A Meta Model Analysis of Exchange Rate Determination[*]

by

Chrystalleni Aristidou[†], Kevin Lee[†] and Kalvinder Shields[††]

Abstract

A novel approach to modelling exchange rates is presented based on a set of models distinguished by the drivers of the rate and regime duration. The models are combined into a 'meta model' using model averaging and non-nested hypothesis-testing techniques. The meta model accommodates periods of stability and slowly-evolving or abruptly-changing regimes involving multiple drivers. Estimated meta models for five exchange rates provide a compelling characterisation of their determination over the last forty years or so, identifying 'phases' during which the influences from policy and financial market responses to news succumb to equilibrating macroeconomic pressures and vice versa.

Keywords: Exchange Rates, Model Averaging, Structural Breaks, Non-Nested Testing

JEL Classification: C51, F31, F47.

1

1 Introduction

Obstfeld and Rogoff (2009) cite the weak relationship between the exchange rate and the rest of the economy as one of the major puzzles in international macroeconomics: the so called "exchange rate disconnect puzzle".[1] Engel *et al.* (2008) provide a useful framework which considers the exchange rate to depend on fundamental drivers and expected future rates and which highlights some of the sources of the disconnect. The framework accommodates any model incorporating Uncovered Interest Parity where, in this case, the drivers are the set of variables chosen to account for the behaviour of the interest rate differential. In practice, if the focus of attention is on the short-term, the set of variables chosen to account for interest rate movements are those best able to capture the effects of policy responses or financial market responses to news. When attention is on the macroeconomy's longer-term adjustment to its steady-state level, variables that capture broader equilibrating pressures on interest rates are considered more appropriate.[2] Of course, in reality, both sets of influences on the exchange rate could play a role at any one time, with their relative importance likely to change over time depending on the extent of business cycle shocks and turbulence in the financial markets and the scale and speed of changes in countries' longer term macroeconomic

---

[1]The disconnect lies behind the difficulties involved in forecasting exchange rates which have been well-rehearsed since Meese and Rogoff's seminal (1983) paper where point predictions from a driftless random walk model were no worse than those from more sophisticated models. See Rossi (2013) for a review on the literature on exchange rate predictability.

[2]On the basis of a survey of US foreign exchange traders, Cheung and Chinn (2001) report that conventional macroeconomic pressures are thought to be important for exchange rate movements by 1% of traders at the intraday horizon, but by 59% of traders in the medium run (i.e. up to 6 months) and by 88% of traders in the long run (i.e. over six months). Comparisons with the results of earlier surveys also lead them to conclude that these rankings of variables change substantially over time.

[1]

outlook. As Engel *et al.* point out, the forward-looking nature of exchange rate determination compounds these difficulties, shifting the influence from the current value of the interest rate fundamentals to their expected future paths. This means the relative weight of the different sets of influences can change in anticipation of changes in future policy or macroeconomic outlook as well as in response to contemporaneous changes.

This inherent instability poses difficulties in applied work and is an explanation for why no single exchange rate model performs well in explaining or forecasting different currencies over different samples. Researchers have attempted to accommodate structural instability in exchange rate models through single-equation time-varying parameter models (see Wolff (1987) or Schinasi and Swamy (1989), for example) and through Markov-switching models (as in Engel (1994) for example). Model averaging, in which a variety of models are estimated - recursively or with a rolling window - and then combined with time-varying weights has also been employed. This can be approached as a full Bayesian exercise - as in Wright (2008) or Byrne et al. (2017) for example - with the weights defined by an estimated posterior probability that the model holds true, or following a more standard forecast-combination approach in which, at each time, all models are given equal weight or a weight based on 'out-of-sample' performance in a recent training period (see, for example, Sarno and Valente (2009)).

In this paper, we adopt a model averaging approach to deal with the inherent structural instability in exchange rate determination but we emphasise the 'regime uncertainty' surrounding the length of time for which a set of fundamentals exerts its influence as well as the 'model instability' surrounding the choice of fundamentals. This follows the suggestion of Pesaran and Timmermann (2007) to apply model averaging techniques to alternative models of the same type but estimated over different

[2]

estimation windows. We allow for uncertainty across model fundamentals as in the literature then, but we also pay explicit attention to the duration of the period over which the different fundamentals are relevant. This distinguishes the approach from those exchange rate papers where the time-variation is introduced implicitly through the recursive nature of the modelling or through the application of a simple rolling window. The more implicit approach may be reasonable in forecasting exercises but it could obscure important regime shifts when the model averaging exercise is conducted to make economically-meaningful inferences. We use the term 'meta modelling' to flag our emphasis on regime instability when compared to more usual model averaging methods.[3]

We also introduce a novel approach to constructing the time-varying weights in our model averages by adopting non-nested hypothesis-testing methods. Here, a characterisation of the data generating process based on a particular combination of fundamentals continues until there is evidence to reject it in favour of a new characterisation. Non-nested testing methods are involved as the new characterisation could be based on a very different combination of fundamentals. The approach has the advantage that, to the extent that it is warranted by the data, it builds in a degree of stability in the characterisation over time by taking the current model as the maintained hypothesis. Given the volatility of exchange rates, this is a feature that is often missing from models driven purely by Bayesian updating or weights based on forecast performance over the recent past and this undermines those models' ability to provide an economic narrative to explain the changes over time.[4]

---

[3]Lee et al. (2013, 2015) provide descriptions of the conduct of monetary policy in the UK and US based on estimated "meta-Taylor rules" for the two countries, obtained using similar methods to those of this paper, where the duration of different policy regimes is an important focus of interest.

[4]See Timmermann (2006) and Aiolfi et al. (2011) for discussion of the approaches taken to model averaging in the forecasting context.

[3]

In the next section, we briefly comment on some traditional models of exchange rate determination to motivate the use of different fundamentals in different models and our characterisation of these as reflecting policy or financial market responses to news or equilibrating macroeconomic pressures. Section 3 elaborates on the model averaging approach that we adopt to construct our meta model. The methods are applied to monthly data for the exchange rates of five currencies against the US dollar spanning over the last forty or fifty years in Section 4. Exchange rate determination in the countries is characterised here according to a series of phases in which there is an ebb and flow between the pressures on the exchange rate from policy and financial market responses to news and those from longer-term macroeconomic adjustments. Section 5 provides concluding remarks.

2 Exchange Rate Fundamentals and Structural Uncertainty

There are four structural models of exchange rate determination frequently found in the literature which we characterise as being more or less relevant during periods of economic turbulence or stability.[5] Exchange rate movements during periods experiencing extreme policy responses or financial market responses to news - 'news' pressures - might be best characterised by models based on the Uncovered Interest Rate Parity (UIP) relationship

$$s_t = s_{t+1}^e - (r_t - r_t^*), \tag{2.1}$$

where S_t is the nominal exchange rate at t, defined as a home price of a unit of foreign currency, R_t and R_t^* are the nominal interest rates paid on domestic and foreign assets during period t respectively, the 'e' superscript indicates expectations (formed at time t) and lower case variables denote logarithms. These include models based

---

[5]Rossi (2013) provides detailed descriptions of these models and the evidence relating to them.

on *interest rate parity fundamentals* in which, iterating forwards, taking expectations and assuming that the expected future interest rate differential follows a simple $AR(p)$ specification, we can write

$$s_t - s_{t-1} = \lambda + \sum_{i=0}^{p} \lambda_i (r_{t-i} - r^*_{t-i}) + \eta_{IRP,t}, \tag{2.2}$$

where the λ's are parameters and $\eta_{IRP,t}$ represent stationary innovations. Alternatively, working with the determinants of the interest rate as expressed in the Taylor rule (i.e. inflation Δp_t and the output gap x_t) and assuming these influences effect domestic and foreign interest rates, a model based on Taylor rule fundamentals would be written as

$$s_t - s_{t-1} = \gamma + \gamma_r \, r_{t-1} - \gamma_{r*} \, r^*_{t-1} + \gamma_{\Delta p} \, \Delta p_t - \gamma_{\Delta p*} \, \Delta p^*_t + \gamma_x \, x_t - \gamma_{x*} \, x^*_t + \gamma_q \, q_t + \eta_{TR,t}, \tag{2.3}$$

where $q_t = s_t + p^*_t - p_t$ is the log of the real exchange rate, the γ's are parameters and $\eta_{TR,t}$ again represent stationary innovations.

In less turbulent times, the future path of interest rates will reflect broader macroeconomic conditions and might be better captured by Mark's (1995) approach to modelling the exchange rate in which deviations of the nominal exchange rate from its equilibrium are gradually eliminated over time according to

$$s_t - s_{t-1} = \beta \, (f_{t-1} - s_{t-1}) \tag{2.4}$$

where $f_t - s_t$ is the deviation of the time-t equilibrium exchange rate, f_t, from the actual rate. The *Purchasing Power Parity (PPP)* hypothesis provides a candidate for the equilibrium level of the exchange rate based on the 'law of one price' so we can write

$$s_t - s_{t-1} = \phi + \beta(p_{t-1} - p^*_{t-1} - s_{t-1}) + \eta_{PPP,t} \tag{2.5}$$

[5]

where $\eta_{PPP,t}$ again reflects stationary innovations. Alternatively, the monetary model of the exchange rate characterises the equilibrium exchange rate as depending on relative money supplies, relative income levels and the interest rate differential and can motivate a model of the form

$$s_t - s_{t-1} = \rho_0 + \beta \left[\sum_{i=0}^{p} \rho_{mi}(m_{t-1-i} - m^*_{t-1-i}) + \sum_{i=0}^{p} \rho_{yi}(y_{t-1-i} - y^*_{t-1-i}) - s_{t-1} \right] + \eta_{MM,t}$$
(2.6)

Modelling structural uncertainty   The four models outlined in (2.2), (2.3), (2.5) and (2.6) are all relatively standard in the literature. Our discussion emphasises that any one of them, or a combination of them, could be more or less relevant in different circumstances and over different sample windows. In what follows then, at time T, there are $N \times J$ models that can be used to characterise recent changes in the exchange rate, described by

$$M_{ijT}: \qquad s_t - s_{t-1} = \boldsymbol{\alpha}_{ijT} \mathbf{X}_{it} + e_{ijT,t} \ , \qquad i = 1, ..., N; \quad j = j_{\min}, ..., j_{\max}; \quad t = T - j, ..., T,$$
(2.7)

where $J = j_{\max} - j_{\min} + 1$ and the complexity of the subscripts reflects the flexibility of the modelling framework. Here, model M_{ijT} is assumed to explain the change in the exchange rate over the period $T - j, ..., T$, and allowing j to vary means we contemplate models that might be relevant only for the most recent past or back to $j_{\max}$ periods in the past. The model involves $\mathbf{X}_{it}$ which is the ith set of N alternative sets of explanatory variables driving the exchange rate; these represent the fundamentals proposed by interest rate parity, the Taylor rule, PPP and the monetary models respectively in the case of the models in (2.2)-(2.6) and, for models of the form in (2.4), the lagged exchange rate level also.

[6]

3 Meta Modelling

The uncertainty surrounding the determination of exchange rates is reflected by the idea that the change in the exchange rate observed at any particular time T could be explained by any one of $N \times J$ different models according to (2.7). The meta modelling approach accommodates this uncertainty by using a weighted average of the alternative models in (2.7). The approach starts from a Bayesian Model Averaging (BMA) formula but is classical in nature avoiding the (often problematic) Bayesian assumption that the model includes the true data generating process (dgp) and avoiding the need to specify prior probabilities for the unknown parameters in the models or for the models themselves. Indeed, the estimated meta model simply aims to characterise exchange rate movements taking account of the possibility of changes in the relative importance of the fundamentals at different times. As we shall see, the model weights in our preferred meta model are updated in each period on the basis of non-nested hypothesis tests, accommodating the possibility of a structural break by switching to an alternative structural model if there is evidence to reject the previously-held null. The meta model *could* reflect the true dgp if we know one of the fundamental models under consideration holds true at all times or if, for example, exchange rate decisions are made by different groups - each focused on different fundamentals - and the weights capture the proportions of individuals in the respective groups as these change over time. But the meta modelling approach is also consistent with the true dgp being distinct from all of the underlying structural models considered. Here the weights simply convey the real-time adequacy of the underlying structural models in characterising recent exchange rate movements, and a reduction of the weight on a model because of its rejection does not imply acceptance of its alternative but simply reflects the shortcomings of the previously-held null.

[7]

3.1 Model Averaging

The basis of the meta modelling approach is the BMA formula:

$$\Pr(\boldsymbol{\theta}_T \mid \mathbf{Z}_T) = \sum_{i=1}^{N} \sum_{j=j_{\min}}^{j_{\max}} \Pr(\boldsymbol{\theta}_T \mid \mathbf{Z}_T, M_{ijT}) \times \Pr(M_{ijT} \mid \mathbf{Z}_T) \qquad (3.8)$$

where $\mathbf{Z}_T = (\mathbf{z}_1, ...\mathbf{z}_T)$ represents the data available at T, with $\mathbf{z}_t = (s_t, \mathbf{X}_{it} \; \forall \; i)$, and $\boldsymbol{\theta}_T = (\beta_{ijT} \;\; \forall \; i, j)$ represents the unknown parameters capturing the influence of all the fundamentals under consideration. The $\Pr(\boldsymbol{\theta}_T \mid \mathbf{Z}_T)$ describes our understanding of the parameters of interest and M_{ijT} represent the various models described at (2.7). The BMA formula decomposes the uncertainties accommodated within $\Pr(\boldsymbol{\theta}_T \mid \mathbf{Z}_T)$ into a weighted average of the conditional distributions, $\Pr(\boldsymbol{\theta}_T \mid M_{ijT}, \mathbf{Z}_T)$, using as weights the model probabilities $\Pr(M_{ijT} \mid \mathbf{Z}_T)$. A strict Bayesian requires a prior distribution for the unknown parameters of all the models to evaluate the conditional distributions. Alternatively, if no meaningful prior distribution is available, one can make the more classical assumption that

$$\boldsymbol{\theta}_T \mid \mathbf{Z}_T, M_{ijT} \;\; \overset{a}{\sim} \;\; N(\widehat{\boldsymbol{\theta}}_T, \widehat{\mathbf{V}}_T)$$

for $\Pr(\boldsymbol{\theta}_T \mid \mathbf{Z}_T, M_{ijT})$, where $\widehat{\boldsymbol{\theta}}_T$ is the familiar maximum likelihood estimate of the parameters under M_{ijT}, and $\widehat{\mathbf{V}}_T$ is the asymptotic covariance matrix of $\widehat{\boldsymbol{\theta}}_T$. This assumption treats $\boldsymbol{\theta}_T$ as a random variable at the inference stage so that $\Pr(\boldsymbol{\theta}_T \mid \mathbf{Z}_T, M_{ijT})$ in (3.8) is approximated by $N(\widehat{\boldsymbol{\theta}}_T, \widehat{\mathbf{V}}_T)$ and standard inference can be carried out for each model in turn.

3.2 The Model Weights

Turning to the model weights $\Pr(M_{ijT} \mid \mathbf{Z}_T)$, we note that in the context of exchange rate determination, where the models under consideration are unlikely to be exhaustive even allowing for structural breaks, the strict Bayesian requirement to assign prior

[8]

probabilities to all models at each point seems unrealistic, or at least very demanding. The alternative 'frequentist' model averaging approaches found in the literature are reviewed in Steel (2020) noting that here weights are chosen to deliver parameters with desirable properties under repeated sampling. The relative advantages of the different frequentist approaches considered in the review make little or no reference to models of different sample lengths though (and, indeed, the desirable properties are typically related to the asymptotic properties of the estimators). This reinforces our use of the term 'meta modelling', with its focus on regime uncertainty and the choice of sampling window, to distinguish it from the more usual context for model averaging.

The Meta (Non-Nested Testing) Approach A pragmatic approach to deriving model weights in these circumstances is to allow these to evolve over time, updating the weights in each period to reflect new evidence on whether the previously-held view continues to be valid or whether an alternative new-born model is now appropriate. Since the new-born model could involve an entirely different set of fundamentals to those of the previously-held model, the evidence involves non-nested hypothesis-testing (NNT) methods which are relevant when one model cannot be obtained from the other by imposition of parameter restrictions or through a limiting process.

The meta-NNT approach can be formalised by writing, for any T and for all $i = 1, ... n$ and $j = j_{\min}, ..., j_{\max} - 1$,

$$\Pr(M_{i,j,T-1} \mid \mathbf{Z}_{T-1}) \quad \rightarrow \quad \begin{cases} \Pr(M_{i,j+1,T} \mid \mathbf{Z}_T) & \begin{array}{l} \text{if the null } M_{i,j+1,T} \text{ is not rejected in favour of } M \\ \text{for } g = j_{\min},...,j \text{ and } k = 1, . \end{array} \\[2em] \Pr(M_{k,g,T} \mid \mathbf{Z}_T) & \begin{array}{l} \text{if the null } M_{i,j+1,T} \text{ is rejected in favour of } M_{k,g} \\ \text{for } g = j_{\min},...,j \text{ and } k = 1, .., \end{array} \end{cases}$$

so that the weight assigned at time $T-1$ to the model containing the ith fundamentals

[9]

and based on data $T - 1 - j$ to $T - 1$ is either transferred to the model with the same fundamentals based on one additional observation - i.e. data $T - 1 - j$ to T - or to a new model based on the shorter sample of data $T - g$ to T containing any one of the alternative sets of fundamentals based on a non-nested test. If a model is rejected in favour of more than one alternative, the weight can be split equally among the alternative models.[6]

In transferring weights, our interest is whether the most recent observation confirms or flags shortcomings on our currently-held characterisation of the data. A natural statistic on which to base the test between the models is the ratio of the squared residuals obtained for the final observation of the two competing models, denoted $\eta_{ij,kg,T}$ say. Here, a large (absolute) value of the residual from the null model casts doubt on its continued relevance, but this is judged relative to the performance of the realistic alternative models. In the case where the alternative is the same behavioural model but with changed parameters based on a shorter sample period, the alternative is nested within the null and the statistic provides a standard F-test of structural instability, itself a likelihood ratio test under the assumption of normally-distributed errors. But, more generally, neither model is nested within the other and non-nested testing procedures are required. The 'Cox test' of two competing non-nested models involves modifying the likelihood ratio test statistic to obtain a statistic with known asymptotic distribution. The modification is required because, taking one model as the null, the alternative is misspecified and its estimated likelihood will depend on the parameters of the null model.[7] In most cases, the required modification renders

---

[6]Alternatively, as illustrated in the empirical exercise below, the weights could be reallocated according to the strength of the rejection (denoted the 'meta-NNTp' approach).

[7]Pesaran (1974) describes the modification required to take into account the misspecification in the case of two non-nested linear regression models estimated over a common sample and derives a statistic which is asymptotically normally distributed with zero mean and calculable finite variance.

[10]

the distribution of the statistic analytically intractable so that simulation methods are required.

The simulation exercise involved here is computationally demanding but relatively straightforward. Here the previously-held model has a clear status as the null and so can be used to simulate R artificial samples of the exchange rate, $s_t^{(r)}$, $r = 1, .., R$, for $t = T - j, ..., T$ using the estimated parameters of model $M_{i,j,T}$ and making random draws from a Normal distribution with mean zero and variance equal to that estimated under $M_{i,j,T}$. For each artificial sample, the models $M_{i,j,T}$ and $M_{k,g,T}$ can be estimated and the ratio of the squared residuals obtained for the final observation of the two competing models, $\eta_{ij,kg,T}^{(r)}$, can be calculated. The set of simulated $\eta_{ij,kg,T}^{(r)}$ statistics provides the appropriate distribution against which to compare the observed $\eta_{ij,kg,T}$ under the null that model $M_{i,j,T}$ is true. Finding that this value lies in the upper 5%, say, of the simulated distribution provides significant evidence to reject the model in favour of the new alternative. Carrying out this exercise at each point in time, holding in turn each model with non-zero probability as the null and comparing it to all realistic alternative models, provides the means to update the weights over time.

Alternative Frequentist Model Averaging Approaches The meta-NNT approach is related to Hansen et al.'s (2011) idea of a Model Confidence Set (MCS) in which a test is applied to a set of competing models and models are eliminated if they perform poorly by some user-specified criterion. The MCS is the set of (equally weighted) models which are not rejected as statistically inferior. In the meta-NNT approach, as we move through the sample, the weight from each model characterising exchange rate determination in one period is effectively transferred to the models in

See Pesaran and Weeks (2003) for a review of the non-nested testing literature.

[11]

its MCS in the next period.

The use of NNT in allocating weights has the advantage of building in a degree of stability in the weights over time through the 'protection' provided to the null. Lee et al's (2015) approach to defining the 'meta' weights also builds in a degree of stability by updating time-t weights at $t+1$ according to the probability of observing the time $t+1$ outcome based on the time-t model, where this latter probability is assumed proportional to the squared estimated residual at the end of the sample.[8] Compared to the meta-NNT approach, the updating criterion in this meta model is more closely related to the approach to defining weights found in the forecast combination literature. Here, weights depend on the historical forecasting performance of the different models, sometimes discounting into the past or focusing on the "most recent best" (MRB) forecasts; see, for example, Diebold and Pauly (2007) or Sarno and Valente (2009) for discussion.

Comparison of the meta-NNT and meta-MRB approaches will provide insights on the role of the updating criterion. Comparison with a more standard model averaging approach in which weights are updated as above but based on the MRB performance of models estimated in a rolling window of fixed-sample-length would further isolate the 'meta' contribution of accommodating regime change. And an exercise in which weights are based only on MRB performance of a rolling model average, with no updating, would reveal the role of the smoothing. These exercises are considered in the empirical work below.

The Meta-NNT Model The meta-NNT model characterising exchange rate determination over the whole sample $T = \underline{T}, ..., \overline{T}$ consists of the set of individual

---

[8] The weights also accommodate the possibility of new regimes being born with a fixed probability; see Lee at al (2015) for details.

[12]

estimated models of the form given by (2.7) plus the associated weights obtained using the non-nested testing procedure described above. Denoting the weights by $w_{i,j,T} = \Pr(M_{i,,j,T} \mid \mathbf{Z}_T)$, the meta-NNT model can be written as

$$\overline{M}_{\cdot,\cdot,\cdot} = \{M_{i,j,T}, \, w_{i,,j,T} \quad , \quad i = 1, ..., N, \quad j = j_{\min}, .., j_{\max}, \quad T = \underline{T}, ..., \overline{T}\}$$

(3.10)

which attaches weights to all the possible models in (2.7) defined according to the definition of exchange rate fundamentals and to the different regime lengths.[9]

Changes in the size of the weights over time provide useful information on how exchange rate determination has evolved therefore. For example, the *duration statistic,* $D_T = \sum_i \sum_j \, j \times w_{i,j,T}$ provides a time-T indication of the duration of the exchange rate regime in place at that time (whatever the nature of the regime). Similarly, the *behavioural model weight* statistic, $F_{i,T} = \sum_j w_{i,j,T}$, $\quad i = 1, .., N$, provides a time-T summary of the usefulness of the ith of the alternative exchange rate models.[10]

4   Characterising Exchange Rate Determination for Five Currencies

We now apply the modelling approach to the analysis of the determination of five exchange rates over the last forty or fifty years; namely, the U.S. dollar (USD) exchange rates for the Canadian dollar, Danish krone, Japanese yen, Swedish krona and British pound. The data are measured monthly and are as provided in Rossi (2013), with the start dates for the analysis varying across countries to accommodate the differences in the dates at which the currency prices are considered to be floating

---

[9]The meta-MRB model takes the form of (3.10) but with weights determined according to the most-recent-best performance as described above and might be denoted $\widetilde{M}_{\cdot,\cdot,\cdot}$, say. The corresponding rolling window model average would then be denoted $\widetilde{M}_{\cdot,s,\cdot}$ with window size s.

[10]The Shannon entropy statistic $H_T = -\sum_i \sum_j \log(w_{ijT}) \times w_{ijT}$, as used in information theory, provides a useful summary measure of the extent of model uncertainty experienced at T.

[13]

- as described in column (1) of Table 1 - but all running to 2010:06.[11] These data are derived originally from Datastream but were collated by Rossi to provide a set of variables that are reasonably comparable across countries. The choice of our five rates was based on the availability of a long run of data, and the results for these rates presented by Rossi provide a useful setting from which to judge our own results.

To be clear on definitions, the data for nominal exchange rates s_t are the end-of-month observations of the rate expressed as the price of one US dollar. Interest rates r_t are three-month Treasury Bill rates, output y_t is measured by monthly industrial production figures and the output gap x_t is calculated as the percentage deviations of actual industrial production from the trend defined by applying a simple moving average to a forecast-augmented industrial output series.[12] Prices p_t are measured by CPI and we use relatively liquid measures of the money supply m_t in each country (e.g. M1 data for the US). Series are seasonally-adjusted using one-sided moving averages with equal weights over the previous twelve months.

A plot of the (logarithm of the) five exchange rates, and corresponding price and interest rate differentials, are provided in the on-line Appendix and show reasonably clear similarities in the movements of each country's exchange rate and its prices relative to those of US over the forty or fifty years of the data sample. For instance, broadly speaking, the Canadian dollar rate s_t^{CAN} rises to the mid-eighties, falls through to the early nineties, rises again through to early 2000's, drops sharply

---

[11]The early years of data in some countries include observations during regimes of highly managed exchange rates. But these early observations provides a convenient way to initiate the modelling and, of course, the modelling strategy is specifically designed to disregard these observations in later years - by moving to shorter samples - as dictated by the data.

[12]Specifically, at each period T, an AR(2) model was estimated for the output series and used to produce forecasts for $T+1, .., T+12$. Trend output at T was identified as the value of the 24-month moving average centred at T applied to the extended series.

[14]

to 2007/8, rises briefly and then falls again at the end of the sample. Exactly the same description applies to Canadian relative prices $p_t^{CAN} - p_t^{US}$. In contrast, and again broadly speaking, the Japanese Yen falls gradually throughout the sample. But so too do Japanese relative prices. As shown in column (2) of Table 1, the simple correlations between each country's exchange rate and its relative prices is high in all five countries, averaging 0.77, showing the importance of broad price pressures for exchange rate determination. On the other hand, these relationships are not one-for-one and divergencies in the movements between the two series appear to persist. Simple ADF tests applied to the entire sample of data show, for all five currencies, that the nominal exchange rate and relative price series are both I(1) and, importantly, that the real exchange rate $q_t = e_t + p_t^{US} - p_t$ is also I(1). In short, price pressures do appear to impact on the exchange rate but, given the periodic and permanent shifts in the series, it seems unlikely that exchange rate determination will be fully captured by a stable PPP or monetary model.

Despite these broad patterns, each country's relative price movements are much smoother over time than those of its exchange rate. Column (4) of Table 1 shows the variance of the change in the exchange rates relative to the variance of the change in relative prices is very large in every country, averaging 67 time larger across these five countries. The volatility of exchange rates is much more in line with the volatility of the interest rate differentials, with the ratio of these variances averaging 1.6 across our five respective countries. This suggests that the asset market pressures captured by the IRP and Taylor Rule models could provide a more important influence on exchange rates over short horizons. On the other hand, the simple correlations in each country between exchange rate changes and the interest rate differential over the whole sample, as reported in column (2) of Table 1, have an average of just -0.09 making it very unlikely that the IRP and Taylor rule models could provide the basis

[15]

for explaining exchange rate movements over the sample in all the countries.

An intuitive account that is consistent with these statistics is that there are equilibrating macroeconomic pressures to move exchange rates towards establishing PPP. But there are also factors that change the relationship between exchange rates and relative prices permanently, and there are jumps and volatile movements in the exchange rates arising in response to news from global markets that are best represented by an IRP or Taylor Rule relationship. The relative strengths of these various pressures varies over time and the meta model allows them all to have an effect, with individual models having non-zero weight while their influence is apparent in the data.

4.1 The Meta-NNT Models

Our modelling work began by estimating, for each country, 12 versions of our four fundamental models based on three years of available data running up to the beginning of the period of analysis reported in column (1) of Table 1; e.g. up to $1965 : 5$ for Canada. The different versions used data ranging between 24 months and 36 months prior to the beginning of the period of analysis, providing estimates of $M_{i,j,T}$ for $i = PPP, MM, IRP, TR$, for $j = 24, ..., 36$. In this first iteration of the modelling, equal weights was given to all $4 \times 12 = 48$ models obtained for each country. The data window was then extended by one month and 52 models were estimated for each country but in this case the weights were assigned to each model following the procedure in (3.9). This iterated procedure then continued for every T up to the end of the period of analysis in $2010m6$. The estimated models and model weights obtained in this way provided the estimated 'meta-NNT model' for each country.

Figure 1 provides graphical representations of the key features of the meta-NNT models obtained for Canada; the equivalent figures for the other countries are provided in the online Appendix. The figure shows the weighted average of the sample length

[16]

- i.e. the duration statistic described earlier - as it evolves over time. The figure demonstrates clearly the extent of the regime instability that has to be captured by exchange rate models: there are runs of data during which the average sample length grows by one month each month, implying no breaks are occurring and weights are simply passed on to the corresponding models which are one month longer; but these runs rarely last for more than two years with the average sample length frequently dropping below 30 months. For Canada between 1968:4-2010:6, the average sample length falls below 30 months on 28 occasions - and often remains there for some time - meaning that there is a distinct break every 19 months or so. Similar instabilities are observed in the other countries and, as documented in column (5) of Table 1, the intervals between distinct breaks is very similar (17-19 months) in Denmark, Japan and Sweden although they are a little less frequent in the UK (with intervals between breaks averaging around 28 months).

The figure also provides some insights on the nature of the breaks in the Canadian case by plotting smoothed measures of the joint weights of the models driven by macroeconomic pressures (i.e. the PPP and Monetary models) and of the models driven by the response to news (i.e. IRP and Taylor rule models), where the smoothing averages over the previous 24 months.[13] While the smoothing obscures the impact of some very short-lived episodes, it provides a clear picture of the broad patterns in the weights and reflects the timing of substantial shifts in behaviour. Broadly speaking, it appears that the PPP and (less frequently) the Monetary model dominate during the sixties and seventies and then again between the mid-1990's and mid-2000's, while the IRP and Taylor Rule models dominate from the mid-80's to the mid-90's. The figure also defines some 'phases' during which macroeconomic pressures succumb to the 'news' pressures and vice versa. A phase in which the response to news dominates is

---

[13]Detailed plots of the underlying weights are provided in the on-line Appendix.

[17]

defined to end when there is a peak in the weights assigned to the IRP and Taylor Rule models. The phase of growing macroeconomic pressures that follows corresponds to a period when the weight on the PPP and Monetary models is increasing, and it ends when the weight on these models peaks.[14] The outcome for Canada is presented as shading in Figure 1 and divides the sample into thirteen phases, each lasting 42 months on average and confirming the earlier broad characterisation of growing influence of news responses in the early 80's and early parts of the 2000's, and the long period between the mid-80's and mid-90's when equilibrating macroeconomic pressures grew to dominate. A similar phasing can be obtained for the other countries, and the summary statistics in column (6) of Table 1 show that these countries experienced similar episodes (albeit with rather shorter phases in Japan and Sweden). However, the timing and nature of the phases are country-specific and reflect the particular circumstances of each country.

4.2 Comparison with Alternative Frequentist Model Averages

The meta-NNT models build in stability in the weights over time through the protection of the null hypothesis that any model with non-zero weights remains relevant, with the same weight, next period. Columns (a) and (b) of Table 2 provides statistics to compare the meta-NNT model with two alternative models that consider the same range of models as the meta-NNT model but with alternative ways of accommodating this stability. Specifically, the meta-NNTp model applied the same modelling strategy as meta-NNT but, when the null is rejected, the weights are reassigned according to the strength of the rejection.[15] And, as discussed earlier, the meta-MRB model

---

[14]Some judgement is exercised in the definition when a change in direction is very short-lived.

[15]Here, rather than reallocating weights equally to all models against which the null is rejected at the 1% level, weights are reallocated among these models in (inverse) proportion to the rejection

focuses on the most-recent-best performance of the models and updates the previous period's weights according to the likelihood of observing the most recent observation as in Lee et al. (2015).

Table 2 column (a) reports, for each currency, the simple correlation between the weights assigned over the sample to the PPP+Monetary models by the meta-NNT model and the meta-NNTp model, and also the degree of agreement over the phases identified by the two models as captured in a 2×2 contingency table. The average correlation over the five currencies is 0.82, while the average Kuipers Score is 0.59 and both the static and dynamic versions of the Pesaran-Timmerman (2009) tests show significantly in every case. These statistics reflect a relatively high degree of consistency across the two approaches then, so that the use of the rejection probability in reallocating weights has an effect but does not substantially alter the results. Column (b) compares the meta-NNT and meta-MRB models and shows an average correlation of 0.44 and average Kuipers score of 0.59. The correlations are reasonably high then and there is some consistency in the timing of phases, although this is statistically significant in only three cases according to the static PT test and only for the GBP according to the more powerful dynamic PT statistic.

Columns (c) and (d) compare the meta-NNT models with models obtained following the more usual rolling model averaging (RMA) approach in which all the estimated models are based on a rolling fixed sample length (of 66 observations in this case). Column (c) reports on the model in which the weights are updated at $t + 1$ on the basis of the probability of observing the time $t + 1$ outcome as in meta-MRB, while column (d) assigns weights purely on the basis of this time $t + 1$ probability without reference to the previous period's weights. The correlations and Kuipers scores

probability, with a model rejecting at 0.2% level gaining five times more weight than a model rejecting at 1%, for example.

[19]

of these columns show that these models have a quite different characterisation of exchange rates to those of the meta models. The average correlations across the five currencies are -0.18 and 0.12 in columns (c) and (d) respectively, while the Kuipers scores are close to zero or negative with no statistical significance for any currency.

Underlying these results is the observation that, at any time, none of the sub-set of models considered by the RMA (all of which are restricted to have the same sample length) performs especially well relative to the others in explaining the most recent outcome. This is as would be expected if there are frequent structural breaks. In the case of the dynamic RMA model of column (c) involving updated weights, this means that the models are typically unable to disrupt the weighting profile from the previous period and, if the weights become concentrated on one model, it remains unchanged over time (so that, as it turns out, the number of phases observed during our data period is fewer than seven in all five of our currencies). This is in contrast to the meta models where, after a break, a model estimated on a short sample can substantially outperform those based on longer samples. For the static RMA model of column (d) where weights are based only on the most recent performance with no updating, the fact that none of the models performs especially well relative to the others means that none of the structural models dominates for any length of time, the sum of weights on the PPP+Monetary models changes rapidly over time, and it is difficult to distinguish any particular phases in the data. Clearly then, *while the use of NNTp or MRB in updating weights has some impact on the characterisation of exchange rate determination in the meta models, it is the ability of the meta models to accommodate structural breaks that dominates the nature of this characterisation.*

[20]

5 Conclusion

There is inherent instability over time in the process determining exchange rates and it is not surprising that explaining exchange rate movements and forecasting them is difficult in these circumstances. The model averaging underlying the meta model of this paper provides a very flexible approach to dealing with this inherent instability in real time. The approach accommodates regime uncertainty as well as model uncertainty, doing this in a way that can account for periods of stability, periods in which policy evolves gradually and episodes of abrupt changes in regime. The results of the paper show that, for the five currencies considered, the meta-NNT models provide sensible characterisations of exchange rate movements over the last 40-50 years, reflecting the ebb and flow of macroeconomic and 'news' pressures on exchange rates. The timing of the phases of the different pressures are country-specific, reflecting countries' individual experiences. But there is a striking similarity in the frequency of structural breaks (occurring every 17-28 months on average) and the duration of the phases in which macroeconomic or 'news' pressures dominate (lasting 29-42 months on average). Comparison with alternative model averaging approaches shows that it is the meta model's ability to accommodate structural change that is central to its success in characterising the data, although the use of non-nested testing - as opposed to updating using most-recent-best criteria - is also important in building some useful stability in the evolution of model weights over time.

Figure 1: Weighted Average Sample Size and Smoothed Weights for PPP and Monetary Models for Canada: 1968m4 - 2010m6
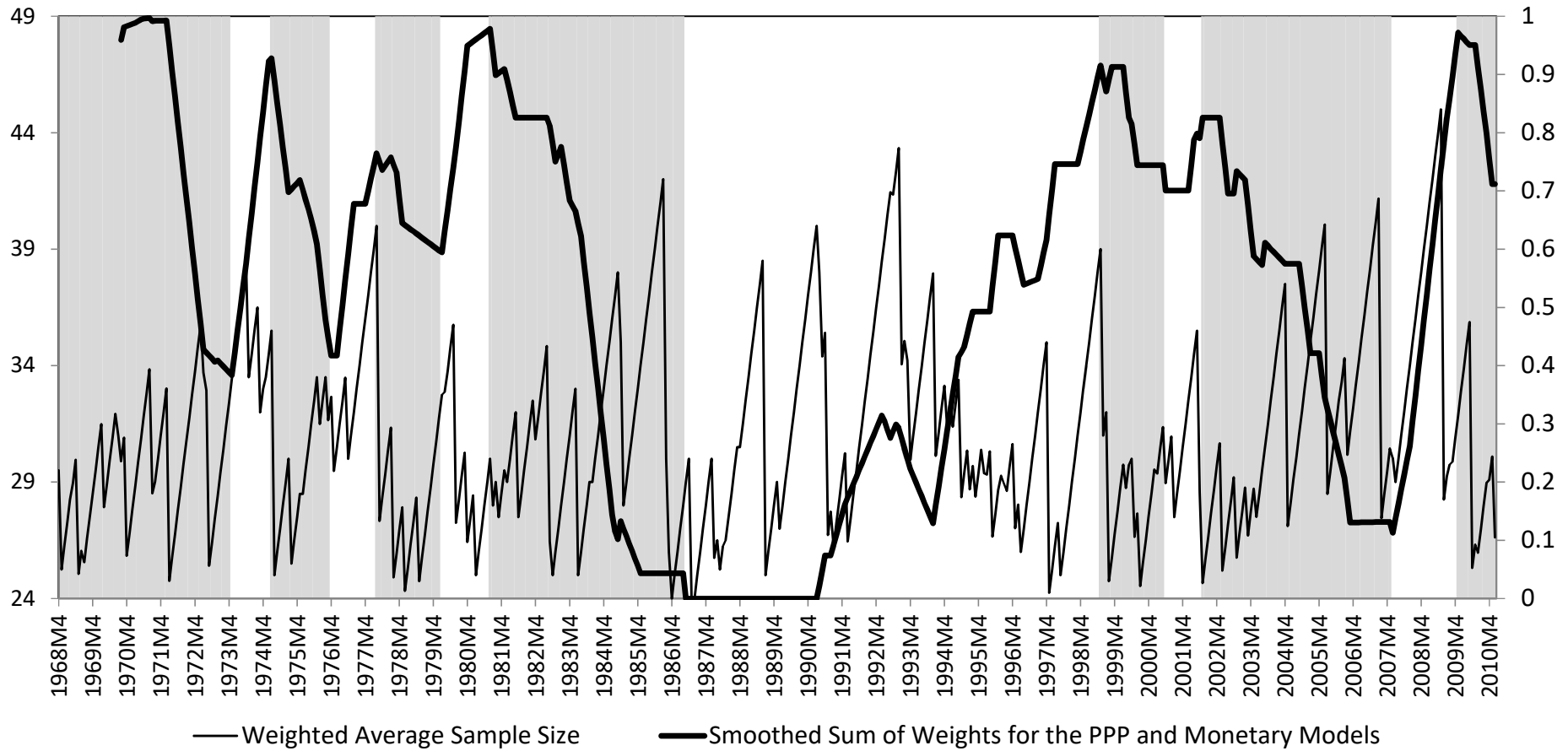
——— Weighted Average Sample Size   ——— Smoothed Sum of Weights for the PPP and Monetary Models

Table 1: Exchange Rate Summary Statistics

	(1)	(2)	(3)	(4)	(5)	(6)
	Period of Analysis (n)	$\rho(\ s_t,\ p_t^{US}-p_t)$	$ADF(s_t)$	$V(\Delta s_t)\big/ V(\Delta(p_t-p_t^{US}))$	No. breaks (av. duration)	No. phases (av. Duration)
		$\rho(\ s_t,\ r_t-r_t^{US})$	$ADF(s_t-p_t^{US}-p_t)$	$V(\Delta s_t)\big/ V(\Delta(r_t-r_t^{US}))$		
Canada	1968:4-2016:6 (518)	$0.89^{(**)}$ 0.00	0.31 0.20	27.52 1.28	28 (18 months)	13 (14 months)
Denmark	1979:3-2016:6 (387)	$0.76^{(**)}$ -0.11	0.40 0.17	90.14 1.82	24 (16 months)	10 (39 months)
Japan	1969:3-2016:6 (507)	$0.86^{(**)}$ -0.16	0.47 0.32	82.10 2.11	28 (18 months)	18 (28 months)
Sweden	1980:2-2016:6 (376)	$0.84^{(**)}$ -0.05	0.28 0.34	62.51 1.03	23 (16 months)	14 (27 months)
UK	1978:7-2016:6 (393)	$0.48^{(**)}$ -0.11	0.09 0.06	70.79 1.92	15 (26 months)	10 (39 months)

Notes: $\rho(.\,,.)$ refers to the correlation between two variables; ADF(.) refers to the p-value of the ADF test applied to the variable (with a constant in the underlying ADF regression and extent of augmentation chosen by AIC with max lag =12). In tests, superscripts [*] and [**] indicate significance at the 5% and 1% level respectively. V(.) refers to the variance of the variable; 'No. breaks' refers to number of occasions in which average sample length drops below 30 months (see text for details); 'No. phases' refers to phases defined by the occurrence of peaks/troughs (see text for details).

Table 2: The Relationship between the Meta-NNT and Alternative Models

	(1) Meta-NNTp	(2) Meta-MRB	(3) RMA - dynamic	(4) RMA - static
Canada Meta-NNT	$\rho = 0.936^{(**)}$	$\rho = 0.713^{(**)}$	$\rho = -0.299$	$\rho = 0.487^{(**)}$
	$KS = 0.298^{(**,**)}$	$KS = 0.117^{(*,-)}$	$KS = -0.024$	$KS = 0.419^{(**,-)}$
Denmark Meta-NNT	$\rho = 0.828^{(**)}$	$\rho = 0.483^{(**)}$	$\rho = -0.391$	$\rho = -0.213$
	$KS = 0.835^{(**,**)}$	$KS = 0.262^{(*,-)}$	$KS = -0.343$	$KS = -0.128$
Japan Meta-NNT	$\rho = 0.883^{(**)}$	$\rho = 0.341^{(**)}$	$\rho = -0.121$	$\rho = 0.071$
	$KS = 0.538^{(**,**)}$	$KS = -0.024$	$KS = -0.150$	$KS = -0.010$
Sweden Meta-NNT	$\rho = 0.557^{(**)}$	$\rho = 0.122^{(**)}$	$\rho = 0.134^{(**)}$	$\rho = 0.185^{(**)}$
	$KS = 0.626^{(**,**)}$	$KS = 0.090$	$KS = -0.060$	$KS = -0.002$
UK Meta-NNT	$\rho = 0.891^{(**)}$	$\rho = 0.452^{(**)}$	$\rho = -0.202$	$\rho = 0.100$
	$KS = 0.763^{(**,**)}$	$KS = 0.343^{(**,-)}$	$KS = 0.043$	$KS = 0.064$

Notes: Alternative models are 'meta' or 'rolling model average' (RMA) and are based on non-nested tests (NNT) or most-recent-best (MRB) performance. See text for details.

ρ refers to the correlation between the probability of PPP and Monetary models according to the meta-NNT model and the same probability according to the alternative models. Superscripts $^{( )}$ indicate significance in a one-tailed test of zero correlation.

KS refers to the Kuipers Score measuring the alignment of phases according to the meta-NNT model and the alternative models. Superscripts $^{(a,b)}$ indicate significance in a test of zero association using the (a) static and (b) dynamic PT tests respectively. In all cases, '*' and '**' show significance at the 5% and 1% levels respectively. See text for details.

References

Aiolfi, M., C. Capistrán, and A.Timmermann (2011), Forecast combinations, in M.P. Clements and D.F. Hendry (eds.) *The Oxford Handbook of Economic Forecasting*

Byrne, J. P., D. Korobilis, and P. J. Ribeiro (2016), Exchange rate predictability in a changing world, *Journal of International Money and Finance*, 62, 1-24.

Cheung, Y-W.and M.D. Chinn. (2001), Currency traders and exchange rate dynamics: a survey of the US market, *Journal of International Money and Finance*, 20, 439–471.

Diebold, F. and P. Pauly (2007), Structural change and the combination of forecasts, *Journal of Forecasting*, 6, 21-40.

Engel, C., (1994), Can the Markov switching model forecast exchange rates?, *Journal of International Economics*, 36, 1, 151-65.

Engel, C., N.C. Mark and K.D. West, (2008). Exchange rate models are not as bad as you think,. *NBER Macroeconomics Annual*, 22, 381-441.

Garratt, A. and K. Lee (2010). Investing under model uncertainty: decision based evaluation of exchange rate forecasts in the US, UK and Japan. *Journal of International Money and Finance*, 29, 403-422

Hansen, P.R., A. Lunde and J.M. Nason (2011), The model confidence set, *Econometrica*, 79, 2, 453–497.

Lee, K, J. Morley and K. Shields. (2015), The meta Taylor rule, *Journal of Money, Credit and Banking*, 47, 1, 73-98.

Lee, K., N. Olekalns and K. Shields (2013), Meta Taylor rules for the UK and Australia; Accommodating Regime Uncertainty in Monetary Policy Analysis using Model Averaging Methods, *Manchester School*, 81, S3, 28-53.

[22]

Mark, N.C. (1995), "Exchange rates and fundamentals: Evidence on long-horizon predictability", *American Economic Review*, 85, 201-218.

Meese, R. A., and K. Rogoff (1983), Empirical exchange rate models of the seventies:do they fit out-of sample? *Journal of International Economics*, 3-24.

Obstfeld, M. and K. Rogoff (2009), The six major puzzles in international macroeconomics: Is there a common cause?, in B.S. Bernanke and K. Rogoff (eds.) *NBER Macroeconomics Annual,* 15, 339-412.

Pesaran, M.H. (1974), On the general problem of model selection, *Review of Economic Studies*, 41, 2, 153-171

Pesaran, M.H., and A. Timmermann (2007). Selection of estimation window in the presence of breaks. *Journal of Econometrics* 137, 134-161.

Pesaran, M. H and A. Timmermann (2009), Testing dependence among serially correlated multicategory variables, *Journal of the American Statistical Association*, 104, 325-337.

Pesaran, M.H. and M. Weeks (2003),Non-nested hypothesis testing: An overview, in B. Baltagi (ed), *A Companion to Theoretical Econometrics*, Wiley.

Rossi, B. (2013), Exchange rate predictability, *Journal of Economic Literature*, 51, 4, 1063-1119.

Sarno, L. and G. Valente (2009) Exchange rates and fundamentals: Footloose or evolving relationship?, Journal of the European Economic Association, 7, 786-830.

Schinasi, G.J., and PA.V.B. Swamy. (1989), The out-of-sample forecasting performance of exchange rate models when coefficients are allowed to change, *Journal of International Money and Finance*, 8, 3, 375-90.

[23]

Steel, M., (2020), Model averaging and its use in economics, *Journal of Economic Literature*, 58, 3, 644-719.

Timmermann, A. (2006). Forecast combination. In: G.Elliot, et al. (eds). *Handbook of Economic Forecasting.* Amsterdam: Elsevier, 1, 135-196.

Wolff, C.C.P. (1987). Time-varying parameters and the out-of-sample forecasting performance of structural exchange rate models. *Journal of Business and Economic Statistics*, 5, 1, 87-97.

Wright, J.H. (2008), Bayesian model averaging and exchange rate forecasts, *Journal of Econometrics*, 146, 2, 329-41.