Fundamentals, Regimes and Exchange Rate Forecasts: Insights from a Meta Exchange Rate Model*

by

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

Abstract

A 'meta' model of the exchange rate combines a range of models distinguished by the drivers of the rate and by regime duration. Alternative model weights are proposed, including those obtained from a novel non-nested hypothesis-testing technique that accommodates periods of stability and slowly-evolving or abruptly-changing regimes involving multiple drivers. Focusing on density forecasts, the meta models perform well, demonstrating that all the sets of fundamentals considered can be useful for forecasting when the model is estimated over an appropriate time frame, but that the ability to exploit the changing relevance of different sets of fundamentals over time is important too.

Keywords: Exchange Rates, Model Averaging, Non-Nested Testing, Forecasting

JEL Classification: C51, F31, F47.

---

*[†]Central Bank of Cyprus, [††]University of Nottingham, UK, [†††]University of Melbourne, Australia. Version dated December 2021. The work has benefitted from helpful comments from the editor and two referees, and from conversations with Tony Garratt, Hyeyoen Kim, James Mitchell, James Morley, Ben Wong and seminar participants at the Bank of England and Reserve Bank of Australia. Corresponding author: Kevin Lee, School of Economics, University of Nottingham, Nottingham, NG7 2RD, UK. Email: kevin.lee@nottingham.ac.uk . Tel: +44 (0)115 8468386.

1 Introduction

The difficulties involved in forecasting exchange rates have been well-rehearsed. Meese and Rogoff's seminal (1983) paper observed that point predictions from a driftless random walk model were more accurate than those from more sophisticated models. Rossi (2013) provides a review of the literature describing the subsequent efforts to improve the point forecasts of exchange rate models, motivated by choice of different driving variables as predictors, applying novel time series and panel econometric techniques, using different forecast evaluation methods, and investigating the exchange rates between different currencies and over very long and/or specific historical sample periods. Rossi concludes that Meese and Rogoff's finding "does not seem to be entirely or convincingly overturned" noting, like Cheung, Chinn, and Pascual (2005) before her, that none of the predictors, models or tests she considers are consistently successful across countries or time periods in terms of their out-of-sample predictive performance.

This conclusion broadly carries over to more sophisticated forecasting exercises in which exchange rate models are used to produce density forecasts, to predict future events (such as appreciations or depreciations) or to make portfolio or other investment decisions. For example, in an analysis of density forecasts using nine exchange rates over the last thirty years, Abbate and Mercellino (2018) conclude that "there are difficulties in finding a model that performs uniformly better than the others across forecast periods, horizon or exchange rate". Similar conclusions are drawn in Gaglianone and Marins (2017) - "no single model accounts for the entire density properly for all forecast horizons" - and Cheung at al. (2017) who, based on five currencies and sample periods including various changes in the global environment, note that structural models outperform random walk using direction of change criteria but "the outperformance is not dramatically in excess of what would be expected on random chance .... as in our previous study, the best model and specification tend to be specific to the currency and out-of-sample forecasting period".[1]

[1]More positive results are found in Garratt and Lee (2010) and Cuaresma et al. (2018) where combinations of forecasts from exchange rate models appeared useful in some trading strategies involving the USD, euro, UKP and Yen.

Underlying the difficulties in forecasting is an apparent 'disconnect' between the exchange rate and its fundamental drivers. For example, Engel *et al.* (2008) describe a framework in which the fundamentals that explain exchange rates are the set of variables that account for the behaviour of the interest rate differential contemporaneously and into the future. Time variation in the relationship between the current and future interest rate differentials and their drivers will then appear as a disconnect between the exchange rate and its fundamentals, appearing as instability in the form of an exchange rate equation.[2] Similarly, Bacchetta and Van Wincoop (2004) describe a situation in which differences in the perceived importance of different macro variables across investors can cause different fundamentals to become 'scapegoats', each exerting disproportionate influence on the exchange rate at different times.

The complexity introduced through time variation in the relevance of different fundamentals is compounded by the observation that different fundamentals may be important for exchange rate determination over different horizons. For example, Bacchetta and Van Wincoop (2006) note how different exchange rate risk exposure and differential information on fundamentals across investors can result in financial market responses to news having a dominant short-run effect on exchange rate movements at the expense of long-run fundamentals, depending on the extent to which the news is differentiated or common across investors. Cheung and Chinn (2001) provide straightforward survey evidence on the different fundamentals relevant at different horizons: conventional macroeconomic pressures are thought to be important for exchange rate movements by 1% of US foreign exchange traders at the intraday horizon, 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). They also report time-variation in the relevance of the fundamentals as the relative importance of variables at different forecast horizons reported in their most recent survey was very different to the results of earlier surveys.

There have been a variety of approaches taken in the applied literature to deal with the structural instabilities in exchange rate models. Changes in the relevance of different sets

[2] See, for example, Molodtsova and Papell's (2013) discussion of structural instability in the Taylor rule determinants of interest rates, particularly during the Financial Crisis.

of fundamentals can be accommodated through Markov-switching models (as in Engel (1994) or Hauzenberger and Huber (2020), for example) or within a single-equation time-varying parameter model in which the underlying model encompasses the potentially relevant sets of fundamentals.[3] Alternatively, time variation in the influence of different sets of fundamentals can be accommodated through model averaging, where a number of structural models are estimated - recursively or with a rolling window - and then combined with time-varying weights. This can be approached as a Bayesian exercise - as in Wright (2008), Byrne et al. (2017) or Aastveit 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).[4] This latter paper explicitly addresses the extent to which changes in the relevance of different sets of fundamentals impacts on exchange rate forecasting performance. They report that, if an investor could select the appropriate model quarter-by-quarter, then economic fundamentals can explain future exchange rates 'with a remarkable degree of accuracy'. But they also note that it is not straightforward to select the appropriate model in each quarter on the basis of the information available at the time, and that conventional model selection criteria fail to detect the frequent shifts in the appropriate model which would be necessary to capture the evolving dynamics between exchange rates and their fundamentals.

In this paper, we adopt a model averaging approach to deal with the inherent structural instability in exchange rate models, but we pay particular attention to finding the time frame for which a given set of fundamentals might be relevant and to the model selection criterion that provides weights to the alternative models. Specifically, on the timing and duration of different exchange rate regimes, we follow the suggestion of Pesaran and

[3]In practice, time-varying-parameter models have tended to focus on a specific set of fundamentals; see Wolff (1987), Schinasi and Swamy (1989) or Rossi (2006), for example. The inclusion of alternative sets of fundamentals will encounter degreees of freedom problems without strong restrictions on the form of the time-variation of parameters.

[4]See also Kouwenberg et al. (2017) for a related approach employing a backward-elimination-of-regressors method for model selection.

Timmermann (2007) to apply model averaging techniques to alternative models estimated over different estimation windows.[5] The approach recognises that, when it is uncertain whether or when a break has occurred in a relationship, there is a trade-off between using short samples and long samples of data in a rolling estimation exercise. This is because longer samples improve the precision of parameter estimates and forecasts in the absence of breaks but are corrupted for longer in the presence of breaks. 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. We use the term 'meta modelling' to highlight our emphasis on regime uncertainty compared to more usual model averaging exercise.[6]

On the model selection criterion, we introduce a novel approach to constructing the time-varying weights in our meta model based on non-nested hypothesis-testing (NNT) methods.[7] Here, the weight on a particular set of fundamentals remains with that set as the sample extends until there is evidence to reject it in favour of a shorter model with the same fundamentals or with an alternative set of fundamentals. Non-nested testing methods are involved in the latter case because neither model is nested within the other.[8] The approach builds in a degree of stability in the characterisation of exchange rates over time by taking the profile of weights on the different models at any point in time as the maintained hypothesis and makes changes in the next period only if this profile is

[5]Clark and McCracken (2009) demonstrate that forecasting gains can arise from the use of recursive or rolling window schemes, or indeed a combination of the two, in this way. The benefits are illustrated in Pesaran et al.'s (2009) global VAR analysis, and in Clark and McCracken's (2010) study of US macroeconomic variables, for example.

[6]Lee et al. (2013, 2015) describe estimated "meta-Taylor rules" for the UK and US, obtained using similar methods to those of this paper, to characterise the different monetary policy regimes observed in those countries since the mid-seventies.

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

[8]The 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 peform poorly by some user-specified criterion. The MCS is the set of models which are not rejected as statistically inferior. In this paper, as we move through the sample, the weight from each model characterising exchange rate determination in one period is transferred to the models in its MCS in the next period.

rejected by the data. The profile of weights are more stable over time than those obtained by Bayesian updating or based on forecast performance over the recent past. But they can still accommodate the possibility of slowly-evolving or abruptly-changing regimes too, either based on the same set of fundamentals but taking into account a break or based on a separate set of fundamentals.

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. The meta model's flexibility in capturing the timing of regime change is crucial in capturing, and helping explain, this aspect of exchange rate determination. The section then provides evidence of the meta models' predictive power in a multi-step forecasting exercise focusing on density forecasts, associated event probability forecasts and the use of forecasts in a simple investment strategy. We find that it is important to take account of changing regimes, although a rolling window is as good as using the meta approach as far as forecasting is concerned. The ability to switch between fundamentals turns out to be very important though and the choice of weights through our NNT procedure is found to be very successful at long forecast horizons. Section 5 provides concluding remarks.

2 Exchange Rate Fundamentals and Structural Uncertainty

Our approach to modelling is based on the idea that different fundamentals, or a combination of fundamentals, could be more or less relevant for the determination of exchange rates in different circumstances. The meta modelling approach deals with this uncertainty through model averaging but emphasises the uncertainty over the duration of any particular characterisation of exchange rate determination by considering models over different sample sizes as well as with the different fundamentals. So, at time T, there might be $N \times J$ models that can potentially be used to characterise recent changes in the exchange rate, described by

$$M_{ijT}: \qquad s_t - s_{t-1} = \boldsymbol{\alpha}_{ijT}\mathbf{X}_{it} + \varepsilon_{ijT} \qquad i = 1, ..., N; \quad j = j_{\min}, ..., j_{\max}; \quad t = T - j, ..., T, \tag{2.1}$$

where $J = j_{\max} - j_{\min} + 1$. 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 very recent past or back to $j_{\max}$ periods in the past. The model involves $\mathbf{X}_{it}$ which is the i^{th} set of N alternative sets of explanatory variables driving the exchange rate and ε_{ijT} are associated random innovations.

In considering the potential drivers of the exchange rate, Engel et al. (2008) note that many familiar exchange rate models can be motivated by the Uncovered Interest Rate Parity (UIP) relationship which captures the equilibrium outcome of the arbitrage process between holding domestic and foreign bonds. Here, any differential in interest rates across countries must be offset by expected exchange rate changes to eliminate the scope for arbitrage. The UIP relationship in log-linear form is

$$s_t = s_{t+1}^e - (r_t - r_t^*), \tag{2.2}$$

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. Iterating forwards and taking expectations, the exchange rate depends on the entire future time path of interest rates

$$s_t = \sum_{h=0}^{\infty} \left[(r_{t+h}^e - r_{t+h}^{*e}) \right]. \tag{2.3}$$

Various models explaining exchange rate determination can be obtained from (2.2) and (2.3) depending on the choice of the fundamentals driving the interest rates and the time frame over which we want to consider their future path. For example, using (2.2), one might focus on the changes in the interest rate initiated by policy responses or financial market responses to news - ‘news pressures’ - during short periods of economic turbulence. Or, using (2.3), the choice might focus on the longer-term during periods of stability when the future path of interest rates might reflect broader macroeconomic conditions. These conditions define equilibrium interest rates and the path taken towards them and are mirrored in turn by the path taken by the exchange rate to its corresponding equilibrium. This idea is captured by Mark’s (1995) approach to modelling the exchange rate in which deviations of the nominal exchange rate from its equilibrium are persistent but gradually eliminated over time according to

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

where $f_t - s_t$ is the deviation of the time-t equilibrium exchange rate, f_t, from the actual rate. These alternative models motivate the choice of variables $\mathbf{X}_{it}$ in (2.1) and below we briefly outline four models frequently found in the literature to specify this choice; Rossi (2013) provides more detailed descriptions of the models and the empirical evidence relating to them.

Interest rate parity fundamentals (IRP). In the case where ‘news’ pressures dominate, one possibility for modelling exchange rates is to focus directly on recent interest rate movements. Recognising that the presence of transactions costs, risk premia and speculative effects provide for the possibility of permanent deviations from UIP and assuming that the expected future interest rate differential follows a simple $AR(p)$ specification, we can use (2.2) to write

$$s_t - s_{t-1} = \lambda + \sum_{i=0}^{p} \lambda_i (r_{t-i} - r^*_{t-i}) + \varepsilon_{IRP,t}, \tag{2.5}$$

where the λ’s are parameters and $\varepsilon_{IRP,t}$ consists of expectational error and any random variation in transactions costs, risk premia or speculative effects.

Taylor rule fundamentals (TR). The simple time series representation for the interest rate differential embedded within (2.5) can be replaced by a more forward-looking approach based on the determinants of the interest rate as expressed in the Taylor rule, specifying the central bank's policy rule as setting interest rates with reference to inflation, the output gap and, potentially, the real exchange rate. If a corresponding policy rule holds in the foreign country, then interest rate differentials will depend on inflation rates and gap measures at home and abroad and the real exchange rate so that, again using (2.2), we might write

$$s_t - s_{t-1} = \gamma + \gamma_r\ r_{t-1} - \gamma_{r*}\ r^*_{t-1} + \gamma_{\Delta p}\ \Delta p_t - \gamma_{\Delta p*}\ \Delta p^*_t + \gamma_x\ x_t - \gamma_{x*}\ x^*_t + \gamma_q\ q_t + \varepsilon_{TR,t}, \tag{2.6}$$

where Δp_t is the inflation rate, x_t is the output gap, $q_t = s_t + p^*_t - p_t$ is the log of the real exchange rate, and where the γ's are parameters $\varepsilon_{TR,t}$ again consists of expectational errors and the effects of risk premia, speculation and other financial market effects that are assumed to be random.[9]

Purchasing power parity fundamentals (PPP). The PPP hypothesis provides a candidate for the equilibrium level of the exchange rate in the longer horizon perspective captured by (2.4). In PPP, the equilibrium exchange rate is based on the idea that the price of a good would be the same in two countries when quoted in a common currency. Information disparities, transportation costs or the effects of tariff and non-tariff barriers are likely to create deviations from (absolute) PPP in the short run but, if these influences are constant over time, then the (weaker) concept of 'relative PPP' holds and we can write

$$s_t - s_{t-1} = \phi + \beta(p_{t-1} - p^*_{t-1} - s_{t-1}) + \varepsilon_{PPP,t} \tag{2.7}$$

where $\varepsilon_{PPP,t}$ again reflects stationary innovations (assuming no permanent changes in transportation costs, etc.).

Monetary model fundamentals (MON). The monetary model of the exchange rate provides an alternative characterisation of the equilibrium exchange rate based on the idea that the equilibrium nominal exchange rate is the relative price of two currencies and is determined by relative money supplies, relative income levels and the interest rate differential. Using the UIP assumption in (2.2) to eliminate the interest rate differential,

[9] See Engel and West (2006) and Molodtsova and Papell (2009) for exmaples of this approach.

an operational model for exchange rate determination is obtained assuming a simple autoregressive specification for each of the relative money supplies and relative outputs and inserting this into the dynamic specification at (2.4). If these relative magnitudes follow a random walk with drift, for example, we have

$$s_t - s_{t-1} = \rho_0 + \beta\left[\sum_{i=0}^{p}\rho_{mi}(m_{t-1-i} - m^*_{t-1-i}) + \sum_{i=0}^{p}\rho_{yi}(y_{t-1-i} - y^*_{t-1-i}) - s_{t-1}\right] + \varepsilon_{MM,t} \tag{2.8}$$

3 The Meta-NNT Model

The $N \times J$ different models explaining the exchange rate at time T in (2.1) can be combined in a weighted average to obtain a meta model similar to a typical model average but also taking account of changes in the relevant sample length at each point in time. The weights can change over time and can accommodate gradual shifts (as transportation costs gradually decline, say) or very abrupt shifts (in response to a sudden change in macroeconomic policy for example). As mentioned earlier, the advantages of model averaging in forecasting are now well documented, but the practice is much less widely employed in structural modelling even though the statistical arguments to support the approach are equally valid in inference and prediction.

The meta-modelling approach. The basis of the meta modelling approach is the Bayesian Model Averaging (BMA) formula (see Hoeting et al., 1999):

$$\Pr(\boldsymbol{\theta}_T \mid \mathbf{Z}_T) = \sum_{i=1}^{N}\sum_{j=j_{\min}}^{j_{\max}} \Pr(\boldsymbol{\theta}_T \mid \mathbf{Z}_T,\, M_{ijT}) \times \Pr(M_{ijT} \mid \mathbf{Z}_T) \tag{3.9}$$

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 $\boldsymbol{\theta}_T = (\beta_{ijT}\quad \forall\ i, j)$ represents the 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 (2.1). The BMA formula deals with the uncertainties accommodated within $\Pr(\boldsymbol{\theta}_T \mid \mathbf{Z}_T)$ by decomposing it into a weighted average of the conditional distributions, $\Pr(\boldsymbol{\theta}_T \mid M_{ijT}, \mathbf{Z}_T)$, using as weights the model probabilities $\Pr(M_{ijT} \mid \mathbf{Z}_T)$ to accommodate uncertainty over the nature of the fundamentals and regime uncertainty.

As discussed in Garratt et al. (2003), one can adopt a classical stance within this Bayesian framework if the conditional distribution $\Pr(\boldsymbol{\theta}_T \mid \mathbf{Z}_T, M_{ijT})$ is assumed to be described by

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

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 (3.9) is approximated by $N(\widehat{\boldsymbol{\theta}}_T, \widehat{\mathbf{V}}_T)$ and standard inference can be carried out for each model in turn.

The NNT weights. In practice, there is a wide choice of model weights $\Pr(M_{ijT} \mid \mathbf{Z}_T)$ that can be used to obtain the meta model.[10] In the 'meta-NNT' model, we propose a novel and pragmatic approach to deriving model weights where we allow them 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 methods which are relevant when one model cannot be obtained from the other by imposition of parameter restrictions or through a limiting process. The approach can be formalised by writing, for any T and for all $i = 1, ...n$ and $j = j_{\min}, ..., j_{\max} - 1$,

$$\Pr(M_{i,j,T-1} \mid \mathbf{Z}_{T-1}) \quad \rightarrow \quad \begin{cases} \Pr(M_{i,j+1,T} \mid \mathbf{Z}_T) & \begin{array}{r}\text{if the null } M_{i,j+1,T} \text{ is not rejected in favour of } M_{k,g,T} \\ \text{for } g = j_{\min},...,j \text{ and } k = 1,..,n,\end{array} \\ \Pr(M_{k,g,T} \mid \mathbf{Z}_T) & \begin{array}{r}\text{if the null } M_{i,j+1,T} \text{ is rejected in favour of } M_{k,g,T} \\ \text{for } g = j_{\min},...,j \text{ and } k = 1,..,n,\end{array} \end{cases} \tag{3.10}$$

so that the weight assigned at time $T-1$ to the model containing the i^{th} fundamentals 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

[10] See Steele (2020) for a comprehensive review.

than one alternative, the weight can be split equally among the alternative models. In transferring the weights, the tests should be conducted comparing the null to successively shorter samples so the weights can be shifted down sequentially where the evidence is that a model based on a shorter sample outperforms a model based on a longer one. This will happen when there is a change in economic behaviour, but this sequencing of the transfer of weights allows for some ambiguity on the timing of the change and the length of the new short sample. The updated weights provide a summary of the probabilistic relevance of the potential determinants of the exchange rate at each point in time.

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.[11] 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. 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. However, in most cases, the required modification renders the distribution of the statistic analytically intractable so that simulation methods are required.

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

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 the observed 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.

A strength of the meta-NNT approach. Models can be judged according to their *relevance* to the task at hand, their *consistency* with applicable theory and real-world institutions, and their *statistical adequacy*.[12] Even if a model is designed specifically to deliver exchange rate forecasts and achieves this satisfactorily (so it is relevant and adequate), consistency with theory and real world experience is important for confidence in the model and for the results to be integrated into our understanding of exchange rate determination. The meta-NNT model provides a useful vehicle for capturing the various theories put forward to explain the exchange rate 'disconnect' described in the introduction and will reflect the true data generating process (dgp) 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. Moreover, the meta modelling approach is consistent with our understanding of exchange rate determination even if the true dgp is distinct from all of the underlying structural models considered. Here, the weights just convey the real-time adequacy of the

[12]See Pesaran and Smith's (1995) discussion on the evaluation of macroeconometric models.

underlying structural models in characterising recent exchange rate movements; but the characterisation still contains useful insights on the underlying theory even in this case.

4 The Meta-NNT Models for Five Exchange Rates

We apply the meta-NNT modelling approach in forecasting exercises for five exchange rates over the last forty or fifty years; namely, the U.S. dollar exchange rates for the Canadian dollar, Danish krone, Japanese yen, Swedish krona and British pound. The data we use are measured monthly and the early part of the data, to 2010 : 6, are as provided in Rossi (2013), with the series' sample start dates varying across countries - as provided in column (1) of Table 1. These data are derived originally from Datastream but were collated by Rossi to provide a set of variables that are reasonably comparable across countries. We have then extended the series to 2019 : 12 based on the Rossi data definitions. 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 the HP filter to a forecast-augmented industrial output series.[13] 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.

Table 1 provides some summary statistics to provide an overview of the time series properties of the five exchange rates and their determinants. Column (2) reports the simple correlations between s_t and $p_t^{US} - p_t$, and between Δs_t and $r_t - r_t^{US}$ for each

[13] Specifically, at each period T, forecasts of future industrial production are obtained using a simple AR(2) model and trend output is identified as the time-T value of the HP filter applied to the extended series. See Garratt et al. (2008) for details.

currency. The former correlations range between -0.43 and -0.90, averaging -0.69 across the five countries, reflecting very similar patterns in the movements of the exchange rate and relative prices in every country considered over the forty- or fifty-year time frame of our samples of data. The large correlations reflect the fact that, across all five countries, the exchange rate is high (relative to its sample mean) at times when domestic prices are low compared to US prices (again relative to the sample mean), and the exchange rate is low when the domestic to US price level is high. Figure 1(a) illustrates this pattern in the case of Canada.[14] This supports the view embedded in the PPP and monetary models that price pressures are important for exchange rate determination over relatively long horizons. However, the relationships are not one-for-one and divergencies in the movements between the two series persist: simple ADF tests applied to the entire sample of data show, for all five countries, 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). Of course ADF tests have relatively low power but, as shown in column (3) of Table 1, the unit root tests statistics are not close to rejection (apart from the UK). In short, price pressures appear to impact on the exchange rate but, given the periodic and permanent shifts in the series, exchange rate determination will not be fully captured by a stable PPP or monetary model.

Column (4) of Table 1 shows that, despite the long run relationships that exist between the exchange rate and relative prices, the relationship over the short run is much less clear with exchange rate movements being much more volatile than those of relative prices. On the other hand, the ratio of the variance of Δs_t to the variance of $r_t - r_t^{US}$ is in the range 0.76 - 1.70 across the five countries so that the volatility of exchange rates is very much in line with the volatility of the interest rate differentials.[15] These patterns are again illustrated in the case of Canada in Figure 1(b), showing exchange rate and interest rate volatility to be of similar order of magnitude and much larger than relative price volatility for most of the sample. However, the diagram also demonstrates the importance of regime

[14] For brevity, we illustrate our results with reference to Canada here and subsequently; comparable figures for the other countries are provided in an Annex.

[15] The correlations of column (2) show that there is no discernable relationship between the levels of Δs_t and $r_t - r_t^*$ over the sample as a whole.

change as exchange rate volatility rises considerably at the end of the sample, after the Global Financial Crisis, while interest rate volatility falls to reflect the zero lower bound constraints faced by countries' monetary policy makers.

An intuitively reasonable account that is consistent with the statistics of Table 1 is that there are equilibrating macroeconomic pressures to move exchange rates towards establishing PPP over long horizons but that there are also factors (some periodic, some more continuous) that change the relationship between exchange rates and relative prices permanently. The monetary model of the exchange rate might capture these permanent shifts to the extent that the factors are reflected in productivity and money growth difference between countries but it is likely that these factors will cause structural breaks in estimated relations. At the same time, overlaying these forces are the jumps and volatile movements in the exchange rates arising in response to news from global markets, the nature of which determines whether the pressures are best represented by a simple IRP relationship or a more forward-looking Taylor Rule. The relative strengths of these various pressures varies over time, with macroeconomic pressures likely to show most clearly at times of stability and the 'news' pressures likely to dominate at times of economic turbulence. No single linear equation is likely to be able to capture all these influences. But the meta-NNT model allows them all to have an effect with individual models having non-zero weight while their influence remains apparent in the data.

4.1 Estimated Models of Exchange Rate Determination

Our modelling work begins by estimating five models for each country using data running over the first three years of their respective samples; e.g. the 36 observations between $1963:3-1966:2$ for Canada. The five models include each of our four fundamental models plus a random walk with drift. We then considered twelve alternative sample sizes, dropping one month from the beginning of the sample each time, until each model was estimated using only the 24 most recent observations; e.g. $1964:3-1966:2$ for Canada. This provides estimates of $M_{i,j,T}$ for $i = PPP, MM, IRP, TR, RW$, for $j = 24, ..., 35$ and $T = 1966:2$ then and, in this first iteration of the modelling, we assign equal weights to all $5 \times 12 = 60$ models obtained for each country.[16] The data window was then extended to cover one extra observation, to $1966:3$ for Canada, and, in this second iteration, $5 \times 13 = 65$ models were estimated for each country. The weights assigned to each model were determined in this case following the procedure in (3.10). Specifically, each model was held in turn as the null and compared to all of the models of shorter length in turn through a set of non-nested tests. The weights were transferred to the alternative if the null was rejected but remained with the null model otherwise.[17] This iterated procedure then continued for every T up to $2019:12$, extending the range of sample lengths at each iteration. The estimated models and model weights obtained in this way provided the estimated 'meta-NNT model' for each country.

Tables 2 and 3 provide statistics to give a sense of the outcome of this modelling strategy set alongside the outcome of two alternative modelling strategies: a 'meta-Most Recent Best (MRB)' strategy and a 'Fixed Window (FW)-NNT' strategy. The meta-MRB strategy continues to roll through the data considering models distinguished by their choice of fundamental and their choice of sample length. Here, however, weights are updated at each point according to the models' abilities to explain the final observation in the current sample. Specifically, time-t weights are updated 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. This approach, explained in detail in Lee et al. (2015), is similar to the approach often adopted in the forecast-combination literature where weights are based on the historical forecasting performance of the alternative models, although the meta-MRB approach builds in a degree of stability in model weights by updating rather than using the probabilities themselves. As a second point of comparison, the FW-NNT strategy updates model weights according to the NNT procedure, as in the meta-NNT model, but

[16]The four behavioural models take the form in (2.5)-(2.8) with $p = 2$. For TR, to conserve degreees of freedom for the small sample regressions, we impose equal and opposite signs on the domestic and foreign variables and apply Taylor's original assumption that $\frac{\gamma_{\Delta p}}{\gamma_x} = \frac{1.5}{0.5}$.

[17]Our modelling also involves a specification search rule in which the individual models are checked for dynamic stability. In the case that a model becomes dynamically unstable (so that its forecasts are unreliable), its associated weight is reallocated across all other models with non-zero weights.

abstracts from the choice of sample length by using a fixed rolling window of five years.

Table 2 reports the R^2 statistics associated with the estimated models, calculated as the ratio of the 'explained' to 'total' sum of squares, where the 'explained' sum of squares is based on the weighted average of the fitted values from each of the models estimated at each point in the sample. The statistics demonstrate the importance of taking account of variable-length regimes with the R^2's of the meta models averaging 0.091, around twice as high as the Fixed Window models across all countries. As illustrated in Figure 2 for the meta-NNT model for Canada, there are typically some prolonged periods of stability - with the weighted sample length , $D_T = \sum_i \sum_j (j \times w_{i,j,T})$, rising above four years on a couple of occasions - but there are also frequent drops in the sample length reflecting regular breaks among the models. Defining a 'substantial break' to occur when the weighted average of the sample length drops into the region $[24, 29]$ months, Canada experiences a substantial break every 12 months on average according to the meta-NNT model, and this sort of experience is mirrored in all countries (with an average period between substantial breaks of 16 months across the five countries). There is a corresponding consistency across countries in the length of time between substantial breaks according to the meta-MRB model, although the average period here is just 7 months, providing even stronger evidence in favour of accommodating frequent changes of regimes of varying duration in modelling exchange rates.

Table 2 also notes the contributions of the different fundamental models to the overall model fit. Again, there is a degree of consistency across countries (and indeed across all three modelling strategies) with the greatest contribution to the overall fit coming from the PPP and the Taylor Rule models respectively and with the IRP and Monetary models typically making a smaller contribution. The results from the two meta models are again broadly similar in this respect (with the contribution from the PPP model averaging 30% across the countries' meta-NNT models and 33% in the meta-MRB, for example) and are distinct from with the Fixed Window model (with an average PPP contribution at 40% for example). Figure 2 shows, for each country, the time variation in the weights allocated to models with different fundamentals, tracing out the sum of the weights allocated to the PPP and Monetary models. The typical pattern observed across countries shows considerable variation over time, with the weights ranging between $0.2 - 0.8$, sometimes rising or falling slowly over time, sometimes more abruptly,[18] and reflecting phases during which exchange rate determination is dominated by the macroeconomic pressures of PPP+MON rather than the news and financial market pressures captured by IRP+TR, and phases where the dominance is reversed.

The statistics of Table 3 report the correlations between the weights on the models observed over time across the three modelling strategies and show that there is degree of agreement on the timing of the phases between the meta-NNT and Meta-MRB models but much less agreement between the Meta-NNT and FW-NNT models. This again supports the view that a rolling fixed window model misses important features of the data and that it is important to accommodate frequent changes of regimes of varying duration in modelling exchange rates. Calculation of the corresponding correlations *across countries* shows little commonality in the timing of the various phases of dominance. Of course, all of these exchange rates are with respect to USD so one might expect some features of the exchange rate movements to show in all the currencies. But the lack of alignment demonstrates that the timing and nature of the phases are complex and country-specific, reflecting the particular circumstances of each country.[19]

In summary:

1. the fundamentals that determine exchange rates differ across countries and across time;

2. exchange rates are influenced both by the macro pressures underlying the Purchasing

[18] The volatility in the weights on the fundamentals is noticably less than that of the weighted sample lengths, indicating that many of the observed breaks relate to the same fundamental model but estimated over shorter samples.

[19] The meta analysis of the Taylor rules in UK and US provided in Lee et al. (2013, 2015) illustrates the detailed and country-specific timing of changes in monetary policy in the two countries. If the exchange rate is influenced by IRP or TR models, this will translate into country-specific regime change in the exchange rate equations too. Similar comments apply to the timing of changes in the risks associated with countries' assets - as discussed, for example, in Ismailov and Rossi (2018) - and countries' trade arrangements and transportation costs which influence the PPP and Monetary drivers of its exchange rate.

Power Parity and Monetary models and by the news and financial market pressures underlying the Taylor Rule and IRP models, and can be characterised by phases during which the two sets of pressures are more or less important;

3. the timing of changes in regimes is different across countries; and

4. regime duration varies over time and across countries, although there are similarities in the frequency of substantial breaks across countries (occuring every16 months on average)

4.2 The Models' Forecasting Performance

Tables 4-6 provide details of the forecasting performance of the models discussed above. The focus of attention is on forecasting exercises in which the models are used to produce density forecasts, to predict future events (such as appreciations or depreciations) and to make portfolio decisions. This requires stochastic simulation methods in which the estimated models M_{ijT} obtained at T are used to generate a range of potential future values of the exchange rate for $T+h$, for $h = 1, 3, 6, 9, 12, 18, 24, 36$, and the forecast densities, event probabilities and portfolio returns are derived from this generated range. Specifically, having obtained the estimates of $\boldsymbol{\alpha}_{ijT}$ in (2.1) for each of the M_{ijT} at T, we simulate the future values of $s_{T+h}^{(r)}$ using the model recursively, where the superscript 'r' refers to the r^{th} replication of the simulation. In each simulation, we take lagged values of the variables and the future values of the fundamentals as given,[20] and we incorporate stochastic variation by drawing errors $\varepsilon_{ijT+h}^{(r)}$ from a Normal distribution with variance equal to the estimated equation's error variance. If there is only one model involved, repeating this for $r = 1, .., R_{ijT}$ provides the forecast density at each horizon for the model; calculating the proportion of replications in which an event occurs (e.g. "the exchange rate appreciates over the 12 month forecast horizon") provides the event probability forecast;

[20]Rossi (2013) describes this as the 'contemporaneous, realised fundamental' model specification and contrasts it with the 'contemporaneous, forecasted fundamental' and 'lagged fundamental' models. Summarising the empirical evidence, she concludes that it is the choice of fundamentals that matters in determining the strength of the predictability of a model rather than the specification.

and maximising the expected gain from applying an investment strategy to the simulated series provides a decision-based forecast outcome for the model. If more than one model is involved, the various forecasts are obtained by pooling the simulated futures from the different models choosing the number of replications of each model in proportion to the weights w_{ijT}.[21]

Table 4 compares the density forecast performance of the various models, using the first 60 months of each countries' sample as a 'training' period and the subsequent months as the 'out-of-sample' forecast evaluation period. The table reports the average log score obtained over the evaluation period based on the meta-NNT model expressed as a ratio to the average log-score based on alternative models.[22] Numbers in excess of unity indicate that the meta-NNT provides a better density forecast therefore. The corresponding Diebold-Mariano tests, calculated using these log scores, are also reported to test the null of equal forecast performance.

The first panel of the Table compares the meta-NNT model with a simple random walk model.[23] The results show that the uncertainty surrounding the random walk forecasts are relatively poorly calibrated and that the meta-NNT's forecasts are significantly and comprehensively superior to the random walk model in all countries and at all forecast horizons.[24] The meta-NNT's density forecast performance is also superior to that of the fixed window random-walk-with-drift model in the second panel. The extent of the over-performance is much reduced compared to the straight RW model but remains clear, especially at longer horizons (e.g. 70% of the reported statistics are significant at $h \geq 9$). In contrast, the third panel reports that the meta-NNT and FW-NNT models have

[21]The forecasts obtained in this way take into account stochastic and model uncertainty but not the parameter uncertainty that surrounds the estimated $\boldsymbol{\alpha}_{ijT}$. In principle, this could be accommodated though an additional simulation stage generating artificial 'histories' but it is not possible in practice given the heavy computational demands of the meta-NNT model estimation.

[22]If the time-T forecast density function of the h-period-ahead forecast is denoted $\widehat{f}_h(.)$, the 'log score' is the log of the value of the density evaluated at the outcome $-\log\left[\widehat{f}_h(s_{T+h})\right]$.

[23]The time-T measure of the variance used for simulating density forecasts from the RW model is $(s_T - s_{T-1})^2$.

[24]Neither model - or any of the other models considered - outperforms the other in terms of point forecasts, confirming the standard Meese and Rogoff finding.

comparable forecast performance, with neither significantly better or worse than the other. This indicates that the strength of the meta-NNT forecast performance in the second panel derives from its ability to switch between model fundamentals then rather than its ability to choose regimes of different lengths. Finally, in the fourth panel, the ratios show that the meta-MRB performs significantly better than meta-NNT at the shortest $h = 1$ forecast horizon, but that the meta-NNT forecast performance improves relative to meta-MRB at longer horizons and is significantly better in 75% of the reported results at $h \geq 18$.

Table 5 and 6 focus on particular aspects of the density forecasts capturing more 'economic' forecast evaluation criteria. Table 5 reports on the models' abilities to forecast the direction of change (i.e. overall appreciation or depreciation) over the various forecast horizons; and Table 6 reports on the outcomes of an investment strategy which, each month, buys (or sells) one unit of the foreign currency when the model predicts the currency will appreciate (or depreciate) over the forecast horizon. Here, the evaluation of the models' forecasting performance depends not only on the accuracy and degree of certainty surrounding the forecasts but also on the relevance of the $\Delta s_T = 0$ threshold in Table 5 and, additionally, on the size of the pay-off ($= \Delta s_{T+h}$) in Table 6.

Table 5 shows the ratio of correctly-forecasted direction of change from the pairs of models described in Table 4, with a value in excess of unity again indicating that the meta-NNT provides a better forecast than the comparator models. A large majority of the reported results are greater than unity in the comparisons with the RW and FW-RWD, confirming the superiority of the meta-NNT model in these cases, although the size and significance of the ratios are typically less than in Table 4. As in Table 4, there is relatively little to choose between the meta-NNT and FW-NNT models according to this criterion, although there is some evidence of dominance of the meta-NNT model at longer horizons. And the comparisons between the meta-NNT and meta-MRB models also reflects the findings in Table 4 that the meta-MRB is found to have superior forecasting power at shorter horizons, but the meta-NNT model performs better at the longer horizons. Similar conclusions are drawn from Table 6, which reports the net gains from the investments recommended by the model pairs (so that a positive statistic indicates that the meta-NNT model provides the most profitable forecasts).

Taken together, the results provide some useful insights for exchange rate forecasting (and exchange rate modelling more generally):

1. Structural models involving model averaging deliver exchange rate density forecasts which are significantly better than those of simple random-walk or random-walk-with-drift models, both based on statistical log-score criteria and the specific economic criteria considered. This superiority is based on the appropriateness of the uncertainty/confidence with which they make their forecasts rather than the accuracy of the point forecasts.

2. As far as forecasting is concerned, the gains from model averaging arise from the models' ability to switch between fundamentals rather than from their ability to accommodate the establishment of new regimes of varying length. Having said that, in terms of the models' in-sample fit, the meta models perform considerably better than the corresponding rolling fixed window models so that they are more consistent with real world experience and can be better integrated into our understanding of exchange rate determination.

3. The NNT procedure for allocating model weights delivers models that perform relatively well at long forecast horizons while the MRB procedure performs better at very short horizons. While it might be a statement of the obvious, different models will be relevant for different tasks and it is entirely reasonable that the relatively rapid regime changes associated with MRB are found to be useful for short term forecasting while the longer-lived phases associated with NNT produce better longer-term forecasts. Having said that, there is some commonality between the regimes picked out by the meta-NNT and meta-MRB models in estimation and this reflects a useful consistency with real world experience from both models.

4. The use of a variety of forecast evaluation criteria helps judge models' forecasting performances from different perspectives (sometimes helping to hone in on a particular model, sometimes exposing that no model can provide useful guidance to decision-makers). The 'economic' evaluation criteria we consider here are heavily

influenced by the location of the forecast densities, and the improvements in density forecasting - and potentially complicated multi-modal shapes of the densities - obtained through model averaging play only a minor role in forecast performance in this context.

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 the uncertainty over the fundamentals driving the exchange rate, doing this in a way that can account for periods of stability, periods in which the emergence of new regimes takes place gradually over time and episodes of abrupt changes in regime. The novel NNT approach to allocating weights introduced in the paper matches well the way in which an observer of the exchange rate market might characterise how the market behaves. This consistency with real world experience is a strength of the modelling when compared to the more data-driven meta-MRB (or related Bayesian) models. The results of the paper show that, for the five currencies considered, the meta 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, for the meta models, there is a similarity in the frequency of structural breaks and choice of drivers that gives some confidence in the validity of the estimated models. The meta models are able to exploit relevant economic information and provide a systematic improvement in density forecasting performance over random walk models, whether judged by purely statistical criteria or in a more economically-meaningful context, with the precise choice of the approach for choosing weights (NNT or MRB) depending on the forecast horizon.

References

Aastveit, K.A., F. Ravazzoloz and H.K van Dijk (2018), Time-varying uncertainty and exchange rate predictability, *Working Paper*, Centre for Applied Macro and Petroleum Economics, BI Norwegian Business School

Abbate, A. andM. Marcellino (2018),.Point, Interval and Density Forecasts of Exchange Rates with Time Varying Parameter Models, *Journal of the Royal Statistical Society Series A* 181, 155-179.

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*

Bacchetta, P. and E. Van Wincoop (2004), Scapegoat Model of Exchange-Rate Fluctuations, *American Economic Review*, vol. 94, pp. 114-118.

Bacchetta, P. and E. Van Wincoop (2006), Can Information Heterogeneity Explain the Exchange Rate Determination Puzzle?, *American Economic Review*, 96, 552-576

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.

Cheung, Y-W, M.D. Chinn and A.G. Pascual (2005), Empirical exchange rate models of the nineties: are any fit to survive? *Journal of International Money and Finance*, 24, 1150-1175.

Cheung, Y-W., M.D. Chinn, A. Pascual and Y. Zhang, (2017), Exchange Rate Prediction Redux: New Models, New Data, New Currencies, *European Central Bank Working Paper*, no. 2018.

Clark. T. and M.W. McCracken (2009), Improving forecast accuracy by combining recursive and rolling forecasts. *International Economic Review*, 50, 363-395.

Clark. T. and M.W. McCracken (2010), Averaging forecasts from VARs with uncertain instabilities. *Journal of Applied Econometrics*, 25, 5-29.

Cuaresma, J. C., I. Fortin and J. Hiouskova, (2018), Exchange Rate Forecasting and the Performance of Currency Portfolios, *Journal of Forecasting*, 37, 519-540.

Engel, C. (1994), Can the Markov Switching Model Forecast Exchange Rates?, *Journal of International Economics,* 36, 1, 151-165.

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.

Engel, C. and K.D. West (2008), Taylor Rules and the Deutschmark-Dollar real Exchange Rate, *Journal of Money, Credit and Banking*, 38, 1175-1194.

Gaggglianone, W. and J.T.M. Marins (2017), Evaluation of Exchange Rate Point and Density Forecasts: An Application to Brazil, *International Journal of Forecasting*, 33, 707-728.

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

Garratt, A., K. Lee, E. Mise and K. Shields (2008), Real time representations of the output gap, *Review of Economics and Statistics*, 2008, 90, 4, 792-804.

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

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

Harzenberger, N. and F. Huber, (2020), Model Instability in Predictive Exchange Rate Regressions, *Journal of Forecasting*, 39, 168-186.

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

Ismailov, A. and B. Rossi, (2018), Uncertainty and Deviations from Uncovered Interest Parity, *Journal of International Money and Finance*, 88, 242-259.

Kouwenberg, R., A. Markiewicz, R. Verhoeks and R. Zwinkels (2017), Model Uncertainty and Exchange Rate Forecasting, *Journal of Financial and Quantitative Analysis*, 52, 341-363.

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.

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.

Molodtsova, T. and D.H. Papell (2009), Exchange Rate Predictability with Taylor Rule Fundamentals, *Journal of International Economics*, 77, 167-180.

Molodtsova, T. and D.H. Papell (2013), Taylor Rule Exchange Rate Forecasting during the Financial Crisis, *NBER International Seminar on Macroeconomics*, University of Chicago Press, vol. 9(1), 55 - 97.

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

Pesaran, M.H., T. Schuermann and L.V. Smith (2009), Forecasting economic and financial variables with global VARs. *International Journal of Forecasting*, 25, 642-675.

Pesaran, M.H. and R. Smith (1985), Evaluation of Macroeconometric Models, *Economic Modelling*, 2, 125-134.

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

Rossi. B. (2006), Are Exchange Rates Really Random Walks? Some Evidence Robust to Parameter Instability, *Macroeconomic Dynamics*, 10, 1, 20-38.

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, 4, 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, 375-390.

Steel, M., (2020), Model Averaging and its Use in Economics, *Journal of Economic Literature*, 58, 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, 329-341.

Table 1: Summary Statistics

	(1) Sample (n)	(2) Corr(s, p^{US}-p) / Corr(Δs, r-r^{US})	(3) ADF(s) / ADF(s+p^{US}-p)	(4) V(Δs)/V(Δ(p^{US}-p)) / V(Δs)/V(r-r^{US})
Canada	1964m4 – 2019m12	-0.75***	0.36	5.50
	(669)	0.01	0.66	0.76
Denmark	1975m3 – 2019m12	-0.61***	0.22	11.07
	(538)	-0.10	0.18	1.26
Japan	1965m3 – 2019m12	-0.90***	0.51	11.02
	(658)	-0.08	0.71	1.48
Sweden	1976m2 – 2019m12	-0.75***	0.28	10.63
	(527)	-0.05	0.42	0.91
UK	1974m6 – 2019m12	-0.43***	0.06*	8.94
	(547)	-0.08	0.36	1.70

Notes: Column (2) gives the correlation between the two variables in parentheses; Column (3) gives the p-value of the ADF test applied to the variable with a constant in the underlying ADF regression and the extent of augmentation chosen by the Akaike Information Criterion with a maximum lag of 12; V(.) in Column (4) refers to the variance of the variable. . *, **, and *** indicate significance at 10%, 5% and at the 1% level, respectively.

Table 2: Properties of Models

	R-squared	Percentage of R-squared statistic due to:					Av Duration between Breaks
		PPP	IRP	TR	MON	RW&Drift	
CANADA							
Meta-NNT	0.093	35%	19%	22%	14%	10%	12 months
FW-NNT	0.057	48%	12%	20%	10%	11%	-
Meta-MRB	0.101	32%	19%	15%	27%	7%	6 months
DENMARK							
Meta-NNT	0.078	32%	18%	24%	13%	12%	14 months
FW-NNT	0.046	37%	21%	16%	15%	10%	-
Meta-MRB	0.088	38%	20%	17%	16%	9%	7 months
JAPAN							
Meta-NNT	0.085	25%	13%	24%	20%	18%	29 months
FW-NNT	0.049	39%	12%	18%	20%	11%	-
Meta-MRB	0.088	32%	9%	24%	22%	12%	8 months
SWEDEN							
Meta-NNT	0.093	32%	13%	23%	20%	13%	11 months
FW-NNT	0.052	38%	17%	15%	14%	15%	-
Meta-MRB	0.095	33%	14%	19%	23%	11%	7 months
UK							
Meta-NNT	0.098	27%	16%	27%	21%	8%	16 months
FW-NNT	0.055	40%	17%	26%	10%	7%	-
Meta-MRB	0.101	33%	18%	24%	19%	6%	7 months

Notes: 'Duration between breaks' refers here to the period between occasions in which the weighted average sample size drops below 29 months.

Table 3: Correlations of Structural Models across Models

	Corr(Meta-NNT, FW-NNT)		Corr(Meta-NNT, Meta-MRB)	
	PPP+MON	IRP+TR	PPP+MON	IRP+TR
CANADA	0.118***	0.156***	0.332***	0.422***
DENMARK	0.542***	0.440***	0.429***	0.436***
JAPAN	0.221***	0.186***	0.312***	0.281***
SWEDEN	0.014	0.188***	0.108**	0.219***
UK	0.138***	0.145***	0.235***	0.277***

Notes: This table presents the correlations of the sum of the probability weights attached to the PPP and MON models together, and the IRP and TR models together, between the Meta-NNT and the FW-NNT models, and the Meta-NNT and the Meta-MRB models, respectively for each country. *, **, and *** indicate significance at 10%, 5% and at the 1% level, respectively.

Table 4: Forecasting Performance: Ratio of Average Log Scores

Meta-NNT vs RW	**Forecast Horizon**							
	1	3	6	9	12	18	24	36
CANADA	1.310***	1.281***	1.282***	1.306***	1.324***	1.318***	1.265***	1.164***
DENMARK	1.055***	1.033***	1.025***	1.040***	1.050***	1.042***	1.036***	1.036***
JAPAN	1.309***	1.257***	1.239***	1.269***	1.280***	1.250***	1.214***	1.133***
SWEDEN	1.307***	1.273***	1.260***	1.264***	1.266***	1.267***	1.227***	1.190***
UK	1.305***	1.275***	1.283***	1.307***	1.325***	1.331***	1.305***	1.276***

Meta-NNT vs FW-RWD	Forecast Horizon							
	1	3	6	9	12	18	24	36
CANADA	1.007**	1.012**	1.011	1.020*	1.031**	1.051**	1.061**	1.107***
DENMARK	1.002	1.004	1.000	1.003	1.007	0.995	1.028	1.103***
JAPAN	1.002	1.005	1.008	1.035**	1.032*	1.045*	1.062**	1.075**
SWEDEN	1.001	0.999	0.996	0.999	1.000	1.008	1.015	1.052***
UK	1.001	1.010*	1.021**	1.033***	1.056***	1.085***	1.130***	1.148***
	1.007**	1.012**	1.011	1.020*	1.031**	1.051**	1.061**	1.107***

Meta-NNT vs FW-NNT	Forecast Horizon							
	1	3	6	9	12	18	24	36
CANADA	1.002	1.007	1.006	1.011	1.014	1.013	1.006	1.009
DENMARK	0.998	0.994	0.982	0.974	0.967	0.955	0.968	0.807
JAPAN	1.002	0.994	0.993	1.014	1.005	0.999	1.001	0.980
SWEDEN	0.998	0.996	0.992	0.991	0.984	0.996	0.992	1.023
UK	0.998	1.001	1.005	1.017	1.019	1.020	1.029*	1.011

Meta-NNT vs Meta-MRB	Forecast Horizon							
	1	3	6	9	12	18	24	36
CANADA	0.993**	1.004	1.015*	1.021**	1.016***	1.016	1.032**	1.050*
DENMARK	0.994**	0.995	0.993	0.993	1.000	0.998	1.015**	0.817***
JAPAN	0.992***	0.996	0.999	1.005	1.003	1.016	1.045**	1.074***
SWEDEN	0.996	0.999	1.007	1.011	1.018	1.035**	1.047**	1.043**
UK	0.994**	1.011**	1.029***	1.026**	1.030***	1.034**	1.032**	1.026

Notes: This table presents the ratio of the average log scores of the meta-NNT model relative to the FW-NNT, meta-MRB, FW-RWD and the RW models, for various respective forecast horizons. Ratios larger than one indicate that the meta-NNT model outperforms the respective models. The table also presents the results from the Diebold and Mariano (DM) test of equal forecasting performance where *, **, and *** indicate significance at 10%, 5% and at the 1% level, respectively.

Table 5: Forecasting Direction of Change in s_t: Ratio of Proportion of Correct Signs

Meta-NNT vs RW	**Forecast Horizon**							
	1	3	6	9	12	18	24	36
CANADA	1.060	1.033	1.074**	1.133***	1.105***	1.118***	1.160***	1.150***
DENMARK	1.020	1.084*	1.090**	1.142***	1.099**	1.069	1.066	1.143***
JAPAN	1.020	1.060	1.046	1.032	1.015	0.969	0.946	0.871***
SWEDEN	0.972	1.017	1.054	1.037	1.019	0.980	0.956	0.996
UK	1.050	1.053	1.104**	1.058	1.056	0.985	1.004	0.988
Meta-NNT vs FW-RWD	Forecast Horizon							
	1	3	6	9	12	18	24	36
CANADA	0.967	0.972	0.965	0.959	0.921	0.905**	1.019***	0.973
DENMARK	1.022	1.111**	1.078*	1.123***	1.099**	1.126***	1.166***	1.143***
JAPAN	0.988	1.074*	1.036	1.009	1.000	0.948	0.940	0.844***
SWEDEN	0.981	1.027	1.022	1.000	1.023	1.096**	1.081	1.047
UK	1.063	1.163***	1.252***	1.315***	1.382***	1.529***	1.448***	1.518***
Meta-NNT vs FW-NNT	Forecast Horizon							
	1	3	6	9	12	18	24	36
CANADA	1.029	1.030	1.056	1.107	1.103***	1.141***	1.285***	1.314***
DENMARK	0.996	0.990	0.986	1.007	0.963	0.949	0.948	1.104**
JAPAN	0.965	1.021	1.040	1.050	1.101*	1.165***	1.224***	1.148***
SWEDEN	0.941	0.993	0.986	0.964	0.989	1.020	1.076	1.458***
UK	0.990	1.044	1.057	1.105**	1.048	0.929	0.868***	0.940
Meta-NNT vs Meta-MRB	Forecast Horizon							
	1	3	6	9	12	18	24	36
CANADA	0.839***	0.915**	0.960	1.030	1.008	1.011	1.054	1.061*
DENMARK	0.843**	0.960	0.927**	0.959	0.912**	1.045	0.993	1.000
JAPAN	0.817***	0.923**	0.927**	0.889***	0.899***	0.901***	0.938	1.004
SWEDEN	0.771***	0.893***	0.986	0.978	0.992	1.008	1.008	1.134***
UK	0.831***	1.040	1.080*	1.084*	1.011	0.977	0.953	0.992

Notes: This table presents the ratios between the proportion of correctly predicted changes in the direction of s_t of the meta-NNT model relative to the FW-NNT, meta-MRB, FW-RWD and the RW models, for the various respective forecast horizons. Ratios larger than one indicate that the meta-NNT model outperforms the respective models. The table also presents the results from the Pesaran-Timmerman (PT) test of the ability of the meta-NNT to forecast the direction of change correctly relative to each of the respective models where *, **, and *** indicate significance at 10%, 5% and at the 1% level, respectively.

Table 6: Relative Net Gains of Alternative Investment Strategies

Meta-NNT vs RW	**Forecast Horizon**							
	1	3	6	9	12	18	24	36
CANADA	0.246	0.661	-0.488	0.233	-0.214	0.681	0.837	1.088**
DENMARK	0.864	1.440*	1.485*	0.947	0.685	0.374	0.580	0.511
JAPAN	-0.098	-0.081	-0.174	-0.007	-0.523	0.326	1.143*	-0.854
SWEDEN	-0.525	0.605	0.596	0.921	-0.511	-0.225	0.336	-1.453
UK	0.944	1.938**	1.884**	1.560	1.787*	0.513	2.028***	-0.142

Meta-NNT vs FW-RWD	Forecast Horizon							
	1	3	6	9	12	18	24	36
CANADA	-0.401	0.541	-0.471	0.016	-0.486	0.306	0.810	0.763*
DENMARK	1.200*	0.909	1.529*	1.291*	1.079	-0.269	0.759	0.075
JAPAN	0.510	0.290	0.162	0.153	0.176	-0.842	-1.245	-0.230
SWEDEN	0.192	1.055	0.417	1.148	-0.075	-0.182	-0.379	-1.257
UK	0.977	2.417**	2.896***	1.924*	2.064**	0.916	1.115	-0.305

Meta-NNT vs FW-NNT	Forecast Horizon							
	1	3	6	9	12	18	24	36
CANADA	0.223	0.722	0.146	0.333	-0.231	0.284	-0.088	0.222
DENMARK	0.302	-0.355	-0.249	-0.548	0.606	-0.644	0.564	1.385**
JAPAN	0.173	0.031	0.909	0.634	0.264	-0.737	-0.582	-0.453
SWEDEN	-0.033	0.977	0.011	-0.487	-0.954	-0.185	0.125	0.906
UK	0.207	1.227*	0.438	1.339	0.403	0.003	0.097	-0.337

Meta-NNT vs Meta-MRB	Forecast Horizon							
	1	3	6	9	12	18	24	36
CANADA	-1.308**	0.356	-0.313	-0.158	-0.808**	0.036	-0.065	0.053
DENMARK	-1.769**	-0.451	0.542	0.319	-0.178	0.042	-0.225	0.006
JAPAN	-2.048***	-1.037	-1.310*	-0.680*	-1.238**	0.393	0.975	0.034
SWEDEN	-2.556***	-0.291	0.585	-0.305	-0.421	-1.456	-0.120**	-0.915
UK	-1.802***	2.130**	0.427	0.473	0.381	-0.953*	-0.369	-0.359

Notes: This table presents the net gains of the investment strategy outlined in the text of the meta-NNT model relative to the FW-NNT, meta-MRB, FW-RWD and the RW models, for the various respective forecast horizons. The table also presents the results from the Diebold and Mariano (DM) test of equal forecasting performance where *, **, and *** indicate significance at 10%, 5% and at the 1% level, respectively.

Figure 1a: Dollar Exchange Rate, Price Differential and Interest Rate Differential for Canada

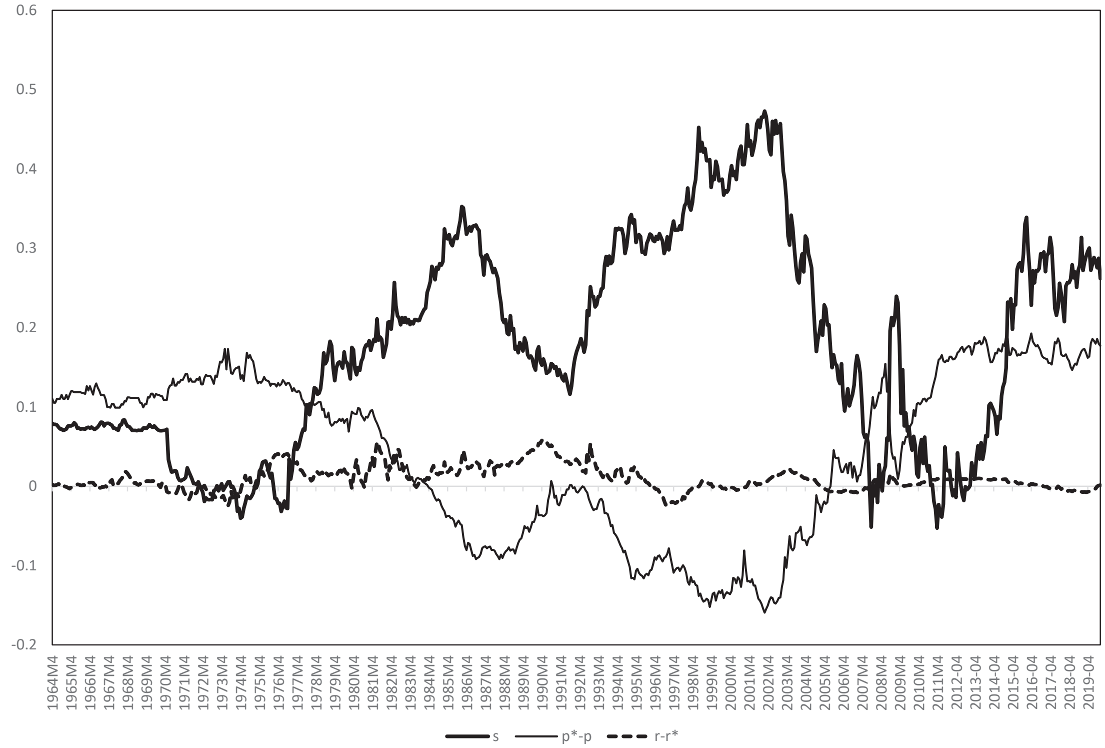


Figure 1b: Variation in Dollar Exchange Rate, Price Differential and Interest Rate for Canada

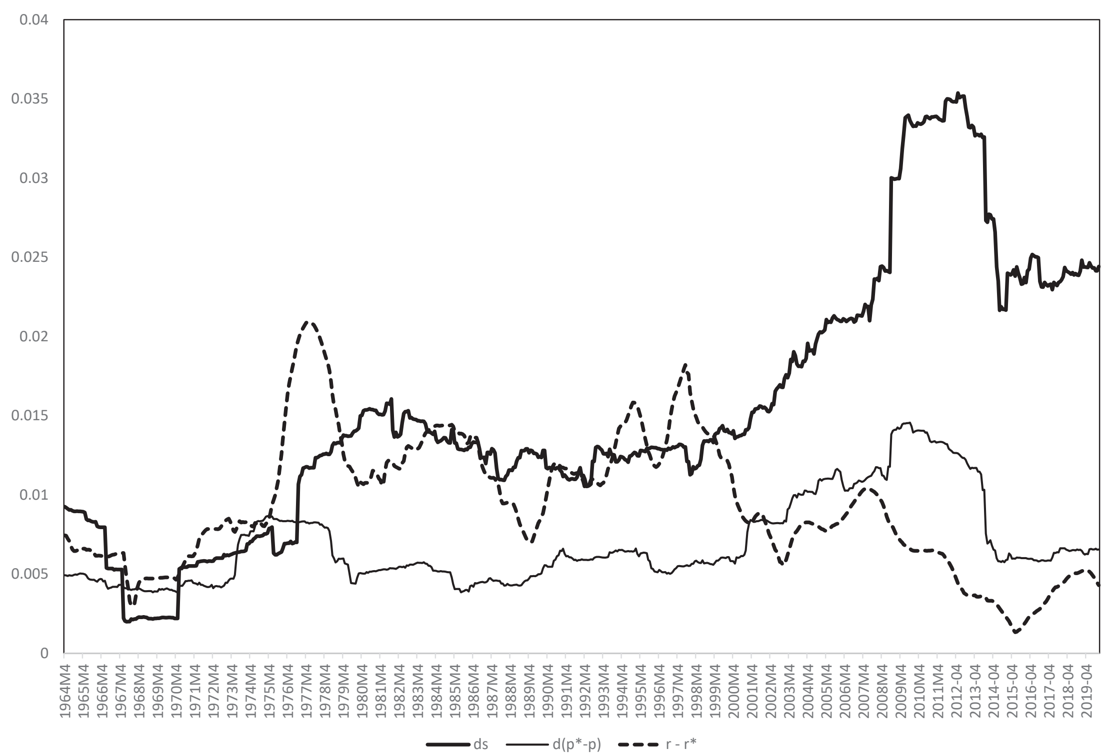
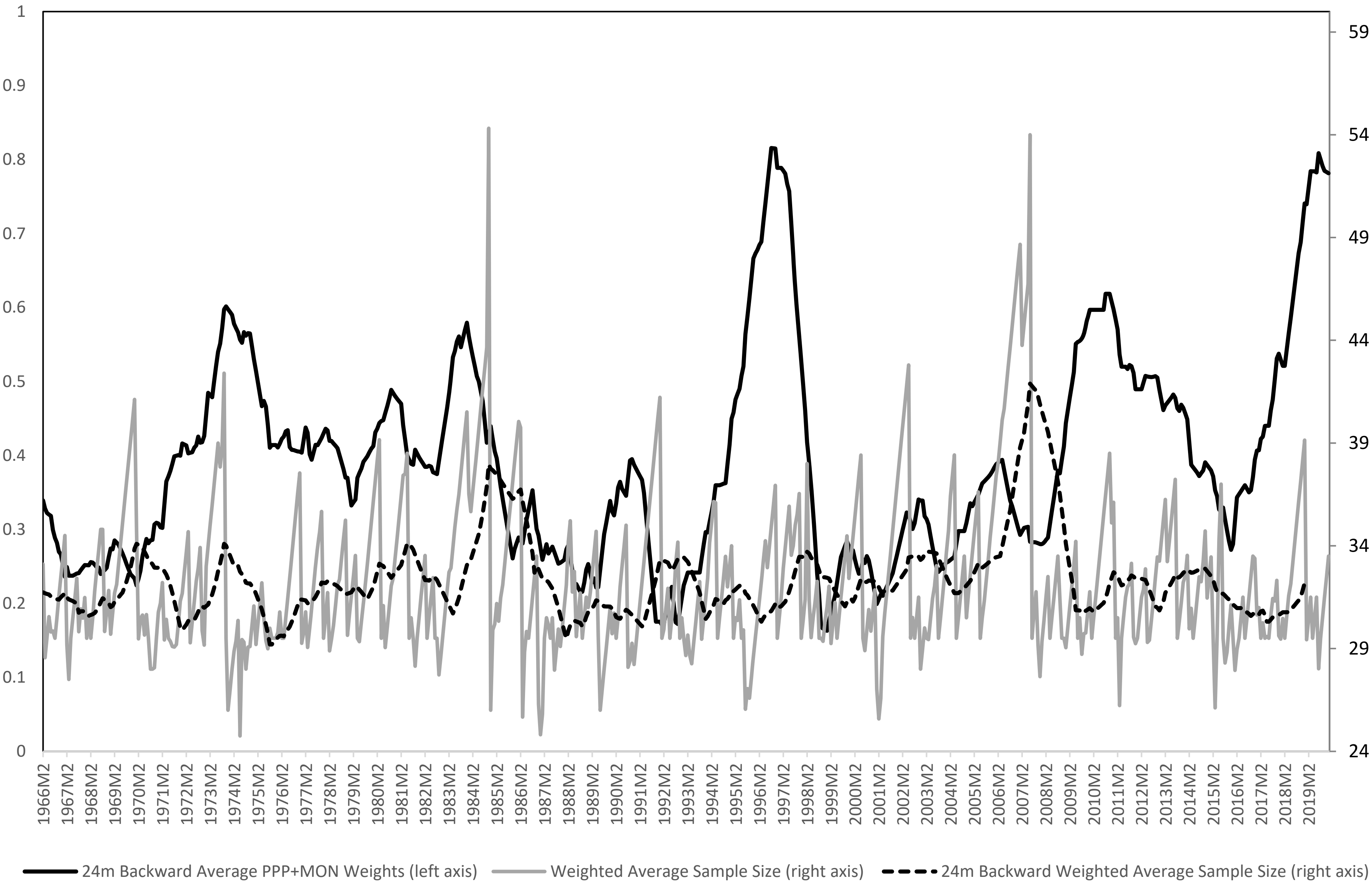


Fig. 2: Meta-NNT: Smoothed Sum of Weights for PPP and Monetary Models and Weighted Average Sample Size: Canada: 1966m2- 2019m12



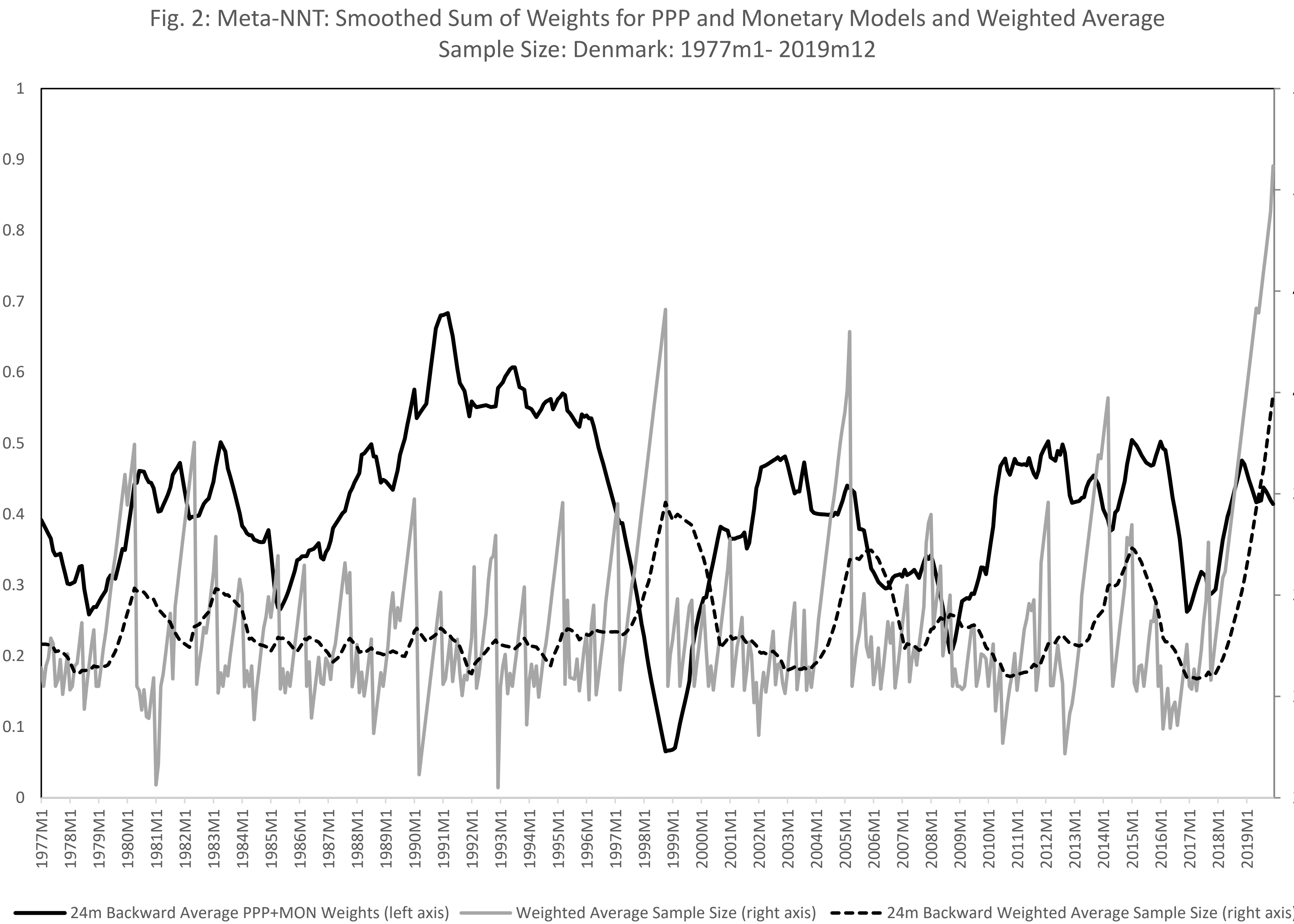


Fig. 2: Meta-NNT: Smoothed Sum of Weights for PPP and Monetary Models and Weighted Average Sample Size: Denmark: 1977m1- 2019m12
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
24
29
34
39
44
49
54
59
1977M1
1978M1
1979M1
1980M1
1981M1
1982M1
1983M1
1984M1
1985M1
1986M1
1987M1
1988M1
1989M1
1990M1
1991M1
1992M1
1993M1
1994M1
1995M1
1996M1
1997M1
1998M1
1999M1
2000M1
2001M1
2002M1
2003M1
2004M1
2005M1
2006M1
2007M1
2008M1
2009M1
2010M1
2011M1
2012M1
2013M1
2014M1
2015M1
2016M1
2017M1
2018M1
2019M1
24m Backward Average PPP+MON Weights (left axis)
Weighted Average Sample Size (right axis)
24m Backward Weighted Average Sample Size (right axis)

Fig. 2: Meta-NNT: Smoothed Sum of Weights for PPP and Monetary Models and Weighted Average Sample Size: Japan: 1967m1- 2019m12

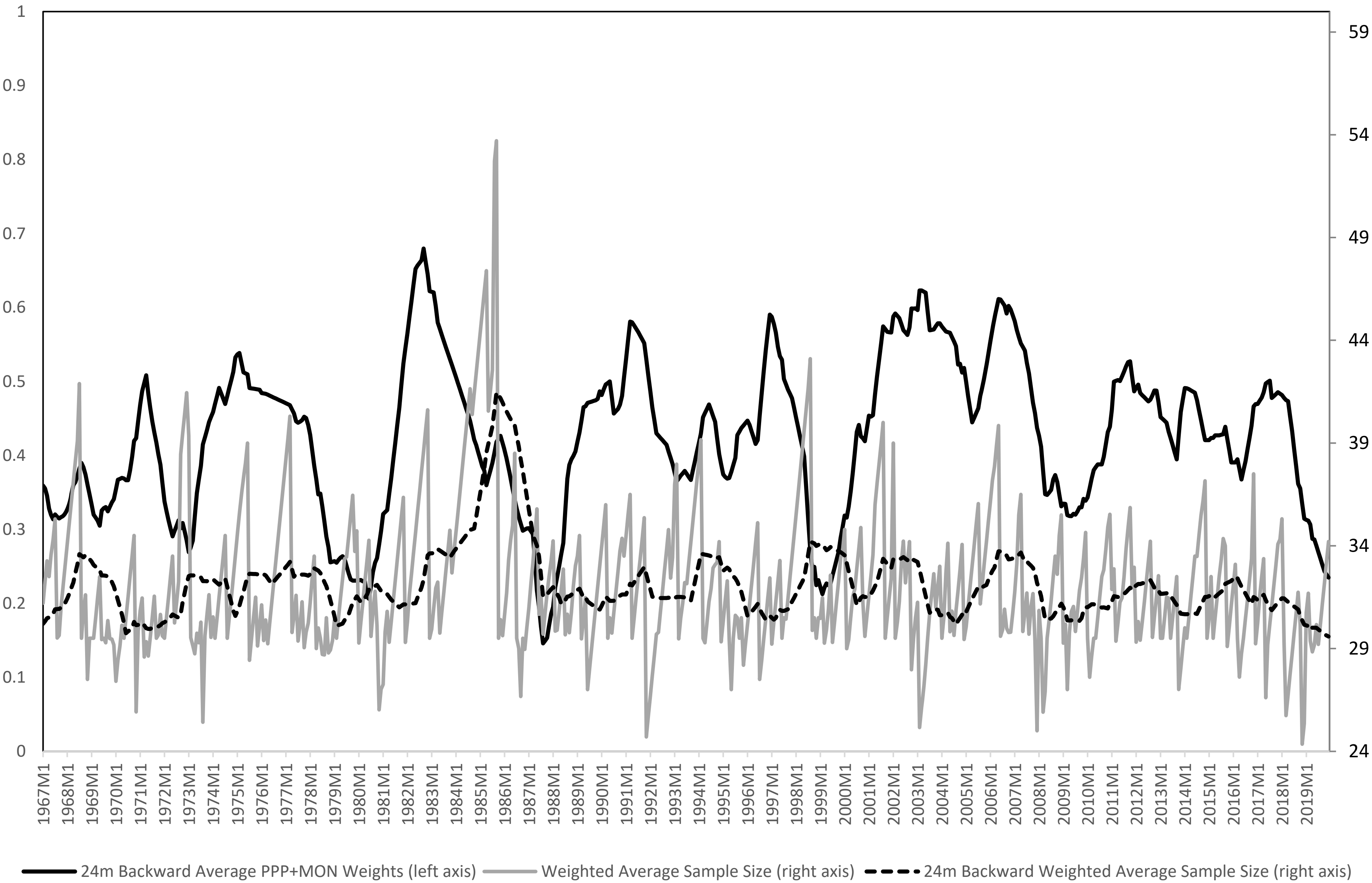


Fig. 2: Meta-NNT: Smoothed Sum of Weights for PPP and Monetary Models and Weighted Average Sample Size: Sweden: 1977m12 - 2019m12

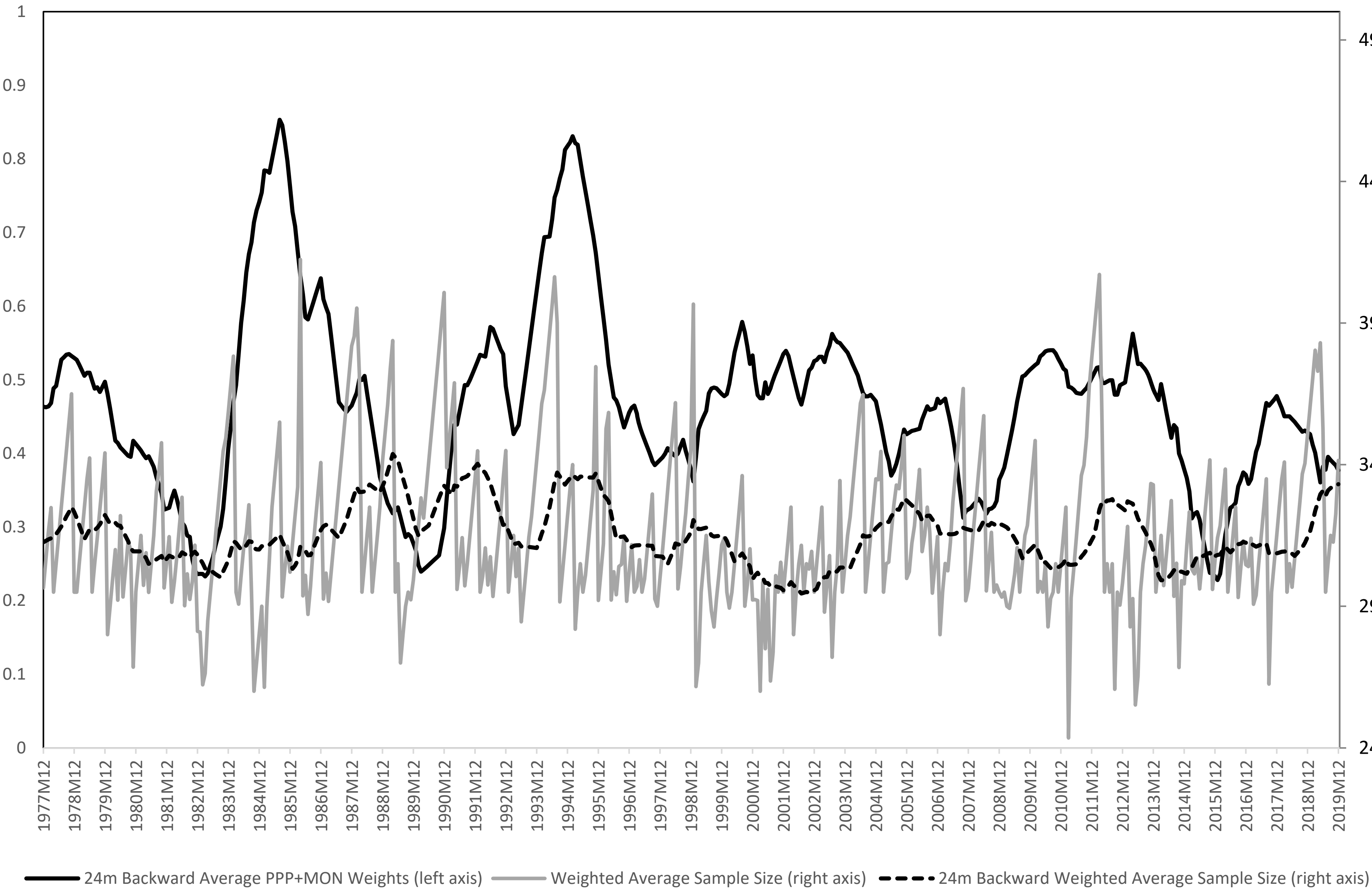


Fig. 2: Meta-NNT: Smoothed Sum of Weights for PPP and Monetary Models and Weighted Average Sample Size: UK: 1976m4 - 2019m12