$ is the mean square forecasting error of model j based on the T_F out-of-sample observations, that is,

$$MSFE_k(h) = \sum_{j=0}^{T_F-h} (y_{T_0+j+h} - y^f_{T_0+j+h})^2/(T_F - h), \tag{5}$$

where y_t refers to the exchange rate in period t, y_t^f is the corresponding forecast, obtained with data up to period $t-h$, and T_0 is the last in-sample observation, so that $T_0 + h$ is the first observation to be forecast in the subsample used to obtain the model weights.

This implies that for a given prior on the model space, the posterior distribution of $\mathbf{y}_h$ can be obtained as a weighted average of the model-specific estimates weighted by the posterior probability of the respective models. If the cardinality of the model space is computationally tractable, (4) can be obtained directly and (1) can be computed. In particular, the expected value of $\mathbf{y}_h$, $\mathrm{E}(\mathbf{y}_h|\mathbf{Y})$, the point forecast, can be computed as follows

$$\mathrm{E}(\mathbf{y}_h|\mathbf{Y}) = \sum_{m=1}^{M} \mathrm{E}(\mathbf{y}_h|\mathbf{Y}, M_m)\mathrm{P}(M_m|\mathbf{Y}), \tag{6}$$

Several methods can be used in order to approximate the expression in (2) when the cardinality of the model space makes the problem intractable. The *leaps and bounds* algorithm, the use of Markov chain Monte Carlo model composite (MC^3) methods or the use of Occam's window are possible methods of setting bounds to the number of models to be evaluated when computing (2).[1] In the empirical application presented below, the number of models is given by all possible combinations of 16 potential variables in three potential functional forms (given by multivariate time series models in the form of vector autoregressions in levels, VAR, in first differences, DVAR or error correction models, VEC) with a lag length between one and six lags. This results in $6 \times 3 \times 2^{16} = 1,179,648$ possible models. In this paper we propose a simple sampling method based on the stratified sampler proposed by Sala-i-Martin *et alia* (2004) in order to evaluate the sums in (2).

In the same fashion, posterior inclusion probabilities for the different variables can be obtained by summing the posterior probability of models containing each variable. This measure captures, thus, the relative importance of the different variables as predictors of exchange rate movements.

[1]See the influential papers by Raftery (1995) and Raftery *et alia* (1997) for a discussion of these methods and a general introduction to Bayesian model averaging.

2.2 Sampling procedure

Sala-i-Martin *et alia* (2004) propose a simple sampling procedure aimed at evaluating the expression given by (2) when the cardinality of the model space makes the computation of the total sum in the denominator intractable. Here we use a generalization of this procedure based on sampling from the model space assigning more weight to models which tend to deliver relatively good predictions of the exchange rate.

The stratified sampling procedure is carried out as follows. Let the total sample be divided into the following subsamples: observations 1 to T_0 correspond to the first in-sample set and observations $T_0 + 1$ to $T_0 + T_F$ are used to obtain the forecasting errors. Using the prior distribution of models, $P(M_j)$, a model specification (given by a set of variables entering the model, a specification of the relationship among variables - VAR, DVAR or VEC - and a lag length) is chosen and estimated for the first in-sample period, $(1, T_0)$. A prediction for the exchange rate observation corresponding to period $T_0 + h$ is computed, the model is reestimated for the sample $(1, T_0 + 1)$ and a new forecast is obtained for the observation corresponding to $T_0 + 1 + h$. This procedure is repeated until the prediction for period $T_0 + T_F$ is achieved. The corresponding squared forecast errors are obtained for this model, a new model is sampled and the procedure is repeated. This is done a large number of times. In order to avoid sampling many models with poor forecasting ability, the sampling probabilities are updated every N replications using the posterior model probabilities computed up to that replication. The corresponding updated sampling probabilities are a linear combination of the prior and the posterior model probabilities weighted by a factor ω and $(1 - \omega)$, respectively. The full procedure is repeated until convergence is achieved in the object to be estimated (in our case, $\mathrm{E}(\mathbf{y}_h|\mathbf{Y})$).

3 Empirical results

3.1 Data and models

The forecasting exercise is carried out for the exchange rate of the euro (EUR) against the US dollar (USD), Japanese yen (JPY), British pound (GBP) and Swiss franc (CHF). The full dataset spans the period January 1980 - January 2006 at monthly frequency, where the euro exchange rate prior to 1999 refers to synthetic euro data. Table 1 presents the definitions of the variables used as potential covariates in the multivariate time series models. Each variable is considered both for the domestic (euro area) and corresponding foreign economy, so that 16 potential variables are considered for each currency. This set contains the usual macroeconomic variables implied by the monetary model of exchange rate determination, financial market series such as stock and earning indices and sentiment indicators. The source of the data is Datastream in all cases.

Include Table 1 here

The multivariate time series specifications used have the following form. For a given group of k variables (which are grouped together with the exchange rate variable in the vector X_t, which is thus of dimension $k+1$) and lag length P, the VAR specification is given by

$$X_t = \Gamma_0 + \sum_{i=1}^{P} \Gamma_i X_{t-i} + \varepsilon_t, \tag{7}$$

where Γ_0 is a $(k+1)\times 1$ vector of intercept terms, Γ_i for $i = 1, \ldots, P$ are $(k+1)\times(k+1)$ matrices of parameters and ε_t is assumed to be a vector disturbance with expected value zero and variance-covariance matrix Σ. The DVAR specification is given by

$$\Delta X_t = \Gamma_0 + \sum_{i=1}^{P} \Gamma_i \Delta X_{t-i} + \varepsilon_t, \tag{8}$$

where Δ is the first difference operator, $\Delta = (1 - L)$, where L is the lag operator. The VEC specification is

$$\Delta X_t = \Gamma_0 + \alpha\theta X_{t-1} + \sum_{i=1}^{P} \Gamma_i \Delta X_{t-i} + \varepsilon_t, \tag{9}$$

where θ is a $1\times(k+1)$ vector which identifies the cointegration relationship (for simplicity we consider exclusively VEC models with a single cointegration relationship) and α is a $(k+1)\times 1$ vector of adjustment parameters to the long-run relationship given by the cointegrating vector.

The model averaging procedure is carried out using the sampling procedure described above. In particular, we assume a prior inclusion probability of 0.25 for each variable considered, which implies that the prior expected number of included variables in the multivariate model is 4. A uniform prior is assumed over the lag length, ranging from one to six, and also a uniform prior is assumed over the model specification (VAR, DVAR or VEC).[2] It should be noticed that the model space also includes univariate autoregressive

[2] For an inclusion probability π for each variable, the probability of a model including s variables is $\pi^s(1-\pi)^{S-s}$, where S is the total number of variables considered. If the prior probability attached to a lag length l is given by π_l (assumed equal across lag lengths) and the prior probability attached to a model specification (VAR, DVAR, VEC) is 1/3, the prior probability of model i with s variables is $P(M_i) = \frac{1}{3}\pi_l \pi^s (1-\pi)^{S-s}$.

time series models (both in levels and first differences), which correspond to sampling zero variables from the potential set. The results are reported for 100,000 replications of the stratified sampler, where the sampling probabilities are updated every 100 replications with a mixing factor $\omega = 0.8$.[3] We will consider that the sampling procedure has converged if the absolute difference between the posterior expected value of the exchange rate vector defined for the period (T_{F+h}, T) in replication g and replication $g + 500$ is smaller than 0.01%. The first in-sample period is defined to be between January 1980 and December 1990 (T_0 in the notation above) and the period January 1991 - December 1998 ($T_0 + T_F$ in the notation above) will be used to obtain the model weights and named the "test sample". Based on these weights, the period January 1999 - January 2006 is used to evaluate the out-of-sample forecasting ability of the model averaging technique. Forecasts are obtained for horizons ranging from one month to one year ahead.

3.2 Forecasting results

The results of the forecasting exercise are presented in Table 2 and Table 3. For each currency and each forecasting horizon these tables presents the relative root mean square forecasting error (RMSFE) of the model average as compared to the random walk model (no-change forecast) for the period January 1999 - January 2006, together with the relative root mean square forecasting error for the same period corresponding to the individual (sampled) model with the lowest forecasting error in the period January 1991 - December 1998 (named "*Best Model*" in the table). The ratios are built so that values above one imply better forecasting ability of the random walk. Direction of change statistics (DOC), defined as the proportion of times that a depreciation or appreciation was correctly forecast are also provided in the tables. We also provide the Diebold-Mariano test statistic (Diebold and Mariano, 1995) and the binomial test statistic for the null hypothesis that the direction of change probability is equal to 0.5 (hypothesis corresponding to the random walk model). The identity of the best sampled model is given in the Table by its specification (AR, VAR, DVAR or VEC), followed by the identity of the variables included in the model and the lag length. The names of the variables corresponds to the names given in Table 1, the superscript d indicates the domestic economy (euro area) and the superscript f indicates the corresponding foreign economy.

Include Table 2 and Table 3 here

[3] As in Sala i Martin *et alia* (2004), we will impose a minimal sampling probability for a variable of 0.1 and a maximum sampling probability of 0.8, so as to avoid that certain variables are sampled too seldom (see the Technical Appendix to Sala i Martin *et alia* (2004)).

The results presented in Table 2 and 3 indicate a very relevant improvement in terms of RMSFE when combining forecasts as compared to the model that would have been chosen based on the forecasting ability in the test sample. This is true for all forecasting horizons in the case of the EUR/USD exchange rate, and all but one in the case of the EUR/JPY (8 months ahead). The single best model improves over the model averaged forecast at two and three forecasting horizons for the EUR/CHF and EUR/GBP rates respectively (at horizons 9 and 10 in the first case and 5, 9 and 11 in the second case). Despite improvements over single best models, the combined forecasts do not appear systematically and significantly better than random walk forecasts. Only for the EUR/CHF forecasts do the BMA predictions improve over the random walk benchmark and pass the Diebold-Mariano test. On the other hand, the evidence against the best single models gathered by comparing forecasts with the random walk model is overwhelming: the Diebold-Mariano test favours the no-change forecast in most of the comparisons.

The results concerning DOC give clear evidence of the supremacy of BMA forecasts for the EUR/USD exchange rate over the forecasts produced by the individual best model. In a couple of cases, the BMA forecasts present DOC forecasts for this exchange rate which appear significantly over the 0.5 benchmark. The DOC results for the EUR/JPY rate also favour he BMA technique strongly: with the exception of the 8 and 9 months-ahead forecasts, averaging improves directional forecasts, rendering the accuracy of long-run forecasts (above 9 months-ahead) significantly above 0.5.

Correctly predicted directions of change in the case of average forecasts appear significantly better than the "toss of coin" benchmark in short-run forecasts (2 and 3 months ahead) for the EUR/GBP exchange rate, where improvements over the single best model appear in most cases, and in long-run forecasts (12 months ahead) for the EUR/CHF rate. In this last case the proportion of correctly forecast directions of change for the BMA technique, furthermore, appears greater than 0.7, although the results for other forecasting horizons in this exchange rate do not tend to systematically support model averaging.

Include Figure 1, Figure 2 and Figure 3 here

In Figure 1 we present the posterior inclusion probabilities of the variables in Table 1, Figure 2 presents the inclusion probabilities of the different model specifications and Figure 3 the posterior probabilities of the lag length parameter for the different exchange rates at the forecasting horizons considered. For the interpretation of Figure 1, it should be stressed that the prior model inclusion probabilities of the different variables equals

0.25 (see section 3.1 above). This implies that posterior inclusion probabilities above 0.25 in Figure 1 indicate that, after observing the data, we consider the inclusion of that variable in the model more probable than a priori assumed. Several interesting conclusions can be drawn from these figures. Probably the most striking feature is the heterogeneiy observed across exchange rates. While no variable can be labelled robust (not even in the sense of attaining posterior inclusion probabilities above the prior) for the EUR/USD and EUR/GBP exchange rates, some variables attain high posterior inclusion probabilities for the EUR/JPY and EUR/CHF. This implies that the big bulk of posterior inclusion probability for the EUR/USD and the EUR/GBP is concentrated on univariate specifications of the exchange rate, regardless of the fact that the individual models that receive the highest weight in the BMA procedure tend to be multivariate models. The importance of univariate models mirrors itself in extremely low posterior inclusion probabilities for error correction models for thee exchange rates, as shown in Figure 2. Among vector autoregressive models, the posterior inclusion probabilities of models in levels versus models in first difference tends to be both currency-specific and forecasting horizon-specific. Interestingly, both in the EUR/USD and the EUR/JPY exchange rate, for short horizons the statistical evidence changes from favoring models in first differences to models in levels in a relatively monotonic manner.

The results for the EUR/JPY exchange rate provide evidence of the importance of including money supply and (eurozone) industrial production variables in error correction specifications for medium-term forecasts (8 and 9 months ahead). As can be seen by comparing the results with those in Table 2, this result is driven by the superior performance of the individual best model, whose predictions appear systematically better than those of the averaged alternative.

For the case of the EUR/CHF exchange rate, robust variables (with very high posterior inclusion probability) are present only for prediction horizons over 6 months ahead. Within this range of medium to long-term forecasts, the results concerning the relative importance of the different macroeconomic variables changes depending on the forecasting horizon considered. While eurozone industrial production and the corresponding Swiss stock market index seem to be relevant for the quality of 7 to 11 months-ahead forecasts, earning indices, the Swiss short term interest rate and the eurozone sentiment indicator appear as robust predictors for one year ahead forecasts. It is noticeable, however, that the improvements brought about by model averaging for these horizons are not systematic, as shown in Table 1.

The inclusion probabilities for the different lag lengths of the time series models entertained, presented in Figure 3, favor relatively small models for all exchange rates considered, with lag lengths over three months having negligible posterior inclusion probabilities for all exchange rates studied.

4 Conclusions

We analyze the performance of Bayesian model averaged exchange rate forecasts using weights based on the out-of-sample predictive likelihood. Our results for the EUR/USD, EUR/JPY, EUR/CHF and EUR/GBP exchange rates indicate that accounting explicitly for model uncertainty when constructing predictions of exchange rates tends to lead to improvement over the use of the single best forecasting model in terms of predictive accuracy. While the forecasting error of the combined forecast tends to be systematically smaller than that of the individual model which would have been chosen based on predictive accuracy in a test sample, the results of the averaged forecasts perform poorly compared to the traditional random walk benchmark. Although improvements over this baseline are observed at some forecasting horizons and exchange rates, none of them is significant using the Diebold-Mariano test.

The results are more promising when comparing forecasts in terms of direction of change. In this case, the BMA alternative tends to improve over the single best model and presents direction of change statistics significantly over 50% in several cases.

The evaluation of posterior inclusion probabilities of the variables used in the Bayesian model averaging exercise reveals that the set of univariate specifications appears still important for prediction, in particular for the EUR/USD and EUR/GBP exchange rates. None of the variables in the set attains a posterior inclusion probability exceeding the prior inclusion probability for these two exchange rates, and error correction models present negligible inclusion probabilities. The variables which appear relevant for the EUR/JPY and EUR/CHF exchange rates, however, appear to be dependent on the forecasting-horizon considered. Furthermore, the evidence of out-of-sample performance of (best) models based on these variables does not seem too convincing. These results reinforce the common view that finding fundamental variables which can be used for exchange rate prediction is a hard task to undertake, and that the criticism of exchange rate models embodied in Meese and Rogoff's (1983) results is still intact after almost 25 years.

References

[1] Bates, J. M. and C. W. J. Granger (1969). The Combination of Forecasts, *Operations Research Quarterly*, 20, 451468.

[2] Diebold, F.X. and R.S. Mariano (1995). Comparing Predictive Accuracy. *Journal of Business and Economic Statistics*, 13, 253-263.

[3] Dornbusch R. (1976). Expectations and Exchange Rate Dynamics. *Journal of Political Economy*, 84, 11611176.

[4] Eklund, J. and S. Karlsson (2007). Forecast Combination and Model Averaging Using Predictive Measures, *Econometric Reviews*, 26, 329 363.

[5] Faust, J., Rogers, J. H. and J. H. Wright (2003). Exchange Rate Forecasting: The Errors Weve Really Made, *Journal of International Economics*, 60, 35-60.

[6] Frenkel J. (1976). A Monetary Approach to the Exchange Rate: Doctrinal Aspects and Empirical Evidence. *Scandinavian Journal of Economics*, 78, 200224.

[7] Granger, C. W. J. and P. Newbold (1974). Experience with Forecasting Univariate Time Series and the Combination of Forecasts, *Journal of the Royal Statistical Society. Series A*, 137, 131165.

[8] Groen, J. J. J. (2002). Cointegration and the Monetary Exchange Rate Model Revisited, *Oxford Bulletin of Economics and Statistics*, 64, 361-380.

[9] Hendry, D. F. and M. P. Clements (2004). Pooling of Forecasts, *The Econometrics Journal*, 7, 1-31.

[10] Chu C.-S. J. and H.-M. Lu (2006). Random Walk Hypothesis in Exchange Rate Reconsidered, *Journal of Forecasting*, 25, 275-290.

[11] Jacobson T. and S. Karlsson (2004). Finding Good Predictors for Inflation: a Bayesian Model Averaging Approach, *Journal of Forecasting*, 23, 479-496.

[12] Kapetanios G., Labhard V. and S. Price (2006). Forecasting Using Predictive Likelihood Model Averaging *Economics Letters*, 91, 373379.

[13] Leamer, E. E. (1978). *Specification Searches.* New York, John Wiley & Sons.

[14] MacDonald, R. and M. P. Taylor (1993). The Monetary Approach to the Exchange Rate: Rational Expectations, Long-Run Equilibrium and Forecasting, *IMF Staff Papers*, 40, 89107.

[15] MacDonald, R. and M. P. Taylor (1994). The Monetary Model of the Exchange Rate: Long-Run Relationships, Short-Run Dynamics, and How to Beat a Random Walk, *Journal of International Money and Finance*, 13, 276-90.

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

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

[18] Raftery, A.E. (1995). Bayesian Model Selection in Social Research. *Sociological Methodology*, 25, 111-196.

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

[20] Sala i Martin, X. Doppelhofer, G. and R. Miller (2004). Determinants of Long-Term Growth: A Bayesian Averaging of Classical Estimates (BACE) Approach, *American Economic Review*, 94, 813-835.

[21] Schwarz, G. (1978) Estimating the Dimension of a Model. *Annals of Statistics*, 6, 461-464.

[22] Timmermann, A. (2006). Forecast Combinations, in Elliot, G., Granger, C. W. J. and A Timmermann (eds.) *Handbook of Economic Forecasting*, 1. Amsterdam, Elsevier North Holland.

[23] Wright, J.H. (2003). Bayesian Model Averaging and Exchange Rate Forecasts, *Board of Governors of the Federal Reserve System, International Finance Discussion Papers No. 779.*

Variable	Definition and transformation
Exchange rate (y)	
Sentiment indicator (x_1)	Standardized leading indicator indices: Ifo-Index (euro area), ISM index (US), CBI index (UK), Kof index (Switzerland), Leading economic indicator (Japan)
Money supply (x_2)	M0 or M1 depending on availability, logged
Short-term nominal interest rate (x_3)	3 month interbank offer rate
Industrial production (x_4)	Industrial production index, logged
Stock index (x_5)	Stock market, Datastream Indices-Standardized Indices, logged
Earning index (x_6)	Earnings from stock market, Datastream Indices-Standardized Indices, logged
Long-term nominal interest rate (x_7)	10 year bond rate (benchmark)
Short-term real interest rate (x_8)	3 month interbank offer rate minus inflation rate

Table 1: Variables in the Bayesian model averaging exercise

Steps ahead	RMSFE (BMA)	D-M test (BMA)	RMSFE (Best Model)	D-M test (Best Model)	DOC (BMA)	Binomial test p=0.5 (BMA)	DOC (Best Model)	Binomial test p=0.5 (Best Model)	Best Model Specification
EUR/USD									
1	0.990	-0.299	1.063	1.183	0.571	1.309	0.536	0.655	VAR(x_2^d,x_2^f,x_5^d,x_7^d), lags= 3
2	1.010	0.418	1.175	2.490	0.482	-0.329	0.422	-1.427	VAR(x_2^d,x_5^d,x_7^d,x_8^d), lags= 2
3	1.012	0.352	1.337	2.680	0.537	0.663	0.415	-1.546	VEC(x_2^d,x_4^d,x_6^d), lags= 5
4	1.006	0.560	1.348	2.456	0.420	-1.444	0.370	-2.333	VAR(x_2^d,x_4^d), lags= 2
5	1.008	0.692	1.518	3.279	0.488	-0.224	0.425	-1.342	VAR(x_1^d,x_2^d,x_2^f,x_4^d,x_7^d,x_8^d), lags= 5
6	1.009	0.550	1.326	2.263	0.506	0.113	0.443	-1.013	VEC(x_2^f,x_4^d,x_6^d,x_8^f), lags= 5
7	1.011	0.456	1.270	3.268	0.551	0.906	0.397	-1.812	DVAR(x_2^f,x_4^d,x_5^f,x_6^d,x_7^f), lags= 6
8	1.010	0.413	1.468	2.608	0.532	0.570	0.429	-1.254	VAR(x_1^d,x_2^d,x_2^f,x_3^d,x_4^d,x_5^d), lags= 5
9	1.007	0.296	1.334	3.701	0.592	1.606	0.289	-3.671	VEC(x_2^f,x_4^d,x_5^d,x_6^d), lags= 6
10	1.002	0.102	1.677	3.394	0.533	0.577	0.360	-2.425	VAR(x_2^d,x_4^d,x_6^d), lags= 6
11	1.001	0.029	1.888	3.830	0.595	1.627	0.365	-2.325	VEC(x_2^d,x_4^d), lags= 5
12	0.999	-0.032	1.695	3.193	0.562	1.053	0.411	-1.522	VAR(x_2^d,x_4^d), lags= 5
EUR/JPY									
Steps ahead	RMSFE (BMA)	D-M test (BMA)	RMSFE (Best Model)	D-M test (Best Model)	DOC (BMA)	Binomial test p=0.5 (BMA)	DOC (Best Model)	Binomial test p=0.5 (Best Model)	Best Model Specification
1	0.992	-0.225	1.013	0.340	0.560	1.091	0.548	0.873	DVAR(x_5^d), lags= 2
2	1.016	0.604	1.172	1.901	0.482	-0.329	0.446	-0.988	VAR(x_2^d,x_4^d), lags= 2
3	1.001	0.056	1.374	2.634	0.512	0.221	0.390	-1.988	VAR(x_2^d,x_4^d,x_7^f,x_8^f), lags= 5
4	1.008	0.403	1.348	2.456	0.481	-0.333	0.370	-2.333	VAR(x_2^d,x_4^d), lags= 2
5	1.008	1.036	1.529	2.991	0.475	-0.447	0.363	-2.460	VAR(x_2^d,x_4^d), lags= 5
6	1.006	0.363	1.696	3.669	0.544	0.788	0.354	-2.588	VEC(x_2^d,x_4^d), lags= 5
7	1.009	0.362	1.294	3.100	0.551	0.906	0.410	-1.585	VEC(x_2^f,x_4^d,x_6^d,x_7^f), lags= 6
8	1.280	1.868	1.134	1.101	0.390	-1.937	0.597	1.709	VEC(x_2^d,x_2^f,x_4^d,x_6^d), lags= 1
9	1.084	0.714	1.103	0.853	0.579	1.376	0.632	2.294	VEC(x_2^d,x_2^f,x_4^d,x_6^d), lags= 1
10	1.002	0.068	1.333	2.362	0.613	1.963	0.400	-1.732	VAR(x_2^d,x_4^d,x_5^d,x_6^d), lags= 2
11	1.000	0.016	1.321	2.307	0.595	1.627	0.338	-2.790	VEC(x_1^d,x_2^d,x_4^d,x_8^d), lags= 6
12	0.999	-0.032	1.429	2.239	0.616	1.990	0.425	-1.287	VEC(x_2^d,x_4^d), lags= 3

The D-M test statistic is asymptotically distributed as a standard normal under the null hypothesis of no difference in predictive accuracy.

The binomial test statistic is asymptotically distributed as a standard normal under the null hypothesis of $p = 0.5$.

Table 2: Forecasting results for euro/US dollar and euro/Japanese yen exchange rates

EUR/CHF									
Steps ahead	RMSFE (BMA)	D-M test (BMA)	RMSFE (Best Model)	D-M test (Best Model)	DOC (BMA)	Binomial test p=0.5 (BMA)	DOC (Best Model)	Binomial test p=0.5 (Best Model)	Best Model Specification
1	1.025	0.581	1.198	2.235	0.560	1.091	0.464	-0.655	DVAR(x_2^f,x_5^d,x_8^f), lags= 2
2	0.996	-0.405	1.017	0.138	0.542	0.768	0.614	2.086	VAR(x_6^d,x_6^f), lags= 1
3	0.974	-0.692	1.005	0.036	0.537	0.663	0.598	1.767	VAR(x_6^d,x_6^f), lags= 1
4	0.971	-1.471	1.000	-0.003	0.556	1.000	0.617	2.111	VAR(x_6^d,x_6^f), lags= 1
5	0.978	-0.975	1.291	1.235	0.463	-0.671	0.625	2.236	VAR(x_3^d,x_6^d,x_6^f,x_8^f), lags= 1
6	0.925	-1.788	0.994	-0.032	0.506	0.113	0.494	-0.113	VAR(x_6^d,x_6^f), lags= 1
7	1.069	1.137	1.650	2.805	0.513	0.226	0.449	-0.906	VAR(x_4^d,x_5^f,x_7^f), lags= 4
8	0.970	-0.179	1.089	0.448	0.623	2.165	0.662	2.849	VAR(x_1^d,x_6^d,x_6^f,x_8^f), lags= 1
9	1.121	1.452	1.119	0.898	0.553	0.918	0.487	-0.229	VEC(x_4^d,x_5^f,x_7^f), lags= 3
10	1.268	2.566	1.183	1.643	0.427	-1.270	0.573	1.270	VEC(x_2^d,x_4^d,x_7^d,x_8^f), lags= 1
11	1.288	3.565	1.322	3.595	0.324	-3.022	0.324	-3.022	VEC(x_4^d,x_5^f,x_7^f), lags= 1
12	0.998	-0.010	0.894	-0.727	0.740	4.096	0.616	1.990	VAR(x_1^d,x_2^d,x_6^d,x_6^f,x_8^f), lags= 1

EUR/GBP									
Steps ahead	RMSFE (BMA)	D-M test (BMA)	RMSFE (Best Model)	D-M test (Best Model)	DOC (BMA)	Binomial test p=0.5 (BMA)	DOC (Best Model)	Binomial test p=0.5 (Best Model)	Best Model Specification
1	1.002	0.046	1.016	0.431	0.548	0.873	0.464	-0.655	DVAR(x_3^d,x_5^d,x_5^f), lags= 2
2	0.988	-0.581	1.030	1.087	0.651	2.744	0.542	0.768	DVAR(x_2^f,x_3^d,x_5^f), lags= 3
3	1.000	0.009	1.161	2.134	0.598	1.767	0.488	-0.221	VEC(x_2^f,x_5^f,x_7^d), lags= 3
4	1.032	0.862	1.211	2.480	0.543	0.778	0.481	-0.333	VEC(x_2^f,x_5^d,x_7^d), lags= 3
5	1.036	1.137	1.024	0.679	0.513	0.224	0.538	0.671	VEC(x_2^f,x_3^d), lags= 1
6	1.035	1.196	1.107	2.173	0.494	-0.113	0.468	-0.563	DVAR(x_6^f,x_7^d), lags= 3
7	1.014	0.576	1.166	2.388	0.538	0.679	0.436	-1.132	VEC(x_4^f,x_6^f,x_7^d,x_8^f), lags= 4
8	1.006	0.286	1.095	1.960	0.532	0.570	0.494	-0.114	DVAR(x_5^f,x_6^f,x_7^d,x_8^f), lags= 3
9	1.010	0.402	1.000	-0.004	0.539	0.688	0.526	0.459	VEC(x_2^f,x_3^d), lags= 1
10	1.002	0.102	1.131	2.163	0.507	0.115	0.413	-1.501	VEC(x_1^d,x_2^f,x_3^d,x_4^d,x_6^d,x_6^f,x_7^d,x_8^d), lags= 1
11	1.019	0.648	0.973	-1.023	0.473	-0.465	0.541	0.697	VEC(x_2^f,x_3^f,x_5^f,x_6^f), lags= 1
12	1.019	0.847	1.166	2.026	0.452	-0.819	0.425	-1.287	VAR(x_1^d,x_3^f,x_7^d,x_7^f,x_8^f), lags= 1

The D-M test statistic is asymptotically distributed as a standard normal under the null hypothesis of no difference in predictive accuracy.

The binomial test statistic is asymptotically distributed as a standard normal under the null hypothesis of $p = 0.5$.

Table 3: Forecasting results for euro/Swiss franc and euro/British pound exchange rates

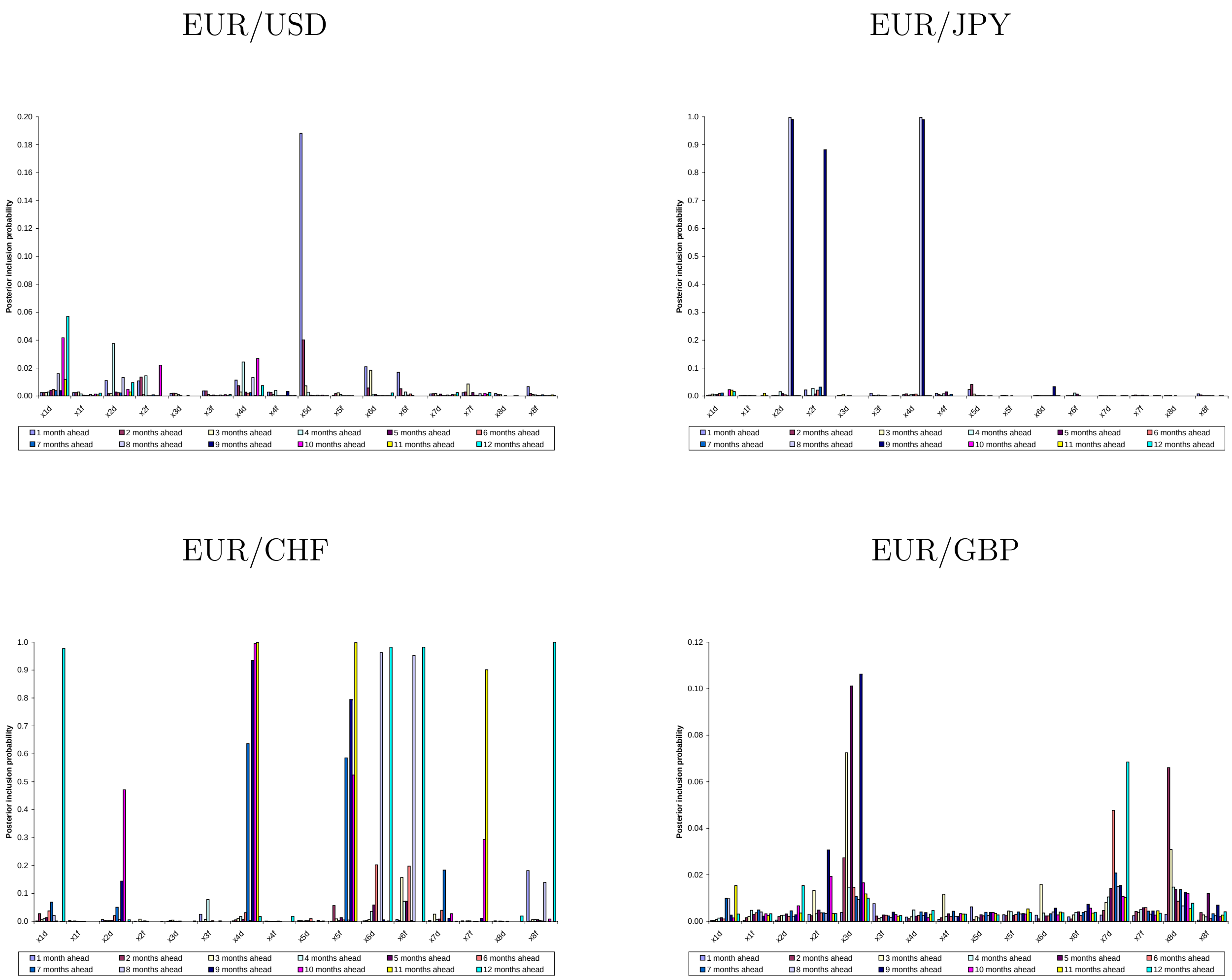
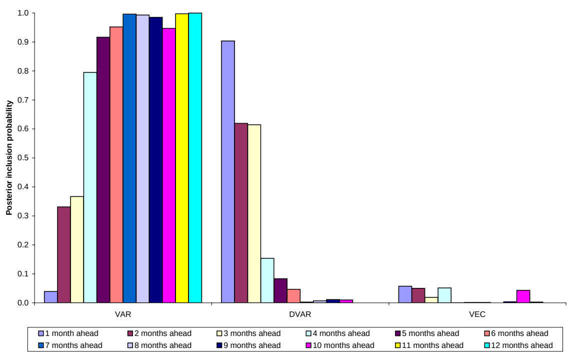
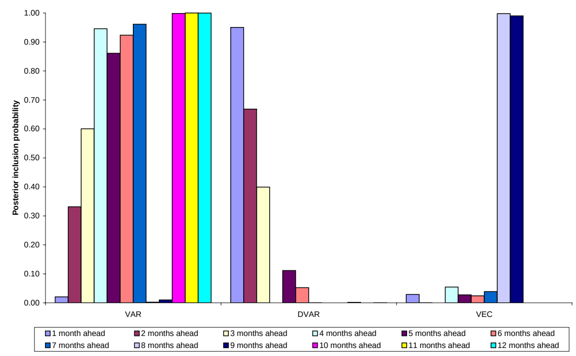


Figure 1: Posterior inclusion probability, variables

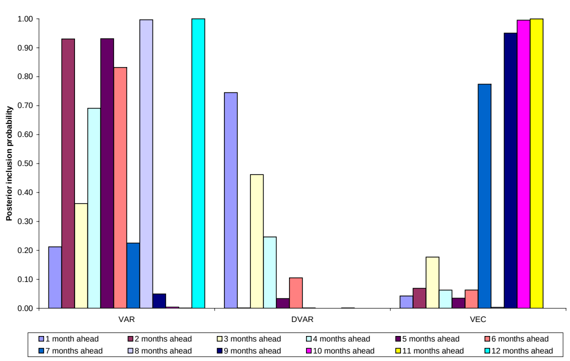
EUR/USD



EUR/JPY



EUR/CHF



EUR/GBP

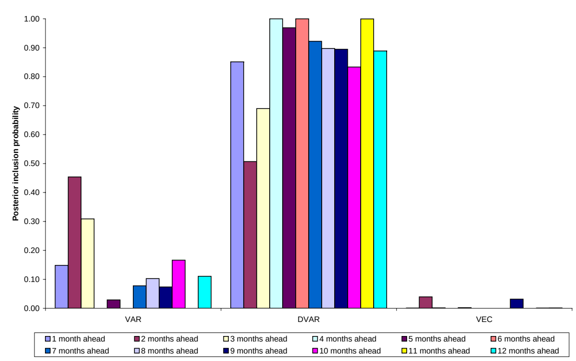


Figure 2: Posterior inclusion probability, model specification

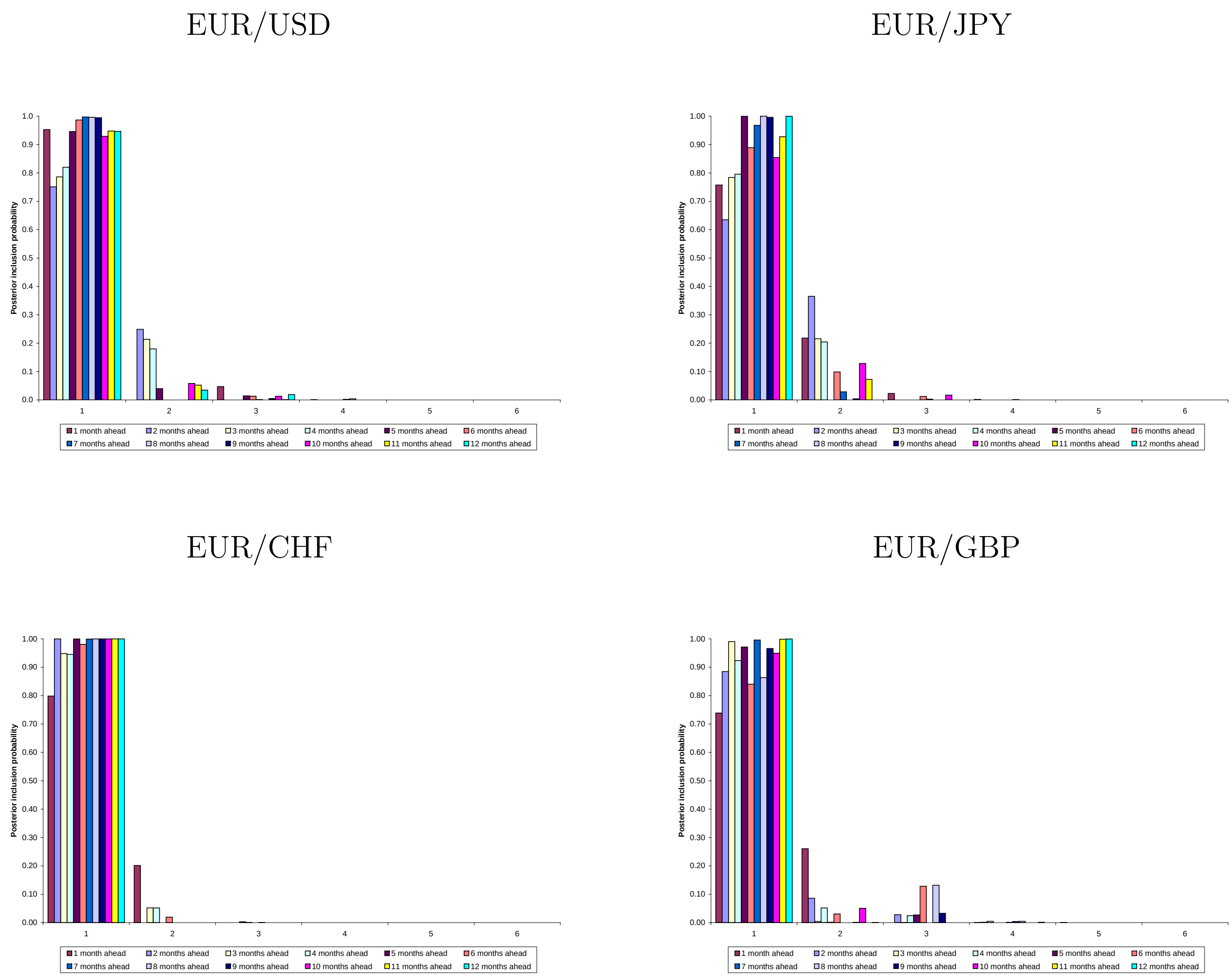


Figure 3: Posterior inclusion probability, lag length

University of Innsbruck – Working Papers in Economics and Statistics
Recent papers

2007-25 **Andrea M. Leiter, Harald Oberhofer and Paul A. Raschky:** Productive disasters? Evidence from European firm level data.
2007-24 **Jesus Crespo Cuaresma:** Forecasting euro exchange rates: How much does model averaging help?
2007-23 **Matthias Sutter, Martin Kocher and Sabine Strauß:** Individuals and teams in UMTS-license auctions.
2007-22 **Jesus Crespo Cuaresma, Adusei Jumah and Sohbet Karbuz:** Modelling and Forecasting Oil Prices: The Role of Asymmetric Cycles.
2007-21 **Uwe Dulleck and Rudolf Kerschbamer:** Experts vs. discounters: Consumer free riding and experts withholding advice in markets for credence goods.
2007-20 **Christiane Schwieren and Matthias Sutter:** Trust in cooperation or ability? An experimental study on gender differences. *Conditionally accepted for publication in: Economics Letters.*
2007-19 **Matthias Sutter and Christina Strassmair:** Communication, cooperation and collusion in team tournaments – An experimental study.
2007-18 **Michael Hanke, Jürgen Huber, Michael Kirchler and Matthias Sutter:** The economic consequences of a Tobin-tax – An experimental analysis.
2007-17 **Michael Pfaffermayr:** Conditional beta- and sigma-convergence in space: A maximum likelihood approach.
2007-16 **Anita Gantner:** Bargaining, search, and outside options. *Revised version forthcoming in: Games and Economic Behavior.*
2007-15 **Sergio Currarini and Francesco Feri:** Bilateral information sharing in oligopoly.
2007-14 **Francesco Feri:** Network formation with endogenous decay.
2007-13 **James B. Davies, Martin Kocher and Matthias Sutter:** Economics research in Canada: A long-run assessment of journal publications. *Revised version forthcoming in: Canadian Journal of Economics.*
2007-12 **Wolfgang Luhan, Martin Kocher and Matthias Sutter:** Group polarization in the team dictator game reconsidered. *Revised version forthcoming in: Experimental Economics.*
2007-11 **Onno Hoffmeister and Reimund Schwarze:** The winding road to industrial safety. Evidence on the effects of environmental liability on accident prevention in Germany.
2007-10 **Jesus Crespo Cuaresma and Tomas Slacik:** An “almost-too-late” warning mechanism for currency crises.
2007-09 **Jesus Crespo Cuaresma, Neil Foster and Johann Scharler:** Barriers to technology adoption, international R&D spillovers and growth.
2007-08 **Andreas Brezger and Stefan Lang:** Simultaneous probability statements for Bayesian P-splines.
2007-07 **Georg Meran and Reimund Schwarze:** Can minimum prices assure the quality of professional services?.
2007-06 **Michal Brzoza-Brzezina and Jesus Crespo Cuaresma:** Mr. Wicksell and the global economy: What drives real interest rates?.
2007-05 **Paul Raschky:** Estimating the effects of risk transfer mechanisms against floods in Europe and U.S.A.: A dynamic panel approach.
2007-04 **Paul Raschky and Hannelore Weck-Hannemann:** Charity hazard - A real hazard to natural disaster insurance.
2007-03 **Paul Raschky:** The overprotective parent - Bureaucratic agencies and natural hazard management.
2007-02 **Martin Kocher, Todd Cherry, Stephan Kroll, Robert J. Netzer and Matthias Sutter:** Conditional cooperation on three continents.
2007-01 **Martin Kocher, Matthias Sutter and Florian Wakolbinger:** The impact of naïve advice and observational learning in beauty-contest games.

University of Innsbruck

Working Papers in Economics and Statistics

2007-24

Jesus Crespo Cuaresma

Forecasting euro exchange rates: How much does model averaging help?

Abstract

We analyze the performance of Bayesian model averaged exchange rate forecasts for euro/US dollar, euro/Japanese yen, euro/Swiss franc and euro/British pound rates using weights based on the out-of-sample predictive likelihood. The paper also presents a simple stratified sampling procedure in the spirit of Sala i Martin et alia (2004) to obtain model weights based on predictive accuracy. Our results indicate that accounting explicitly for model uncertainty when constructing predictions of euro exchange rates leads to improvements in predictive accuracy as measured by the mean square forecast error. While the forecasting error of the combined forecast tends to be systematically smaller than that of the individual model that would have been chosen based on predictive accuracy in a test sample, random walk forecasts cannot be beaten significantly in terms of squared forecast errors. Direction of change statistics, on the other hand, are significantly improved by Bayesian model averaging.

ISSN 1993-4378 (Print)
ISSN 1993-6885 (Online)