Affine Models of Currency Pricing: Accounting for the Forward Premium Anomaly[*]

David Backus,[†] Silverio Foresi,[‡] and Chris Telmer[§]

First Draft: June 1994
This Version: July 10, 1998

Abstract

One of the most puzzling features of currency prices is the tendency for high interest rate currencies to appreciate. Some have attributed this *forward premium anomaly* to a time-varying risk premium, but theory has been largely unsuccessful in producing a risk premium with the requisite properties. We characterize the risk premium in both general and affine arbitrage-free settings and describe the features a theory must have to account for the anomaly. In affine models, the anomaly requires either that state variables have asymmetric effects on state prices in different currencies or that we abandon the requirement that interest rates be strictly positive.

JEL Classification Codes: F31, G12, G15.

Keywords: forward and spot exchange rates; risk premiums; pricing kernels.

[†]Stern School of Business, New York University and NBER; dbackus@stern.nyu.edu.

[‡]Salomon Smith Barney; sforesi@sbi.com.

[§]Graduate School of Industrial Administration, Carnegie Mellon University; telmer@aleast.gsia.cmu.edu.

1 Introduction

Perhaps the most puzzling feature of currency prices is the tendency for high interest rate currencies to appreciate, when one might guess, instead, that investors would demand higher interest rates on currencies expected to fall in value. This departure from uncovered interest parity, which we term the *forward premium anomaly*, has been documented in dozens — and possibly hundreds — of studies, and has spawned a second generation of papers attempting to account for it. One of the most influential of these is Fama (1984), who attributed the behavior of forward and spot exchange rates to a time-varying risk premium. Fama showed that the implied risk premium on a currency must (i) be negatively correlated with its expected rate of depreciation and (ii) have greater variance.

We refer to this feature of the data as an anomaly because asset pricing theory to date has been notably unsuccessful in producing a risk premium with the requisite properties. Attempts include applications of the capital asset pricing model to currency prices (Frankel and Engel, 1984; Mark, 1988), statistical models relating risk premiums to changing second moments (Cumby, 1988; Domowitz and Hakkio, 1985; Hansen and Hodrick, 1983), and consumption-based asset pricing theories, including departures from time-additive preferences (Backus, Gregory, and Telmer, 1993; Bansal, Gallant, Hussey, and Tauchen, 1995; and Bekaert, 1996), from expected utility (Bekaert, Hodrick, and Marshall, 1997), and from frictionless trade in goods (Hollifield and Uppal, 1997).

We study the anomaly in the context of affine models, whose popularity in bond pricing has spread to currencies in recent years. Notable applications of affine models to currencies include Ahn (1997), Amin and Jarrow (1991), Bakshi and Chen (1997), Bansal (1997), Frachot (1996), Nielsen and Saá-Requejo (1993), and Saá-Requejo (1994). We consider a generalization of the models used in these papers, an adaptation to currencies of Duffie and Kan's (1997) class of affine yield models, and outline the conditions on such models needed to reproduce the forward premium anomaly. Theory and examples indicate that the anomaly places strong restrictions on the structure and parameter values of affine models: either state variables have asymmetric effects on state prices in different currencies or interest rates are negative with positive probability. We quantify both with parameter values estimated from monthly data on the dollar-pound rate.

We proceed as follows. In Section 2, we summarize the properties of dollar exchange rates and one-month eurocurrency interest rates. These properties serve as an anchor to the theory that follows. In Section 3, we describe the relations among pricing kernels, currency prices, and interest rates dictated by general arbitrage-free environments and our specific class of affine models. In Section 4, we develop and

estimate several affine models and show how different features and parameter values bear on the forward premium anomaly.

2 Properties of Currency Prices and Interest Rates

Properties of exchange rates and eurocurrency interest rates have been widely documented, but a review focuses our attention on the issues to be addressed and provides a quantitative benchmark for subsequent theory. Accordingly, we summarize the properties of spot and forward exchange rates for the US dollar versus the remaining G7 currencies and interest rates for the same currencies. Here and elsewhere, s_t is the logarithm of the dollar price of one unit of foreign currency and f_t is the logarithm of the dollar price of a one-month forward contract: a contract arranged at date t specifying payment of $\exp(f_t)$ dollars at date $t+1$ and receipt of one unit of foreign currency.

In Table 1, we report sample moments for depreciation rates of the dollar, $s_{t+1} - s_t$, continuously-compounded one-month eurocurrency interest rates, r_t, and forward premiums, $f_t - s_t$. Panel A is concerned with depreciation rates. For the currencies in our sample, mean depreciation rates are smaller than their standard deviations, typically by a factor of about eight. In this sense, volatility is one of the most striking features of currency prices. Apparently little of this volatility is predictable from past depreciation rates: estimated autocorrelations are less than 0.1 for all six currencies. Panels B and C are concerned with interest rates and interest rate differentials. Unlike currency prices, both interest rates and their differentials are highly persistent. They also exhibit less variability, both absolutely and relative to their means.

One way to think about this evidence is to relate it to the expectations hypothesis: that forward rates are expected future spot rates. We have no reason to think the expectations hypothesis is an accurate description of the world, but it provides a useful benchmark to which we can compare the data. We express the hypothesis in logarithmic form as $f_t = E_t s_{t+1}$ or $f_t - s_t = E_t s_{t+1} - s_t$, where E_t denotes the expectation conditional on date-t information. Although we do not observe expected future spot rates, we can get an indication of the accuracy of the expectations hypothesis by comparing mean forward premiums with mean depreciation rates across currencies. We see in Figure 1 (based on entries from Table 1) that while the two means are not the same, their differences are small relative to their cross-sectional variation. Currencies with large forward premiums, on average, are also those against which the dollar has depreciated the most. In other words, currencies with average interest rates higher than the dollar have typically fallen in value relative to the dollar.

This sanguine view of the expectations hypothesis changes dramatically when we turn from cross-section to time-series evidence — that is, from unconditional moments to conditional moments. A huge body of work has established, for the extant flexible exchange rate period, that forward premiums have been negatively correlated with subsequent depreciation rates for exchange rates between most major currencies. Canova and Marrinan (1995), Engel (1996), and Hodrick (1987) provide exhaustive references to the literature. The most common evidence comes from regressions of the form

$$s_{t+1} - s_t = a_1 + a_2(f_t - s_t) + \text{residual}. \tag{1}$$

The expectations hypothesis implies a regression slope $a_2 = 1$, yet most studies estimate a_2 to be negative. Thus they find not only that the expectations hypothesis provides a poor approximation to the data, but that its predictions of future currency movements are in the wrong direction. We report similar evidence in Table 2, where estimates of a_2 range from -0.073 for the lira to -1.840 for the pound. Since the forward premium equals the interest differential (covered interest parity), the regressions indicate that currencies with high interest rate differentials (relative to their means) have low rates of depreciation (also relative to their means). All of these estimates are at least two standard errors from the value of one indicated by the expectations hypothesis. Although the R^2s are small (the largest, for the Canadian dollar, is 0.034), equation (1) can be used to construct profitable investment strategies. Backus, Gregory, and Telmer (1993) and Bekaert and Hodrick (1992) show that while such strategies are not riskless, they have positive and statistically significant average excess returns.

Evidence of negative correlation between forward premiums and depreciation rates has survived, so far, a number of attempts to reverse it. One issue is the logarithmic version of the expectations hypothesis and equation (1). If we define the expectations hypothesis in terms of forward and spot exchange rates, rather than their logarithms, the evidence is virtually the same. See, for example, Backus, Gregory, and Telmer (1993, Table 2). A second issue is stability. Although estimates of a_2 vary substantially over time, they remain consistently negative. Bekaert and Hodrick (1993), for example, find that estimates based on data subsequent to Fama's (1984) sample are more strongly negative than those based on the entire sample. Data from the early 1990s moderates this conclusion, but does not invalidate it. A third issue concerns measurement error and bid-ask spreads. Bekaert and Hodrick (1993) and Bossaerts and Hillion (1991) argue, however, that neither of these factors has a material effect on the sign or magnitude of estimates of a_2. A fourth issue concerns the exchange-rate regime. Flood and Rose (1996) find that negative slope parameters are less apparent for currencies covered by the Exchange Rate Mechanism of the European Monetary System. In fact, the evidence for exchange rates in the ERM is mixed: estimates of a_2 are close to one for the German mark and the

3

French franc, but large and negative for the mark and the Dutch guilder. Flood and Rose estimate a typical ERM slope parameter of 0.58, which is significantly different from one but nevertheless positive. For floating exchange rate regimes they report, as others do, negative values for a_2.

While the evidence apparently contradicts the expectations hypothesis, Fama (1984) notes that it is consistent with a time-varying risk premium. In Fama's interpretation, the forward premium, $f_t - s_t$, includes a risk premium p_t as well as the expected rate of depreciation q_t:

$$
\begin{aligned}
f_t - s_t &= (f_t - E_t s_{t+1}) + (E_t s_{t+1} - s_t) \\
&\equiv p_t + q_t.
\end{aligned}
\tag{2}
$$

The risk premium $q_t = f_t - E_t s_{t+1}$ is the (log-linearized) expected return from buying dollars in the forward market. The cross-section evidence (Table 1 and Figure 1) suggests that risk premiums are small on average, but the time series evidence implies they are highly variable. Since the population regression coefficient is

$$
a_2 = \frac{Cov(q, p+q)}{Var(p+q)} = \frac{Cov(q,p) + Var(q)}{Var(p+q)},
\tag{3}
$$

it is clear that a constant risk premium p generates $a_2 = 1$. To generate a negative value of a_2 we need $Cov(q,p) + Var(q) < 0$. Fama notes that this requires (i) negative covariance between p and q and (ii) greater variance of p than q. We refer to these requirements as Fama's necessary conditions. They serve as hurdles that any theoretical explanation of the anomaly must surpass.

3   A Theoretical Framework

With these properties of the data in mind, we consider theories that might account for them. The challenge is to account simultaneously for currency prices and prices of fixed income securities denominated in both currencies. A model of the dollar-pound rate, for example, must account for the properties of interest rates in dollars and pounds, as well as those of the exchange rate between the two currencies. From a theoretical perspective, this challenge places demands on a model's internal consistency. It gains greater force in quantitative applications, when parameter values chosen to imitate (say) movements in exchange rates must be reconciled with properties of interest rates.

We find it useful to consider currency prices in a fairly general theoretical setting before turning to the more structured environment of affine models. We describe

models in terms of *pricing kernels*: stochastic processes governing prices of state-contingent claims. Existence of a pricing kernel (or, equivalently, of risk-neutral probabilities) is guaranteed in any economic environment that precludes arbitrage opportunities. The beauty of this result is its simplicity. It requires only that market prices of traded assets not permit combinations of trades that produce positive payoffs in some states with no initial investment — a departure from covered interest rate parity, for example. The framework encompasses, among other things, the possibility that agents trade on different information, or that some agents harbor "irrational" beliefs. In the rest of this section, we relate properties of currency prices to those of pricing kernels in two currencies and examine the relation between pricing kernels and the forward premium anomaly.

3.1 Pricing Kernels and Currency Prices

Currencies are largely a matter of units: we can quote prices in dollars or pounds, and the exchange rate is the ratio of the two. In similar fashion, we show that we can compute the depreciation rate from the ratio of pricing kernels in two currencies. With complete markets, absence of arbitrage determines unique pricing kernels whose ratio equals the depreciation rate. With incomplete markets, pricing kernels are no longer unique. We show, however, that we can choose them to satisfy the same condition.

Consider assets denominated in either domestic currency ("dollars") or foreign currency ("pounds"). The dollar value v_t of a claim to the stochastic cash flow of d_{t+1} dollars one period later satisfies

$$v_t = E_t\left(m_{t+1}d_{t+1}\right),$$ (4)

or

$$1 = E_t\left(m_{t+1}R_{t+1}\right),$$ (5)

where $R_{t+1} = d_{t+1}/v_t$ is the gross one-period return on the asset. We refer to m as the dollar pricing kernel. In economies with a representative agent, m is the nominal intertemporal marginal rate of substitution and (5) is one of the agent's first-order conditions. More generally, there exists a positive random variable m satisfying the pricing relation (5) for returns R on all traded assets if the economy admits no pure arbitrage opportunities. When the economy has a complete set of markets for state-contingent claims, m is the unique solution to (5), but otherwise there is a range of choices of m that satisfy the pricing relation for returns on all traded assets. These issues, and the relevant literature, are reviewed by Duffie (1992).

The pricing kernel m and the pricing relation (5) are the basis of modern theories of bond pricing: given a pricing kernel, we use (5) to compute prices and yields for

bonds of all maturities. The price of a one-period bond, for example, is $b_t^1 = E_t m_{t+1}$, and the (one-period) short rate r_t is

$$r_t = -\log b_t^1 = -\log E_t m_{t+1}. \tag{6}$$

We return to this equation shortly.

When we consider assets with returns denominated in pounds, we might adopt an analogous approach and use a random variable m^* to value them. Alternatively, we could convert mark returns into dollars and value them using m. The equivalence of these two procedures gives us a connection between exchange rate movements and pricing kernels in the two currencies, m and m^*. If we use the first approach, pound returns R_t^* satisfy

$$1 = E_t \left(m_{t+1}^* R_{t+1}^* \right). \tag{7}$$

If we use the second approach, with $S = \exp(s)$ denoting the dollar spot price of one pound, dollar returns on this asset are $R_{t+1} = (S_{t+1}/S_t)R_{t+1}^*$ and

$$1 = E_t \left[m_{t+1}(S_{t+1}/S_t)R_{t+1}^* \right].$$

If the pound asset and currencies are both traded, the return must satisfy both conditions:

$$E_t \left(m_{t+1}^* R_{t+1}^* \right) = E_t \left[m_{t+1}(S_{t+1}/S_t)R_{t+1}^* \right].$$

This equality ties the rate of depreciation of the dollar to the random variables m and m^* that govern state prices in dollars and pounds. Certainly this relation is satisfied if $m_{t+1}^* = m_{t+1}S_{t+1}/S_t$. This choice is dictated when the economy has a complete set of markets for currencies and state-contingent claims. With incomplete markets, the choices of m and m^* satisfying (5,7) are not unique, but we can choose them to satisfy the same equation:

Proposition 1 *Consider stochastic processes for the depreciation rate, S_{t+1}/S_t, and returns R_{t+1} and R_{t+1}^* on dollar and pound denominated assets. If these processes do not admit arbitrage opportunities, then there exist pricing kernels m and m^* for dollars and pounds that satisfy both*

$$m_{t+1}^*/m_{t+1} = S_{t+1}/S_t \tag{8}$$

and the pricing relations (5,7).

Proof. Consider dollar returns on the complete set of traded assets, including the dollar returns $(S_{t+1}/S_t)R_{t+1}^*$ on pound-denominated assets. If these returns do not admit arbitrage opportunities, then there exists a positive random variable m_{t+1} satisfying (5) for dollar returns on each asset (Duffie 1992, Theorem 1A and

6

extensions). For any such m, the choice $m_{t+1}^* = m_{t+1}S_{t+1}/S_t$ automatically satisfies (7). ∎

The proposition tells us that of the three random variables — m_{t+1}, m_{t+1}^*, and S_{t+1}/S_t — one is effectively redundant and can be constructed from the other two. Most of the existing literature uses the domestic pricing kernel m (or its equivalent expressed as state prices or risk-neutral probabilities) and the depreciation rate. We start instead with the two pricing kernels, which highlights the essential symmetry of the theory between the two currencies.

The essence of the proposition is that we can always choose an m^* that satisfies (7) and (8). Suppose, instead, that m^* satisfies (7), but departs from (8) by an arbitrary error η:

$$(S_{t+1}/S_t) = (m_{t+1}^*/m_{t+1})e^{\eta_{t+1}}.$$

Since we can convert dollar cash flows to pounds, and vice versa, this error must satisfy

$$
\begin{aligned}
E_t\left(m_{t+1}^* R_{t+1}^*\right) &= E_t\left[m_{t+1}(S_{t+1}/S_t)R_{t+1}^*\right] &= E_t\left(m_{t+1}^* e^{\eta_{t+1}} R_{t+1}^*\right) \qquad (9)\\
E_t\left(m_{t+1} R_{t+1}\right) &= E_t\left[m_{t+1}^*(S_t/S_{t+1})R_{t+1}\right] &= E_t\left(m_{t+1} e^{-\eta_{t+1}} R_{t+1}\right)
\end{aligned}
$$

for all feasible returns R and R^*. Given such an error, we can satisfy the proposition by choosing a different foreign pricing kernel $\hat{m}_{t+1}^* = m_{t+1}^* \exp(\eta_{t+1})$.

Missing from the proposition is a position on purchasing power parity. Although it presumes frictionless trade in (some) assets, including currencies, it applies equally to environments in which costs of shipping goods across countries lead to departures from purchasing power parity (Hollifield and Uppal, 1997) and those in which purchasing power parity holds exactly (Bakshi and Chen, 1997).

3.2 Forward Rates and Risk Premiums

We can now relate the risk premium defined by Fama to properties of the two pricing kernels. Consider a forward contract specifying at date t the exchange at $t+1$ of one pound and $F_t = \exp(f_t)$ dollars, with the forward rate F_t set at date t as the notation suggests. This contract specifies a net dollar cash flow at date $t+1$ of $F_t - S_{t+1}$. Since it involves no payments at date t, pricing relation (4) implies

$$0 = E_t\left[m_{t+1}(F_t - S_{t+1})\right].$$

Dividing by S_t and applying Proposition 1, we find

$$(F_t/S_t)E_t(m_{t+1}) = E_t\left(m_{t+1}S_{t+1}/S_t\right) = E_t(m_{t+1}^*).$$

7

Thus the forward premium is

$$f_t - s_t = \log E_t m_{t+1}^* - \log E_t m_{t+1}. \tag{10}$$

This equation and definitions of the short rate [equation (6) for domestic rate and an analogous relation for the foreign rate], give us $f_t - s_t = r_t - r_t^*$, the familiar covered interest rate parity condition.

Now consider the components of the forward premium. The expected rate of depreciation is, from (8),

$$q_t \equiv E_t s_{t+1} - s_t = E_t \log m_{t+1}^* - E_t \log m_{t+1}. \tag{11}$$

Thus we see that the first of Fama's components is the difference in conditional means of the logarithms of the pricing kernels. The risk premium is, from (2,10),

$$p_t = (\log E_t m_{t+1}^* - E_t \log m_{t+1}^*) - (\log E_t m_{t+1} - E_t \log m_{t+1}), \tag{12}$$

the difference between the "log of the expectation" and the "expectation of the log" of the pricing kernels m and m^*.

With additional structure we can be more specific about the factors that affect the risk premium. Many popular models of bond and currency prices, including the affine models we examine shortly, start with conditionally log-normal pricing kernels: $\log m_{t+1}$ and $\log m_{t+1}^*$ are conditionally normal with (say) means (μ_{1t}, μ_{1t}^*) and variances (μ_{2t}, μ_{2t}^*). With this structure, one-period bond prices are

$$E_t m_{t+1} = \exp(\mu_{1t} + \mu_{2t}/2)$$
$$E_t m_{t+1}^* = \exp(\mu_{1t}^* + \mu_{2t}^*/2)$$

and the risk premium is

$$p_t = (\mu_{2t}^* - \mu_{2t})/2. \tag{13}$$

Fama's conditions require, in this case, (i) negative correlation between differences in conditional means and conditional variances of the two pricing kernels and (ii) greater variation in one-half the difference in the conditional variances. We need, in short, a great deal of variation in conditional variances.

If the conditional distributions of $\log m$ and $\log m^*$ are not normal, the risk premium also depends on moments of order three and higher. For an arbitrary distribution, equation (11) tells us (again) that only the means affect the expected rate of depreciation. The risk premium is given, in general, by (12), but if all of the conditional moments of $\log m_{t+1}$ exist, $\log E_t m_{t+1}$ can be expanded as

$$\log E_t m_{t+1} = \sum_{j=1}^{\infty} \kappa_{jt}/j!, \tag{14}$$

8

where κ_{jt} is the jth cumulant for the conditional distribution of $\log m_{t+1}$. Equation (14) is an expansion of the cumulant generating function (the logarithm of the moment generating function) evaluated at one; see, for example, Stuart and Ord (1987, chs 3,4). Cumulants are closely related to moments, as we see from the first four: $\kappa_{1t} = \mu_{1t}$, $\kappa_{2t} = \mu_{2t}$, $\kappa_{3t} = \mu_{3t}$, and $\kappa_{4t} = \mu_{4t} - 3(\mu_{2t})^2$. The notation is standard, with μ_{1t} denoting the conditional mean of $\log m_{t+1}$ and μ_{jt}, for $j > 1$, denoting the jth central conditional moment. For the normal distribution, cumulants are zero after the first two, so equation (14) gives us a way of quantifying the impact of departures from normality. If the foreign kernel has a similar representation, the forward premium is

$$f_t - s_t = \sum_{j=1}^{\infty} (\kappa_{jt}^* - \kappa_{jt})/j!,$$

and the risk premium is

$$p_t = \kappa_{-1,t}^* - \kappa_{-1,t}, \tag{15}$$

where

$$\kappa_{-1,t} \equiv \sum_{j=2}^{\infty} \kappa_{jt}/j!, \qquad \kappa_{-1,t}^* \equiv \sum_{j=2}^{\infty} \kappa_{jt}^*/j!.$$

We refer generically to the sums $\kappa_{-1,t}$ and $\kappa_{-1,t}^*$ as "higher-order cumulants."

With equations (15) and (11) describing risk premiums and expected rates of depreciation, we have

Remark 1 *If conditional moments of all order exist for the logarithms of the two pricing kernels, m and m^*, then Fama's necessary conditions for the forward premium anomaly imply (i) negative correlation between differences in conditional means, $\mu_{1t}^* - \mu_{1t}$, and differences in higher-order cumulants, $\kappa_{-1,t}^* - \kappa_{-1,t}$, and (ii) greater variation in the latter. A necessary and sufficient condition is a negative covariance between $q_t = \mu_{1t}^* - \mu_{1t}$ and $f_t - s_t = \mu_{1t}^* - \mu_{1t} + \kappa_{-1,t}^* - \kappa_{-1,t}$.*

This characterization of the risk premium suggests an interpretation for the failure of GARCH-M models, which model the risk premium as a function of the conditional variance of the depreciation rate. Studies by Bekaert (1995), Bekaert and Hodrick (1993), and Domowitz and Hakkio (1985) document strong evidence of time-varying conditional variances of depreciation rates, but little that connects the conditional variance to the risk premium p. One view of this failure is that GARCH-M models violate our sense of symmetry: an increase in the conditional variance of the depreciation rate increases risk on both sides of the market, and hence carries no presumption in favor of one currency or the other. Our framework indicates why. The conditional variance of the depreciation rate is

$$Var_t(s_{t+1} - s_t) = Var_t(\log m_{t+1}^* - \log m_t),$$

the conditional variance of the difference between the logarithms of the two kernels. The risk premium, on the other hand, is half the difference in the conditional variances [equation (13)] and possibly higher moments [equation (15)], which need bear no specific relation to the conditional variance of the depreciation rate. GARCH-M models, to put it simply, focus on a different conditional variance.

3.3 The General Affine Model

Further progress requires more structure. We explore an adaptation to currencies of the Duffie and Kan (1996) class of affine yield models, whose relevant properties are summarized in Appendix A. The linearity of these models makes it relatively to easy to explore the implications of different structures and parameter values for the forward premium anomaly.

This class of affine currency models starts with a vector z of state variables following the law of motion

$$z_{t+1} = (I - \Phi)\theta + \Phi z_t + V(z_t)^{1/2}\varepsilon_{t+1} \tag{16}$$

where $\{\varepsilon_t\} \sim \text{NID}(0, I)$, Φ is stable with positive diagonal elements, and V is diagonal with typical element

$$v_i(z) = \alpha_i + \beta_i^\top z.$$

Further conditions on the parameters guarantee that the state z never leaves the region defined by nonnegative values of the volatility functions v_i; see Appendix A for details. Given z, pricing kernels have the form

$$-\log m_{t+1} = \delta + \gamma^\top z_t + \lambda^\top V(z_t)^{1/2}\varepsilon_{t+1} \tag{17}$$
$$-\log m_{t+1}^* = \delta^* + \gamma^{*\top} z_t + \lambda^{*\top} V(z_t)^{1/2}\varepsilon_{t+1}, \tag{18}$$

and the depreciation rate is

$$s_{t+1} - s_t = (\delta - \delta^*) + (\gamma - \gamma^*)^\top z_t + (\lambda - \lambda^*)^\top V(z_t)^{1/2}\varepsilon_{t+1}. \tag{19}$$

Thus, log kernels and depreciation rates are conditionally normal with conditional means and variances that are linear in the state z.

Application of equation (8) of Proposition 1 comes without loss of generality in this environment. The presumption of the model is that predictable and unpredictable movements in log bond prices and depreciation rates are spanned by the state z and the innovations $V(z)^{1/2}\varepsilon$. From this, any error in (8) is affine, and

10

Proposition 1 might be satisfied only by an alternative affine choice of foreign pricing kernel,

$$-\log \hat{m}^*_{t+1} \quad = \quad (\delta^* + \delta') + (\gamma^* + \gamma')^\top z_t + (\lambda^* + \lambda')^\top V(z_t)^{1/2}\varepsilon_{t+1},$$

for some arbitrary nonzero choice of $(\delta', \gamma', \lambda')$. Equation (9) implies, however, that m^* and $\hat{m}^*$ price all traded assets the same way. In this sense, the two kernels are observationally equivalent, and there is no loss of generality in applying Proposition 1.

We turn now to the possibility of using models in this class to account for the forward premium anomaly.

4   Accounting for the Anomaly

Remark 1 suggests that it should be relatively easy to construct examples that reproduce the anomaly: we simply arrange for differences in first and second moments of pricing kernels to move in opposite directions. Consider a model like Engel and Hamilton's (1990) in which the conditional distributions of two pricing kernels alternate between two log-normal regimes. If the difference in conditional means of the pricing kernels is higher in regime 1, and one-half the difference in conditional variances is higher in regime 2, and varies more than the difference in means, then the model will reproduce the anomaly.

A greater challenge is to construct a model that mimics the properties of currency prices and interest rates more generally. We approach this problem with affine models, which have several advantages in this context. First, we have, as a profession, more than two decades' experience with affine models in pricing fixed income securities. Much of this experience can be transferred directly to currency pricing. Second, conditional means and variances of logarithms of pricing kernels are linear functions of a vector of state variables. As a result, we can easily compare their properties to Fama's necessary conditions for the forward premium anomaly. Finally, we will see shortly that many of the models in this class automatically generate the contrary movements in the conditional mean and variance of pricing kernels suggested by Fama's condition (i) in log-normal settings.

4.1   Example 1: Two-Currency Cox-Ingersoll-Ross

An obvious starting point is a two-currency version of Cox, Ingersoll, and Ross (1985). Our version is adapted from Sun's (1992) discrete-time translation.

The model is based on two state variables, indexed by $i = 1, 2$, that obey independent "square-root" processes

$$z_{it+1} = (1 - \varphi_i)\theta_i + \varphi_i z_{it} + \sigma_i z_{it}^{1/2} \varepsilon_{it+1}, \qquad (20)$$

with $0 < \varphi_i < 1$, $\theta_i > 0$, and $\{\varepsilon_{it}\} \sim \text{NID}(0, 1)$. The unconditional mean of z_i is θ_i, the autocorrelation is φ_i, the conditional variance is $\sigma_i^2 z_{it}$, and the unconditional variance is $\sigma_i^2 \theta_i / (1 - \varphi_i^2)$. A variant of (20),

$$z_{it+1} - z_{it} = (1 - \varphi_i)(\theta_i - z_{it}) + \sigma_i z_{it}^{1/2} \varepsilon_{it+1},$$

is a direct analog of the continuous-time original (Cox, Ingersoll, and Ross, 1985, eq 17). A salient feature of (20) is the square-root term in the innovation, whose conditional variance falls to zero as z_i approaches zero. In continuous time, this feature and the *Feller condition*,

$$\frac{2(1 - \varphi_i)\theta_i}{\sigma_i^2} \geq 1, \qquad (21)$$

guarantee that z_i remains positive. In discrete time, z_i can turn negative with a large enough negative realization of ε_i. This happens with positive probability, but the probability approaches zero as the time interval goes to zero (Sun, 1992).

In the standard one-factor Cox-Ingersoll-Ross model, z_1 (say) is the dollar short rate and the pricing kernel is

$$-\log m_{t+1} = (1 + \lambda_1^2/2)z_{1t} + \lambda_1 z_{1t}^{1/2} \varepsilon_{1t+1}, \qquad (22)$$

a special case of (18). The coefficient of z_1 makes it the one-period rate of interest. The parameter λ_1 controls the covariance of the kernel with movements in interest rates and thus governs the risk of long bonds and the average slope of the yield curve.

This structure is an example of the conditionally log-normal pricing kernels described in Section 3. Moreover, equation (22) builds in an inverse relation between the conditional mean and variance of the logarithm of the pricing kernel, as required by Fama's condition (i). The conditional mean and variance,

$$
\begin{aligned}
E_t \log m_{t+1} &= -(1 + \lambda_1^2/2)z_{1t} \\
Var_t \log m_{t+1} &= \lambda_1^2 z_{1t},
\end{aligned}
$$

are both linear in the state variable z_1. The short rate is therefore

$$r_t = -\log E_t m_{t+1} = -\left(E_t \log m_{t+1} + \frac{1}{2} Var_t \log m_{t+1} \right) = z_{1t}, \qquad (23)$$

as claimed earlier.

A natural extension to two currencies is to posit an analogous pricing kernel for valuing foreign-currency cash flows. The pricing kernel m^* for (say) pounds is based on a second state variable z_2 and follows

$$- \log m^*_{t+1} = (1 + \lambda_2^2/2)z_{2t} + \lambda_2 z_{2t}^{1/2} \varepsilon_{2t+1}. \tag{24}$$

Then the pound short rate is $r_t^* = z_{2t}$ and the forward premium is

$$f_t - s_t = z_{1t} - z_{2t}.$$

If we impose (purely for convenience) the symmetry restriction $\lambda_1 = \lambda_2$, we can write the expected depreciation rate as $q_t = (1 + \lambda_1^2/2)(z_{1t} - z_{2t})$ and the risk premium as $p_t = -(\lambda_1^2/2)(z_{1t} - z_{2t})$. Thus the linearity of the conditional mean and variance translate into forward premium components that are linear functions of the differential $z_1 - z_2$. More important, this structure automatically generates the negative correlation between p and q of Fama's condition (i): since equation (22) implies an inverse relation between the conditional mean and variance of $\log m_{t+1}$, and the two pricing kernels are independent, the difference in conditional means is inversely related to the difference in conditional variances.

This model cannot, however, satisfy Fama's condition (ii) or reproduce the anomalous regression slope. If we regress the depreciation rate on the forward premium in this model, the slope is

$$a_2 = 1 + {\lambda_1}^2/2.$$

The slope is not only positive, and therefore inconsistent with the anomaly, it exceeds one, and is therefore inconsistent even with the Flood and Rose (1996) evidence for the ERM.

4.2   Example 2: Negative Factors

We turn next to a generalization of example 1 that is capable of satisfying both of Fama's necessary conditions, but abandons the trademark positive interest rates of the Cox-Ingersoll-Ross model. We also add a factor that is common to both currencies, which allows us to account for nonzero correlation of interest rates across currencies but does not otherwise bear on the ability of the model to account for the anomaly.

Consider a two-currency world based on three independent state variables, a common state variable z_0 and (as before) currency-specific state variables z_1 and z_2,

each following (20). Pricing kernels in the domestic and foreign currency are

$$-\log m_{t+1} = (1 + \lambda_0^2/2)z_{0t} + (-1 + \lambda_1^2/2)z_{1t} + \lambda_0 z_{0t}^{1/2}\varepsilon_{0t+1} + \lambda_1 z_{1t}^{1/2}\varepsilon_{1t+1}$$
$$-\log m_{t+1}^* = (1 + \lambda_0^2/2)z_{0t} + (-1 + \lambda_2^2/2)z_{2t} + \lambda_0 z_{0t}^{1/2}\varepsilon_{0t+1} + \lambda_2 z_{2t}^{1/2}\varepsilon_{2t+1},$$

with $\{\varepsilon_{it}\}$ independent standard normal random variables. The depreciation rate is therefore

$$s_{t+1} - s_t = (-1 + \lambda_1^2/2)(z_{1t} - z_{2t}) + \lambda_1(z_{1t}^{1/2}\varepsilon_{1t+1} - z_{2t}^{1/2}\varepsilon_{2t+1}), \tag{25}$$

where we have again imposed the symmetry condition, $\lambda_1 = \lambda_2$, for analytic convenience.

The new ingredients are the common factor z_0 and the coefficients of z_1 and z_2 in the pricing kernels. Since the common factor z_0 and its innovation ε_0 affect both pricing kernels the same way, they have no effect on currency prices or interest differentials. They therefore have no effect on the slope parameter a_2 that characterizes the forward premium anomaly. The other change, however, bears directly on the anomaly. The coefficients of (z_0, z_1, z_2) and z_2 imply short-term interest rates of $r_t = z_{0t} - z_{1t}$ and $r_t^* = z_{0t} - z_{2t}$. The forward premium is $-(z_{1t} - z_{2t})$ and expected depreciation is $(-1 + \lambda_1^2/2)(z_{1t} - z_{2t})$. The regression slope is therefore

$$a_2 = 1 - \lambda_1^2/2, \tag{26}$$

which is always less than one, and negative for large enough values of λ_1.

We thus have a model that can account for the forward premium anomaly. The cost is that we can no longer maintain strictly positive interest rates. Whether this is a serious difficulty for the model is largely an empirical issue that we address later in this section. If a small probability of negative interest rates leads to an affine model that is realistic in other respects, we might regard this a small cost paid for the convenience of linearity. Duffie and Singleton (1997) and Pearson and Sun (1994) make similar arguments in extending the Cox-Ingersoll-Ross model of bond pricing.

4.3 A Proposition

Examples 1 and 2 illustrate a general feature of affine models based on independent interest-rate factors. If we restrict ourselves to models in which foreign and domestic interest rates depend on independent factors, so that the only connection between them is a common factor that affects both rates the same way, then we cannot simultaneously account for the anomaly and maintain strictly positive interest rates.

We refer to this generalization of example 2 as an independent factor model. As in example 2, the state vector has three independent components, but here we allow them to be vectors: a vector z_0 of common factors that affect interest rates in both currencies the same way and vectors z_1 and z_2 of currency-specific factors. If we express the components of parameter vectors with similar subscripts, the independent factor structure we have in mind places these restrictions on the general affine model of Subsection 3.3:

- The three components of z are independent: Φ is block-diagonal with elements Φ_i for $i = 0, 1, 2$.

- The common factor affects both pricing kernels the same way: $\gamma_0 = \gamma_0^*$ and $\lambda_0 = \lambda_0^*$.

- The currency-specific factors affect the pricing kernel of only one currency: $\gamma_2 = \lambda_2 = 0$ and $\gamma_1^* = \lambda_1^* = 0$.

Ahn (1998) and Amin and Jarrow (1991) consider similar structures.

As result of this structure, these models exhibit the same tension between the anomaly and positive interest rates that we observed in examples 1 and 2:

Proposition 2 *Consider the affine class of currency models described by equations (16,17,18,19) with the independent factor structure described above. If such a model implies positive bond yields for all admissible values of the state variables, then it cannot generate a negative value of the slope parameter α_2 from forward premium regressions.*

A proof is given in Appendix B, but the intuition follows examples 1 and 2. The affine models permitted in Proposition 2 are based on state variables that are unbounded in one direction. In both examples, state variables z_i assume all positive values with positive probability. Such state variables have two effects on the short rate, one operating through the mean of the pricing kernel, the other through the variance. An increase in the conditional mean tends to raise the short rate, while an increase in the conditional variance lowers it. If the mean effect is larger, as it is in the Cox-Ingersoll-Ross model, then the short rate is unbounded above. The anomaly requires instead that the effect of the variance must be larger, and thus that increases in variance be associated with decreases in the short rate. But since the conditional variance is unbounded above, the short rate will be negative for large enough values of the state variable.

4.4 Example 3: Interdependence with One Factor

Example 2 and Proposition 2 indicate that one approach to explaining the anomaly is to abandon the requirement of strictly positive interest rates. Another is to posit interdependence between factors: domestic and foreign interest rates depend in different ways on the same state variable.

A one-factor example shows how this might work. Consider a model based on a single state variable z_1 obeying (20) with pricing kernels

$$
\begin{aligned}
-\log m_{t+1} &= (1 + \lambda_1^2/2)z_{1t} + \lambda_1 z_{1t}^{1/2}\varepsilon_{1t+1} \\
-\log m_{t+1}^* &= (\gamma_2 + \lambda_2^2/2)z_{1t} + \lambda_2 z_{1t}^{1/2}\varepsilon_{1t+1}.
\end{aligned}
$$

The model is interdependent in the sense that z_1 affects the two pricing kernels differently if $(1, \lambda_1) \neq (\gamma_2, \lambda_2)$. In this setting, short rates are $r_t = z_{1t}$ and $r_t^* = \gamma_2 z_{1t}$, so the forward premium is $f_t - s_t = (1 - \gamma_2)z_{1t}$. Both interest rates are strictly positive if $\gamma_2 > 0$. The depreciation rate is

$$
s_{t+1} - s_t = \left[1 - \gamma_2 + (\lambda_1^2 - \lambda_2^2)/2\right] z_{1t} + (\lambda_1 - \lambda_2)z_{1t}^{1/2}\varepsilon_{1t+1},
$$

so expected depreciation is $q_t = [1 - \gamma_2 + (\lambda_1^2 - \lambda_2^2)/2]z_{1t}$. The slope of the forward premium regression is therefore

$$
a_2 = 1 + \frac{\lambda_1^2 - \lambda_2^2}{2(1 - \gamma_2)}, \tag{27}
$$

which differs from one if $\lambda_1^2 \neq \lambda_2^2$. If $\gamma_2 = 1$ the forward premium is zero: the model has no forward premium and thus no forward premium anomaly. For other values, the model implies an inverse relation between the forward premium and the interest differential if $2(1 - \gamma_2)$ and $\lambda_1^2 - \lambda_2^2$ have opposite signs and the latter is larger in absolute value. Ahn (1997) and Frachot (1996) describe a similar examples in continuous time.

4.5 Example 4: Interdependence with Two Factors

We can easily extend example 3 to a more realistic two-factor setting in which the correlation between foreign and domestic interest rates is imperfect. Consider a model based on two state variables, z_1 and z_2, obeying identical independent square root processes (20), and pricing kernels

$$
\begin{aligned}
-\log m_{t+1} &= (1 + \lambda_1^2/2)z_{1t} + (\gamma_2 + \lambda_2^2/2)z_{2t} + \lambda_1 z_{1t}^{1/2}\varepsilon_{1t+1} + \lambda_2 z_{2t}^{1/2}\varepsilon_{2t+1} \\
-\log m_{t+1}^* &= (\gamma_2 + \lambda_2^2/2)z_{1t} + (1 + \lambda_1^2/2)z_{2t} + \lambda_2 z_{1t}^{1/2}\varepsilon_{1t+1} + \lambda_1 z_{2t}^{1/2}\varepsilon_{2t+1}.
\end{aligned}
$$

Bakshi and Chen (1997), Nielsen and Saá-Requejo (1993, pp 9-10), and Saá-Requejo (1994, p 17) describe similar models. Our version is symmetric in the sense that the unconditional distributions of the two pricing kernels are the same, but interdependent in the sense that the state variables z_1 and z_2 potentially affect the two kernels in different ways.

Consider the implications for the anomaly. Short rates in this model are

$$
\begin{aligned}
r_t &= z_{1t} + \gamma_2 z_{2t} \\
r_t^* &= \gamma_2 z_{1t} + z_{2t}.
\end{aligned}
$$

The forward premium,

$$
f_t - s_t = r_t - r_t^* = (1 - \gamma_2)(z_{1t} - z_{2t}).
$$

and depreciation rate,

$$
s_{t+1} - s_t = \left[1 - \gamma_2 + (\lambda_1^2 - \lambda_2^2)/2\right](z_{1t} - z_{2t}) + (\lambda_1 - \lambda_2)(z_{1t}^{1/2}\varepsilon_{1t+1} - z_{2t}^{1/2}\varepsilon_{2t+1}),
$$

imply a regression slope of

$$
a_2 = 1 + \frac{\lambda_1^2 - \lambda_2^2}{2(1 - \gamma_2)},
$$

a repeat of equation (27). As in example 3, appropriate choice of parameters allows us to generate a negative value. The critical feature in this regard is, again, that state variables affect the two kernels differently.

In example 4, interdependence takes a particularly striking form. Suppose (without loss of generality) that $\gamma_2 < 1$. From the short rate equations, we might say that z_1 is the "dollar factor," since it has a greater effect on the dollar short rate than z_2. For similar reasons we might refer to z_2 as the pound factor. But the anomaly (in fact, any value of a_2 less than one) implies $\lambda_2^2 > \lambda_1^2$, implying that innovations in the "pound factor" have greater influence on the dollar kernel than do innovations in the dollar factor. It's as if (to use a concrete example) US money growth had a larger influence than British monetary policy on dollar interest rates, but a smaller influence on the dollar pricing kernel.

4.6   Estimates of Examples 2 and 4

Examples 2, 3, and 4 suggest that the set of affine currency models that are consistent with the forward premium anomaly is potentially large. In this section, we report estimates of examples 2 and 4 and comment on their ability to account for both the anomaly and features of currencies and interest rates in general.

Our approach is to estimate the parameters of examples 2 and 4 by GMM using one moment condition for each parameter. The spirit is similar to Constantinides' (1992) study of interest rates, but the formalism of GMM provides us with standard errors for the estimated parameters. In each example, we use data on the dollar-pound exchange rate, the dollar short rate, and the dollar-pound forward premium. All of the relevant sample moments are reported in Tables 1 and 2.

In Table 3, we report estimates of parameter values and standard errors for example 2 (the independent factor model) and example 4 (the interdependent factor model). The two models illustrate two ways to account for the anomaly: negative interest rates and interdependence, respectively. For example 2, we choose parameters to reproduce these moments: the mean, variance, and autocorrelation of the dollar short rate; the variance and autocorrelation of the dollar-pound forward premium; the variance of the dollar-pound depreciation rate; and the regression slope a_2 of equation (1) that we use to characterize the anomaly.

These estimates allow us to assess the quantitative impact of Proposition 2. Since all three state variables range between zero and infinity, short rates ($r_t = z_{0t} - z_{1t}$ and $r_t^* = z_{0t} - z_{2t}$) are negative with positive probability, as required by the proposition. We note in Appendix C that the state variables have (approximately) gamma distributions. With estimated parameters, we find that the probability of negative r is less than 10^{-5}. The difficulty, instead, is that the model exhibits extreme distributional properties for the currency-specific factors and the forward premium. The problem can be traced directly to the anomaly. The regression coefficient dictates the large value of λ_1^2 that we see in Table 3; see equation (26). But the variance of the depreciation rate dictates a small value of $\lambda_1^2 \theta_1$:

$$Var(s_{t+1} - s_t) \geq \lambda_1^2 \theta_1;$$

see equation (25). Since the monthly standard deviation of the depreciation rate is about 3% (see Table 1), we must choose an extremely small value for θ_1. This value, in turn, implies that we violate the Feller condition (21) by more than two orders of magnitude:

$$\frac{2(1 - \varphi_1)\theta_1}{\sigma_1^2} = 0.003 \not\geq 1.$$

This ratio governs the distributional properties of z_1, and implies extreme values for its skewness and kurtosis; see Appendix C. None of these features is peculiar to the dollar-pound rate: parameter estimates are similar for the other currencies described in Tables 1 and 2.

Table 3 also lists estimates for the interdependent factor model, example 4. There is one difference in the moments used to estimate parameter values: we drop the autocorrelation of the short rate, since the model cannot reproduce different values

for the autocorrelation of the short rate and the forward premium. The resulting estimates do not have the difficulty noted for those of example 2. The Feller ratio,

$$\frac{2(1 - \varphi_1)\theta_1}{\sigma_1^2} = 3.455 \geq 1,$$

easily satisfies the Feller condition. As a result, the unconditional distributions of interest rates do not exhibit the extreme behavior we noted for example 2.

5   Final Remarks

We have studied the relation between interest rates and currency prices in theory and data. The relation shows up in theory as a restriction on the joint behavior of currencies and state prices in arbitrage-free environments. In the data, it appears in the form of the forward premium anomaly, in which interest rate differentials help to predict future movements in currency prices. We find that the anomaly imposes further conditions on affine models: either interest rates must be negative with positive probability or the effects of one or more factors on pricing kernels must differ across currencies (what we term interdependence). Our estimates suggest that within the class of affine models, those with interdependence offer the best hope of accounting for the properties of currency prices and interest rates in general.

We are left with two outstanding issues. The first is whether affine models with a small number of state variables are capable of approximating the properties of currency prices and a more comprehensive set of fixed income securities. Ahn (1997) and Saá-Requejo (1994) have made some progress along these lines, extending the analysis to yields on bonds with longer maturities. The second is the economic foundations of pricing kernels that reproduce the anomaly. We have followed a "reverse engineering" strategy in which pricing kernels are simply stochastic processes that account for observed asset prices, but one might reasonably ask what kinds of behavior by private agents and policy makers might lead to such pricing kernels. Several possibilities are outlined by Alvarez and Atkeson (1996), Bakshi and Chen (1997), Stulz (1987), and Yaron (1995), who develop dynamic general equilibrium models in which interest rates and currency prices reflect preferences, endowments, and monetary policies. Perhaps further work will connect pricing kernels in these models to properties of interest rates, currency prices, and monetary aggregates.

19

A Affine Models

We outline a class of affine yield models adapted from Duffie and Kan (1996) and translated into discrete time. This section explains the relation between Duffie and Kan's class of affine models and our discrete-time analog and sets the stage for the affine currency models of Sections 3 and 4.

Expressed in discrete time, Duffie and Kan's affine models are based on a k-dimensional vector of state variables z that follow

$$z_{t+1} - z_t = (I - \Phi)(\theta - z_t) + V(z_t)^{1/2}\varepsilon_{t+1}, \tag{28}$$

where $\{\varepsilon_t\} \sim \text{NID}(0, I)$, $V(z)$ is a diagonal matrix with typical element

$$v_i(z) = \alpha_i + \beta_i^\top z,$$

β_i has nonnegative elements, and Φ is stable with positive diagonal elements. State prices are governed by a pricing kernel of the form

$$-\log m_{t+1} = \delta + \gamma^\top z_t + \lambda^\top V(z_t)^{1/2}\varepsilon_{t+1}. \tag{29}$$

The process for z requires that the volatility functions v_i be positive. We define the set D of *admissible* states as those values of z for which volatility is positive:

$$D = \{z : v_i(z) \geq 0 \text{ all } i\}.$$

Duffie and Kan (1996, Section 4) show that z remains in D if the process satisfies

Condition A *For each i:*

(a) *for all $z \in D$ satisfying $v_i(z) = 0$ (the boundary of positive volatility), the drift is sufficiently positive: $\beta_i^\top(I - \Phi)(\theta - z) > \beta_i^\top \beta_i/2$; and*

(b) *if the jth component of β_i is nonzero for any $j \neq i$ then $v_i(z)$ and $v_j(z)$ are proportional to each other (their ratio is a positive constant).*

We refer to models characterized by (28,29) and satisfying Condition A as the *Duffie-Kan class of affine models*.

Our description of these models in Subsection 3.3 differs from Duffie and Kan's in a number of respects. First, Duffie and Kan write (28) as

$$z_{t+1} - z_t = (I - \Phi)(\theta - z_t) + \Sigma V(z)^{1/2}\varepsilon_{t+1}, \tag{30}$$

which includes a matrix Σ that is missing in our version. We show that our choice is innocuous by reducing their model to ours. Assume Σ is invertible (this is convenient but not essential) and define $z' = \Sigma^{-1}z$. If we substitute for z, equation (30) becomes

$$z'_{t+1} - z'_t = (I - \Phi')(\theta' - z'_t) + V'(z')^{1/2}\varepsilon_{t+1},$$

with $v'_i(z') = \alpha_i + \beta_i'^\top z'$, $\Phi' = \Sigma^{-1}\Phi\Sigma$, $\theta' = \Sigma^{-1}\theta$, and $\beta_i'^\top = \beta_i^\top\Sigma$. Equation (29) becomes

$$-\log m_{t+1} = \delta + \gamma'^\top z'_t + \lambda^\top V'(z'_t)^{1/2}\varepsilon_{t+1}.$$

with $\gamma'^\top = \gamma^\top\Sigma$. Thus we have effectively eliminated Σ from the model.

A second difference from Duffie and Kan is the assumption that the volatility parameters β_i are nonnegative, which comes without loss of generality. Define the matrix $\beta = (\beta_1, \ldots, \beta_k)$ with β_{ij} denoting the jth element of β_i and the (i,j)th element of $\beta^\top$. Note that we can choose the diagonal elements of β to be nonnegative: if $\beta_{ii} < 0$ for any i, we replace z_i and $-z_i$ and β_{ii} with $-\beta_{ii}$ and change the other parameters in the model accordingly. This produces a matrix β with positive diagonal elements. Condition A(b) tells us that if β has nonzero off-diagonal elements, then they are proportional to diagonal elements and hence positive as well.

The class of affine currency models we describe in Subsection 3.3 simply adds to this environment a second pricing kernel for the foreign currency.

B   Proof of Proposition 2

We prove Proposition 2, starting with some preliminary results on affine models in which bond yields are strictly positive.

In the affine models of equations (28,29), bond prices are log-linear functions of the state variables z. If b_t^n denotes the price at date t of a claim to one dollar in all states at date $t + n$, then

$$-\log b_t^n = A(n) + B(n)^\top z_t$$

for some parameters $\{A(n), B(n)\}$. Since bond yields are $y_t^n = -n^{-1}\log b_t^n$, they are linear in z:

$$y_t^n = A'(n) + B'(n)z_t,$$

where $A'(n) = n^{-1}A(n)$ and $B'(n) = n^{-1}B(n)$. We use the pricing relation (4) to generate parameters recursively:

$$A(n+1) \quad = \quad A(n) + \delta + B(n)^\top(I - \Phi)\theta - \frac{1}{2}\sum_{j=1}^{k}(\lambda_j + B(n)_j)^2\,\alpha_j$$

21

$$B(n+1)^\top = \left(\gamma^\top + B(n)^\top \Phi \right) - \frac{1}{2} \sum_{j=1}^{k} \left(\lambda_j + B(n)_j \right)^2 \beta_j^\top,$$

starting with $A(0) = 0$ and $B(0) = 0$. We say that a model is *invertible* if there exist k maturities for which the matrix

$$B = \left[B'(n_1) \cdots B'(n_k) \right]$$

is nonsingular. The assumption of invertibility is not restrictive: if a model is not invertible, we can construct an equivalent invertible model with a smaller state vector. Analogously, define the vector $A^\top = [A'(n_1), \ldots, A'(n_k)]$.

Consider now a subclass of affine models in which bond yields are always positive:

Lemma 1 *Consider the Duffie-Kan class of affine models. If the model is invertible and bond yields are positive for all admissible states z, then β is diagonal with strictly positive elements.*

In words: the volatility functions have the univariate square root form

$$v_i(z) = \alpha_i + \beta_{ii} z_i,$$

with strictly positive β_{ii}. As a consequence, β has full rank. This rules out both pure Gaussian factors like $\beta_i^\top = (0, 0, \ldots, 0)$ and multivariate factors like $\beta_i^\top = (1, 1, \ldots, 1)$.

Proof. Suppose, in contradiction to the lemma, that β has less than full rank. Then there exists a nonzero vector h satisfying $\beta^\top h = 0$. For any admissible z, $z' = z + \rho h$ is also admissible for any real ρ since it generates the same values for the volatility functions. Now consider bond yields. For yields to be positive we need bond prices to be less than one. If y denotes a vector of yields for a set of maturities for which B is invertible, then we need

$$y = A + B^\top z \geq 0$$

for all admissible z. Since $z' = z + \rho h$ is also admissible, we have

$$y = A + B^\top z + \rho B^\top h.$$

By assumption, B is invertible so $B^\top h \neq 0$. Thus we can choose ρ to make yields as negative as we like, thereby violating the premise of the lemma. We conclude that β has full rank. Condition A(b) then tells us that β must be diagonal. ∎

A second result is that in this environment (univariate volatility functions), $Var(z)$ has nonnegative elements:

22

Lemma 2 *Consider the Duffie-Kan class of affine models in which β is diagonal with strictly positive elements β_{ii}. Then $\text{Var}(z)$ has all positive elements.*

The proof hinges on Condition A(a), Duffie and Kan's multivariate analog of the Feller condition. Since β is diagonal with positive elements, the condition implies that for each i

$$\sum_{j=1}^{k} \kappa_{ij}(\theta_j - z_j) \geq \beta_{ii}/2 > 0.$$

for all admissible z satisfying $v_i(z) = 0$, where $K = I - \Phi$ has elements κ_{ij}. The new ingredient relative to the univariate Feller condition is the effect of variables $z_j, j \neq i$, on the drift of z_i. The structure placed on β means that the set of admissible z's includes values of z_j that are arbitrarily large. The condition therefore implies $\kappa_{ij} \leq 0$ for all $j \neq i$. The admissible set also includes $z_j = \theta_j$, so $\kappa_{ii} > 0$. Since by assumption $\Phi = I - K$ has positive diagonal elements, $0 < \kappa_{ii} < 1$. This additional step is an artifact of our discrete time approximation: if κ_{ii} were greater than one, we would simply choose a smaller time interval.

We have established that the elements of Φ are nonnegative. We now show that the unconditional variance of z, which we denote by the matrix Ω, has no negative elements. Since z is a first-order autoregression with stable Φ, its variance is the solution to

$$\Omega = \Phi\Omega\Phi^{\top} + V(\theta),$$

where $V(\theta)$ is a diagonal matrix with positive elements $v_i(\theta_i) = \alpha_i + \beta_{ii}\theta_i$. Since Φ is stable, we can compute Ω iteratively using

$$\Omega_{j+1} = \Phi\Omega_j\Phi^{\top} + V(\theta),$$

starting with $\Omega_0 = 0$. We see that at each stage the elements of Ω_{j+1} are sums of products of nonnegative numbers, so we conclude that the elements of Ω are nonnegative. $\blacksquare$

We come at last to a proof of Proposition 2. The proposition is based on the model described in Subsection 4.3. It has three independent state variables or factors: a common factor z_0 and currency-specific factors z_1 and z_2. The common factor has, by construction, no influence on currency prices or the forward premium. It therefore has no influence on the anomaly, and we can disregard it.

With this simplification, interest rates in the two currencies are

$$
\begin{aligned}
r_t &= (\delta - \omega) + (\gamma - \tau)^{\top} z_{1t} \\
r_t^* &= (\delta^* - \omega^*) + (\gamma^* - \tau^*)^{\top} z_{2t},
\end{aligned}
$$

where $\omega = \sum_j \lambda_j^2 \alpha_j^1$, $\omega^* = \sum_j \lambda_j^{*2} \alpha_j^2$, $\tau = \sum_j \lambda_j^2 \beta_j^1 / 2$, and $\tau^* = \sum_j \lambda_j^{*2} \beta_j^2 / 2$. The forward premium is therefore

$$f_t - s_t = (\gamma - \tau)^\top z_{1t} - (\gamma^* - \tau^*)^\top z_{2t}$$

and the depreciation rate is

$$s_{t+1} - s_t = (\delta - \delta^*) + \gamma^\top z_{1t} - \gamma^{*\top} z_{2t} + \lambda^\top V^1 (z_{1t})^{1/2} \varepsilon_{1t+1} - \lambda^{*\top} V^2 (z_{2t})^{1/2} \varepsilon_{2t+1}.$$

The anomaly therefore requires

$$Cov(s_{t+1} - s_t, f_t - s_t) = (\gamma - \tau)^\top Var(z_1) \gamma + (\gamma^* - \tau^*)^\top Var(z_2) \gamma^* < 0. \qquad (31)$$

The question is whether this is consistent with interest rates that are always positive.

The condition that interest rates are positive for all admissible states places restrictions on the parameters. For the "dollar" short rate r we need the elements of $\gamma - \tau$, and hence of γ, to be nonnegative, since $\tau > 0$ and each element of z_1 is unbounded above. By Lemma 2, $Var(z_1)$ has nonnegative elements, so the bilinear form

$$(\gamma - \tau)^\top Var(z_1) \gamma,$$

is nonnegative. Identical reasoning applies to the second term in (31). We conclude that the model cannot reproduce the anomaly with strictly positive interest rates. ∎

C   Distribution of State Variables

We clarify the role played by the Feller condition in determining higher moments of state variables following square root processes, an issue that arises in estimates of the independent factor model (Section 4.6).

Consider a state variable z following the square root process (20). The unconditional distribution of z is Gamma with density

$$f(z) = [b^a \Gamma(a)]^{-1} z^{a-1} e^{-z/b}.$$

and parameters $a, b > 0$. This statement is exact in continuous time, approximate in discrete time. The mean and variance of a Gamma random variable are $ab = \theta$ and $ab^2 = \theta \sigma^2 / (1 - \varphi^2)$, which defines the parameters as

$$\begin{aligned} a &= (1 - \varphi^2) \theta / \sigma^2 \\ b &= \sigma^2 / (1 - \varphi^2). \end{aligned}$$

In the continuous-time limit, $1 - \varphi^2 \rightarrow 2\kappa = 2(1 - \varphi)$, so $a \rightarrow 2(1 - \varphi)\theta/\sigma^2$, the ratio in the Feller condition, inequality (21). This ratio governs the distribution's higher moments:

$$\gamma_1 = \frac{E(z - \theta)^3}{[Var(z)]^{3/2}} = \frac{2}{a^{1/2}} \qquad \text{(skewness)}$$

$$\gamma_2 = \frac{E(z - \theta)^4}{[Var(z)]^2} = \frac{6}{a}. \qquad \text{(kurtosis)}$$

Parameter values implying $a > 1$ satisfy the Feller condition. Smaller values violate the condition and generate large values of the skewness and kurtosis measures, γ_1 and γ_2.

References

Ahn, Dong-Hyun, 1997, "Common factors and local factors: Implications for term structures and exchange rates," unpublished manuscript, University of North Carolina.

Ahn, Dong-Hyun, 1998, "Generalized squared-autoregressive-independent-variable term structure model," unpublished manuscript, University of North Carolina.

Amin, Kaushik, and Robert Jarrow, 1991, "Pricing foreign currency options under stochastic interest rates," *Journal of International Money and Finance* 10, 310-329.

Alvarez, Fernando, and Andrew Atkeson, 1996, "Liquidity, inflation, interest rates and exchange rates," unpublished manuscript, University of Pennsylvania, March.

Backus, David, Allan Gregory, and Chris Telmer, 1993, "Accounting for forward rates in markets for foreign currency," *Journal of Finance* 48, 1887-1908.

Bakshi, Gurdip, and Zhiwu Chen, 1997, "Equilibrium valuation of foreign exchange claims," *Journal of Finance* 52, 799-826.

Bansal, Ravi, 1997, "An exploration of the forward premium puzzle in currency markets," *Review of Financial Studies* 10, 369-304.

Bansal, Ravi, A. Ronald Gallant, Robert Hussey, and George Tauchen, 1995, *Journal of Econometrics* 66, 251-287.

Bekaert, Geert, 1995, "The time variation of expected returns and volatility in foreign exchange markets," *Journal of Business and Economic Statistics* 13, 397-408.

Bekaert, Geert, 1996, "The time-variation of risk and return in foreign exchange markets: A general equilibrium perspective," *Review of Financial Studies* 9, 427-470.

Bekaert, Geert, and Robert Hodrick, 1992, "Characterizing predictable components in excess returns on equity and foreign exchange markets," *Journal of Finance* 47, 467-509.

Bekaert, Geert, and Robert Hodrick, 1993, "On biases in the measurement of foreign exchange risk premiums," *Journal of International Money and Finance* 12, 115-138.

Bekaert, Geert, Robert Hodrick, and David Marshall, 1997, "The implications of first-order risk aversion for asset market risk premiums," *Journal of Monetary Economics* 40, 3-39.

Bossaerts, Peter, and Pierre Hillion, 1991, "Market microstructure effects of government intervention in the foreign exchange market," *Review of Financial Studies* 4, 513-541.

Canova, Fabio, and Jane Marrinan, 1995, "Predicting excess returns in financial markets," *European Economic Review* 39, 35-69.

Constantinides, George, 1992, "A theory of the nominal term structure of interest rates," *Review of Financial Studies* 5, 531-552.

Cox, John, Jonathan Ingersoll, and Stephen Ross, 1985, "A theory of the term structure of interest rates," *Econometrica* 53, 385-407.

Cumby, Robert, 1988, "Is it risk? Explaining deviations from uncovered interest parity," *Journal of Monetary Economics* 22, 279-299.

Domowitz, Ian, and Craig Hakkio, 1985, "Conditional variance and the risk premium in the foreign exchange market," *Journal of International Economics* 19, 47-66.

Duffie, Darrell, 1992, *Dynamic Asset Pricing Theory*, Princeton, NJ: Princeton University Press.

Duffie, Darrell, and Rui Kan, 1996, "A yield-factor model of interest rates," *Mathematical Finance* 6, 379-406.

Duffie, Darrell, and Kenneth Singleton, 1997, "An econometric model of the term structure of interest rate swap yields," *Journal of Finance* 52, 1287-1321.

Engel, Charles, 1996, "The forward discount anomaly and the risk premium: A survey of recent evidence," *Journal of Empirical Finance* ??, 123-191.

Engel, Charles, and James Hamilton, 1990, "Long swings in the dollar: Are they in the data and do markets know it?" *American Economic Review* 80, 689-713.

Fama, Eugene, 1984, "Forward and spot exchange rates," *Journal of Monetary Economics* 14, 319-338.

Flood, Robert, and Andrew Rose, 1996, "Fixes: Of the forward discount puzzle," *Review of Economics and Statistics*.

Frachot, Antoine, 1996, "A reexamination of the uncovered interest rate parity hypothesis," *Journal of International Money and Finance* 15, 419-437.

Frankel, Jeffrey, and Charles Engel, 1984, "Do asset demand functions optimize over the mean and variance of real returns? A six currency test," *Journal of International Economics 17*, 309-323.

Hansen, Lars, and Robert Hodrick, 1983, "Risk averse speculation in the forward foreign exchange market," in J.A. Frenkel, ed., *Exchange Rates and International Macroeconomics*, Chicago: University of Chicago Press, 113-142.

Hansen, Lars, and Ravi Jagannathan, 1991, "Implications of security market data for models of dynamic economies," *Journal of Political Economy 99*, 225-262.

Hodrick, Robert, 1987, *The Empirical Evidence on the Efficiency of Forward and Futures Foreign Exchange Markets*, New York: Harwood Academic Publishers.

Hollifield, Burton, and Raman Uppal, 1997, "An examination of uncovered interest rate parity in segmented international commodity markets," *Journal of Finance* 52, 2145-2170.

Mark, Nelson, 1988, "Time varying betas and risk premia in the pricing of forward foreign exchange contracts," *Journal of Financial Economics 22*, 335-354.

Nielsen, Lars, and Jesús Saá-Requejo, 1993, "Exchange rate and term structure dynamics and the pricing of derivative securities," unpublished manuscript, INSEAD, May.

Pearson, Neil, and Tong-Sheng Sun, 1994, "Exploiting the conditional density in estimating the term structure: An application to the Cox, Ingersoll, and Ross model," *Journal of Finance 54*, 1279-1304.

Saá-Requejo, Jesús, 1994, "The dynamics and the term structure of risk premia in foreign exchange markets," unpublished manuscript, INSEAD, May.

Stuart, Alan, and J.K. Ord, 1987, *Kendall's Advanced Theory of Statistics, Volume 1*, New York: Oxford University Press.

Stulz, René, 1987, "An equilibrium model of exchange rate determination and asset pricing with nontraded goods and imperfect competition," *Journal of Political Economy 95*, 1024-1040.

Sun, Tong-sheng, 1992, "Real and nominal interest rates: A discrete-time model and its continuous-time limit," *Review of Financial Studies 5*, 581-611.

Yaron, Amir, 1995, "Liquidity shocks and international asset pricing: A theoretical and empirical investigation," unpublished manuscript, Carnegie Mellon University, October.

Table 1

Properties of Currency Prices and Interest Rates

Currency	Mean	Std Deviation	Autocorrelation
A. Depreciation Rate, $s_{t+1} - s_t$			
British Pound	−0.0017	0.0342*	0.084
Canadian Dollar	−0.0015	0.0122*	0.057
French Franc	−0.0005	0.0328*	−0.002
German Mark	0.0021	0.0340*	−0.015
Italian Lira	−0.0038	0.0334*	0.049
Japanese Yen	0.0044	0.0324*	0.067
B. One-Month Interest Rate, r_t			
American Dollar	0.0069*	0.0030*	0.957*
British Pound	0.0093*	0.0027*	0.915*
Canadian Dollar	0.0081*	0.0028*	0.965*
French Franc	0.0091*	0.0035*	0.755*
German Mark	0.0053*	0.0020*	0.969*
Italian Lira	0.0122*	0.0045*	0.743*
Japanese Yen	0.0046*	0.0020*	0.914*
C. Forward Premium, $f_t - s_t = r_t - r_t^*$			
British Pound	−0.0024*	0.0027*	0.900*
Canadian Dollar	−0.0014*	0.0014*	0.842*
French Franc	−0.0023*	0.0032*	0.660*
German Mark	0.0017*	0.0029*	0.953*
Italian Lira	−0.0056*	0.0045*	0.724*
Japanese Yen	0.0021*	0.0029*	0.888*

Entries are sample moments of depreciation rates, $s_{t+1} - s_t$, one-month eurocurrency interest rates, r_t, and forward premiums, $f_t - s_t$. The data are monthly, last Friday of the month, from the Harris Bank's *Weekly Review: International Money Markets*

and Foreign Exchange, compiled by Richard Levich at New York University's Stern School of Business. The data are available by anonymous ftp: aleast.gsia.cmu.edu in directory /dist/fx. Dates t run from July 1974 to November 1994 (245 observations). An asterisk (*) indicates a sample moment at least twice its Newey-West standard error. The letters s and f denote logarithms of spot and one-month forward exchange rates, respectively, measured in dollars per unit of foreign currency, and r denotes the continuously-compounded one-month yield. Mean is the sample mean, St Dev the sample standard deviation, and Autocorr the first autocorrelation.

Table 2

Forward Premium Regressions

Currency	a_1	a_2	Std Er	R^2
British Pound	-0.0062 (0.0027)	-1.840 (0.847)	0.0339	0.0213
Canadian Dollar	-0.0036 (0.0009)	-1.575 (0.460)	0.0120	0.0341
French Franc	-0.0021 (0.0031)	-0.674 (0.827)	0.0328	0.0042
German Mark	0.0033 (0.0025)	-0.743 (0.805)	0.0340	0.0041
Italian Lira	-0.0042 (0.0039)	-0.073 (0.453)	0.0335	0.0001
Japanese Yen	0.0080 (0.0024)	-1.711 (0.643)	0.0320	0.0230

Entries are statistics from regressions of the depreciation rate, $s_{t+1} - s_t$, on the forward premium, $f_t - s_t$:

$$s_{t+1} - s_t = a_1 + a_2(f_t - s_t) + \text{residual},$$

where s and f are logarithms of spot and forward exchange rates, respectively, measured as dollars per unit of foreign currency. The data are described in the notes to Table 1. Dates t run from July 1974 to November 1994 (245 observations). Numbers in parentheses are Newey-West standard errors (3 lags) and Std Er is the estimated standard deviation of the residual.
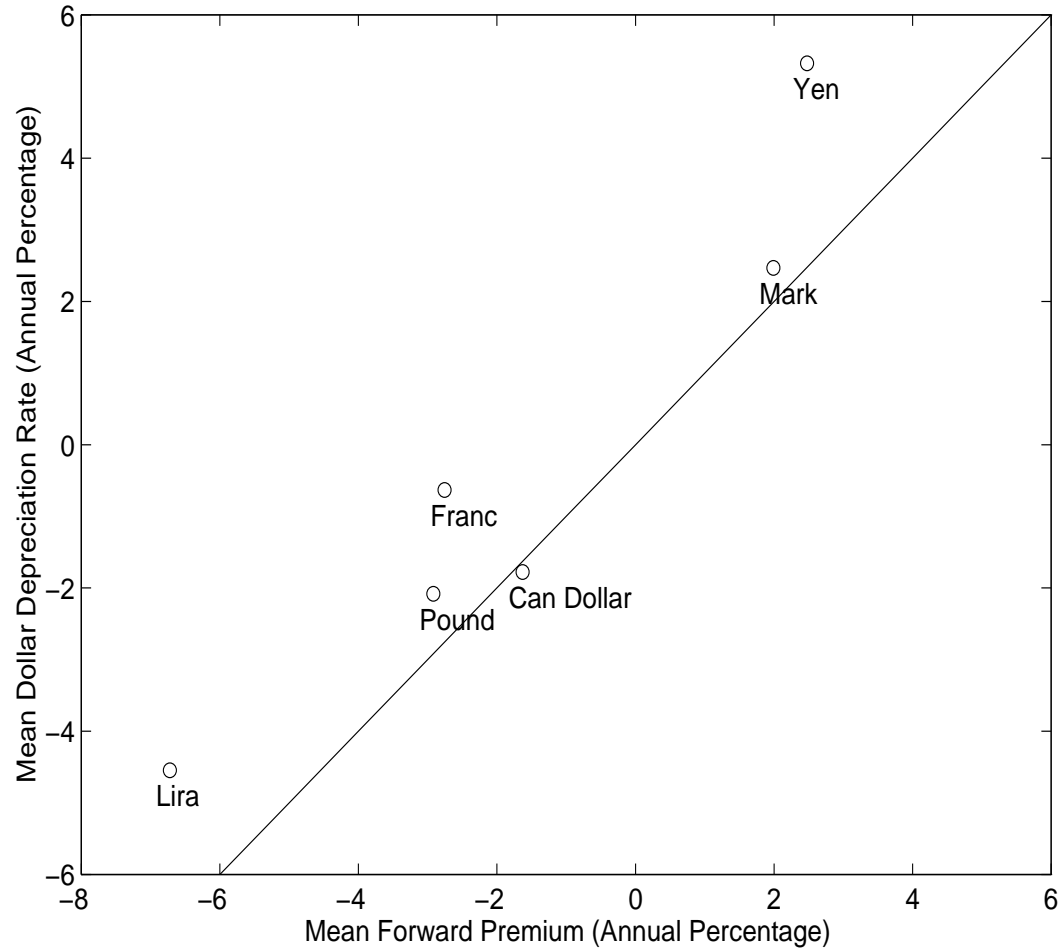
Table 3

Estimates of Examples 2 and 4

Parameter	Estimate	Standard Error
A. Example 2: Independent Factor Model		
θ_0	0.007	6.567×10^{-4}
σ_0	0.003	0.009
φ_0	0.992	0.041
θ_1	1.004×10^{-4}	3.614×10^{-5}
σ_1	0.081	0.022
φ_1	0.919	0.032
$|\lambda_1|$	2.383	0.384
B. Example 4: Interdependent Factor Model		
θ_1	0.005	4.950×10^{-4}
σ_1	0.017	0.004
φ_1	0.919	0.037
γ_1	0.331	0.080
λ_1	-5.886	2.184
λ_2	-5.541	2.202

Entries are exactly identified GMM estimates of the parameters of examples 2 and 4 in the text based on data for the dollar-pound rate. Example 2 is estimated to reproduce these moments: the mean, variance, and autocorrelation of the dollar short rate; the variance and autocorrelation of the dollar-pound forward premium; the variance of the dollar-pound depreciation rate; and the regression slope a_2 of equation (1). Example 4 is similar. We drop the condition based on the autocorrelation of the the short rate, since the (symmetric) model cannot reproduce different values for the autocorrelation of the short rate and the forward premium. Standard errors are computed by the Newey-West method (12 lags).

Figure 1

Mean Depreciation Rates and Forward Premiums

The figure plots mean depreciation rates of the dollar against mean forward premiums. The numbers are taken from Table 1 and multiplied by 1200 to convert them to annual percentages. Circles represent values for different currencies. The solid line represents equal values for mean depreciation rates and forward premiums. (a "45-degree line").