Affine Models of the Joint Dynamics of Exchange Rates and Interest Rates

Bing Anderson

Peter J. Hammond

Cyrus A. Ramezani *

September 17, 2006

*Anderson, bianders@calpoly.edu, and Ramezani, cramezan@calpoly.edu, California Polytechnic State University, Orfalea College of Business, San Luis Obispo, CA 93407; Hammond, hammond@stanford.edu, Stanford University, Department of Economics, 579 Serra Mall, Stanford, CA 94305. We are indebted to an anonymous referee and the editor, Paul Malatesta, for many helpful suggestions that have improved the paper significantly. We thank readers on the doctoral dissertation committee of Bing Anderson (Bing Han at that time) at Stanford - Darrell Duffie, Peter Glynn, Michael Saunders - for insightful feedbacks. The paper also benefits from comments by Kimberly Anderson, Zhiwu Chen, Robert Elliott, Hua He, Maria Kasch-Haroutounian, Qi Li, and participants of the 58th European Meeting of the Econometric Society, the European Financial Management Association 2003 annual conference, and the 2003 Hawaii International Conference on Business. Anderson acknowledges research grants from the California Polytechnic State University and the University of Calgary.

Affine Models of the Joint Dynamics of Exchange Rates and Interest Rates

September 17, 2006

ABSTRACT

This paper extends the affine class of term structure models to describe the joint dynamics of exchange rates and interest rates. In particular, the issue of how to reconcile the low volatility of interest rates with the high volatility of exchange rates is addressed. The incomplete market approach of introducing exchange rate volatility that is orthogonal to both interest rates and the pricing kernels is shown to be infeasible in the affine setting. Models in which excess exchange rate volatility is orthogonal to interest rates but not orthogonal to the pricing kernels are proposed, and validated via Kalman filter estimation of maximal five-factor models for six country pairs.

I. Introduction

Modeling exchange rate movements as diffusion processes dates back to Biger and Hull (1983) and Garman and Kohlhagen (1983). They use geometric Brownian motion with constant exchange rate volatility, along with constant interest rates. As better interest rate models become available, efforts are made to extend these models to include exchange rate dynamics. For example, Amin and Jarrow (1991) modify the Heath-Jarrow-Morton (1992) model of forward interest rates to incorporate exchange rate processes. Later, Nielsen and Saà-Requejo (1993) and Saà-Requejo (1994) generalize the Cox-Ingersoll-Ross (1985) model to a multicurrency environment. More recent models of the joint dynamics of exchange rates and interest rates are seen in Bakshi and Chen (1997), who take a general equilibrium approach, and in Brandt and Santa-Clara (2002), who propose an incomplete market framework.

Recently, the most generally used interest rate models are members or variants of the Duffie-Kan (1996) class of affine term structure models, due to their analytical tractability. This paper studies how to extend the affine class of term structure models to describe the joint dynamics of exchange rates and interest rates.

It is well known that exchange rate volatilities are much higher than the corresponding interest rate volatilities. This necessitates models which partially dissociate exchange rates from interest rates, without violating the fundamental pricing equations that relate exchange rates, pricing kernels, and interest rates. One way of dissociation is the incomplete market approach, which introduces exchange rate volatility that is orthogonal to both interest rates and the pricing kernels. In this paper, such an approach is shown to be infeasible in the affine setting. A different approach, in which excess exchange rate volatility is orthogonal to interest rates, but not orthogonal to the pricing kernels, is proposed here.

These theoretical discussions are collected in Section II of this paper. Dai and Singleton (2000) classify N-factor affine term structure models into $N+1$ subfamilies, according to the number of state variables that directly drive the conditional volatilities of all the N state variables. For example, N-factor models in which n state variables directly drive the conditional volatilities form a subfamily, denoted as $\mathbb{A}_n(N)$. Within each subfamily, there exists a maximal model that nests all models in the subfamily. To validate the theoretical approach proposed in our paper, in Section III, maximal $\mathbb{A}_n(5)$ affine models are fitted for six country pairs. For each country pair, in the best-fit model, excess exchange rate volatility turns out to be orthogonal to interest rates but not to the pricing kernels, without being specified as such a priori. The outcome of the model fitting also sheds light on the issue of common versus local factors (see Ahn (2004), and Mosburger and Schneider (2005)). Section IV concludes.

II. Theory

A. Fundamental Pricing Equations

We begin with a review of the fundamental pricing equations that govern exchange rates, pricing kernels, and interest rates. The setup is discrete-time. More on discrete-time dynamic asset pricing can be found in Duffie (2001). To simplify notation, we limit our discussion to a pair of countries: domestic and foreign.

Let M denote the domestic pricing kernel (also called the state-price deflator, state-price kernel, stochastic discount factor, or state-price density) and M^* the foreign one. Let r_t be the domestic short interest rate for the period from time t to time $t+1$, with r_t^* as its foreign counterpart. Let S be the exchange rate, defined as the price of one unit of foreign currency in units of domestic currency. Let s denote the natural logarithm of S.

Within the two individual countries, the fundamental pricing equation is

$$(1) \qquad E_t\left[\frac{M_{t+1}}{M_t}\, e^{r_t}\right] = 1 = E_t\left[\frac{M^*_{t+1}}{M^*_t}\, e^{r^*_t}\right] \quad .$$

One unit of domestic currency at time t, invested in a domestic interest-bearing account, will become e^{r_t} at time $t+1$. The time t price of the time $t+1$ payoff, e^{r_t}, is found by pricing this amount using the domestic pricing kernel, as $E_t\left[\frac{M_{t+1}}{M_t}\, e^{r_t}\right]$. On the other hand, e^{r_t} at time $t+1$ comes from one unit of domestic currency at time t, hence its time t price is 1. This explains the left half of (1). The right half of (1) is explained similarly using the foreign currency.

Between countries, Backus, Foresi, and Telmer (2001) show that the following fundamental pricing equation has to be satisfied:

$$(2) \qquad E_t\left[\frac{M_{t+1}}{M_t}\,\frac{S_{t+1}}{S_t}\, e^{r^*_t}\right] = 1 = E_t\left[\frac{M^*_{t+1}}{M^*_t}\,\frac{S_t}{S_{t+1}}\, e^{r_t}\right] \quad .$$

A domestic investor who starts with one unit of domestic currency at time t, converts it to the foreign currency at the exchange rate S_t, then saves it in a foreign interest-bearing account, and finally converts back to domestic currency at time $t+1$ at the exchange rate S_{t+1}, will end up having $\frac{S_{t+1}}{S_t}\, e^{r^*_t}$ domestic currency units at time $t+1$. Then one derives the left half of (2) by reasoning similar to that used to derive (1). The right half of (2) is explained in much the same way, considering a foreign investor who starts with one unit of the foreign currency.

Pricing equation (1) relates interest rate processes to pricing kernels. Equation (2) further relates exchange rate processes to both pricing kernels and interest rates. A key question is: given (1), how should exchange rate processes be specified in a model, in order to satisfy (2)?

With (1) given, a sufficient but not a necessary condition for (2) to hold is:

$$(3) \qquad \left(\frac{M^*_{t+1}}{M^*_t}\right) \,/\, \left(\frac{M_{t+1}}{M_t}\right) \;=\; S_{t+1} \,/\, S_t \quad .$$

Thus, specifying S according to (3) gives one answer to the above question.

Are there other ways to specify exchange rates? Because (3) is a sufficient but not a necessary condition for (2), there should be ways to specify S that satisfy (2) but not (3). Then, which of the two approaches of specifying exchange rates, according to (3) and not according to (3), should one take? Or might both work? We address these questions in the next two subsections.

B. Volatilities of Exchange Rates and Interest Rates

Backus, Foresi, and Telmer (2001) study two special cases of affine models involving exchange rates. In both cases, exchange rates are specified according to (3). Their models have difficulties in accounting for the empirical characteristics of exchange rate and interest rate dynamics. The discussion on page 297 of their paper suggests that the difficulties they encounter are due to the much higher volatility of the exchange rate, compared to that of the interest rates. Equation (1) links the interest rates to the pricing kernels; (3) in turn links the pricing kernels to the exchange rate in a very rigid fashion. Hence, the close association of the exchange rate and the interest rates via (1) and (3) in these models seems to explain why the models have difficulties in reconciling the high volatility of the exchange rate with the low volatility of the interest rates. In other words, once the exchange rate and interest rates are associated in the way as they are in these models, they must both have high volatilities, or both have low volatilities.

These analyses seem to indicate that one should specify exchange rates using processes which satisfy (2) but not (3), in order to account for the empirical fact that exchange rate volatilities are much higher than the corresponding interest rate volatilities. This incomplete market approach is seen in Brandt and Santa-Clara (2002).

As a convention on notation, a subscript, such as M_t, is used for time in a discrete-time setting, whereas an argument of a function is used for time in a continuous-time setting—$M(t)$ for instance. Adapted to this notational convention, equation (24) on page 176 of the Brandt and Santa-Clara (2002) paper specifies the exchange rate as

$$S(t) = \frac{M^*(t)}{M(t)} O(t) \quad ,$$

where $O(t)$ "is a martingale that is orthogonal to" $M(t)$, $M^*(t)$, "and all domestic and foreign assets". The continuous-time equivalent of (3) would specify the exchange rate as $S(t) = \frac{M^*(t)}{M(t)}$. The introduction of the extra $O(t)$ process, and consequently the specification of the exchange rate not according to (3), do partially dissociate the exchange rate process from the interest rate processes. As a result, this model is able to accommodate high exchange rate volatility without running into difficulties with the low interest rate volatility. However, from a foreigner's point of view, the same exchange rate expressed as one unit of domestic currency in units of the foreign currency is

$$\frac{1}{S(t)} = \frac{M(t)}{M^*(t)} \frac{1}{O(t)}$$

where $1/O(t)$ by the same reasoning also needs to be a martingale satisfying similar orthogonality conditions. Although it might be possible in special cases, it is generally not true that when $O(t)$ is a martingale, $1/O(t)$ is also a martingale.

Are these difficulties particular to the specifications of the Brandt and Santa-Clara model, or do they exist more generally, for the entire approach of specifying the exchange rate according to (2) but not (3)? We answer this question next, within the affine setting.

C. Constraints on Exchange Rate Dynamics

Before proceeding further, we need to clarify what we mean by "within the affine setting". For single-country interest rate term structure models, the meaning of "the affine setting" is straightforward. From Duffie and Kan (1996), the vector X of N state variables (sometimes also referred to as factors) evolves according to an affine diffusion under the risk-neutral measure, and short interest rate r is an affine function of the state variables. Dai and Singleton (2000) impose conditions on the parameters to ensure admissibility and identification.

The simplest way to extend the above framework to include exchange rates, is to specify log of the exchange rate, s_t, also as an affine function of the state variables:

$$s_t = a + b^T X_t \quad , \tag{4}$$

where a is a scalar and b is an N-vector of constant coefficients. In this way, the analytical tractability of the affine term structure models in pricing bonds and derivatives readily extends to the exchange rates. Also, this is a rather basic requirement, in that if s_t can no longer be expressed as an affine function of the state variables, the setting can hardly be called affine anymore.

In the extended models, the law governing state variable dynamics remains the same as in the single-country models; short interest rate in each country remains an affine function of the state variables; the restrictions Dai and Singleton (2000) placed on parameters remain intact for the state variable dynamics and for the short interest rate of one of the countries. No restrictions are placed on the additional parameters introduced for specifying short interest rates in the rest of the countries and the exchange rates. In other words, we take the canonical single-country model of Dai and Singleton, together with all their restrictions on the parameters, then add one affine function of the state variables for each additional country's short interest rate, and one affine function of the state variables for each log exchange rate process. We place no restrictions on the parameters in these affine functions we add on. We continue to call these extended models canonical.

Obviously, with the state variable dynamics and all constraints on its parameters remaining intact, admissibility of the model, guaranteed by Dai and Singleton in the single-country setting, has not been upset in such an extension to the multi-country setting. However, we need to check that the extended model, with no constraints on any of the additional parameters introduced in the extension, is still identified. We do so by demonstrating that there exists at least one estimation procedure, for which all the parameters in the extended model are identified. Such an estimation procedure is heuristically found by following the steps of the extension process itself.

Before any extension, for the single-country canonical affine term structure model, identification is guaranteed by Dai and Singleton (2000). One can first estimate this single-country model. As a by-product, one obtains estimates of the state variable time series. After the extension to a multi-country setting, for each additional country, one regresses its short interest rate against the state variable time series one already has, to estimate parameters in the affine function of the state variables for the short interest rate of that country. Similarly, parameters in (4) are uniquely obtained via regression. Parameters for the market prices of risk can be determined by using the model to price coupon bonds or swap coupon rates and minimizing the pricing errors.

Although this estimation procedure is not the best possible (for example, the state variable dynamics is based on only one country's term structure, not all exchange rates and term structures in all countries as it ideally would be), it is a valid procedure in which all the data are utilized, and this estimation procedure identifies all the parameters for the extended model. Therefore, the conditions we impose on the parameters of the extended canonical models are sufficient for both the admissibility and the identification of these multi-country models.

The specification of log exchange rates as (4) is very general. To see this, one only needs to recall that for the entire single-country affine class of term structure models, short interest rate is specified as an affine function of the state variables exactly like (4). If such a specification can describe all the different short interest rate dynamics within the affine class of term structure models, it certainly can accommodate a wide range of exchange rate dynamics, too.

To our knowledge, none of the previous exchange rate models in the literature specify exchange rate dynamics in this way. Instead, existing models use (3) as a starting point in specifying exchange rates. Some specify exchange rates as (3) itself (for example, Backus, Foresi, and Telmer (2001)), some as (3) with additional terms (for example, Brandt and Santa-Clara (2002)). In some sense, this is not surprising. All exchange rate and interest rate dynamics have to obey the fundamental pricing equations (1) and (2). It is much easier to verify that the dynamics satisfy (1) and (2) when exchange rates are specified with (3) as a starting point. This naturally invites a question about our specification: in order to satisfy (1) and (2), how do the exchange rates in our specification relate to (3)?

To answer this question, we first construct a certain domestic asset, the price of which at time t, according to the left half of (1), is 1. Pricing this domestic asset with the right half of (2) gives a set of constraints on the parameters of the affine model. Similarly, we can construct a particular foreign asset, which has a price of 1 at time t according to the right half of (1), and when priced with the left half of (2) gives a second set of constraints on the model parameters. Comparing these two sets of constraints, and taking into account the restrictions on the model parameters, we arrive at the conclusion that all exchange rate dynamics in our extended canonical affine models must satisfy (3). In other words, in any canonical affine model, an exchange rate process specified in the format of (4) also conforms to (3).

This result is summarized in Proposition 1, proved in the discretized affine setting in Appendix A. Here, we work with the discretized instead of the continuous-time affine models, because of the analytical tractability of the former. Nevertheless, the insights obtained in discrete time readily generalize to continuous time.

Proposition 1 *All canonical affine models of exchange rates and interest rates must conform to (3).*

Proposition 1 focuses our attention back to models which specify exchange rates according to (3). To make such a specification work empirically, we need to find a way to partially dissociate the exchange rate process from the interest rates. The method is to partially break the link between the interest rates and the pricing kernels in (1), therefore dissociate interest rates from the exchange rate, despite the tight link between the exchange rate and pricing kernels via (3). We do so by introducing factors which affect the pricing kernels but not the interest rates.

In the affine setting, the short interest rates are specified as affine functions of the state variables. As can be seen in more details in the next subsection, volatility of the pricing kernels are determined by a *different* set of affine functions of the state variables. If we set some coefficients in the affine functions for interest rates to zero, it is possible for some state variables to contribute to the pricing kernel dynamics and consequently the exchange rate dynamics, but not directly to the dynamics of short interest rates. Henceforth, these state variables which contribute to exchange-rate but not interest-rate dynamics will be referred to as "extra". Once exchange rate and interest rate movements have been partially dissociated by these "extra" state variables, exchange rate volatility in a model acquires the freedom to become higher than interest rate volatility. In other words, these "extra" state variables introduce excess exchange rate volatilities that are orthogonal to the interest rates, but not orthogonal to the pricing kernels.

D. Canonical Model Specification

The discussions so far lead us to consider models which satisfy (3) and also have "extra" state variables for exchange rate dynamics alone. We refer to these models as canonical, also because they conform to equations (1) and (2), two canons of international finance. Next, we wrap up this section on theory with a specification of the canonical model in continuous time.

In the most general one-country continuous-time affine setting, there is an equivalent martingale measure under which the vector of state variables $X(t)$ follows an affine diffusion process

$$dX(t) = \mu(X(t),t)dt + \sigma(X(t),t)dB(t)$$

where the elements of both the vector $\mu(X(t),t)$ and the matrix $(\sigma(X(t),t)\sigma(X(t),t)^T)$ are affine functions of the state variables $X(t)$ (see Chapter 7 of Duffie (2001) for more details) and $B(t)$ is standard Brownian motion. The density process $\xi(t)$ of the equivalent martingale measure follows

(5) $$d\xi(t) = -\xi(t)\eta(t)dB(t)$$

where $\eta(t)$ is the market-price-of-risk process.

From Duffie (2001), Section 6F, the relationship between the density process $\xi(t)$ of the equivalent martingale measure and the pricing kernel $M(t)$ is

(6) $$\xi(t) = \exp\left\{\int_0^t r(s)ds\right\} \frac{M(t)}{M(0)} \quad .$$

In our setting with more than one country, by (5) and (6), it can be shown that country i's pricing kernel $M^{(i)}(t)$ is

$$\frac{dM^{(i)}(t)}{M^{(i)}(t)} = -r^{(i)}(t)dt - \eta^{(i)}(t)\, dB(t) \tag{7}$$

where $\eta^{(i)}(t)$, the market-price-of-risk process for that country, may be defined as

$$\eta^{(i)}(t) = \left(\sqrt{\Sigma(t)}\lambda^{(i)}\right)' \tag{8}$$

with $\lambda^{(i)}$ being an N-vector of constants.

With $\eta^{(i)}(t)$ so defined, the dynamic process followed by the state variables $X(t)$ under the actual probability measure is

$$dX(t) = \Psi(\bar{X} - X(t))dt + \sqrt{\Sigma(t)}\, dB(t) \tag{9}$$

where $\Sigma(t)$ is an $N \times N$ diagonal matrix with

$$\Sigma_{ii}(t) = g_{i0} + \sum_{j=1}^{N} g_{ij} X_j(t) \quad . \tag{10}$$

For each country i, its short interest rate is

$$r^{(i)}(t) = \rho_{i0} + \sum_{j=1}^{N} \rho_{ij} X_j(t) \tag{11}$$

where each ρ_{ij} is constant, and some of the ρ's can be zeros. Especially, the ρ's corresponding to the "extra" state variables are all zeros in (11), so that these state variables do not enter the interest rate dynamics directly. However, the "extra" state variables may contribute to the dynamics of the pricing kernels and the exchange rate via non-zero g coefficients in (10). The exchange rate $S^{(ij)}$, defined as the price of one unit of currency i in units of currency j, is obtained via (4). Namely, using $s^{(ij)}$ to denote the natural logarithm of $S^{(ij)}$, we have

$$s^{(ij)}(t) = a^{(ij)} + b^{(ij)T} X(t) \tag{12}$$

where $a^{(ij)}$ is a scalar and $b^{(ij)}$ is an N-vector of constant coefficients.

As explained in the previous subsection, we impose conditions from Dai and Singleton (2000) on the parameters related to the state variable dynamics (9) — the coefficients g_{ij} , the $N \times N$ matrix Ψ, and the N-vector $\bar{X}$. We also impose the Dai and Singleton restrictions on the ρ's in (11) for one of the countries, let's say, $\rho_{10}, \rho_{11}, \rho_{12}, \ldots, \rho_{1N}$ for the first country. For the parameters ρ_{ij} with $i > 1$, that is, parameters in (11) for the short interest rates of the rest of the countries, we impose no constraints. There are also no constraints on the parameters $a^{(ij)}$ and $b^{(ij)}$ in (12).

III. Empirical Analysis

The theoretical structure proposed above suggests that affine models of the joint dynamics of exchange rates and interest rates need to have "extra" state variables which contribute to exchange-rate but not directly to interest-rate dynamics.

One way to empirically validate the theoretical considerations is to show that models with these "extra" state variables provide better fit of observed data than those without. Given the poor empirical performance of models without the "extra" state variables, such an improvement, however, will hardly come as a surprise.

A related, but more interesting and more challenging, question is: if we don't restrict any of the ρ's in (11) to be zero, and therefore don't specify any "extra" state variables a priori, will the best-fit models turn out to have these "extra" state variables nevertheless?

Currently, state-of-the-art empirical analyses with one-country affine term structure models employ three state variables. Recently, Mosburger and Schneider (2005) suggest that even for two-country affine term structure models, three state variables in total for the two term structures are sufficient. For the UK-US data in particular, they discover that all three state variables are common factors for both countries instead of local factors particular to an individual country. This partially explains why three factors are sufficient.

Taking these results into consideration while allowing some extra room for flexibility, we decide to fit five-factor canonical affine models, or $\mathbb{A}_n(5)$ in the terminology of Dai and Singleton (2000), for the joint dynamics of exchange rates and interest rates. They systematically study $\mathbb{A}_n(3)$ for one-country term structure models. To our knowledge, no study exists that examines $\mathbb{A}_n(4)$ or beyond. Hence, successfully fitting these $\mathbb{A}_n(5)$ models and studying their performance is interesting in itself.

The Kalman filter has been used by many authors to estimate term structure models (Babbs and Nowman (1999); Duan and Simonato (1999); Duffee (1999); De Jong (2000); Dewachter and Maes (2001); Chen and Scott (2003); among others), and has been shown to have good small-sample properties (Duffee and Stanton (2001)). Because we express the exchange rate in (4) as an affine function of the state variables, similar to the way short interest rates are defined and the way zero-coupon bond yields are solved in affine term structure models, the existing Kalman-filter quasi-maximum likelihood (QML) estimation techniques for one-country term structure models can be applied to the joint dynamics models here with virtually no modification.

To reduce the number of parameters that need to be estimated, in addition to the usual assumption of a diagonal covariance matrix for measurement errors, Brennan and Xia (2003) and Tang and Xia (2005) further use just one parameter for the entire measurement error covariance matrix, either assuming all the variances to be constant, or assuming them to be inversely proportional to the maturities. We use a method that is similar in spirit but allows more flexibilities in specifying the variances. We divide the interest-rate maturities into three non-overlapping groups of short-, medium-, and long-terms, with short-term being up to and including 1 year, and medium term up to and including 10 years. We assume the variances of measurement errors are constant within each maturity group, which leaves room for variation across groups. We use ω_s, ω_m, ω_l to denote the standard deviations of measurement errors of the short-, medium-, and long-maturities, respectively. In addition, the standard deviation of the measurement errors for exchange rate is denoted as ω_e.

A. The Data

Both exchange rate and interest rate data are downloaded from Datastream. The countries involved in this study are the United States, the United Kingdom, Germany, and Japan. Daily data of 1-, 2-, 3-, 4-, 5-, 6-, 7-, 8-, 9-, 10-, 11-, and 12-month LIBOR, and of 1-, 2-, 3-, 4-, 5-, 6-, 7-, 8-, 9-, 10-, 12-, 15-, 20-, 25-, and 30-year plain-vanilla fixed-for-floating interest-rate swap coupon rates, are obtained for the time period extending from May 1, 1998 to August 5, 2005. Some of the swap time series are not available prior to May 1, 1998. Data from every Wednesday are then used to construct weekly time series for all maturities.

We use 1-, 2-, 4-, 6-, 8-, 10-, and 12-month LIBOR, and 2-, 4-, 6-, 8-, 10-, 12-, 15-, 20-, 25-, and 30-year swap coupon rates for model estimation. The rest of the weekly interest-rate time series are reserved for testing the out-of-sample performance of the fitted models.

Because of the special structure of counterparty risks in swap contract execution, the default-related component in swap coupon rates has been shown to be rather small (Duffie and Huang (1996)). As is customary in the literature, we ignore the default risk in LIBOR and swap. Because zero-coupon bond yields are solved as affine functions of the state variables in affine models, in estimation of these models using Kalman filters, it is convenient to work with zero-coupon bond yields. We use iteration to fit Nelson-Siegel (1987) zero-coupon bond yield curves to the LIBOR-swap data.

B. Models Estimated

Out of the four countries, one can construct six country pairs. When Dai and Singleton (2000) carry out an empirical analysis of the $\mathbb{A}_n(3)$ family of one-country canonical affine term structure models, they focus on $\mathbb{A}_1(3)$ and $\mathbb{A}_2(3)$. The reason for ignoring $\mathbb{A}_0(3)$ is that it implies constant conditional volatilities of the state variables, which is apparently counter-factual. The reason for not empirically fitting $\mathbb{A}_3(3)$ is that it cannot account for negative unconditional correlations among the state variables, and the U.S. interest rate data they use likely necessitate such negative correlations. The problem with $\mathbb{A}_0(3)$ also applies to $\mathbb{A}_0(5)$, and it does not depend on which country's data we are fitting the model to. However, the problem of $\mathbb{A}_3(3)$, and similarly, of $\mathbb{A}_5(5)$, may not exist for all countries' data, as some may not call for the negative correlations. Therefore, we decide to fit the maximal $\mathbb{A}_1(5)$, $\mathbb{A}_2(5)$, $\mathbb{A}_3(5)$, $\mathbb{A}_4(5)$, and $\mathbb{A}_5(5)$ models, for each of the six country pairs.

Even for models with only three state variables, given the number of parameters involved and the non-trivial shape of the likelihood function surface in the parameter space, all one can hope is to find a good local maximum of the likelihood function, instead of the global maximum. Preliminary experimentation with different starting parameter values suggests that there is a positive correlation between the likelihood value at the starting point of the optimization search, and the likelihood value at the resultant maximum. Therefore, for each of the models we estimate, within a reasonable space of the parameter values, we randomly sample 100,000 points of parameter combinations and evaluate the likelihood function at each point. The 10 points with the best likelihood values are selected as the starting points for search of a parameter combination that maximizes the likelihood function, using the Nelder-Mead optimization algorithm. The best among the 10 results, unless it is something obviously unreasonable, is reported as the final estimate of the model parameters. Selecting the best starting points, as well as optimization from each starting point, each takes about a day to run on a personal computer. Given that we estimate 5 models each for 6 pairs of countries, the total computational time for the empirical analysis is approximately an entire year, if performed on one computer.

Table 1 shows the log likelihood values, Akaike Information Criteria (AIC), Bayesian Information Criteria (BIC), and the number of parameters estimated, for the maximal $\mathbb{A}_1(5)$ through $\mathbb{A}_5(5)$ models of the six country pairs. For these models, log likelihood values, AIC, and BIC, all indicate that the best fit is the maximal $\mathbb{A}_1(5)$ model for each of the country pairs, with the exception of the Japan - U.S. pair which has the maximal $\mathbb{A}_2(5)$ model as its best fit. As mentioned in the Introduction, from Dai and Singleton (2000), the value of n in $\mathbb{A}_n(5)$ indicates the number of state variables that directly drive the conditional volatilities of the state variables themselves, via Σ in (9). Because it is the square root of Σ that enters (9), all the n state variables are guaranteed by model construction to be non-negative. There is a trade-off: the bigger the value of n, the richer the conditional volatility structure of the state variables, but at the same time the less flexible the model becomes in specifying conditional correlations of the state variables. Our results indicate that in this trade-off, flexibility in specifying the conditional correlations is more important than flexibility in specifying the conditional volatilities of the state variables.

We focus on the best-fit model of each country pair for the remainder of this section. For each of the six best-fit models, the time series of smoothed estimates of the state variables are constructed, which consequently yield the model-fitted values of each of the observed data time series. For the interest rates of each country, for each week in the time series and each maturity across the term structure, the absolute value of the error between the fitted value and the observed is calculated. The mean of these absolute fitting errors (MAE) and the corresponding root mean squared errors (RMSE) are reported in Table 2, in units of basis points. In Table 2, we also report the MAE and RMSE of the log exchange rate, expressed as a fraction of the mean magnitude of the log exchange rate itself, to facilitate comparison between models for different country pairs. Clearly, the models provide reasonable fits of interest rates and exchange rates in most cases, although the fit on interest rates are generally better than on exchange rates. The reason, besides the fact that exchange rates are more volatile, is that there are 17 observed interest rate time series for each country, corresponding to the 17 different maturities we use across the term structure, but there is only 1 observed time series for the exchange rate. The Kalman filter estimation gives equal attention to each of the observed time series. As a result, the interest rate term structures weigh more in the likelihood function than the exchange rate.

In Table 3, we examine how fitting errors vary across different maturities. To save space, we present results for each country taken from only one model, instead of all the 12 appearances of the countries in the six models. The results in Table 3 show that absolute fitting errors are generally larger for long maturities than short and medium maturities.

We do not use the 3-, 5-, 7-, 9-, and 11-month LIBOR, and the 1-, 3-, 5-, 7-, and 9-year swap coupon rates when fitting the models. Zero-coupon bond yields are extracted from these interest rate data, for testing the out-of-sample performance of the fitted models. Similar to Table 3, statistics on the out-of-sample fitting errors are presented in Table 4. For every currency, comparable maturities from Tables 3 and 4 tend to have comparable fitting errors. It is reassuring that the models perform just as well out of sample as in sample.

For the best-fit model of each of the six country pairs, asymptotic standard errors of the parameter estimates are calculated using the covariance matrix for quasi-maximum likelihood estimation proposed by White (1982). The parameter estimates, as well as their standard errors, are presented in Table 5 for the Japan - US pair, the only country pair for which $\mathbb{A}_2(5)$, instead of $\mathbb{A}_1(5)$, is the best fit. The state variables contribute directly to the short interest rates via the parameters ρ, and to the exchange rate dynamics via the parameters b.

To see which state variables enter each of the exchange rate and interest rate dynamics, in Table 6, for each of the ρ and b parameters, we record if that parameter estimate is significantly different from zero at the 1% level, for all six country pairs. Consistent with the theory, we can see from Table 6 that in the best-fit models, there are indeed "extra" state variables which contribute to the exchange rate dynamics, but not directly to interest rate dynamics, even without specifying the models as such a priori. The "extra" state variable is X_3 in the Japan-Germany pair, X_4 for the Germany-UK pair, X_3 for the Germany-US pair, and X_5 for the Japan-UK pair. In the Japan-US pair, X_4 is a state variable that enters the exchange rate dynamics, but not directly the short interest rate of Japan, and X_5 is a state variable that enters the exchange rate dynamics, but not directly the short interest rate of the US. Similarly, in the UK-US pair, X_2 is a state variable that enters the exchange rate dynamics, but not directly the short interest rate of the US.

What happens to the UK short interest rate in the best-fit model for the UK-US pair? From Table 6, all state variables have non-zero contribution to the UK short interest rate, which at a first glance seems to be an exception to the general pattern in that table and to our theoretical framework as well. Although one exception out of 12 countries in the table may seem negligible, we investigate this case further.

Table 7 details the parameter estimates, as well as their standard errors, for this UK - US pair. In Figure 1 through 3, we plot the observed log exchange rate in US dollars per British pound, and the smoothed estimates of two state variables X_2 and X_5 in the UK - US model.

From Table 6, there are two state variables, X_2 and X_5, which do not contribute directly to the US interest rate. Thus one or both of these two variables are likely responsible for the excess exchange rate volatility. Also can be seen from Table 6 is that X_5 does not have a significant contribution to the exchange rate dynamics. So it is X_2 alone which is responsible for modeling the excess exchange rate volatility. Because X_2 also enters the UK short interest rate, it brings much of the high exchange rate volatility into the UK short interest rate in the model, which may create a problem. X_5 is interesting because it enters neither the US short interest rate, nor the log exchange rate, but instead just the UK short interest rate, as seen in Table 6. This means X_5 may provide an additional dimension of freedom in modeling the UK short interest rate, which can be used to cancel out the excess exchange rate volatility brought in by X_2. This conjecture is supported by the correlation coefficient between X_2 and X_5 of -0.34. But a more definitive support comes from the correlation coefficient between the log exchange rate and $(\rho_{2,GBP} \cdot X_2 + \rho_{5,GBP} \cdot X_5)$. After all, it is neither X_2 alone nor X_5 alone, but $(\rho_{2,GBP} \cdot X_2 + \rho_{5,GBP} \cdot X_5)$, that enters the UK short interest rate in the model. The correlation coefficient between the log exchange rate and $(\rho_{2,GBP} \cdot X_2 + \rho_{5,GBP} \cdot X_5)$ is a mere -0.04. In other words, although X_2 by itself brings excess exchange rate volatility into the UK short interest rate, much of this excess volatility is canceled out by X_5, so that when we look at the net impact of X_2 and X_5 together on the UK interest rate, very little excess exchange rate volatility has been brought into the interest rate. In that sense, even for the UK short interest rate in the US-UK pair, we do still have an “extra” state variable that contributes to exchange rate dynamics but not directly to the interest rate. It is an “extra” state variable synthesized out of X_2 and X_5, in this rather intriguing way.

Ahn (2004) shows that the existence of factors or state variables local to a particular country, as opposed to factors common to both countries, is necessary for investors to benefit from international diversification of a bond portfolio. Mosburger and Schneider (2005), on the other hand, fit three-factor two-country affine models to the UK - US data, and conclude that all the three factors in their models are common instead of local factors. This paper directly contributes to the resolution of this issue. From Table 6, it can be seen that common factors or common state variables are indeed more common than most would think, supporting the findings of Mosburger and Schneider (2005). Nevertheless, local factors do exist in many, although not all, cases, consistent with Ahn (2004).

IV. Conclusion

This paper addresses issues in extending the affine class of term structure models to a multi-country setting to describe the joint dynamics of exchange rates and interest rates. We emphasize the need to adequately model the high volatility of exchange rates vis-à-vis the low volatility of interest rates. Even though this need itself may not come as a great surprise to researchers in this field, exactly how to correctly model the volatilities is not clear from the previous works.

Contrary to existing beliefs in the literature, we show that the feasible choices for exchange rate specification in the extended affine models are surprisingly narrow. Namely, in order to satisfy the fundamental pricing equation (2), one has to specify exchange rates according to (3). To adequately model the volatility of exchange rates and interest rates despite of such an exchange rate specification, we propose to partially dissociate interest rates and exchange rates through fundamental pricing equation (1), by introducing "extra" state variables, and hence excess exchange rate volatilities, that are orthogonal to interest rates but not orthogonal to the pricing kernels. When we fit maximal $\mathbb{A}_n(5)$ two-country extended affine models to data of six country pairs, all of the best-fit models turn out to have these "extra" state variables, without being specified as such a priori.

A very general class of analytically tractable models for the joint dynamics of exchange rates and interest rates can open doors to many interesting future investigations.

In particular, the canonical model we formulate is general for multiple countries. Using multi-country Cox-Ingersoll-Ross models, Hodrick and Vassalou (2002) suggest that multi-country models explain the dynamics of interest rates and exchange rates better than two-country models. It will be interesting to see how their findings generalize in maximal multi-country affine models. The impact of third-country factors may also be unveiled — for example, whether the US interest rate has any role to play in the exchange rate between the Japanese yen and the British pound.

In addition, better exchange rate models can more accurately price derivatives on exchange rates. Choi and Hauser (1990) find that the term structures of interest rates can impact currency option prices. Taking into account both exchange rates and interest rates when pricing currency derivatives is a task the models proposed in this paper can easily handle. There also exist derivatives that naturally involve both exchange rates and interest rates and therefore genuinely require a joint dynamics model to price — for example, cross-currency spread options and cross-currency swaptions.

In this paper, what we have extended to a multi-country setting with exchange rates, are the "completely affine" term structure models. The completely affine models are only a special case of the "essentially affine" models of Duffee (2002). Due to the more sophisticated specification of the market price of risk process in the essentially affine models, the pricing kernels in the essentially affine setting may no longer assume the log-linear form of the completely affine setting. Thus, extending our work to the essentially affine setting could be a project for future research that promises to be both very interesting and very challenging.

Appendix A. Proof of Proposition 1

The Duffie–Kan (1996) class of continuous-time affine interest rate term structure models was discretized in Backus, Foresi, and Telmer (2001). There is a vector X of N state variables which evolves according to the law

$$X_{t+1} - X_t = \Psi(\bar{X} - X_t) + \sqrt{\Sigma_t}\,\varepsilon_{t+1} \quad . \tag{A-1}$$

Here ε is an N-vector of independent Normal $(0,1)$ disturbances, and Σ_t is an $N \times N$ diagonal matrix with elements given by the affine functions $\Sigma_{ii,t} = g_{i0} + \sum_{j=1}^{N} g_{ij} X_{j,t}$. Conditions from Dai and Singleton (2000) are imposed on the parameters—the coefficients g_{ij}, the $N \times N$ matrix Ψ, and the N-vector $\bar{X}$, as explained in Section II.C of the main text.

The domestic and foreign pricing kernels are respectively specified by

$$-\log\left(\frac{M_{t+1}}{M_t}\right) = \delta + \gamma^T X_t + \lambda^T \sqrt{\Sigma_t}\,\varepsilon_{t+1} \quad , \tag{A-2}$$

$$-\log\left(\frac{M^*_{t+1}}{M^*_t}\right) = \delta^* + \gamma^{*T} X_t + \lambda^{*T} \sqrt{\Sigma_t}\,\varepsilon_{t+1} \quad , \tag{A-3}$$

where δ and δ^* are constant scalar parameters, and γ, γ^*, λ, and λ^* are N-vectors of constant parameters. Also, equations (A-1), (A-2), and (A-3) share the same set of disturbance ε_{t+1}.

In the affine setting, pricing kernels take the form of (A-2), or equivalently, (A-3) (Backus, Foresi, and Telmer (2001)). This log-linear form of pricing kernels for the affine setting is standard (see, for example, page 511 of Duffee (2006)). It should be emphasized that (A-2) and (A-3) do not in any way imply that the pricing kernels are unique in either country. All (A-2) says is that there exists a domestic pricing kernel as specified by (A-2). It does not exclude the possibility that there exists a different pricing kernel $\tilde{M}$, where $-\log\left(\frac{\tilde{M}_{t+1}}{\tilde{M}_t}\right) = \tilde{\delta} + \tilde{\gamma}^T X_t + \tilde{\lambda}^T \sqrt{\Sigma_t}\,\varepsilon_{t+1}$, for the domestic country. Same can be said for (A-3) and the foreign country. This proof does not require the uniqueness of the pricing kernel in any country. In fact, the proof does not even require that the remaining pricing kernels in a country to be of the log-linear form. As long as there exists one pricing kernel in the domestic country that can be specified as (A-2) and one foreign pricing kernel that can be specified as (A-3), this proof stands.

The short interest rates r and r^* are affine functions of the state variables, which need to satisfy the fundamental pricing equation (1). In the case of r, substituting (A-2) into (1) gives

$$\text{(A-4)} \qquad E_t\left[\frac{M_{t+1}}{M_t}e^{r_t}\right] = E_t\left[\exp(-\delta-\gamma^T X_t - \lambda^T\sqrt{\Sigma_t}\,\varepsilon_{t+1}+r_t)\right] = 1\ .$$

Recall that, if J is a Normal random variable with mean μ_J and standard deviation σ_J, then $E[\exp(cJ)] = \exp(c\mu_J + c^2\sigma_J^2/2)$ for any constant c. The only Normal random variables in (A-4) are the components of ε_{t+1}. Therefore, equation (A-4) becomes

$$\exp\left\{-\delta-\gamma^T X_t + r_t + \sum_{i=1}^{N}\lambda_i^2(g_{i0}+\sum_{j=1}^{N}g_{ij}X_{j,t})/2\right\} = \exp\{0\} = 1\ \ .$$

Hence

$$\text{(A-5)} \qquad r_t = \delta+\gamma^T X_t - \sum_{i=1}^{N}\lambda_i^2(g_{i0}+\sum_{j=1}^{N}g_{ij}X_{j,t})/2$$

and similarly

$$\text{(A-6)} \qquad r_t^* = \delta^* + \gamma^{*T} X_t - \sum_{i=1}^{N}\lambda_i^{*2}(g_{i0}+\sum_{j=1}^{N}g_{ij}X_{j,t})/2\ \ .$$

In our extension of the affine term structure models, log of the exchange rate, s_t, is an affine function of the state variables:

$$s_t = a + b^T X_t \quad ,$$

where a is a scalar and b is an N-vector of constant coefficients.

First consider a domestic asset K that pays off, at time $(t+1)$, in units of domestic currency,

$$\text{(A-7)} \qquad K_{t+1} = \exp\{\delta+\gamma^T X_t + \lambda^T\sqrt{\Sigma_t}\,\varepsilon_{t+1}\}$$

Its price at time t, in units of the domestic currency, is obtained as

$$\begin{aligned} & E_t[\tfrac{M_{t+1}}{M_t}K_{t+1}] \\ = \ & E_t[\exp\{-\delta-\gamma^T X_t - \lambda^T\sqrt{\Sigma_t}\,\varepsilon_{t+1}+\delta+\gamma^T X_t+\lambda^T\sqrt{\Sigma_t}\,\varepsilon_{t+1}\}] \\ = \ & \exp\{0\} = 1 \quad . \end{aligned}$$

If a foreign investor invests one unit of the foreign currency in K at time t , his time $(t+1)$ payoff in units of the foreign currency is $\frac{S_t}{S_{t+1}}K_{t+1}$. The time t foreign-currency price of this payoff is obtained by $E_t[\frac{M^*_{t+1}}{M^*_t}\frac{S_t}{S_{t+1}}K_{t+1}]$. However, we already know that the payoff $\frac{S_t}{S_{t+1}}K_{t+1}$ comes from an investment of one unit of the foreign currency at time t, thus the foreign-currency price at time t of this investment is one:

$$E_t[\frac{M^*_{t+1}}{M^*_t}\frac{S_t}{S_{t+1}}K_{t+1}] = 1 \quad .$$

Substituting (A-3), (4), (A-7), and (A-1) into the above equation, and once again applying the formula $E[\exp(cJ)] = \exp(c\mu_J + c^2\sigma_J^2/2)$ on ε_{t+1}, we have

$$\begin{aligned}
& E_t[\tfrac{M^*_{t+1}}{M^*_t}\tfrac{S_t}{S_{t+1}}K_{t+1}] \\
= \; & E_t[\exp\{-\delta^* - \gamma^{*T}X_t - \lambda^{*T}\sqrt{\Sigma_t}\,\varepsilon_{t+1} - b^T(X_{t+1}-X_t) \\
& \qquad +\delta + \gamma^T X_t + \lambda^T\sqrt{\Sigma_t}\,\varepsilon_{t+1}\}] \\
= \; & E_t[\exp\{-\delta^* - \gamma^{*T}X_t - \lambda^{*T}\sqrt{\Sigma_t}\,\varepsilon_{t+1} - b^T\Psi(\bar{X}-X_t) - b^T\sqrt{\Sigma_t}\,\varepsilon_{t+1} \\
& \qquad +\delta + \gamma^T X_t + \lambda^T\sqrt{\Sigma_t}\,\varepsilon_{t+1}\}] \\
= \; & E_t[\exp\{-(\delta^*-\delta) - (\gamma^*-\gamma)^T X_t - b^T\Psi(\bar{X}-X_t) + (\lambda - b - \lambda^*)^T\sqrt{\Sigma_t}\,\varepsilon_{t+1}\}] \\
= \; & \exp\{-(\delta^*-\delta) - (\gamma^*-\gamma)^T X_t - b^T\Psi(\bar{X}-X_t) \\
& \qquad + \textstyle\sum_{i=1}^N(\lambda_i - b_i - \lambda_i^*)^2(g_{i0} + \sum_{j=1}^N g_{ij}X_{j,t})/2\} \\
= \; & 1 \; = \; \exp\{0\} \quad .
\end{aligned}$$

To simplify notation, the exponent is denoted as

$$f(X_t) = -(\delta^*-\delta) - (\gamma^*-\gamma)^T X_t - b^T\Psi(\bar{X}-X_t) + \sum_{i=1}^N(\lambda_i - b_i - \lambda_i^*)^2(g_{i0} + \sum_{j=1}^N g_{ij}X_{j,t})/2 \quad .$$

Because the equality $\exp\{f(X_t)\} = \exp\{0\}$ has to hold for any value X_t can take at t, it must be true that the coefficient for every element of vector X_t in $f(X_t)$ is 0. Otherwise, the values of X_t can be varied to find a contradiction to the equality. Consequently, the remaining constant term in $f(X_t)$ must also be 0. Thus $N+1$ equations are obtained :

(A-8) $$b^T\Psi\bar{X} - (\delta - \delta^*) = \sum_{i=1}^N(\lambda_i - b_i - \lambda_i^*)^2 g_{i0}/2 \quad ,$$

and

(A-9) $$-(b^T\Psi)_j - (\gamma_j - \gamma_j^*) = \sum_{i=1}^{N}(\lambda_i - b_i - \lambda_i^*)^2 g_{ij}/2 \qquad \text{for} \quad j = 1,2,\ldots,N.$$

Similarly, consider a foreign asset K^* that pays off, at time $(t+1)$, in units of the foreign currency,

(A-10) $$K_{t+1}^* = \exp\{\delta^* + \gamma^{*T} X_t + \lambda^{*T}\sqrt{\Sigma_t}\ \varepsilon_{t+1}\} \quad .$$

The price of K^* at time t, in units of the foreign currency, is

$$\begin{aligned}
& E_t[\tfrac{M_{t+1}^*}{M_t^*} K_{t+1}^*] \\
= \ & E_t[\exp\{ -\delta^* - \gamma^{*T} X_t - \lambda^{*T}\sqrt{\Sigma_t}\ \varepsilon_{t+1} + \delta^* + \gamma^{*T} X_t + \lambda^{*T}\sqrt{\Sigma_t}\ \varepsilon_{t+1}\}] \\
= \ & \exp\{0\} \ = 1 \quad .
\end{aligned}$$

If a domestic investor invests one unit of the domestic currency in K^* at time t , his time $(t+1)$ payoff in units of the domestic currency is $\frac{S_{t+1}}{S_t}K_{t+1}^*$. The time t domestic-currency price of this payoff is obtained by $E_t[\frac{M_{t+1}}{M_t}\frac{S_{t+1}}{S_t}K_{t+1}^*]$. However, we already know that the payoff $\frac{S_{t+1}}{S_t}K_{t+1}^*$ comes from an investment of one unit of the domestic currency at time t, thus the domestic-currency price at time t of this investment is one:

$$E_t[\frac{M_{t+1}}{M_t}\frac{S_{t+1}}{S_t}K_{t+1}^*] = 1 \quad .$$

Substituting (A-2), (4), (A-10), and (A-1) into the above equation, we have

$$\begin{aligned}
& E_t[\tfrac{M_{t+1}}{M_t}\tfrac{S_{t+1}}{S_t} K_{t+1}^*] \\
= \ & E_t[\exp\{ -\delta - \gamma^T X_t - \lambda^T\sqrt{\Sigma_t}\ \varepsilon_{t+1} + b^T(X_{t+1} - X_t) \\
& \qquad + \delta^* + \gamma^{*T} X_t + \lambda^{*T}\sqrt{\Sigma_t}\ \varepsilon_{t+1}\}] \\
= \ & E_t[\exp\{ -\delta - \gamma^T X_t - \lambda^T\sqrt{\Sigma_t}\ \varepsilon_{t+1} + b^T\Psi(\overline{X} - X_t) + b^T\sqrt{\Sigma_t}\,\varepsilon_{t+1} \\
& \qquad + \delta^* + \gamma^{*T} X_t + \lambda^{*T}\sqrt{\Sigma_t}\ \varepsilon_{t+1}\}] \\
= \ & E_t[\exp\{ (\delta^* - \delta) + (\gamma^* - \gamma)^T X_t + b^T\Psi(\overline{X} - X_t) + (b + \lambda^* - \lambda)^T\sqrt{\Sigma_t}\,\varepsilon_{t+1}\}] \\
= \ & \exp\{ (\delta^* - \delta) + (\gamma^* - \gamma)^T X_t + b^T\Psi(\overline{X} - X_t) \\
& \quad + \textstyle\sum_{i=1}^N(\lambda_i - b_i - \lambda_i^*)^2(g_{i0} + \sum_{j=1}^N g_{ij}X_{j,t})/2\} \\
= \ & 1 \ = \ \exp\{0\} \quad .
\end{aligned}$$

Considering once again the coefficients of the elements of X and the constant term inside the exponent, we obtain the following $N+1$ equations:

$$\text{(A-11)} \qquad b^T\Psi\bar{X} - (\delta - \delta^*) = -\sum_{i=1}^{N}(\lambda_i - b_i - \lambda_i^*)^2 g_{i0}/2 \quad ,$$

and

$$\text{(A-12)} \qquad -(b^T\Psi)_j - (\gamma_j - \gamma_j^*) = -\sum_{i=1}^{N}(\lambda_i - b_i - \lambda_i^*)^2 g_{ij}/2 \qquad \text{for} \quad j = 1,2,\ldots,N.$$

Comparing (A-8) with (A-11), and (A-9) with (A-12), we immediately have

$$\text{(A-13)} \qquad b^T\Psi\bar{X} - (\delta - \delta^*) = 0 \quad ,$$

$$\text{(A-14)} \qquad -(b^T\Psi)_j - (\gamma_j - \gamma_j^*) = 0 \quad \text{for} \quad j = 1,2,\ldots,N,$$

and

$$\text{(A-15)} \qquad \sum_{i=1}^{N}(\lambda_i - b_i - \lambda_i^*)^2 g_{i0}/2 = 0 \quad ,$$

$$\text{(A-16)} \qquad \sum_{i=1}^{N}(\lambda_i - b_i - \lambda_i^*)^2 g_{ij}/2 = 0 \quad \text{for} \quad j = 1,2,\ldots,N.$$

We define an $N \times (N+1)$ matrix $G = [g_{ij}]$, with $i = 1,2,\ldots,N$, $j = 0,1,2,\ldots,N$ and constants g_{ij} the same ones as in the definition of Σ_t. For models in each subfamily $\mathbb{A}_n(N)$, among the N state variables in X, there are n of them which are guaranteed to be always non-negative by construction. Without loss of generality, arrange these always-non-negative state variables to be the first n components of X. According to the conditions of Dai and Singleton (2000), $g_{i0} = 0$ for $1 \leq i \leq n$ and $g_{i0} = 1$ for $n < i \leq N$. Substituting these values of g_{i0} into (A-15), we have

$$\text{(A-17)} \qquad b_i = \lambda_i - \lambda_i^* \quad \text{for} \;\; n < i \leq N \;\; .$$

Therefore, from equation (A-16) we have

$$\text{(A-18)} \qquad \sum_{i=1}^{n}(\lambda_i - b_i - \lambda_i^*)^2 g_{ij}/2 = 0 \quad \text{for} \quad j = 1,2,\ldots,n.$$

Also according to Dai and Singleton's conditions, the submatrix of G, $\tilde{G} = [g_{ij}]$, with $i = 1,2,\ldots,n$, $j = 1,2,\ldots,n$, is an identity matrix, and thus of rank n. Hence from (A-18),

$$\text{(A-19)} \qquad b_i = \lambda_i - \lambda_i^* \quad \text{for} \;\; 1 \leq i \leq n \;\; .$$

Together with (A-17), we have

(A-20) $$b_i = \lambda_i - \lambda_i^* \quad \text{for} \;\; 1 \leq i \leq N \;\; .$$

Applying (A-1), (A-13), (A-14), and (A-20), we obtain

$$\begin{aligned}
& s_{t+1} - s_t \\
= \; & b^T(X_{t+1} - X_t) \\
= \; & b^T\Psi(\bar{X} - X_t) + b^T\sqrt{\Sigma_t}\,\varepsilon_{t+1} \\
= \; & b^T\Psi\bar{X} - b^T\Psi X_t + b^T\sqrt{\Sigma_t}\,\varepsilon_{t+1} \\
= \; & (\delta - \delta^*) + (\gamma - \gamma^*)^T X_t + (\lambda - \lambda^*)^T\sqrt{\Sigma_t}\,\varepsilon_{t+1} \\
= \; & \log\left(\frac{M_{t+1}^*}{M_t^*}\right) - \log\left(\frac{M_{t+1}}{M_t}\right) \quad .
\end{aligned}$$

Q.E.D.

In total, three conditions from Dai and Singleton (2000) are used in the proof: $g_{i0} = 0$ for $1 \leq i \leq n$, $g_{i0} = 1$ for $n < i \leq N$, and the submatrix of G, $\tilde{G} = [g_{ij}]$, with $i = 1,2,\ldots,n$, $j = 1,2,\ldots,n$, is an identity matrix. The conditions are used solely for the purpose of deriving (A-20) from (A-15) and (A-16). These conditions can be weakened and the proof still stands.

We can require, instead, only that $g_{ij} \geq 0$, for $i = 1,2,\ldots,N$, $j = 0,1,2,\ldots,N$. In addition, we assume that the specification in (A-1) is not degenerate, in that for each diagonal element of Σ_t, $\Sigma_{ii,t} = g_{i0} + \sum_{j=1}^N g_{ij} X_{j,t}$, at least one of the coefficients $g_{i0}, g_{i1}, g_{i2}, \ldots, g_{iN}$ is not zero. Otherwise, $\Sigma_{ii,t} = 0$ and X_i in (A-1) becomes a deterministic linear combination of the rest of the state variables. Below, we show how to derive (A-20) from (A-15) and (A-16), under these weakened conditions.

Given $g_{ij} \geq 0$, for $i = 1,2,\ldots,N$, $j = 0,1,2,\ldots,N$, (A-15) and (A-16) imply

(A-21) $$(\lambda_i - b_i - \lambda_i^*)^2 g_{ij} = 0 \quad \text{for} \;\; i = 1,2,\ldots,N \;, \;\; j = 0,1,2,\ldots,N.$$

Suppose, for k, (A-20) does not hold. That is, $b_k \neq \lambda_k - \lambda_k^*$. From (A-21), we have, for $j = 0, 1, 2, \ldots, N$, $(\lambda_k - b_k - \lambda_k^*)^2 g_{kj} = 0$. Because $b_k \neq \lambda_k - \lambda_k^*$, it has to be true that $g_{kj} = 0$, for $j = 0, 1, 2, \ldots, N$. This implies $\Sigma_{kk,t} = g_{k0} + \sum_{j=1}^{N} g_{kj} X_{j,t} = 0$, and is inconsistent with the condition above that the specification in (A-1) is not degenerate. Therefore, it has to be the case that $b_k = \lambda_k - \lambda_k^*$, for $k = 1, 2, \ldots, N$, which is equation (A-20).

References

Ahn, D. H. "Common Factors and Local Factors: Implications for Term Structures and Exchange Rates." *Journal of Financial and Quantitative Analysis*, 39 (2004), 69-102.

Amin, K. I., and R. A. Jarrow. "Pricing Foreign Currency Options Under Stochastic Interest Rates." *Journal of International Money and Finance*, 10 (1991), 310-329.

Babbs, S. H., and K. B. Nowman. "Kalman Filtering of Generalized Vasicek Term Structure Models." *Journal of Financial and Quantitative Analysis*, 34 (1999), 115-130.

Backus, D. K.; S. Foresi; and C. I. Telmer. "Affine Term Structure Models and the Forward Premium Anomaly." *Journal of Finance*, 56 (2001), 279-304.

Bakshi, G. S., and Z. Chen. "Equilibrium Valuation of Foreign Exchange Claims." *Journal of Finance*, 52 (1997), 799-826.

Biger, N., and J. Hull. "The Valuation of Currency Options." *Financial Management*, 12

(1983), 24-28.

Brandt, M. W., and P. Santa-Clara. “Simulated Likelihood Estimation of Diffusions with an Application to Exchange Rate Dynamics in Incomplete Markets.” *Journal of Financial Economics*, 63 (2002), 161-210.

Brennan, M. J., and Y. Xia. “International Capital Markets and Foreign Exchange Risk.” *Review of Financial Studies*, 19 (2006), 753-795.

Chen, R., and L. O. Scott. “Multi-Factor Cox-Ingersoll-Ross Models of the Term Structure: Estimates and Tests from a Kalman Filter Model.” *Journal of Real Estate Finance and Economics*, 27 (2003), 143-172.

Choi, J. J., and S. Hauser. “The Effects of Domestic and Foreign Yield Curves on the Value of Currency American Call Options.” *Journal of Banking and Finance*, 14 (1990), 41-53.

Cox, J. C.; J. E. Ingersoll; and S. A. Ross. “A Theory of the Term Structure of Interest Rates.” *Econometrica*, 53 (1985), 385-407.

Dai, Q., and K. J. Singleton. “Specification Analysis of Affine Term Structure Models.” *Journal of Finance*, 55 (2000), 1943-1978.

De Jong, F. “Time Series and Cross-Section Information in Affine Term-structure Models.” *Journal of Business and Economic Statistics*, 18 (2000), 300-314.

Dewachter, H., and K. Maes. “An Affine Model for International Bond Markets.” Working Paper, Catholic Univ. of Leuven (2001).

Duan, J. C., and J. G. Simonato. “Estimating and Testing Exponential-Affine Term Structure Models by Kalman Filter.” *Review of Quantitative Finance and Accounting*, 13 (1999), 111-135.

Duffee, G. R. “Estimating the Price of Default Risk.” *Review of Financial Studies*, 12 (1999), 197-226.

Duffee, G. R. “Term Premia and Interest Rate Forecasts in Affine Models.” *Journal of Finance*, 57 (2002), 405-443.

Duffee, G. R. “Term Structure Estimation without Using Latent Factors.” *Journal of Financial Economics*, 79 (2006), 507-536.

Duffee, G. R., and R. H. Stanton. “Estimation of Dynamic Term Structure Models.” Working Paper, UC Berkeley (2001).

Duffie, D. *Dynamic Asset Pricing Theory*, 3rd ed. Princeton, New Jersey: Princeton University Press (2001).

Duffie, D., and M. Huang. “Swap Rates and Credit Quality.” *Journal of Finance*, 51 (1996), 921-949.

Duffie, D., and R. Kan. “A Yield-Factor Model of Interest Rates.” *Mathematical Finance*, 6 (1996), 379-406.

Garman, M. B., and S. W. Kohlhagen. “Foreign Currency Option Values.” *Journal of In-*

ternational Money and Finance, 2 (1983), 231-237.

Heath, D.; R. A. Jarrow; and A. Morton. "Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claims Valuation." *Econometrica*, 60 (1992), 77-105.

Hodrick, R. J., and M. Vassalou. "Do We Need Multi-Country Models to Explain Exchange Rate and Interest Rate and Bond Return Dynamics?" *Journal of Economic Dynamics and Control*, 26 (2002), 1275-1299.

Mosburger, G., and P. Schneider. "Modelling International Bond Markets with Affine Term Structure Models." Working Paper, Univ. of Vienna (2005).

Nelson, C. R., and A. F. Siegel. "Parsimonious Modeling of Yield Curves." *Journal of Business*, 60 (1987), 473-489.

Nielsen, L. T., and J. Saà-Requejo. "Exchange Rate and Term Structure Dynamics and the Pricing of Derivative Securities." Working Paper, INSEAD (1993).

Saà-Requejo, J. "The Dynamics and the Term Structure of Risk Premia in Foreign Exchange Markets." Working Paper, INSEAD (1994).

Tang, H., and Y. Xia. "An International Examination of Affine Term Structure Models and the Expectations Hypothesis." *Journal of Financial and Quantitative Analysis*, forthcoming.

White H. "Maximum Likelihood Estimation of Misspecified Models." *Econometrica*, 50 (1982), 1-25.

Country Pair	n in $\mathbb{A}_n(5)$	Log Likelihood	AIC	BIC	Parameter Count
DEM JPY	1	78057*	-155992	-155752	61
DEM JPY	2	76280	-152436	-152192	62
DEM JPY	3	73763	-147400	-147152	63
DEM JPY	4	76267	-152406	-152154	64
DEM JPY	5	75957	-151784	-151528	65
DEM GBP	1	78843*	-157564	-157324	61
DEM GBP	2	75853	-151582	-151338	62
DEM GBP	3	78106	-156086	-155838	63
DEM GBP	4	76434	-152740	-152488	64
DEM GBP	5	74086	-148042	-147786	65
DEM USD	1	77266*	-154410	-154170	61
DEM USD	2	77000	-153876	-153632	62
DEM USD	3	76636	-153146	-152898	63
DEM USD	4	74836	-149544	-149292	64
DEM USD	5	72867	-145604	-145348	65
JPY GBP	1	77786*	-155450	-155210	61
JPY GBP	2	74693	-149262	-149018	62
JPY GBP	3	77226	-154326	-154078	63
JPY GBP	4	77081	-154034	-153782	64
JPY GBP	5	74751	-149372	-149116	65
JPY USD	1	74294	-148466	-148226	61
JPY USD	2	79406*	-158688	-158444	62
JPY USD	3	72021	-143916	-143668	63
JPY USD	4	73485	-146842	-146590	64
JPY USD	5	76488	-152846	-152590	65
GBP USD	1	76084*	-152046	-151806	61
GBP USD	2	73110	-146096	-145852	62
GBP USD	3	75952	-151778	-151530	63
GBP USD	4	75214	-150300	-150048	64
GBP USD	5	72757	-145384	-145128	65

Table 1. Comparison of Estimated Models. Log likelihood values, Akaike and Bayesian Information Criteria, and number of parameters, for the maximal $\mathbb{A}_1(5)$ through $\mathbb{A}_5(5)$ models estimated, of the six country pairs. An asterisk next to the log likelihood value indicates the best model for each country pair. Throughout, countries are represented by the international standard three-letter codes for their currencies.

Country	MAE	RMSE	Country	MAE	RMSE	FX MAE	FX RMSE
JPY	16.49	28.49	DEM	5.58	9.50	0.0438	0.0506
DEM	7.11	11.06	GBP	9.03	14.45	0.1259	0.1531
DEM	12.40	21.02	USD	8.14	13.89	0.5135	0.6085
JPY	16.45	28.26	GBP	6.00	10.02	0.0351	0.0462
JPY	15.60	27.18	USD	4.95	7.06	0.0242	0.0317
USD	12.89	21.16	GBP	8.26	12.84	0.1819	0.2010

Table 2. Statistics on Absolute Fitting Errors. The means of the absolute fitting errors (MAE) and RMSE for the best-fit model of each country pair. Each row represents a different model for a different country pair. The MAE and RMSE reported next to the country itself are for the interest rates of that country, expressed in basis points. "FX MAE" and "FX RMSE" are for the log exchange rates between the two currencies.

Country	MAE DEM	RMSE DEM	MAE JPY	RMSE JPY	MAE GBP	RMSE GBP	MAE USD	RMSE USD
1-month	2.35	3.40	1.36	3.04	2.03	2.89	3.18	5.12
2-month	1.47	2.81	0.98	2.20	1.24	2.44	1.80	4.28
4-month	2.48	4.05	1.20	2.16	2.39	3.61	3.26	5.31
6-month	2.17	2.96	0.82	1.39	1.74	2.14	2.68	3.64
8-month	1.35	1.67	0.41	0.56	1.21	1.52	1.52	1.96
10-month	0.87	1.33	0.49	0.67	1.02	1.48	1.24	1.77
12-month	1.86	2.49	0.85	1.28	2.49	2.88	2.67	3.37
2-year	6.02	7.27	5.19	7.29	10.92	13.87	11.23	13.93
4-year	5.21	6.34	16.50	19.36	14.51	16.76	12.10	15.94
6-year	2.66	3.22	25.90	30.84	12.25	14.03	8.82	10.77
8-year	2.10	2.63	29.78	36.48	9.96	12.45	13.84	16.97
10-year	4.10	5.00	30.47	38.50	9.14	11.36	18.13	21.72
12-year	6.16	7.48	31.15	40.33	9.06	11.47	20.69	24.90
15-year	9.11	11.14	31.75	42.14	10.88	13.92	24.14	29.78
20-year	12.72	15.82	30.82	41.84	13.76	17.80	29.74	37.09
25-year	15.64	19.77	28.96	40.04	16.48	21.63	31.13	39.73
30-year	18.51	22.84	28.48	39.89	21.30	26.32	33.01	41.47

Table 3. Statistics on Absolute Fitting Errors for Different Maturities. The means of the absolute fitting errors (MAE) and RMSE for different in-sample maturities, in basis points, for each of the four countries. Both the GBP and the USD statistics are taken from the model for the GBP/USD pair. The JPY statistics are from the model for the JPY/USD pair, and DEM from the DEM/JPY pair.

Country	MAE DEM	RMSE DEM	MAE JPY	RMSE JPY	MAE GBP	RMSE GBP	MAE USD	RMSE USD
3-month	2.28	4.32	1.32	2.62	2.42	4.20	2.97	6.29
5-month	2.43	3.68	1.09	1.81	2.20	2.92	3.28	4.82
7-month	1.79	2.29	0.56	0.91	1.48	1.77	2.11	2.73
9-month	1.01	1.31	0.37	0.47	1.05	1.46	1.36	1.77
11-month	1.20	1.74	0.66	0.97	1.61	1.98	1.82	2.40
1-year	2.90	3.63	1.68	2.55	7.88	9.34	5.91	7.79
3-year	6.30	7.60	10.62	12.93	13.70	15.91	15.35	19.10
5-year	3.91	4.81	21.91	25.76	13.67	15.42	8.99	11.16
7-year	1.76	2.14	28.46	34.36	10.85	12.97	11.24	13.92
9-year	3.03	3.70	30.24	37.66	9.24	11.57	16.27	19.60

Table 4. Statistics on Out-of-Sample Fitting Errors. The means of the absolute fitting errors (MAE) and RMSE for different out-of-sample maturities, in basis points, for each of the four countries. Both the GBP and the USD statistics are taken from the model for the GBP/USD pair. The JPY statistics are from the model for the JPY/USD pair, and DEM from the DEM/JPY pair.

Index i	1	2	3	4	5
Ψ_{1i}	1.4981 (0.0426)	-0.0229 (0.0096)	0	0	0
Ψ_{2i}	-0.0062 (0.0072)	0.9034 (0.0235)	0	0	0
Ψ_{3i}	-0.1207 (0.0301)	-0.0004 (0.0088)	1.4669 (0.0014)	0.0869 (0.0226)	-0.0352 (0.0206)
Ψ_{4i}	-0.3066 (0.0444)	-0.1074 (0.0387)	-0.0523 (0.0156)	0.0435 (0.0239)	-0.0109 (0.1119)
Ψ_{5i}	-0.1304 (0.1871)	-0.1029 (0.3405)	-0.0471 (0.0091)	-0.0583 (0.0878)	0.3708 (0.0131)
$\bar{X}_i$	1.6048 (0.0483)	0.6731 (0.0438)	0	0	0
g_{i0}	0	0	1	1	1
g_{1i}	1	0	0	0	0
g_{2i}	0	1	0	0	0
g_{3i}	0.0000 (0.0000)	0.6762 (0.0796)	0	0	0
g_{4i}	1.6726 (0.1476)	0.2121 (0.0445)	0	0	0
g_{5i}	0.9438 (0.6747)	0.0278 (0.3905)	0	0	0
$\rho_{0,JPY}$	0.0024 (0.0027)				
$\rho_{i,JPY}$	0.0508 (0.0013)	0.1303 (0.0010)	0.0952 (0.0015)	0.0016 (0.0026)	0.0215 (0.0040)
$\rho_{0,USD}$	0.2685 (0.0058)				
$\rho_{i,USD}$	-0.0230 (0.0011)	-0.3569 (0.0050)	0.1112 (0.0017)	0.0212 (0.0018)	0.0048 (0.0033)
$\lambda_{i,JPY}$	-0.0138 (0.1075)	-0.1886 (0.0227)	0.0411 (0.0111)	0.4788 (0.3134)	0.9606 (0.2997)
$\lambda_{i,USD}$	3.2499 (0.0410)	0.0026 (0.0183)	0.3675 (0.0128)	-0.5887 (0.0654)	0.5251 (0.1930)
a	0.0479 (0.0826)				
b_i	0.0464 (0.0028)	0.1325 (0.0009)	0.0835 (0.0029)	0.1880 (0.0694)	-0.2599 (0.0756)
	ω_s	ω_m	ω_l		
JPY	0.00021 (0.00000)	0.00284 (0.00002)	0.00412 (0.00004)		
USD	0.00055 (0.00001)	0.00104 (0.00001)	0.00066 (0.00001)		
ω_e	0.13683 (0.00079)				

Table 5. Parameter Estimates of the Best-Fit Model for the Japan - US Pair. The best-fit model for the Japan - US pair belongs to the $\mathbb{A}_2(5)$ subfamily. Asymptotic standard errors of the estimated parameters are given in parentheses. Some of the parameters are specified by the model as 0 or 1, and thus appear in the table without accompanying standard errors.

Index i	0	1	2	3	4	5
$\rho_{i,JPY}$	*	*	-	-	*	*
$\rho_{i,DEM}$	*	*	*	-	*	-
b_i	*	*	-	*	*	*
$\rho_{i,DEM}$	*	*	*	-	-	*
$\rho_{i,GBP}$	*	*	*	-	-	*
b_i	*	*	*	-	*	*
$\rho_{i,DEM}$	*	*	-	-	*	*
$\rho_{i,USD}$	*	*	*	-	-	-
b_i	*	*	-	*	*	-
$\rho_{i,JPY}$	*	*	*	*	*	-
$\rho_{i,GBP}$	*	*	-	*	*	-
b_i	*	*	*	*	-	*
$\rho_{i,JPY}$	-	*	*	*	-	*
$\rho_{i,USD}$	*	*	*	*	*	-
b_i	-	*	*	*	*	*
$\rho_{i,USD}$	*	*	-	*	*	-
$\rho_{i,GBP}$	*	*	*	*	*	*
b_i	*	*	*	*	*	-

Table 6. Contributions of the State Variables to Exchange Rate and Interest Rate Dynamics. For the best-fit model of each country pair, the estimates of the ρ and b parameters are tested to see which state variables have significant direct contributions to each of the exchange rate and interest rate dynamics. An asterisk shows that the corresponding parameter is statistically significantly different from zero at the 1% level. A "-" indicates non-significance. Because $s_t = a + b^T X_t$, where the vector b contains elements b_1 through b_5, results for the parameter a is reported in the space for b_0 in the table.

Index i	1	2	3	4	5
Ψ_{1i}	1.0452 (0.0101)	0	0	0	0
Ψ_{2i}	-0.1041 (0.0431)	2.4677 (0.0208)	-0.0653 (0.0784)	0.0252 (0.0580)	0.0133 (0.0587)
Ψ_{3i}	-0.2743 (0.0343)	0.0077 (0.1169)	0.2760 (0.0168)	0.1310 (0.2013)	-0.0390 (0.0533)
Ψ_{4i}	-0.0284 (0.0971)	0.0267 (0.0918)	0.0302 (0.2823)	2.1886 (0.0238)	-0.0164 (0.1592)
Ψ_{5i}	-0.0008 (0.1007)	0.0070 (0.1375)	0.0034 (0.0688)	-0.0029 (0.2549)	0.8266 (0.0069)
$\bar{X}_i$	0.5179 (0.0186)	0	0	0	0
g_{i0}	0	1	1	1	1
g_{1i}	1	0	0	0	0
g_{2i}	0.0364 (0.0186)	0	0	0	0
g_{3i}	0.3884 (0.0101)	0	0	0	0
g_{4i}	0.2566 (0.0391)	0	0	0	0
g_{5i}	0.9765 (0.1676)	0	0	0	0
$\rho_{0,USD}$	-0.0368 (0.0031)				
$\rho_{i,USD}$	0.4228 (0.0043)	0.2395 (0.1065)	0.1764 (0.0624)	0.3617 (0.0357)	0.0157 (0.0527)
$\rho_{0,GBP}$	0.1700 (0.0035)				
$\rho_{i,GBP}$	0.0940 (0.0003)	0.1674 (0.0470)	0.1521 (0.0441)	0.1950 (0.0163)	0.1398 (0.0459)
$\lambda_{i,USD}$	-0.2166 (0.0115)	0.1318 (0.1099)	-0.4287 (0.1362)	0.8845 (0.0519)	0.4668 (0.3378)
$\lambda_{i,GBP}$	-0.5680 (0.0104)	-1.2550 (0.1611)	-0.2810 (0.3173)	0.8917 (0.2831)	0.7061 (0.1003)
a	0.0775 (0.0054)				
b_i	-0.0787 (0.0019)	0.1826 (0.0030)	-0.3572 (0.0027)	0.0838 (0.0073)	-0.0032 (0.0061)

	ω_s	ω_m	ω_l
USD	0.00049 (0.00001)	0.00165 (0.00001)	0.00349 (0.00002)
GBP	0.00030 (0.00000)	0.00163 (0.00001)	0.00157 (0.00003)
ω_e	0.88810 (0.00006)		

Table 7. Parameter Estimates of the Best-Fit Model for the US - UK Pair. The best-fit model for the US - UK pair belongs to the $\mathbb{A}_1(5)$ subfamily. Asymptotic standard errors of the estimated parameters are given in parentheses. Some of the parameters are specified by the model as 0 or 1, and thus appear in the table without accompanying standard errors.

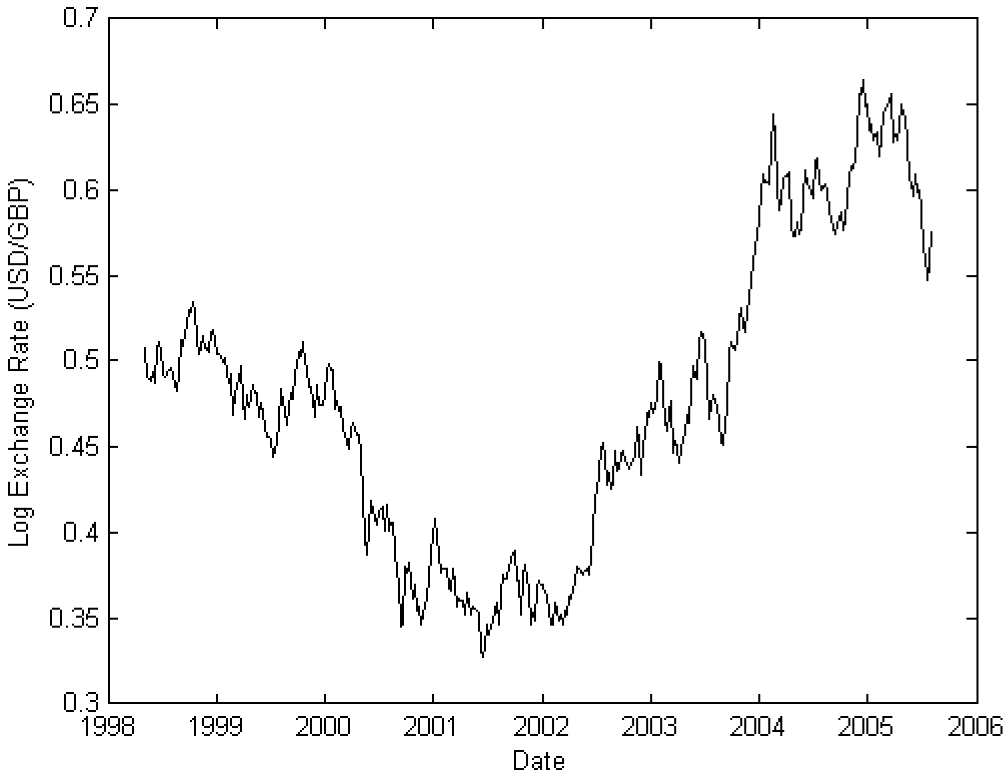

Figure 1. Log Exchange Rate between US Dollar and British Pound. Weekly log exchange rate between the US dollar and the British pound, in dollars per pound, taken every Wednesday from May 1, 1998 to August 5, 2005.

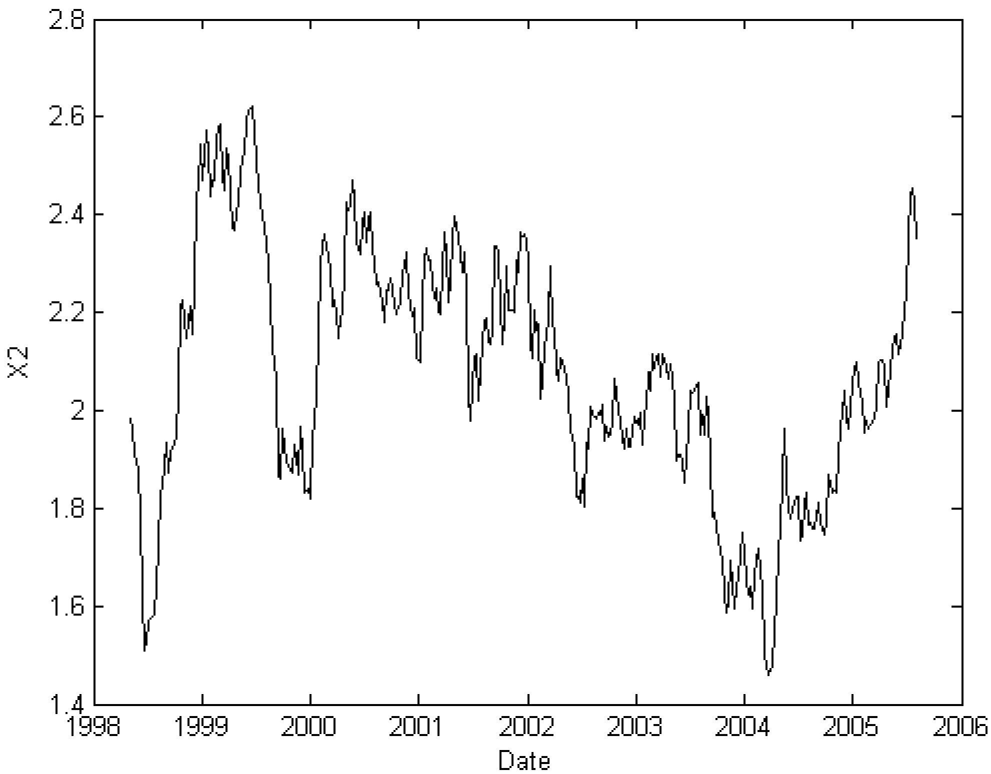

Figure 2. Smoothed Estimates of the State Variable X_2. Smoothed estimates of the state variable X_2, in the best-fit model $\mathbb{A}_1(5)$ for the UK - US pair. Estimates are weekly for every Wednesday from May 1, 1998 to August 5, 2005.

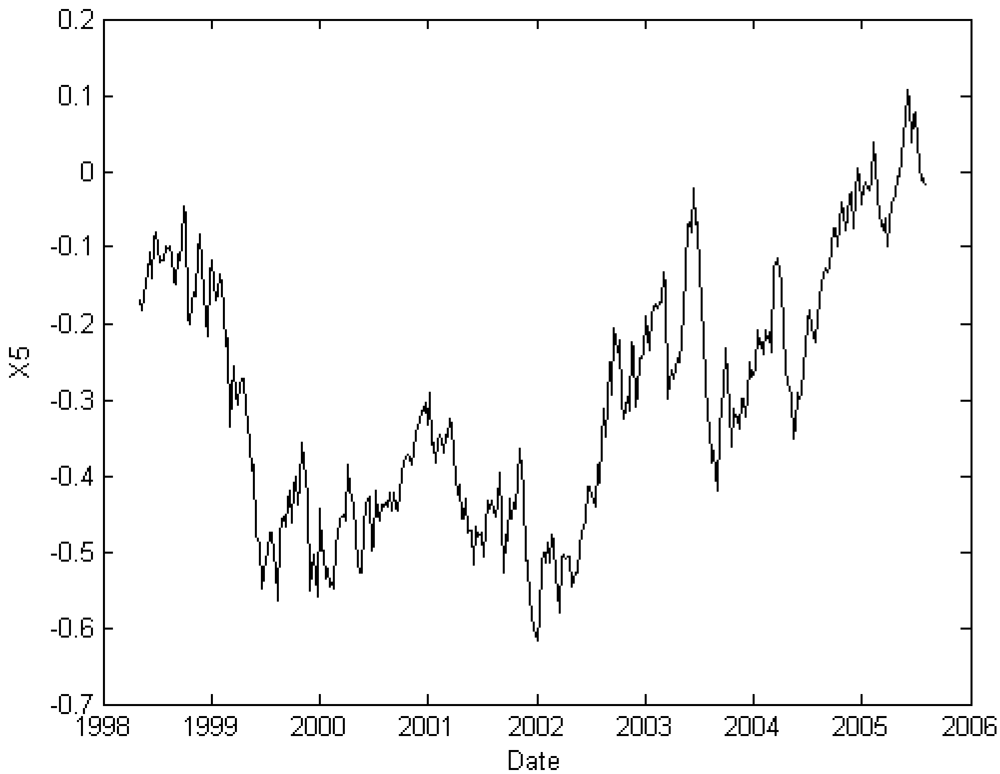

Figure 3. Smoothed Estimates of the State Variable X_5. Smoothed estimates of the state variable X_5, in the best-fit model $\mathbb{A}_1(5)$ for the UK - US pair. Estimates are weekly for every Wednesday from May 1, 1998 to August 5, 2005.