Birkbeck Working Papers in Economics & Finance

ISSN 1745-8587

School of Economics, Mathematics and Statistics

BWPEF 0616

Investing Under Model Uncertainty: Decision Based Evaluation of Exchange Rate and Interest Rate Forecasts in the US, UK and Japan

Anthony Garratt
Birkbeck, University of London

Kevin Lee
University of Leicester

November 2006

Investing Under Model Uncertainty: Decision Based Evaluation of Exchange Rate and Interest Rate Forecasts in the US, UK and Japan*

by

Anthony Garratt and Kevin Lee

Birkbeck College and University of Leicester, UK

Abstract

We evaluate the forecast performance of a range of theory-based and atheoretical models explaining exchange rates and interest rates in US, UK and Japan. The decision-making environment is fully described for an investor who optimally allocates portfolio shares to domestic and foreign assets. Methods necessary to compute and use forecasts in this context are proposed, including the means of combining density forecasts to deal with model uncertainty. An out-of-sample evaluation exercise covering the 1990's is described, comparing statistical criteria with decision-based criteria. The theory-based models are found to perform relatively well when their forecasts are judged by their economic value.

Keywords: Model Averaging, Buy and Hold, Exchange rate and interest rate forecasts.

JEL Classifications: C32, C53, E17

*We are grateful to Stephen Hall, Lucio Sarno and participants at the Society for Computational Economics 2005 conference for helpful comments on earlier versions of the papers. Of course, any errors remain our responsibility. Corresponding author: Kevin Lee, Department of Economics, University of Leicester, University Road, Leicester LE1 7RH, UK; web www.le.ac.uk/economics/kcl2.

1 Introduction

Recent years have seen a growing interest in the decision-based approach to the evaluation and comparison of forecasts. Here, forecast accuracy is judged according to its economic value to an individual given an explicitly defined decision-making context. This reflects the recognition that models should be judged according to their purpose and that the statistical criteria used to evaluate models, typically based solely around point forecasts and measured using mean squared forecasting error (MSE), are unlikely to provide information on the economic value of their forecasts.[1] The preponderance of studies employing the decision-based approach to forecast evaluation are in the area of applied finance where the decision-making context is relatively straightforward to describe.[2] But they remain relatively rare even here and model evaluation in the context of the analysis of exchange rates, for example, still focuses primarily on statistical criteria.[3]

One explanation for why the decision-based approach remains relatively unusual is because of the technical difficulties that arise in this type of analysis. For example, even where it is possible to fully articulate the decision-making context, the task of calculating the economic value of a model's forecasts can be daunting where many variables are involved (and it is interesting to note that even those papers that have undertaken a decision-based evaluation have typically focused on single variables). Similarly, in a real world decision-making context involving a number of variables, there is rarely consensus on the appropriate model(s) to be employed so that methods for accommodating model uncertainty in the analysis, in addition to the more usual stochastic and parameter uncertainties, are required. And, given that the economic worth of a forecast depends on aspects of the model which will differ from one analysis to another, there is no generally accepted decision-based criterion with which to judge models. This contrasts with the widespread acceptance of the MSE as an imperfect but reasonable statistical criterion that can be used to judge forecast performance in any analysis.

In this paper, we motivate and describe the methods necessary for an investor to

[1]See Granger and Pesaran (2000a,b) for an overview of this discussion.

[2]See, for example, , Boothe (1983, 1987), Leitch and Tanner (1991), West et al. (1993), Pesaran and Timmerman (1995), Kandel and Stambaugh (1996), Barberis (2000) and Abhyanker et al. (2005).

[3]See, for example, Meese and Rogoff (1983), Mark (1995), Mark and Sul (2001), Berkowitz and Giorganni (2001), Faust, Rogers and Wright (2003) Clarida, Sarno, Taylor and Valente (2003), Killian and Taylor (2003) and Cheung, Chinn and Pascual (2005) among others.

compute and use forecasts in choosing the proportion of his portfolio to be invested in domestic and foreign assets. The methods are based on simulation exercises and are straightforward to implement meaning that the technical issues involved in decision-based forecast evaluation can be overcome relatively easily. We apply the methods to a range of theory-based and atheoretical models explaining exchange rates and interest rates in the US, UK and Japan. The exercise involves calculating multivariate predictive density forecasts, combining density forecasts to allow for model averaging, and identifying and interpreting the appropriate decision-based criterion with which to judge the models. An out-of-sample evaluation exercise is described, comparing statistical criteria with decision-based criteria. It demonstrates that the conclusions drawn on the basis of the alternative criteria are quite different, with the theory-based models found to perform relatively well when their forecasts are judged by their economic value.

The plan of the paper is as follows. In section 2 we describe the investment decision and the methods required to use and evaluate forecasts from a number of individual models and/or from a model average. Section 3 outlines the candidate set of models for the exchange rate and interest rates on which the investment decision might be made. Section 4 describes the estimation of the models using US, UK and Japanese data for the period 1981m1-1997m12 and evaluates their forecasting performance using statistical criteria. Section 5 describes the decision-based forecast evaluation, judging the models' performance according to the utility derived from the associated investment strategies. Section 6 concludes.

2 The Investment Decision

The decision problem here is one in which an investor chooses at time T how much of his wealth to invest at home or abroad.[4] We assume that identical domestic and foreign assets are available, both maturing in each period, and their returns measured in local currency at time t are r_t and r_t^*, respectively. In the simplest case, the investment decision might be whether to invest all initial wealth at home during the decision (forecast) period $T+1$ to $T+H$ or to invest it abroad, based on a straight comparison of end-of-decision-period wealth, W_{T+H}, obtained under the alternative strategies (ignoring risk). In this very simple "home vs. abroad" case, attention focuses on the difference between end-of-

[4] The investment decision problem is similar to that in Kandel and Stambaugh (1996), Barberis (2000) and Abhyankar et al. (2004), among others.

decision-period wealth under the two strategies. Normalising so $W_T = 1$, the difference is given by

$$\begin{aligned} D_{T+H} &= \prod_{h=1}^{H}(1+r_{T+h}) - \prod_{h=1}^{H}(1+r^*_{T+h})\frac{E_{T+H}}{E_T} \\ &= \exp(\sum_{h=1}^{H} r_{T+h}) - \exp(\sum_{h=1}^{H} r^*_{T+h} + \Delta_H\ e_{T+H}), \end{aligned} \tag{1}$$

where E_t denotes the spot (end-of-period) nominal bilateral exchange rate describing the domestic price of the foreign currency, $e_t = \log(E_t)$, and where we use the approximations $\log(1+r) \approx r$ and $\log(E_{T+H}/E_T) \approx e_{T+H} - e_T$. The decision requires the investor to evaluate $\mathrm{E}(D_{T+H} \mid \Omega_T)$ with expectations based on information available at time T, Ω_T, choosing to invest at home if this expected value is positive and abroad if it is negative. However, even in this very straightforward case, it will not be possible to base the decision on simple point forecasts of the r_{T+h}, r^*_{T+h} $(h = 1, .., H)$ and e_{T+H} series given the non-linearities built into D_{T+H}.[5] Rather the investor will generally need to evaluate the entire joint probability distribution of the forecast values of r_{T+h}, r^*_{T+h} $(h = 1, .., H)$ and e_{T+H} to evaluate $\mathrm{E}(D_{T+H} \mid \Omega_T)$.

A more realistic decision-making context might be where the investor chooses the proportion of his portfolio allocated to the foreign asset (ω) at the outset and then, continuing to restrict attention to a 'buy-and-hold' strategy, retains this portfolio until the end of the decision horizon. Again normalising with unit initial wealth, the wealth of the investor at the end of the decision horizon can be expressed as

$$W_{T+H}(\omega) = (1-\omega)\exp(\sum_{i=h}^{H} r_{T+h}) + \omega\exp(\sum_{i=h}^{H} r^*_{T+h} + \Delta_H\ e_{T+H}) \tag{2}$$

where the dependence of W_{T+H} on ω is made explicit. Further extending the exercise to accommodate risk aversion in the investor's decision making, we might also assume that the investor derives utility from W_{T+H} according to the standard constant relative risk aversion (CRRA) power utility function,

$$\nu(W_{T+H}) = \frac{W_{T+H}^{1-A}}{1-A}, \tag{3}$$

[5] Recall, for example, that if a vector of variables $\mathbf{X} \sim N(\boldsymbol{\mu}, \Sigma)$, then $\mathrm{E}(\exp(\boldsymbol{\tau}'\mathbf{X})) = \exp(\boldsymbol{\tau}'\boldsymbol{\mu}) + \boldsymbol{\tau}'\Sigma\boldsymbol{\tau}$ where $\boldsymbol{\tau}'\mathbf{X}$ is some linear combination of the variables in $\mathbf{X}$. Hence, if $\mathbf{X}$ contained the series influencing D_{T+H}, including r_{T+h}, r^*_{T+h} $(h = 1, .., H)$ and e_{T+H}, the investor would need to know the variance and covariances of the forecast variables at each point in time as well as their means.

where A is the coefficient of risk aversion.[6] In this case, the investor's problem at time T can be written as

$$\max_{\omega} \left\{ \mathrm{E} \left[\nu(W_{T+H}(\omega)) \mid \Omega_T \right] \right\}.$$

The decision problem will again require the investor to use the entire joint probability distribution of the forecasts values of r_{T+h}, r^*_{T+h} $(h = 1, .., H)$ and e_{T+H}, so that they can first calculate the expected utility obtained for any given portfolio share, and then identify the optimal portfolio share as that which maximises the expected utility across all portfolio shares.

2.1 The Probability Density Function of the Forecast Values

The key to decision-making here is the probability density function of the forecast values of the exchange rate and domestic and foreign interest rates over the decision horizon. Denoting $\mathbf{z}_t = (z_{1t}, z_{2t}, ..., z_{nt})'$ to be an $n \times 1$ vector of variables of interest (including at least r_t, r^*_t and e_t here) and $\mathbf{Z}_T = (\mathbf{z}_1, \mathbf{z}_2, ..., \mathbf{z}_T)'$ to be the available observations at the end of period T, we are interested in the probability density function of $\mathbf{Z}_{T+1,H} = (\mathbf{z}_{T+1}, \mathbf{z}_{T+2}, ..., \mathbf{z}_{T+H})'$ conditional on $\mathbf{Z}_T$; that is $\Pr(\mathbf{Z}_{T+1,H} \mid \mathbf{Z}_T)$, sometimes termed the "predictive density function". The decision problem can then be written as

$$\max_{\omega} \left\{ \int \nu(W_{T+H}(\omega)) \Pr\left(\mathbf{Z}_{T+1,H} \mid \mathbf{Z}_T\right) d\mathbf{Z}_{T+1,H} \right\}. \tag{4}$$

The form of the density function $\Pr(\mathbf{Z}_{T+1,H} \mid \mathbf{Z}_T)$ depends on the types of uncertainty that surround the forecast and the approach taken to characterising and estimating the function. The types of uncertainty that might influence the forecasts include: the *stochastic uncertainty* associated with a model; the *parameter uncertainty* associated with estimated model parameters; and the *model uncertainty* surrounding the choice of model itself. The first two of these are routinely taken into account in forecasting, but model uncertainty is less frequently considered. This is despite the fact that this latter source of uncertainty is potentially more important in decision-making if there is little consensus on how the variables are determined (as is the case with international investment decisions, for example, since there is little agreement on the processes underlying exchange rate or interest rate determination).

[6] Campbell and Viceria (2002) argue in favour of power utility functions as they have the attractive property that absolute risk aversion declines with wealth whilst relative risk aversion remains constant.

The approach taken to characterising and estimating the density function varies according to judgements on the role of economic theory in econometric modelling and pragmatic decisions on the use of prior knowledge. Draper (1995) and Hoeting *et al.* (1999), for example, describe the "Bayesian Model Averaging" approach which elegantly accommodates all three forms of uncertainty described above in a comprehensive, fully Bayesian approach to estimating $\Pr(\mathbf{Z}_{T+1,\ H} \mid \mathbf{Z}_T)$. On the other hand, Garratt *et al.* (2003) [GLPS] argue that there are practical difficulties involved in the choice of priors for models, and in the choice of priors for the parameters of any given model, in the context of forecasting that involves high-dimensional models. GLPS therefore use approximations to certain probabilities so that their approach adopts a classical stance in a Bayesian framework.

To be more specific, assume that there are m different models, denoted M_i, $i = 1, ..., m$, characterized by a probability density function of $\mathbf{z}_t$ defined over the estimation period $t = 1, ..., T$, as well as the forecast period $t = T+1, ..., T+H$, in terms of a $k_i \times 1$ vector of unknown parameters, $\boldsymbol{\theta}_i$, assumed to lie in the parameter space, Θ_i. Model M_i is then defined by

$$M_i : \quad \{f_i\left(\mathbf{z}_1, \mathbf{z}_2, ..., \mathbf{z}_T, \mathbf{z}_{T+1}, \mathbf{z}_{T+2}, ..., \mathbf{z}_{T+H}; \boldsymbol{\theta}_i\right), \quad \boldsymbol{\theta}_i \in \Theta_i\}, \tag{5}$$

where $f_i(.)$ is the joint probability density function of past and future values of $\mathbf{z}_t$. The "Bayesian model averaging" formula writes

$$\Pr\left(\mathbf{Z}_{T+1,H} \mid \mathbf{Z}_T\right) = \sum_{i=1}^{m} \Pr\left(M_i \mid \mathbf{Z}_T\right) \Pr(\mathbf{Z}_{T+1,H} \mid \mathbf{Z}_T, M_i). \tag{6}$$

Both posterior probabilities on the right-hand-side of (6) can be rewritten, treating models and their parameters as intermediate inputs, in terms of prior probabilities and integrated likelihoods. For example, we can write

$$\Pr(\mathbf{Z}_{T+1,H} \mid \mathbf{Z}_T, M_i) = \int_{\boldsymbol{\theta}_i} \Pr\left(\boldsymbol{\theta}_i \mid \mathbf{Z}_T, M_i\right) \Pr(\mathbf{Z}_{T+1,H} \mid \mathbf{Z}_T, M_i, \boldsymbol{\theta}_i) d\boldsymbol{\theta}_i, \tag{7}$$

in which $\Pr\left(\boldsymbol{\theta}_i \mid \mathbf{Z}_T, M_i\right)$, the posterior probability of $\boldsymbol{\theta}_i$ given model M_i, is proportionate to $\Pr\left(\boldsymbol{\theta}_i \mid M_i\right) \Pr(\mathbf{Z}_T \mid M_i, \boldsymbol{\theta}_i)$, the product of the prior on $\boldsymbol{\theta}_i$ given M_i and the likelihood function of model M_i.[7] If meaningful priors exist, therefore, the application of Monte Carlo integration techniques to the elements underlying (6) provides a systematic approach to the estimation of $\Pr\left(\mathbf{Z}_{T+1,H} \mid \mathbf{Z}_T\right)$.

[7] See Draper (1995) or Garratt et al (2003) for details.

If, on the other hand, there are difficulties in obtaining meaningful priors, then it might be reasonable to employ approximations to some of the key probabilities in the BMA formula to estimate $\Pr(\mathbf{Z}_{T+1,H} \mid \mathbf{Z}_T)$. Specifically, in (7), GLPS suggest using the assumption

$$\boldsymbol{\theta}_i \mid \mathbf{Z}_T, M_i \overset{a}{\sim} N(\widehat{\boldsymbol{\theta}}_{iT}, T^{-1}\widehat{\mathbf{V}}_{\theta_i}) \tag{8}$$

for the posterior probability of $\boldsymbol{\theta}_i$ given M_i, where $\widehat{\boldsymbol{\theta}}_{iT}$ is the maximum likelihood estimate of $\boldsymbol{\theta}_{i0}$, the true value of $\boldsymbol{\theta}_i$ under M_i, and $T^{-1}\widehat{\mathbf{V}}_{\theta_i}$ is the asymptotic covariance matrix of $\widehat{\boldsymbol{\theta}}_{iT}$ conditional on M_i. And in (6), Draper (1995) suggests the use of the familiar Schwarz Bayesian information criterion to obtain model weights w_{iT}:

$$\Pr(M_i \mid \mathbf{Z}_T) = \frac{\exp(SBC^*_{iT})}{\sum_{j=1}^{m}\exp(SBC^*_{jT})} \tag{9}$$

where $SBC^*_{iT} = SBC_{iT} - \max_j(SBC_{jT})$, $SBC_{iT} = LL_{iT} - \left(\frac{k_i}{2}\right)\ln(T)$ is the Schwarz Bayesian information criterion, and LL_{iT} is the maximized value of the log-likelihood function for model M_i calculated on the basis of the sample running to period T. Alternatively, following Burnham and Anderson (1998), one could use Akaike weights, using AIC in place of SBC in (9).[8] These assumptions allow $\Pr(\mathbf{Z}_{T+1,H} \mid \mathbf{Z}_T)$ to be estimated straightforwardly using (6) and (7)and based on ML estimation of the candidate models.

The above discussion shows that there might be a variety of alternative predictive densities available to a decision-maker, including model-specific densities, $\Pr(\mathbf{Z}_{T+1,H} \mid \mathbf{Z}_T, M_i)$, $i = 1, ..., m$, and densities obtained through model averaging. Pesaran and Skouras (2000) suggest a decision-based criterion function for the evaluation of a predictive density function which, in the context of (4), is given by

$$\begin{aligned} \Psi &= \mathrm{E}_P\left[\nu(W_{T+H}(\omega^\dagger) \mid \Omega_T\right] \\ &= \int \nu(W_{T+H}(\omega^\dagger))P(\mathbf{Z}_{T+1,H})\, d\mathbf{Z}_{T+1,H} \end{aligned} \tag{10}$$

where $\omega^\dagger$ is the chosen optimal value of ω for the given predictive density and $\mathrm{E}_P[.]$ is the expectations operator with respect to $P(\mathbf{Z}_{T+1,H})$, the "true" probability density function of $\mathbf{Z}_{T+1,H}$ conditional on Ω_T. This can be viewed as the average utility obtained using

[8] The SBC weights are asymptotically optimal if the data generation process lies in the set of models under consideration, but the AIC weights are likely to perform better when the models represent approximations to a complex data generation process. Fernandez, Ley and Steel (2001) note that the choice of uninformed priors implies Bayes factors which behave asymtotically like SBC.

the given predictive density function when large samples of forecasts and realisations are available. The criterion function for the evaluation of the predictive density function clearly depends on the decision-making context, as captured by the utility function $\nu(.)$. Pesaran and Skouras show that the form of this criterion function is independent of the parameters of the underlying utility function only in the special case of the "LQ problem" involving a single decision variable (where the utility function is quadratic and constraints (if they exist) are linear). In that special case, the criterion is proportional to the MSE so that the purely statistical measure is appropriate. However, even the multivariate version of the LQ problem involves the parameters of the utility function so that, generally, statistical and decision-based forecast evaluation criteria are markedly different.

2.2 The linear VAR case

To illustrate these ideas more practically, assume that each of the models M_i can be written in the VAR form

$$\mathbf{z}_t = \sum_{i=1}^{p} \Phi_i \mathbf{z}_{t-i} + \mathbf{a}_0 + \mathbf{a}_1 t + \mathbf{v}_t,\ t = 1, 2, ..., T, T+1, .., T+\ H, \tag{11}$$

where Φ_i is an $n \times n$ matrix of parameters, $\mathbf{a}_0$, and $\mathbf{a}_1$are $n \times 1$ parameter vectors and $\mathbf{v}_t$ is assumed to be a serially uncorrelated *iid* vector of shocks with zero means and a positive definite covariance matrix, Σ. Using this model, an estimate of the probability distribution function of the forecasts can be obtained using stochastic simulation techniques.

Specifically, suppose that the ML estimators of the parameters in (11) Φ_i, $i = 1, \dots, p$, $\mathbf{a}_0$, $\mathbf{a}_1$ and Σ are denoted by $\hat{\Phi}_i$, $i = 1, \dots, p$, $\hat{\mathbf{a}}_0$, $\hat{\mathbf{a}}_1$ and $\hat{\Sigma}$, respectively. Then the point estimates of the h-step ahead forecasts of $\mathbf{z}_{T+h}$ conditional on Ω_T, denoted by $\hat{\mathbf{z}}_{T+h}$, can be obtained recursively by

$$\hat{\mathbf{z}}_{T+h} = \sum_{i=1}^{p} \hat{\Phi}_i \hat{\mathbf{z}}_{T+h-i} + \hat{\mathbf{a}}_0 + \hat{\mathbf{a}}_1 (t+h),\ h = 1, 2, \dots, \tag{12}$$

where the initial values, $\mathbf{z}_T, \mathbf{z}_{T-1}, \dots, \mathbf{z}_{T-p+1}$, are given. Hence, abstracting from parameter uncertainty for the time being, we can obtain an estimate of $\Pr(\mathbf{Z}_{T+1,\ H} \mid \mathbf{Z}_T, M_i)$ using stochastic simulation, obtaining forecast values of $\mathbf{z}_{T+\ H}$ using

$$\mathbf{z}_{T+h}^{(r)} = \sum_{i=1}^{p} \hat{\Phi}_i \mathbf{z}_{T+h-i}^{(r)} + \hat{\mathbf{a}}_0 + \hat{\mathbf{a}}_1 (t+h) + \mathbf{v}_{T+h}^{(r)}, \quad h = 1, 2, ..., H \quad \text{and} \quad r = 1, 2, ..., R, \tag{13}$$

where superscript '(r)' refers to the r^{th} replication of the simulation algorithm, and $\mathbf{z}_T^{(r)} = \mathbf{z}_T$, $\mathbf{z}_{T-1}^{(r)} = \mathbf{z}_{T-1}, \ldots$, $\mathbf{z}_{T-p+1}^{(r)} = \mathbf{z}_{T-p+1}$ for all r. The $\mathbf{v}_{T+h}^{(r)}$'s can be drawn either by parametric methods based on $\hat{\Sigma}$ or by non-parametric methods based on the estimated residuals on which $\hat{\Sigma}$ is calculated (see GLPS for more details).

These simulation exercises provide estimates of $\Pr(\mathbf{Z}_{T+1,\ H} \mid \mathbf{Z}_T, M_i)$ which can be used as predictive densities assuming a particular model is appropriate, or which can be used in a model averaging exercise. For any particular density, the simulations also allow us to evaluate $\mathrm{E}[\nu(W_{T+\ H}) \mid \Omega_T]$ in (4) for a range of values of ω (in practice calculating $\nu(W_{T+\ H}(\omega_0))$ in each replication for various values of ω_0 and calculating the mean value across replications). The investor's decision then simply involves choosing the ω associated with the maximum value of the simulated expected wealth.

In evaluating the alternative prediction densities, based on alternative models or model averages, the sample counterpart of the criterion function in (10) is

$$\overline{\Psi} = \frac{1}{N} \sum_{s=1}^{N} \nu(W_{T+\ H+s}(\omega^{\dagger})) \tag{14}$$

calculated recursively for $s = 1, .., N$ for each predictive density (with associated optimal share $\omega^{\dagger}$) and over the out-of-sample forecast evaluation period $T+s$,.., $T+H+s$. This provides an estimate of the realised utility to the decision-maker of using the predictive distribution function. In practice an absolute standard for forecast evaluation is not available because the true probability density function of the forecast variable is not known. But calculating loss differentials, comparing the economic value of outcomes based on alternative predictive distributions, is straightforward and a choice between the two can simply depend on whether the differential is positive or negative. If one predictive distribution function is given the status of a 'null', then the choice can be cast in terms of whether the loss differential is significantly greater than zero. The asymptotic distribution of the loss differential can be derived in the case of LQ problem (see Diebold and Mariano (1995)) but the nature of the test needs to be investigated on a case-by-case basis for other problems. This is relatively straightforward in the linear VAR case discussed here, however, since the distributional properties of the criterion function under the null can also be obtained through simulation.

3 The Candidate Set of Models

The exercise described above requires that we forecast the exchange rate and domestic and foreign rates of return. In what follows, we consider twelve alternative models on which forecasts of e_t, r_t and r_t^* can be based, each of which can be accommodated within the linear framework of (11). The twelve candidate models represent combinations obtained from four alternative sub-models of exchange rate determination explaining $\mathbf{z}_t^E$ (containing e_t and possibly other variables) and three alternative sub-models explaining $\mathbf{z}_t^R$ (containing r_t and r_t^* and possibly other variables). The sub-models relate to alternative theories of the exchange rate and interest rate determination. The twelve combined models explaining $\mathbf{z}_t = (\mathbf{z}_t^E, \mathbf{z}_t^R)$ provide forecasts of the variables of interest.

In the modelling exercises, we assume the variables in $\mathbf{z}_t$ to be $I(1)$ so that the candidate set of models can be written in the vector error correction (VECM) form:

$$\Delta \mathbf{z}_t = \mathbf{a} + \sum_{i=1}^{p} \Gamma_i \Delta \mathbf{z}_{t-i} + \boldsymbol{\alpha}\boldsymbol{\beta}' \mathbf{z}_{t-1} + \mathbf{u}_t, \tag{15}$$

using alternative cointegrating vectors $\boldsymbol{\beta}$ as suggested by the theory associated with the alternative models. Each model that we construct in the VECM form, (15), therefore represents a restricted version of (11) chosen to reflect a particular view on exchange rate or interest rate determination. The separate estimation of the sub-models for $\mathbf{z}_t^E$ and $\mathbf{z}_t^R$ impose block-diagonality in (11) providing a representation of the form:

$$\begin{pmatrix} \Delta \mathbf{z}_t^E \\ \Delta \mathbf{z}_t^R \end{pmatrix} = \mathbf{a} \begin{pmatrix} a^E \\ a^R \end{pmatrix} + \sum_{i=1}^{p} \begin{pmatrix} \Gamma_{11i} & 0 \\ 0 & \Gamma_{22i} \end{pmatrix} \begin{pmatrix} \Delta \mathbf{z}_{t-i}^E \\ \Delta \mathbf{z}_{t-i}^R \end{pmatrix} + \boldsymbol{\alpha} \begin{pmatrix} \beta^{E\prime} & 0 \\ 0 & \beta^{R\prime} \end{pmatrix} \mathbf{z}_{t-1} + \begin{pmatrix} \mathbf{u}_t^E \\ \mathbf{u}_t^R \end{pmatrix}.$$

This latter set of restrictions is for expositional convenience, allowing us to focus on the relative performance of the separate exchange rate and interest rate formulations.[9]

Exchange Rate Models The set of models that we consider for predicting exchange rates is:

- M_A : Autoregressive model of e_t in differences [AR(p)]
- M_E : Efficient Market Hypothesis [EMH]
- M_M : Monetary Fundamentals model [MF]

[9] Exercises in which the exchange rate and interest rate variables are modelled simultaneously, allowing for feedbacks across $\mathbf{z}_t^E$ and $\mathbf{z}_t^R$, showed that these feedbacks are relatively unimportant empirically.

- M_P : Purchasing Power Parity [PPP]

The most simple model that we consider is an autoregressive model of the change in exchange rates, so that $\mathbf{z}_t^E = (e_t)$ and it is assumed $\boldsymbol{\beta}^{\mathbf{E}'} = 0$. This specification is widely used and constitutes a benchmark against which to judge the other three, more structural, models of exchange rate determination. In the EMH, we define $\mathbf{z}_t^E = (e_t,\, f_t)'$ where f_t is the logarithm of the three month forward (end-of-period) nominal bilateral exchange rate and we assume the cointegrating vector is given by $\boldsymbol{\beta}^{E\prime} = (1, -1)$. This model relates to the literature on foreign exchange market efficiency which tests whether the forward rate is an optimal predictor of the future spot exchange rate. Although the empirical evidence is mixed regarding the optimality of the forward rate as a predictor of the spot rate, evidence in Clarida and Taylor (1997) suggests, for example, that some information is contained in the term structure of the forward rate. Moreover the EMH specification we adopt does not require efficient markets to hold at all points in time.

In the MF model, $\mathbf{z}_t^E = (e_t,\, x_t)'$, where x_t represents a 'fundamentals' term, given by $x_t = (m_t - m_t^*) - (y_t - y_t^*)$, and m_t and y_t denote the log-levels of the domestic money supply and real income respectively, the '*' superscript indicates the corresponding foreign variable, and $\boldsymbol{\beta}^{\mathbf{E}'} = (1,\, -1)$. This specification has a long tradition in the analysis of exchange rate determination (Frenkel, 1976; Mussa, 1976, 1979; Frenkel and Johnson, 1978), and has recently been the subject of much debate (as in Mark (1995), Mark and Sul (2001) and Berkowitz and Giorgianni (2001), for example). And, finally, in the PPP model, $\mathbf{z}_t^E = (e_t,\, p_t - p_t^*)'$, where p_t and p_t^* denote the logarithm of the domestic and foreign price level respectively, and $\boldsymbol{\beta}^{\mathbf{E}'} = (1,\, -1)$ so that the real exchange rate is stationary. Like the EMH, this theory is often viewed as an arbitrage condition in international goods and is considered to be an integral part to many open economy views of the world. The literature considering the empirical validity of PPP is well developed and the conclusions are mixed, but there is some recent evidence that it may hold in the long-run (see Garratt et al. (2003) for example).

Interest Rate Models The set of models that we consider for predicting foreign and domestic interest rates is:

- M_V : Vector-Autoregressive model of r_t and r_t^* in differences [VAR(p)]

- M_I : Interest Rate Parity [IRP]

- M_T : Term Structure models plus IRP [TS]

The simplest model we consider is an unrestricted bivariate VAR in the difference in domestic and foreign interest rate, so that $\mathbf{z}_t^R = (r_t, r_t^*)'$, where r_t and r_t^* are the continuously compounded return on identical domestic and foreign assets of one month maturity, and $\boldsymbol{\beta}^{R\prime}= \mathbf{0}$. The IRP model defines $\mathbf{z}_t^R = (r_t,\, r_t^*)'$ and sets $\boldsymbol{\beta}^{R\prime} = (1, -1)$ so that there is interest rate parity in the long-run. The form of this equation arises from the UIP arbitrage condition, which implies that if there no expected change in the exchange rate in the long-run, then we would expect returns to domestic and foreign assets be equalised (note that the inclusion of the intercept in (15) allows for the possibility of a non-zero, but constant, risk premium). See Garratt *et al.* (2003a) for further details.

Finally here, the third model of interest rates defines $\mathbf{z}_t^R = (r_{lt},\, r_t, r_{lt}^*, r_t^*,)'$ where r_{lt} and r_{lt}^* are the returns on long-term domestic and foreign assets. The TS model assumes the cointegrating relations to take the following form

$$\boldsymbol{\beta}^{R\prime} = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 1 & 0 & -1 \end{pmatrix}.$$

This form ensures long-run parity between the rates on domestic short- and long-term assets, between the foreign short-and long-term assets, and between domestic and foreign assets. The model is motivated by arbitrage once more and assumes constant risk premia for domestic and foreign countries and between long- and short-term interest rates. This type of specification is similar to those used by Campbell and Shiller (1987,1991) and Fauvel *et al.* (1999), for example .

Combining Exchange Rate and Interest Rate Models The combinations of the four exchange rate and three interest rate models are summarised in Table 1.

Table 1: The Candidate Set of Models

	Interest Rates		
Exchange Rate	VAR (M_V)	IRP (M_I)	TS+IRP (M_T)
AR (M_A)	M_{AV}	M_{AI}	M_{AT}
EMH (M_E)	M_{EV}	M_{EI}	M_{ET}
MF (M_M)	M_{MV}	M_{MI}	M_{MT}
PPP (M_P)	M_{PV}	M_{PI}	M_{PT}

4 Exchange Rate and Interest Rate Models for the US, UK and Japan

4.1 Data

In our empirical work, we use monthly data for the US, UK and Japan over the period 1981m1-2001m12 (252 observations) and consider two separate exercises based on the decision to invest in the US or UK, and the decision to invest in the US or Japan. Variables employed in the analysis include short term 3-month nominal interest rates (r_t and r_t^*), long term 10-year government bond yields (r_{lt} and r_{lt}^*), money supply (m_t and m_t^*), industrial production (y_t and y_t^*) and consumer prices (p_t and p_t^*) in the three countries. We also consider the one month spot- and forward- nominal exchange rates (denoted by e_t and f_t respectively) for Sterling-Dollar and Yen-Dollar. All the data used in the analysis are in natural logarithms and the precise definitions, sources and transformations are described in the Data Appendix. The main sample period is 1981m1-1997m12, but also considers data upto four years later for out-of-sample model evaluation.

Figures 1-8 plot the levels and first differences of the exchange rates, the level of short term interest rates and their differentials, plus the excess returns computed as $\Delta e_{t+1} - (r_t - r_t^*)$. Figures 1-4 show the exchange rates to be volatile, possibly non-stationary in levels (confirmed by unit root tests). For the out of sample forecasting period 1998m1 onwards, the exchange rates show no clear patterns with the Pound-Dollar rate first depreciating but then appreciating back to approximately its 1997 level, whilst the Yen-Dollar exchange rate shows a slightly more pronounced appreciation which is then mostly reversed. Figures 5-6 suggest non-stationarity of the interest rates for the sample period, with similar looking downward trends in all three rates demonstrating some co-movement. The differentials are mostly positive for the US-Japan case and negative for the US-UK case. The differentials also look downward trended in the first half of the sample. The excess returns in Figures 7 and 8 are volatile and do not exhibit any clear patterns.

4.2 Estimation

Our empirical analysis began by testing the assumption in Section 3 that all variables are $I(1)$. In every case, we failed to reject the null of a unit root in levels but rejected the null in first differences. We therefore proceeded in the analysis assuming all variables are $I(1)$.[10] Next, we selected the lag length to be used in our forecasting models by estimating a sequence of unrestricted VAR(p), $p = 0, 1, 2, ...12$ for each set of variables employed in the models M_A to M_T in Table 1 for both the US-UK and US-Japan data sets and over the sample period 1981m1-1997m12. The lag selection criteria used was the likelihood ratio test and the lag length chosen was twelve with a few exceptions. We therefore use a lag length of twelve for all models in both data sets, except for the models M_A and M_E for the US-UK data set and model M_E for the US-Japan data set, where we use a lag length of six.

Turning to the cointegrating properties of our two data sets, Table 2 reports for the period 1981m1-1997m12 pairwise cointegration tests and the tests of the over-identifying restrictions on the six long-run relationships implied by the theories underlying the candidate set of models described in the previous section. The results in Table 2 present some evidence in favour of the long-run relationships used in our candidate set of models. However, like much of the literature, the support is mixed and we are faced with a trade off between empirical fit and a form which reflects known theories of exchange rate and interest rate determination. For the long-run relationships that involve the exchange rate (the first three in the tables), the p-value for Johansen's trace statistic indicates the presence of a cointegrating relationship in all cases.[11] For the long-run relationships involving interest rates, the trace statistic provides mixed evidence for cointegration. The US term structure relationship and the interest parity relationship between US-Japan rates show weak evidence in favour of cointegration[12] but there is more evidence rejecting cointegration in the other cases. Note however that when we impose cointegration and test for the over-identifying long-run restrictions on the interest rate relations suggested

[10]The results are available from the authors on request.

[11]A noteworthy observation of these results is the cointegration of the exchange rate with the constructed fundamentals term, an issue that has received much attention in the literature, see for example Berkowitz and Giorgianni (2001).

[12]Note ADF and PP statistics, which we do not report, also give evidence in favour of stationarity in $r_t - r_t^l$ for the US and $r_t - r_t^*$ for US-Japan.

by the various theories, we cannot reject the null in nearly every case. There is also some evidence in favour of the over identifying long-run restrictions for the efficient markets hypothesis and the fundamentals relationship in the exchange rate analysis, but there is little evidence for the long-run restrictions implied by purchasing power parity.

The error-correction models outlined in (15) and summarised in Table 1 can be estimated, assuming the long-run restrictions suggested by the various theories hold, to form the basis of our forecasting models. For reasons of parsimony, we do not report the full estimation details for all models. However, we document some basic diagnostics for the Δe_t, Δr_t, Δr_t^* equations for all the models in Table 3. The table indicates that the estimated interest rate models show a reasonable degree of in-sample fit, with $\overline{R}^2$ of around 0.2 or 0.3, but the exchange rate models fit less well (in line with most empirical findings in the literature). The diagnostic tests for the exchange rate and foreign interest rate models perform relatively well but, despite the earlier work on choice of lag length, there is evidence of significant serial correlation and ARCH effects for domestic US interest rates. In short, the models appear broadly adequate although none clearly outperforms the others. These results echo those of the cointegration tests which provide some support for the cointegrating restrictions suggested by the various economic theories but which showed that the evidence is by no means clear-cut. These ambiguities lie at the heart of the model uncertainty experienced in investment decision-making because the empirical models each have their strengths but none of the models seems entirely satisfactory on purely statistical grounds and certainly none unambiguously dominates the others.

Probabilitistic statements on the likely relevance of models can be made on the basis of the weights given in (9). Table 4 and Figures 9-10 report on the model weights, w_{iT}, based on the AIC and SBC statistics and calculated according to the formula (9), and the corresponding formula for AIC, for $T = 1989m12$, $1990m3, ..., 1997m12$; i.e. an evaluation period that covers most of the nineties.[13] To obtain these statistics, the twelve

[13] As our candidate set of models are not nested, system-based criteria are not directly comparable. Instead, the reported AIC and SBC statistics of Table 4 are based on the equations explaining Δe_t, Δr_t and Δr_t^* in each model, taking the equations for these three series in isolation from the system in which they are embedded. For example, the criteria are based on equation log likelihoods calculated by $LL = \frac{-n}{2}\left\{1 + \log(2\Pi\widetilde{\sigma}^2)\right\}$ where $\widetilde{\sigma}^2 = \frac{\mathbf{e}'\mathbf{e}}{n}$ and $\mathbf{e}$ are the equation residuals. Such a decomposition of a system's likelihood effectively assumes the covariances between the variables of interest and the other variables in the system are negligible. While this is unlikely to be true in practice, these approximations

models were each estimated over the period 1981m1-1989m12 and then recursively, at three month intervals, through to 1981m1-1997m12 (making 33 recursions in total). As a reference against which to compare the performance of the twelve models, a 'random walk' model for Δe_t, Δr_t, Δr_t^* was also estimated and is referred to as model M_{RW}. Table 4 reports the average value of the model weights obtained over the evaluation period and shows that, in the US-UK case for example, some support was obtained for nine of the thirteen models at some point during the evaluation period when the weights are based on AIC (and a similar number for the US-Japan exercise). Taking the period as a whole, the most successful model in both the US-UK and US-Japan cases is model M_{AT} combining an autoregressive model for the exchange rate with a term structure model of interest rates, with an average weight of 56% and 42% respectively. However, Figures 9-10 show that there is considerable variability in the model weights calculated over different recursions, so that the average weight assigned to M_{AT} is based on very high weights at the end of the evaluation period offsetting very low weights at the beginning of the period. Selection according to the SBC places a very high premium on parsimony and, as it turns out, weights based on SBC are completely dominated by the reference random walk model M_{RW} in both the US-UK and US-Japan case.[14] Taken together, then, these statistics reflect the extent of the model uncertainty, indicating that it is relatively difficult to choose between models on purely-statistical selection criteria and that model weights can be quite sensitive to movements in the values of the likelihoods over time.

4.3 Statistical Evaluation of Forecasting Performance

The forecasting performance of the models can be evaluated statistically by calculating the standard root mean squared error (RMSE) relating to the forecasts of (cumulative) excess returns, defined for forecast horizon H at time T as $c_T(H) = \sum_{h=1}^{H}[\Delta e_{T+h+1} - (r_{s,T+h} - r_{s,T+h}^*)]$. The RMSE are calculated for each model and reported in Table 5 for forecast horizons $H = 1, 3, 6, 12, 24, 36$ and 48. The table also reports the RMSEs obtained using a weighted average of forecasts from all the models, with equal weights (i.e. $\frac{1}{13}$) and weights based on AIC and SBC as in (9). The reported statistics in the

allow model comparison across alternative systems.

[14] If the random walk model is excluded, the simplest model M_{AV} is picked out for UK-US, with a weight of 97%, and M_{AV} and M_{EV} are highlighted for US-Japan, with weights 52% and 33%.

table are again averages based on the RMSEs obtained in 33 recursions covering the evaluation period $T = 1989m12$, $1990m3, ...,$ $1997m12$ at three-monthly intervals.

Table 5 indicates that simple atheoretical models perform relatively well in terms of RMSEs at short horizons but that more sophisticated theory-based models perform better at $H = 24$, 36, and 48. The models including autoregressive models for the exchange rate and VARs for interest rates perform relatively well in both the US-UK and US-Japan cases for $H = 1$, 3, 6, and 12 although, in fact, almost all models are outperformed by the simple random walk reference model M_{RW} at these horizons. This is perhaps not too surprising given that the models other than M_{RW} are heavily parameterised and, as shown in Clements and Hendry (2005), using RMSE as a criterion penalises models for including variables with low associated t-values even if the model is misspecified by their exclusion.[15] However, at longer horizons, the models involving interest rate parity and the term structure begin to perform relatively well (certainly for the US-UK case), as do models involving PPP in the exchange rate equations. Indeed, almost all the models involving PPP outperform the random walk models by $H = 48$. So, despite their size, the theory-based models appear to perform well by statistical criteria at longer horizons.

The results for the forecasts obtained through model-averaging reflect these patterns in some respects so that, for example, the simple equal-weight average has a higher RMSE than that for M_{RW} at short horizons, but falls below that for M_{RW} subsequently (at $H = 24$ for US-UK and $H = 6$ for US-Japan). On the other hand, there is some evidence that model averaging serves to reduce the forecast error compared to the individual models. For example, taking all the forecasts horizons together, the RMSE of the equal-weights average model is just 89% of the average of the RMSE across the thirteen models for US-UK forecasts and 91% for US-Japan forecasts. Similar, but less pronounced, gains are obtained using the AIC-weighted average.[16] This finding is in line with the view, expressed in the review of Clemen (1989) and more recently by Harvey and Newbold (2005) for example, that combinations of forecasts typically perform well in a statistical sense and can outperform the forecasts of a single model even if this is the true (but estimated) data generating process.

[15]More precisely, Clements and Hendry show that forecasting stationary processes using a model that retains all variables with an expected $(t - value)^2 > 2$ will dominate in terms of one-step ahead forecast accuracy measured by RMSE.

[16]Given that SBC systematically chooses M_{RW}, the SBC-average and M_{RW} figures are equal throughout.

In brief, then, a statistical evaluation of the models in terms of their diagnostic statistics or in-sample fit provides relatively little guidance on the appropriateness of the various models for use in investment decisions or on the gains to be made from the various models. In terms of forecasting performance measured by RMSE, simple models appear to perform well over shorter horizons, while theory-based models work well at longer horizons, and model-averaging is useful if there is ambiguity over the true model. It remains to be seen whether a more clear-cut picture emerges on the usefulness of the models' forecasts when they are judged more directly in the context of the objectives of the investment decision.

5 Forecast Evaluation by US Investors

Section 2 described the decision made by a buy-and-hold investor with a given horizon to be one of solving the problem in (4) to choose the proportion of her/his portfolio that should be devoted to domestic and foreign assets. This choice requires the implementation of the simulation-based procedure described in Section 2.2 to obtain the probability distribution of the future values of $\mathbf{Z}$, $\Pr(\mathbf{Z}_{T+1,\ H}\ |\mathbf{Z}_T, M_i, \boldsymbol{\theta}_i)$, with which to evaluate (and then maximise) expected future utility. A description of the algorithm used to compute the optimal portfolio shares, based on the discussion of Section 2.2, is given in the Appendix.

Having computed the portfolio shares we are then able to conduct what is the main focus of this paper; namely, an ex-post forecasting exercise which uses the optimal portfolio shares and evaluates, given observed outcomes for exchange rates and interest rates, the end-of-period wealth and utility (for each investment horizon) if the investor had allocated their portfolio using these portfolio shares. The evaluation takes the form of comparing the utility ratios of the models and the model-averages relative to a "passive" benchmark strategy of holding wealth entirely at home.

5.1 The Optimal Portfolio Shares

Tables 6a and 6b report the optimal portfolio share allocated by a US investor through the nineties over the investment horizons $H = 1, 3, 6, 12, 24, 36, 48$ and for three different values of the coefficient of risk aversion, $A = 2$, 5, and 10, employing various alternative models of exchange rate and interest rate determination. Table 6a relates to the choice between US and UK assets and Table 6b relates to the US-Japan choice. The statistics are again generated in the recursive manner described above. Hence, the models are estimated first for the period 1981m1-1989m12 and the optimal portfolio shares decided based on the forecasts obtained from the various models. The process is then repeated moving forward three months, recomputing the model weights (for the average models) and wealth and utility forecasts to obtain new optimal shares for each model. This process is repeated for each recursion until we have results for 33 recursions covering the evaluation period $T = 1989m12$, $1990m3, ...,$ $1997m12$ at three-monthly intervals. The statistics reported in Table 6 relate to the average portfolio share across the 33 recursions.

To illustrate the range of outcomes obtained across the models, the optimal shares are reported in Table 6a for a US-UK investor using models M_{EV}, M_{MV}, or M_{AV} as the basis of the portfolio allocation decision plus the outcomes based on the equal-weighted, AIC-weighted or SBC-weighted average of forecasts from the thirteen candidate models. In Table 6b, the shares associated with models M_{EV}, M_{EI} and M_{AT} are reported alongside the shares based on model averages for the US-Japan case.

There are a number of interesting features of the statistics reported in the tables. As expected, given the uncertainties associated with the exchange rate, the proportion of wealth in foreign assets falls as the risk aversion parameter rises. So, for example, if we simply average the figures in the columns of Table 6a, the share allocated to UK assets falls from 43% when $A = 2$ to 33% when $A = 5$ and to 23% when $A = 10$. The share of UK assets based on forecasts from the equal-weight average model is 31% when averaged over the various investment horizons and for $A = 5$. Here, where the AIC weights were distributed relatively widely across the models, this is reasonably close to the 26% suggested by the AIC-weighted average model.[17] The average results accommodate considerable heterogeneity in outcome across the various models, however, as shown in the three examples provided in the table. Hence, again taking the averages across the investment horizons and with $A = 5$, model M_{EV} suggests a holding of 45% while M_{MV} suggests 62% and M_{AT} suggests just 13%. Similar patterns are observed in Table 6b. Here, averaging across the columns in the Table, the shares of Japanese assets are 63%, 40% and 29% for $A = 2$, 5 and 10 respectively. The share suggested by

[17] Although the 'SBC average' is actually simply the random walk model M_{RW} in this case, with a weight of unity, the optimal share here is, at 21%, close to the results of the other average models too.

the equal-weight average model is 45%, looking across the various investment horizons and for $A = 5$. But this is rather lower than all the other figures reported in the table, which suggest shares of 58% (M_{EV}), 55% (M_{EI}), 32% (M_{AT}), 42% (AIC-average) and 47% (SBC-average), again highlighting the heterogeneity in outcomes across the thirteen models that were considered.

5.2 Economic Evaluation of Forecasts using Investor Utility

Table 7 provides an economic evaluation of the forecast performance of the various models from the perspective of an investor with risk-aversion parameter $A = 5$. In line with the criterion function described in (14), the table describes the end-of-period utilities that would have been obtained over the period 1989m12-1997m12 if the investor had chosen the optimal portfolio shares suggested by the thirteen models, plus the averaged models, in real time. The utilities are expressed as a ratio to the utility that would have been achieved if the investor had followed a passive investment strategy in which no modelling took place and the whole portfolio was invested at home, so that the return (and utility level) was determined entirely by the US interest rate. As before, the statistics reported in the table relate to the average outcome over 33 recursions in the evaluation period (i.e. setting $N = 33$ in (14)). As a means of judging the effects of changing the risk-aversion parameter, Table 8 provides a set of illustrative results for the end-of-period utility outcomes for six models with $A = 2$, 5 and 10.

The finding that virtually all the utility ratios presented in Tables 7 and 8 exceed unity indicates that an active investment strategy would have outperformed the passive one using any model and at almost every investment horizon. As an indication of whether these figures are 'significant' or not, a simulation exercise was undertaken. Here, starting at $T = 1989m12$ and then for the subsequent 33 quarterly observations, 1000 replications of artificial data on the interest rates and exchange rates were generated for $T + h$, $h = 1, ..., 48$, using estimated random walk (with drift) models for these series. In each replication, the end-of-period utility obtained following the passive investment strategy was calculated for the various investment horizons and this was averaged over the 33 recursions. The simulations provide the distribution of utilities that would be obtained over different data realisations under the assumption that interest rates follow a random walk and US investors always invest at home. By comparing the model-based utilities that are reported in Table 7 against the 95th percentile of the simulated distribution, we can judge whether the model-based utilities are 'unusually' or significantly high. As the table shows, the largest utilities are indeed generally significant according to this criterion.

It is clear not only that an active investment strategy outperforms a passive one, but also that the active investment strategy is substantially improved when the forecasts of some models are used rather than others. So, for the US-UK investor for example, the models incorporating the efficient markets hypothesis for exchange rate determination, M_{EV}, M_{EI} and M_{ET}, clearly lead to the highest utility outcomes for shorter investment horizons ($H = 1$, ...,$H = 12$) while models incorporating the monetary fundamentals model of the exchange rate, M_{MV}, M_{MI} and M_{MT}, dominate for longer investment horizons. Comparison across the interest rate models appears to place the VAR model ahead of the IRP models which, in turn, dominate the Term Structure models, but the differences are small compared to the differences due to the choice of exchange rate model. End-of-period utility ratios are systematically higher for the results relating to US-Japan, indicating that a passive strategy would have been even less attractive in this case, and the efficient market models again clearly outperform the alternatives, with these three ranked in the order M_{EV}, M_{EI} and M_{ET}.

The utility outcomes for the equal-weight average model are comparable to, and indeed slightly lower than, the average outcomes across the thirteen individual models, so that the equal-weight model outperforms some models but does not perform as well as others in terms of utility outcomes. Similar comments apply to the AIC-weighted average in the UK case, although the AIC-weights are large on the efficient markets models in the Japanese case, so this average model performs well compared to most of the individual models. Taken together, however, there is little evidence here to suggest that the advantages of model averaging that accrue in terms of reduced forecast RMSE carry over to the decision-based forecast evaluations. Moreover, the AIC and SBC on which the averaging weights might be based appear to have little relevance to the forecasting performance judged according to utility-maximisation. For example, the single model that would be chosen according to AIC is model M_{AT} for both the US-UK and US-Japan cases. As shown in Table 8, however, investments based on forecasts from this model are actually outtperformed by the passive investment strategy in the US-UK case and are not strong relative to other models for the US-Japan case. The random walk model that is strongly supported by SBC performs relatively well in the Japanese case, but performs less well than the average model in the UK case.

6 Discussion and Conclusion

Generally speaking, judgements on the forecasting performance of the various models appear to be quite different depending on whether the evaluation is based on a statistical approach or a decision-based approach. According to the statistical view based on RMSEs, the simple random walk model M_{RW} performs best at short horizons, while the models involving interest rate parity, the term structure and PPP perform well at longer horizons. In contrast, according to the decision-based criteria, the models incorporating the efficient markets hypothesis dominate for shorter investment horizons while, at least for the UK, those incorporating the monetary fundamentals model dominate for longer investment horizons.

In a similar vein, we found that in-sample prediction criteria like AIC and SBC provide little insight on which models will perform well in terms of utility maximisation in the context of our investment decision. Equally, although our results conform with the frequently-observed finding that modelling averaging is useful when judged by the statistical criterion, we found that model averaging was unhelpful when decision-based criterion are used.

It is worth observing that the forecast evaluation based on the statistical criteria appear to support the atheoretic models of interest rate and exchange rate determination (along with artificial model averaging). The theory-based models, on the other hand, are relatively well-supported by the decision-based criteria. Indeed, theory might have suggested also that models incorporating the efficient markets hypothesis would perform particularly well at short horizons, as indicated by the decision-based criteria. Moreover, this criterion also selected the monetary fundamentals model for exchange rates for which there was strongest support in the tests of the theory-based long-run restrictions in the cointegrating vectors.

The results of this empirical exercise provide a clear illustration that decision-based forecast evaluation can differ markedly from that provided by general statistical evaluation criteria. While it will remain unusual for the decision-making environment to be fully articulated, it is clear that, when it is possible, models and their forecasts should be evaluated according to the purpose to which they will be used. The exercise also show that the technical issues involved in decision-based forecast evaluation can be readily addressed using the methods outlined in the paper, based on relatively straightforward simulation exercises, even where many variables and/or model uncertainty are involved.

References

Abhyankar, A., Sarno, L., and G. Valente (2005), "Exchange Rates and Fundamentals: Evidence on the Economic Value of Predictability", *Journal of International Economics*, 66(2), 325-348.

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

Baberis,N. (2000), "Investing for the Long Run when Returns Are Predictable", *Journal of Finance*, Vol. LV, N0. 1, 225-264.

Berkowitz, J. and L. Giorgianni (2001), "Long-Horizon Exchange Rate Predictability?", *Review of Economics and Statistics*, 83(1), 81-91.

Campbell, J.Y. and R.J. Shiller (1987), "Cointegration Tests and Present Value Models", *Journal of Political Economy*, 95, 1002-1088.

Campbell, J.Y. and R.J. Shiller (1991), "Yield Spreads and Interest Rate Movements: A Birds Eye View", *Review of Economic Studies*, 58, 495-514.

Campbell, J.Y. and L.M., Viceira (2002), *Strategic Asset Allocation: Portfolio Choice for Long-Term Investors*, Claredon Lectures in Economics, Oxford University Press: Oxford.

Cheung, Y., M.D. Chinn and A.G. Pascual (2005), "Empirical Exchange Rate Models of the Nineties: Are They Fit to Survive?", *Journal of International Money and Finance,* 24 (7), 1150-1175.

Clarida, R.H. and M.P. Taylor (1997), "The Term Structure of Forward Premiums and the Forecastability of Spot Exchange Rates", *Review of Economics and Statistics*, 89, 353-361.

Clarida, R.L., L. Sarno, M.P. Taylor and G. Valente (2003), "The Out of Sample Success of Term Structure Models as Exchange Rate Predictors: a step beyond", *Journal of International Economics*, 60, 61-83.

Clemen, R.T. (1989), "Combining Forecasts: A Review and Annotated Bibliography with Discussion, *International Journal of Forecasting*, 5, 559-608.

Clements, M. and D. Hendry (2005), "Information in Economic Forecasting", *Oxford Bulletin of Economics and Statistics*, 67 (Supplement), 713-753.

Draper, D. (1995), "Assessment and Propagation of Model Uncertainty," *Journal of Royal Statistical Society Series B*, 58, 45-97.

Faust, J., J.H. Rogers and J.H Wright (2003), "Exchange Rate Forecasting: the error we've really made", Journal of International Economics, 60, 35-59.

Fauvel, Y., A. Paquet and C. Zimmerman (1999), "A Survey on Interest Rate Forecasting", CREFE Working Paper 87, University of Quebec.

Fernandez, C., Ley, E. and Steel, M. F. J. (2001), "Benchmark Priors for Bayesian Model Averaging," *Journal of Econometrics*, 100, 381-427.

Garratt, A., Lee, K., Pesaran, M. H. and Shin, Y. (2003), Forecast Uncertainties in Macroeconometric Modelling: An Application to the UK Economy, *Journal of American Statistical Association, Applications and Case Studies*, 98, 464, 829-838.

Granger, C. W. J. and Pesaran, M. H. (2000a), "A Decision Theoretic Approach to Forecast Evaluation," in Statistics and Finance: An Interface, eds. W. S. Chan, W. K. Li and H. Tong, London: Imperial College Press, pp. 261-278.

Granger, C. W. J. and Pesaran, M. H. (2000b), "Economic and Statistical Measures of Forecast Accuracy," *Journal of Forecasting*, 19, 537-560.

Hai, Mark and Wu (1997), "Understanding Spot and Forward Exchange Rate Regressions", *Journal of Applied Econometrics*, 12, 715-734

Harvey, D.I. and P. Newbold (2005), "Forecast Encompassing and Parameter Estimation", *Oxford Bulletin of Economics and Statistics*, 67 (Supplement), 815-835.

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

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

Kandel, S. and R.F. Stanbaugh (1996), "One the Predictability of Stock Returns: An Asset-Allocation Perspective", *Journal of Finance*, 51, 385-424.

Leitch, G. and J.E. Tanner (1991), "Economic Forecasts Evaluation: Profits Versus the Conventional Measures", *American Economic Review*, 81, 580-590.

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

Mark, N.C. and D. Sul (2001),"Nominal Exchange Rates and Monetary Fundamentals: Evidence from a Small Post Bretton Woods Panel", *Journal of International Economics,* 53, 29-52.

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

Pesaran, M. H. and Timmermann, A. (1995), "Predictability of Stock Returns: Robustness and Economic Significance," *Journal of Finance*, 50, 1201-1228.

Pesaran, M.H. and S. Skouris (2002),"Decision-based Methods for Forecast Evaluation", in Clements, M.P. and Hendry, D.F. (eds.), *A Companion to Economic Forecasting,* Oxford: Blackwell.

Sarno, L. and G. Valente (2003), "Comparing the Accuracy of Density Forecasts form Competing Models", Mimeo.

Wallis, K. F. (1999), "Asymmetric Density Forecasts of Inflation and the Bank of England's Fan Chart," *National Institute Economic Review*, 106-112.

West, K.D., H.J. Edison and D. Cho (1993), "A Utility-Based Comparison of Some Models of Exchange Rate Volatility", *Journal of International Economics*, 35, 23-45.

Appendix A: The Data

The sources and transformations for the data are as follows:

e_t : the natural logarithm of the UK Sterling and Japanese Yen per US Dollar nominal spot exchange rate. Source: International Financial Statistics (IFS), codes 112AGZF and 158AEZF respectively.

f_t : For UK/US, the natural logarithm of the UK Sterling/US one month forward rate. Source: Bank of England, code XUMLDS1. For US-Japan, the natural logarithm of the 1 month forward Yen per dollar exchange rate collated from three sources: (i) for 1979m1-1992m8, the data is from Hai et al. (1997) available from the *JAE Data Archive* (ii) for 1997m1-2003m6, the data was collected from Datastream, code USJP1F; and (iii) for 1992m9-1996m12, we constructed the data assuming covered interest parity i.e. $f_t = r_t - r_t^* + e_t$.

r_t : US (domestic) three month treasury bill rates expressed as a monthly rate, $r_t = 1/12 \times \ln[1 + (R_t/100)]$ where R_t is the annualised rate. Source: IFS, code 11260C.

r_t^* : for foreign short term interest rates, monthly rates $r_t^* = 1/12 \times \ln[1 + (R_t^*/100)]$ where R_t^* is the annualised three month Treasury bill rates for the UK (Source: IFS, code 11160C) and the three month discount rate for Japan (Source: IFS, code 15860ZF).

r_{lt} : US (domestic) 10 year government bond rates, expressed as a monthly, $r_{lt} = 1/12 \times \ln[1 + (R_{lt}/100)]$ where R_{lt} is the annualised rate. Source: IFS, code 11161ZF.

r_{lt}^* : for foreign long term interest rates, monthly rates $r_{lt}^* = 1/12 \times \ln[1 + (R_{lt}^*/100)]$ where R_{lt}^* is the annualised long term government bond rate for both the UK and Japan. Source: IFS, codes 11261ZF and 15861ZF respectively.

y_t : the natural logarithm of US industrial production, constant 1995 prices, 1995=100. Source: IFS, code 11166 CZF.

y_t^* : the natural logarithm of UK and Japanese industrial production, constant 1995 prices, 1995=100. Source: IFS, codes 11266 CZF and 15866 CZF respectively.

p_t : the natural logarithm of US (domestic) consumer prices, index 1995=100. Source: IFS, code 11164ZF.

p_t^* : the natural logarithm of UK and Japanese (foreign) consumer prices, index 1995=100. Source: IFS, codes 11264ZF and 15864ZF respectively.

m_t : the natural logarithm of US (domestic) narrow money (M1 seasonally adjusted). Source: IFS, code 11159MA.

m_t^* : the natural logarithm of UK and Japanese (foreign) narrow money (M0 seasonally adjusted). Source: IFS, codes 11259MC ZF and 15834BZF respectively.

Appendix B: Computation of the Optimal Portfolio Weights

1. For a given model, M_i, with a fixed set of parameters, $\boldsymbol{\theta}_i$, we generate a sequence of forecasts for $r_{T+h}^{(r)}$, $r_{T+h}^{*(r)}$ and $\Delta_H e_{T+H}^{(r)}$, for $h = 1, .., H$ and $r = 1, ..., R$ (where $R = 50,000$ and $i = 1, ..., 13$) based on draws from a distribution of errors. These are non -parametric draws in our case.

2. For each replication r, we compute the value of $W_{T+H}^{(r,\omega)}$ using equation (**??**), where ω has 101 values $\omega = 0, ..., 1$ in step lengths of 0.01. Hence, we have a total of $R \times 101$ values of $W_{T+H}^{(r,\omega)}$ for each forecast horizon H. The forecast horizons considered in the paper are $H = 1, 3, 6, 12, 24, 36$, and 48.

3. We translate $W_{T+H}^{(r,\omega)}$ into the utility $v(W_{t+H}^{(r,\omega,A)})$, using CRRA utility defined in equation (**??**), for each level of risk aversion $A = 2, 5$ and 10. Then we compute, for the given A, H and ω,:

$$\frac{1}{R}\sum_{r=1}^{R} v(W_{T+H}^{(r,\omega,A)})$$

4. The optimal portfolio weight, for each forecast horizon H and level of risk aversion A, is the value of ω which maximizes the above expression; i.e. the maximum utility over the 101 different values of the portfolio weight ω.

5. Repeat for all models $i = 1, ..., 13$ and the average models, AIC, SBC and equal weights.

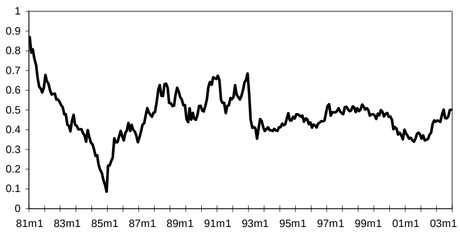

Figure 1: Logarithm Pound-Dollar Spot Exchange Rate

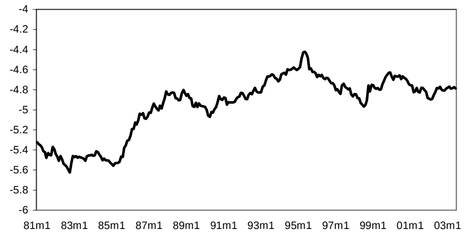

Figure 2: Logarithm Yen-Dollar Spot Exchange Rate

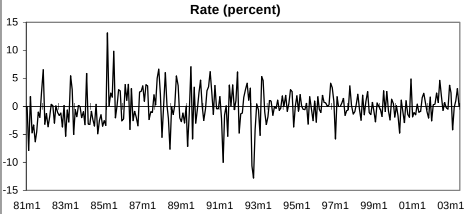

Figure 3: Change in the Pound-Dollar Spot Exchange Rate (percent)

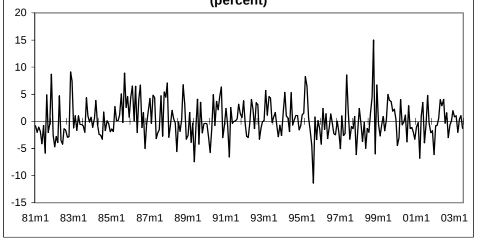

Figure 4: Change in the Yen-Dollar Spot Exchange Rate (percent)

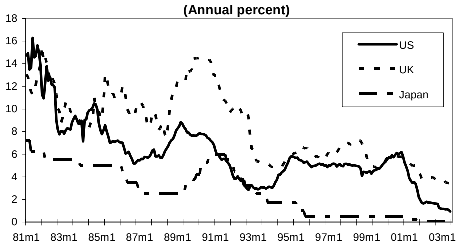

Figure 5: Short Term Interest Rates (Annual percent)

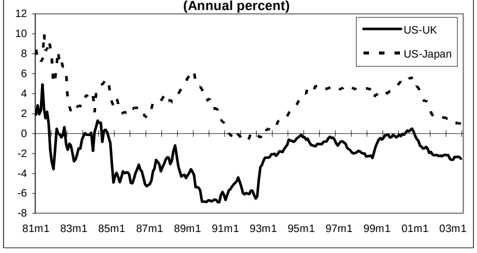

Figure 6: Short Term Interest Rates Differentials (Annual percent)

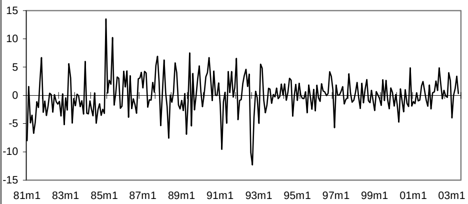

Figure 7: Excess Returns Pound-Dollar (percent)

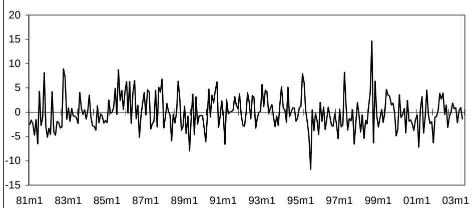

Figure 8: Excess Returns Yen-Dollar (percent)

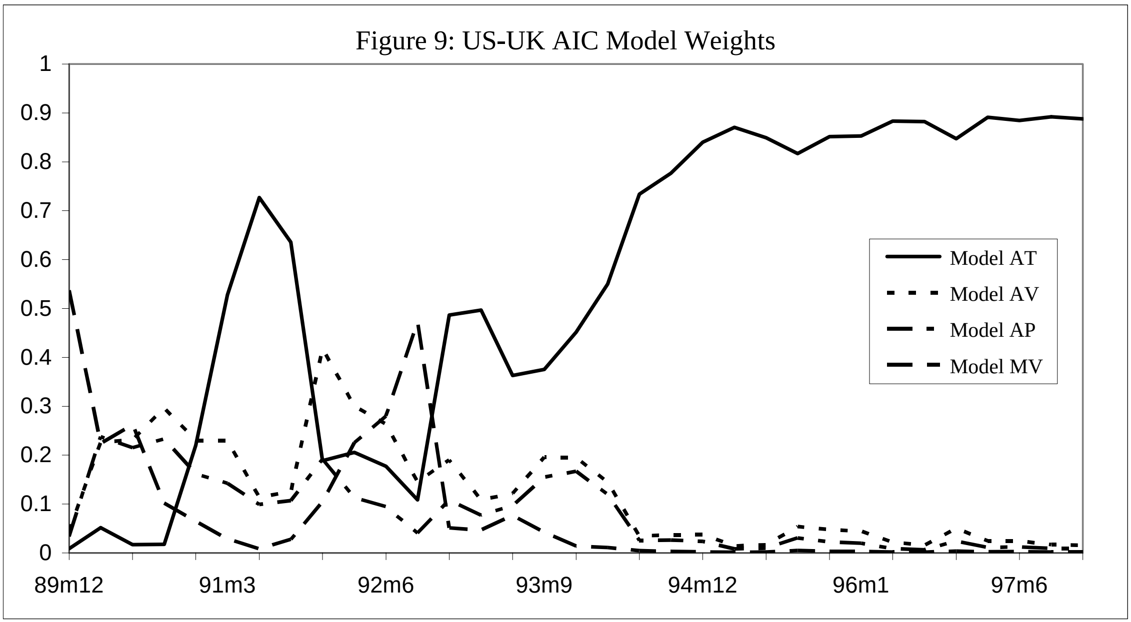
Figure 9: US-UK AIC Model Weights
Model AT
Model AV
Model AP
Model MV
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
89m12
91m3
92m6
93m9
94m12
96m1
97m6

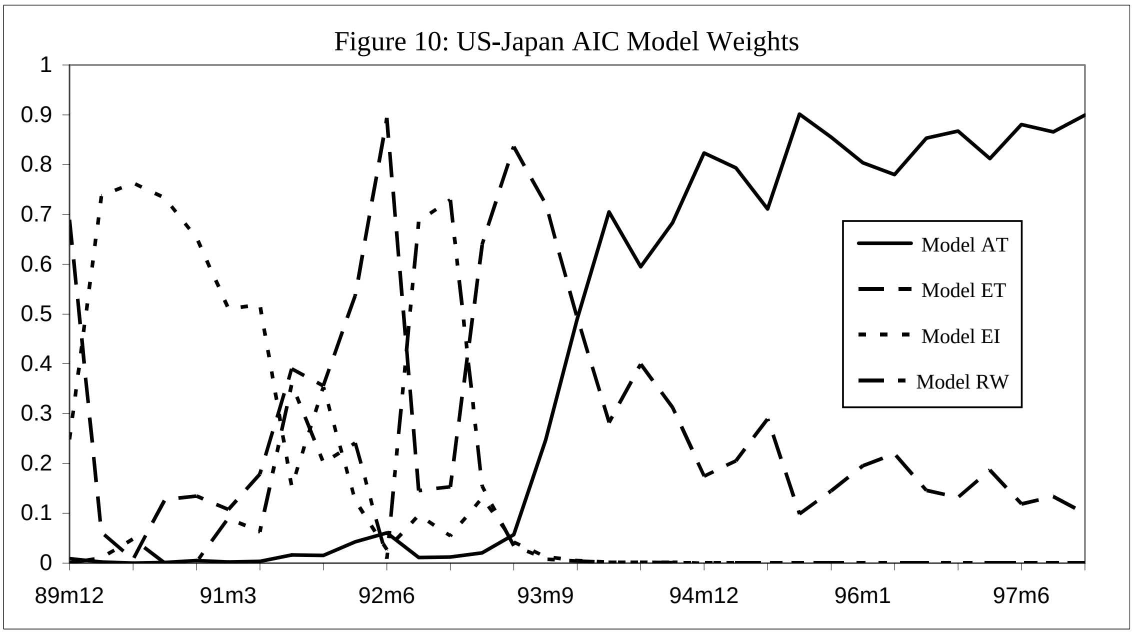
Figure 10: US-Japan AIC Model Weights
Model AT
Model ET
Model EI
Model RW
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
89m12
91m3
92m6
93m9
94m12
96m1
97m6

Table 2 : **Cointegration Rank and Over Identifying Tests on the Long Run Relationships**

(a) US-UK

Long Run	**P-value on Trace Statistic**	$(1,-1)$ **Over identifying Restriction Test**
$e_t - f_t$	$r=0: 0.045$ $r \leq 1: 0.168$	$\chi^2(1) = 5.12$ [.023]
$x_t - e_t$	$r=0: 0.006$ $r \leq 1: 0.115$	$\chi^2(1) = 0.09$ [.760]
$e_t - (p_t - p_t^*)$	$r=0: 0.026$ $r \leq 1: 0.300$	$\chi^2(1) = 9.98$ [.002]
$r_t - r_t^*$	$r=0: 0.360$ $r \leq 1: 0.608$	$\chi^2(1) = 6.29$ [.023]
$r_{lt} - r_{st}$	$r=0: 0.062$ $r \leq 1: 0.163$	$\chi^2(1) = 1.55$ [.213]
$r_{lt}^* - r_{st}^*$	$r=0: 0.482$ $r \leq 1: 0.452$	$\chi^2(1) = 0.81$ [.367]

(b) US-Japan

Long Run	**P-value on Trace Statistic**	$(1,-1)$ **Over identifying Restriction Test**
$e_t - f_t$	$r=0: 0.001$ $r \leq 1: 0.790$	$\chi^2(1) = 4.05$ [.044]
$x_t - e_t$	$r=0: 0.006$ $r \leq 1: 0.196$	$\chi^2(1) = 5.70$ [.017]
$e_t - (p_t - p_t^*)$	$r=0: 0.002$ $r \leq 1: 0.208$	$\chi^2(1) = 20.2$ [.000]
$r_t - r_t^*$	$r=0: 0.071$ $r \leq 1: 0.093$	$\chi^2(1) = 2.26$ [.133]
$r_{lt} - r_{st}$	$r=0: 0.062$ $r \leq 1: 0.163$	$\chi^2(1) = 1.55$ [.213]
$r_{lt}^* - r_{st}^*$	$r=0: 0.358$ $r \leq 1: 0.615$	$\chi^2(1) = 3.06$ [.080]

Notes: The underlying model used when calculating the p-values for the Johansen's log-likelihood-based trace statistic testing for the rank r is a VAR model of lag order 12 (except $e_t - f_t$, where 6 is used) where we allow for the presence of a linear deterministic time trend in the data for the first three long-run relationships involving the exchange rate, but assume no deterministic trend is present for the relationships involving interest rates. All test statistics are computed for the period 1981m1-1997m12 (204 observations).

Table 3: **Model Equation Diagnostics**

(a) US-UK

Model	LL	$\overline{R}^2$	$S.E.$	χ^2_{SC}	χ^2_H	χ^2_{ARCH}
M_A: Eq for Δe_t	403.2	0.003	0.0341	3.75 [0.71]	9.22 [0.68]	6.96 [0.32]
M_E: Eq for Δe_t	406.4	-0.002	0.0342	2.13 [0.91]	30.57 [0.24]	8.40 [0.21]
M_M: Eq for Δe_t	417.1	0.037	0.0335	20.14 [0.06]	44.35 [0.70]	11.50 [0.49]
M_P: Eq for Δe_t	411.9	-0.012	0.0344	9.29 [0.68]	6.91 [0.86]	90.21 [0.03]
M_V: Eq for Δr_t	1350.6	0.226	0.0003	45.68 [0.00]	95.27 [0.00]	50.12 [0.00]
Eq for Δr^*_t	1359.5	0.299	0.0003	17.65 [0.13]	54.95 [0.21]	16.45 [0.17]
M_I: Eq for Δr_t	1350.8	0.223	0.0003	45.85 [0.00]	96.74 [0.00]	48.17 [0.00]
Eq for Δr^*_t	1360.9	0.304	0.0003	17.66 [0.13]	57.59 [0.28]	22.68 [0.03]
M_T: Eq for Δr_t	1388.1	0.368	0.0003	27.14 [0.01]	161.2 [0.00]	52.51 [0.00]
Eq for Δr^*_t	1379.2	0.319	0.0003	18.58 [0.10]	97.27 [0.00]	19.09 [0.09]
RW: Eq for Δe_t	399.8	0.00	0.0340	9.53 [0.65]	-	9.79 [0.63]
Eq for Δr_t	1311.6	0.00	0.0004	25.30 [0.01]	-	56.99 [0.00]
Eq for Δr^*_t	1310.4	0.00	0.0004	25.14 [0.01]	-	25.66 [0.02]

(b) US-Japan

Model	LL	$\overline{R}^2$	$S.E.$	χ^2_{SC}	χ^2_H	χ^2_{ARCH}
M_A: Eq for Δe_t	411.4	0.021	0.0333	16.19 [0.18]	17.88 [0.81]	12.77 [0.39]
M_E: Eq for Δe_t	410.5	0.007	0.0335	6.54 [0.37]	21.61 [0.71]	7.10 [0.31]
M_M: Eq for Δe_t	417.7	0.011	0.0334	16.71 [0.16]	40.52 [0.83]	9.32 [0.68]
M_P: Eq for Δe_t	415.9	-0.006	0.0337	15.92 [0.19]	59.68 [0.16]	17.12 [0.15]
M_V: Eq for Δr_{st}	1342.9	0.165	0.0004	57.53 [0.00]	88.06 [0.00]	47.18 [0.00]
Eq for Δr^*_t	1511.3	0.133	0.0001	16.11 [0.19]	33.71 [0.21]	23.44 [0.02]
M_I: Eq for Δr_t	1343.8	0.168	0.0004	59.97 [0.00]	87.82 [0.00]	46.45 [0.00]
Eq for Δr^*_t	1514.6	0.156	0.0002	14.10 [0.29]	36.33 [0.93]	22.54 [0.03]
M_T: Eq for Δr_t	1379.5	0.313	0.0003	39.05 [0.00]	127.94 [0.04]	62.61 [0.00]
Eq for Δr^*_t	1540.8	0.235	0.0001	16.09 [0.19]	85.97 [0.93]	22.77 [0.03]
RW: Eq for Δe_t	403.1	0.00	0.0336	17.84 [0.12]	-	9.79 [0.63]
Eq for Δr_t	1311.6	0.00	0.0004	25.30 [0.01]	-	56.99 [0.00]
Eq for Δr^*_t	1483.9	0.00	0.0001	32.52 [0.01]	-	25.66 [0.02]

Notes: RW denotes a random walk 'benchmark' model; models $M_A - M_T$ are described in Table 1. For model comparison and diagnosis, LL is the Log Likelihood, S.E. is the standard error of the regression, SC tests for the presence of serial correlation in the residuals, H tests for heteroscedasticity and ARCH tests for autoregressive conditional heteroscedasticity. P-values are given in [ ] brackets and the period of estimation is 1981m1-1997m12.

Table 4: **Model Weights for the Period 1989m12-1997m12 based on AIC and SBC**

Model	US-UK		US-Japan	
	w_{it}^{AIC}	w_{it}^{SBC}	w_{it}^{AIC}	w_{it}^{SBC}
M_{AV}	0.1223	0.0	0.0020	0.0
M_{AI}	0.0799	0.0	0.0059	0.0
M_{AT}	0.5566	0.0	0.4190	0.0
M_{EV}	0.0185	0.0	0.0452	0.0
M_{EI}	0.0134	0.0	0.1566	0.0
M_{ET}	0.0358	0.0	0.2910	0.0
M_{MV}	0.0793	0.0	0.0	0.0
M_{MI}	0.0374	0.0	0.0	0.0
M_{MT}	0.0565	0.0	0.0001	0.0
M_{PV}	0.0	0.0	0.0	0.0
M_{PI}	0.0	0.0	0.0	0.0
M_{PT}	0.0	0.0	0.0001	0.0
M_{RW}	0.0	1.0	0.0800	1.0

Notes: The weights reported here are the average calculated from the recursive regressions ran over 1981m1-1989m12 through to 1981m12-1997m12. The models $M_{AV} - M_{RW}$ are as defined in Table 1.

Table 5: **Root Mean Squared Errors for Cumulative Excess Returns**

(a) US-UK

Model	$H=1$	$H=3$	$H=6$	$H=12$	$H=24$	$H=36$	$H=48$
M_{AV}	.0252	.0367	.0497	.0669	.0803	.0937	.1052
M_{AI}	.0252	.0367	.0497	.0668	.0787	.0889	.0975
M_{AT}	.0252	.0368	.0500	.0681	.0847	.0986	.1091
M_{EV}	.0256	.0397	.0561	.0784	.1092	.1273	.1442
M_{EI}	.0256	.0397	.0559	.0777	.0991	.1195	.1313
M_{ET}	.0256	.0397	.0562	.0789	.1029	.1251	.1374
M_{MV}	.0305	.0435	.0589	.0800	.1018	.1218	.1344
M_{MI}	.0305	.0435	.0589	.0799	.1006	.1169	.1248
M_{MT}	.0305	.0435	.0591	.0800	.1005	.1147	.1196
M_{PV}	.0298	.0378	.0485	.0618	.0707	.0760	.0814
M_{PI}	.0298	.0445	.0626	.0845	.0977	.1004	.0986
M_{PT}	.0298	.0445	.0626	.0845	.0980	.1015	.1007
M_{RW}	.0253	.0368	.0492	.0661	.0827	.0992	.1136
Equal-weight Av.	.0264	.0378	.0510	.0685	.0799	.0899	.0946
AIC Average	.0271	.0388	.0520	.0700	.0826	.0964	.1082
SBC Average	.0253	.0368	.0492	.0661	.0827	.0992	.1136

(b) US-Japan

Model	$H=1$	$H=3$	$H=6$	$H=12$	$H=24$	$H=36$	$H=48$
M_{AV}	.0234	.0423	.0584	.0782	.1195	.1637	.2008
M_{AI}	.0234	.0423	.0586	.0795	.1262	.1758	.2151
M_{AT}	.0234	.0423	.0586	.0801	.1279	.1787	.2201
M_{EV}	.0259	.0460	.0663	.0867	.1275	.1661	.2050
M_{EI}	.0259	.0461	.0663	.0866	.1281	.1691	.2084
M_{ET}	.0259	.0461	.0664	.0875	.1299	.1725	.2138
M_{MV}	.0264	.0471	.0645	.0926	.1440	.1939	.2231
M_{MI}	.0264	.0472	.0649	.0943	.1514	.2079	.2424
M_{MT}	.0264	.0472	.0650	.0950	.1534	.2112	.2462
M_{PV}	.0257	.0444	.0617	.0870	.1337	.1718	.2011
M_{PI}	.0257	.0444	.0613	.0819	.1243	.1644	.1952
M_{PT}	.0257	.0443	.0611	.0804	.1164	.1500	.1763
M_{RW}	.0228	.0427	.0598	.0801	.1183	.1567	.1919
Equal-weight Av.	.0238	.0427	.0589	.0782	.1164	.1567	.1839
AIC Average	.0251	.0450	.0630	.0830	.1210	.1605	.1976
SBC Average	.0228	.0427	.0598	.0801	.1183	.1567	.1919

Notes: Reported statistics are average RMSEs for cumulated returns, computed as follows: (i) define cumulated excess returns for forecast horizon H at time T as $c_T(H) = \sum_{h=1}^{H}[\Delta e_{T+h+1} - (r_{s,T+h} - r^*_{s,T+h})]$; (ii) calculate the RMSE of $c_T(H)$ for 33 quarterly recursions 1981m1-T, T=1989m12,...,1997m12; (iii) take the average of the 33 RMSE ratios for each model and horizon

Table 6: **Optimal Portfolio Shares Allocated to Foreign Assets**
(percentages)

(a) US-UK

Example 1; Model M_{EV}				Equal Weights Av. Model		
$A=2$	$A=5$	$A=10$		$A=2$	$A=5$	$A=10$
58.7	45.2	32.4	$H=1$	42.4	34.9	27.6
55.4	49.6	41.9	$H=3$	38.9	32.3	24.9
53.2	48.9	38.9	$H=6$	35.8	28.4	21.7
52.3	48.4	34.6	$H=12$	37.3	29.7	20.6
51.9	45.6	25.8	$H=24$	39.3	29.7	19.5
51.6	41.1	21.3	$H=36$	40.6	30.1	19.9
51.0	37.9	19.1	$H=48$	43.1	31.1	20.9

Example 2; Model M_{MV}				AIC Average Model		
$A=2$	$A=5$	$A=10$		$A=2$	$A=5$	$A=10$
63.0	61.2	54.3	$H=1$	38.6	31.2	24.6
57.5	52.9	43.7	$H=3$	36.1	27.7	19.0
56.7	49.2	40.6	$H=6$	33.0	22.1	15.3
63.7	57.4	45.1	$H=12$	34.1	22.2	15.1
75.4	68.3	54.6	$H=24$	36.9	25.4	17.3
74.9	71.7	58.9	$H=36$	37.7	26.1	18.1
75.0	71.9	60.9	$H=48$	38.5	26.4	18.7

Example 3; Model: M_{AT}				SBC AverageModel		
$A=2$	$A=5$	$A=10$		$A=2$	$A=5$	$A=10$
32.8	24.9	17.3	$H=1$	34.1	16.7	8.4
29.2	20.0	11.1	$H=3$	34.7	17.2	8.7
23.5	12.2	6.1	$H=6$	35.3	17.7	8.9
22.0	9.2	4.6	$H=12$	36.1	19.2	9.9
18.9	7.6	3.9	$H=24$	37.4	22.0	12.0
19.0	7.6	3.8	$H=36$	39.1	25.4	15.0
19.9	7.9	3.9	$H=48$	40.9	28.7	18.9

(b) US-Japan

Example 1; Model M_{EV}				Equal Weights Av. Model		
$A=2$	$A=5$	$A=10$		$A=2$	$A=5$	$A=10$
55.1	54.09	48.1	$H=1$	49.0	43.6	35.4
71.1	61.6	50.4	$H=3$	49.0	39.2	28.6
72.3	65.2	50.4	$H=6$	48.7	37.6	24.3
70.9	61.9	43.6	$H=12$	46.0	33.0	19.4
73.5	53.7	34.4	$H=24$	42.4	26.1	14.4
75.0	55.1	30.8	$H=36$	41.0	24.0	13.0
76.1	54.4	29.9	$H=48$	41.7	24.2	13.5

Example 2; Model M_{EI}				AIC Average Model		
$A=2$	$A=5$	$A=10$		$A=2$	$A=5$	$A=10$
54.7	53.8	47.9	$H=1$	57.1	53.5	44.7
72.1	61.5	50.1	$H=3$	66.8	60.9	48.4
73.2	65.2	49.7	$H=6$	71.9	63.1	44.6
72.5	61.2	41.5	$H=12$	71.1	56.4	36.1
75.7	52.4	30.2	$H=24$	72.2	47.5	27.1
77.2	47.9	25.3	$H=36$	71.3	43.7	23.4
77.5	44.1	23.0	$H=48$	71.4	41.5	22.2

Example 3; Model: M_{AT}				SBC AverageModel		
$A=2$	$A=5$	$A=10$		$A=2$	$A=5$	$A=10$
45.7	42.2	35.5	$H=1$	66.9	40.7	20.4
46.8	39.0	28.5	$H=3$	67.3	41.5	20.8
53.6	38.8	21.3	$H=6$	67.5	42.5	21.4
57.2	34.8	17.5	$H=12$	68.5	44.2	22.8
52.4	24.7	12.5	$H=24$	70.4	48.3	26.2
49.3	21.4	11.1	$H=36$	72.7	52.9	30.8
48.9	20.5	10.7	$H=48$	74.7	57.3	35.9

Notes: The statistics relate to the optimal share held on average across 33 quarterly recursions over the period 1989m12-1997m12.

Table 7: **End-Period Utility Ratios from Investments,** $A = 5$
(relative to benchmark strategy of 100% portfolio in US assets)

(a) US-UK

Model	$H = 1$	$H = 3$	$H = 6$	$H = 12$	$H = 24$	$H = 36$	$H = 48$
M_{AV}	1.0101**	1.0159*	1.0020	0.9963	1.0171	0.9814	0.9963
M_{AI}	1.0103**	1.0148**	1.0003	0.9901	1.0064	0.9844	0.9924
M_{AT}	1.0107**	1.0140**	0.9973	0.9843	0.9871	0.9817	0.9847
M_{EV}	1.0125**	1.0369**	1.0729	1.0891	1.0913	0.9994	1.0412
M_{EI}	1.0128**	1.0371**	1.0739	1.0901	1.0746	0.9987	1.0364
M_{ET}	1.0120**	1.0380**	1.0752	1.0866*	1.0346	0.9875	1.0200
M_{MV}	1.0019	1.0273**	1.0686**	1.0670	1.1607	1.1122	1.2058*
M_{MI}	1.0019	1.0275**	1.0637**	1.0669	1.1527**	1.1269**	1.2164**
M_{MT}	1.0017	1.0283**	1.0557**	1.0658	1.0981	1.1062**	1.1939**
M_{PV}	1.0048**	1.0119**	1.0156**	1.0001	1.0469**	1.0196	1.0539
M_{PI}	1.0048**	1.0121**	1.0154**	0.9975	1.0121**	1.0075	1.0441**
M_{PT}	1.0047**	1.0122**	1.0154**	0.9967	1.0014	1.0011	1.0149**
M_{RW}	1.0056**	1.0045	1.0085	1.0040	1.0141	0.9706	0.9921
Equal-weight Av.	1.0063**	1.0194**	1.0302**	1.0246	1.0455	1.0165	1.0536
AIC Average	1.0025	1.0083	1.0071	0.9898	1.0624	0.9964	1.0170
SBC Average	1.0056**	1.0045	1.0085	1.0040	1.0141	0.9706	0.9921

(b) US-Japan

Model	$H = 1$	$H = 3$	$H = 6$	$H = 12$	$H = 24$	$H = 36$	$H = 48$
M_{AV}	1.0107**	1.0437**	1.0734**	1.1067**	1.1249	1.0915	1.0267
M_{AI}	1.0107**	1.0435*	1.0665**	1.0784**	1.0407	0.9625	0.9030
M_{AT}	1.0106*	1.0430**	1.0684**	1.0744**	1.0379	0.9494	0.8977
M_{EV}	1.0041	1.0172	1.0813**	1.1763**	1.3134	1.3210	1.2786
M_{EI}	1.0037	1.0161	1.0772**	1.1619**	1.2349	1.1419	1.0601
M_{ET}	1.0043	1.0116	1.0708**	1.1479**	1.2135	1.1051	1.0219
M_{MV}	0.9987	1.0002	0.9894	1.0022	1.0017	1.0566**	1.4137**
M_{MI}	0.9986	0.9987	0.9889	1.0002	1.0000	1.0000	1.0058
M_{MT}	0.9984	0.9966	0.9883	1.0003	1.0000	1.0000	1.0000
M_{PV}	1.0105**	1.0203**	1.0356**	1.0836**	1.1379**	1.1533**	1.1412**
M_{PI}	1.0103**	1.0196**	1.0288	1.0589	1.0511**	1.0060	0.9760
M_{PT}	1.0105**	1.0191	1.0326**	1.0607	1.0565**	0.9995**	0.9723
M_{RW}	1.0092**	1.0289**	1.0618**	1.1292**	1.2575**	1.3365	1.2621
Equal-weight Av.	1.0056**	1.0177**	1.0373**	1.0738**	1.0917**	1.0537	1.0266
AIC Average	1.0026	1.0390**	1.0891**	1.1517**	1.2556**	1.2022	1.1391
SBC Average	1.0092*	1.0289**	1.0618**	1.1292**	1.2575**	1.3365	1.2621

Notes: Statistics are average end-period utility ratios, expressed relative to that obtained when no modelling is undertaken and 100% of portfolio is held in US assets, calculated over 33 quarterly recursions 1981m1-T, T=1989m12,...,1997m12. Superscripts "*" and "**" indicate significance at the 10% and 5% levels respectively as explained in the text.

Table 8: **End-Period Utility Ratios from Home-Overseas Investments**
(relative to benchmark strategy of 100% portfolio in US assets)

(a) US-UK

Example 1: Model M_{EV}				Equal Weights Av. Model		
$A=2$	$A=5$	$A=10$		$A=2$	$A=5$	$A=10$
1.0009**	1.0125**	1.0474**	$H=1$	1.0009**	1.0063**	1.0186**
1.0037**	1.0369**	1.1344	$H=3$	1.0035**	1.0194**	1.0533**
1.0087**	1.0729	1.2370	$H=6$	1.0064**	1.0302**	1.0723
1.0107	1.0891	1.2174	$H=12$	1.0038	1.0246	1.0519
1.0129	1.0913	1.1455	$H=24$	1.0109	1.0455	1.0760
1.0015	0.9994	0.9998	$H=36$	1.0030	1.0165	1.0439
1.0153**	1.0412	1.0514	$H=48$	1.0153*8	1.0536**	1.1109**

Example 2: Model M_{MV}				AIC Average Model		
$A=2$	$A=5$	$A=10$		$A=2$	$A=5$	$A=10$
0.9997**	1.0019**	1.0135	$H=1$	1.0000	1.0025	1.0116
1.0053**	1.0273**	1.0944**	$H=3$	1.0009	1.0088	1.0205
1.0124**	1.0686	1.1772	$H=6$	1.0011	1.0071	1.0110
1.0116**	1.0670	1.2049	$H=12$	0.9938	0.9898	0.9968
1.0251**	1.1607	1.3977	$H=24$	1.0091	1.0624	1.1101
1.0234**	1.1122	1.3212	$H=36$	0.9937	0.9946	1.0052
1.0418**	1.2058**	1.6206**	$H=48$	1.0011	1.0170	1.0463

Example 3: Model M_{AT}				SBC AverageModel		
$A=2$	$A=5$	$A=10$		$A=2$	$A=5$	$A=10$
1.0021**	1.0107**	1.0043**	$H=1$	1.0022**	1.0056**	1.0069**
1.0037**	1.0140**	1.0148**	$H=3$	1.0026**	1.0045	1.0063
0.9995	0.9973	0.9978	$H=6$	1.0058	1.0085	1.0120
0.9901	0.9843	0.9823	$H=12$	1.0037	1.0040	1.0094
0.9914	0.9871	0.9854	$H=24$	1.0022	0.9706	1.0283
0.9883	0.9817	0.9801	$H=36$	0.9890	0.9921	0.9697
0.9900	0.9846	0.9826	$H=48$	0.9968	1.0481	1.0072

(b) US-Japan

	Example 1: Model M_{EV}			Equal Weights Av. Model		
	$A=2$	$A=5$	$A=10$	$A=2$	$A=5$	$A=10$
$H=1$	1.0008	1.0041	1.0067	1.0017**	1.0056**	1.0101
$H=3$	0.9986	1.0172	1.0840	1.0037**	1.0177**	1.0433**
$H=6$	1.0109**	1.0813**	1.2260**	1.0085**	1.0373**	1.0604**
$H=12$	1.0266**	1.1763**	1.3988**	1.0185**	1.0738**	1.1162**
$H=24$	1.0389	1.3134	1.6905	1.0215**	1.0917**	1.1414**
$H=36$	1.0289	1.3210	1.5394	1.0061	1.0537	1.0811
$H=48$	1.0029	1.2786	1.5227	0.9893	1.0266*	1.0657

	Example 2: Model M_{EI}			AIC Average Model		
	$A=2$	$A=5$	$A=10$	$A=2$	$A=5$	$A=10$
$H=1$	1.0006	1.0037	1.0063	1.0007**	1.0026	1.0046
$H=3$	0.9984	1.0161	1.0817	1.0072**	1.0039**	1.0953
$H=6$	1.0101**	1.0772**	1.2148	1.0170**	1.0891**	1.1837**
$H=12$	1.0242**	1.1619**	1.3601	1.0294**	1.1517**	1.2802**
$H=24$	1.0284	1.2349	1.4873	1.0368	1.2556**	1.4984**
$H=36$	1.0121	1.1419	1.2161	1.0189	1.2022	1.3099
$H=48$	0.9838	1.0601	1.1247*	0.9912	1.1391	1.2506

	Example 3: Model M_{AT}			SBC AverageModel		
	$A=2$	$A=5$	$A=10$	$A=2$	$A=5$	$A=10$
$H=1$	1.0020**	1.0106**	1.0292**	1.0043**	1.0092**	1.0107**
$H=3$	1.0080*	1.0430	1.1003	1.0085**	1.0289**	1.0343**
$H=6$	1.0169*	1.0684	1.0897	1.0164**	1.0618**	1.0739**
$H=12$	1.0184**	1.0744*	1.0928**	1.0285**	1.1292**	1.1570**
$H=24$	1.0012**	1.0379**	1.0540	1.0352*	1.2575**	1.3635**
$H=36$	0.9550*	0.9494**	0.9536	1.0228	1.3365	1.6179
$H=48$	0.9260**	0.8977	0.8890	0.9943	1.2621	1.6911

Notes: See notes to Table 7.