The Contribution of Economic Fundamentals to Movements in Exchange Rates

Nathan S. Balke*
Department of Economics, Southern Methodist University
Dallas, TX 75275
and
Research Department, Federal Reserve Bank of Dallas
Office: (214) 768-2693
Fax: (214) 768-1821
E-mail: nbalke@mail.smu.edu

Jun Ma
Department of Economics, Finance and Legal Studies
Culverhouse College of Commerce & Business Administration
University of Alabama
Tuscaloosa, AL 35487-0024
Phone: 205-348-8985
Email: jma@cba.ua.edu

Mark E. Wohar
Department of Economics
Mammel Hall 332S
University of Nebraska at Omaha
Omaha, NE 68182-0286
Phone: 402-554-3712
Fax: 402-554-2853
Email: mwohar@mail.unomaha.edu

October 31, 2011

*Corresponding Author.

The Contribution of Economic Fundamentals to Movements in Exchange Rates

Abstract

A puzzle in the international finance literature has been that fundamental variables do not help forecast the future exchange rate change better than the random walk benchmark. Recently Engel and West (2005, 2006) show that such a result can be explained by the present value model of the exchange rate if the discount factor for the expectation of future fundamental variables is close to one and the fundamental variables are I(1). The approach we take in this paper allows us not only to directly estimate the discount factor but also to study the expectation dynamics that are important in evaluating these exchange rate models. Our estimates are based on a century of data for the US and UK. We employ a Bayesian approach to account for both uncertainty about the specification of the underlying state space model as well as parameter uncertainty. First, we show that the degree of model uncertainty is great and that the implied contribution of observed monetary fundamentals is imprecisely estimated. Second, we deal with the weak identification by bringing additional information to bear on the analysis. This additional information comes in the form of data on interest rate and price differentials and prior information about PPP half-lives and the semi-elasticity of money demand. In general, we find that monetary fundamentals (money and output differentials) and money demand shifters contribute most to movements in exchange rates while uncovered interest parity risk premium contribute to a lesser degree.

1. Introduction

The well-known papers by Meese and Rogoff (1983a, 1983b) showed that a simple random walk model for exchange rates can beat various time series and structural models in terms of out-of-sample forecasting performance. Although some of the subsequent literature on exchange rate predictability find evidence in favor of beating the random walk benchmark, most of those results do not hold up to scrutiny. The extant literature has found the linkage between the nominal exchange rate and fundamentals to be weak (Cheung, Chinn, and Pascual 2005; and Sarno, 2005). This weak linkage has become known as the "exchange rate disconnect puzzle".

Engel and West (2005) took a new line of attack in this analysis and demonstrate that this so-called disconnect between fundamentals and nominal exchange rates can be reconciled within a rational expectations model. The Engel and West (2005) model implies that the exchange rate is the present discounted value of expected economic fundamentals. Specifically,

$$s_t = f_t + E_t \left[\sum_{j=1}^{\infty} \psi^j \cdot \Delta f_{t+j} \right] + R_t \qquad (1.1)$$

where s_t is the spot exchange rate, f_t is the current value of observed fundamentals (for example money growth and output growth differentials), and ψ is the discount factor. The term R_t includes current and expected future values of unobserved fundamentals (risk premia, money demand shocks, etc) as well as perhaps "nonfundamental" determinants of exchange rate movements.

The "exchange rate disconnect puzzle" reflects the fact that fluctuations in $s_t - f_t$ can be "large" and persistent, while the promise of the present value approach is that this disconnect can be explained by the expectations of future fundamentals. The potential empirical success of the Engel and West model hinges on two major assumptions. First, fundamentals are non-stationary.

Second, the factor used to discount future fundamentals is "large" (between 0.9 and unity). Nonstationary fundamentals impart nonstationarity to exchange rates while a large discount factor gives greater weight to expectations of future fundamentals relative to current fundamentals. As a result, current fundamentals are only weakly related to exchange rates as exchange rates appear to follow an approximate random walk. The first assumption of nonstationary fundamentals has been supported by empirical work (Engel and West, 2005; Engel, Mark, and West, 2007), however, only recently has there been direct evidence in support for the second assumption of a large discount factor (Sarno and Sojli, 2009).

The key research question that still remains is to what extent can expectations of future fundamentals explain exchange rate movements? The challenge in evaluating the present value model is that not only the expected future fundamentals are not observed but other economic fundamentals, i.e. the R_t in equation (1.1), are also not observed. Indeed, Engel and West (2005) acknowledge that the kind of decompositions based on forecasting observed fundamentals such as those applied to stock prices (see Campbell and Shiller (1988)) is made difficult by the presence of unobserved fundamentals.

In this paper, we use a simple monetary model of exchange rates to specify explicitly the relationship between economic fundamentals and exchange rates. To sharpen our focus on expectations about future fundamentals, we use a state space model to conveniently model the relationship between observed fundamentals and the unobserved predictable components of fundamentals. We integrate the state space model into the present value model of the exchange rate to show the links between the predictable component of fundamentals and exchange rate fluctuations. We use annual data on the pound to the dollar exchange rate, money, output, prices, and interest rates for the UK and US from 1880 to 2010. The directly observed

fundamental in our model is money supply differentials between the UK and US minus output differentials between the UK and the US. This variable has been the primary focus of the literature's examination of fundamentals' contribution to exchange rate movements. For example, Mark (1995) evaluate the ability of this variable to forecast the future exchange rate movements for a set of countries including the United States, Canada, Germany, Japan, and Switzerland since the end of the Bretton Woods regime. Rapach and Wohar (2002) construct this variable for 14 industrial countries covering a period of more than a century and study the cointegration relationship between the exchange rates and the fundamentals. Mark and Sul (2001) further demonstrate that the panel data techniques are able to find more evidence of predictability of this variable to the future exchange rate movements. More recently, Cerra and Saxena (2010) conduct a comprehensive study of a very large dataset consisting of 98 countries and find more evidence that this fundamental variable help forecast the future exchange rate movements.

We show the difficulties of using only information on observed fundamentals and exchange rates to infer the expectations about future fundamentals. Using Bayesian model averaging across different specifications of a state space model, we show that the posterior distribution of the contribution of observed fundamentals to the variance of exchange rates is bimodal, with roughly equal weight placed on close to a zero contribution and on close to a 100% contribution. The reason for great uncertainty about the relative contributions of observed fundamentals is that in the data the predictable component of changes in observed fundamentals is relatively small compared to the unpredictable component--most of the information about future fundamentals is contained in exchange rates rather than observable fundamentals. This

makes identification of the separate contribution of expectations of future observed fundamentals problematic.

To solve this identification problem, we bring additional information to bear on the analysis. Within the context of the basic monetary model of exchanger rates, data on interest rate and price differentials provide information about two previously unobserved fundamentals—money demand shifters and a risk premium (deviations from uncovered interest parity). Another source of information is prior information about key parameters in the state space model. Specifically, prior information about the half-life of deviations from purchasing power parity helps to identify expected future deviations from purchasing power parity while prior information about the semi-elasticity of money demand which determines the value of the discount factor. Adding this additional information, results in sharper inference about the relative contribution of various fundamentals. We find that monetary fundamentals, money minus output differentials and particularly money demand shifters, explain the bulk of exchange rate movements. Fluctuations in the risk premium play a lesser role.

Our findings have very important implications to the fundamental exchange rate models that relate the exchange rate fluctuations to the economic fundamentals such as the output and monetary factors. Our results indicate that these economic fundamentals, either directly observed or indirectly inferred, contribute to the exchange rate movement in a substantial way. The large literature that finds it hard for economic models to produce a better out-of-sample forecasts for the exchange rate than a random walk may simply be due to the predictable component of fundamentals is small relative to the unpredictable component. That is, a simple regression cannot detect the small signals buried under the volatile noise.

The rest of the paper is organized as follows. In section 2, we outline the simple monetary model of exchange rates used by Engel and West (2005) to show how the spot exchange rate can be written as a function of expectations of future fundamentals, some of which are observed and some of which are unobserved. In section 3, we develop a state space model to describe the dynamics of the predictable component of observed fundamentals and embed it in the simple rational expectations monetary model of exchange rates. In section 4, we demonstrate using Bayesian model averaging that there is substantial uncertainty about the quantitative contribution of observed fundamentals to exchange rate movements. In section 5, we use additional information to obtain tighter inferences about relative contributions of observed and unobserved fundamentals. Section 6 concludes.

2. The Monetary Exchange Rate Model

We start with the classical monetary model as below (all variables are in logarithm, and asterisk denotes foreign variable):

$$m_t - p_t = \phi y_t - \lambda i_t + v_t^{md} \tag{2.1}$$

$$m_t^* - p_t^* = \phi y_t^* - \lambda i_t^* + v_t^{*md} \tag{2.2}$$

The variable definitions used in our model are the natural log of money supply (m), natural log of price level (p), and nominal interest rate (i). Variables with asterisk represent the foreign country. The terms v_t^{md} and v_t^{*md} represent unobserved variables that shift money demand.

We integrate into the model a generalized Uncovered Interest Parity (UIP) condition that allows for a time-varying risk premium, r_t^{uip}, since in general the UIP fails and the equilibrium model would imply non-trivial risk premium (see Engel (1996)):

$$i_t - i_t^* = E_t s_{t+1} - s_t + r_t^{uip} \tag{2.3}$$

where s_t is the natural log of the exchange rate. To complete model, we add the Purchasing Power Parity (PPP) relationship:

$$s_t = p_t - p_t^* + r_t^{ppp}.$$ (2.4)

Since the PPP in general only holds in the long run (Rogoff (1996)), the variable r_t^{ppp} picks up all deviations from the PPP.[1]

Combining (2.1) through (2.4) together, we can derive a stochastic difference equation that describes how exchange rate would depend on observed monetary fundamentals and an unobserved remainder. Similar results may be found in standard textbooks such as Mark (2001). The algebra can be manipulated so as to express the exchange change rate determination in terms of its deviation from observed fundamentals, similar to the stock price decomposition by Campbell and Shiller (1988):

$$s_t - f_t = \psi \cdot E_t[s_{t+1} - f_{t+1}] + \psi \cdot E_t[\Delta f_{t+1}] + \psi \cdot r_t$$ (2.5)

where, $f_t \equiv (m_t - m_t^*) - \phi(y_t - y_t^*)$ is the observed monetary fundamental, and $\psi = \dfrac{\lambda}{1 + \lambda}$ is the so-called discount factor. In the following estimation exercise, we set $\phi = 1$ as in Rapach and Wohar (2002) and Mark (1995). The unobserved term, $\psi \cdot r_t$, consists of the unobserved money demand shifter as well as deviations from both uncovered interest rate parity and purchasing power parity: $\psi \cdot r_t = \psi r_t^{uip} + (1 - \psi)r_t^{ppp} - (1 - \psi)r_t^{md}$ where $r_t^{md} = v_t^{md} - v_t^{*md}$.

Recently, Engel and West (2005, 2006) show that this asset price formulation for the exchange rates is very general and can be derived from a variety of monetary policy models including the Taylor rule and may include more fundamental information under that type of

---

[1] One way to avoid the short-run deviation from the PPP is to choose to estimate the model at a very low frequency at which the friction would go away and PPP would hold.

monetary policy rules. One advantage of writing the exchange rate determination in terms of its deviation from its current fundamentals is that we can study which variable will most likely respond to such deviations, and therefore determining which factor is more responsible for these deviations.

To further solve the model, we can iterate eq. (2.5) forward. Under the assumption of no explosive solution, the model can be solved as below:

$$s_t - f_t = E_t \left[\sum_{j=1}^{\infty} \psi^j \cdot \Delta f_{t+j} \right] + E_t \left[\sum_{j=0}^{\infty} \psi^{j+1} \cdot r_{t+j} \right] \tag{2.6}$$

where, $s_t - f_t$ is the deviation of current exchange rate from its current observed monetary fundamental. Eq. (2.6) is similar to the present discounted value formula for the exchange rate derived in Engel and West (2005, 2006), and this equation states that any deviation of current exchange rate from its observed fundamentals should reflect the variation of the present discounted value of agent's expected future economic fundamentals.

Engel and West (2005, 2006) argue that the fundamentals may be primarily driving the exchange rate but the close-to-unity discount factor ψ leads to near random walk behavior for the exchange rate and therefore it has been extremely hard to forecast any future exchange rate movement using the information of fundamentals. To see this intuitively, we may re-write (2.5) to express the one-period-ahead exchange rate change as below:

$$E_t \Delta s_{t+1} = \frac{(1-\psi)}{\psi} \cdot (s_t - f_t) - r_t \tag{2.7}$$

If the discount factor ψ approaches unity, the current fundamentals would forecast a very small fraction of the near future exchange rate change, because essentially the fundamental information has too small a loading compared with the remainder. A straightforward extension of (2.7) shows that a similar result holds for the multi-period-ahead exchange rate change.

3. Decomposing the contribution of observed fundamentals and unobserved shocks

One obstacle in evaluating the above exchange rate model is that what matters in explaining current deviation of exchange rate from its fundamentals are agent's expectations of future fundamentals and the remainder but these expectations are not directly observable. The state-space model offers a convenient framework in which we can model the expectations as latent factors and allow them to have flexible dynamics. In doing this, we can extract agent's expectations using the Kalman filter and decompose the current deviation of exchange rate from its fundamentals into the contributions of future observed fundamentals and current and future unobserved remainder. Furthermore, instead of assuming that the discount factor is close to unity as in Engel and West (2005, 2006), we can directly estimate the discount factor and provide further statistical evidence for the Engel and West model.

Denote the expectations by $E_t[\Delta f_{t+1}] = g_t$ and $E_t[r_{t+1}] = \mu_t$. Then the realized variables are the sum of their conditional expectations and the realized shocks:

$$\Delta f_t = g_{t-1} + \varepsilon_t^f \tag{3.1}$$

$$r_t = \mu_{t-1} + \varepsilon_t^r \tag{3.2}$$

Where the realized shocks (or forecast errors) ε_t^f and ε_t^r are white noise. At the same time, the expectations processes may embody important dynamics, and here we assume AR processes for them:

$$(1 - \phi_g(L)L)g_t = \varepsilon_t^g \tag{3.3}$$

$$(1 - \phi_\mu(L)L)\mu_t = \varepsilon_t^\mu \tag{3.4}$$

where ε_t^g and ε_1^μ are expectation shocks. The four shocks $\begin{bmatrix} \varepsilon_t^g & \varepsilon_t^\mu & \varepsilon_t^f & \varepsilon_t^r \end{bmatrix}$ are allowed to be contemporaneously correlated but they are serially uncorrelated.

Using equation (2.6) and evaluating expectations, we can write the exchange rate relative to observed current monetary fundamental as:

$$s_t - f_t = B_1(L)g_t + B_2(L)\mu_t + \psi\mu_{t-1} + \psi\varepsilon_t^r, \tag{3.5}$$

where $E_t\left[\sum_{j=1}^{\infty}\psi^j \cdot \Delta f_{t+j}\right] = B_1(L)g_t$ and $E_t\left[\sum_{j=1}^{\infty}\psi^{j+1} \cdot r_{t+j}\right] = B_2(L)\mu_t$. The contribution of

observed fundamentals, $E_t\left[\sum_{j=1}^{\infty}\psi^j \cdot \Delta f_{t+j}\right]$, will in fact depend on current (and possibly lagged)

values of the unobserved component g_t. The coefficients of the lag polynomials, $B_1(L)$ and

$B_2(L)$, depend on the values of ψ, $\phi_g(L)$, and $\phi_\mu(L)$. In general, the larger the value of ψ

and the more persistent is g_t, the larger are the coefficients of $B_1(L)$.

Given the observed monetary fundamentals, Δf_t, we can write the model in a state space form with the measurement equation:

$$\begin{bmatrix} \Delta f_t \\ s_t - f_t \end{bmatrix} = \begin{bmatrix} L & 0 & 1 & 0 \\ B_1(L) & B_2(L)+\psi L & 0 & \psi \end{bmatrix}\begin{bmatrix} g_t \\ \mu_t \\ \varepsilon_t^f \\ \varepsilon_t^r \end{bmatrix} \tag{3.6}$$

and the transition equation:

$$\begin{bmatrix} g_t \\ \mu_t \\ \varepsilon_t^f \\ \varepsilon_t^r \end{bmatrix} = \begin{bmatrix} \phi_g(L) & 0 & 0 & 0 \\ 0 & \phi_\mu(L) & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} g_{t-1} \\ \mu_{t-1} \\ \varepsilon_{t-1}^f \\ \varepsilon_{t-1}^r \end{bmatrix} + \begin{bmatrix} \varepsilon_t^g \\ \varepsilon_t^\mu \\ \varepsilon_t^f \\ \varepsilon_t^r \end{bmatrix}. \tag{3.7}$$

As none of the state variables are directly observed and must be inferred from the two observed time series, identification of the state space model will depend on the dynamics of the observed time series and its variance/covariance matrix (see Appendix A). Balke and Wohar (2002) and Ma and Wohar (2011) show that inference about the relative contribution of the unobserved components may be very weak if the observed fundamental does not provide a lot of direct information about the relative sizes of the predictable component of fundamentals. For example, if the variance of innovations to the predictable components, σ_g^2, is small relative to the unpredictable component, σ_f^2, then observations of Δf_t provide very little information about g_t, leaving only $s_t - f_t$ to infer both g_t and μ_t. One can see this by rewriting the state space model as a VARMA:

$$\begin{pmatrix} 1 - \phi_g(L)L & 0 \\ 0 & (1 - \phi_g(L)L)(1 - \phi_\mu(L)L) \end{pmatrix} \begin{pmatrix} \Delta f_t \\ s_t - f_t \end{pmatrix} =$$

$$\begin{pmatrix} L & 0 & (1 - \phi_g(L)L) & 0 \\ B_1(L)(1 - \phi_\mu(L)L) & (B_2(L) + \psi)(1 - \phi_g(L)L) & 0 & (1 - \phi_g(L)L)(1 - \phi_\mu(L)L) \end{pmatrix} \begin{pmatrix} \varepsilon_t^g \\ \varepsilon_t^\mu \\ \varepsilon_t^f \\ \varepsilon_t^r \end{pmatrix} \quad (3.8)$$

If $\sigma_g^2 \big/ \sigma_f^2$ is "small", then observations of Δf_t are not sufficient to identify $\phi_g(L)$--this polynomial is canceled out in the Δf_t equation. As there could be numerous combinations of $\phi_g(L)$ and $\phi_\mu(L)$ would yield the same autoregressive dynamics for $s_t - f_t$, $(1 - \phi_g(L)L)(1 - \phi_\mu(L)L)$, whether there is sufficient information to identify $\phi_g(L)$ and $\phi_\mu(L)$ would depend on the moving average dynamics of $s_t - f_t$. These may not be sufficiently rich to identify $\phi_g(L)$ and $\phi_\mu(L)$.

Taking annual observations of the nominal exchange rate, relative money supply, and relative real GDP between the US and UK and UK and US interest rates from 1880 to 2010 (see appendix C for details of data construction), the top panel of Figure 1 plots the log UK-US exchange rate, s_t, and the log level of the observed monetary fundamental, f_t, while the lower panel of Figure 1 plots the realized fundamentals growth (Δf_t) along with the deviation of current exchange rate from the observed monetary fundamentals ($s_t - f_t$). The deviation, $s_t - f_t$, is quite persistent and volatile while the realized fundamentals growth is much less persistent and less volatile. Figure 1 suggests that the persistent component g_t is likely to be "small" relative to Δf_t which in turn implies that it may be difficult to separately identify g_t and μ_t from data on $s_t - f_t$ alone.

We address this potential identification problem in two ways. First, in following section, we estimate decompositions of exchange rate deviations from monetary fundamentals for five separate non-nested models that reflect different specifications of $\phi_g(L)$ and $\phi_\mu(L)$. We use Bayesian model averaging to account for model uncertainty in attributing the source of exchange rate fluctuations. Second, in section 5, we bring additional information to bear on the analysis. This additional information comes in the form of data on interest rate differentials and price level differentials. We use the simple monetary model outlined above to link the additional observations to components in the unobserved remainder term. This additional information allows us to break up the remainder term into its constituent components: r_t^{uiip}, r_t^{ppp}, and r_t^{md}. The other source of information is prior information about key parameters in the state space model—specifically, the half-life of deviations from PPP and the semi-elasticity of money demand λ which determines the value of the discount factor ψ in equation (2.6).

4. Bayesian model averaging of alternative state space models

As suggested above, most of the information about the predictable component of the observed monetary fundamentals might actually be in the exchange rate, $s_t - f_t$, rather than in observed monetary fundamentals growth itself, Δf_t. This suggests that the model is weakly identified as essentially a single data series ($s_t - f_t$) is used to identify two components (g_t and μ_t). To evaluate to what extent this potential identification problem holds in practice, we consider five alternative, non-nested models. Each of these five models gives rise to an ARMA(4,4) model for $s_t - f_t$ but will imply very different exchange rate decompositions. We will take as the benchmark model an AR(2) for both $\phi_g(L)L$ and $\phi_\mu(L)L$ with the innovations in the four components $\begin{bmatrix} \varepsilon_t^g & \varepsilon_t^\mu & \varepsilon_t^f & \varepsilon_t^r \end{bmatrix}$ to be correlated with one another. The alternative four models are different combinations of AR models for $\phi_g(L)L$ and $\phi_\mu(L)L$ that still yields an ARMA(4,4) model for $s_t - f_t$. These include AR(4) and AR(0), AR(3) and AR(1), AR(1) and AR(3), AR(0) and AR(4) as well as AR(2) and AR(2) for $\phi_g(L)L$ and $\phi_\mu(L)L$ respectively.

We take a Bayesian approach to account for uncertainty about the specification of the underlying state space model. To evaluate these alternative models, we first estimate the posterior distribution of the parameters for each of the five competing models. The posterior distribution of the parameters given the data and a particular model, we will denote as $P[\theta_m \mid \mathbf{Y}_T, M = m]$, where m is the model and θ_m is the parameter vector of model m. The posterior probability of model m is then:

$$P[m|Y] = \frac{B(m)h(m)}{\sum_{m \in M} B(m)h(m)},$$ (4.1)

where M is the set of models, $h(m)$ is the prior probability of model m,

$B(m) = \int L(\mathbf{Y}_T, \theta_m)h(\theta_m)d\theta_m$, $L(\mathbf{Y}_T, \theta_m)$ is the likelihood, and $h(\theta_m)$ is the prior density of

the parameters. As we are interested in determining the relative contribution of fundamentals to

exchange rate decomposition we calculate the variance decomposition, $V(\theta_m, m)$, implied by a

given model m and parameter vector, θ_m. Given the posterior distribution of the parameters, we

can obtain the posterior distribution of the variance decomposition for a given model,

$P[V(\theta_m, m)|m, \mathbf{Y}_T]$. Finally, we can employ Bayesian model averaging to account for model

uncertainty on the posterior distribution of the variance decomposition.

Given the nonlinear (in parameters) structure of the model, there is no closed form

solution for the posterior distribution given standard priors; therefore, we use a Metropolis-

Hastings Markov Chain Monte Carlo (MH-MCMC) to approximate the posterior distribution of

the parameters given the model (see Appendix B for details). We can easily construct the

posterior distribution of the variance decomposition from the results of the MH-MCMC as well

the posterior probabilities for the five models given the data. In this section, we consider the

case of very diffuse priors so that the likelihood function is the principal determinant of the

posterior distribution.[2] The posterior distribution is based on 500,000 draws from the MH-

MCMC after a burn-in period of 500,000 draws.

---

[2] Formally, for each of the models the individual autoregressive parameters have a prior distribution of N(0,100), the prior distribution ψ is U(0,1), the variances in Q are distributed U(0,1000) while the co-variances in Q are distributed U(-1000,1000). Draws in which F matrix implies nonstationarity are rejected as are draws where the Q matrix is not positive definite. These prior distributions ensure that for this model and data, the acceptance in the Metropolis-Hastings sampler depends only on the likelihoods; thus, when comparing models the likelihoods are going to be decisive.

Figure 2 presents for the benchmark model (AR(2), AR(2)) the histogram of the posterior distribution of parameters (the histograms for the other four models are available upon request). Consider the posterior distribution of the autoregressive coefficients ϕ_{g1} and ϕ_{g2}. These parameters play a key role in determining the contribution of observed fundamentals to exchange rate movements. The values of ϕ_{g1} and ϕ_{g2} determine the persistence of the predictable component of observed fundamental growth, g_t, In general, the more persistent is g_t, the larger its effect on exchange rates. From Figure 2, the posterior distributions of these parameters are quite disperse—substantially more disperse than implied by just the standard errors from standard maximum likelihood estimation (available upon request).[3] On the other hand, the joint probability distribution of ϕ_{g1} and ϕ_{g2} has mass concentrated close to $\phi_{g1} + \phi_{g2} = 1$, see the three-dimensional histogram of ϕ_{g1} and ϕ_{g2} and the associated heat plot in Figure 3. Thus, while there is substantial uncertainty about individual parameters, there is little uncertainty about the overall persistence of the predictable component of the fundamentals in the benchmark model.

Figure 4 displays for all five models the posterior distribution of the contributions of the predictable component of the fundamentals, g_t, to the variance of deviations of the exchange rate from observed monetary fundamentals, $s_t - f_t$. For the benchmark model, the posterior distribution of the observed fundamental's contribution to exchange rate variance is concentrated around 100 percent but does show a small secondary mode close to zero. Model 2, in which $\phi_g(L)$ is an AR(4) and $\phi_\mu(L)$ is an AR(0), also implies that the observed fundamentals

---

[3] Note also the posterior distribution for some of parameters is skewed or multi-modal suggesting that using standard maximum likelihood methods for inference would miss these features.

explains nearly all the variance of exchange rates.[4] Models 4 and 5, on the other hand, imply

that observed monetary fundamentals explain almost none of the variance of exchange rates

($\phi_g(L)$ is an AR(1) and $\phi_\mu(L)$ is an AR(3) for Model 4 and AR(0) and AR(4), respectively, for

Model 5). Model 3 ($\phi_g(L)$ is an AR(3) and $\phi_\mu(L)$ is an AR(1)), suggests a bimodal distribution

with probability mass concentrated on the extremes.

Which of the five models does the data prefer? From the MCMC posterior distribution,

we can construct the posterior distribution of the log likelihoods, $log(L(\mathbf{Y}_T,\theta_m))$, for each

model. Figure 5 displays the cumulative distribution of the posterior distribution of log

likelihoods for the five models.[5] From Figure 5, the CDF of the posterior likelihood values are

quite close to one another, and while the benchmark model appears to stochastically dominate

models 3, 4, and 5 it does not dominate model 2. Table 1 presents the posterior probabilities of

the five models. While the benchmark model has the highest posterior probability, it is not

substantially higher than Model 5.

Figure 6 plots the histogram for the observed monetary fundamental's variance

decomposition of $s_t - f_t$ once we account for model uncertainty. Here the variance

decomposition for each model is weighted by its posterior. Taking into account of model

uncertainty suggests that a bimodal distribution for the contribution of observed monetary

fundamentals on the variance of $s_t - f_t$, with the probability mass concentrated on either a zero

contribution or 100 percent contribution. This suggests that using data on exchange rates and

---

[4] We use the term "observed" monetary fundamentals because we have direct observations on f = ((m-m*) –(y-y*)), but we do not have direct observations on r^{uip},or r^{md},

[5] The maximum log likelihood values for the five models were: -889.44, -884.81, -889.87, -889.61, and -890.18, respectively.

observed monetary fundamentals alone is not sufficient to determine whether to what extent exchange rates are driven by monetary fundamentals.

5. Prior information and data on interest differentials and money demand

As the previous section suggests, using data on observed monetary fundamentals, Δf_t, and exchange rates does not yield a precise decomposition of sources of exchange rate movements. In this section, we use a combination of additional data, restrictions implied by the model, and prior information about key parameters to sharpen inferences about the sources of exchange rate fluctuations. Specifically, we break the remainder, $\psi \cdot r_t$, into its constituent parts: an unobserved money demand shifter r_t^{md}, deviations from uncovered interest parity r_t^{uip}, and deviations from purchasing power parity r_t^{ppp}. These along with the observed monetary fundamentals will help determine exchange rate movements.

We assume each of the four components of exchange rates consists of a predictable and unpredictable component:

$$\Delta f_{t+1} = g_t + \varepsilon_{t+1}^f, \qquad \phi_g(L)g_{t+1} = \varepsilon_{t+1}^g \tag{5.1}$$

$$r_{t+1}^{ppp} = \mu_t^{ppp} + \varepsilon_{t+1}^{ppp}, \qquad \phi_{\mu,ppp}(L)\mu_{t+1}^{ppp} = \varepsilon_{t+1}^{\mu,ppp} \tag{5.2}$$

$$r_{t+1}^{uip} = \mu_t^{uip} + \varepsilon_{t+1}^{uip} \qquad \phi_{\mu,uip}(L)\mu_{t+1}^{uip} = \varepsilon_{t+1}^{\mu,uip} \tag{5.3}$$

$$r_{t+1}^{md} = \mu_t^{md} + \varepsilon_{t+1}^{md} \qquad \phi_{\mu,uip}(L)\mu_{t+1}^{uip} = \varepsilon_{t+1}^{\mu,uip} \tag{5.4}$$

We specify $\phi_g(L)$, $\phi_{\mu,ppp}(L)$, $\phi_{\mu,uip}(L)$, $\phi_{\mu,md}(L)$ to be AR(1)'s to keep the model parsimonious. The state vector is:

$$S_t = \begin{bmatrix} g_t & g_{t-1} & \mu_t^{uip} & \mu_{t-1}^{uip} & \mu_t^{ppp} & \mu_{t-1}^{ppp} & \mu_t^{md} & \mu_{t-1}^{md} & \varepsilon_t^f & \varepsilon_t^{uip} & \varepsilon_t^{ppp} & \varepsilon_t^{md} \end{bmatrix}'$$ and the

transition equation to be

$$S_t = FS_{t-1} + \Sigma_t \tag{5.5}$$

with

$$
F = \begin{bmatrix}
\phi_g & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \phi_{uip} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \phi_{ppp} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \phi_{md} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix} \tag{5.6}
$$

and

$$\Sigma_t = \begin{bmatrix} \varepsilon_t^g & 0 & \varepsilon_t^{\mu,uip} & 0 & \varepsilon_t^{\mu,ppp} & 0 & \varepsilon_t^{\mu,md} & 0 & \varepsilon_t^f & \varepsilon_t^{uip} & \varepsilon_t^{ppp} & \varepsilon_t^{md} \end{bmatrix}'.$$

We use data on interest rate and price level differentials, to construct two additional observations equations. Our observation vector now consists of four variables: relative money demand ($f_t - (p_t - p_t^*)$), interest rate differentials ($i_t - i_t^*$), along with the growth rate in observed monetary fundamentals (Δf_t) and the exchange rate relative to current observed monetary fundamental ($s_t - f_t$). The relative money demand equation is given by:

$$f_t - (p_t - p_t^*) = -\frac{\psi}{1-\psi}(i_t - i_t^*) + r_t^{md}. \tag{5.7}$$

The uncovered interest rate condition is given by:

$$i_t - i_t^* = E_t\left(s_{t+1} - f_{t+1}\right) + E_t\Delta f_{t+1} - \left(s_t - f_t\right) + r_t^{uip}. \tag{5.8}$$

Recall that the growth rate of observed monetary fundamentals is

$$\Delta f_t = g_{t-1} + \varepsilon_t^f, \tag{5.9}$$

and the exchange rate equation is given by

$$s_t - f_t = E_t \left[\sum_{j=1}^{\infty} \psi^j \cdot \Delta f_{t+j} \right] + E_t \left[\sum_{j=0}^{\infty} \psi^{j+1} \cdot r_{t+j} \right], \tag{5.10}$$

where $\psi \cdot r_t = \psi r_t^{uip} + (1-\psi)r_t^{ppp} - (1-\psi)r_t^{md}$.

Taking expectations and writing equations (5.7)-(5.10) in terms of the state variables, S_t, yields the following measurement equation:

$$\begin{bmatrix} f_t - (p_t - p_t^*) \\ i_t - i_t^* \\ \Delta f_t \\ s_t - f_t \end{bmatrix} = \begin{bmatrix} -\frac{\psi}{1-\psi}\left((H_d + H_{\Delta f})F - H_d + H_{uip}\right) + H_{md} \\ (H_d + H_{\Delta f})F - H_d + H_{uip} \\ H_{\Delta f} \\ H_d \end{bmatrix} S_t \tag{5.11}$$

where

$$H_{\Delta f} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}, \quad H_{uip} = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

$$H_{md} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$H_d = \begin{bmatrix} B_{11} & B_{12} & B_{21} & B_{22} & B_{31} & B_{32} & B_{41} & B_{42} & 0 & \psi & 1-\psi & -(1-\psi) \end{bmatrix},$$

$$\begin{bmatrix} B_{11} & B_{12} \end{bmatrix} = \begin{bmatrix} \psi \cdot (1 - \psi \phi_g)^{-1} & 0 \end{bmatrix},$$

$$\begin{bmatrix} B_{21} & B_{22} \end{bmatrix} = \begin{bmatrix} \psi^2 \cdot (1 - \psi \phi_{uip})^{-1} & \psi \end{bmatrix},$$

$$\begin{bmatrix} B_{31} & B_{32} \end{bmatrix} = \begin{bmatrix} (1-\psi)\psi \cdot (1 - \psi \phi_{ppp})^{-1} & (1-\psi) \end{bmatrix},$$

$$\begin{bmatrix} B_{41} & B_{42} \end{bmatrix} = \begin{bmatrix} (1-\psi)\psi \cdot (1 - \psi \phi_{md})^{-1} & (1-\psi) \end{bmatrix}.$$

Equations (5.5), (5.6), and (5.11) describe the state space model which includes observations on relative money demand, interest rate differentials, growth rate of observed monetary fundamentals, and the exchange rate.

We can also bring prior information to bear on the estimation of the model parameters. One of the unobserved factors represents deviations from purchasing power parity. We do not use observations on PPP deviations as they are just a linear combination of two other observation variables: $(f_t - (p_t - p_t^*))$ and $(s_t - f_t)$. There is, however, a large literature on PPP deviations that we can draw upon to provide information about μ_t^{ppp}. We assume a prior distribution for ϕ_{ppp} so that the half-life for PPP deviations is similar to that found in the literature (see Rogoff, 1996). Specifically, we assume a Beta(10,2) distribution for the prior distribution of ϕ_{ppp}. Figure 7 displays the prior distribution for ϕ_{ppp} as well as the prior distribution of PPP half-lives implied by ϕ_{ppp}.

One of the other key parameters in the model is $\psi = \dfrac{\lambda}{1+\lambda}$ where λ is the semi-elasticity of interest rates on money demand. We set the prior distribution of λ to be a normal distribution with mean equal to 25 and standard deviation of 15, truncated at zero. We chose this prior based on studies of long-run money demand that typically estimate the semi-elasticity of money to be in the range from around 10 to 40.[6] Figure 8 displays the prior distribution of λ and the implied prior distribution for ψ.[7]

Results:

Figure 9 displays posterior distribution of autoregressive parameters for the four predictable components: g_t, μ_t^{uip}, μ_t^{ppp}, and μ_t^{md}. The autoregressive parameters suggest that all

---

[6] Bilson (1978) estimates λ to be 60 in the monetary model, whereas Frankel (1979) finds λ to be equal to 29 and Stock and Watson (1993, 802, table 2, panel I) give a value of λ equal to 40. A more recent study by Haug and Tam (2006) suggest values ranging from around 10 to 20.
[7] Prior distributions on the other autoregressive parameters is fairly diffuse: $N(0.5,(1.5)^2)$. The variances in Q are distributed U(0,1000) while the co-variances in Q are distributed U(-1000,1000). Draws in which F matrix implies nonstationarity are rejected as are draws where the Q matrix is not positive definite.

the unobserved predictable components are fairly persistent. Figure 10 displays posterior

distribution of λ and ψ . The posterior distribution for λ suggests a semi-elasticity of around 25,

similar to the mode of the prior distribution, but it is substantially tighter than the prior

distribution for λ . The posterior distribution for the discount factor, ψ , suggests a value around

.96 not far from the mode of the prior, but again substantially tighter.

Figure 11 displays the posterior distribution of variance decomposition of $(s_t - f_t)$.[8] We

can see that the contributions of the risk premium (r^{uip}) and of deviations from PPP (r^{ppp}) are

small. However, the joint contribution of monetary fundamentals, g and r^{md}, is very large. The

greatest contribution comes from the money demand shifter, r^{md}, while the contribution of

predictable component of directly observed monetary fundamentals, g, is more modest. Thus,

the exchange rate disconnect in our context appears due to fluctuations in money demand.

Figure 12 displays historical decomposition of the <u>level</u> of the exchange rates. The

historical decompositions display the contribution of each of the states over the entire sample.

To calculate these decompositions, the parameters of the state space model are drawn from their

posterior distribution using Metropolis-Hasting MCMC. For each parameter draw $\theta^{(i)}$, we draw

$S_T^{(i)} = [S_1^{(i)}, ... , S_t^{(i)}, ... , S_T^{(i)}]$ from the conditional posterior distribution for the unobserved states,

$P(S_T \mid \theta^{(i)}, Y_T)$ using the "filter forward, sample backward" approach of Carter and Kohn

(1994). The contribution of the states in time period t for a given parameter and state draw is

$H(\theta^{(i)})S_t^{(i)}$. The historical decompositions reported in Figure 12-16 are based on the 5th, 95th ,

and 50th percentiles of the sample distribution of 500,000 draws from the Metropolis Hasting

---

[8] These values can be above 100% due to covariances.

Markov Chain. Note because the Q matrix (variance-covariance matrix in the transition equation) in the state space model is not diagonal, the states are in general correlated with one another and the historical decomposition are in general not orthogonal.

The top left panel displays the historical decomposition of the monetary fundamentals, both observed, f_t, and unobserved money demand shifters, r_t^{md}. The individual contributions of deviations from uncovered interest rate parity, r_t^{uip}, and deviations from PPP, r_t^{ppp}, are also presented in Figure 12. From Figure 12, one observes that most of movement in the UK/US exchange rate, especially in the later part of the sample, appear due to monetary fundamentals, and in particular to money demand shifters, r_t^{md}. The risk premium, r_t^{uip}, while not as important as either monetary fundamental, does contribute to exchange rate fluctuates in a few instances. Deviations from PPP, on the other hand, contribute very little to exchange rate fluctuations. The results suggest that monetary factors explain the long run swings of the exchange rate but deviations from uncovered interest parity, r_t^{uip}, appear to explain some of the short-run movements in exchange rates. The lower right graph in Figure 12 shows the total contribution of all the factors to movements of the exchange rate. This graph is a check to see that the contributions of the factors add up to the exchange rate itself.

We can also examine historical decomposition of the other observable variables. Figure 13 displays the historical decomposition of UK/US interest rate differentials. Here monetary factors, particularly the unobserved money demand shifter, contribute to movements in the UK/US interest rate differential. The uncovered interest rate premium has a modest contribution to interest rate differential movements while PPP deviations factor have only a minor contribution.

Figure 14 displays the historical decomposition of money demand differentials $(f_t - (p_t - p_t^*))$. Again, monetary factors (Δf_t and r_t^{md}) contribute to most of the fluctuations in $(f_t - (p_t - p_t^*))$ with the money demand shifter, r_t^{md}, being a large contributor. The money demand shifts has two effects here. There is a direct effect by shifting money demand. There is also an indirect in that r_t^{md} affects s_t which in turn affects interest rate differentials, $i_t - i_t^*$, which in turn leads to changes in money demand differentials. Deviations from uncovered interest parity and deviations from PPP yield only modest contributions.

Our original conjecture was that the predictable component of observed monetary fundamentals, g_t, was relatively small compared to observed fundamentals, Δf_t. Figure 15 presents the historical decomposition of Δf_t. From Figure 15, it is clear that fluctuations in the predictable component of Δf_t are relatively small. This is consistent with the results in section 4, where inference about the relative importance of the predictable component was plagued by weak identification when using only $s_t - f_t$ and Δf_t as observables. When we add additional information, the contribution of predictable component of Δf_t is small both for exchange rates and for Δf_t itself.

Recall that the sum of $(s_t - f_t)$ and money demand differentials $(f_t - (p_t - p_t^*))$ equal deviations from PPP. As a check of the model, we calculate historical decomposition of PPP deviations which are displayed in Figure 16. The model assumes that PPP deviations are driven entirely by r_t^{ppp}. From Figure 15, we observe that indeed the fluctuations in PPP deviations are driven entirely by r_t^{ppp}.

6. Conclusion

In this paper, we use the asset pricing approach proposed by Engel and West (2005) to quantify the contribution of monetary fundamentals to exchange rate movements. We show that using information on just directly observed monetary fundamentals, $m_t - m_t^* - (y_t - y_t^*)$, and exchange rates is plagued by weak identification of the expected future fundamentals. We solve the problem by using the restrictions implied by a simple rational expectations version of the monetary model of exchanges rates, additional data on interest rate and price differentials, and prior information about key parameters in the model. Adding this additional information results in sharper inference about the relative contribution of various fundamentals. We find that directly observed monetary fundamentals and money demand shifters contribute most to movements in exchange rates and movements in interest rate differentials to a lesser degree. The results suggest that monetary fundamentals, as defined in this paper, appear to explain long-run movements in exchange rates (consistent with the monetary approach to exchange rate determination) while risk premium associated with deviations from uncovered interest parity explains some of the short-run movements in exchange rates.

Table 1. Posterior Model Probabilities for Competing Models

Model	Posterior probability
Benchmark – AR(2), AR(2)	0.38
Model 2 – AR(4), AR(0)	0.09
Model 3 – AR(3), AR(1)	0.11
Model 4 – AR(1), AR(3)	0.15
Model 5 – AR(0), AR(4)	0.27

Note. The first AR(p) refers to the autoregressive process for $\phi_g(L)L$ while the second AR(p) refers to the autoregressive process for $\phi_\mu(L)L$.

References

Balke, N. S., and M. E. Wohar. 2002. Low Frequency Movements in Stock Prices: A State-Space Decomposition. *Review of Economics and Statistics* 84: 649-667.

Bilson, John F. O. 1978. "The Monetary Approach to the Exchange Rate: Some Empirical Evidence." *IMF Staff Papers* 25 (March): 48–75.

Campbell J. and R. Shiller, 1988, The dividend-price ratio and expectations of future dividends and discount factors, *Review of Financial Studies*, 1, 195-228.

Cerra, V. and S. C. Saxena, 2010, "The Monetary Model Strikes Back: Evidence from the World," *Journal of International Economics*, 81, 184-196.

Cheung, Yin-Wong, Menzie D. Chinn, and Antonio Garcia Pascual, 2005, "Empirical Exchange Rate Models of the Nineties: Are Any Fit to Survive?," *Journal of International Money and Finance*, Vol. 24, No. 7 (November): pp. 1150-1175.

Cochrane J., 2008, "State-space vs. VAR models for stock returns", unpublished working paper, University of Chicago.

Engel, Charles, and Kenneth D. West, 2005, "Exchange Rates and Fundamentals," *Journal of Political Economy*, Vol. 113, No. 3 (June): 485-517.

Engel, Charles, Nelson C. Mark, and Kenneth D. West, 2007, "Excahge Rate Models Are Not as Bad as You Think." In NBER Macroeconomics Annual 2007, edited by Acemoglu, Daron, Kenneth Rogoff, and Michael Woodford, pp. 381 – 411. Cambridge, MA: MIT Press.

Engel, C. 1996, "The Forward Discount Anomaly and the Risk Premium: A Survey of Recent Evidence," *Journal of Empirical Finance*, 3: 123–191.

Frankel, Jeffrey A. 1979. "On the Mark: A Theory of Floating Exchange Rates Based on Real Interest Differentials." *American Economic Review*. 69 (September): 610–22.

Granger, Clive W.J. and Paul Newbold, 1986, Forecasting Economic Time Series, Second Edition, Orlando: Academic Press.

Haug, Alfred A. and Julie Tam, 2007, "A Closer Look at Long-run U.S. Money Demand: Linear or Nonlinear Error-correction with M0, M1, OR M2?" Economic Inquiry, 45, 2 (April) 363-376.

Ma, Jun and Charles R. Nelson, 2010, "Valid Inference for a Class of Models Where Standard Inference Performs Poorly; Including Nonlinear Regression, ARMA, GARCH, and

Unobserved Components", working paper, Economic Series 256, Institute for Advanced Studies, Vienna. http://www.ihs.ac.at/publications/eco/es-256.pdf.

Ma, Jun, and Mark E. Wohar, 2011, "How Much Do Expected Returns and Expected Dividend Growth Contribute To Movements in Stock Returns? Issues of Weak Identification Make Existing Estimates Unreliable," working paper.

Morley, James, Charles R. Nelson, and Eric Zivot, 2003, Why Are Unobserved Component and Beveridge-Nelson Trend-Cycle Decompositions of GDP Are So Different?, *The Review of Economics and Statistics*, 85, 235-243.

Mark, Nelson C., 1995, "Exchange Rates and Fundamentals: Evidence on Long-Horizon Predictability." *American Economic Review*, Vol. 85, No.1 (March): pp. 201-218.

Mark, N. C., D. Sul, 2001, "Nominal exchange rates and monetary fundamentals: evidence from a small post-Bretton Woods panel," *Journal of International Economics* 53, 29-52.

Mark, Nelson C., 2001, International Macroeconomics and Finance, Blackwell Publishing.

Meese, Richard A., and Kenneth Rogoff, 1983a, "Empirical Exchange Rate Models of the Seventies: Do They Fit Out of Sample?" *Journal of International Economics*, Vol. 14, No. 1-2 (February): pp. 3-24.

Meese, Richard A., and Kenneth Rogoff, 1983b, "The Out-of-Sample Empirical Failure of Exchange Rate Models: Sampling Error or Misspecification," In Exchange Rates and International Macroeconomics, Ed. Jacob A. Frankel, 67-105. Chicago, Chicago University Press

Nelson, Charles R., and Richard Startz, 2007, The Zero-Information-Limit Condition and Spurious Inference in Weakly Identified Models, *Journal of Econometrics*, 138, 47-62.

Rapach, D. and M. Wohar, (2002) "Testing the Monetary Model of Exchange Rate Determination: New Evidence From A Century of data," *Journal of International* Economics 58 (December): 359-385.

Rogoff, K., 1996, "The Purchasing Power Parity Puzzle," *Journal of Economic Literature*, 34, pp. 647-668.

Stock, J. H., and M. W. Watson, 1993. "A Simple Estimator of Cointegrating Vectors in Higher Order Integrated Systems." *Econometrica* 61 (July):783–820.

Sarno, Lucio, 2005, "Viewpoint: Towards a Solution to the Puzzles in Exchange Rate Economics: Where Do We Stand?" *Canadian Journal of Economics*, Vol. 38, No .3 (August): pp. 673-708.

Sarno, Lucio, and Elvira Sojli, 2009, "The Feeble Link between Exchange Rates and Fundamentals: Can We Blame the Discount Factor?" *Journal of Money, Credit and Banking*, Vol. 41, No. 2-3 (March – April): pp. 437 – 442.

Appendix A.

We use the AR(2) specification for the latent expectations to illustrate the identification. Specifically we can derive:

$$E_t \left[\sum_{j=1}^{\infty} \psi^j \cdot \Delta f_{t+j} \right] = \begin{bmatrix} 1 & 0 \end{bmatrix} \cdot \psi \cdot \left(I - \psi F_g \right)^{-1} \cdot \begin{bmatrix} g_t \\ g_{t-1} \end{bmatrix} \tag{A.1}$$

$$E_t \left[\sum_{j=0}^{\infty} \psi^{j+1} \cdot r_{t+j} \right] = \psi \cdot \mu_{t-1} + \psi \cdot \varepsilon_t^r + \begin{bmatrix} 1 & 0 \end{bmatrix} \cdot \psi^2 \cdot \left(I - \psi F_\mu \right)^{-1} \cdot \begin{bmatrix} \mu_t \\ \mu_{t-1} \end{bmatrix} \tag{A.2}$$

Where, $F_g = \begin{bmatrix} \phi_{g1} & \phi_{g2} \\ 1 & 0 \end{bmatrix}$, $F_\mu = \begin{bmatrix} \phi_{\mu 1} & \phi_{\mu 2} \\ 1 & 0 \end{bmatrix}$; (A.1) corresponds to the contribution of expected

future fundamentals to current deviation of exchange rate from its fundamentals; (A.2) denotes the contribution of expected future remainder to current deviation of exchange rate from its fundamentals .

Assume that the 1 by 2 row vectors $\begin{bmatrix} 1 & 0 \end{bmatrix} \cdot \psi \cdot \left(I - \psi F_g \right)^{-1} = \begin{bmatrix} B_{11} & B_{12} \end{bmatrix}$, and

$\begin{bmatrix} 1 & 0 \end{bmatrix} \cdot \psi^2 \cdot \left(I - \psi F_\mu \right)^{-1} = \begin{bmatrix} B_{21} & B_{22} \end{bmatrix}$. Then, we can set up the following state-space model for the exchange rate model:

Measurement Equations:

$$\begin{bmatrix} \Delta f_{t+1} \\ d_{t+1} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 & 1 & 0 \\ B_{11} & B_{12} & B_{21} & B_{22} + \psi & 0 & \psi \end{bmatrix} \begin{bmatrix} g_{t+1} \\ g_t \\ \mu_{t+1} \\ \mu_t \\ \varepsilon_{t+1}^f \\ \varepsilon_{t+1}^r \end{bmatrix} \tag{A.3}$$

Transition Equations:

$$
\begin{bmatrix} g_{t+1} \\ g_t \\ \mu_{t+1} \\ \mu_t \\ \varepsilon^f_{t+1} \\ \varepsilon^r_{t+1} \end{bmatrix} = \begin{bmatrix} \phi_{g1} & \phi_{g2} & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \phi_{\mu1} & \phi_{\mu2} & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} g_t \\ g_{t-1} \\ \mu_t \\ \mu_{t-1} \\ \varepsilon^f_t \\ \varepsilon^r_t \end{bmatrix} + \begin{bmatrix} \varepsilon^g_{t+1} \\ 0 \\ \varepsilon^\mu_{t+1} \\ 0 \\ \varepsilon^f_{t+1} \\ \varepsilon^r_{t+1} \end{bmatrix}
\tag{A.4}
$$

Where, the variance-covariance matrix of the vector of four shocks

$V_{t+1} = \begin{bmatrix} \varepsilon^g_{t+1} & \varepsilon^\mu_{t+1} & \varepsilon^f_{t+1} & \varepsilon^r_{t+1} \end{bmatrix}$ is:

$$
\Omega = Var(V_{t+1}) = \begin{bmatrix} \sigma^2_g & - & - & - \\ \sigma_{g\mu} & \sigma^2_\mu & - & - \\ \sigma_{gf} & \sigma_{\mu f} & \sigma^2_f & - \\ \sigma_{gr} & \sigma_{\mu r} & \sigma_{fr} & \sigma^2_r \end{bmatrix}
$$

Because of the economic constraint (2.6) we only need to explicitly model d and Δf, and then r can be backed out as a residual. This appendix shows that the above state-space model implies a general VARMA reduced-form representation for $\begin{bmatrix} \Delta f_{t+1} & d_{t+1} \end{bmatrix}'$ and derives the specific mapping between the structural and reduced-form model. Following Morley, Nelson, and Zivot (2003), we show in the Appendix that the above structural model is identified.

This appendix presents the mapping between the state-space model and its reduced-form VARMA representation, and discusses relevant identification issues. Plug the transition equation into the measurement equation to obtain:

$$
\begin{pmatrix} \phi_g(L) & 0 \\ 0 & \phi_g(L)\phi_\mu(L) \end{pmatrix} \begin{pmatrix} \Delta f_{t+1} \\ d_{t+1} \end{pmatrix} = \begin{pmatrix} L & 0 & \phi_g(L) & 0 \\ (B_{11}+B_{12}L)\phi_\mu(L) & (B_{21}+(B_{22}+\psi)L)\phi_g(L) & 0 & \phi_g(L)\phi_\mu(L) \end{pmatrix} \begin{pmatrix} \varepsilon^g_{t+1} \\ \varepsilon^\mu_{t+1} \\ \varepsilon^f_{t+1} \\ \varepsilon^r_{t+1} \end{pmatrix}
\tag{A.5}
$$

Where, $\phi_g(L) = (1 - \phi_{g1}L - \phi_{g2}L^2)$ and $\phi_\mu(L) = (1 - \phi_{\mu1}L - \phi_{\mu2}L^2)$. Denote the LHS of (A.5) by:

$$\begin{pmatrix} x_{1t+1} \\ x_{2t+1} \end{pmatrix} = \begin{pmatrix} \phi_g(L) & 0 \\ 0 & \phi_g(L)\phi_\mu(L) \end{pmatrix} \begin{pmatrix} \Delta f_{t+1} \\ d_{t+1} \end{pmatrix}$$

Then:

$$\begin{pmatrix} x_{1t+1} \\ x_{2t+1} \end{pmatrix} = \left(C + DL + EL^2 + FL^3 + GL^4 \right) \cdot V_{t+1} \tag{A.6}$$

Where, $V_{t+1} = \begin{pmatrix} \varepsilon^g_{t+1} & \varepsilon^\mu_{t+1} & \varepsilon^f_{t+1} & \varepsilon^r_{t+1} \end{pmatrix}$, and its variance is Ω.

$$C = \begin{pmatrix} 0 & 0 & 1 & 0 \\ B_{11} & B_{12} & 0 & 1 \end{pmatrix},$$

$$D = \begin{pmatrix} 1 & 0 & -\phi_{g1} & 0 \\ -B_{11}\phi_{\mu1} + B_{12} & -B_{21}\phi_{g1} + (B_{22} + \psi) & 0 & -\phi_{\mu1} - \phi_{g1} \end{pmatrix},$$

$$E = \begin{pmatrix} 0 & 0 & -\phi_{g2} & 0 \\ -B_{11}\phi_{\mu2} - B_{12}\phi_{\mu1} & -B_{21}\phi_{g2} - (B_{22} + \psi)\phi_{g1} & 0 & -\phi_{\mu2} - \phi_{g2} + \phi_{\mu1}\phi_{g1} \end{pmatrix},$$

$$F = \begin{pmatrix} 0 & 0 & 0 & 0 \\ -B_{12}\phi_{\mu2} & -(B_{22} + \psi)\phi_{g2} & 0 & \phi_{\mu1}\phi_{g2} + \phi_{\mu2}\phi_{g1} \end{pmatrix},$$

$$G = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \phi_{\mu2}\phi_{g2} \end{pmatrix}.$$

From the eq. (A.2), we can derive the second moments of its RHS:

$$Var\begin{pmatrix} x_{1t+1} \\ x_{2t+1} \end{pmatrix} = C\Omega C' + D\Omega D' + E\Omega E' + F\Omega F' + G\Omega G' \tag{A.7}$$

$$Cov\left(\begin{pmatrix} x_{1t+1} \\ x_{2t+1} \end{pmatrix}, \begin{pmatrix} x_{1t} \\ x_{2t} \end{pmatrix} \right) = D\Omega C' + E\Omega D' + F\Omega E' + G\Omega F' \tag{A.8}$$

$$Cov\left(\begin{pmatrix} x_{1t+1} \\ x_{2t+1} \end{pmatrix}, \begin{pmatrix} x_{1t-1} \\ x_{2t-1} \end{pmatrix} \right) = E\Omega C' + F\Omega D' + G\Omega E' \tag{A.9}$$

$$Cov\left(\begin{pmatrix} x_{1t+1} \\ x_{2t+1} \end{pmatrix}, \begin{pmatrix} x_{1t-2} \\ x_{2t-2} \end{pmatrix}\right) = F\Omega C' + G\Omega D' \qquad (A.10)$$

$$Cov\left(\begin{pmatrix} x_{1t+1} \\ x_{2t+1} \end{pmatrix}, \begin{pmatrix} x_{1t-3} \\ x_{2t-3} \end{pmatrix}\right) = G\Omega C \qquad (A.11)$$

Therefore, by Granger and Newbold's Theorem (1986), the structure of the second moments

implies that the $\begin{pmatrix} x_{1t+1} \\ x_{2t+1} \end{pmatrix}$ has a reduced-form VMA(4) process. The AR parameters of

$\phi_g(L) = \left(1 - \phi_{g1}L - \phi_{g2}L^2\right)$ and $\phi_\mu(L) = \left(1 - \phi_{\mu 1}L - \phi_{\mu 2}L^2\right)$ can be identified by the AR structure of

$\begin{pmatrix} \Delta f_{t+1} \\ d_{t+1} \end{pmatrix}$. The parameters left in the state-space model as set up in (A.3) and (A.4) are 10 variance

and covariance parameters and these parameters can be identified by the moving average terms

of the $\begin{pmatrix} x_{1t+1} \\ x_{2t+1} \end{pmatrix}$ as shown above. Therefore, the state-space model is identifiable.

Appendix B. Details of Markov Chain Monte Carlo.

Each of the three models has an empirical state space model of the form:

$$Y_t^{obs} = H(\theta)S_t,$$ (B1)

$$S_t = F(\theta)S_{t-1} + V_t, \quad V_t \sim MVN \ (0, Q(\theta))$$ (B2)

where Y_t^{obs} is the vector of observable time series, S_t is the vector of unobserved state variables,

and θ is the vector of structural parameters. The predictive log likelihood of the state space

model is given by:

$$
\begin{aligned}
log(L(\mathbf{Y}_T, \theta)) = \sum_{t=1}^{T} \{ &-.5 log(det(H(\theta)'P_{t|t-1}H(\theta))) \\
&-.5(Y_t - H(\theta)S_{t|t-1})'(H(\theta)'P_{t|t-1}H(\theta))^{-1}(Y_t - H(\theta)S_{t|t-1})\}
\end{aligned}
$$ (B3)

where $S_{t|t-1}$ and $P_{t|t-1}$ are the conditional mean and variance of S_t from the Kalman filter.

Given a prior distribution over parameters, $h(\theta)$, the posterior distribution, $P(\theta \mid \mathbf{Y_T})$, is

$$P(\theta \mid \mathbf{Y_T}) \propto L(\mathbf{Y_T}, \theta)h(\theta).$$ (B4)

Because the log-likelihood is a nonlinear function of the structural parameter vector, it is not

possible to write an analytical expression for the posterior distribution. As a result, we use

Bayesian Markov Chain Monte Carlo methods to estimate the posterior distribution of the

parameter vector, θ. In particular, we employ a Metropolis-Hasting sampler to generate draws

from the posterior distributions. The algorithm is as follows:

(i)   Given a previous draw of the parameter vector, $\theta^{(i-1)}$, draw a candidate vector θ^c from the

distribution $g(\theta \mid \theta^{(i-1)})$.

(ii) Determine the acceptance probability for the candidate draw,

$$\alpha(\theta^c,\theta^{(i-1)}) = min\left[\frac{L(\mathbf{Y_T},\theta^{(c)})h(\theta^{(c)})}{L(\mathbf{Y_T},\theta^{(i-1)})h(\theta^{(i-1)})}\frac{g(\theta^{(i-1)}\,|\,\theta^c)}{g(\theta^c\,|\,\theta^{(i-1)})},1\right].$$

(iii) Determine a new draw from the posterior distribution, $\theta^{(i)}$.

$\theta^{(i)} = \theta^c$ with probability $\alpha(\theta^c,\theta^{(i-1)})$

$\theta^{(i)} = \theta^{(i-1)}$ with probability $1 - \alpha(\theta^c,\theta^{(i-1)})$.

(iv) Return to (i).

Starting from an initial parameter vector and repeating enough times, the distribution parameters draws, $\theta^{(i)}$, will converge to the true posterior distribution.

In our application, $\theta^c = \theta^{(i-1)} + v$, where v is drawn from a multivariate t-distribution with 50 degrees of freedom and a covariance matrix Σ. We set Σ to be a scaled value of the Hessian matrix of $-log(L(\mathbf{Y_T},\theta))$ evaluated at the maximum likelihood estimates. We choose the scaling so that around 50 percent of the candidate draws are accepted. We set a burn-up period of 500,000 draws and then sampled the next 500,000 draws. We can also obtain the posterior distributions for the unobserved states. Given a parameter draw, we draw from the conditional posterior distribution for the unobserved states, $P(\mathbf{S_T}\,|\,\Theta^{(i)},\mathbf{Y_T})$. Here we use the "filter forward, sample backwards" approach proposed by Carter and Kohn (1994) and discussed in Kim and Nelson (1999).

Appendix C. Details on Data Construction

 This appendix describes the sources of the data used in the text.

 The US/UK nominal exchange rate comes from Taylor (2001): specifically the pre-1948 data are from the statistical volumes of Brian Mitchell; and the series after 1948 are period average observations taken from the IMF's *International Financial Statistics* (*IFS*).

 UK real national income. Data series of $1880 - 1948$ are real GDP taken from Bordo et al. (1998) that are originally from Mitchell (1988); and data series after 1948 are real GDP taken from the *IFS*. When combining two series we adjust those observations of the earlier period by

using the formula: $y_t \times \dfrac{y_{1948, IFS}}{y_{1948, Bordo}}$ for $1880 \le t \le 1948$.

 US real national income. Data series of $1880 - 1948$ are real GNP taken from Bordo et al. (1998) that are originally from Balke and Gordon (1986); and data series after 1948 are real GDP taken from the *IFS*. When combining two series we adjust those observations of the earlier

period by using the formula: $y_t \times \dfrac{y_{1948, IFS}}{y_{1948, Bordo}}$ for $1880 \le t \le 1948$.

 UK money suppy. Data series of $1880 - 1966$ are net money supply (M2) taken from Bordo et al. (1998) that are originally from Sheppard (1986); and data series after 1966 are money plus quasi-money taken from the *IFS*. When combining two series we adjust those

observations of the earlier period by using the formula: $y_t \times \dfrac{y_{1966, IFS}}{y_{1966, Bordo}}$ for $1880 \le t \le 1966$.

 US money supply. Data series of $1880 - 1971$ are money supply (M2) taken from Bordo et al. (1998) that are originally from Balke and Gordon (1986); and data series after 1971 are money plus quasi-money taken from the *IFS*. When combining two series we adjust those

observations of the earlier period by using the formula: $y_t \times \dfrac{y_{1971, IFS}}{y_{1971, Bordo}}$ for $1880 \le t \le 1971$.

 UK price level. Data series of $1880 - 1988$ are from Rapach and Wohar (2002); and data series after 1988 are CPI taken from the *IFS*. When combining two series we adjust those

observations of the earlier period by using the formula: $y_t \times \dfrac{y_{1988, IFS}}{y_{1988, RW}}$ for $1880 \le t \le 1988$.

US price level. Data series of 1880 – 1948 are from Rapach and Wohar (2002); and data series after 1948 are CPI taken from the *IFS*. When combining two series we adjust those observations of the earlier period by using the formula: $y_t \times \dfrac{y_{1948,IFS}}{y_{1948,RW}}$ for $1880 \leq t \leq 1948$.

UK interest rate. Prime bank bill rates from NBER macro history database.

US interest rate. Commercial paper rates from NBER macro history database.

Figure 1. UK-US Exchange Rate and Observed Monetary Fundamentals

Log exchange rate (s) and observed level of monetary fundamentals (f)

Legend:
- log exchange rate (s)
- observed monetary fundamentals (f = m-m*-(y-y*))

s(t)-f (t) and f(t)-f(t-1)

Legend:
- s(t)-f(t)
- f(t)-f(t-1)
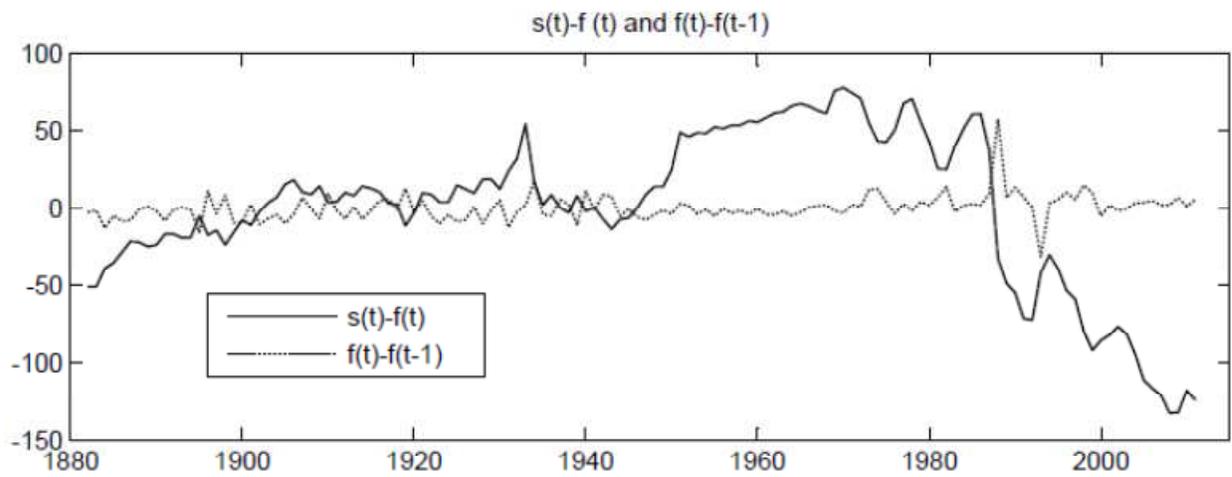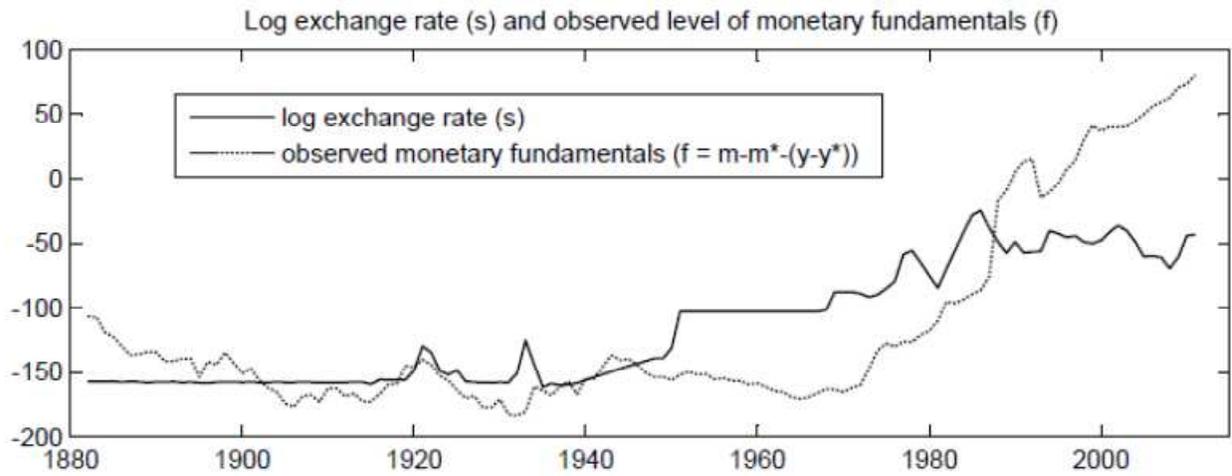
Figure 2. Histograms of the posterior distribution of parameters in Benchmark (AR(2), AR(2)) Model
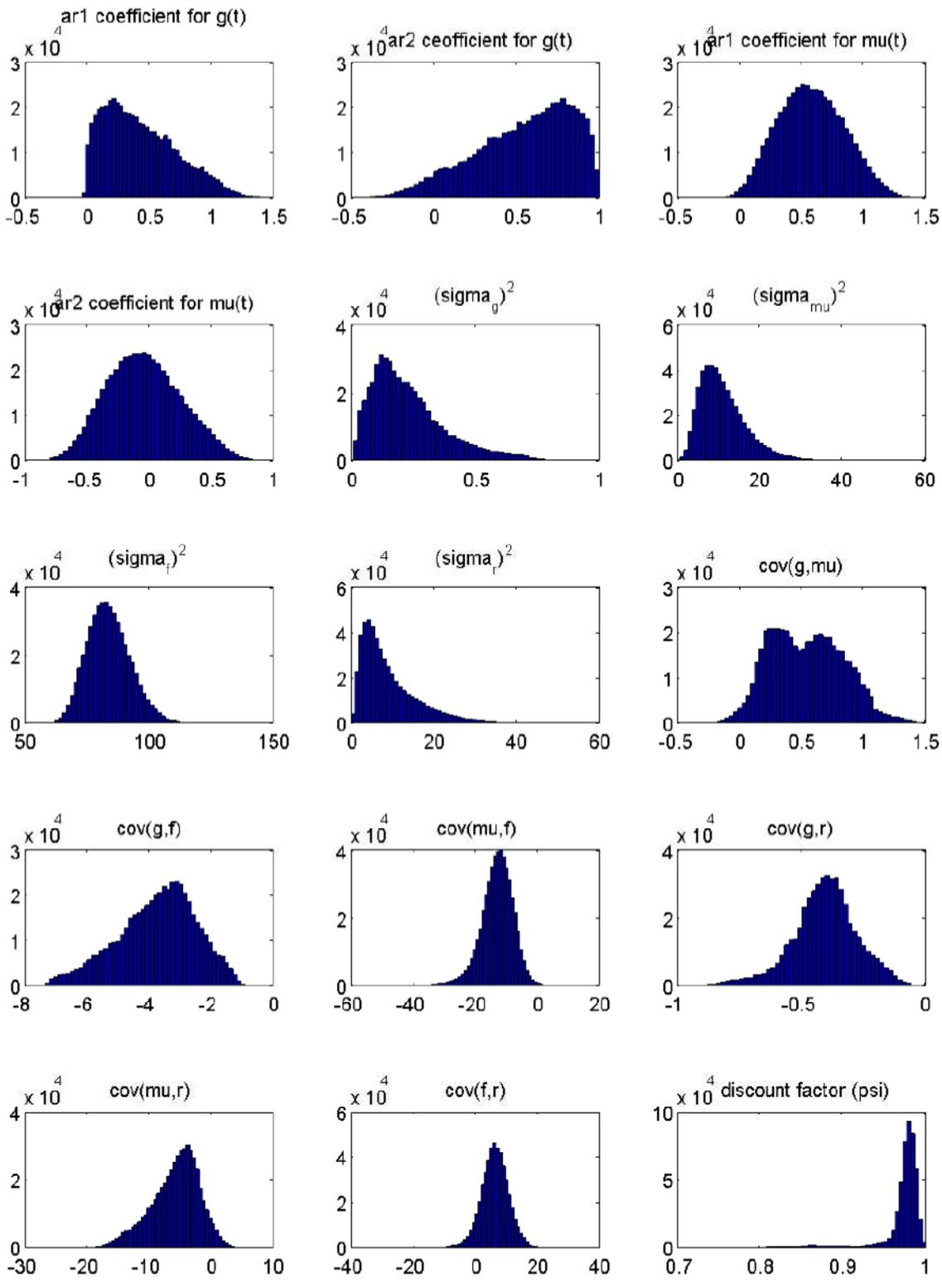
Figure 3. Histograms of joint posterior distribution of phi $_{g1}$ and phi $_{g2}$

Histogram of joint distribution of phi $_{g1}$ and phi $_{g2}$

Heat plot of histogram of joint distribution of phi $_{g1}$ and phi $_{g2}$
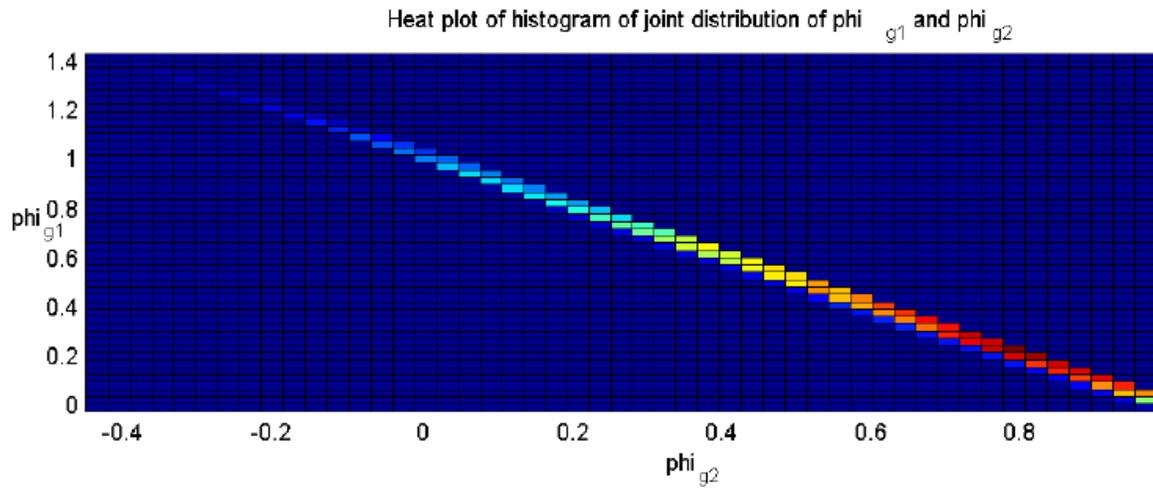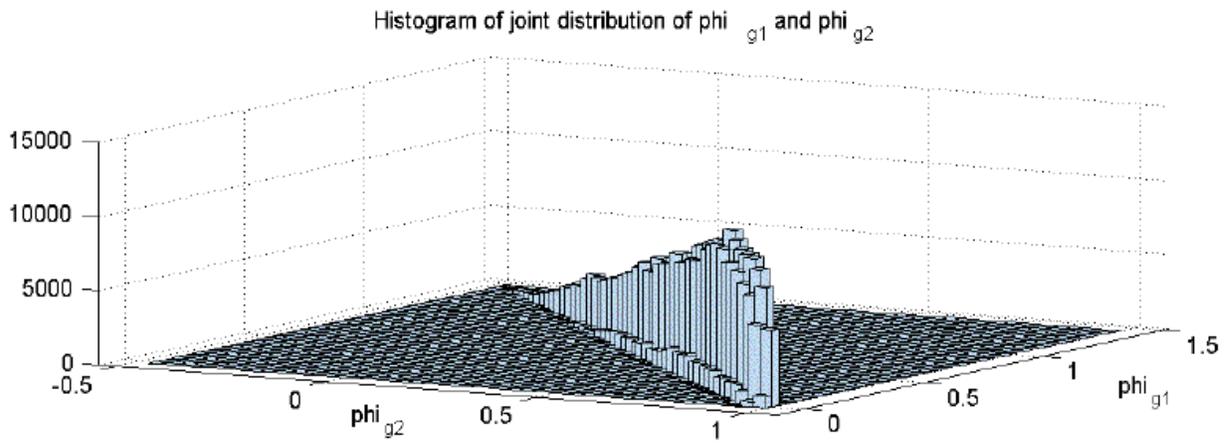
Figure 4. Histograms of Posterior Distribution of
Variance decomposition of deviations of exchange rate from observed monetary fundamental (s(t)-f(t))

Figure 5.

Cumulative Distribution of Log Likelihoods Implied by Posterior Distribution of Parameters
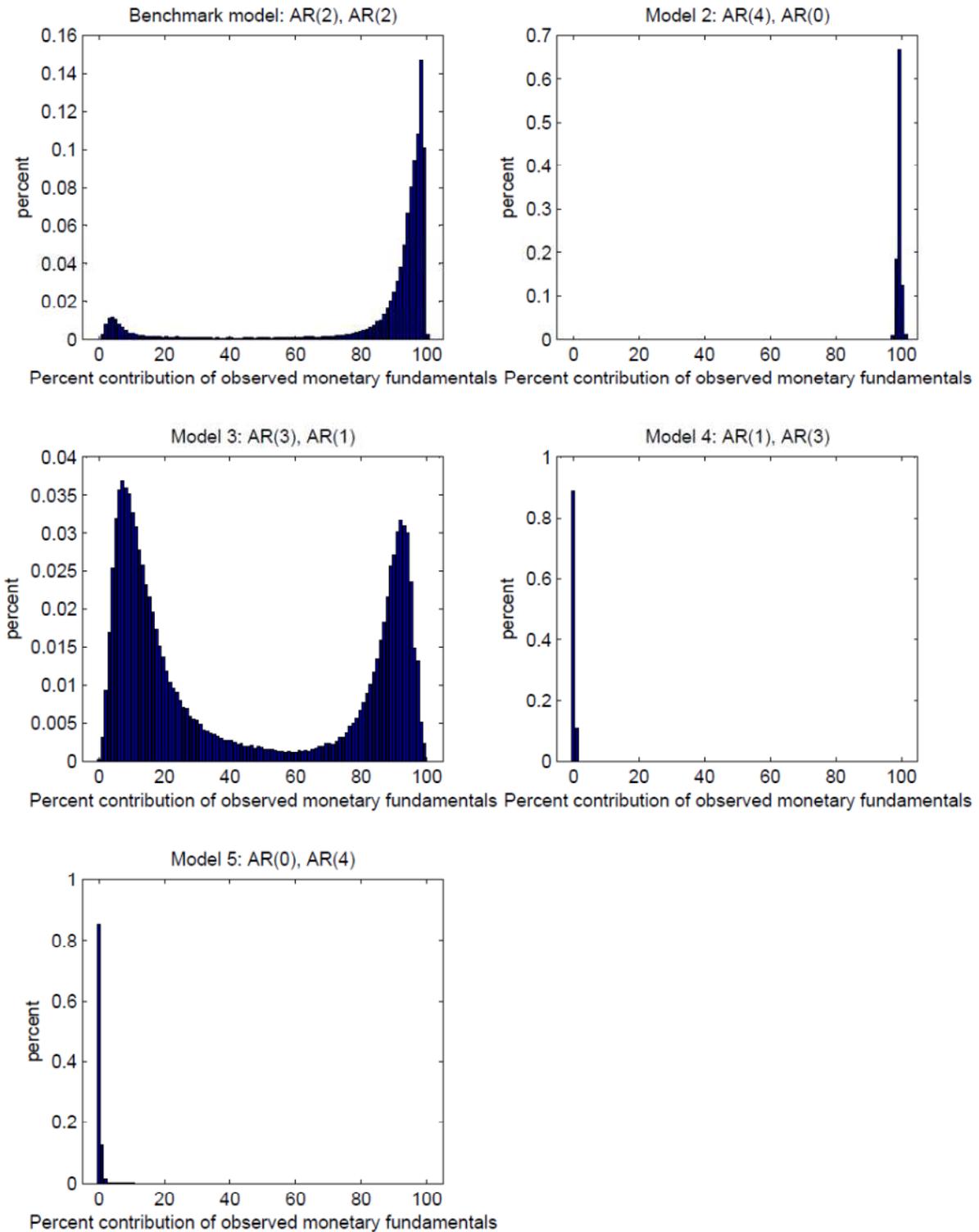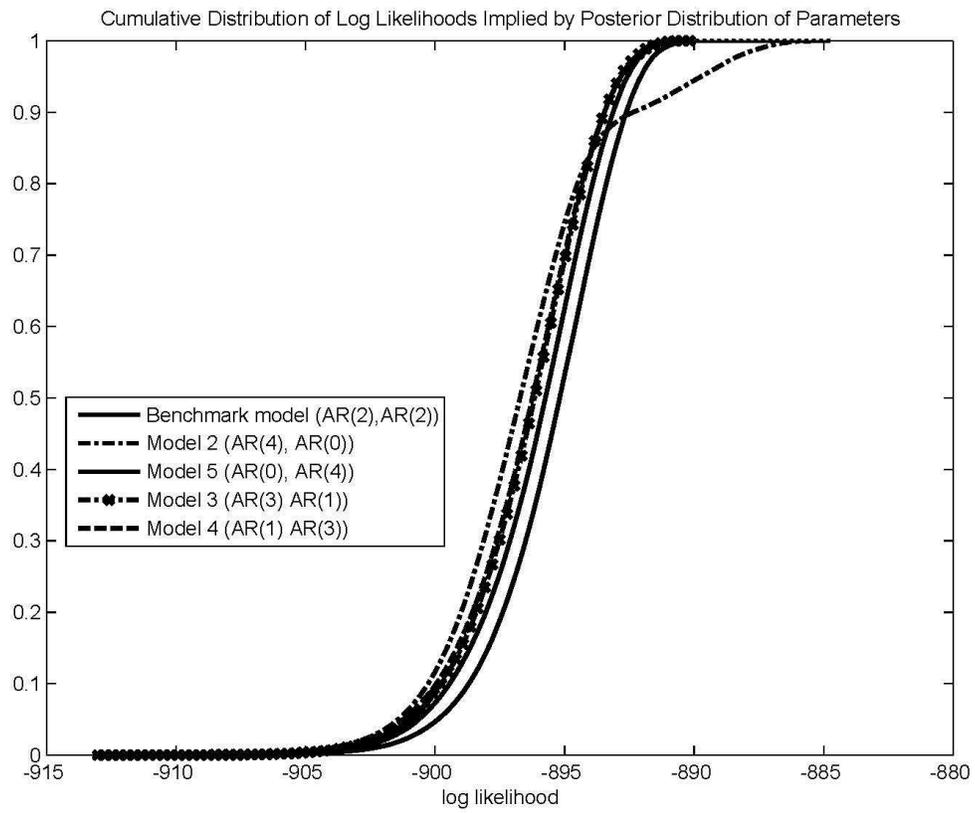
Figure 6. Posterior distribution of variance decomposition of s(t)-f(t)
Histogram based on Bayesian model averaging
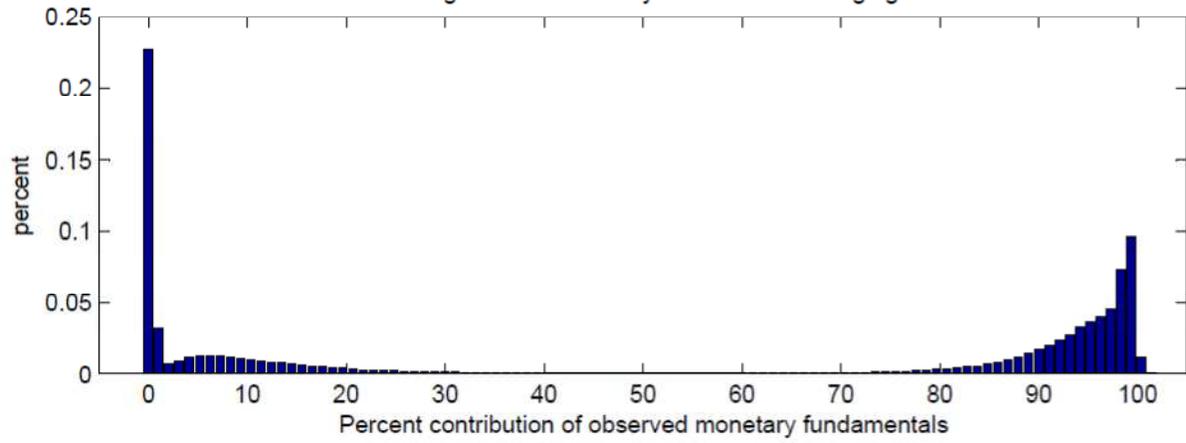
Figure 7. Prior distriutions of AR coefficient and half-life of PPP deviations
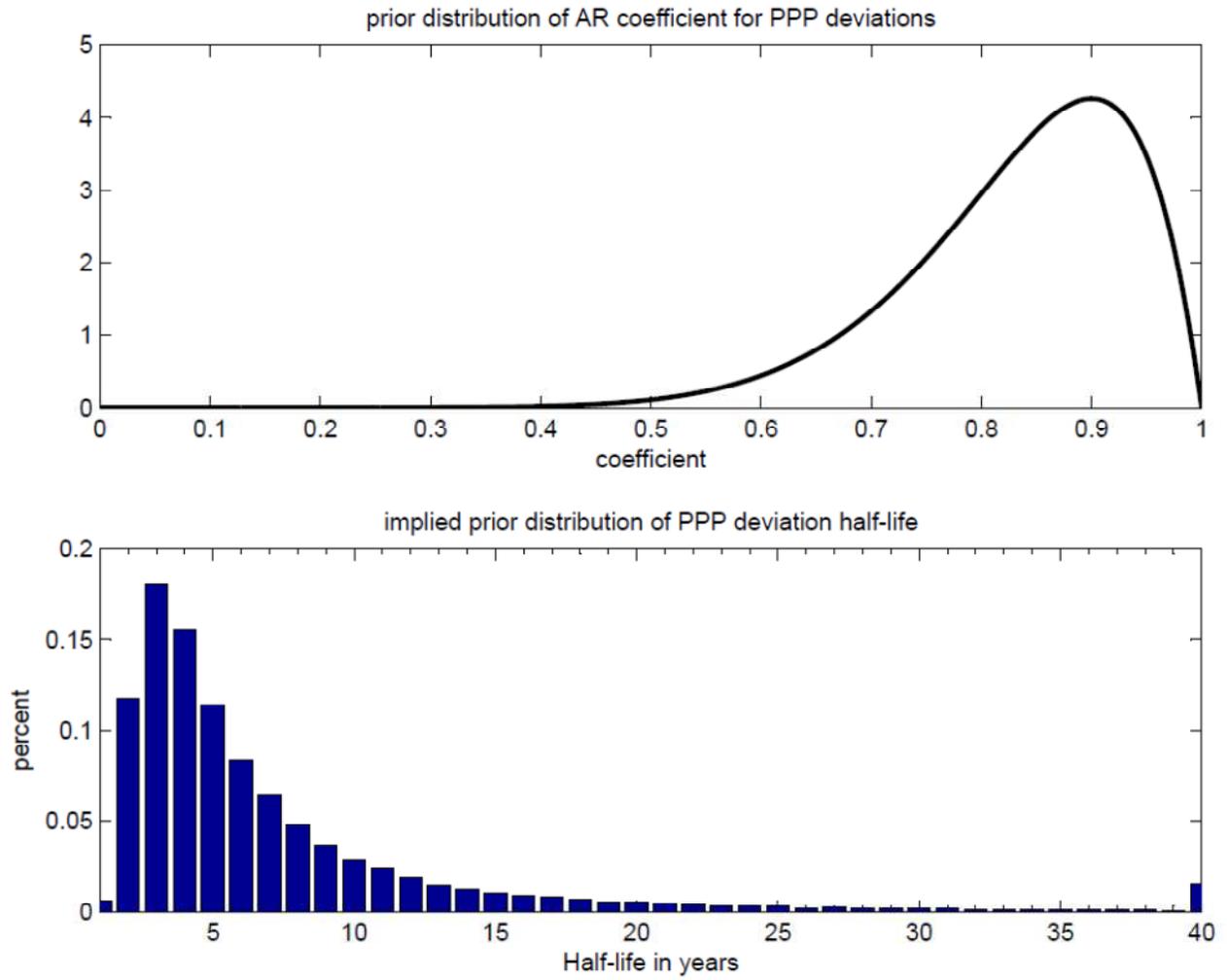
Figure 8. Prior distributions of semi-elasticity of money demand and psi
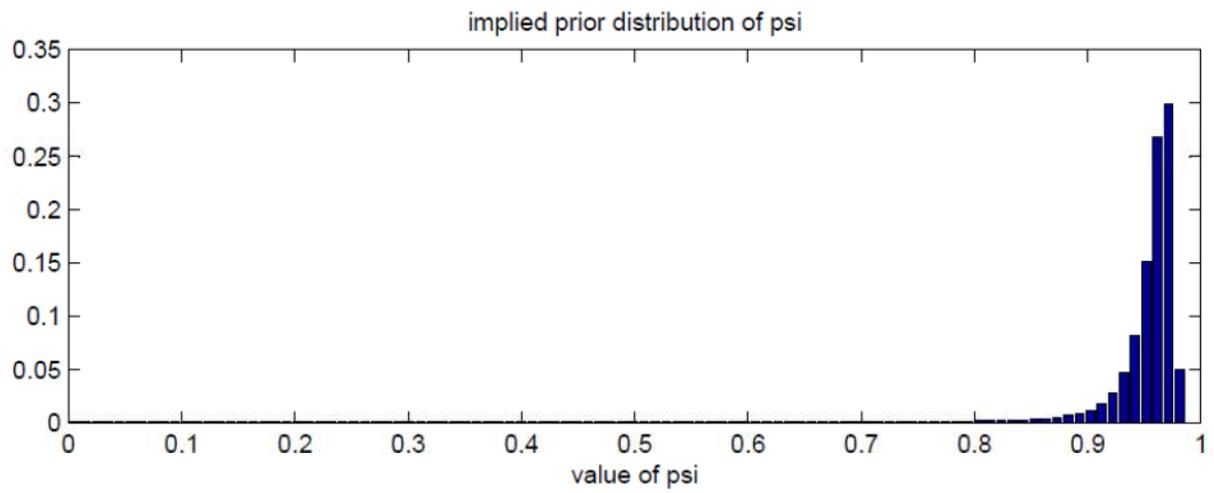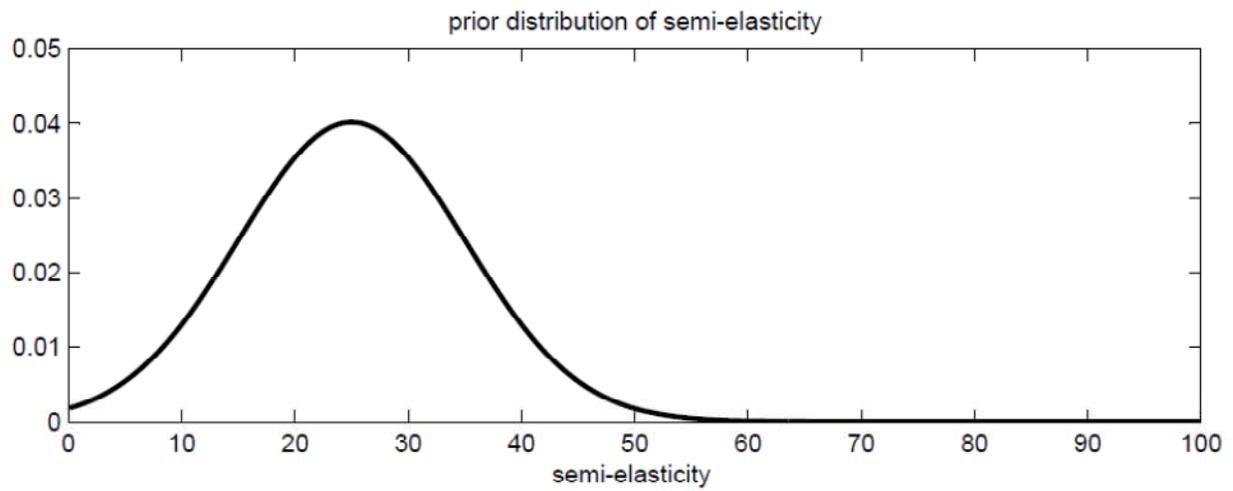
Figure 9. Posterior distribution of autoregressive parameters of the unobserve factors
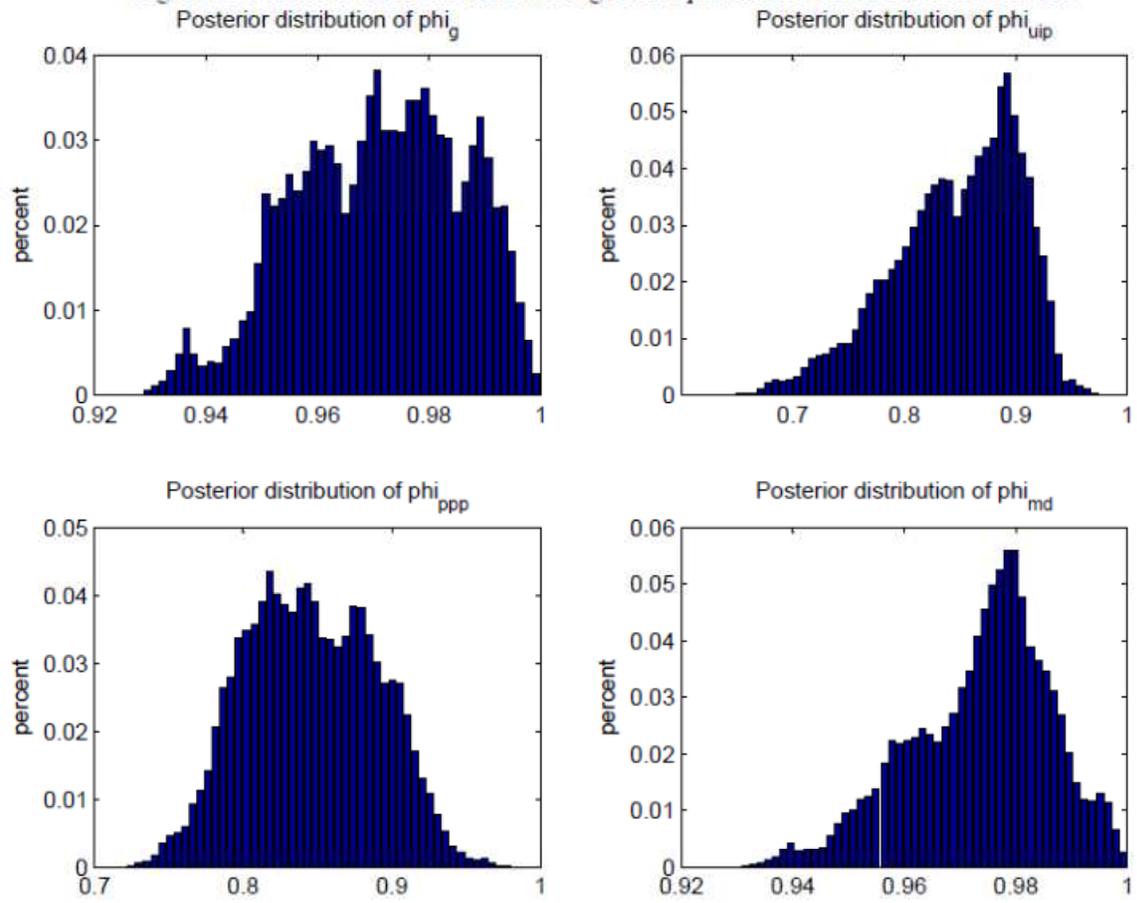
Figure 10. Posterior Distributions of interest rate semi-elasticity and discount factor
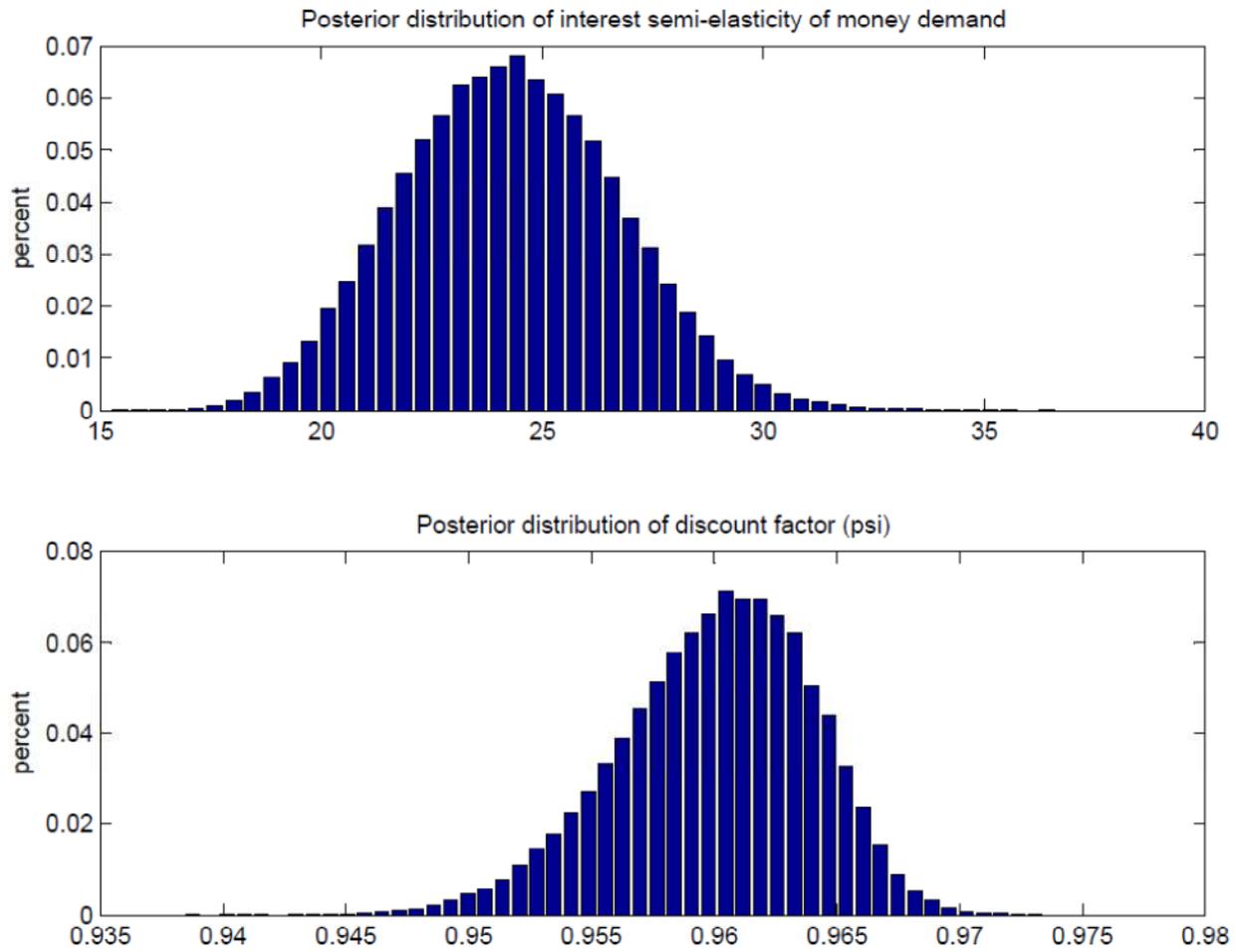
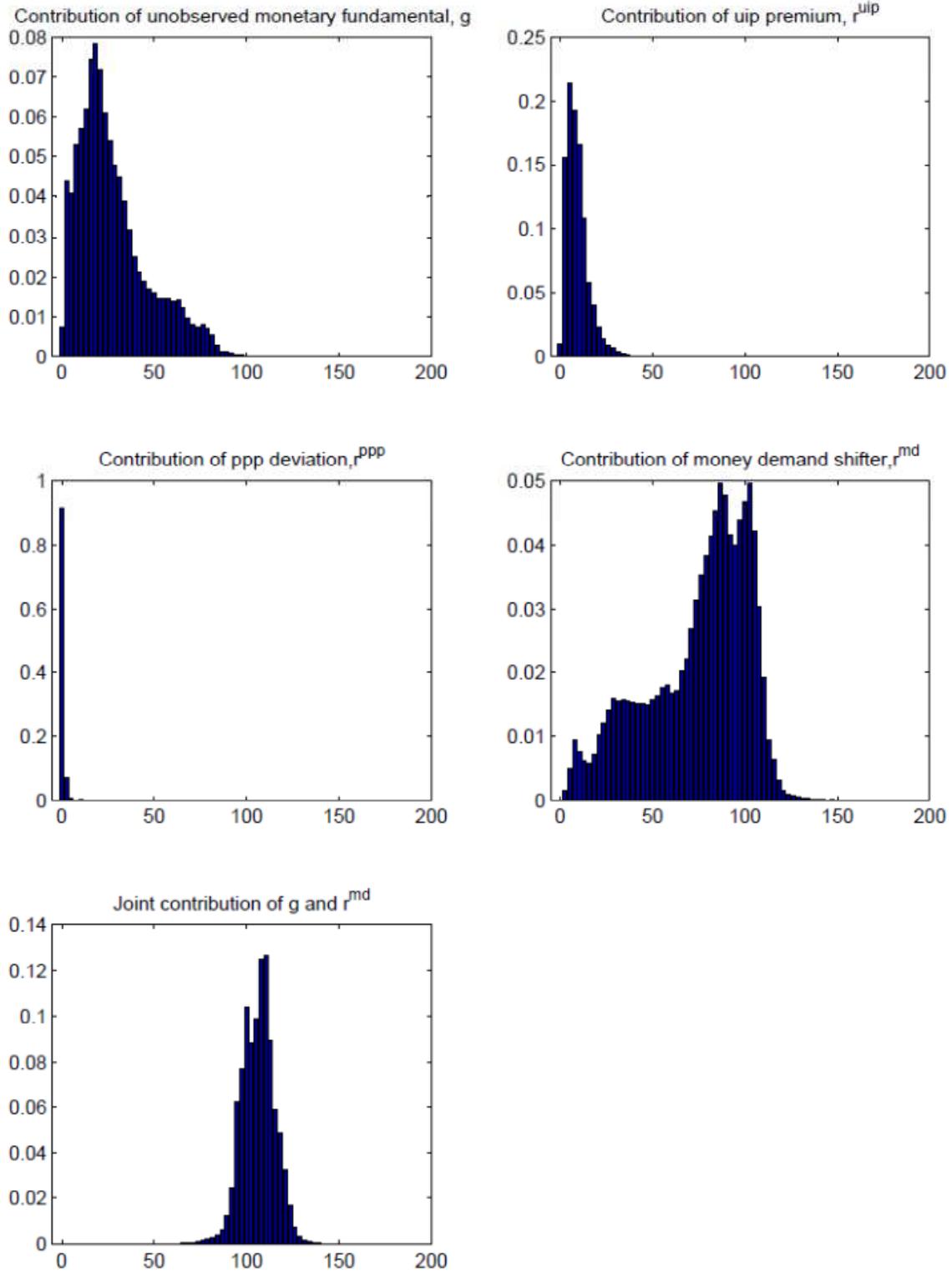Figure 11. Posterior distribution of variance decomposition of s(t)-f(t)

Figure 12. Historical Decomposition of UK-US Exchange Rate
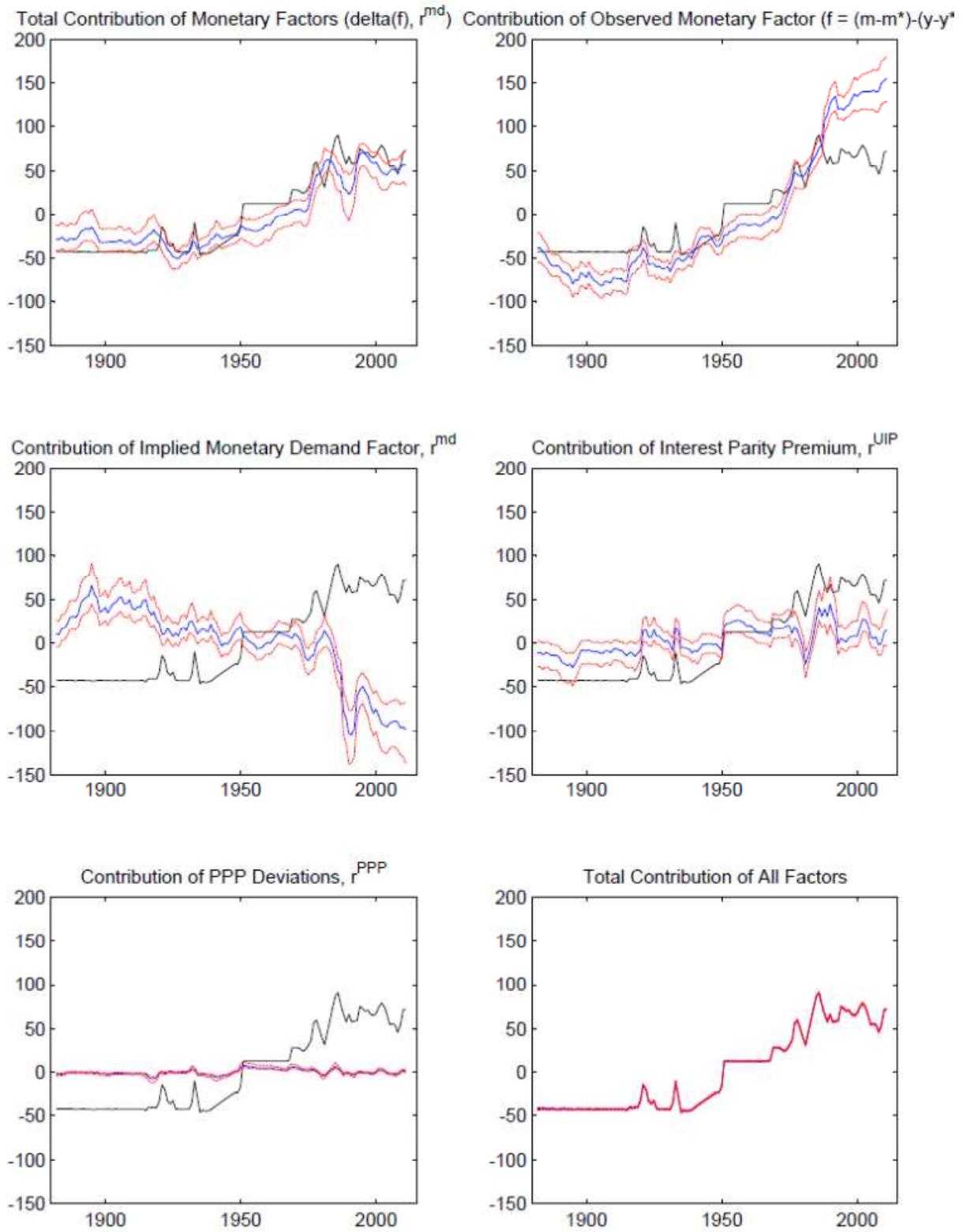### Median, 5th, and 95th percentiles of the posterior distribution

Total Contribution of Monetary Factors (delta(f), r^{md})

Contribution of Observed Monetary Factor ($f = (m-m^*)-(y-y^*$

Contribution of Implied Monetary Demand Factor, r^{md}

Contribution of Interest Parity Premium, r^{UIP}

Contribution of PPP Deviations, r^{PPP}

Total Contribution of All Factors

Figure 13. Historical Decomposition of UK-US Interest Rate Differential
Median, 5th, and 95th percentiles of the posterior distribution

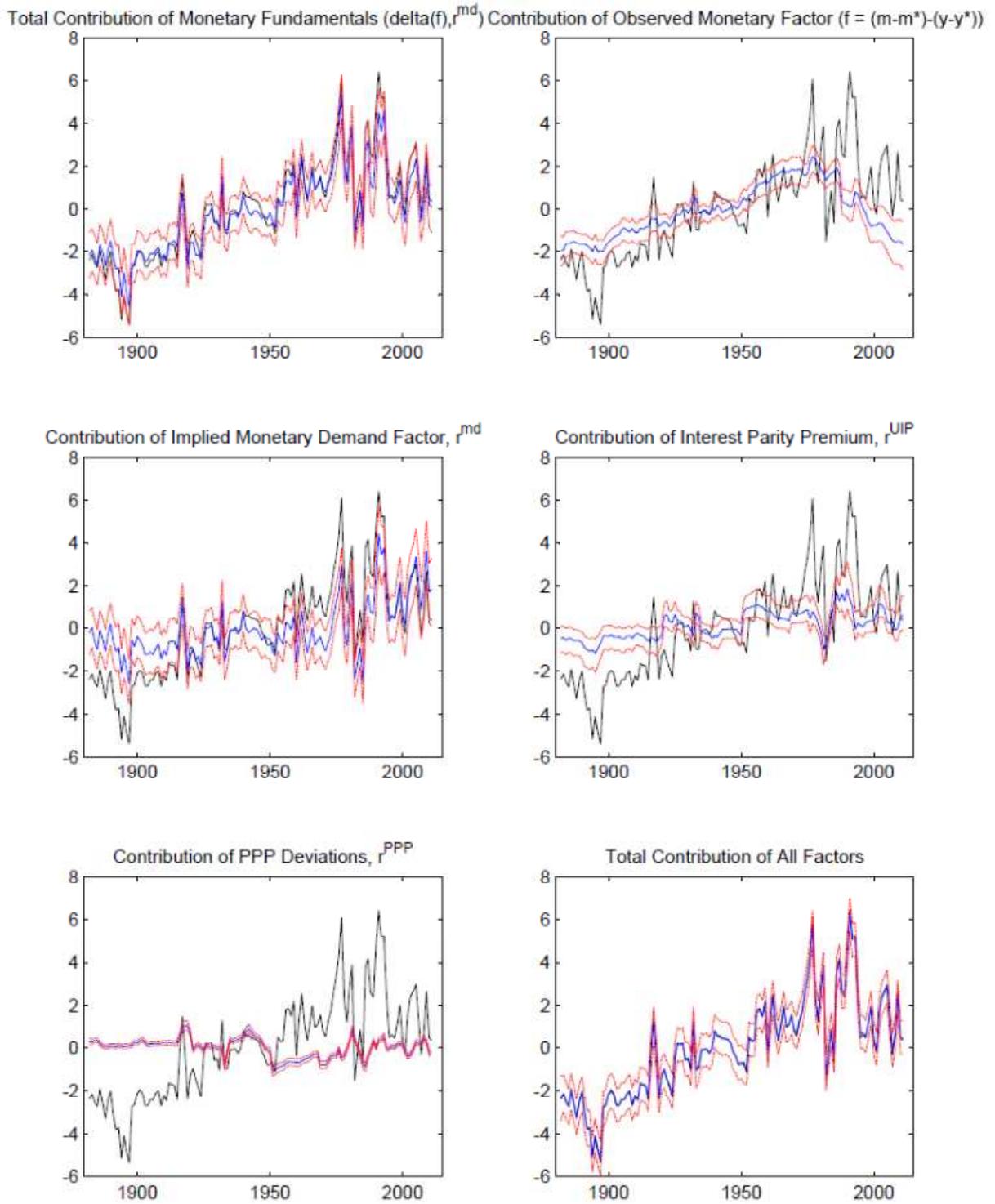Total Contribution of Monetary Fundamentals (delta(f),r^{md}) Contribution of Observed Monetary Factor (f = (m-m*)-(y-y*))

Contribution of Implied Monetary Demand Factor, r^{md}

Contribution of Interest Parity Premium, r^{UIP}

Contribution of PPP Deviations, r^{PPP}

Total Contribution of All Factors

Figure 14. Historical Decomposition of UK-US Money Demand (m-y-p) Differentials
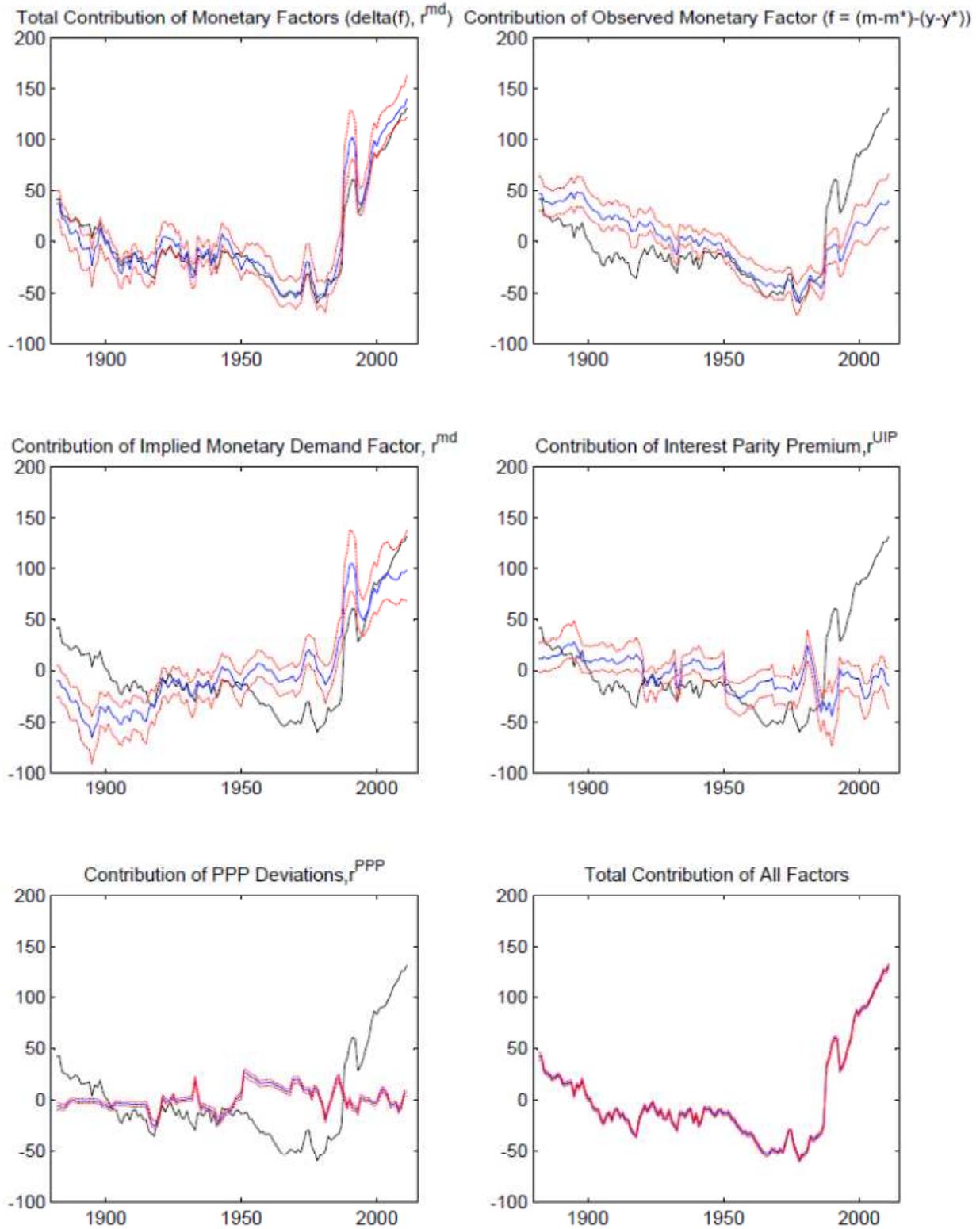Median, 5th, and 95th percentiles of the posterior distribution

Figure 15. Historical Decomposition of Observable Monetary Fundamental Growth (delta f)
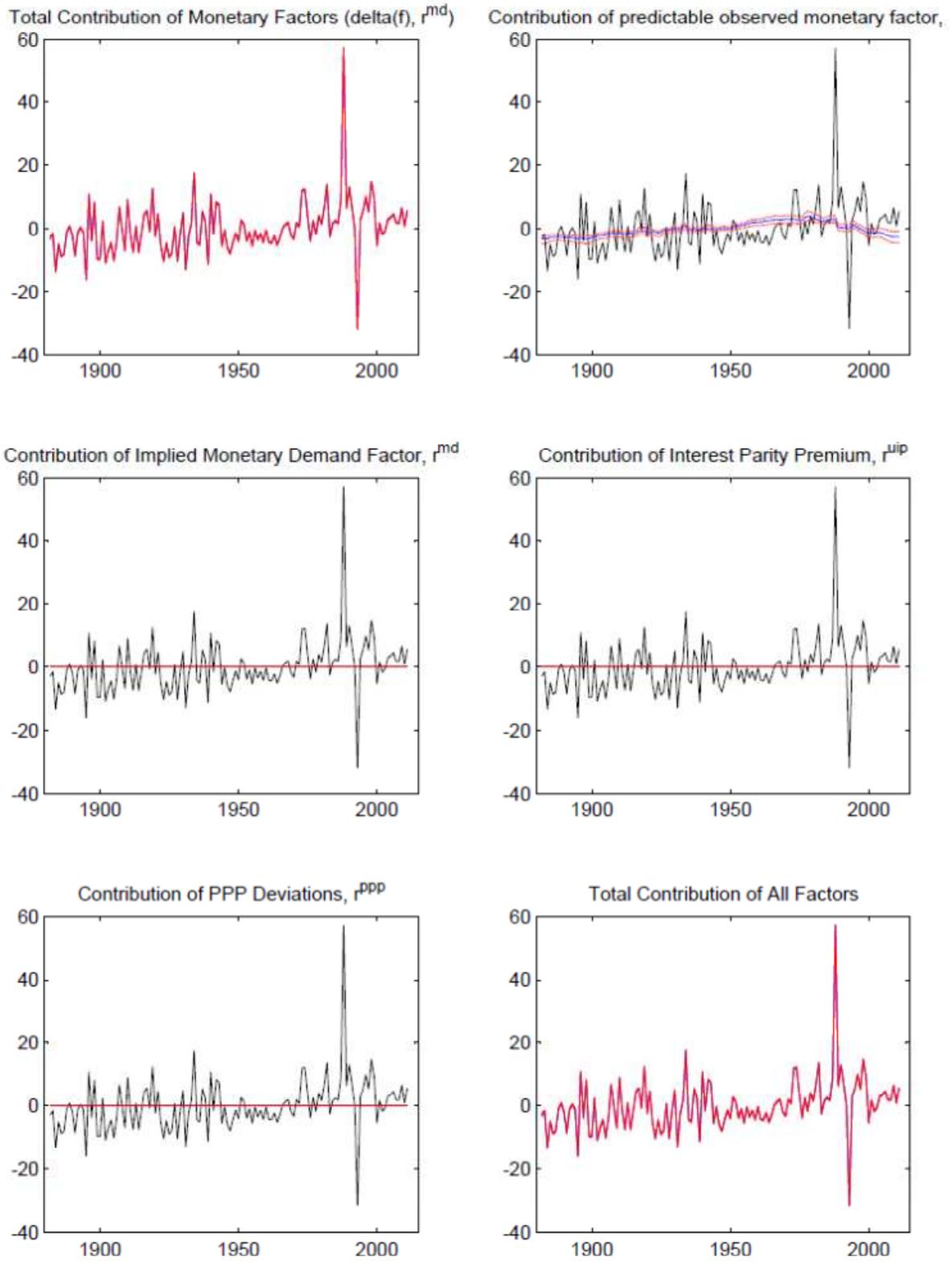Median, 5th, and 95th percentiles of the posterior distribution

Figure 16. Historical Decomposition of UK-US Deviations from Purchasing Power Parity
Median, 5th, and 95th percentiles of the posterior distribution