

# Monetary Policy Rules and Exchange Rates: A Structural VAR Identified by No Arbitrage\*

Sen Dong<sup>†</sup>

Columbia University and Lehman Brothers

This Draft: June 7, 2006

\*I thank Asim Ansari, Amitabh Arora, Jean Boivin, Mikhail Chernov, Sid Dastidar, Marc Giannoni, Po-Hsuan Hsu, Raghu Iyengar, Wei Jiang, Michael Johannes, Marco Naldi, Monika Piazzesi, Tano Santos, and seminar participants at Columbia University for helpful comments. I especially thank Andrew Ang, Geert Bekaert, and Robert Hodrick for detailed comments and encouragement. All remaining errors are my own.

<sup>†</sup>Columbia Business School, 311 Uris Hall, 3022 Broadway, New York, NY 10027; Phone: (646) 354 8611; Fax: (212) 662 8474; Email: sd2068@columbia.edu; WWW: <http://www.columbia.edu/~sd2068>.

## **Abstract**

Expected exchange rate changes are determined by interest rate differentials across countries and risk premia, while unexpected changes are driven by innovations to macroeconomic variables, which are amplified by time-varying market prices of risk. In a model where short rates respond to the output gap and inflation in each country, I identify macro and monetary policy risk premia by specifying no-arbitrage dynamics of each country's term structure of interest rates and the exchange rate. Estimating the model with US/German data, I find that the correlation between the model-implied exchange rate changes and the data is over 60%. The model implies a countercyclical foreign exchange risk premium with macro risk premia playing an important role in matching the deviations from Uncovered Interest Rate Parity. I find that the output gap and inflation drive about 70% of the variance of forecasting the conditional mean of exchange rate changes.

# 1 Introduction

This paper studies the role of macro variables in explaining the foreign exchange risk premium and the dynamics of exchange rates. One puzzling observation in foreign exchange markets is the tendency of high yield currencies to appreciate. This departure from Uncovered Interest Rate Parity (UIRP) implies a volatile foreign exchange risk premium, which many consumption and money-based general equilibrium models based on Lucas (1982) fail to generate (see, e.g., Hodrick, 1989; Backus, Gregory, and Telmer, 1993; Bansal et al., 1995; Bekaert, 1996).<sup>1</sup> Moreover, previous studies find that exchange rate movements are largely disconnected from macro fundamentals (see Meese, 1990; Frankel and Rose, 1995). In this paper, I study the dynamics of exchange rates with macro variables in a no-arbitrage term structure model. I find that macro risk premia can generate deviations from UIRP and after taking into account risk premia, macro variables and exchange rates are connected more tightly than previous studies have found.

I incorporate macro variables as factors in a two-country term structure model by assuming that central banks set short term interest rates in response to the output gap and inflation and using a factor representation for the stochastic discount factor (SDF). In the model, interest rate differentials across countries and risk premia determine expected exchange rate changes, which is a typical finance perspective on the returns of risky assets. In addition, similar to monetary models of exchange rates, such as Dornbusch (1976) and Frankel (1979), innovations to macro fundamentals drive unexpected exchange rate changes. Thus, a key feature of the model in this paper is the way that macro shocks are mapped into exchange rate movements.

In my no-arbitrage setting, the exposure of exchange rate changes to macro innovations is amplified by time-varying market prices of risk, which I identify with term structure data. Ignoring these risk premia or assuming constant market prices of risk, as many empirical studies based on monetary models or New Open Economy Macroeconomics (NOEM) models

---

<sup>1</sup> UIRP states that the conditional expected value of the depreciation rate of a currency relative to another is equal to the interest rate differential between the two currencies. A related concept is the Unbiasedness Hypothesis (UH), which states that the logarithm of the forward exchange rate is an unbiased predictor of the logarithm of the future spot exchange rate. Because of covered interest rate arbitrage, the interest rate differential equals the difference between the forward exchange rate and the spot exchange rate. Hence, UIRP and the UH are equivalent concepts and I will use them interchangeably.

do, may lead to the conclusion that exchange rates are disconnected from macro fundamentals, even when there is a tight link between macro risks and exchange rate dynamics.

Since shocks to macro variables directly affect exchange rates in the model, and monetary policy shocks are important in driving exchange rate movements (see, e.g., Clarida and Galí, 1994; Eichenbaum and Evans, 1995), I use a structural Vector Autoregression (VAR) to model the joint dynamics of the output gap, inflation, and the short term interest rate. From the structural VAR, I identify shocks to the output gap, inflation, and monetary policy with standard recursive identification assumptions and derive the dynamics of state variables. In particular, I assume that the short term interest rate follows a backward-looking monetary policy rule, from which I identify monetary policy shocks. Importantly, the backward-looking policy rule for the US places weights only on the US output gap and US inflation, whereas for Germany, the policy rule also reacts to the US interest rate.<sup>2</sup>

I estimate the model with US/German data over the post-Bretton Woods era with Markov Chain Monte Carlo (MCMC) methods. The estimation reveals three major results. First, the time-varying macro risk premia are important in generating deviations from UIRP, where the model matches the data. The model implies a countercyclical foreign exchange risk premium, which is mostly driven by the output gap and inflation. After attributing the total deviation from UIRP to the risk premium associated with each macro shock, I find that monetary policy shocks, and in particular, German monetary policy shocks, account for more than 90% of the deviation from UIRP in the data.

Second, the time-varying market prices of macro risks are essential to link exchange rate movements to macro variables. In the model, macro variables account for about 38% of the variation of exchange rate changes, much higher than  $R^2$ s of around 10% found in previous studies (see, e.g., Engel and West, 2004; Lubik and Schorfheide, 2005). The correlation between the model-implied exchange rate changes and the data is over 60%. I find that the output gap and inflation account for about 70% of the variance of forecasting the conditional

---

<sup>2</sup> Recently, Engel and West (2004) and Mark (2005) also explore the empirical implication for exchange rates if central banks follow Taylor (1993) rules for setting interest rates instead of focusing on stock of monetary aggregates as in traditional monetary models of exchange rates. However, neither of these studies impose no-arbitrage conditions and price the term structure of interest rates. Instead, both of them derive the exchange rate dynamics by assuming that UIRP holds. They find that their model cannot match the variance of the exchange rate and macro variables explain a small proportion of exchange rate movements.

mean of exchange rate changes. About 50% of the variance of forecasting exchange rate changes is due to monetary policy shocks, especially, US monetary policy shocks.

Third, the model produces economically reasonable responses of the exchange rate to various macro shocks. These impulse responses are dependent on the current state of the economy, since the effects of macro shocks to the exchange rate are amplified by the time-varying market prices of macro risks. I find that the responses of the exchange rate exhibit overshooting to monetary policy shocks and are consistent with the deviations from UIRP. Impulse responses under constant market prices of risk differ significantly from responses produced by the model with time-varying market prices of risk.

This paper is related to the literature addressing the forward premium anomaly using two-country term structure models of interest rates. Early work includes Amin and Jarrow (1991) and Nielsen and Saá-Requejo (1993). Although Bansal (1997) and Backus, Foresi, and Telmer (2001) advocate tying the SDF in the term structure models to monetary policy, output growth, and other economic variables, subsequent studies (e.g., Dewachter and Maes, 2001; Han and Hammond, 2003; Leippold and Wu, 2004) only introduce more latent factors or explicit exchange rate latent factors. Models with only latent factors do not address the underlying question of what economic factors are responsible for driving the variation in interest rates, exchange rates, risk premia, and the deviations from UIRP. In comparison, this paper builds upon the recent literature linking the term structure of interest rates with macro variables (e.g., Ang and Piazzesi, 2003; Dai and Philippon, 2004; Ang, Dong, and Piazzesi, 2004; Bikbov and Chernov, 2005) and explicitly links the exchange rate and the term structure of interest rates to macroeconomic variables using a structural VAR, monetary policy rules, and no-arbitrage conditions.

The rest of the paper is organized as follows. In Section 2, I outline the model and show how to price bonds and exchange rates under no-arbitrage conditions. I discuss the data used in the paper and the econometric estimation methodology in Section 3. In Section 4, I report empirical results. Section 5 concludes.

## 2 The Model

In this section, I present a model with both observable macro variables and latent factors, where the central bank sets the short term interest rate by following a monetary policy rule and bonds and exchange rates are priced under no-arbitrage conditions.

### 2.1 Stochastic Discount Factors

I denote the US as the domestic country and Germany as the foreign country. Let the vector of state variables be  $X_t = [Z_t \ Z_t^* \ f_t \ f_t^*]^\top$ , where  $Z_t = [g_t \ \pi_t]^\top$  with  $g_t$  standing for the US output gap and  $\pi_t$  standing for the US inflation rate. I denote German variables with an asterisk. The factors  $f_t$  and  $f_t^*$  are latent and I identify shocks to  $f_t$  ( $f_t^*$ ) as monetary policy shocks of the US (Germany) using additional assumptions for the factor dynamics, which I outline below.

In the absence of a generally accepted equilibrium model for asset pricing, many researchers adopt flexible factor models as they impose only no-arbitrage conditions.<sup>3</sup> In this paper, I use a factor representation for the SDF to model the exchange rate and the term structure. I specify the SDF for country  $i$  as

$$\begin{aligned} M_{t+1}^i &= \exp(m_{t+1}^i) \\ &= \exp\left(-r_t^i - \frac{1}{2}\lambda_t^{i\top}\lambda_t^i - \lambda_t^{i\top}\varepsilon_{t+1}\right), \end{aligned} \quad (1)$$

where  $r_t^i$  is the short term interest rate of country  $i$  and  $\varepsilon_{t+1}$  is a  $6 \times 1$  vector of shocks to  $X_{t+1}$ .  $\lambda_t^i$  is a  $6 \times 1$  vector of the time-varying market prices of  $\varepsilon$  risk assigned by country  $i$ 's investors.

I follow Dai and Singleton (2002) and Duffee (2002) and assume that the market prices of risk are affine functions of  $X_t$ :

$$\lambda_t^i = \lambda_0^i + \lambda_1^i X_t, \quad (2)$$

---

<sup>3</sup> For example, Ang and Piazzesi (2003), Dai and Philippon (2004), Ang, Dong, and Piazzesi (2004) and Bikbov and Chernov (2005) use factor models with both observed and latent factors to study how yields respond to macroeconomic variables. Hördahl, Tristani, and Vestin (2004) and Rudebusch and Wu (2004) impose more structure, but use a reduced-form SDF, which is not consistent with the intertemporal marginal rate of substitution underlying the Euler equation. In contrast, Bekaert, Cho, and Moreno (2004) derive the term structure model with the SDF implicit in the IS curve for the macro model, but only for the case of constant market prices of risk.

where  $\lambda_0^i$  is a  $6 \times 1$  vector and  $\lambda_1^i$  is a  $6 \times 6$  matrix with the following parameterization:

$$\lambda_1^i = \begin{bmatrix} \lambda_{zz}^i & 0_{2 \times 1} & \lambda_{zf}^i & 0_{2 \times 1} \\ 0_{2 \times 2} & \lambda_{z^*z^*}^i & 0_{2 \times 1} & \lambda_{z^*f^*}^i \\ \lambda_{fz}^i & 0_{1 \times 2} & \lambda_{ff}^i & 0_{1 \times 1} \\ 0_{1 \times 2} & \lambda_{f^*z^*}^i & 0_{1 \times 1} & \lambda_{f^*f^*}^i \end{bmatrix}. \quad (3)$$

The parameterization of  $\lambda_1^i$  implies that the market prices of the US (German) risk are linear functions of only US (German) factors. Hence, the risk premia of US (German) macro variable shocks are time-varying in US (German) factors only. The parameterization of  $\lambda_t^i$  also implies that the US and German SDFs are correlated. This is important because Brandt, Cochrane, and Santa-Clara (2005) show that the volatility of exchange rates and the volatility of SDFs based on asset market data imply that SDFs are highly correlated across countries.

## 2.2 Exchange Rate Dynamics

The law of one price implies that the rate of depreciation of one currency relative to another is related to the SDFs of the two countries. Let  $P_t^n$  be the price of an  $n$ -period domestic bond at time  $t$ , so that

$$E_t(M_{t+1}P_{t+1}^{n-1}) = P_t^n. \quad (4)$$

The price of the same bond denominated in foreign currency is  $P_t^n/S_t$ , where  $S_t$  is the spot exchange rate (US Dollars per Deutsche Mark). Under no arbitrage, we must have

$$E_t(M_{t+1}^*P_{t+1}^{n-1}/S_{t+1}) = P_t^n/S_t.$$

If markets are complete, or  $M^i$  is the minimum variance SDF in country  $i$ , we have

$$\frac{S_{t+1}}{S_t} = \frac{M_{t+1}^*}{M_{t+1}}. \quad (5)$$

Equation (5) has been derived in various papers (see, e.g., Bekaert, 1996; Bansal, 1997; Backus, Foresi, and Telmer, 2001; Brandt, Cochrane, and Santa-Clara, 2005), and it must hold in equilibrium.<sup>4</sup> With the definition for  $M_{t+1}$  and  $M_{t+1}^*$  in equation (1), taking natural

<sup>4</sup> Graveline (2006) takes a different but equivalent approach. He chooses to specify a SDF that prices payoffs denominated in domestic currency and the stochastic process of the exchange rate. Under no-arbitrage constraint, the domestic SDF and the exchange rate process imply a SDF that prices payoffs denominated in foreign currency. When markets are complete or the SDFs are the minimum variance SDFs, the two approaches are equivalent.

logarithms of both sides of equation (5) yields the depreciation rate

$$\Delta s_{t+1} \equiv s_{t+1} - s_t = m_{t+1}^* - m_{t+1} \quad (6)$$

$$\begin{aligned} &= r_t - r_t^* + \frac{1}{2}(\lambda_t^\top \lambda_t - \lambda_t^{*\top} \lambda_t^*) + (\lambda_t^\top - \lambda_t^{*\top})\varepsilon_{t+1} \\ &= \mu_t^s + \sigma_t^s \varepsilon_{t+1}, \end{aligned} \quad (7)$$

where  $s$  is the natural logarithm of  $S$ ,  $\Delta$  stands for first difference,  $\mu_t^s = r_t - r_t^* + \frac{1}{2}(\lambda_t^\top \lambda_t - \lambda_t^{*\top} \lambda_t^*)$  and  $\sigma_t^s = \lambda_t^\top - \lambda_t^{*\top}$ .

At first glance, equation (7) suggests that a regression with time-varying coefficients and volatility can characterize the depreciation rate. However,  $\mu_t^s$  and  $\sigma_t^s$  in equation (7) are not free parameters but are severely constrained, as they are functions of the market prices of risk related to macro variables, which price the entire term structure in each country.

It is worthwhile to note several features of the depreciation rate process in equation (7). First, the conditional expected value of the exchange rate change,  $\mu_t^s$ , is composed of the foreign exchange risk premium and the interest rate differential between the US and Germany. From the perspective of a US investor, the excess return from investing in foreign exchange markets is  $s_{t+1} - s_t - r_t + r_t^*$ . Thus, the one-period expected excess return or the foreign exchange risk premium is

$$rp_t = E_t(s_{t+1} - s_t - r_t + r_t^*) = (\lambda_t^\top - \lambda_t^{*\top})\lambda_t - \frac{1}{2}(\lambda_t^\top - \lambda_t^{*\top})(\lambda_t^\top - \lambda_t^{*\top})^\top, \quad (8)$$

where the quadratic term is a Jensen's inequality term. Time-varying risk premium  $rp_t$  implies deviations from UIRP. If investors are risk-neutral, i.e.,  $\lambda_t = \lambda_t^* = 0$ , the expected rate of depreciation is simply  $r_t - r_t^*$  and UIRP holds. If  $\lambda_t$  and  $\lambda_t^*$  are constants, risk premia are constant and UIRP also holds.

The time-varying conditional mean,  $\mu_t^s$ , implies predictable variation in returns in foreign exchange markets. This is consistent with the empirical finding that the forward rate is not an unbiased predictor of the future spot rate. However, the weak predictability of exchange rate changes implies that  $\mu_t^s$  can explain only a small portion of the variation of the depreciation rate. Shocks to macro variables may be important in explaining the variation of the depreciation rate, after taking into account risk premia.



The innovations to the depreciation rate share the same shocks to  $X_t$  as implied by the no-arbitrage condition in equation (5). It is intuitive and appealing to think that shocks to the output gap, inflation, and monetary policy also drive variations in the exchange rate, which is the case in monetary models of exchange rates. In comparison, traditional finance exchange rate models using only latent factors do not have this economically meaningful interpretation.

Moreover, in equation (7), the market prices of risk are not only important in determining the conditional mean of exchange rate changes, but also directly affect the conditional volatility of exchange rate changes. The exchange rate exposure to macro innovations is amplified by the prices of risk,  $\sigma_t^s = \lambda_t^\top - \lambda_t^{*\top}$ . This exposure to macro risk also varies over time. Thus, in the model, exchange rates are heteroskedastic.

In comparison, existing structural models of exchange rates, such as the monetary models and NOEM models, typically assume that UIRP holds and dictate a static and linear relationship between macro variables and exchange rates. Compared to the time-varying mapping from macro shocks to exchange rate movements in equation (7), a static and linear relationship may be misspecified and may lead to the conclusion that exchange rates are disconnected from macro fundamentals, even when there is a tight link between them. I confirm this conjecture in Section 4.3.

Since bond returns do not necessarily span returns in foreign exchange markets, a common modelling assumption is to introduce an exchange rate factor that is orthogonal to bond market factors (see, e.g., Brandt and Santa-Clara, 2002; Leippold and Wu, 2004; and Graveline, 2006). I purposely choose not to do this here. My goal is to attribute as much variation of the exchange rate as possible to the output gap, inflation, and monetary policy shocks without resorting to latent exchange rate factors. This approach ties my hands by not assigning a specific latent factor to explain the exchange rate dynamics, but instead I break the singularity of the depreciation rate equation by assigning only an IID measurement error to equation (7). In Appendix A, I show that under common modelling assumptions, an additive IID measurement error to  $\Delta s$  can approximate the effect of foreign exchange factors on the exchange rate in a parsimonious way and does not affect the inference of issues such as the forward premium anomaly. The model-implied  $\Delta s$  then represents the maximum explanatory power of the output gap, inflation, and monetary policy shocks on the exchange rate dynamics.

## 2.3 Short Rates and Factor Dynamics

In this section, I specify the short rates for the US and Germany as linear functions of the output gap, inflation, and a latent factor. Following Ang, Dong, and Piazzesi (2004), I relate the short rate equation to a backward-looking monetary policy rule. The innovation to the latent factor is identified as the monetary policy shock. I derive the dynamics of the state variable  $X_t$  from a structural VAR describing the joint dynamics of the output gap, inflation, and short term interest rate. This structural VAR serves three purposes. First, it identifies shocks to the output gap, inflation, and monetary policy. Second, it helps to reduce the number of parameters of the model through imposing economically meaningful restrictions. Third, it relates a factor model to a typical monetary economics model. In Appendix B, I derive the factor dynamics from the structural VAR. I focus on the short rate equations below.

### US Short Rate Equation

The Taylor (1993) rule captures the notion that central banks set short term interest rates in response to movements in the output and inflation. I follow Ang, Dong, and Piazzesi (2004) and assume that the short rate equation for the domestic country is an affine function of the output gap, inflation and a latent factor  $f_t$ :

$$r_t = \gamma_0 + \gamma_1^\top Z_t + f_t, \quad (9)$$

where the latent factor,  $f_t$ , follows the process

$$f_t = \mu_f + \phi^\top Z_{t-1} + \rho f_{t-1} + \varepsilon_{t,MP}, \quad (10)$$

where  $\varepsilon_{t,MP} \sim \text{IID } \mathcal{N}(0, \sigma_f^2)$  is a monetary policy shock.

Equation (9) is a factor representation of the US short rate. The latent factor  $f_t$  can be interpreted as the effect of the lagged US macro factors and the US monetary policy shock  $\varepsilon_{t,MP}$  on  $r_t$  while controlling for the persistence of the short rate. The latent factor also allows the model to capture movements in the US term structure not directly captured by the output gap and inflation. To see this, I substitute equation (10) into equation (9) and apply equation (9) at time  $t - 1$ , which leads to a equivalent representation to equations (9) and (10):

$$r_t = (1 - \rho)\gamma_0 + \mu_f + \gamma_1^\top Z_t + (\phi - \rho\gamma_1)^\top Z_{t-1} + \rho r_{t-1} + \varepsilon_{t,MP}. \quad (11)$$

As the Fed directly controls the level of the short term interest rate, equation (11) has a structural interpretation of the Fed's reaction function, in which the short rate is a combination of a systematic reaction function of the central bank,  $\gamma_1^\top Z_t + (\phi - \rho\gamma_1)^\top Z_{t-1} + \rho r_{t-1}$ , and the monetary policy shock,  $\varepsilon_{t,MP}$  (see, Bernanke and Blinder, 1992; Christiano, Eichenbaum, and Evans, 1996).

### German Short Rate Equation

For Germany, which is a relatively small economy compared to the US, the central bank takes US monetary policy as an external constraint and may react systematically to US monetary conditions, summarized by the US short rate. This could be the result of two effects: First, since the US is a large country, higher US interest rates tend to increase German interest rates due to the integration of the global capital markets. This effect treats the US factors as global factors, to which Germany has some exposure. Second, the Bundesbank may respond to an increase in the US short rate by increasing its own short rate to avoid the inflationary effect of the devaluation of the Deutsche Mark. There is empirical evidence supporting the second effect. For example, Clarida, Galí, and Gertler (1998) find that the US Federal Funds rate significantly affects the Bundesbank's reaction function. Hence, I assume that the Bundesbank reacts to the German output gap, inflation, and the US short rate. I write the German short rate equation as

$$r_t^* = \gamma_0^* + \gamma_1^{*\top} Z_t^* + \gamma_2^* r_t + f_t^*, \quad (12)$$

where  $Z_t^* = [g_t^* \ \pi_t^*]^\top$ , with  $g^*$  and  $\pi^*$  representing the German output gap and inflation, respectively.  $r_t$  is the US short rate and  $f_t^*$  is a latent factor that follows the process

$$f_t^* = \mu_{f^*} + \phi^{*\top} Z_{t-1}^* + \rho^* f_{t-1}^* + \varepsilon_{t,MP}^*, \quad (13)$$

where  $\varepsilon_{t,MP}^* \sim \text{IID } \mathcal{N}(0, \sigma_{f^*}^2)$  can be identified as the German monetary policy shock.

Substituting equation (13) into equation (12) and applying equation (12) at time  $t - 1$ , I obtain a monetary reaction function for the Bundesbank similar to the backward-looking monetary policy rule for the US (see equation 11):

$$r_t^* = (1 - \rho^*)\gamma_0^* + \mu_{f^*} + \gamma_1^{*\top} Z_t^* + (\phi^* - \rho^*\gamma_1^*)^\top Z_{t-1}^* + \gamma_2^* r_t - \rho^*\gamma_2^* r_{t-1} + \rho^* r_{t-1}^* + \varepsilon_{t,MP}^*. \quad (14)$$

Whereas equation (12) expresses the German short rate as a linear function of factors,  $Z_t^*$ ,  $f_t^*$  and  $r_t$ , which is in turn a linear function of  $Z_t$  and  $f_t$ , equation (14) is a backward-looking monetary policy rule, in which the Bundesbank sets  $r^*$  by systematically responding to the German output gap, inflation, and the US short rate. By comparing the two equations, we can interpret the latent factor  $f_t^*$  as representing the effect of the lagged German macro factors, the US short rate, and the German monetary policy shock  $\varepsilon_{t,MP}^*$  on  $r_t^*$  while controlling for the persistence of the German short rate.

### Dynamics of the State Variables

I derive the dynamics of state variables  $X_t$  from a structural VAR of  $[Z_t, Z_t^*, r_t, r_t^*]^\top$ . Full details are provided in in Appendix B. Briefly, by introducing  $f$  and  $f^*$ , I map the identified VAR for  $[Z_t, Z_t^*, r_t, r_t^*]^\top$  into the reduced-form VAR of  $[Z_t, Z_t^*, f_t, f_t^*]^\top$  while maintaining the structural interpretation of the short rate equations as backward-looking monetary policy rules. The two latent factors have a clear economic interpretation, as they allow the monetary policy shocks to be identified by controlling for the effect of lagged macro variables and lagged short rates in the backward-looking monetary policy rules. With a mapping of the reduced form parameters to the structural VAR parameters, the dynamics of  $X_t$  follow a reduced-form VAR:

$$\begin{aligned}
 X_t &= \mu + \Phi X_{t-1} + \Sigma \varepsilon_t & (15) \\
 \begin{bmatrix} Z_t \\ Z_t^* \\ f_t \\ f_t^* \end{bmatrix} &= \begin{bmatrix} \mu_z \\ \mu_{z^*} \\ \mu_f \\ \mu_{f^*} \end{bmatrix} + \begin{bmatrix} \Phi_{zz} & 0 & \Phi_{zf} & 0 \\ \Phi_{z^*z} & \Phi_{z^*z^*} & \Phi_{z^*f} & \Phi_{z^*f^*} \\ \phi^\top & 0 & \rho & 0 \\ 0 & \phi^{*\top} & 0 & \rho^* \end{bmatrix} \begin{bmatrix} Z_{t-1} \\ Z_{t-1}^* \\ f_{t-1} \\ f_{t-1}^* \end{bmatrix} \\
 &+ \begin{bmatrix} \Sigma_z & 0 & 0 & 0 \\ 0 & \Sigma_{z^*} & 0 & 0 \\ 0 & 0 & \sigma_f & 0 \\ 0 & 0 & 0 & \sigma_{f^*} \end{bmatrix} \begin{bmatrix} \varepsilon_t \\ \varepsilon_t^* \\ \varepsilon_{t,MP} \\ \varepsilon_{t,MP}^* \end{bmatrix},
 \end{aligned}$$

where the zeros in  $\Phi$  are the results of the assumption in the underlying structural VAR, in which the German variables do not affect the US variables, but that the US short rate affects the German macro variables. The block diagonal  $\Sigma$  matrix is the result of the structural monetary policy shocks in equations (11) and (14), and the standard recursive identification assumption

for the structural VAR (see Appendix B for details).

I can express the US short rate in terms of  $X_t$  as

$$r_t = \delta_0 + \delta_1^\top X_t, \quad (16)$$

where as in equation (9)

$$\delta_0 = \gamma_0, \quad (17)$$

$$\delta_1 = [\delta_{1z}^\top, 0, \delta_{1f}, 0]^\top = [\gamma_1^\top, 0, 1, 0]^\top. \quad (18)$$

For Germany, the short rate equation takes the form

$$r_t^* = \delta_0^* + \delta_1^{*\top} X_t. \quad (19)$$

The dependence of the German short rate on the US short rate, as in equation (12), implies that the German short rate has an exposure to the US output gap, inflation, and  $f$ . Given the US short rate equation in equation (16), no arbitrage then imposes the following constraints on the German short rate coefficients on  $X_t$ :

$$\delta_0^* = \gamma_0^* + \gamma_0 \gamma_2^* \quad (20)$$

$$\delta_1^* = [\delta_{1z}^{*\top}, \delta_{1z^*}^{*\top}, \delta_{1f}^*, \delta_{f^*}^*]^\top = [\boxed{\gamma_1 \gamma_2^*}, \gamma_1^*, \boxed{\gamma_2^*}, 1]^\top. \quad (21)$$

In particular, no arbitrage imposes the following constraint on  $\delta_{1,z}^*$ , the coefficient on  $Z_t$ :

$$\delta_{1,z}^* = \delta_{1,z} \delta_{1,f}^*, \quad (22)$$

as indicated by boxes in equation (21). This constraint is imposed in the estimation.

## 2.4 Pricing Long Term Bonds

In this section, I describe the pricing of long term bonds. Estimating the system with long term yields is important for two reasons. First, long term yields help identify market prices of risk. As equation (7) shows, the market prices of risk play an important role in determining the depreciation rate. Without the term structure information, the market prices of risk are not well identified with only short rates and the exchange rate. Second, the model's fit to long term

yields serves as an over-identifying test, given that the SDFs must be constrained to be highly correlated to price the exchange rate (Brandt, Cochrane, and Santa-Clara, 2005).

Using the SDFs introduced in Section 2.1, I can compute the price of an  $n$ -period zero coupon bond of country  $i$  recursively by the Euler equation in equation (4). Ang and Piazzesi (2003) show that equations (1), (2), and (15) – (19) imply that the yield on an  $n$ -period zero coupon bond for country  $i$  is

$$y_t^{i,n} = a_n^i + b_n^{i\top} X_t. \quad (23)$$

The scalar  $a_n^i$  and the  $6 \times 1$  vector  $b_n^i$  are given by  $a_n^i = -A_n^i/n$  and  $b_n^i = -B_n^i/n$ , where  $A_n^i$  and  $B_n^i$  satisfy the following recursive relations:

$$A_{n+1}^i = A_n^i + B_n^{i\top} (\mu - \Sigma \lambda_0^i) + \frac{1}{2} B_n^{i\top} \Sigma \Sigma^\top B_n^i - \delta_0^i, \quad (24)$$

$$B_{n+1}^{i\top} = B_n^{i\top} (\Phi - \Sigma \lambda_1^i) - \delta_1^{i\top}, \quad (25)$$

with  $A_1^i = \delta_0^i$  and  $B_1^i = \delta_1^i$ .

From the difference equations (24) and (25), we can see that the constant market price of risk parameter  $\lambda_0^i$  only affect the constant yield coefficient  $a_n^i$  while the parameter  $\lambda_1^i$  also affects the factor loading  $b_n^i$ . The parameter  $\lambda_0^i$  therefore only affects average term spreads and average expected bond returns, while  $\lambda_1^i$  controls the time variation in term spreads and expected bond returns. If there are no risk premia, i.e.,  $\lambda_0^i = 0$  and  $\lambda_1^i = 0$ , a local version of the Expectations Hypothesis (EH) holds. Note that the US bond prices depend only on US factors because of the parameterization of  $\Phi$  and  $\Sigma$  in equation (15), and the market prices of risk in equation (3). The German bond prices depend on both US and German factors.

Together, the SDFs in equation (1), the market prices of risk in equation (2), the factor dynamics in equation (15), and the short rate equations (16) and (19) lead to a Duffie and Kan (1996) affine term structure model, but with both latent and observable macro variables.

### 3 Data and Econometric Methodology

This section describes the data and econometric methodology used to estimate the model. I relegate all technical issues to Appendix C.

### 3.1 Data

Over time, there has been a fundamental shift in the way the central banks of the US and Germany conduct monetary policy. Clarida, Galí, and Gertler (1998) provide some guidance on when controlling inflation became a major focus of those central banks. For the Bundesbank, Clarida, Galí, and Gertler (1998) pick March 1979, the time Germany entered the European Monetary System. They choose October 1979 for the Federal Reserve, when Volcker clearly signalled his intention to rein in inflation. They also experiment with a post-1982 sample period, as the operating procedures of the Fed prior to 1982 focused on meeting non-borrowed aggregate reserve targets instead of managing short interest rates in the post-Volcker era. I focus on a sample period starting from 1983 to avoid the possible change of monetary policy regimes in the US and Germany. The sample period ends on December 1998, since the Euro became the common currency for the European Union in January 1999. Hence, the dataset used in this paper has 192 monthly observations.

To estimate the model, I use continuously compounded yields of maturities 1, 3, 12, 24, 36, 48, and 60 months for the US and maturities 1, 3, 12, 36, and 60 months for Germany. The US bond yields of 12-month maturity and longer are from the CRSP Fama-Bliss discount bond file, while the 1-month and 3-month rates are from the CRSP Fama risk-free rate file. For Germany, I use the 1-month EuroMoney deposit rate (after converting to a continuously compounded rate) as the 1-month interest rate for Germany. For yields with maturities longer than 3 months, I update the Jorion and Mishkin (1991) German yields dataset used in Bekaert, Wei, and Xing (2003) with the EuroMoney deposit rates for the period of 1996-1997 and the German zero yield curve data (after 1997) from Datastream.

I follow Bekaert and Hodrick (2001) and compute the US Dollar/Deutsche Mark exchange rate from the quoted sterling exchange rates obtained from Datastream, which are closing middle rates provided by Reuters. The exchange rate and the zero coupon bond yields are sampled at the end of month.

I take seasonally adjusted CPI and Industrial Production Index for the US and Germany from the International Finance Statistics database. The German Industrial Production Index falls and rises over 10% in June 1984. I follow Engel and West (2004) and replace it with the

average of the neighboring months. For both the US and Germany, I compute the inflation rate as the 12-month change in CPI.<sup>5</sup> To estimate the output gap, I apply the Hodrick and Prescott (1997) filter to the monthly series of Industrial Production Index using a smoothing parameter of 129,600 as suggested by Ravn and Uhlig (2002). I demean the gap measure and impose the zero mean constraint in estimation. Moreover, I divide both the inflation measure and the output gap measure by 12 so that both sides of the short rate equation are monthly quantities.

Figure 1 plots the output gap and inflation data used in the estimation. There is no clear break in the data during the sample period. The US output gap is much smoother than that of Germany. The US economy also has less volatile inflation than Germany over the sample period of this paper.

## 3.2 Identification and Estimation

The structural VAR is not sufficient to exactly identify all the parameters, because latent factors can be arbitrarily scaled and rotated to produce observationally equivalent systems. To exactly identify the latent factor processes, I pin the mean of each latent factor to be zero and ensure that the model matches the mean of the short rate in the data (see Appendix C).

I estimate the model with MCMC. There are three main reasons why I use a Bayesian estimation method. First, the Bayesian estimation method can handle the nonlinear observation equation of the depreciation rate without any approximation. Previous papers, such as Han and Hammond (2003) and Inci and Lu (2004), use the maximum likelihood estimator and an extended Kalman filter, and they must linearize the observation equation for the depreciation rate. This approximation essentially violates the no-arbitrage condition used to derive the depreciation rate process, which could introduce large errors.

Second, the Bayesian estimation method is computationally much more tractable. The model in this paper is high-dimensional and nonlinear in parameters. The parameters are also constrained, for example, in equation (22). These complexities make a likelihood function with latent factors hard to optimize. The Bayesian estimation method also infers the latent factors

---

<sup>5</sup> I use core CPI (all items excluding food and energy) for the US. Core CPI of Germany is not available over the sample period.



from all yields and the depreciation rate by assigning a measurement error to each observation equation.

Third, the Bayesian estimation method provides a posterior distribution of the parameters and the time-series paths of latent factors  $f_t$  and  $f_t^*$ . From the posterior distribution, I can compute the finite sample moments of yields and the best estimates (the posterior mean) of the model-implied depreciation rate and the foreign exchange risk premium.

The Bayesian estimation method also allows the use of prior distribution to incorporate additional information into the parameter estimation. I follow the literature of Bayesian estimation of monetary policy rules (e.g., Justiniano and Preston, 2004; Lubik and Schorfheide, 2005) and impose positivity of  $\delta_{1,z^i}^i$  and  $\Phi_{f^i}$  and assume a normal prior distribution for  $\Phi_{f^i}$  with mean [.02, .04, .95] and a diagonal covariance matrix with [.01, .02, 10] on the diagonal.<sup>6</sup>

As in equation (7), the innovations to the depreciation rate are the same shocks to  $X_t$  as implied by the no-arbitrage condition. However, previous studies treat the innovations to the depreciation rate separately from the innovations to the state variables. For example, papers using Kalman filter techniques, such as Han and Hammond (2003) and Dewachter and Maes (2001), simply take the conditional variance of the depreciation rate in their estimation and assume that the innovations to the depreciation rate are unrelated to the factor dynamics. This approach overlooks the fact that the same shocks to the state variables drive the depreciation rate as implied by the no-arbitrage condition.

In this study, I ensure that the shocks to macro variables in equation (7) enter the depreciation rate. With the state variable process defined in equation (15), I can write equation (7) as

$$\begin{aligned}\Delta s_{t+1} &= \mu_t^s + \sigma_t^s \varepsilon_{t+1} \\ &= \mu_t^s + \sigma_t^s \Sigma^{-1}(X_{t+1} - \mu - \Phi X_t),\end{aligned}\tag{26}$$

---

<sup>6</sup> Compared to the priors used in Justiniano and Preston (2004) and Lubik and Schorfheide (2005) (90% interval [0.12, 0.87] for the long-run response to output and [1.09, 1.89] for long-run response to inflation), the priors in this paper are much less informative. Assuming  $\Phi_{f^i f^i}$  is fixed at .95, the prior distribution for  $\Phi_{f^i z^i}$  has a 90% interval [0, 3.69] for the long-run response to output gap and [0, 5.45] for long-run response to inflation. The prior distribution on  $\Phi_{f^i f^i}$  is also noninformative with a variance of 10. Hence, the only informative priors used in this paper are the positivity of the short rate's responses to output gap and inflation and the stationarity of the VAR.

where the vector of unobservable shocks  $\varepsilon_{t+1}$  is replaced by the implied residuals using the dynamics of the VAR. To break the stochastic singularity for the depreciation rate equation, I assume an additive measurement error for  $\Delta s$ ,

$$\Delta s_t = \widehat{\Delta s}_t + \eta_t^{\Delta s}, \quad (27)$$

where  $\eta_t^{\Delta s} \sim \text{IID } \mathcal{N}(0, \sigma_{\eta, \Delta s}^2)$ .

I also assume that all yields are observed with errors, which avoids the arbitrary choice of selecting a few yields to be measured without errors as in Chen and Scott (1993). Hence, the observation equations for bond yields are

$$y_t^{(n)} = \widehat{y}_t^{(n)} + \eta_t^{(n)}, \quad (28)$$

$$y_t^{*(n)} = \widehat{y}_t^{*(n)} + \eta_t^{*(n)}, \quad (29)$$

where  $\widehat{y}_t^{(n)}$  and  $\widehat{y}_t^{*(n)}$  are the model-implied yields from equation (23),  $\eta_t^{(n)} \sim \text{IID } \mathcal{N}(0, \sigma_{\eta, n}^2)$ , and  $\eta_t^{*(n)} \sim \text{IID } \mathcal{N}(0, \sigma_{\eta, n}^{*2})$ .

## 4 Empirical Results

### 4.1 Parameter Estimates

In Panel A of Table 1, I report the posterior mean and standard deviation of the factor dynamics of equation (15). The first (second) row of  $\Phi$  shows that the US output gap (inflation) can be forecasted by the lagged US output gap (inflation) and lagged latent factor,  $f_{t-1}$ . Lagged US variables significantly enter the German output gap equation. Consistent with Figure 1, the US output gap process is much smoother than the German output gap process, which is seen by the  $\Phi_{gg}$  estimate (0.93), compared to the estimated  $\Phi_{g^*g^*}$  (0.85). The German macro variables also have larger conditional variances.

Panel B reports the estimates of the short rate equation parameters. The US short rate loads positively on the US output gap and US inflation with coefficients 0.171 and 0.636, respectively. This suggests that the Fed usually increases the short rate when the economy is operating over its potential and is facing an potentially high inflation. In particular, a 1% inflation leads to a 63.6 basis points (bp) contemporaneous increase in the US short rate.

The German short rate also loads positively on the German output gap and German inflation. Consistent with Clarida, Galí, and Gertler (1998), the US short rate significantly enters the German short rate equation. The coefficient on  $f_t$  is 0.177,<sup>7</sup> which suggests that the Bundesbank increases the German short rate by 17.7 bp per 1% increase of the US short rate by the Fed.<sup>8</sup>

To check whether the model in this paper can capture the response to inflation as implied by the data in a simple OLS regression, I compute the model-implied coefficients of the following standard Taylor rule

$$r_t^i = \beta_0^i + \beta_1^i g_t^i + \beta_2^i \pi_t^i + \nu_t^i. \quad (30)$$

The model-implied coefficient is 0.511 (1.157) for the output gap (inflation) for the US. The corresponding OLS regression coefficient is 0.309 (1.068). For Germany, the model-implied coefficient is 0.240 (1.150) for the output gap (inflation), which is also very close to the corresponding OLS regression coefficient 0.148 (1.021). Therefore, the model in this paper can capture the response to inflation implied by the data in an OLS regression of standard Taylor rules.

I report the estimates of the market prices of risk in Panel C of Table 1. The estimates of  $\lambda_0$  and  $\lambda_0^*$  are very close to each other. Moreover, the elements of  $\lambda_1$  and  $\lambda_1^*$  are also estimated to be very close to each other. This is expected from Brandt, Cochrane, and Santa-Clara (2005), who show that the volatility of the exchange rate and the volatility of the SDFs based on asset markets imply that the SDFs must be highly correlated across countries. Therefore, the market prices of risk on the common source of risk assigned by investors in different countries must be very close to each other. In Brandt, Cochrane, and Santa-Clara (2005), the domestic and foreign SDFs load equally on priced shocks, and their estimates of the correlation coefficients between the US SDF and the SDFs of the UK, Germany, and Japan are all above 0.98 with

---

<sup>7</sup> As in Section 2.3, we can write the German short rate equation as  $r_t^* = \delta_0^* + \delta_{1,Z^*}^\top Z_t^* + \delta_{1,r}^* r_t + f_t^* = \delta_0^* + \delta_{1,Z}^{*\top} Z_t + \delta_{1,Z^*}^{*\top} Z_t^* + \delta_{1,f}^* f_t + f_t^*$ , where  $\delta_{1,f}^* = \delta_{1,r}^*$  and  $\delta_{1,Z}^{*\top} = \delta_{1,Z}^\top \delta_{1,r}^*$ . Panel B of Table 1 reports the estimated parameters with the constraints imposed. Therefore, the coefficient on  $f_t$  is the same as the coefficient on  $r_t$ .

<sup>8</sup> The long-run response to US inflation, as in equation (11), is  $\frac{\phi_{f\pi}}{1-\phi_{ff}} + \delta_{1,\pi} = 0.954$ , which is very close to 1. The long run response to German inflation is estimated at  $\frac{\phi_{f^*\pi^*}}{1-\phi_{f^*f^*}} + \delta_{1,\pi^*}^* = 0.441$ , which is smaller than 1. However, German inflation is positively correlated with US inflation and this ignores the inclusion of the US short rate, with its implied policy loading on US inflation, in the German short rate equation.

standard errors of 0.01 or smaller. The correlation between the SDFs of the US and Germany estimated from my model is 0.99 with a standard error smaller than 0.01. Backus, Foresi, and Telmer (2001) also report very close estimates of the market prices of risk on the common factor in their 3-factor interdependence affine term structure model. Market incompleteness or missing priced risk does not significantly reduce the high correlation between SDFs, as Brandt, Cochrane, and Santa-Clara (2005) show.

In Panel D of Table 1, I report the estimates of the standard deviations (per month) of the measurement errors. The standard deviations of the measurement errors are fairly small for all yields. For the US, the estimates range from 7.01 bp for the one-month yield to 4.50 bp for the 60-month yield. The German yield curve is fitted even better, with the standard deviations of the measurement errors all around 4 bp. The standard deviation of the measurement error of  $\Delta s$  has a posterior mean of 3.07% and standard deviation of 0.19%. These estimates suggest that the model provides a very good fit to the yield curves and a reasonable fit to exchange rate dynamics.

In Figure 2, I plot the estimated time series of the latent factors and contrast them with the demeaned short rates. Simple eyeballing suggests that the estimated latent factors do not resemble the exchange rate level or changes, but in common with many term structure models with latent and macro factors, the latent factors closely follow interest rate levels. This suggests that it is not with the latent factors that the model captures the exchange rate dynamics, but with macro risks.

### **Matching Moments of Macro Variables and Yields**

Table 2 reports the first and second unconditional moments of macro variables and yields in the data and implied by the model. I compute the standard errors of the data moments using the General Method of Moments (GMM) with 4 lags. For the moments implied by the model, I report their posterior standard deviation. Panel A shows that the model provides a close match to the means, standard deviations, and autocorrelations of the output gap and inflation for both the US and Germany. Note that the output gaps data are demeaned and I impose the zero mean constraint in the estimation. In summary, Panel A suggests that the factor dynamics in equation (15) produce a good fit to the dynamics of macro variables in the data.

Panels B and C of Table 2 show that the model of this paper provides a close match to the means, standard deviations, and autocorrelations of the US and German yields. Because of the additive IID measurement errors in equations (28) and (29), the standard deviations of model-implied yields are always slightly lower than their data counterparts by construction. The extremely high term spread found in the complete affine models of Backus, Foresi, and Telmer (2001) does not show up in the more flexible essential affine model of this paper.

## 4.2 Model-Implied Exchange Rate Dynamics

The model-implied depreciation rate  $\widehat{\Delta}s$  in this paper represents the maximum explanatory power of the included macro variables. In Panel D of Table 2, I present moments of the model-implied depreciation rate  $\widehat{\Delta}s$  and contrast them with the data. The mean of  $\widehat{\Delta}s$  matches the mean of the data depreciation rate  $\Delta s$ . The standard deviation of  $\widehat{\Delta}s$  is about 2.3 times smaller than the data. This is the result of the additive measurement error. The lower volatility of the model-implied depreciation rate implies that there are factors affecting exchange rates not included in the model. Some of these may be variables like the current account (see Hooper and Morton, 1978), or market incompleteness (see Brandt and Santa-Clara, 2002). The autocorrelation of  $\widehat{\Delta}s$  is 0.206, higher than the 0.021 autocorrelation in the data. This is because  $\mu_t^s$  is very persistent (with an autocorrelation of 0.911), while the model-implied exchange rate changes are not as volatile as the data. The standard deviation of  $\mu_t^s$  is estimated at 0.374, about one-ninth of the standard deviation of the data exchange rate changes  $\Delta s$ , and one-fourth of the standard deviation of the model implied  $\widehat{\Delta}s$ . Thus, most model-implied exchange rate movements are unexpected.

In the top panel of Figure 3, I plot the posterior mean of the model-implied depreciation rate  $\widehat{\Delta}s$ , which is a function of macro variables, together with the data depreciation rate.<sup>9</sup> The correlation between  $\Delta s$  and  $\widehat{\Delta}s$  is 0.616, and the  $R^2$  of regressing  $\Delta s$  onto  $\widehat{\Delta}s$  (with a constant)

---

<sup>9</sup> The model-implied exchange rate change,  $\widehat{\Delta}s$ , successfully reproduces several large movements of the US Dollar/Deutsche Mark exchange rate. One interesting episode is the 8.75% depreciation of the Deutsche Mark in October 1992, which is associated with the turmoil of the British Pound's withdrawal from the European Monetary System. The model-implied exchange rate change in October 1992 is -4.69%. After decomposing the  $\widehat{\Delta}s$  into the contributions of each shock to  $X_t$  in this month, the US (German) monetary policy shock accounts for 21.8% (44.3%) of the 4.69% depreciation. In addition, the German inflation shock accounts for 38.4%.

is 38%. Therefore, a significant proportion of exchange rate movements is explained by macro fundamentals. In comparison, empirical studies based on monetary models or NOEM models find that macro variables can only explain about 10% of the variation of exchange rate changes in the data. For example, Engel and West (2004) estimate a monetary model using monthly US/German data and report a correlation of 10% between the model-implied exchange rate changes and the data. Lubik and Schorfheide (2005) look at US/Euro exchange rate and find that their estimated model explains 10% of the variation of the exchange rate changes in the data. The model in this paper improves the fit to the exchange rate by deriving the depreciate rate process under no arbitrage. Unlike Engel and West (2004) and Lubik and Schorfheide (2005), UIRP does not hold and thus the exchange rate exposure to macro innovations is amplified and varies over time due to the time-varying market prices of risk. I explore the importance of time-varying risk premia with a simulation study below.

The model in this paper links exchange rate changes, instead of the exchange rate level, to macro fundamentals. However, I can compute the model-implied exchange rate levels from  $\widehat{\Delta s}$ . In the bottom panel of Figure 3, I plot the level of the exchange rate in the data and that implied by the model. I define the model-implied exchange rate level as the cumulative sum of the posterior mean of the exchange rate changes,  $\widehat{S}_t = S_1 + \exp(\sum_{i=2}^t \widehat{\Delta s}_i)$ , where  $S_1$  is the data exchange rate level at the beginning of the sample period. As we can see in Figure 3, the model-implied exchange rate level,  $\widehat{S}$ , shares the same trend as the exchange rate level in the data  $S$  and largely tracks the movement of  $S$ . The correlation between  $\widehat{S}$  and  $S$  is 0.874.

Panel D of Table 2 also reports the first and second moments of the exchange rate level in the data and implied by the model.  $\widehat{S}$  matches the autocorrelation of the data exchange rate level, but has a slightly lower mean. The volatility of  $\widehat{S}$  is about half that of the data, as  $\widehat{\Delta s}$  has a smaller standard deviation than the depreciation rate in the data. Although not directly comparable to this paper due to different modeling assumptions, Engel and West (2004) find that the exchange rate level implied by their model has a variance of about 1/16 of that observed in the data, while Mark (2005) finds that the model-implied exchange rate level is much more volatile than the data.

### 4.3 How Important are Time-Varying Market Prices of Risk?

In monetary models and NOEM models, the exchange rate is a linear function of macro variables, which we can rewrite in first differences as

$$\Delta s_t = \beta_0 + \beta_1 \Delta r_t + \beta_1^* \Delta r_t^* + \beta_2 \Delta g_t + \beta_2^* \Delta g_t^* + \beta_3 \Delta \pi_t + \beta_3^* \Delta \pi_t^* + \nu_t, \quad (31)$$

where  $\nu_t$  is an IID residual. In these models, domestic and foreign macro variables typically enter the exchange rate equation in differences (see, e.g., Meese and Rogoff, 1983; Engel and West, 2004). Hence, the coefficients in equation (31) are typically constrained, i.e.,  $\beta_i = -\beta_i^*$  for  $i = 1, 2, 3$ .

When estimating the regression (31) on the data used in this paper, I find an  $R^2$  of 10.1% for the unconstrained regression and an  $R^2$  of 8.2% for the constrained regression, respectively. We can view the  $R^2$  of unconstrained regression as the upper bound for the  $R^2$  that any model dictating a static linear relationship between the exchange rate and macro variables can produce. Thus, although macro factors account for 38% of the variation of exchange rate changes in the model, we cannot capture this link between macro factors and exchange rates using a standard macro regression like equation (31). I conduct simulation exercises to illustrate how important this time-varying mapping is and why the static linear regressions as in (31) typically find little evidence of macroeconomic determination of exchange rates.

I simulate two datasets using the posterior mean of parameters listed in Table 1. First, I simulate a very long sample (10,000 monthly observations) to compute the population regression coefficients. Second, I simulate 1,000 samples each with 192 monthly observations. By construction, the simulated depreciation rate is completely determined by macro variables since it is generated using equation (7) without measurement error. I find that the mean, standard deviation, and autocorrelation of the simulated data closely match those of the real data. The only exception is that the model-generated  $\Delta s$  has lower volatility than observed in the data, as the model-implied depreciation rate  $\widehat{\Delta s}$  accounts for about 38% of variation of the data depreciation rate. This is expected from the match of the model-implied exchange rate to the data in Table 2.

In Table 3, I report the coefficient estimates and the  $R^2$  of the regression based on (31)

using the simulated samples. In Panel A, I report the results for unconstrained regressions. The population  $R^2$  is only 45.2%, even though the depreciation rate is generated under the null that the depreciation rate is completely determined by macro variables. This suggests that even with a large amount of observations, the linear regression cannot capture the time-varying link between macro shocks and the exchange rate.

The small sample regression results suggest that the coefficient estimates have large standard deviations. The mean  $R^2$  is 44.4% with a standard deviation of 12.7%. When the regression is constrained, the coefficients on the output gap differential and the interest rate differential are statistically significant. If the US output gap (short rate) increases faster than the German output gap (short rate) and  $\Delta g - \Delta g^*$  ( $\Delta i - \Delta i^*$ ) is above its mean, the dollar tends to appreciate relative to its mean. In the constrained regression, the  $R^2$  decreases to 38.0% for the population regression and the mean  $R^2$  of small sample regressions is 29.2%. Taking into account the fact that the model explains about 38% of the depreciation rate variation in the data, the  $R^2$ s of the simulation studies roughly translate into 16.9% ( $= 0.38 \times 0.444$ ) and 11.1% ( $= 0.38 \times 0.292$ ) for the unconstrained and constrained regressions, respectively. This is a similar order of magnitude to the  $R^2$  of 10.1% (8.2%) for the unconstrained (constrained) regression when applying equation (31) to the data.

To further illustrate the importance of the time-varying market prices of risk, I simulate another dataset with  $\lambda_1^i$  set equal to zero. In this case, the market prices of risk are constant and the depreciation rate is linked to macro variables in a static and linear fashion. In Panel C of Table 3, I report the regression results using the second set of simulated samples with no time-varying risk premia. The standard deviation of the coefficient estimates in small samples are much smaller and closer to the long sample estimates than the case when  $\lambda_1^i$  is not constrained to be zero as in Panels A and B. The  $R^2$  has a mean of 94.1% and a standard deviation of 0.02, but is biased upward in small sample compared to the 87.5%  $R^2$  for the population regression. Overall, this suggests that the relation between exchange rate changes and macro variables can be well captured by the regression based on equation (31), but only in the absence of time-varying risk premia. However, the signs of coefficients in Panel C do not make much economic sense, which suggests that the model cannot pick up the true relation between macro variables and exchange rate movements without the time-varying market prices of risk.



In conclusion, time-varying market prices of risk are important in the mapping of macro shocks to unexpected exchange rate movements. The static linear relationship dictated by the monetary models and NOEM models overlooks this important feature. Consequently, empirical studies based on these linear models tend to find little evidence that exchange rates are related to macro fundamentals, even though there may be a tight link between macro variables and exchange rates.

#### 4.4 The Foreign Exchange Risk Premium

In this section, I study the time series property of the model-implied foreign exchange risk premium. Fama (1984) shows that the deviations from UIRP translate into two necessary conditions on the moments of the foreign exchange risk premium. First, the foreign exchange risk premium  $rp_t$  is negatively correlated with the interest rate differential  $r_t - r_t^*$ . Second, the foreign exchange risk premium  $rp_t$  has a larger variance than  $r_t - r_t^*$ :

$$\text{Corr}(rp_t, r_t - r_t^*) < 0, \quad (32)$$

$$|\text{Cov}(rp_t, r_t - r_t^*)| > \text{Var}(r_t - r_t^*). \quad (33)$$

In Panel A of Figure 4, I plot the model-implied conditional mean of the exchange rate change  $\mu^s$  and its components – the interest rate differential  $r - r^*$  and the foreign exchange risk premium  $rp$ . The Deutsche Mark is expected to appreciate most of the time in the sample period, as  $\mu^s$  tends to be greater than zero. This is in line with the ex post pattern of the US Dollar/Deutsch Mark in Figure 3. The foreign exchange risk premium  $rp$  moves closely with  $\mu^s$ , while the interest rate differential is much smoother than both  $rp$  and  $\mu^s$ . The standard deviation of  $rp$  is 1.44% per year, larger than 1.35% per year for  $\mu^s$ . Moreover, the foreign exchange risk premium  $rp$  is negatively correlated with the interest rate differential  $r - r^*$  with a correlation coefficient equal to  $-0.40$ . These statistics satisfy the Fama (1984) conditions in equations (32) and (33).

The negative correlation between  $r - r^*$  and  $rp$  suggests that on average, the currency carry trade (borrow in currencies with low interest rates and invest in currencies with high interest rates) is profitable, especially in the early 1990's during which the US interest rates

are falling and are low relative to German interest rates. Panel B of Figure 4 suggests that the expected carry trade profit is related to the countercyclical property of the foreign exchange risk premium. The foreign exchange risk premium has peaks and troughs, and it moves closely with the output gap differential,  $g^* - g$ . The correlation between  $rp$  and  $g^* - g$  is 67.8%. Therefore, when the US economy is sluggish relative to Germany, i.e., when  $g^* - g$  is positive, the dollar is expected to depreciate and the carry trade of borrowing in the dollar and investing in the Deutsche Mark commands a higher expected return. Hence, the foreign exchange risk premium is countercyclical.

In Panel C of Figure 4, I plot the foreign exchange risk premium,  $rp$ , together with the inflation differential between Germany and the US,  $\pi^* - \pi$ . The inflation differential is negative most of the time during the sample period. The foreign exchange risk premium  $rp$  tends to move together with  $\pi^* - \pi$ . The correlation between them is 54.4%, which is lower than the correlation between  $rp$  and the output gap differential. Nevertheless, this suggests that a large proportion of the variation of  $rp$  is due to shocks to the output gap and inflation.

## 4.5 Deviations From the Unbiasedness Hypothesis

According to the UH, the forward premium (equal to the interest rate differential with covered interest rate parity arbitrage) is an unbiased predictor of future exchange rate changes. It is often empirically tested by a regression of the following form:

$$\frac{1}{n}(s_{t+n} - s_t) = \alpha_n + \beta_n(y_t^n - y_t^{*n}) + v_{t,t+n}. \quad (34)$$

If the UH (or UIRP) holds,  $\beta_n$  should be equal to one.

Empirical studies find that  $\beta_n$  is negative (see Hodrick, 1987; Engel, 1996), which means when the interest rate differential is greater than its sample average, currency depreciation is greater than its sample average (i.e.,  $\Delta s$  is below its sample average). This departure from UIRP is known as forward premium anomaly, as many asset pricing models fail to meet the Fama (1984) conditions in equations (32) and (33) for a negative  $\beta_n$  coefficient.

Recent studies find that latent factor term structure models can generate negative  $\beta_n$  (see, e.g., Bansal, 1997; Dewachter and Maes, 2001; Han and Hammond, 2003; Ahn, 2004).

However, these studies use only latent factors and cannot attribute the deviations from UIRP to macro fundamentals. In comparison, VAR studies (e.g., Eichenbaum and Evans, 1995) find monetary policy shocks induce a departure from UIRP, but their finding only implies a deviation from UIRP conditional on a monetary policy shock. Since their VARs do not explicitly specify and identify the risk premium, they cannot quantify how important monetary policy shocks are for the unconditional deviation from UIRP studied in the literature, which is the deviation of  $\beta_1$  coefficient from 1. In this section, I quantify the relative contributions of macro shocks to the deviations from UIRP.

I compute the model-implied  $\beta_n$  coefficient by running the UH regressions in equation (34) using interest rates and exchange rate changes fitted using the time series of latent factors and parameters drawn in each iteration of the MCMC estimation. I compute the fitted exchange rate changes over horizons beyond one month as the sum of fitted one month changes.

In Panel A of Table 4, I report the UH test results. The point estimate of UH test coefficient estimated with the data is negative at the one-month horizon, then gradually turns positive. This is consistent with the finding in Chinn and Meredith (2005). The model-implied coefficients show similar pattern. For example, the one-month horizon coefficient is -0.014 for the model and -0.262 for the data.<sup>10</sup> For the 12-month horizon, the model implied coefficient is 0.218, compared to 0.517 estimated with the data.

I can quantify the relative contributions of macro shocks to the deviation from UIRP by decomposing and attributing the total deviation to the risk premium associated with each macro shock. Since  $E_t(\Delta s_{t+1}) = r_t - r_t^* + rp_t$ , I can express the slope coefficient  $\beta_1$  in equation (34) as

$$\beta_1 = \frac{Cov(r_t - r_t^* + rp_t, r_t - r_t^*)}{Var(r_t - r_t^*)} = 1 + \frac{Cov(rp_t, r_t - r_t^*)}{Var(r_t - r_t^*)}. \quad (35)$$

It is easy to see that the slope coefficient  $\beta_1$  can be negative as found in the data only if the

---

<sup>10</sup> The  $\beta_1$  based on the simple regression in equation (34) is not statistically significant from one. However, this does not mean that UIRP holds. A Bayes factor test strongly rejects the hypothesis of constant risk premium, under which UIRP holds, in favor of time-varying risk premium. Failing to reject the UH in a univariate regression may due to the low power of the test. Chinn and Meredith (2005) and Bekaert, Wei, and Xing (2003) use either longer sample or VAR-based tests and reject the short horizon UH using US/German data. Recently, Dittmar and Thornton (2004) shows that adding macro information increases the power of EH tests. In Section 4.4, the output gap differential and inflation differential clearly contain information of the foreign exchange risk premium. Therefore, I add  $g^* - g$  and  $\pi^* - \pi$  to the right hand of the UH regression. I find that the short horizon UH is rejected at the 10% level in the expanded regression test with Hodrick (1992) standard errors.

Fama (1984) conditions in equations (32) and (33) are satisfied.

Note that the total foreign exchange risk premium is the sum of the risk premium assigned to each macro shock, i.e.,  $rp_t = \sum_j rp_t^j$ , where  $rp_t^j = (\lambda_{t,j}^\top - \lambda_{t,j}^{*\top})\lambda_{t,j}$  and  $\lambda_{t,j}^i$  is the element on the  $j$ -th row of  $\lambda_t^i$  for  $j = \{g, \pi, g^*, \pi^*, f, f^*\}$ . Since the covariance operator is linear, we have

$$\beta_1 = 1 + \frac{Cov(rp_t, r_t - r_t^*)}{Var(r_t - r_t^*)} = 1 + \frac{\sum_j Cov(rp_t^j, r_t - r_t^*)}{Var(r_t - r_t^*)}. \quad (36)$$

In Panel B of Table 4, I present the decomposition results.<sup>11</sup> The risk premium associated with the US output gap shocks contributes little to the deviation of  $\beta_1$  from 1, while the German output gap risk premium makes the deviation less severe by contributing a positive number to  $\beta_1$ . The US (German) inflation risk premium contributes about 12% (13%) of the deviation from UIRP. Risk premia associated with monetary policy shocks are most important in driving  $\beta_1$  from 1. The risk premium associated with US monetary policy shocks leads to a deviation of  $-0.187$  of  $\beta_1$  from 1, about 18% of the total deviation. The risk premium associated with German monetary policy shocks have the lion's share on  $\beta_1$  by contributing  $-0.831$  to the total deviation of  $-1.055$ . Eichenbaum and Evans (1995) find that US monetary policy shocks induce a departure from UIRP. The results in this paper reveal that in terms of contribution to the deviations from UIRP, German monetary policy shocks are more important than US monetary policy shocks, and that inflation shocks also play an important role.

## 4.6 What Drives Exchange Rate Dynamics?

I now investigate the source of variation in exchange rate changes. Specifically, I compute variance decompositions of  $\mu_t^s$ ,  $rp_t$ ,  $\Delta s_t$ , and  $fp_t^n$  from the model. I ignore measurement error in the depreciation rate when computing the variance decompositions. As  $\mu_t^s$ ,  $rp_t$ , and  $\Delta s_t$  are nonlinear functions of  $X_t$ , I compute the variance decompositions using Monte Carlo simulations conditional on each observation in my sample as in Iwata and Wu (2005). I simulate the model by drawing random shocks  $\varepsilon_{t+i}$  ( $i = 1, \dots, 120$ ) from  $\mathcal{N}(0, I)$ . Conditional

---

<sup>11</sup> I compute the decomposition of the deviation of  $\beta_1$  from 1 computed using the estimated model. The Jensen's term as in equation (8) has minimal effect (less than 1%) on the deviation of  $\beta_1$  from 1, thus I do not include it in the decomposition.

on an observation of  $X_t$ , I compute the evolution of  $X$  using equation (15). I use equations (7) and (8) and compute the forecast errors,  $V_{t+n} - E_t(V_{t+n})$  for  $V = \mu^s, rp$ , and  $\Delta s$ , due to each component of  $\varepsilon$  for  $n = 1, 3, 12, 60, 120$ . This process is repeated 1,000 times. The variance decompositions are computed as the percentage of the sample variance of the forecast error due to each element of  $\varepsilon$  of the variance of the total forecast errors. The simulation-based variance decomposition method produces the same results for linear systems as standard VAR technique.<sup>12</sup>

Panel A of Table 5 reports the variance decompositions of  $\mu^s$  for various forecast horizons. At short horizons, shocks to output gaps and inflations are responsible for more than 75% of the variation in  $\mu^s$ . The latent factors  $f_t$  and  $f_t^*$ , whose shocks are monetary policy shocks, explains more variation in  $\mu^s$  as the horizon increases, ranging from 24.1% for the 1-month horizon to about 34.9% for the 120-month horizon. The last column of Panel A reports the total mean squared forecast errors (MSE), which increases as the forecast horizon increases and converges to the unconditional variance of  $\mu^s$ .

In Panel B of Table 5, I report the variance decompositions of the foreign exchange risk premium  $rp_t$ , whose variance indicates deviations from UIRP, for various forecast horizons. Together, the output gap and inflation account for about 80% (70%) of the variation of the foreign exchange risk premium over the 1 (120) month forecast horizon. This is consistent with Panels B and C of Figure 4. Note that this result is not in conflict with the finding in Section 4.5 that risk premia associated with monetary policy shocks are responsible for most of the deviation of  $\beta_1$  from 1 at the one-month horizon, as it is the covariance between  $rp_t$  and  $r_t - r_t^*$  that determines the deviation of  $\beta_1$  from 1. The MSE increases over forecast horizons and converges to the unconditional variance of  $rp$ . Moreover, the MSE of forecasting  $rp_t$  is larger than that of  $mu_t^s$  for all horizons and the unconditional variance of  $rp_t$  is also larger than  $\mu_t^s$ , which implies that  $rp_t$  is negatively correlated with  $r_t - r_t^*$ . This is consistent with the negative slope coefficient in the UH test regression and the statistics reported in Section 4.4.

I report the variance decompositions of  $\Delta s$  for various forecast horizons in Panel C of Table

---

<sup>12</sup> An alternative approach is linearizing  $\mu_{t+n}^s$ ,  $rp_{t+n}$ , and  $\Delta s_{t+n}$  by Taylor expansion and applying the VAR techniques for variance decompositions. I find that it delivers variance decomposition results similar to the simulation approach used in this paper.

5. The MSEs over various horizons are close to each other and also close to the unconditional variance of  $\Delta s$ , because most of the variation of  $\Delta s$  is due to unexpected movements. The output gap and inflation each accounts for about 10-17% variation of  $\Delta s$ . Monetary policy shocks (innovations to  $f_t$  and  $f_t^*$ ) are responsible for about 60% (50%) of the variance of forecasting  $\Delta s$  for short (long) horizons. US monetary policy shocks alone account for about 40% of the variance of forecast errors. This finding is consistent with previous studies. For example, Clarida and Galí (1994) find that above 40% of the variance of forecasting the change in the real US Dollar/Deutsche Mark exchange rate for various horizons (ranging from 1 to 20 quarters) is due to monetary shocks. Eichenbaum and Evans (1995) also find that shocks to US monetary policy contribute 43% to the variability of the US Dollar/Deutsche Mark exchange rate in their benchmark specification. However, VARs studies cannot gauge how important the risk premium is in driving the variation of exchange rates, while the departure from UIRP implies volatile risk premia. Panel B and Panel C suggest that  $rp$  unconditionally accounts for about 10% of the total variation of  $\Delta s$ . However, over short forecast horizons, the foreign exchange risk premium is less important. For example, over a 1-month horizon,  $rp$  accounts for only 1.3% of the total variation of exchange rate changes.

In Panel D of Table 5, I report the unconditional variance decomposition of the forward premium,  $fp_t^n = f_t^n - s_t = y_t^n - y_t^{*n}$ . I compute the unconditional forecast variance using a horizon of 240 months. Monetary policy shocks account for more than 55% of the variation of the forward premium across all maturities. The German output gap is also important in driving variations in the forward premium, accounting for 40% (30%) of the variation of the forward premium in the short (long) run. The US output gap and inflation account for little variation in the forward premium, as they enter both the US interest rates and German interest rates and their effects tend to cancel out.

The last column of Panel D of Table 5 reports the proportion of the unconditional variance decompositions of the forward premium due to risk premia. Since  $fp_t^n = y_t^n - y_t^{*n} = a_n + b_n X_t - (a_n^* + b_n^* X_t)$ , I follow Ang, Dong, and Piazzesi (2004) and partition the bond coefficient  $b_n^i$  on  $X_t$  into an EH term and into a risk-premia term:

$$b_n^i = b_n^{i,EH} + b_n^{i,RP},$$

where I compute the  $b_n^{EH}$  bond pricing coefficient by setting the market prices of risk  $\lambda_1^i = 0$ . I

let  $\Omega^{F,h}$  stand for the forecast variance of the factors  $X_t$  at horizon  $h$ , where  $\Omega^{F,h} = \text{var}(X_{t+h} - E_t(X_{t+h}))$ . Since yields are given by  $y_t^{i,n} = a_n^i + b_n^{i\top} X_t$ , the forecast variance of the  $n$ -maturity forward premium at horizon  $h$  is given by  $(b_n - b_n^*)^\top \Omega^{F,h} (b_n - b_n^*)$ . I compute the unconditional forecast variance using a horizon of 240 months.

I decompose the forecast variance of yields as follows:

$$\text{Risk Premia Proportion} = \frac{(b_n^{RP} - b_n^{*,RP})^\top \Omega^{F,h} (b_n^{RP} - b_n^{*,RP})}{(b_n - b_n^*)^\top \Omega^{F,h} (b_n - b_n^*)}.$$

This risk premia proportion reports only the pure risk premia term and ignores any covariances of the risk premia with the state variables.

The column under the heading ‘‘Proportion Risk Premia’’ in Panel E of Table 5 reports the proportion of the forecast variance attributable to time-varying risk premia. The remainder is the proportion of the variance implied by the predictability embedded in the VAR dynamics without risk premia, under the EH. As the maturity increases, the importance of the risk premia increases. Risk premia play important roles in explaining the forward premium over longer maturities. Unconditionally, the pure risk premia proportion of the 60-month forward premium is 28.7%.

## 4.7 Impulse Responses of Exchange rates

In this section, I compute impulse response functions to gauge the effect of various macro shocks on the exchange rate. As  $\Delta s_t$  and  $rp_t$  are nonlinear functions of  $X_t$ , I follow the literature on nonlinear impulse responses (Gallant, Rossi, and Tauchen, 1993; Koop, Pesaran, and Potter, 1996; Potter, 2000) and treat the nonlinear impulse response function as the difference between a pair of conditional expectations while averaging out the future shocks. For example, the response of  $s_{t+h}$  to a current shock  $\nu_t$  is

$$E(s_{t+h} | \Omega_{t-1}, \nu_t) - E(s_{t+h} | \Omega_{t-1}), \quad (37)$$

where  $\Omega_{t-1}$  stands for the set of information available at  $t - 1$ .

I follow Koop, Pesaran, and Potter (1996) and generate random shocks  $\varepsilon_{t+h}$  from  $\mathcal{N}(0, I)$  to compute sample paths of  $s_{t+h}$  for  $h = 1, 2, 3, \dots, 60$ , from the given initial conditions  $\Omega_{t-1}$

and  $\nu_t$ . For example, for responses to 1% shock to  $g$ ,  $\nu_t$  is the scaled column corresponding to  $g$  in  $\Sigma$  in equation (15) so that the shock to  $g$  is 1% (annualized). Each sample path is repeated ten times and the two averages of the 10 paths (with and without  $\nu_t$ ) are recorded as one outcome. This process is repeated 500 times and the conditional expectation in equation (37) is estimated as the average of the outcomes. I compute the impulse response functions conditional on each observation of  $X_t$ . I plot the mean and one-standard-deviation bands for the impulse response functions.

### **Impulse Responses of the Exchange Rate Level**

Figure 5 plots the responses of the exchange rate to 1% shocks to the output gap, inflation, and monetary policy. On average, a 1% shock to the US (German) output gap leads to an immediate 0.6% (0.3%) appreciation for the dollar (Mark). After the shock, the dollar (Mark) keeps appreciating up to 2.2% (1%) after 20 months. Therefore, economic growth strengthens the currency. On average, the Mark depreciates by 0.8% immediately after a 1% shock to German inflation and keep depreciating up to 3.2% after 20 months. In contrast, a 1% shock to US inflation appreciates the dollar, which is not consistent with the Dornbusch (1976) model, as the long run PPP condition in the Dornbusch (1976) model will lead to a depreciation of the dollar in both the short run and the long run. One explanation is related to the monetary policy rule. An increase of US inflation leads to a 67 bp contemporaneous increase of the US short rate, as the Fed sets the short rate following the reaction function in equation (11) in the model. The Fed increases the US short rate in the future from an initial shock to US inflation today due to the policy inertia in equation (11), which may lead to long run appreciation of the dollar. Therefore, bad news about inflation may be good news for the exchange rate. This phenomenon arises in the model of Clarida (2004), who shows that if central banks follow Taylor rules, a shock that pushes up inflation may trigger an aggressive rise in nominal interest rates that causes the nominal exchange rate to appreciate in the short run and long run.

I overlay the impulse response function with  $\lambda_t^i$  set to be its sample mean in thin solid lines in Figure 5.<sup>13</sup> Constant market prices of risk lead to very different impulse responses after the initial shock, which are often in opposite directions compared to responses under time-

---

<sup>13</sup> I find similar impulse response functions based on a separate estimation of the model under constant market prices of risk.



varying risk premia. This discrepancy is because under constant risk premia, the exposure of the exchange rate to the macro shocks is no longer state-dependent, and the conditional mean of  $\Delta s$  is simply the interest rate differential (plus constant risk premia). Again, this suggests that the time-varying market prices of risk are important in driving the dynamics of the exchange rate.

In the last column of Figure 5, a contractionary US (German) monetary policy shock appreciates the dollar (Mark). Both exhibit overshooting as in the model of Dornbusch (1976). The US dollar achieves its maximal response contemporaneously. In comparison, the Mark exhibits a delayed overshooting, as it achieves its maximal response about 20 months after the shock. In a VAR study, Eichenbaum and Evans (1995) find a delayed overshooting of US monetary policy shock but Faust and Rogers (2003) find that the delayed overshooting result is sensitive to identification assumptions. The last column of Figure 5 suggests that the delayed overshooting result is also sensitive to the nature of the risk premia. With constant risk premia, there is no delay in the responses the dollar/Mark exchange rate to monetary policy shocks.

The last row of Figure 5 plots the responses of the exchange rate to 1% shocks to the US short rate, German short rate, and the interest rate differential. I construct a 1% interest rate shock by shocking all of the state variables in proportion to their Cholesky decomposition so that the sum of the shocks leads to a 1% (annualized) interest rate shock. These impulse responses all have wide one-standard-deviation bands, which is in line with the large standard errors and small  $R^2$  in the UH regression tests. On average, a 1% shock to the US short rate leads to a 2.2% appreciation of the dollar and the dollar keeps appreciating up to about 3.3% after 60 months. A 1% shock to the German short rate leads to a 1.8% depreciation of the Deutsche Mark, which may reflect the influence of the increase of the US short rate after the shock. To get a better measure, I construct a 1% shock to the interest rate differential  $r_t - r_t^*$ . In the last plot, a 1% shock to  $r_t - r_t^*$  makes the dollar appreciate about 7% and the appreciation lasts for a long time, which is consistent with the persistent deviations from UIRP observed in the data.

### **Impulse Responses of the Deviations from UIRP**

Under UIRP, the impulse response functions of foreign exchange risk premium,  $rp_t$ , as in

equation (8), should be identically equal to zero. Therefore, large responses of  $rp_t$  represent deviations from UIRP, which the model generates through the prices of risk  $\lambda_t$  and  $\lambda_t^*$ . Figure 6 plots the responses of the foreign exchange risk premium  $rp_t$  to output gap shocks, inflation shocks, and monetary policy shocks conditional on each observation of  $X_t$ .<sup>14</sup> Shocks to the output gap, inflation, and monetary policy all generate nonzero initial responses of  $rp$ . The  $rp$  responses fade out monotonically, but are persistent and converge to zero only around 20 to 30 months. The output gap shocks and German monetary policy shocks have the most persistent effect on  $rp$ , lasting about 60 months. The impulse response results are consistent with the finding in Section 4.5 that deviations from UIRP are larger at short horizons and the UH holds better over longer horizons. Eichenbaum and Evans (1995) and Faust and Rogers (2003) also find that monetary policy shocks lead to deviations from UIRP in the sense that responses of the VAR-implied expectation of ex post excess returns are nonzero. Since their VARs do not price the term structure together with exchange rates, they cannot link term structure responses to exchange rate movements. However, an interesting question is what happens to the term premia when the foreign exchange risk premium responds to shocks that push up the short rate by 1%. I now investigate the responses of  $rp_t$  and the term spread to 1% interest rate shocks.

The third row of Figure 6 plots the responses of the foreign exchange risk premium to a 1% shock to the US short rate  $r_t$ , German short rate  $r_t^*$ , and the interest rate differential  $r_t - r_t^*$ . A 1% shock to the US short rate leads to a decrease in  $rp$ . In comparison, a 1% shock to the German short rate makes the  $rp$  slightly increase, but this increase is not significant from zero. Both cases imply a negative correlation between  $r_t - r_t^*$  and  $rp_t$ . A 1% shock to the interest rate differential produces a  $-0.7\%$  decrease on average in  $rp_t$  and converges to zero in about 12 months. The wide one-standard-deviation bands suggest a large amount of uncertainty and a variable foreign exchange risk premium.

The last row of Figure 6 plots the responses of the term spread,  $sp_t^i = y_t^{i,60} - r_t^i$ , to a 1% short rate shock. Note that the foreign exchange risk premium  $rp_t$  is from the perspective of a US investor. From the perspective of a German investor, the foreign exchange risk premium is  $rp_t^* = -rp_t$ . By comparing the last two rows of Figure 6, we can see that the term spreads

---

<sup>14</sup> I include the Jensen's inequality term in the  $rp_t$  as in equation (8), but excluding the Jensen's inequality term makes very little difference to the impulse response functions.

move in the same direction as the foreign exchange risk premium in response to a 1% short rate shock. For example, a 1% shock to the US (German) short rate  $r_t$  ( $r_t^*$ ) leads to a decrease of the term spread  $sp_t$  ( $sp_t^*$ ) and at the same time decreases the  $rp_t$  ( $rp_t^*$ ). As the German interest rates have a exposure to the US short rate in the model, I compute the responses of the German term spread to a 1% shock to the US short rate. In the last column, we can see that 1% shock to the US short rate increases  $sp_t^*$ , while it also increases (decreases)  $rp_t^*$  ( $rp_t$ ). These results suggest that interest rate risk premia and the foreign exchange risk premium tend to move together as they are driven by the same macro fundamentals.

## 4.8 Deviations From Purchasing Power Parity

In this section, I investigate the model-implied deviations from the relative PPP in expectation. If relative PPP in expectation holds, we have:

$$E_t(s_{t+n} - s_t) = E_t(\pi_{t,t+n} - \pi_{t,t+n}^*), \quad (38)$$

where  $\pi_{t,t+n}$  stands for inflation over the period from  $t$  to  $t + n$ .

Relative PPP in expectation is a weak form PPP among various PPP definitions, as it dictates that expected exchange rate changes are equal to the expected inflation differential. Froot and Ramadorai (2005), among others, assume that relative PPP in expectation holds in the long run as a transversality condition. How good is this assumption? What drives deviations from relative PPP in expectation? To investigate these questions, I define the deviation from relative PPP in expectation over  $n$  months,  $DPPP_n$ , as:

$$DPPP_n \equiv \frac{1}{n} [E_t(s_{t+n} - s_t) - E_t(\pi_{t,t+n} - \pi_{t,t+n}^*)], \quad (39)$$

In Figure 7, I plot the model-implied deviations from relative PPP in expectation over 1, 12, and 120 months. Clearly, for the 1-month horizon, the deviation from relative PPP in expectation is large and volatile, and resembles the conditional mean of exchange rate changes,  $\mu^s$ .  $DPPP$  monotonically becomes less volatile as the horizon gets longer. For the 120-month horizon, the deviation is much smaller, ranging from -2% to 6% per year. It shows some mean reversion property in the 1980's but it does not show a clear tendency to damp to zero in the late 1990's. The deviation from relative PPP implied by the model is in line with the finding in

the literature that short-run deviations from PPP are large and volatile, and that exchange rates tend toward PPP in the long run but the speed of convergence is extremely slow (see, Rogoff, 1996). However, the model implies that the mean of long run deviation,  $E(DPPP_\infty)$ , is 3.29% with a standard deviation of 0.146%.

I investigate the source of variation in deviations from relative PPP over different horizons by computing variance decompositions. In Panel E of Table 5, I report the unconditional variance decompositions of  $DPPP_n$  using a forecast horizon of 120 months. The variance decompositions for  $DPPP_1$  are very close to that of  $\mu^s$ , including the total MSE. This is consistent with the fact that the deviation from PPP is large at short horizons and  $DPPP_1$  resembles  $\mu^s$ , as shown in Figure 7. Consistent with Figure 7, the total MSE drops monotonically as  $n$  increases. Overall, shocks to German output gap and US monetary policy shocks drive most of the variation in DPPP over all horizons. Shocks to US output gap are important for the short horizon deviation from PPP, accounting for about 20% of the forecast variance of  $DPPP_1$ , but are less important over longer horizons.

## 5 Conclusion

This paper builds a no-arbitrage model to understand the dynamics of exchange rates with macro risks and monetary policy. I incorporate macro variables as factors in a term structure model by assuming that central banks set short term interest rates in response to the output gap and inflation and using a factor representation for the stochastic discount factors. In the model, exchange rate changes are the ratio of domestic and foreign stochastic discount factors. Interest rate differentials across countries and risk premia determine expected exchange rate changes, similar to finance models of exchange rates. Innovations to macro variables drive the unexpected exchange rate changes as in monetary models of exchange rates, but scaled by time-varying market prices of risk.

I estimate the model with US/German data. The model implies a countercyclical foreign exchange risk premium with macro risk premia playing an important role in matching deviations from UIRP. After decomposing the total deviation to each macro shock, I find that monetary policy shocks account for about 90% of the deviation from UIRP at short horizons.

I find that more than 70% of the variance of forecasting the foreign exchange risk premium is due to the output gap and inflation of both countries. About 50% of the variance of forecasting exchange rate changes is due to monetary policy shocks. I also compute impulse response functions of the exchange rate to macro shocks. The responses of the exchange rate exhibit over-shooting to monetary policy shocks and are consistent with the deviations from UIRP.

The time-varying market prices of risk are essential to link the macro shocks to unexpected exchange rate movements. The static linear relationship dictated by many monetary models overlooks the relation, and consequently, tends to find little evidence of the relationship between macro variables and exchange rates. I find that the correlation between model-implied exchange rate changes and the data is over 60%. I compute variance decompositions from the model and find that time-varying risk premia unconditionally account for about 10% of the variation of the exchange rate changes and 28.7% of the variation of the long-term forward premium.

Interesting extensions of the model considered in this paper are to impose more structure on the stochastic discount factor and the factor dynamics, for example, allowing the exchange rate to also affect output and inflation. Other interesting extensions include expanding the state variable vector to include other factors that may affect the exchange rates.

# Appendix

## A Exchange Rate Factors

In this appendix, I discuss how statistical inference based on the SDFs introduced in Section 2.1 will be affected if there are factors orthogonal to  $X_t$  affecting exchange rates. I assume that the SDFs have an exchange rate component,  $N_{t+1}^i$ , defined as  $\log(N_{t+1}^i) = -\frac{1}{2}\omega^i \omega^i - \omega^i \xi_{t+1}$ , where  $\xi$  is related to certain exchange rate factors. A common modeling assumption is that  $\xi$  is orthogonal to  $\varepsilon$  in equation (1) (see, e.g., Brandt and Santa-Clara, 2002; Leippold and Wu, 2004).

Brandt and Santa-Clara (2002) point out that the exchange rate risk  $\xi$  and its market prices of risk  $\omega^i$  are not well identified. Some economic sources for  $\omega^i$  could be the current account, terms of trade or market incompleteness. Without loss of generality, I assume  $\xi$  is a scalar summarizing all risks related to exchange rates but orthogonal to  $\varepsilon$  and  $\omega^i$  is a scalar constant. The depreciation rate with the presence of exchange rate factors is then

$$\begin{aligned} \Delta s_{t+1} &\equiv s_{t+1} - s_t = m_{t+1}^* - m_{t+1} + n_{t+1}^* - n_{t+1} \\ &= r_t - r_t^* + \frac{1}{2}(\lambda_t^\top \lambda_t - \lambda_t^{*\top} \lambda_t^*) + (\lambda_t^\top - \lambda_t^{*\top})\varepsilon_{t+1} + \frac{1}{2}(\omega^2 - \omega^{*2}) + (\omega - \omega^*)\xi_{t+1} \\ &= \widehat{\Delta s}_{t+1} + \frac{1}{2}(\omega^2 - \omega^{*2}) + (\omega - \omega^*)\xi_{t+1} \end{aligned} \quad (\text{A-1})$$

where I denote the model-implied (as in Section 2.2) depreciation rate as  $\widehat{\Delta s}_{t+1}$ .

As  $\omega^i$  is constant and only enters the drift of  $\Delta s$  through the Jensen's inequality term, and  $\xi$  is orthogonal to  $\varepsilon$ , they will not affect studying issues that involve only the drift, such as the forward premium anomaly (see Brandt and Santa-Clara, 2002). Therefore, I can treat the unspecified residual of  $\Delta s$ ,  $(\omega - \omega^*)\xi_{t+1}$ , as an IID measurement error of  $\Delta s$ .

The measurement error not only captures model misspecification by serving as a proxy for unspecified exchange rate factors, but also breaks the stochastic singularity of the depreciation rate equation. Without the measurement error, the estimation procedure attempts to fit the depreciation rate based on the tightly restricted equation (6), which may result in implausible estimates. Effectively, I focus on how much variation of  $\Delta s$  in the data can be explained by the output gap, inflation, and monetary policy shocks, while leaving the unexplained variation of  $\Delta s$  to a measurement error.

## B Factor Dynamics from a Structural VAR

To specify the the process for  $Z_t$ , I model the US economy as being represented by the following structural VAR:

$$\begin{bmatrix} A_0^{zz} & A_0^{zr} \\ -\gamma_1^\top & 1 \end{bmatrix} \begin{bmatrix} Z_t \\ r_t \end{bmatrix} = \begin{bmatrix} A_1^z \\ \mu_r \end{bmatrix} + \begin{bmatrix} A_2^{zz} & A_2^{zr} \\ (\phi - \rho\gamma_1)^\top & \rho \end{bmatrix} \begin{bmatrix} Z_{t-1} \\ r_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_t \\ \varepsilon_{t,MP} \end{bmatrix} \quad (\text{B-2})$$

where  $\mu_r = (1 - \rho)\gamma_0 + \mu_f$ , and  $\varepsilon_t \sim \mathcal{N}(0, I)$  is a structural shock to  $Z_t$ . These structural shocks are assumed to be uncorrelated, i.e.,  $E(\varepsilon_t \varepsilon_{t,MP}) = 0$ . The equation in the structural VAR for the short rate is directly from the backward-looking monetary policy rule in equation (11).

Without additional constraints, the structural VAR in equation (B-2) is not identified. I follow Bernanke and Blinder (1992) and assume  $A_0^{zr} = 0$ . That is, monetary policy actions affect output and inflation with a lag. As this paper uses monthly data, this is a plausible identification assumption. Furthermore,  $A_0^{zz}$  is assumed to be a lower-triangular matrix. This is a common assumption in recursive identification schemes used in many monetary structural VAR papers, such as Christiano, Eichenbaum and Evans (1996). With the above identification assumptions, the  $Z_t$  equation in the structural VAR can be written as

$$Z_t = a_1 + a_2^\top Z_{t-1} + a_3 r_{t-1} + \Sigma_z \varepsilon_t, \quad (\text{B-3})$$

where  $a_1 = (A_0^{zz})^{-1} A_1^z$ ,  $a_2^\top = (A_0^{zz})^{-1} A_2^{zz}$ ,  $a_3 = (A_0^{zz})^{-1} A_2^{zr}$ , and  $\Sigma_z = (A_0^{zz})^{-1}$ .  $\Sigma_z$  is a lower-triangular matrix as a result of the lower-triangular  $A_0^{zz}$ . The  $r_t$  equation in (B-2) remains unchanged with  $\varepsilon_{t,MP}$  identified

as the monetary policy shock in the backward-looking monetary policy rule. I refer to equation (B-3) as the identified VAR.

Replacing  $r_{t-1}$  in equation (B-3) with  $Z_{t-1}$  and  $f_{t-1}$  as in equation (9), I can write the joint dynamics of  $Z_t$  and  $f_t$ , the factors driving the US short rate, as

$$\begin{bmatrix} Z_t \\ f_t \end{bmatrix} = \begin{bmatrix} a_1 + a_3\gamma_0 \\ \mu_f \end{bmatrix} + \begin{bmatrix} a_2^\top + a_3\gamma_1^\top & a_3 \\ \phi^\top & \rho \end{bmatrix} \begin{bmatrix} Z_{t-1} \\ f_{t-1} \end{bmatrix} + \begin{bmatrix} \Sigma_z & 0 \\ 0 & \sigma_f \end{bmatrix} \begin{bmatrix} \varepsilon_t \\ \varepsilon_{t,MP} \end{bmatrix}. \quad (\text{B-4})$$

In Section 2.3, I discuss the advantages of using equations (B-4) and (9) instead of their equivalent representation of equations (10) and (B-2).

I assume that the German economy can be represented by the following structural VAR:

$$\begin{aligned} & \begin{bmatrix} B_0^{z^*r} & B_0^{z^*z^*} & B_0^{z^*r^*} \\ -\gamma_2^* & -\gamma_1^{*\top} & 1 \end{bmatrix} \begin{bmatrix} r_t \\ Z_t^* \\ r_t^* \end{bmatrix} \\ &= \begin{bmatrix} B_1^z \\ \mu_{r^*} \end{bmatrix} + \begin{bmatrix} B_2^{z^*r} & B_2^{z^*z^*} & B_2^{z^*r^*} \\ -\rho^*\gamma_2^* & (\phi^* - \rho^*\gamma_1^*)^\top & \rho^* \end{bmatrix} \begin{bmatrix} r_{t-1} \\ Z_{t-1}^* \\ r_{t-1}^* \end{bmatrix} + \begin{bmatrix} \varepsilon_t^* \\ \varepsilon_{t,MP}^* \end{bmatrix}, \end{aligned} \quad (\text{B-5})$$

where  $\mu_{r^*} = (1 - \rho^*)\gamma_0^* + \mu_f^*$ , and  $\varepsilon_t^* \sim \mathcal{N}(0, I)$  is a structural shock to  $Z_t^*$ . I assume that the structural shock  $\varepsilon_t^*$  is uncorrelated with the monetary policy shocks  $\varepsilon_{t,MP}^*$ , i.e.,  $E(\varepsilon_t^* \varepsilon_{t,MP}^*) = 0$ . Furthermore, I assume that  $\varepsilon_t^*$  and  $\varepsilon_{t,MP}^*$  are orthogonal to the US structural shocks  $\varepsilon_t$  and  $\varepsilon_{t,MP}$ , i.e.,  $E(\varepsilon_t \varepsilon_t^*) = 0$  and  $E(\varepsilon_{t,MP} \varepsilon_{t,MP}^*) = 0$ .

I assume the same recursive identification scheme for the German structural VAR as for the US, but I further assume that the German economy is affected by the US short term interest rate only with a lag, so that  $B_0^{z^*r} = 0$ . Hence, the identified VAR for the German macro variable  $Z^*$  takes the form:

$$Z_t^* = b_1 + b_2^\top Z_{t-1}^* + b_3 r_{t-1} + b_4 r_{t-1}^* + \Sigma_{z^*} \varepsilon_t^*, \quad (\text{B-6})$$

where  $b_1 = (B_0^{z^*z^*})^{-1} B_1^z$ ,  $b_2^\top = (B_0^{z^*z^*})^{-1} B_2^{z^*z^*}$ ,  $b_3 = (B_0^{z^*z^*})^{-1} B_2^{z^*r}$ ,  $b_4 = (B_0^{z^*z^*})^{-1} B_2^{z^*r^*}$ , and  $\Sigma_{z^*} = (B_0^{z^*z^*})^{-1}$ .  $\Sigma_{z^*}$  is a lower-triangular matrix as a result of the lower-triangular  $B_0^{z^*z^*}$ .

Substituting  $r_{t-1}$  and  $r_{t-1}^*$  in equation (B-6) with equations (9) and (12), and collecting equations (B-4) and (13), I can write the joint dynamics of  $X_t = [Z_t, Z_t^*, f_t, f_t^*]^\top$  as follows:

$$\begin{aligned} X_t &= \mu + \Phi X_{t-1} + \Sigma \varepsilon_t \quad (\text{B-7}) \\ \begin{bmatrix} Z_t \\ Z_t^* \\ f_t \\ f_t^* \end{bmatrix} &= \begin{bmatrix} \mu_z \\ \mu_{z^*} \\ \mu_f \\ \mu_{f^*} \end{bmatrix} + \begin{bmatrix} a_2^\top + a_3\gamma_1^\top & 0 & a_3 & 0 \\ \boxed{\kappa\gamma_1^\top} & b_2^\top + b_4\gamma_1^{*\top} & \boxed{\kappa} & b_4 \\ \phi^\top & 0 & \rho & 0 \\ 0 & \phi^{*\top} & 0 & \rho^* \end{bmatrix} \begin{bmatrix} Z_{t-1} \\ Z_{t-1}^* \\ f_{t-1} \\ f_{t-1}^* \end{bmatrix} \\ &+ \begin{bmatrix} \Sigma_z & 0 & 0 & 0 \\ 0 & \Sigma_{z^*} & 0 & 0 \\ 0 & 0 & \sigma_f & 0 \\ 0 & 0 & 0 & \sigma_{f^*} \end{bmatrix} \begin{bmatrix} \varepsilon_t \\ \varepsilon_t^* \\ \varepsilon_{t,MP} \\ \varepsilon_{t,MP}^* \end{bmatrix}, \end{aligned}$$

where  $\mu_z = a_1 + a_3\gamma_0$ ,  $\mu_{z^*} = b_1 + b_4\gamma_0^* + \kappa\gamma_0$ ,  $\kappa = (b_3 + b_4\gamma_2^*)$ .

The parameters in equation (B-7) are constrained. In particular, one constraint is

$$\Phi_{z^*z} = \Phi_{z^*f}\gamma_1^\top, \quad (\text{B-8})$$

as indicated by boxes in equation (B-7). The constraint is the result of mapping from the SVAR to the reduced VAR by substituting the lagged short rate in equation (B-6) with the short rate equation (9). This constraint is imposed in the estimation.

## C Econometric Identification

Dai and Singleton (2000) provides identification conditions for latent factor affine term structure models. For the two-country Gaussian model considered in this paper, my identification strategy is to set the mean of the latent

factors  $F_t \equiv [f_t \ f_t^*]$  to zero and pin down  $\delta_{1,f}$  and  $\delta_{1,f^*}^*$ , while letting the conditional variance  $\sigma_f^2$  and  $\sigma_{f^*}^2$  be unconstrained. To ensure that  $F_t$  is mean zero, I parameterize  $\mu = [\mu_z \ \mu_z^* \ \mu_f \ \mu_f^*]^\top$  so that  $\mu_f$  and  $\mu_f^*$  satisfy the equation:

$$e_i^\top (I - \Phi)^{-1} \mu = 0, i = 5, 6$$

where  $e_i$  is a vector of zeros with a one in the  $i$ -th position. I set  $\delta_{1,f} = 1$  and  $\delta_{1,f^*}^* = 1$  as implied by equations (9) and (12). Similar identification strategies have been applied in Ang, Dong, and Piazzesi (2004), Dai and Philippon (2004), and Bikbov and Chernov (2005). I also impose zero mean on output gaps  $g$  and  $g^*$ .

To match the mean of the short rate in the sample, I set  $\delta_0$  in each Gibbs iteration so that

$$\delta_0 = \bar{r} - \delta_1^\top \bar{X}, \quad \delta_0^* = \bar{r}^* - \delta_1^{*\top} \bar{X}, \quad (\text{C-9})$$

where  $\bar{r}$  ( $\bar{r}^*$ ) is the average US (German) short rate in the data and  $\bar{X}$  is the time-series average of the factors  $X_t$ . This means that  $\delta_0$  and  $\delta_0^*$  are not individually drawn as separate parameters, but  $\delta_0$  and  $\delta_0^*$  change their value in each Gibbs iteration because they are functions of  $\delta_1$  and  $\delta_1^*$ .

## D Estimating the Model

I estimate the model by MCMC with a Gibbs sampling algorithm. Lamoureux and Witte (2002), Mikkelsen (2002), Bester (2003), and Johannes and Polson (2005) develop similar Bayesian methods for estimating term structure models, but their settings do not have macro variables. Ang, Dong, and Piazzesi (2004) develop Bayesian methods for estimating term structure models with both latent and macro variables, but their observation equations are all linear.

The parameters of the model are  $\Theta = [\mu, \Phi, \Sigma, \delta_0, \delta_1, \lambda_0, \lambda_1, \delta_0^*, \delta_1^*, \lambda_0^*, \lambda_1^*, \sigma_\eta]$ , where  $\sigma_\eta$  denotes the vector of measurement error volatilities  $[\sigma_\eta^{(n)}, \sigma_{\eta^*}^{(n)}, \sigma_\eta^{\Delta s}]$ . The latent factor vector  $F$  is also generated in each iteration of the Gibbs sampler. I simulate 50,000 iterations of the Gibbs sampler after an initial burn-in period of 10,000 iterations.

I now detail the procedure for drawing each of these variables. I denote the factors  $X = [Z, Z^*, F] = \{X_t\}_{t=1}^T$ , the set of yields for all maturities in data as  $Y = \{y_t^n\}_{t=1}^T$ , and  $Y^* = \{y_t^{*,n}\}_{t=1}^T$ , and  $\Delta s = \{\Delta s_t\}_{t=1}^T$ .

### Drawing the latent factor $F_t$

The factor dynamics (15), together with the observation equations (27) – (28), imply that the model in this paper can be written as a state-space system. The state equation for the system are Gaussian and linear in  $F_t$ , but also involve the macro variables  $Z_t$  and  $Z_t^*$ . The observation equations for the yields are linear in  $F_t$

$$y_t^{(n)} = a_n + b_n^\top X_t + \eta_t^{(n)}, \quad (\text{D-1})$$

$$y_t^{*,(n)} = a_n^* + b_n^{*\top} X_t + \eta_t^{*,(n)}. \quad (\text{D-2})$$

However, the observation equation for  $\Delta s$  is not a linear function of  $F_t$

$$\Delta s_{t+1} = r_t - r_t^* + \frac{1}{2}(\lambda_t^\top \lambda_t - \lambda_t^{*\top} \lambda_t^*) + (\lambda_t^\top - \lambda_t^{*\top}) \Sigma^{-1} (X_{t+1} - \mu - \Phi X_t) + \eta_t^{\Delta s s}. \quad (\text{D-3})$$

As  $\lambda_t = \lambda_0 + \lambda_1 X_t$  and  $\lambda_t^* = \lambda_0^* + \lambda_1^* X_t$ ,  $\Delta s_{t+1}$  is quadratic in  $F_t$  and contains cross terms of  $F_t$  and  $F_{t+1}$ . Classic methods such as MLE and Kalman filter must linearize the  $\Delta s_{t+1}$  equation. The Bayesian estimation method can draw  $F_t$  by single state updating without linearization (see Jacquier, Polson, and Rossi, 1994 and 2004).

The joint posterior for the latent factor vector  $F_t$  can be written as

$$\underbrace{P(F_t | F_{-t}, Z, Z^*, Y, Y^*, \Delta s, \Theta)}_{P_{FY|F_t}} \propto \overbrace{P(\Delta s_t | X_t, X_{t-1}, \Theta) P(\Delta s_{t+1} | X_{t+1}, X_t, \Theta)}^{P_{s|F_t}} P(F_{t+1} | X_t, \mu, \Phi, \Sigma) P(Y_t | X_t, \Theta) P(Y_t^* | X_t, \Theta) P(F_t), \quad (\text{D-4})$$



where  $F_{-t}$  stands for the vector  $F$  except  $F_t$ , and  $P(\cdot | \mu, \Phi, \Theta^-, X)$  is the likelihood function for  $Y_t$  ( $Y_t^*$  and  $\Delta s_t$ ), which is normally distributed from the assumption of normality for the measurement errors in equations (27)–(29).

As the distribution of  $F_t$ ,  $P(F_t | F_{-t}, Y, Y^*, \Delta s, Z, Z^*, \Theta)$ , is not standard, I use Metropolis-Hastings to sample from it. I assume a flat normal prior  $P(F_t)$  for  $F_t$ , which is reasonable given that  $F_t$  is conditionally normal as in equation (15). The normal distribution  $P_{FY|F_t}$  combined with  $P(F_t)$  is a bivariate normal distribution, which I use as a proposal density for  $F_t$ . I denote this proposal density as  $q(F_t) = P(F_t)P_{FY|F_t}$ . I draw a candidate  $F_t$  from  $q(F_t)$  then apply a Metropolis step based on the density of  $P_{s|F_t}$ . The proposal draw  $F_t^{m+1}$  for the  $(m+1)$ -th draw of  $F_t$  is accepted with probability  $\alpha$ , where

$$\begin{aligned} \alpha &= \min \left\{ \frac{P(F_t^{m+1} | \Theta, X, Y, Y^*, \Delta s)}{P(F_t^m | \Theta, X, Y, Y^*, \Delta s)} \frac{q(F_t^m)}{q(F_t^{m+1})}, 1 \right\} \\ &= \min \left\{ \frac{P_{s|F_t^{m+1}} P_{FY|F_t^{m+1}} P(F_t^{m+1})}{P_{s|F_t^m} P_{FY|F_t^m} P(F_t^m)} \frac{P(F_t^m) P_{FY|F_t^m}}{P(F_t^{m+1}) P_{FY|F_t^{m+1}}}, 1 \right\} \\ &= \min \left\{ \frac{P(\Delta s_t | F_t^{m+1}, Z_t, X_{t-1}, \Theta) P(\Delta s_{t+1} | X_{t+1}, F_t^{m+1}, Z_t, \Theta)}{P(\Delta s_t | F_t^m, Z_t, X_{t-1}, \Theta) P(\Delta s_{t+1} | X_{t+1}, F_t^m, Z_t, \Theta)}, 1 \right\}. \end{aligned} \quad (\text{D-5})$$

To draw  $F_t$  at the beginning and the end of the sample, I integrate out the initial and end values of  $F_t$  by drawing from the VAR in (15), following Jacquier, Polson, and Rossi (2004). Since I specify the mean of  $F$  to be zero for identification, I set each generated draw of  $F$  to have a mean of zero.

### Drawing $\mu$ and $\Phi$

Updating  $\mu$  and  $\Phi$  requires Metropolis algorithms as the conditional posteriors are not standard distributions. To draw  $\mu$  and  $\Phi$ , I note that the posterior of  $\mu$  and  $\Phi$  conditional on  $X, Y, Y^*, \Delta s$  and the other parameters is:

$$P(\mu, \Phi | \Theta_-, X, Y, Y^*, \Delta s) \propto P(Y | \Theta, X) P(Y^* | \Theta, X) P(\Delta s | \Theta, X) P(X | \mu, \Phi, \Sigma) P(\mu, \Phi), \quad (\text{D-6})$$

where  $\Theta_-$  denotes the set of all parameters except  $\mu$  and  $\Phi$ . This posterior suggests an Independence Metropolis draw, where I first draw a proposal for the  $(m+1)$ -th value of  $\mu$  and  $\Phi$  ( $\mu^{m+1}$  and  $\Phi^{m+1}$ , respectively) from the proposal density:

$$q(\mu, \Phi) \propto P(X | \mu, \Phi, \Sigma) P(\mu, \Phi),$$

where I specify the prior  $P(\mu, \Phi)$  to be normally distributed (for  $\Phi_f$  and  $\Phi_{f^*}$ , the prior distribution is a Normal distribution with mean [.02, .04, .95] and covariance matrix with diagonal as [.01, .02, 10] and truncated from below at zero. For other parameters, the prior is  $N(0, 1000)$ , consequently,  $q(\mu, \Phi)$  is a natural conjugate normal distribution. The proposal draw  $(\mu, \Phi)^{m+1}$  is then accepted with probability  $\alpha$ , where

$$\begin{aligned} \alpha &= \min \left\{ \frac{P((\mu, \Phi)^{m+1} | \Theta_-, X, Y, Y^*, \Delta s)}{P((\mu, \Phi)^m | \Theta_-, X, Y, Y^*, \Delta s)} \frac{q(\mu^m, \Phi^m)}{q((\mu, \Phi)^{m+1})}, 1 \right\} \\ &= \min \left\{ \frac{P(Y | (\mu, \Phi)^{m+1}, \Theta_-, X) P(Y^* | (\mu, \Phi)^{m+1}, \Theta_-, X) P(\Delta s | (\mu, \Phi)^{m+1}, \Theta_-, X)}{P(Y | (\mu, \Phi)^m, \Theta_-, X) P(Y^* | (\mu, \Phi)^m, \Theta_-, X) P(\Delta s | (\mu, \Phi)^m, \Theta_-, X)}, 1 \right\}. \end{aligned} \quad (\text{D-7})$$

From equation (D-7),  $\alpha$  is just the ratio of the likelihoods of the new draw of  $\mu$  and  $\Phi$  relative to the old draw. I draw  $\mu$  and  $\Phi$  separately for each equation in the VAR system (15). As in Bester (2003), I combine the Independence Metropolis algorithms with Random Walk Metropolis algorithms, but update each row of  $\Phi$  together in both algorithms.

I impose the restriction that  $F_t$  is mean zero for identification. I set  $\mu_i$  to satisfy  $e_i^\top (I - \Phi)^{-1} \mu = 0$  for  $i = 5, 6$ , to ensure that the factor  $F_t$  has zero mean. Hence  $\mu_f$  and  $\mu_{f^*}$  are simply a function of the other parameters in the factor VAR in equation (15).

### Drawing $\Sigma \Sigma^\top$

I draw  $\Sigma_z \Sigma_z^\top$ ,  $\Sigma_{z^*} \Sigma_{z^*}^\top$ ,  $\sigma_f^2$ , and  $\sigma_{f^*}^2$  separately as  $\Sigma \Sigma^\top$  is block diagonal. I focus on  $\Sigma_z \Sigma_z^\top$  since drawing the other variance parameters is similar. I note that the posterior of  $\Sigma_z \Sigma_z^\top$  conditional on  $X, Y, Y^*, \Delta s$  and the other

parameters is:

$$\begin{aligned} & P(\Sigma_z \Sigma_z^\top | \Theta_-, X, Y, Y^*, \Delta s) \\ & \propto P(Y | \Theta, X) P(Y^* | \Theta, X) P(\Delta s | \Theta, X) P(X | \mu, \Phi, \Sigma) P(\Sigma_z \Sigma_z^\top), \end{aligned} \quad (\text{D-8})$$

where  $\Theta_-$  denotes the set of all parameters except  $\Sigma_z$ . This posterior suggests an Independence Metropolis draw. I draw  $\Sigma_z \Sigma_z^\top$  from the proposal density  $q(\Sigma_z \Sigma_z^\top) = P(X | \mu, \Phi, \Sigma) P(\Sigma_z \Sigma_z^\top)$ , which is an Inverse Wishart (*IW*) distribution if I specify the prior  $P(\Sigma_z \Sigma_z^\top)$  to be *IW*, so that  $q(\Sigma_z \Sigma_z^\top)$  is an *IW* natural conjugate. The proposal draw  $(\Sigma_z \Sigma_z^\top)^{m+1}$  for the  $(m + 1)$ th draw is then accepted with probability  $\alpha$ . The accept/reject probability for the draws of  $\lambda s$  is similar to equation (D-7).

### Drawing $\delta_1$ and $\delta_1^*$

I draw  $\delta_1$  using a Random Walk Metropolis step:

$$\delta_1^{m+1} = \delta_1^m + \zeta_{\delta_1} v \quad (\text{D-9})$$

where  $v \sim N(0, 1)$  and  $\zeta_{\delta_1} = .01$  is the scaling factor used to adjust the acceptance rate. The acceptance probability  $\alpha$  for  $\delta_1$  is given by:

$$\begin{aligned} \alpha &= \min \left\{ \frac{P(\delta_1^{m+1} | \Theta_-, X, Y, \Delta s)}{P(\delta_1^m | \Theta_-, X, Y, \Delta s)} \frac{q(\delta_1^m | \delta_1^{m+1})}{q(\delta_1^{m+1} | \delta_1^m)}, 1 \right\} \\ &= \min \left\{ \frac{P(\delta_1^{m+1} | \Theta_-, X, Y, \Delta s)}{P(\delta_1^m | \Theta_-, X, Y, \Delta s)}, 1 \right\}, \end{aligned} \quad (\text{D-10})$$

where the posterior  $P(\delta_1 | \Theta_-, X, Y, \Delta s)$  is given by:

$$P(\delta_1 | \Theta_-, X, Y, \Delta s) \propto P(Y | \delta_1, \Theta_-, X) P(\Delta s | \delta_1, \Theta_-, X) P(\delta_1).$$

Thus, in the case of the draw for  $\delta_1$ ,  $\alpha$  is the posterior ratio of the new and old draws of  $\delta_1$ . I set  $\delta_0$  to match the sample mean of the short rate. I impose a prior that  $\delta_{1,z}^i > 0$ .

Since  $\delta_1^*$  is constrained as in equation (22), I draw  $\delta_{1,3:5}^*$  using a Random Walk Metropolis step similar to  $\delta_1$ , but I impose the constraint that  $\delta_{1,z}^* = \delta_{1,z} \delta_{1,f}^*$  as in equation (22).

### Drawing $\lambda_0, \lambda_1, \lambda_0^*$ and $\lambda_1^*$

I draw  $\lambda s$  with a Random Walk Metropolis algorithm. I assume a flat prior. I draw each parameter separately in  $\lambda_0$  and  $\lambda_0^*$ , and each row in  $\lambda_1$  and  $\lambda_1^*$ . The accept/reject probability for the draws of  $\lambda s$  is similar to equation (D-10).

### Drawing $\sigma_\eta$

I draw the observation variance  $(\sigma_\eta^{(n)})^2$  separately from each yield and for the exchange rate changes. I specify a conjugate prior  $IG(0, 0.00002)$ , so that the posterior distribution of  $\sigma_\eta^2$  is a natural conjugate Inverse Gamma distribution, which can be drawn directly as Gibbs sampling. The prior distribution,  $IG(0, 0.00002)$ , helps keep the algorithm from over-fitting a certain yield.

## E Robustness Checks

In this appendix, I look at several robustness checks. For a model with latent factors, a valid concern is that the fit of a model is due to the latent factors. However, the latent factors in my model have an economic interpretation as the effect of lagged macro variables and short rate and a monetary policy shock on the current value of the short rate. As we can see from Figure 2, it is unlikely that the two latent factors  $f$  and  $f^*$  have been stretched to explain the dynamics of the exchange rate.

I also check whether the results are sensitive to the choice of the output gap measure. It is well-known that the output gap is estimated with large uncertainty. Hence, I re-estimate the model using the output gap produced by quadratic detrending, as in Clarida, Galí, and Gertler (1998 and 2000). I find the results remain qualitatively the same. The correlation between  $\Delta s$  and  $\widehat{\Delta s}$  is 0.619, slightly higher than 0.616 when the HP filter is used to estimate the output gap. The model still matches the UH coefficients estimated with the data. Another concern is that the output gap is estimated with the full sample data and has a looking-forward bias. As a way to check whether my results are robust to this concern, I use the unemployment rate instead of the pre-estimated output gap and re-estimate the model. The results remain qualitatively the same, with the correlation between  $\Delta s$  and  $\widehat{\Delta s}$  equal to 0.620.

As a final check, I re-estimate the model with constant market prices of risk. With constant market prices of risk, the mapping of macro shocks onto exchange rate movements is no longer state-dependent but constant over time (see equation 7). The model-implied UH regression coefficients are unit for all horizons, since the foreign exchange risk premium is constant under the constrained model. Furthermore, the correlation between  $\Delta s$  and  $\widehat{\Delta s}$  is 0.407, and only 16.6% of the variation of the exchange rate changes in data can be explained by the constrained model. The results of the constrained estimation emphasize that the time-varying market prices of risk are responsible for the matching of the deviations from UIRP in the data and the larger proportion of the exchange rate movements explained by the model in this paper.

## References

- [1] Ahn, Dong-Hyun, 2004, "Common Factors and Local Factors: Implications for Term Structures and Exchange Rates," *Journal of Financial and Quantitative Analysis* 39, 1, 69-102.
- [2] Amin, Kaushik, and Jarrow, Robert A., 1991, "Pricing Foreign Currency Options under Stochastic Interest Rates," *Journal of International Money and Finance* 10, 310-29.
- [3] Ang, Andrew, and Monika Piazzesi, 2003, "A No-Arbitrage Vector Autoregression of Term Structure Dynamics with Macroeconomic and Latent Variables," *Journal of Monetary Economics* 50, 4, 745-87.
- [4] Ang, Andrew, Sen Dong, and Monika Piazzesi, 2004, "No-Arbitrage Taylor Rules", Working paper, Columbia University.
- [5] Backus, David K., Allan Gregory, and Chris I. Telmer, 1993, "Accounting for Forward Rates in Markets for Foreign Currency," *Journal of Finance* 48, 1, 1887-1908.
- [6] Backus, David K., Silverio Foresi, and Chris I. Telmer, 2001, "Affine Term Structure Models and the Forward Premium Anomaly," *Journal of Finance* 56, 1, 279-304.
- [7] Bansal, Ravi, A. Ronald Gallant, Robert Hussey, and George Tauchen, 1995, "Nonparametric Estimation of Structural Models for High Frequency Currency Market Data," *Journal of Econometrics* 66, 251-287.
- [8] Bansal, Ravi, 1997, "An Exploration of the Forward Premium Puzzle in Currency Markets," *Review of Financial Studies* 10, 369-403.
- [9] Bekaert, Geert, 1996, "The Time Variation of Risk and Return in Foreign Exchange Markets: A General Equilibrium Perspective," *Review of Financial Studies*, 9, 427-70.
- [10] Bekaert, Geert, and Robert J. Hodrick, 2001, "Expectation Hypotheses Tests," *Journal of Finance* 56, 1357-94.
- [11] Bekaert, Geert, Min Wei, and Yuhang Xing, 2003, "Uncovered Interest Rate Parity and the Term Structure," Working paper, Columbia University.
- [12] Bekaert, Geert, Seonghoon Cho, and Antonio Moreno, 2004, "New-Keynesian Macroeconomics and the Term Structure," Working paper, Columbia University.
- [13] Bernanke, Ben S., and Alan S. Blinder, 1992, "The Federal Funds Rates and the Channels of Monetary Transmission," *American Economic Review* 82, 4, 901-21.
- [14] Bester, Alan C., 2003, "Random Fields and Affine Models of Interest Rates," Working paper, Duke University.
- [15] Bikbov, Ruslan, and Mikhail Chernov, 2005, "No-Arbitrage Macroeconomic Determinants of the Yield Curve," Working paper, Columbia University.
- [16] Brandt, Michael W., and Pedro Santa-Clara, 2002, "Simulated Likelihood Estimation of Diffusions with an Application to Exchange Rate Dynamics in Incomplete Markets," *Journal of Financial Economics* 63, 161-210.
- [17] Brandt, Michael W., John H. Cochrane, and Pedro Santa-Clara, 2005, "International Risk Sharing is Better Than You Think, or Exchange Rates are Too Smooth," Forthcoming, *Journal of Monetary Economics*.
- [18] Chen, Ren-Raw, and Louis Scott, 1993, "Maximum Likelihood Estimation for a Multi-factor Equilibrium Model of the Term Structure of Interest Rates," *Journal of Fixed Income*, 3, 14-31.
- [19] Chinn, Menzie D., and Guy Meredith, 2005, "Testing Uncovered Interest Rate Parity at Short and Long Horizons During the Post-Bretton Woods Era," NBER Working Paper 11077.
- [20] Christiano, Lawrence J., Martin Eichenbaum, and Charles Evans, 1996, "The Effects of Monetary Policy Shocks: Evidence from the Flow of Funds," *Review of Economics and Statistics* 78, 1, 16-34.
- [21] Clarida, Richard, and Jordi Galí, 1994, "Sources of Real Exchange rate Fluctuations: How Important are Nominal Shocks?" *The Carnegie-Rochester Conference Series on Public Policy* 41, 1-55.
- [22] Clarida, Richard, Jordi Galí, and Mark Gertler, 1998, "Monetary Policy Rules in Practice: Some International Evidence," *European Economic Review* 42, 1033-67.
- [23] Clarida, Richard, 2004, "Inflation Targeting and Exchange Rate Dynamics: A Classic Model Reinterpreted", invited lecture at the Central Bank of Mexico, October 2004.

- [24] Dai, Qiang, and Kenneth J. Singleton, 2000, "Specification Analysis of Affine Term Structure Models," *Journal of Finance* 55, 1943-78.
- [25] Dai, Qiang, and Kenneth J. Singleton, 2002, "Expectation Puzzles, Time-Varying Risk Premia, and Affine Term Structure Models of the Term Structure," *Journal of Financial Economics* 63, 415-41.
- [26] Dai, Qiang, and Thomas Philippon, 2004, "Fiscal Policy and the Term Structure of Interest Rates," Working paper, NYU.
- [27] Dewachter, Hans, and Konstantijn Maes, 2001, "An Affine Model for International Bond Markets," Working paper, Catholic University of Leuven.
- [28] Dittmar, Robert D., and Daniel L. Thornton, 2004, "New Evidence on the Expectations Hypothesis of the Term Structure of Bond Yields," Working paper, Federal Reserve Bank of St. Louis.
- [29] Dornbusch, Rudiger, 1976, "Expectations and Exchange Rate Dynamics," *Journal of Political Economy* 84, 1161-76.
- [30] Duffee, Gregory R., 2002, "Term Premia and Interest Rate Forecasts in Affine Models," *Journal of Finance* 57, 406-43.
- [31] Duffie, Darrell, and Rui Kan, 1996, "A Yield-Factor Model of Interest Rate," *Mathematical Finance* 6, 379-406.
- [32] Eichenbaum, Martin, and Charles Evans, 1995, "Some Empirical Evidence on the Effects of Shocks to Monetary Policy on Exchange Rates," *Quarterly Journal of Economics* 110, 4, 975-1009.
- [33] Engel, Charles, 1996, "The Forward Discount Anomaly and the Risk Premium: A Survey of Recent Evidence," *Journal of Empirical Finance* 3, 123-92.
- [34] Engel, Charles, and Kenneth D. West, 2004, "Taylor Rules and the DeutschMark-Dollar Real Exchange Rate," Forthcoming, *Journal of Money, Credit, and Banking*.
- [35] Fama, Eugene, 1984, "Forward and Spot Exchange Rates," *Journal of Monetary Economics* 14, 319-38.
- [36] Faust, Jon, and John H. Rogers, 2003, "Monetary Policy's Role in Exchange Rate Behavior," *Journal of Monetary Economics* 50, 1403-24.
- [37] Frankel, Jeffrey A., 1979, "On the Mark: A Theory of Floating Exchange Rates Based on Real Interest Differentials," *American Economic Review* 69, 601-22.
- [38] Frankel, Jeffrey A., and Andrew K. Rose, 1995, "Empirical Research on Nominal Exchange Rates," in Grossman and Rogoff (eds.), *Handbook of International Economics* Vol. 3, North Holland.
- [39] Froot, Kenneth A., and Tarun Ramadorai, 2005, "Currency Returns, Intrinsic Value, and Institutional-Investor Flows," *Journal of Finance* 60, 3, 1535-66.
- [40] Gallant, A. Ronald, Peter E. Rossi, and George Tauchen, 1993, "Nonlinear Dynamic Structures," *Econometrica* 61, 4, 871-907.
- [41] Graveline, Jeremy J., 2006, "Exchange Rate Volatility and the Forward Premium Anomaly," Working paper, Stanford University.
- [42] Han, Bing, and Peter Hammond, 2003, "Affine Models of the Joint Dynamics of Exchange Rates and Interest Rate," Working paper, University of Calgary.
- [43] Hodrick, Robert J., 1987, *The Empirical Evidence on the Efficiency of Forward and Futures Foreign Exchange Markets* (Harwood Academic Publishers, Chur, Switzerland).
- [44] Hodrick, Robert J., 1989, "Risk, Uncertainty, and Exchange Rates," *Journal of Monetary Economics* 23, 433-59.
- [45] Hodrick, Robert J., 1992, "Dividend Yields and Expected Stock Returns: Alternative Procedures for Inference and Measurement," *Review of Financial Studies* 5, 3, 357-86.
- [46] Hodrick, Robert J., and Edward C. Prescott, 1997, "Postwar U.S. Business Cycles: An Empirical Investigation," *Journal of Money, Credit and Banking* 29, 1, 1-16.
- [47] Hooper, Peter, and John E. Morton, 1978, "Fluctuations in the Dollar: A Model of Nominal and Real Exchange Rate Determination," *Journal of International Money and Finance* 1, 39-56.
- [48] Hördahl, Peter, Oreste Tristani, and David Vestin, 2004, "A Joint Econometric Model of Macroeconomic and Term Structure Dynamics," Forthcoming, *Journal of Econometrics*.

- [49] Inci, Ahmet Can, and Biao Lu, 2004, "Exchange Rates and Interest Rates: Can Term Structure Models Explain Currency Movements?," *Journal of Economic Dynamics and Control* 28, 1595-1624.
- [50] Iwata, Shigeru, and Shu Wu, 2005, "What Macroeconomic Risks Are (Not) Shared by International Investors?," *Journal of Money, Credit and Banking* 37, 1121-41
- [51] Jacquier, Eric, Nicholas. G. Polson, and Peter. E. Rossi, 1994, "Bayesian Analysis of Stochastic Volatility Models." *Journal of Business and Economic Statistics* 12, 4, 371-417.
- [52] Jacquier, Eric, Nicholas. G. Polson, and Peter. E. Rossi, 2004, Bayesian Analysis of Fat-tailed Stochastic Volatility Models with Correlated Errors, *Journal of Econometrics* 122, 185-212.
- [53] Johannes, Micheal, and Nicholas Polson, 2005, "MCMC methods for Financial Econometrics" Working paper, Columbia University.
- [54] Jorion, Philippe, and Frederic Mishkin, 1991, "A Multicountry Comparison of Term Structure Forecasts at Long Horizon," *Journal of Financial Economics* 29, 59-80.
- [55] Justiniano, Alejandro, and Bruce Preston, 2004, "Small Open Economy DSGE Models: Specification, Estimation and Model Fit," Working paper, Columbia University.
- [56] Koop, Gary, M. Hashem Pesaran, and Simon M. Potter, 1996, "Impulse Response Analysis in Nonlinear Multivariate Models," *Journal of Econometrics* 74, 119-47.
- [57] Lamoureux, Christopher G., and Douglas H. Witte, 2002, "Empirical Analysis of the Yield Curve: The Information in the Data Viewed through the Window of Cox, Ingersoll, and Ross," *Journal of Finance* 57, 1479-1520.
- [58] Leippold, Markus, and Liuren Wu, 2004, "Design and Estimation of Multi-Currency Quadratic Models," Working paper, University of Zurich.
- [59] Lubik, Thomas, and Frank Schorfheide, 2005, "A Bayesian Look at New Open Economy Macroeconomics," Forthcoming *NBER Macroeconomics Annual*.
- [60] Lucas, Robert E. Jr., 1982, "Interest Rates and Currency Prices in a Two-Country World," *Journal of Monetary Economics* 10, 335-59.
- [61] Mark, Nelson C., 2005, "Changing Monetary Policy Rules, Learning, and Real Exchange Rate Dynamics," NBER Working Paper 11061.
- [62] Meese, Richard A., and Kenneth Rogoff, 1983, "Empirical Exchange Rate Models of the Seventies: Do They Fit Out of Sample?," *Journal of International Economics* 14, 3-24.
- [63] Meese, Richard A., 1990, "Currency Fluctuations in the Post-Bretton Woods Era," *Journal of Economic Perspectives* 4,1,117-34.
- [64] Mikkelsen, Peter, 2002, "MCMC Based Estimation of Term Structure Models," Working paper, Aarhus School of Business, Denmark.
- [65] Nielsen, Lars T., and Jesús Saá-Requejo, 1993, "Exchange Rate and Term Structure Dynamics and the Pricing of Derivative Securities", Working paper, INSEAD.
- [66] Potter, Simon M., 2000, "Nonlinear Impulse Response Functions," *Journal of Economic Dynamics and Control* 24, 1425-46.
- [67] Ravn, Morten O. and Harald Uhlig, 2002, "On adjusting the HP filter for the frequency of Observations," CEPR Working Paper 2858.
- [68] Rogoff, Kenneth, 1996, "The Purchasing Power Parity Puzzle," *Journal of Economic Perspectives* 34, 2, 647-68.
- [69] Rudebusch, Glenn D., and Tao Wu, 2004, "A Macro-Finance Model of the Term Structure, Monetary Policy, and the Economy," Working paper, Federal Reserve Bank of San Francisco.
- [70] Taylor, John B., 1993, "Discretion versus Policy Rules in Practice," *Carnegie-Rochester Conference Series on Public Policy* 39, 195-214.

Table 1: Parameter Estimates

Panel A: Factor Dynamics

$\mu \times 1,000$		$\Phi$						
		$g$	$\pi$	$g^*$	$\pi^*$	$f$	$f^*$	
$g$	0.173 (0.170)	0.928 (0.016)	-0.060 (0.059)	0 –	0 –	0.120 (0.031)	0 –	
$\pi$	0.079 (0.037)	0.007 (0.011)	0.970 (0.011)	0 –	0 –	0.019 (0.008)	0 –	
$g^*$	-0.087 (0.137)	0.012 (0.006)	0.042 (0.021)	0.853 (0.031)	-0.014 (0.065)	0.068 (0.032)	0.007 (0.065)	
$\pi^*$	0.131 (0.053)	0.001 (0.002)	0.004 (0.008)	-0.001 (0.009)	0.920 (0.023)	0.005 (0.013)	0.057 (0.020)	
$f$	-0.018 (0.013)	0.003 (0.002)	0.007 (0.005)	0 –	0 –	0.978 (0.006)	0 –	
$f^*$	-0.018 (0.012)	0 –	0 –	0.033 (0.005)	0.010 (0.007)	0 –	0.969 (0.009)	
		$\Sigma_Z \Sigma'_Z \times 10^7$				$\Sigma_F \Sigma'_F \times 10^7$		
	$g$	$\pi$	$g^*$	$\pi^*$	$f$	$f$	$f^*$	
$g$	1.897 (0.180)	<b>0.144</b> [0.145]	<b>-0.047</b> [0.140]	<b>-0.235</b> [0.152]	$f$	0.557 (0.101)	<b>-0.393</b> [0.097]	
$\pi$	0.129 (0.051)	0.198 (0.012)	<b>0.278</b> [0.127]	<b>0.172</b> [0.090]	$f^*$	0 –	0.376 (0.073)	
$g^*$	0 –	0 –	14.598 (1.495)	<b>0.249</b> [0.184]				
$\pi^*$	0 –	0 –	0.163 (0.239)	0.683 (0.079)				

Panel B: Short Rate Equation Coefficients

	$\delta_0$	$g$	$\pi$	$g^*$	$\pi^*$	$f$	$f^*$
$r$	0.003 (0.000)	0.171 (0.036)	0.636 (0.160)			1 –	
$r^*$	0.004 (0.000)	0.031 (0.014)	0.114 (0.052)	0.021 (0.014)	0.118 (0.069)	0.177 (0.066)	1 –

Table 1 Continued

Panel C: Market Prices of Risk

US	$\lambda_0$	$\lambda_1$			$\lambda_0$	$\lambda_1$		
		$g$	$\pi$	$f$	$g^*$	$g^*$	$\pi^*$	$f^*$
$g$	-0.546 (0.056)	21.40 (25.18)	-201.08 (15.46)	-32.66 (18.34)	$g^*$ -0.401 (0.090)	26.85 (38.01)	-56.76 (30.78)	10.39 (21.50)
$\pi$	0.170 (0.074)	143.88 (39.91)	-62.65 (19.27)	30.15 (34.81)	$\pi^*$ 0.279 (0.107)	16.94 (28.55)	-63.02 (20.17)	30.56 (25.22)
$f$	-0.029 (0.089)	16.34 (23.92)	-47.82 (16.89)	-79.45 (13.19)	$f^*$ 0.374 (0.101)	10.29 (33.46)	16.47 (21.30)	70.64 (40.01)

Germany	$\lambda_0^*$	$\lambda_1^*$			$\lambda_0^*$	$\lambda_1^*$		
		$g$	$\pi$	$f$	$g^*$	$g^*$	$\pi^*$	$f^*$
$g$	-0.529 (0.056)	21.17 (25.37)	-205.45 (15.37)	-30.87 (18.52)	$g^*$ -0.401 (0.089)	28.03 (37.97)	-59.11 (30.86)	11.03 (21.69)
$\pi$	0.177 (0.075)	146.36 (40.10)	-64.28 (20.24)	31.13 (34.88)	$\pi^*$ 0.275 (0.107)	16.45 (28.47)	-59.77 (20.18)	27.02 (25.07)
$f$	-0.040 (0.088)	12.04 (24.23)	-42.30 (17.13)	-82.83 (13.21)	$f^*$ 0.368 (0.100)	10.12 (33.42)	18.27 (22.28)	66.21 (40.79)

Panel D: Measurement Errors (in basis points)

	$n = 1$	$n = 3$	$n = 12$	$n = 24$	$n = 36$	$n = 48$	$n = 60$	$\Delta s$
$\sigma_{\eta,n}$	7.01 (0.38)	6.04 (0.32)	4.43 (0.25)	3.72 (0.20)	3.82 (0.21)	4.19 (0.23)	4.50 (0.25)	$\sigma_{\eta,\Delta s}$ 307 19
$\sigma_{\eta,n}^*$	4.13 (0.25)	3.92 (0.24)	4.11 (0.24)		4.39 (0.24)		4.71 (0.25)	

NOTE: This table lists parameter estimates for the model. Panel A reports parameter values for the factor dynamics as in equations (15). The upper triangle of the covariance matrix contains the correlation matrix with numbers in bold and GMM standard errors (with 4 lags) in square brackets. For the correlation between latent factors  $f$  and  $f^*$ , I compute the correlation of the draws of  $f$  and  $f^*$  in each iteration and report the posterior mean and standard deviation of the correlation statistic. Panel B reports short rate equation coefficients as in equations (16) and (19). Panel C reports the market prices of risk as in equation (2). The measurement error standard deviations for yields of maturity  $n$  months are reported in Panel D. I use 50,000 iterations after a burn-in number of 10,000. Posterior standard deviation is reported under the posterior mean and in parenthesis. Parameters with standard deviation “-” are restricted to be constants.



Table 2: Fit of the Model

Panel A: Moments of Macro Factors

	Means %		Std.Deviation %		Autocorrelations	
	Data	Model	Data	Model	Data	Model
$g$	0 (0.022)	0 –	0.146 (0.012)	0.203 (0.062)	0.943 (0.035)	0.970 (0.014)
$\pi$	0.309 (0.012)	0.254 (0.053)	0.077 (0.006)	0.120 (0.037)	0.978 (0.012)	0.991 (0.004)
$g^*$	0 (0.033)	0 –	0.223 (0.022)	0.278 (0.065)	0.853 (0.040)	0.889 (0.032)
$\pi^*$	0.192 (0.019)	0.173 (0.033)	0.124 (0.014)	0.178 (0.084)	0.970 (0.021)	0.983 (0.011)

Panel B: Moments of US Yields

	$n = 1$	$n = 3$	$n = 12$	$n = 24$	$n = 36$	$n = 48$	$n = 60$
Means %							
Data	0.478 (0.024)	0.510 (0.025)	0.552 (0.026)	0.586 (0.027)	0.608 (0.026)	0.626 (0.026)	0.636 (0.026)
Model	0.478 –	0.492 (0.001)	0.542 (0.003)	0.584 (0.006)	0.608 (0.010)	0.621 (0.015)	0.628 (0.020)
Standard Deviations %							
Data	0.154 (0.014)	0.160 (0.014)	0.166 (0.016)	0.169 (0.017)	0.167 (0.017)	0.167 (0.018)	0.167 (0.018)
Model	0.152 (0.003)	0.155 (0.003)	0.161 (0.002)	0.165 (0.002)	0.166 (0.003)	0.164 (0.004)	0.161 (0.006)
Autocorrelations							
Data	0.961 (0.019)	0.983 (0.016)	0.978 (0.019)	0.977 (0.019)	0.974 (0.020)	0.974 (0.019)	0.973 (0.020)
Model	0.986 (0.001)	0.986 (0.001)	0.987 (0.001)	0.987 (0.001)	0.987 (0.001)	0.987 (0.001)	0.987 (0.001)

Table 2 Continued

## Panel C: Moments of German Yields

	$n = 1$	$n = 3$	$n = 12$	$n = 36$	$n = 60$
Means %					
Data	0.474 (0.027)	0.475 (0.026)	0.466 (0.024)	0.513 (0.020)	0.540 (0.017)
Model	0.474 –	0.471 (0.001)	0.473 (0.003)	0.509 (0.005)	0.538 (0.009)
Standard Deviations %					
Data	0.171 (0.014)	0.166 (0.014)	0.153 (0.012)	0.125 (0.009)	0.106 (0.008)
Model	0.170 (0.002)	0.166 (0.002)	0.150 (0.002)	0.120 (0.002)	0.103 (0.003)
Autocorrelations					
Data	0.987 (0.010)	0.989 (0.009)	0.983 (0.011)	0.976 (0.012)	0.968 (0.014)
Model	0.993 (0.001)	0.993 (0.001)	0.992 (0.001)	0.990 (0.001)	0.989 (0.001)

## Panel D: Moments of the Exchange Rate

	Means%		Standard Deviations%		Autocorrelations	
	Data	Model	Data	Model	Data	Model
$\Delta s$	0.205 (0.257)	0.167 (0.128)	3.339 (0.183)	1.443 (0.133)	0.021 (0.077)	0.206 (0.072)
$\mu_t^s$		0.238 (0.055)		0.374 (0.043)		0.911 (0.038)
$S$	0.552 (0.017)	0.474 (0.009)	0.108 (0.011)	0.059 (0.005)	0.980 (0.015)	0.983 (0.013)

NOTE: Panel A lists moments of the output gap and inflation in the data and implied by the model. For the model, I construct the posterior distribution of unconditional moments by computing the unconditional moments implied from the parameters in each iteration of the Gibbs sampler. Panels B and C report data and model unconditional moments of  $n$ -month maturity yields for the US and Germany, respectively. I compute the posterior distribution of the model-implied yields using the generated latent factors in each iteration. In Panel D, I report the moments of the level of exchange rate  $S$  and the exchange rate changes,  $\Delta s$ , in data and implied by the model. The moments of the conditional mean of the exchange rate changes,  $\mu_t^s$ , are also reported. I define the model-implied level of exchange rate as the cumulative sum of posterior mean of the exchange rate changes  $\Delta s$ :  $\hat{S}_t = S_1 + \exp(\sum_{i=2}^t \hat{\Delta s}_i)$ , where  $S_1$  is the data exchange rate level at the beginning of the sample period. The statistics of  $S$  and  $\hat{S}$  are not in percentage term. In all panels, the data standard errors (in parentheses) are computed using GMM (with 4 lags) and all moments are computed at a monthly frequency. For the model, we report posterior means and standard deviations (in parentheses) of each moment for the macro data and the yields. For the exchange rate level and changes, moments of the posterior mean are reported. The sample period is January 1983 to December 1998 and the data frequency is monthly.

Table 3: Monetary-Model Regressions Based on Simulated Data

$$\Delta s_t = \beta_0 + \beta_1 \Delta r_t + \beta_1^* \Delta r_t^* + \beta_2 \Delta g_t + \beta_2^* \Delta g_t^* + \beta_3 \Delta \pi_t + \beta_3^* \Delta \pi_t^* + \nu_t$$

Panel A: Unconstrained Regressions

	$\Delta g$	$\Delta g^*$	$\Delta \pi$	$\Delta \pi^*$	$\Delta i$	$\Delta i^*$	$R^2$
Population	-7.483	-1.203	3.568	-10.911	-19.120	8.376	0.452
Mean	-5.113	-1.198	2.543	-6.890	-22.701	15.338	0.444
Stdev	(11.927)	(24.631)	(2.079)	(11.648)	(20.787)	(8.854)	(0.127)

Panel B: Constrained Regressions

	$\Delta g - \Delta g^*$	$\Delta \pi - \Delta \pi^*$	$\Delta i - \Delta i^*$	$R^2$
Population	-4.010	9.426	-18.402	0.380
Mean	-2.877	5.056	-21.446	0.292
Stdev	(1.466)	(13.266)	(6.924)	(0.091)

Panel C: Unconstrained Regressions under  $\lambda_1^i = 0$

	$\Delta g$	$\Delta g^*$	$\Delta \pi$	$\Delta \pi^*$	$\Delta i$	$\Delta i^*$	$R^2$
Population	-39.445	-66.206	-0.117	11.915	40.062	22.177	0.875
Mean	-40.023	-71.033	-0.412	11.211	37.436	26.640	0.941
Stdev	(0.691)	(2.737)	(0.196)	(1.211)	(1.336)	(2.373)	(0.020)

NOTE: This table reports the results of monetary-model-style regressions using simulated data generated by the model. The simulation is based on the posterior mean of parameters listed in Table 1. The row labelled “Population” reports the regression coefficients and  $R^2$  from a long sample (10,000 monthly observations) to represent the population. I also report the mean and standard deviation of small sample estimates based on 1,000 simulations each with 192 monthly observations. Panel A reports the results for unconstrained regressions. Panel B reports the results for constrained regressions ( $\beta_i = -\beta_i^*$  for  $i = 1, 2, 3$ ). In Panel C, I report the results for unconstrained regressions using simulated data generated similarly as for Panels A and B but with  $\lambda_1^i = 0$ .

Table 4: Unbiasedness Hypotheses Tests

PANEL A: UNBIASEDNESS HYPOTHESES							
	$n = 1$	$n = 3$	$n = 12$				
Data	-0.262 (1.120)	-0.205 (1.135)	0.218 (0.985)				
Model	-0.014 (0.830)	0.092 (0.836)	0.517 (0.857)				

PANEL B: DECOMPOSITION OF THE DEVIATION OF $\beta_1$ FROM 1							
$j =$	$g$	$\pi$	$g^*$	$\pi^*$	$f$	$f^*$	sum
$\frac{Cov(rp_t^j, r_t - r_t^*)}{Var(r_t - r_t^*)}$	-0.006	-0.124	0.238	-0.144	-0.187	-0.831	-1.055
%	0.56	11.79	-22.55	13.65	17.77	78.79	100

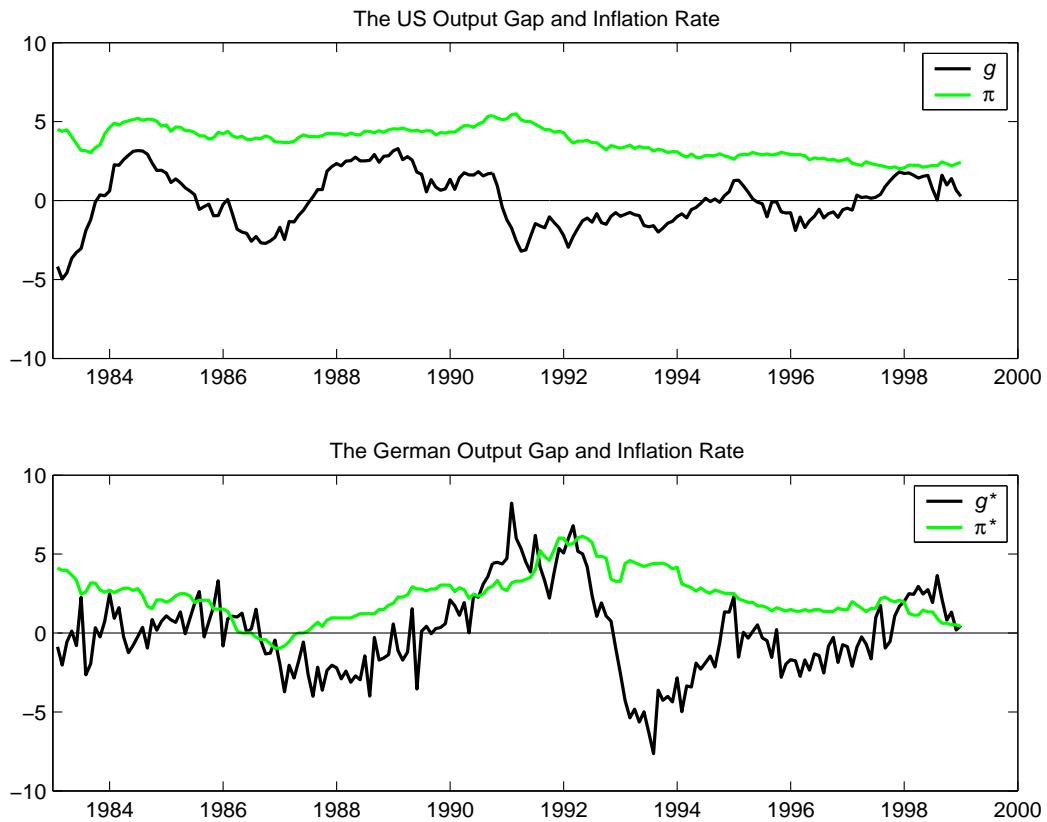
NOTE: Panel A reports unbiasedness hypothesis test coefficients for different horizon  $n$  in months. I compute the posterior distribution of the model coefficient for the UH test using yields and depreciation rates fitted by the generated latent factors in each iteration. Reported in parentheses are Hodrick (1992) standard errors with  $n - 1$  lags for the data regression and posterior standard deviations for the model estimates. Panel B reports the decomposition of the deviation of  $\beta_1$  from 1 computed using the model of this paper and the parameter estimates listed in Table 1. The Jensen's term as in equation (8) has negligible effect and I do not include it in the computation.

Table 5: Variance Decompositions

$N$	$g$	$\pi$	$g^*$	$\pi^*$	$f$	$f^*$	Total MSE $\times 10,000$
Panel A: Conditional Mean, $\mu_t^s$							
1	22.2	5.2	33.2	15.2	18.2	5.9	0.031
3	20.7	5.2	34.5	16.0	18.0	5.5	0.076
12	20.7	4.7	31.4	16.0	22.9	4.3	0.173
60	17.9	2.6	31.1	15.1	29.5	3.8	0.220
120	16.9	2.5	30.2	15.5	31.0	3.9	0.221
$Var(\mu_t^s)$							<b>0.263</b>
Panel B: Deviation From the UIRP, $rp_t$							
1	24.6	6.8	33.5	13.3	10.1	11.6	0.032
3	22.6	6.7	36.6	13.6	9.5	11.0	0.081
12	21.4	6.0	38.4	12.4	11.9	9.9	0.197
60	18.2	3.3	38.9	10.6	19.8	9.1	0.269
120	17.2	3.2	38.5	10.7	21.6	8.9	0.272
$Var(rp_t)$							<b>0.316</b>
Panel C: The Depreciation Rate, $\Delta s_{t+1}$							
1	8.5	8.3	11.6	8.3	47.7	15.6	2.505
3	9.0	8.4	10.8	7.7	48.4	15.8	2.523
12	9.8	7.8	13.1	8.7	48.8	11.9	2.751
60	13.4	10.7	15.2	8.3	37.3	15.1	2.616
120	17.4	9.8	14.3	8.4	33.2	17.0	2.615
$Var(\Delta s_{t+1})$							<b>2.082</b>
Panel D: Forward Premium, $fp_t^n$							
$n$	$g$	$\pi$	$g^*$	$\pi^*$	$f$	$f^*$	Proportion Risk Premia
1	2.1	1.5	40.7	0.7	31.5	23.6	0
3	1.8	1.7	40.7	0.6	32.5	22.7	0.1
12	1.2	2.4	39.1	0.7	36.9	19.7	2.4
36	0.9	4.5	32.5	0.8	46.0	15.4	15.2
60	1.1	6.5	27.9	0.8	50.7	13.1	28.7
Panel E: Deviation from PPP, $DPPP_n$							
$n$	$g$	$\pi$	$g^*$	$\pi^*$	$f$	$f^*$	Total MSE $\times 10,000$
1	19.2	5.8	34.0	5.8	29.8	5.4	0.222
12	16.3	7.8	39.0	0.8	27.2	8.9	0.156
60	9.2	4.1	38.2	0.3	37.8	10.6	0.038
120	6.3	3.0	39.0	0.1	40.0	11.5	0.014

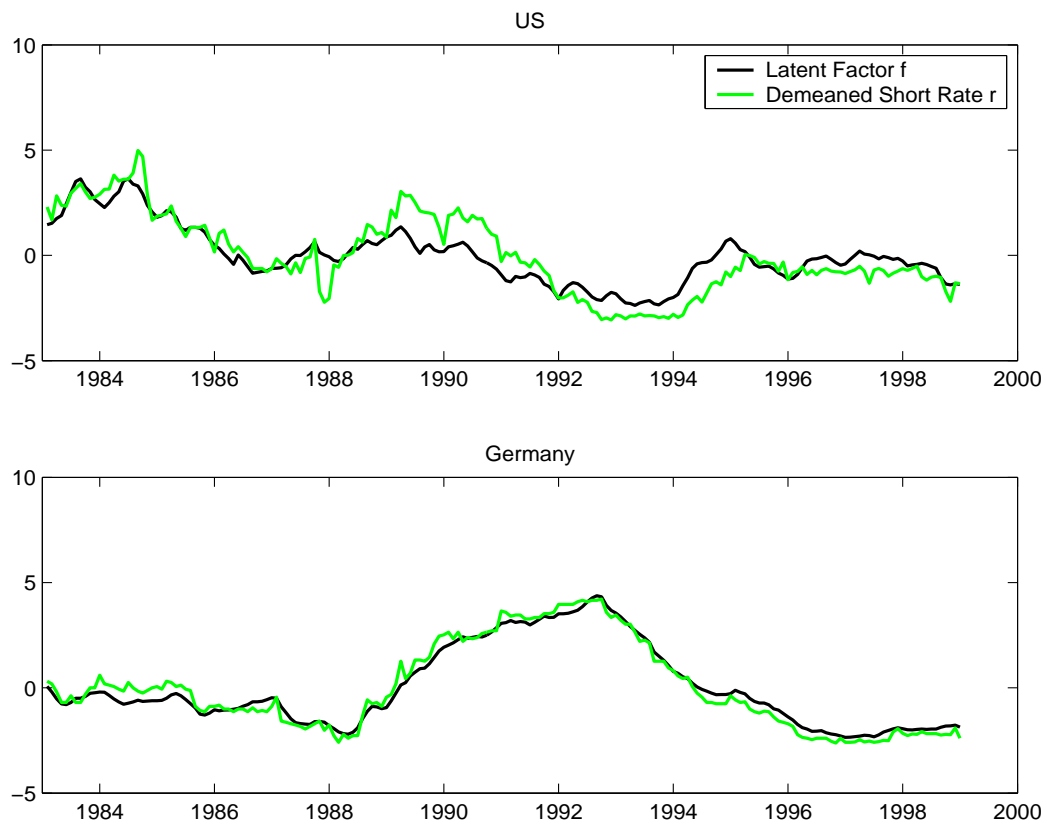
NOTE: This table reports variance decompositions of forecast variance (in percentage) for conditional mean of the depreciation rate,  $\mu_t^s$ , as defined in equation (6) in Panel A; the deviation from uncovered interest rate parity,  $rp_t$ , as defined in equation (8) in Panel B; the depreciation rate,  $\Delta s_{t+1}$ , as defined in equation (6) in Panel C; the forward premium,  $fp_t^n = f_t^n - s_t = y_t^n - y_t^{*n}$ , in Panel D; and the deviation from PPP,  $DPPP_n$ , as defined in equation (39) in Panel E. In Panels A, B, and C, I report variance decompositions computed by simulation as described in Section 4.6. I also report the total mean squared errors (MSE) for forecasts over each horizon  $N$  in the last column and the unconditional variance in the last row. I report the unconditional variance decompositions of forward premia in Panel D. In Panel E, I report variance decompositions computed by simulation as described in Section 4.6 for a forecast horizon of 120 months. All forecast horizons ( $N$ ) and maturities ( $n$ ) are in months. I ignore the observation error for computing variance decompositions. All the variance decompositions are computed using the posterior mean of the parameters listed in Table 1.

Figure 1: Macro Data Used in the Estimation



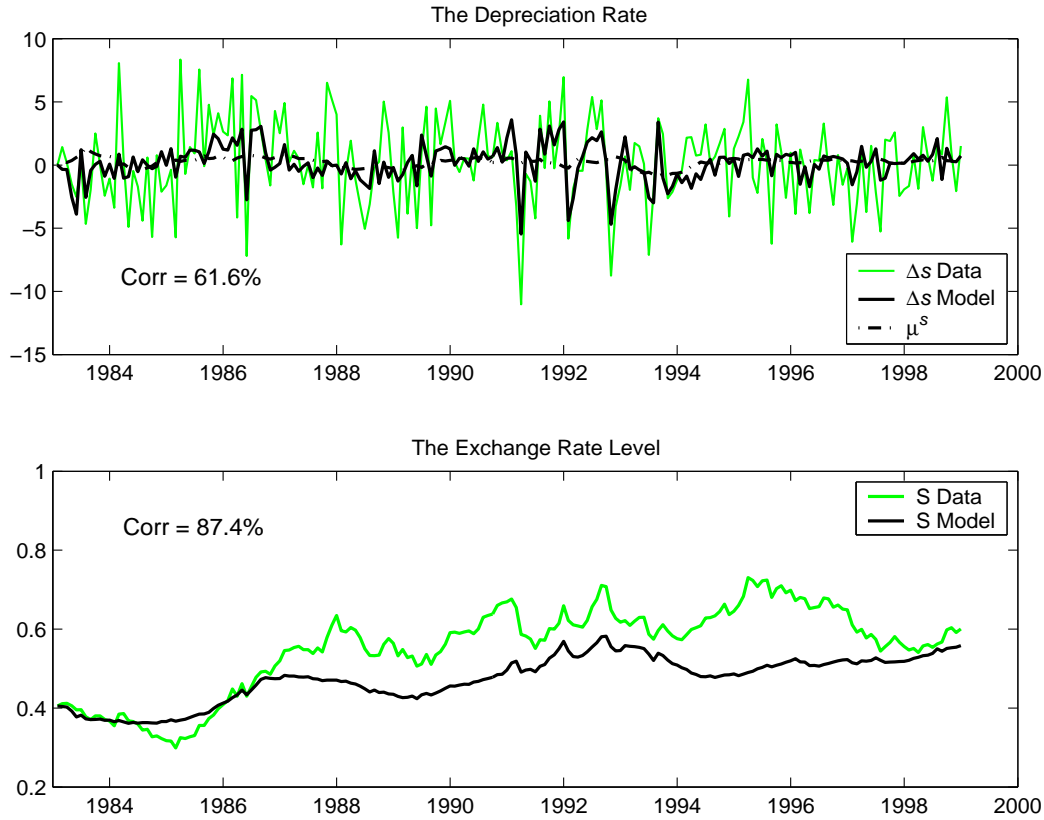
NOTE: I plot the output gap and inflation for the US and Germany used in the estimation. The inflation rate is 12-month change of the Consumer Price Index. The output gap is estimated from Industrial Production Index using HP filter with a smoothing factor of 129,600. The sample period is January 1983 to December 1998 and the data frequency is monthly. The units on the vertical axis are percent.

Figure 2: The Estimated Latent Factors and Demeaned Short Rates



NOTE: In the top panel, I plot the posterior mean of the latent factor  $f_t$  together with the demeaned US short rate from data. I plot the corresponding quantities for Germany in the bottom panel. The latent factors and short rates are all annualized. The units on the vertical axis are percent.

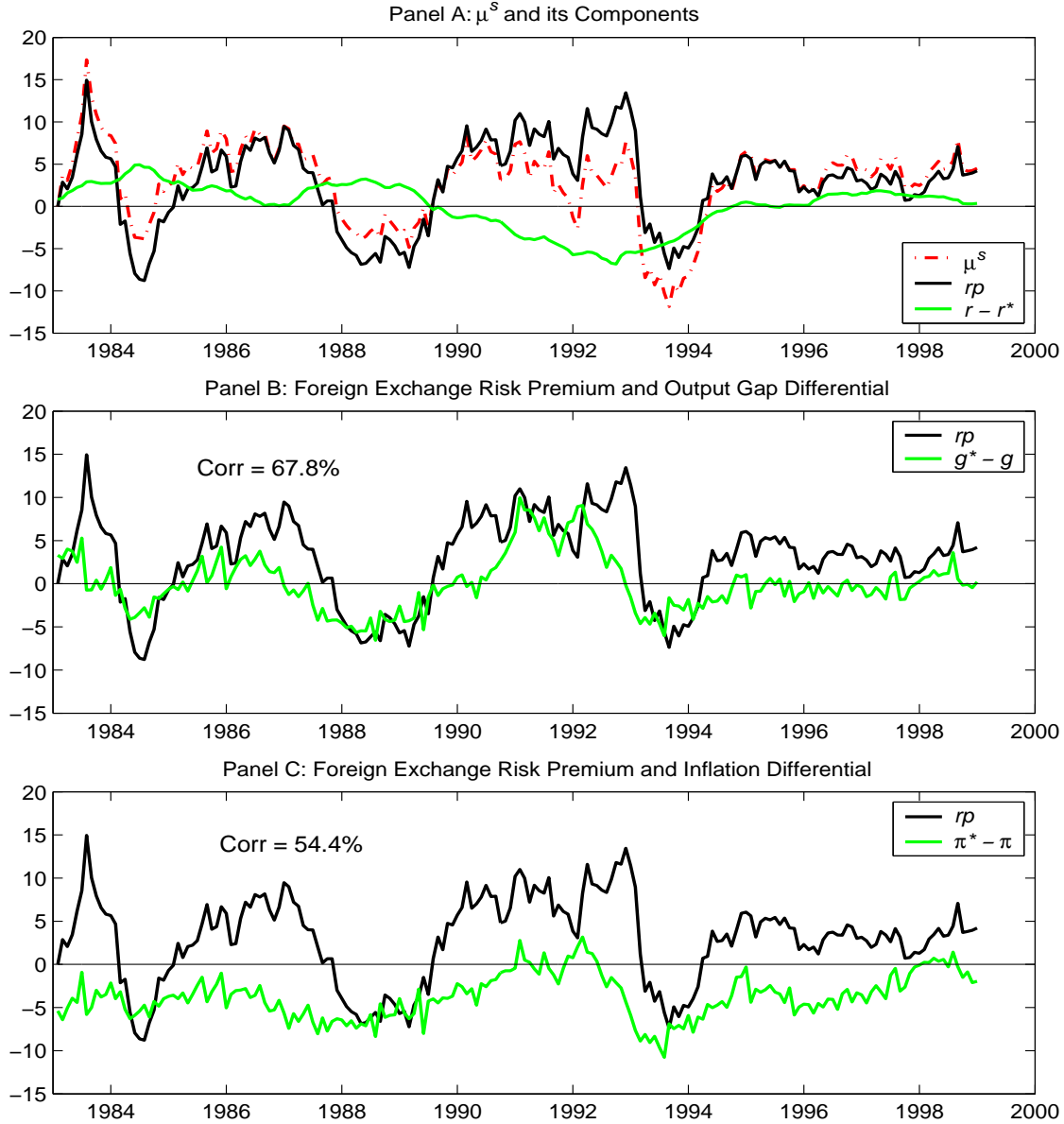
Figure 3: Data and Model-Implied Exchange Rate



NOTE: In the top panel, I plot the posterior mean of the model-implied monthly depreciation rate,  $\widehat{\Delta s}$ , computed using equation (26), exchange rate change in the data,  $\Delta s$ , and the conditional mean of the depreciation rate,  $\mu^s$ . The units on the vertical axis in the top panel are percent. In the bottom panel, I plot the exchange rate level in the data together with the model implied exchange rate level,  $\widehat{S}$  computed using as the exponential of the cumulative sum of posterior mean of model-implied  $\Delta s$  ( $\widehat{S}_t = S_1 + \exp(\sum_{i=2}^t \widehat{\Delta s}_i)$ ), where  $S_1$  is the data exchange rate level at the beginning of the sample period. The frequency of the data is monthly.

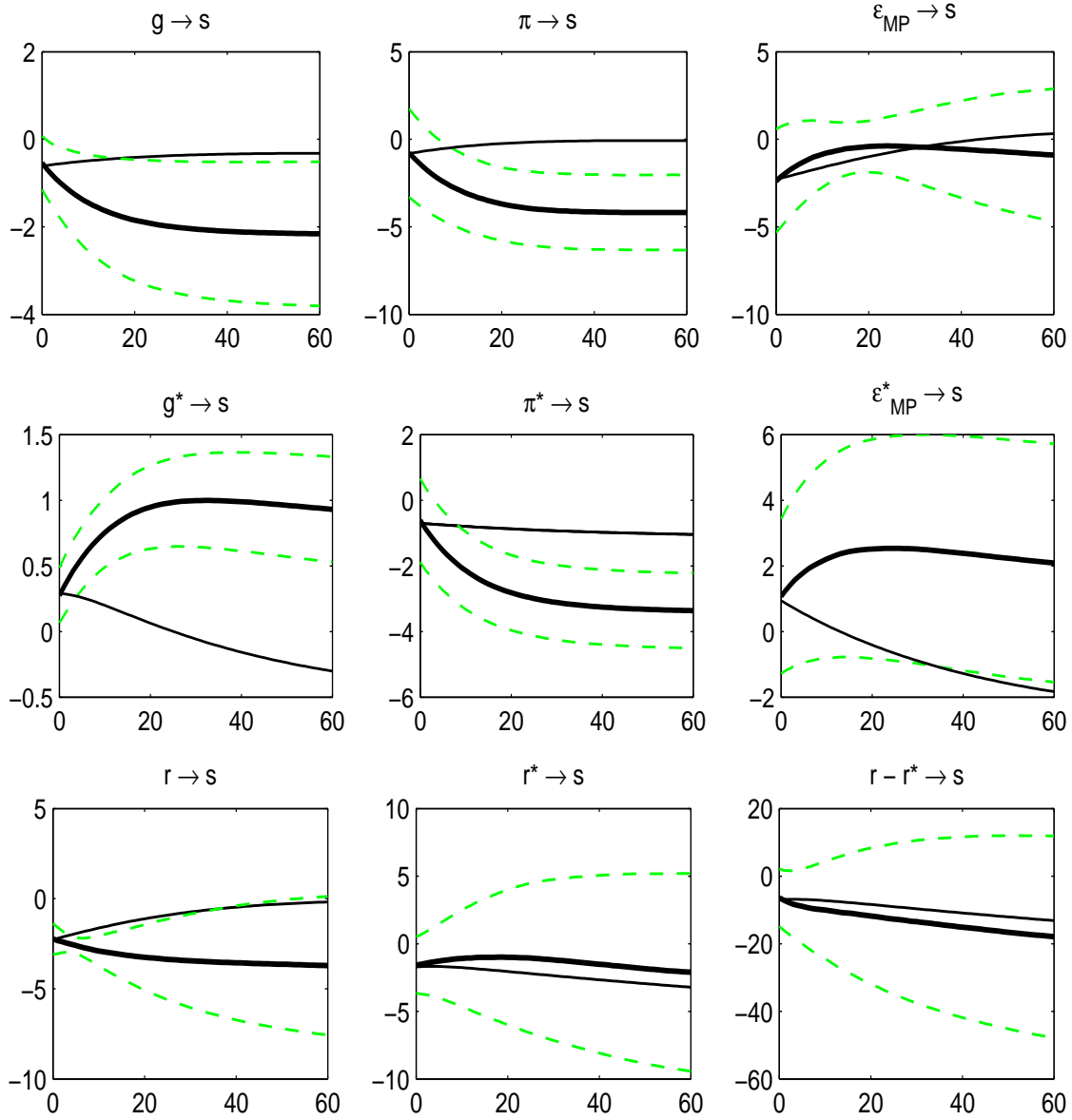


Figure 4: The Foreign Exchange Risk Premium



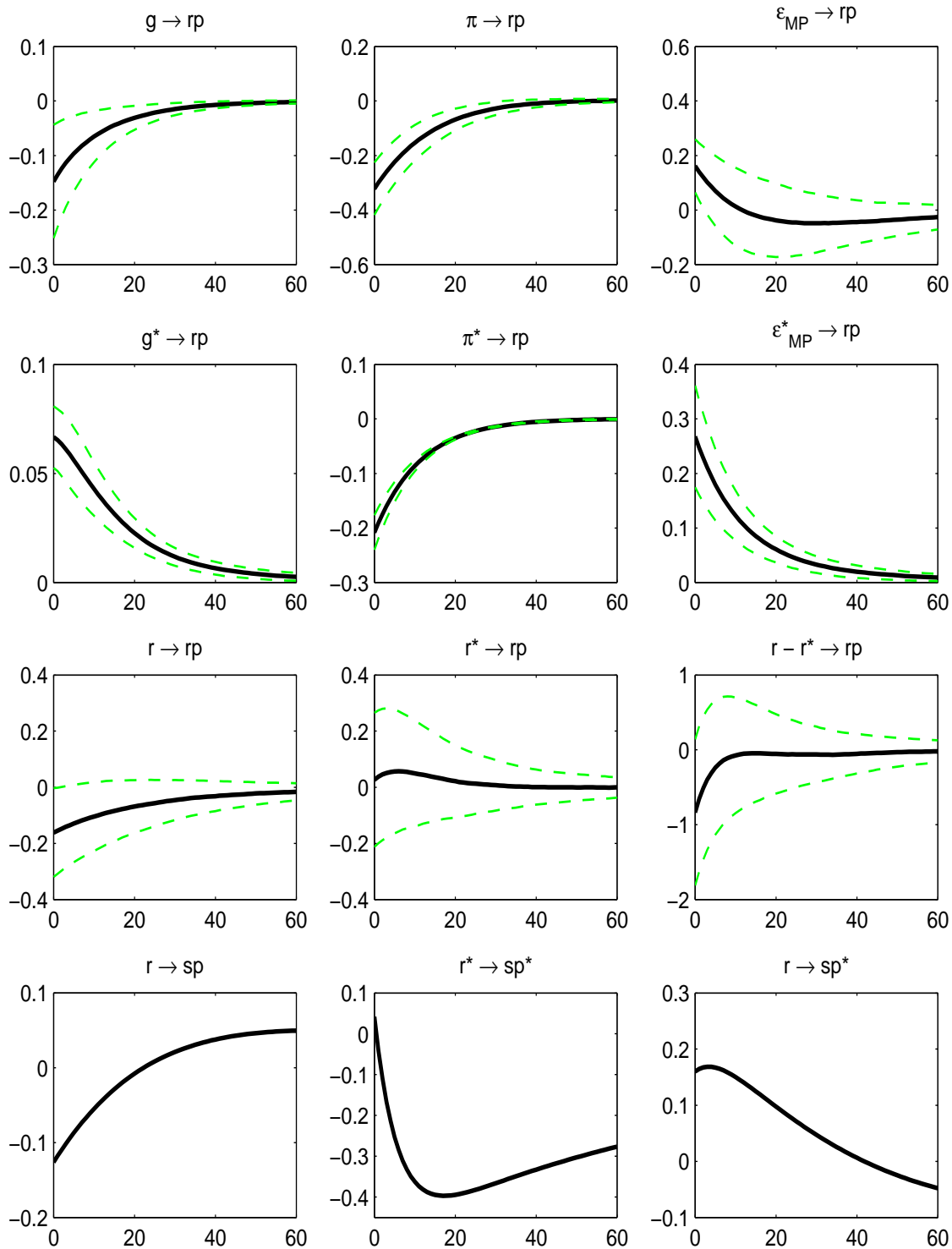
NOTE: In Panel A, I plot the model-implied conditional expected value of exchange rate changes,  $\mu_t^s$ , and its components, the foreign exchange risk premium  $rp_t$  and the interest rate differential  $r_t - r_t^*$ . In Panel B, I plot the foreign exchange risk premium together with the output gap differential between Germany and the US,  $g_t^* - g_t$ . In Panel C, I plot the foreign exchange risk premium together with the inflation differential between Germany and the US,  $\pi_t^* - \pi_t$ . The frequency of the data is monthly. The units on the vertical axis are percent. I plot annualized quantities.

Figure 5: Impulse Responses of the Exchange Rate



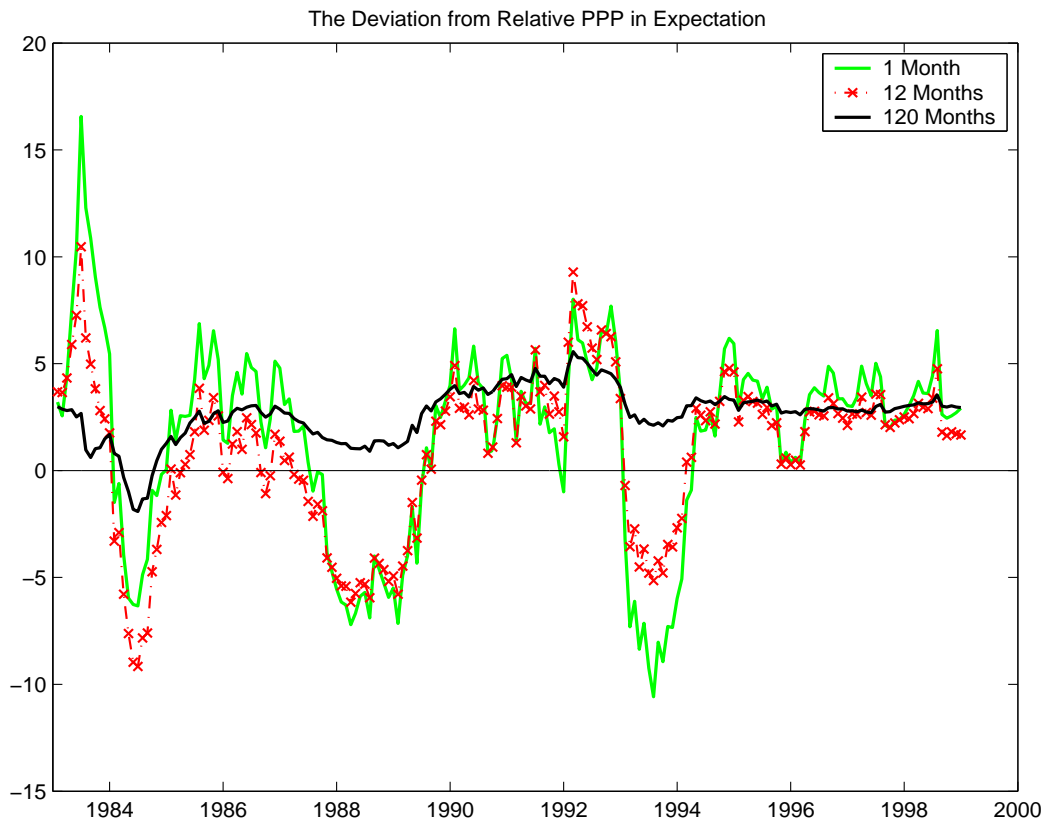
NOTE: I plot the responses of the exchange rate  $s$  to 1% shocks of  $X_t$  and 1% shocks of  $r_t$ ,  $r_t^*$  and  $r_t - r_t^*$  in thick solid lines. The response of  $s$  is conditional on on each observation of  $X_t$  and is computed as the sum of responses of  $\Delta s$  using equation (7). One standard deviation bands are plotted in dashed lines. The responses of  $s$  under constant risk premia are plotted in thin solid lines. One standard deviation bands are plotted in dashed lines. The units on the vertical axis are percent. The horizon axis is lag horizon in months.

Figure 6: Impulse Responses of the Deviation from UIRP and the Term Spread



NOTE: I plot the responses of the deviation from uncovered interest rate parity, i.e., the foreign exchange risk premia  $rp_t$ , to 1% shocks of  $X_t$  and 1% shocks of  $r_t$ ,  $r_t^*$  and  $r_t - r_t^*$  in solid lines.  $rp$  is defined in equation (8). The dashed lines are the one standard deviation bands of the responses of  $rp$  conditional on each observation of  $X_t$ . The last row plots the responses of the term spread,  $sp_t^i = y_t^{i,60} - r_t^i$ , to a 1% shock of  $r_t^i$ . The units on the vertical axis are percent. The horizon axis is lag horizon in months.

Figure 7: The Deviation From Purchasing Power Parity



NOTE: I plot the deviation from relative PPP in expectation over different horizons implied by the model. The deviation from relative PPP in expectation,  $DPPP_n$ , is defined in equation (39). The frequency of the data is monthly. The units on the vertical axis are percent. The quantities in the plot are annualized.