NEW YORK UNIVERSITY
STERN SCHOOL OF BUSINESS
FINANCE DEPARTMENT

Working Paper Series, 1996

Affine Models of Currency Pricing

Backus, David, Silverio Foresi and Chris Telmer

FIN-96-9

NBER Working Paper No 5623

Affine Models of Currency Pricing*

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

April 16, 1996

Abstract

Perhaps the most puzzling feature of currency prices is the tendency for high interest rate currencies to appreciate when the expectations hypothesis suggests the reverse. 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 a general arbitrage-free setting 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 common requirement that interest rates be strictly positive.

JEL Classification Codes: F31 G12 G15.

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

*A predecessor of this paper circulated as "The forward premium anomaly: Three examples in search of a solution," presented at the NBER Summer Institute, July 1994. We thank Ravi Bansal, Geert Bekaert, Wayne Ferson, Burton Hollifield, Andrew Karolyi, Amir Yaron, and Stanley Zin for helpful comments and suggestions, as well as seminar participants at Carnegie Mellon University, Columbia University, Stanford University, the Universities of California at Berkeley and San Diego, the University of Illinois, the University of Southern California, the American Finance Association, the Utah Winter Finance Conference, and the Western Finance Association. Backus thanks the National Science Foundation for financial support.

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

‡Stern School of Business, New York University; sforesi@stern.nyu.edu.

§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 the expectations hypothesis suggests the opposite: that investors will 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 must (i) be negatively correlated with the expected rate of depreciation and (ii) have a 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, 1991; Mark, 1991), from expected utility (Bekaert, Hodrick, and Marshall, 1992), and from frictionless trade in goods (Hollifield and Uppal, 1995). We address more recent attempts to account for the anomaly with affine models later in the paper.

We study the anomaly in both a general theoretical framework and in the more limited class of affine models. We express our general framework in terms of *pricing kernels*: stochastic processes governing prices of state-contingent claims. In this framework, we relate the spot exchange rate to pricing kernels in the two currencies and describe the properties the kernels must have to account for the puzzling behavior of forward and spot exchange rates. The anomaly implies, in general, an inverse relation between differences in conditional means and differences in conditional higher moments of pricing kernels.

We then turn to specific models that might account for the anomaly and other features of currency prices and interest rates. Natural candidates are affine models, in which conditional means and variances of logarithms of pricing kernels are linear functions of state variables. Affine models have been widely used in pricing fixed income securities, their linear structure makes them relatively transparent. Indeed, affine models of currency pricing have become increasingly popular in recent years, with notable applications by Ahn (1995), Amin and Jarrow (1991), Bakshi and Chen (1995), Bansal (1995), Brandt (1994), Nielsen and Saá-Requejo (1993), and Saá-Requejo (1994). We describe, for the Duffie-Kan (1993) class of affine models, the conditions needed to reconcile the anomaly with strictly positive interest rates.

We begin with a summary of the properties of dollar exchange rates and one-month eurocurrency interest rates, which serves as an anchor to the theoretical modeling that follows.

2 Properties of Currency Prices and Interest Rates

The properties of exchange rates and eurocurrency interest rates have been widely documented, but a review serves to focus our attention on the issues to be addressed and provides a quantitative benchmark for theoretical developments. 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. There is also weak evidence that depreciation rates exhibit greater kurtosis than one would find with the normal distribution, but none of our estimated measures of skewness or kurtosis exceeds twice its estimated standard error. Earlier work — Kritzman (1994), for example — suggests that kurtosis may be more apparent over shorter time intervals. There is little evidence of autocorrelation in depreciation rates for any of the six currencies we examine. 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. As with currency prices, none of the measures of skewness and kurtosis are more than twice as large as their standard errors.

One way to think about this evidence is to relate it to the expectations hypothesis: that forward rates are expected future spot rates. We express this 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 (1995): Hodrick (1987) provide exhaustive references to the literature. The most common evidence comes from regressions of the form

$$s_{t+1} - s_t = \alpha_1 + \alpha_2(f_t - s_t) + \text{residual}. \tag{1}$$

The expectations hypothesis implies a regression slope $\alpha_2 = 1$: yet most studies estimate α_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 α_2 range from -0.073 for the lira to -1.840 for the pound. All 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. One might invest: for example: in the currency with the higher interest rate. Bekaert and Hodrick (1992) show that while such strategies are not risklessr: 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 stability. Although estimates of α_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 second issue concerns measurement error and bid/ask spreads. Bossaerts and Hillion (1991) and Bekaert and Hodrick (1993) argue: however: that neither of these factors has a material effect on the sign or magnitude of estimates of α_2. A third issue concerns the exchange-rate regime. Flood and Rose (1994) 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 α_2 are close to one for the German mark and the 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 estimate: as others do: negative values for α_2.

The anomaly has motivated a large and growing number of studies suggesting explanations. Foremost among these is Fama (1984): who labels the difference between the forward

rate and the expected future spot rate a risk premium and proceeds to document its properties. 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 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

$$\alpha_2 = \frac{Cov(q, p+q)}{Var(p+q)} = \frac{Cov(q,p) + Var(q)}{Var(p+q)}, \tag{3}$$

it is clear that a constant risk premium p generates $\alpha_2 = 1$. To generate negative values of α_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.

In summary, Fama interprets the evidence as suggesting a highly variable risk premium that reverses the sign of the slope parameter α_2 in the forward premium regression relative to what it would be under the expectations hypothesis. We refer to this feature of the data as an anomaly because of the large number of unsuccessful attempts to account for it with risk-based theories. In this sense, the term "risk premium" is more a convenient label than an explanation.

3 A Theoretical Framework

The challenge of currency pricing 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.

Before turning to specific models, we find it useful to consider currency prices in a fairly general theoretical setting. We characterize asset prices with a *pricing kernel*: a stochastic process governing prices of state-contingent claims. Existence of such a process (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 adapt this approach to the pricing of currencies, relate the volatility of currency prices to the variability of pricing kernels in two currencies, and examine the relation between pricing kernels and the forward premium anomaly.

3.1 Pricing Kernels

We begin with assets denominated in domestic currency ("dollars"), then move on to those denominated in foreign currency ("pounds"). With respect to dollar assets, consider the dollar value v_t of a claim to the stochastic cash flow of d_{t+1} dollars one period later. The price v and cash flow d satisfy the pricing relation,

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

or

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

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. Denote by b_t^n the price of an n-period zero-coupon bond: the claim to one dollar at date $t+n$ in all states. Since the one-period return on an $(n+1)$-period bond is b_{t+1}^n / b_t^{n+1}, we can compute bond prices recursively from

$$b_t^{n+1} = E_t\left(m_{t+1} b_{t+1}^n\right), \tag{6}$$

starting with $b_t^0 = 1$ (a dollar today costs a dollar). The price of a one-period bond, for example, is $b_t^1 = E_t m_{t+1}$. Continuously-compounded bond yields y are related to prices by $b_t^n \equiv \exp\left(-y_t^n n\right)$. The short rate r_t is the yield y_t^1 on a one-period bond:

$$r_t = -\log b_t^1 = -\log E_t m_{t+1}. \tag{7}$$

We return to this equation when we examine exchange rates.

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

$$1 = E_t\left(m_{t+1}^* R_{t+1}^*\right). \tag{8}$$

If we use the second approach, with $S = \exp(s)$ denoting the dollar spot price of one pound, then

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

If the pound asset and currencies are both traded, there are obvious arbitrage opportunities unless the return satisfies 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, 8) are not unique, but we will see that we can choose them to satisfy the same equation.

We summarize the connection between pricing kernels and currency prices in

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 we can choose the pricing kernels m and m^* for dollars and pounds to satisfy both*

$$m_{t+1}^*/m_{t+1} = S_{t+1}/S_t \tag{9}$$

and the pricing relations (5,8).

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 extensions). For any such m, the choice $m_{t+1}^* = m_{t+1}S_{t+1}/S_t$ automatically satisfies (8). ∎

The intuition is straightforward: if we know prices of state-contingent claims in dollars and pounds, we can compute the implied exchange rate from their ratio. The only ambiguity stems from combinations of state-contingent claims that are not traded.

The proposition tells us that of the three random variables — $m_{t+1}\Gamma m^*_{t+1}\Gamma$ andS_{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 kernelsΓ which highlights the essential symmetry of the theory between the two currencies.

One implication of this symmetric perspective is that pricing kernels appear to be highly correlated. To see thisΓ note that equation (9) implies

$$Var(s_{t+1} - s_t) = Var(\log m^*_{t+1}) + Var(\log m_{t+1}) - 2Cov(\log m^*_{t+1}, \log m_{t+1}). \qquad (10)$$

Estimates of $Var(s_{t+1} - s_t)$ are in the neighborhood of 0.03^2 for most of the exchange rates in Table 1Γ smaller for the Canadian dollar. Estimates of $Var(\log m)$ are typically larger: Backus and Zin (1994Γ Section 6) suggest that 015^2 is a conservative estimate for monthly dollar returns. Estimates from closely related Hansen-Jagannathan (1991) bounds are similar. If pricing kernels for other currencies exhibit comparable variabilityΓ then equation (10) implies that the correlation between the logarithms of the two kernels is 0.98. Larger estimates of $Var(\log m)$ and $Var(\log m^*)$ and smaller estimates of $Var(s_{t+1} - s_t)$ imply larger correlations. The strong correlation between pricing kernels is not an indication of international capital mobility — capital mobility wasΓ in factΓ a premise of Proposition 1. We find it striking neverthelessΓ since it implies that state prices are more highly correlated across currencies than returns. Roughly speakingΓ the two pricing kernels appear to be more highly correlated than their conditional means.

3.2 Forward Rates and Risk Premiums

Given pricing kernels for two currencies and equation (9) for spot exchange ratesΓ we can derive the forward premium and its components from the pricing relation (4). Consider a forward contract specifying at date t the exchange at $t + 1$ of one pound and $F_t = \exp(f_t)$ dollarsΓ 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 $t\Gamma$ the pricing relation implies

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

If we divide by S_t and apply Proposition 1Γ we find

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

Thus the forward premium is

$$f_t - s_t = \log E_t m^*_{t+1} - \log E_t m_{t+1}. \qquad (11)$$

This equation and the definition of the short rate equation (7) give us

$$f_t - s_t = r_t - r_t^*, \tag{12}$$

the familiar covered interest rate parity condition.

Now consider the components of the forward premium. The expected rate of depreciation is from (9)

$$q_t \equiv E_t s_{t+1} - s_t = E_t \log m_{t+1}^* - E_t \log m_{t+1}. \tag{13}$$

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 (2.11)

$$p_t = (\log E_t m_{t+1}^* - E_t \log m_{t+1}^*) - (\log E_t m_{t+1} - E_t \log m_{t+1}), \tag{14}$$

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 later 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

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

and the risk premium is

$$p_t = (\mu_{2t}^* - \mu_{2t})/2. \tag{15}$$

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 depends on higher moments. For an arbitrary distribution equation (13) tells us (again) that only the means affect the expected rate of depreciation. The risk premium is given in general by (14) but if all of the conditional moments of $\log m$ exist $\log E_t m_{t+1}$ can be expanded

$$\log E_t m_{t+1} = \sum_{j=1}^{\infty} \kappa_{jt}/j!, \tag{16}$$

where κ_{jt} is the jth cumulant for the conditional distribution of $\log m_{t+1}$. Equation (16) 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, ch. 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 (16) 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

$$p_t = \kappa^*_{-1,t} - \kappa_{-1,t}, \tag{17}$$

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 moments."

With equations (17) and (13) 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 $f_t - s_t = \mu^*_{1t} - \mu_{1t} + \kappa^*_{-1,t} - \kappa_{-1,t}$.*

Our 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 (15)] and possibly higher moments [equation (17)], which need bear no specific relation to the conditional variance of the depreciation rate. GARCH-M models, to put it simply, focus on the wrong conditional variance.

4 Affine Models with Independent Factors

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 1Γ and one-half the difference in conditional variances is higher in regime 2Γ and varies more than the difference in meansΓ 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. Affine models have a number of clear advantages. FirstΓ conditional means and variances of logarithms of pricing kernels are linear functions of a vector of state variables. SecondΓ we haveΓ as a professionΓ more than a decade's experience with these models in pricing fixed income securities; much of this experience can be transferred directly to currency pricing. FinallyΓ we will see 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.

In this section we consider specific examples of affine models motivated by related work. In the next section we consider the general class of affine currency models.

4.1 A Cox, Ingersoll, and Ross Model for Two Currencies

An obvious starting point is a two-currency version of CoxΓ IngersollΓ and Ross (1985) like Bakshi and Chen (1995). Our version is adapted from Sun's (1992) discrete-time translation.

In discrete timeΓ the Cox-Ingersoll-Ross model can be expressed in two equationsΓ one specifying a "square-root" process for a state variableΓ the other relating the pricing kernel to the state. Let us say that the state variable z follows

$$z_{t+1} = (1-\varphi)\theta + \varphi z_t + \sigma z_t^{1/2}\varepsilon_{t+1}, \tag{18}$$

with $0 < \varphi < 1\Gamma\, \theta > 0\Gamma$ and $\{\varepsilon_t\} \sim \text{NID}(0,1)$. The unconditional mean of z is $\theta\Gamma$ the autocorrelation is $\varphi\Gamma$ the conditional variance is $\sigma^2 z_t\Gamma$ and the unconditional variance is $\sigma^2\theta/(1-\varphi^2)$. With the substitution $\kappa = 1-\varphi\Gamma$ we can write (18) as

$$z_{t+1} - z_t = \kappa(\theta - z_t) + \sigma z_t^{1/2}\varepsilon_{t+1},$$

a direct analog of the continuous-time original. The critical ingredient of (18) is the square-root term in the innovation, whose conditional variance falls to zero as z approaches zero. In continuous time, this feature and the *Feller condition*,

$$(1-\varphi)\theta = \kappa\theta \geq \sigma^2/2, \tag{19}$$

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

The pricing kernel for the discrete-time Cox-Ingersoll-Ross model can be expressed

$$-\log m_{t+1} = (1+\lambda^2/2)z_t + \lambda z_t^{1/2}\varepsilon_{t+1}. \tag{20}$$

The coefficient of z is a normalization, chosen to make z the one-period rate of interest; see Appendix A.1. The parameter λ 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. Note that equation (20) builds in an inverse relation between the conditional mean and variance of the logarithm of the pricing kernel.

This structure is an example of the conditionally log-normal pricing kernels described in Section 3. The conditional mean and variance,

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

are both linear in the state variable z. The short-term rate of interest is

$$r_t = -\log E_t m_{t+1} = -\left(E_t \log m_{t+1} + \frac{1}{2} Var_t \log m_{t+1}\right) = z_t, \tag{21}$$

as claimed earlier.

If m is the dollar pricing kernel, we complete the model by considering a second pricing kernel, m^*, for pounds. If the pound pricing kernel is based on an analogous state variable z^* following an identical but independent process, then the pound short rate is $r_t^* = z_t^*$. The forward premium is

$$f_t - s_t = z_t - z_t^*,$$

with expected depreciation $q_t = (1+\lambda^2/2)(z_t - z_t^*)$ and risk premium $p_t = -\lambda^2/2(z_t - z_t^*)$. Thus the linearity of the conditional mean and variance translate into forward premium components that are linear functions of the differential $z - z^*$. More important, this structure automatically generates the negative correlation between p and q of Fama's condition (i): since equation (20) 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. Bakshi and Chen (1995, eq 47) make a similar observation.

This model cannot, however, reproduce the anomaly. If we regress the depreciation rate on the forward premium in this model, the slope is

$$\alpha_2 = 1 + \lambda^2/2.$$

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 (1994) evidence for the ERM.

There is a simple solution to this problem, but it has a cost. Suppose we replace (20) with

$$-\log m_{t+1} = (-1 + \lambda^2/2)z_t + \lambda z_t^{1/2}\varepsilon_{t+1}, \tag{22}$$

so that the coefficient of z contains -1 rather than $+1$. Short-term interest rates are then $r_t = -z_t$ and $r_t^* = -z_t^*$ and the forward premium is $-(z_t - z_t^*)$. Expected depreciation is $(-1 + \lambda^2/2)(z_t - z_t^*)$. The regression slope is therefore

$$\alpha_2 = 1 - \lambda^2/2,$$

which is always less than one, and negative for large enough values of λ. We have lost, however, the trademark positive interest rates of the Cox-Ingersoll-Ross model.

4.2 Models with Independent Factors

The two-currency model that accounts for the anomaly has two properties that clearly differ from the evidence: interest rates are uncorrelated across currencies and negative with probability one. We consider a generalization in an attempt to resolve, or at least mitigate, both problems.

Our generalization extends the model in two directions: we replace the independent univariate Cox-Ingersoll-Ross models for each currency with independent general affine models, and we introduce a common state variable, independent of currency prices, that affects interest rates in both currencies. The latter allows us to reproduce the positive correlation of interest rates across currencies. The former offers the potential to reduce or eliminate the possibility of negative interest rates. An additional benefit of this class of models is that they use factors parsimoniously. To model the dollar/pound rate, for example, we need only the dollar, pound, and common factors: factors for other currencies are irrelevant. And since the exchange rate between two currencies depends only on the factors governing the two currencies, these models are easily extended to additional currencies by introducing additional currency-specific factors. Perhaps for this reason, Bakshi and Chen (1995) and Bansal (1995) examine models in this class.

Consider then a two-currency world based on three independent vectors of state variables: a common state variable z_0 and currency-specific state variables z_1 and z_2, say. (In an effort to streamline the notation we have replaced z and z^* with z_1 and z_2.) Our class of independent factor models consists of laws of motion

$$z_{it+1} = (I - \Phi_i)\theta_i + \Phi_i z_{it} + V^i(z_{it})^{1/2}\varepsilon_{it+1} \tag{23}$$

for each state variable i and pricing kernels

$$\begin{aligned} -\log m_{t+1} &= \delta + \gamma_0^\top z_{0t} + \gamma_1^\top z_{1t} + \lambda_0^\top V^0(z_{0t})^{1/2}\varepsilon_{0t+1} + \lambda_1^\top V^1(z_{1t})^{1/2}\varepsilon_{1t+1} \\ -\log m^*_{t+1} &= \delta^* + \gamma_0^\top z_{0t} + \gamma_2^\top z_{2t} + \lambda_0^\top V^0(z_{0t})^{1/2}\varepsilon_{0t+1} + \lambda_2^\top V^2(z_{2t})^{1/2}\varepsilon_{2t+1}, \end{aligned} \tag{24}$$

with $\{\varepsilon_{it}\}$ independent standard normal random variables. The autoregressive matrices Φ_i have stable roots with real parts between zero and one and positive diagonal elements. The volatility matrices V^i have typical elements

$$v^i_j(z) = \alpha^i_j + \beta^{i\top}_j z_i.$$

We define *admissible* values of the state variables as those for which volatility functions are nonnegative. Readers will recognize this model as an example of the affine class characterized by Duffie and Kan (1993), who report sufficient conditions for keeping state variables in the admissible set (Condition A described in Appendix A.2 below).

This model builds a lot of structure into asset prices. 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 α_2 that characterizes the forward premium anomaly. We find, as result of this structure, that these models retain one of the weaknesses of the two-currency Cox-Ingersoll-Ross model:

Proposition 2 *Consider the Duffie-Kan class of affine models with independent currency factors summarized by equations (23,24). 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 A.3. The intuition is similar to the two-currency Cox-Ingersoll-Ross model. The affine models permitted in Proposition 2 are based on state variables that are unbounded in one direction. In our version of Cox-Ingersoll-Ross, for example, the state variable z assumes all positive values with positive probability. This state variable has 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.

Proposition 2 indicates that we cannot use this class of models to reconcile the anomaly with strictly positive interest rates but does not in our view rule them out altogether. If a small probability of negative interest rates led to an affine model that was realistic in other respects we might view it as a small cost paid for the convenience of linearity. Duffie and Singleton (1995) and Pearson and Sun (1994) make a similar argument in extending the Cox-Ingersoll-Ross model of bond pricing. We examine this possibility in the next subsection.

4.3 Informal Estimation of an Independent Factor Model

We estimate the parameters of the simplest independent factor model to see how this structure might work in practice. We find for this example that Proposition 2 understates the difficulties. Although estimated parameters imply a small probability that short rates are negative the model differs sharply from the data along other dimensions.

Our model starts with three scalar state variables following square root processes

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

for $i = 0, 1, 2$ and pricing kernels

$$\begin{aligned} -\log m_{t+1} &= \delta + \gamma_0 z_{0t} + \gamma_1 z_{1t} + \lambda_0 z_{0t}^{1/2}\varepsilon_{0t+1} + \lambda_1 z_{1t}^{1/2}\varepsilon_{1t+1} \\ -\log m^*_{t+1} &= \delta^* + \gamma_0 z_{0t} + \gamma_2 z_{2t} + \lambda_0 z_{0t}^{1/2}\varepsilon_{0t+1} + \lambda_2 z_{2t}^{1/2}\varepsilon_{2t+1}, \end{aligned}$$

where $\{\varepsilon_{it}\}$ is independent standard normal. We refer to z_0 as the common factor and z_1 and z_2 as respectively the dollar and pound factors. We assume as well that the parameters related to z_1 and z_2 are the same: $\varphi_1 = \varphi_2 = \varphi$ $\theta_1 = \theta_2 = \theta$ $\sigma_1 = \sigma_2 = \sigma$ $\delta^* = \delta$ $\gamma_1 = \gamma_2 = \gamma$ and $\lambda_1 = \lambda_2 = \lambda$. This presumption of symmetry is pure convenience: it reduces the number of parameters and in the process clarifies the relation between parameter values and features of the data. We use the normalizations

$$\begin{aligned} \gamma_0 &= 1 + \lambda_0^2/2 \\ \gamma &= -1 + \lambda^2/2. \end{aligned}$$

As in the two-currency Cox-Ingersoll-Ross model the latter allows the model to reproduce the anomaly.

Given parameter valuesΓ we can compute properties of currency prices and interest rates in dollars and pounds. Here we do the reverse: we use sample moments for the dollar/pound exchange rate and short rates in dollars and pounds to estimate the model's parameters. We do this informally to illustrate as simply as possible the issues that arise in formal estimation of such models. The sample moments are those reported in Tables 1 and 2. The model's structure allows us to estimate its parameters recursively:

1. The regression slope underlying the anomaly identifies λ. The depreciation rate is

$$s_{t+1} - s_t = (-1 + \lambda^2/2)(z_{1t} - z_{2t}) + \lambda\left(z_{1t}^{1/2}\varepsilon_{1t+1} - z_{2t}^{1/2}\varepsilon_{2t+1}\right).$$

   Short rates are

$$\begin{aligned} r_t &= \delta + z_{0t} - z_{1t} \\ r_t^* &= \delta + z_{0t} - z_{2t}, \end{aligned}$$

   so the forward premium is $f_t - s_t = -(z_{1t} - z_{2t})$. The slope parameter in the forward premium regression is

$$\alpha_2 = 1 - \lambda^2/2.$$

   Since the estimated slope parameter in Table 2 is -1.840Γ we estimate $|\lambda| = 2.38$ (the sign is not identified).

2. We use properties of the forward premium to estimate the parameters governing the currency factors z_1 and z_2Γ which we have assumed are the same. The autocorrelation parameter φ is the autocorrelation of the forward premiumΓ in this case 0900.

3. The variance of z_1 (equivalently z_2) is related to the variance of the forward premium by

$$Var(f - s) = Var(z_1 - z_2) = 2\,Var(z_1).$$

   Since $Var(f - s) = 0.0027^2$ for dollar/pound ratesΓ we have $Var(z_1) = 0.0027^2/2 = 1.82 \times 10^{-6}$. Given $Var(z_1)$Γ we find θ from the variance of the depreciation rate:

$$Var\,(s_{t+1} - s_t) = 2\alpha_2^2\,Var(z_1) + 2\lambda^2\theta. \tag{25}$$

   Our estimates of λ and $Var(z_1)$ imply $\theta = 5.46 \times 10^{-6}$.

4. We compute σ from our estimates of $Var(z_1)$ΓθΓ andφ:

$$Var(z_1) = \frac{\theta\sigma^2}{1 - \varphi^2},$$

   which implies $\sigma = 0.251$.

5. We identify the parameters of the common factor from properties of short-term interest rates. The mean determines θ_0:

$$E(r) = 0.0069 = \theta_0 + \theta,$$

implying $\theta_0 = 6.91 \times 10^{-3}$.

6. The other two parameters of the common factor φ_0 and σ_0 are intertwined. The former is an input into the autocorrelation of z_0:

$$Auto(r) = \varphi_0 \frac{Var(z_0)}{Var(z_0) + Var(z_1)} + \varphi \frac{Var(z_1)}{Var(z_0) + Var(z_1)}.$$

We find $Var(z_0)$ from the variance of the short rate: $Var(r) = Var(z_0) + Var(z_1)$. Given our earlier estimate of $Var(z_1)$ we compute $Var(z_0) = 2.68 \times 10^{-6}$ and $\varphi_0 = 0.996$. Roughly speaking, greater persistence in interest rates than interest differentials implies $\varphi_0 > \varphi$.

7. We use our estimate of $Var(z_0)$ to determine σ_0:

$$Var(z_0) = 2.68 \times 10^{-6} = \frac{\theta_0 \sigma_0^2}{1 - \varphi_0^2},$$

which implies $\sigma_0 = 1.80 \times 10^{-3}$.

We can now provide a quantitative assessment of Proposition 2. Since z_1 and z_2 range between zero and infinity, we see that short rates are negative with positive probability, as required by the theory. Since both z_0 and z_1 have approximately gamma distributions (see Appendix A.4), the probability is easily computed. We find that the probability is less than 10^{-5}.

The difficulty with these parameter values is, instead, that the unconditional distribution has enormous skewness and kurtosis. One symptom of this is the Feller condition: our estimates of $\{\theta, \sigma, \varphi\}$ violate the condition by five orders of magnitude:

$$\frac{2(1-\varphi)\theta}{\sigma^2} = 1.73 \times 10^{-5} \not\geq 1. \tag{26}$$

Holding sample moments fixed, the Feller ratio declines as we increase α_2 toward one, but exceeds one for the dollar/pound rate only when $\alpha_2 > 0.98$. Even the Flood-Rose estimate of 0.58 for the ERM is too small to eliminate the problem. By one interpretation, parameters that violate the Feller condition are infeasible: the state variables are absorbed at zero and the model has a degenerate limiting distribution. Another interpretation is that zero is a reflecting barrier, and that violation of the Feller condition indicates extreme values for higher moments; see Appendix A.4.

This difficulty stems directly from the anomaly. As we see in equation (25)Γ the variance of the depreciation rate depends on both λ and θ. The anomaly dictates a large value of λ. For the model to reproduce the observed volatility of the depreciation rateΓ we therefore need a small value of θ. With small θ we find that large σ is required to reproduce the variability of the forward premiumΓ which violates the Feller condition.

Stated somewhat differentlyΓ the model is squeezed between the anomalyΓ on the one handΓ and the variability of the depreciation rateΓ on the other. The anomaly indicates a large value of $|\lambda|$. But the variance of the depreciation rate restricts the independent variation in the two kernelsΓ and therefore indicates a small value of $\lambda^2\theta$. A small value of θ is the compromiseΓ which leads to a violation of the Feller condition.

5 Affine Models with Interdependence

We turn next to the general class of affine currency models. We show for such a model to account for the anomalyΓ it must exhibit asymmetric interdependence: common factors must influence interest rates differently in the two currencies.

5.1 Two Examples

We illustrate the intuition for interdependence with two examples. One of the simplest examples is based on a single state variable z obeying a process like (18)Γ with pricing kernels

$$\begin{aligned} -\log m_{t+1} &= (1+\lambda^2/2)z_t + \lambda z_t^{1/2}\varepsilon_{t+1} \\ -\log m^*_{t+1} &= (\gamma^* + \lambda^{*2}/2)z_t + \lambda^* z_t^{1/2}\varepsilon_{t+1}. \end{aligned}$$

The model is interdependent in the sense that the same factor z affects both pricing kernels and asymmetric if its effects are different: if $(1,\lambda) \neq (\gamma^*, \lambda^*)$. In this setting short rates are $r_t = z_t$ and $r^*_t = \gamma^* z_t$Γ so the forward premium is $f_t - s_t = (1-\gamma^*)z_t$. Both interest rates are strictly positive if $\gamma^* > 0$. Depreciation is

$$s_{t+1} - s_t = \left[1 - \gamma^* + (\lambda^2 - \lambda^{*2})/2\right] z_t + (\lambda - \lambda^*) z_t^{1/2}\varepsilon_{t+1},$$

so expected depreciation is $q_t = [1-\gamma^*+(\lambda^2-\lambda^{*2})/2]z_t$. The slope of the forward premium regression is therefore

$$\alpha_2 = 1 + \frac{\lambda^2 - \lambda^{*2}}{2(1-\gamma^*)}. \tag{27}$$

If $\gamma^* = 1$ the forward premium is zero: the model has no forward premium and thus no forward premium anomaly. For other valuesΓ the model implies an inverse relation between

the forward premium and the interest differential if $2(1-\gamma^*)$ and $\lambda^2 - \lambda^{*2}$ have opposite signs and the latter is larger in absolute value. Frachot (1994) describes a similar example in continuous time.

This example is asymmetric in two respects. The first is that the two interest rates follow different processes. The second is that the state variable z has different effects on the two pricing kernels and hence on interest rates. A second example indicates that the latter which we refer to as asymmetric interdependence is the key to explaining the anomaly.

Consider then a similar model based on two state variables z_1 and z_2 obeying identical independent square root processes (18) with pricing kernels

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

Ahn (1995) 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 the state variables z_1 and z_2 potentially affect the two kernels in different ways.

Short rates in this model are

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

The forward premium

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

and depreciation rate

$$s_{t+1} - s_t = \left[1 - \gamma^* + (\lambda^2 - \lambda^{*2})/2\right](z_{1t} - z_{2t}) + (\lambda - \lambda^*)(z_{1t}^{1/2}\varepsilon_{1t+1} - z_{2t}^{1/2}\varepsilon_{2t+1}),$$

imply a regression slope of

$$\alpha_2 = 1 + \frac{\lambda^2 - \lambda^{*2}}{2(1-\gamma^*)},$$

as in equation (27). As in the first example appropriate choice of parameters allows us to generate a negative value. The critical feature in this regard is again that each state variable affects the two kernels differently.

These two examples highlight the role of asymmetric interdependence: of state variables that affect pricing kernels differentially. In the second this takes a particularly striking form. Suppose $\gamma^* < 1$ (if $\gamma^* > 1$ the argument is similar). From the short rate equations we might say then 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 α_2 less than one) implies $\lambda^{*2} > \lambda^2$, 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.

5.2 Informal Estimation of the Two-Factor Example

Despite this unusual feature, an informal estimation exercise suggests that this model provides a closer approximation than the independent factor model to the properties of currency prices and short-term interest rates. The model has six parameters and thus requires six sample moments to estimate:

1. We use the relative variance of the forward premium and the dollar short rate to identify γ^*:
$$\frac{Var(f-s)}{Var(r)} = \frac{2(1-\gamma^*)^2}{1+\gamma^{*2}}.$$
This equation has two solutions, but they are observationally equivalent. We choose the smaller root, $\gamma^* = 0.333$.

2. Given γ^*, the mean value of the dollar short rate determines θ:
$$E(r) = (1+\gamma^*)\theta.$$
With a mean short rate of 0.006904, we estimate $\theta = 0.00518$.

3. The autocorrelation of the forward premium identifies $\varphi = 0.900$.

4. We use the standard deviation of the dollar short rate to compute σ. The variance of the short rate is
$$Var(r) = (1+\gamma^{*2})\,Var(z_1),$$
implying $Var(z_1) = 0.00286^2$. Since
$$Var(z_1) = \frac{\sigma^2\theta}{1-\varphi^2},$$
we find $\sigma = 0.0173$. Unlike the independent factor model, the estimates of φ, θ, and σ satisfy the Feller condition.

5. We use properties of currency prices to estimate λ and λ^*. The difference in the squared values is determined by the anomaly: equation (27) implies
$$\lambda^2 - \lambda^{*2} = -3.79.$$

6. Finally, the variance of the depreciation rate identifies $(\lambda - \lambda^*)^2$:

$$Var(s_{t+1} - s_t) = 2(1-\gamma^*)^2\alpha_2^2\, Var(z_1) + 2(\lambda - \lambda^*)^2\theta.$$

Given values for $\alpha_2, Var(z_1)$, and $Var(s_{t+1} - s_t)$, we compute $(\lambda - \lambda^*)^2 = 0.111$.

On the whole, this example fits the data much better than the independent factor model. The Feller condition, for example, is satisfied. Outstanding issues are the magnitudes of λ and λ^*, which play an important role in the shape of the yield curve, and the unexpected asymmetry in the effects of each state variable on short rates and pricing kernels noted earlier.

5.3 Affine Models of Currency Pricing

The general affine currency model is characterized by a law of motion for a vector of state variables,

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

and pricing kernels

$$\begin{aligned} -\log m_{t+1} &= \delta + \gamma^\top z_t + \lambda^\top V(z_t)^{1/2}\varepsilon_{t+1} \\ -\log m^*_{t+1} &= \delta^* + \gamma^{*\top} z_t + \lambda^{*\top} V(z_t)^{1/2}\varepsilon_{t+1}, \end{aligned} \tag{29}$$

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

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

Equations (28–29) characterize what we term the general class of affine currency models. Duffie and Kan's (1993) Condition A guarantees that state variables remain in the region defined by nonnegative volatility in the continuous-time analog; see Appendix A.2.

With this structure, the depreciation rate is

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

Short rates are

$$\begin{aligned} r_t &= (\delta - \omega) + (\gamma - \tau)z_t \\ r^*_t &= (\delta^* - \omega^*) + (\gamma^* - \tau^*)z_t, \end{aligned}$$

where $\omega = \sum_j \lambda_j^2\alpha_j/2$, $\omega^* = \sum_j \lambda_j^{*2}\alpha_j/2$, $\tau = \sum_j \lambda_j^2\beta_j/2 \geq 0$, and $\tau^* = \sum_j \lambda_j^{*2}\beta_j/2 \geq 0$. For interest rates to be positive, we need each z_i bounded (say) below, and therefore $(\gamma - \tau)$ and $(\gamma^* - \tau^*)$ must have nonnegative elements.

Together these two equations imply a negative value of α_2 in the forward premium regression equation (1) if

$$Cov(s_{t+1} - s_t, f_t - s_t) = [(\gamma - \gamma^*) - (\tau - \tau^*)]^\top Var(z)(\gamma - \gamma^*).$$

We see immediately that the anomaly hinges on differences between γ and γ^* and between τ and τ^*. If $Var(z)$ is diagonal as in the examples of Section 5.1 then the ith elements of $\gamma - \gamma^* - \tau + \tau^*$ and $\gamma - \gamma^*$ must have different signs for at least one i. More generally the covariance hinges on differences in the effects of state variables on interest rates and pricing kernels in the two currencies. In this sense the asymmetric effect of the state variables on the two pricing kernels in our two examples is a general requirement of a model that accounts for the anomaly.

This model also makes it clear why models with independent factors in Section 4.2 cannot account for the anomaly with positive interest rates. This model is a special case in which z can be partitioned into independent subvectors: $z = (z_0, z_1, z_2)$. Then $\gamma = (\gamma_0, \gamma_1, 0)$ and $\gamma^* = (\gamma_0, 0, \gamma_2)$. These models can be asymmetric but the "exclusion restrictions" and the assumption that the common factor z_0 affects both kernels the same way limits their interdependence. Proposition 2 is the result.

6 Final Remarks

We have examined the implications for models of currency pricing of the forward premium anomaly: the tendency for currencies with high interest rates to rise subsequently in value. Many regard this feature of the data anomalous because of the many failed attempts to build theoretical models that account for it.

We find instead that it is relatively easy to construct models consistent with the anomaly: we need an inverse relation between the difference in the conditional means of the logarithms of pricing kernels in two currencies and differences in the conditional variances. In the class of affine models this requires either a positive probability of negative interest rates or that some state variables have asymmetric effects on the pricing kernels in the two currencies. Examples of each exist in the literature. Informal estimation suggests that within the class of affine models those with asymmetries offer the best hope of explaining the properties of currency prices and interest rates in general.

We are left with two outstanding issues. The first is whether a closer look finds that affine models with a small number of state variables are capable of approximating the properties of currency prices and fixed income securities in different currencies. Ahn (1995) 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 a stochastic processes that account for observed asset prices; but one might reasonably ask what kinds of behavior by policy makers and private agents might lead to such pricing kernels. One possibility is outlined by Alvarez and Atkeson (1996); Stulz (1987); and Ahron (1995); who develop dynamic general equilibrium models in which interest rates and currency prices reflect monetary policies. Perhaps further work will connect pricing kernels in these models to properties of interest rates; currency prices; and monetary aggregates.

A Mathematical Appendix

A.1 Representations of One-Factor Affine Models

We show that affine models capable of reproducing the anomaly can be characterized by two equations of the form:

$$z_{t+1} = (1-\varphi)\theta + \varphi z_t + \sigma z_t^{1/2}\varepsilon_{t+1} \tag{18}$$

and

$$-\log m_{t+1} = \delta \pm \gamma_0 z_t + \lambda z_t^{1/2}\varepsilon_{t+1} \tag{30}$$

for a fixed positive parameter γ_0 (a normalization). This model differs from Cox-Ingersoll-Ross in two respects: (i) the intercept δ in the relation for the pricing kernel and (ii) the possibility of negative coefficient of z in the same equation. We assume $\beta \neq 0$ since otherwise the conditional variance is constant and the model cannot account for the anomaly.

Consider as an alternative a more heavily parameterized model:

$$z_{t+1} = (1-\varphi)\theta + \varphi z_t + (\alpha + \beta z_t)^{1/2}\varepsilon_{t+1}$$

and

$$-\log m_{t+1} = \delta + \gamma z_t + \lambda(\alpha + \beta z_t)^{1/2}\varepsilon_{t+1},$$

with a Feller-like condition guaranteeing that $\alpha + \beta z$ is always positive in the continuous-time analog. This structure nests the one-factor models of Cox-Ingersoll-Ross (1985) and Pearson and Sun (1994) as special cases.

The goal is to show that the second model can be expressed in the form of the first equations (18–30). The key is that the state variable z is not observable: its role is simply to help characterize the conditional distribution of m_{t+j} for $j \geq 1$. In particular, linear transformations of z leave the distribution of future m's unchanged.

We proceed in steps. Consider first a reparametrization of the second model based on the substitution $z' = \alpha + \beta z$. In terms of this new state variable, the equations for the state and pricing kernel are

$$\begin{aligned} z'_{t+1} &= (1-\varphi)(\alpha + \beta\theta) + \varphi z'_t + \beta z_t'^{1/2}\varepsilon_{t+1} \\ &= (1-\varphi)\theta' + \varphi z'_t + \beta z_t'^{1/2}\varepsilon_{t+1} \end{aligned}$$

and

$$\begin{aligned} -\log m_{t+1} &= \left(\delta - \frac{\alpha}{\beta}\right) + \frac{\gamma}{\beta} z'_t + \lambda z_t'^{1/2}\varepsilon_{t+1} \\ &= \delta' + \gamma' z'_t + \lambda z_t'^{1/2}\varepsilon_{t+1}. \end{aligned}$$

The state equation is now in the same form as (18)Γ but we have an additional parameter γ' in the relation for the kernel. Step two is to eliminate the extra parameter with a normalizationΓ which we do by offsetting changes in γ' by rescaling z'. The only subtly is that the scaling must retain the sign of the state variable. If $\gamma' > 0$Γ define$z'' = (\gamma_0/\gamma')z'$. Then we can rewrite the equations for the state and pricing kernel in terms of z'':

$$\begin{aligned} z''_{t+1} &= (1-\varphi)\left(\frac{\gamma'\theta'}{\gamma_0}\right) + \varphi z''_t + \beta\left(\frac{\gamma'}{\gamma_0}\right)^{1/2} z''^{1/2}_t \varepsilon_{t+1} \\ &= (1-\varphi)\theta'' + \varphi z''_t + \beta'' z''^{1/2}_t \varepsilon_{t+1} \end{aligned}$$

and

$$\begin{aligned} -\log m_{t+1} &= \delta' + \gamma_0 z''_t + \lambda\left(\frac{\gamma_0}{\gamma'}\right)^{1/2} z''^{1/2}_t \varepsilon_{t+1} \\ &= \delta' + \gamma_0 z''_t + \lambda'' z''^{1/2}_t \varepsilon_{t+1}, \end{aligned}$$

which are in the form of (18Γ30). When$\gamma' < 0$ this procedure does not workΓ since we would be taking the square root of $z'' < 0$. We instead define $z'' = -(\gamma_0/\gamma')z'$ and proceed analogously. The equation for the pricing kernel in this case becomes

$$-\log m_{t+1} = \delta' - \gamma_0 z''_t + \lambda'' z''^{1/2}_t \varepsilon_{t+1},$$

which is also in the form of (30).

We use two normalizations in the paper. One is "$\pm\gamma_0$" $= 1 + \lambda^2/2$Γ which results in z being the short rate. The other is "$\pm\gamma_0$" $= -1 + \lambda^2/2$Γ which allows the model to account for the anomaly.

An alternative representation of the one-factor affine model is

$$z_{t+1} = (1-\varphi)\theta + \varphi z_t + (\alpha + \beta z_t)^{1/2}\varepsilon_{t+1}$$

and

$$-\log m_{t+1} = (1 + \lambda^2/2)z_t + \lambda(\alpha + \beta z_t)^{1/2}\varepsilon_{t+1}.$$

This effectively loads the sign change for γ_0 into β and the intercept δ into α. This representation has the property that it reproduces the nonstochastic volatility case when $\beta = 0$.

A.2 Affine Models of Bond Pricing

Duffie and Kan (1993) characterize a class of affine bond pricing models in continuous time. We translate their class of models into discrete time and derive conditions under which bond yields are strictly positive.

Duffie and Kan's affine models are based on a k-dimensional vector of state variables z following

$$z_{t+1} - z_t = (I - \Phi)(\theta - z_t) + V(z)^{1/2}\varepsilon_{t+1}, \tag{31}$$

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

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

and β_i has nonnegative elements. State prices are governed by a pricing kernel of the form

$$-\log m_{t+1} = \delta + \gamma^\top z_t + \lambda^\top V(z_t)^{1/2}\varepsilon_{t+1}. \tag{32}$$

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 (1993Γ 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 (31Γ32) and satisfying Condition A as the*Duffie-Kan class of affine models.*

Our characterization of these models differs from Duffie and Kan's in two respects. FirstΓ Duffie and Kan write (31) as

$$z_{t+1} - z_t = (I - \Phi)(\theta - z_t) + \Sigma V(z)^{1/2}\varepsilon_{t+1}, \tag{33}$$

which includes a matrix Σ that is missing in our version. We show that our choice is innocuous by reducing their model to ours. As in Section A.1Γ the key is that z is not directly observable. Assume Σ is invertible (this is convenient but not essential) and define $z' = \Sigma^{-1}z$. If we substitute for zΓ equation (33) becomes

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

with $v'_i(z') = \alpha_i + \beta_i'^\top z'$Γ $\Phi = \Sigma^{-1}\Phi\Sigma$Γ$\theta' = \Sigma^{-1}\theta$Γ and$\beta_i'^\top = \beta_i^\top\Sigma$. Equation (32) 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.

Our second difference from Duffie and Kan is the assumption that the volatility parameters β_i are nonnegative. 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.

Given this structure bond prices are log-linear in the state variables z. If b_t^n is the price at date t of a claim to one dollar in all states at date $t+n$ then by log-linearity we mean that

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

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:

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

where $A'(n) = n^{-1}A(n)$ and $B'(n) = n^{-1}B(n)$. We use equation (6) to generate parameters recursively:

$$\begin{aligned} A(n+1) &= A(n) + \delta + B(n)^{\top}(I - \Phi)\theta - \frac{1}{2}\sum_{j=1}^{k} (\lambda_j + B(n)_j)^2 \alpha_j \\ B(n+1)^{\top} &= \left(\gamma^{\top} + B(n)^{\top}\Phi\right) - \frac{1}{2}\sum_{j=1}^{k} (\lambda_j + B(n)_j)^2 \beta_j^{\top}, \end{aligned}$$

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

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

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 we define the vector $A^{\top} = [A'(n_1), \ldots, A'(n_k)]$.

We now turn to a smaller class of 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 consequenceΓ 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. SupposeΓ in contradiction to the lemmaΓ thatβ has less than full rank. Then there exists a nonzero vector h satisfying $\beta^\top h = 0$. For any admissible $z\Gamma z' = z + \rho h$ 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 invertibleΓ then we need

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

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

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

By assumptionΓB is invertible so $B^\top h \neq 0$. Thus we can choose ρ to make yields as negative as we likeΓ 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. ∎

Our final result is that in this environment (univariate volatility functions)Γ $Var(z)$ has nonnegative elements:

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

The proof hinges on Condition A(a)Γ Duffie and Kan's multivariate analog of the Feller condition. Since β is diagonal with positive elementsΓ 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 $v_i(z) = 0\Gamma$ 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 $z_j = \theta_j\Gamma$ so$\kappa_{ii} > 0$. Since $\Phi = I - K$ has positive diagonal elementsΓ by assumption$\Gamma\kappa_{ii} < 1$.

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. ■

We have addressed positive bond yields (Lemma 1) and nonnegative covariances (Lemma 2) in affine models for *admissible* values of the state, which need not be the same as values of the state that occur with positive probability. Consider, for example, the bivariate process

$$\begin{aligned} z_{1t+1} &= (1-\varphi)\theta + \varphi z_t + z_{1t}^{1/2}\varepsilon_{1t+1} \\ z_{2t+2} &= z_{1t} + z_{1t}^{1/2}\varepsilon_{2t+1}. \end{aligned}$$

The admissible region for this model includes any real number for z_2, yet by design z_2 cannot be negative (subject to the discrete time approximation). Nevertheless, the covariance between z_1 and z_2 is positive.

A.3 Proof of Proposition 2

Proposition 2 is based on a model with 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 $\omega = \sum_j \lambda_j^2\alpha_j^1/2$, $\omega^* = \sum_j \lambda_j^{*2}\alpha_j^2/2$, $\tau = \sum_j \lambda_j^2\beta_j^1/2$, and $\tau^* = \sum_j \lambda_j^{*2}\beta_j^2/2$. The forward premium is therefore

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

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

$$0 > Cov(s_{t+1} - s_t, f_t - s_t) = (\gamma - \tau)^\top Var(z_1)\gamma + (\gamma^* - \tau^*)^\top Var(z_2)\gamma^*. \qquad (34)$$

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 $\gamma - \tau\Gamma$ and hence of $\gamma\Gamma$ 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. Similar reasoning applies to the second term in (34) so we conclude that the model cannot reproduce the anomaly with strictly positive interest rates. ■

An example indicates that the proposition is limited to the Duffie-Kan class of affine models. Consider a model based on an iid state variable z drawn from the uniform distribution on [0,1] with pricing kernel

$$-\log m_{t+1} = \delta + \gamma z_t + \beta z_t^{1/2}\varepsilon_{t+1},$$

with $\{\varepsilon_t\} \sim \text{NID}(0,1)$ and independent of z. Then the short rate is

$$r_t = \delta + (\gamma - \beta^2/2)z_t.$$

If $\gamma < \beta^2/2$ and $\delta + \gamma - \beta^2/2 = 0$ the short rate varies between 0 and δ and is thus always positive. With a similar model for the foreign interest rate based on an analogous state variable z^* the forward premium is $(\gamma - \beta^2/2)(z - z^*)$ and the forward premium regression has slope

$$\alpha_2 = \frac{\gamma}{\gamma - \beta^2/2},$$

which is negative under the stated conditions. This example is affine in the sense that yields are linear functions of state variables and is capable of accounting for the anomaly with positive interest rates but it is not in the Duffie-Kan class.

A.4 Distribution of State Variables

Consider a state variable z following the square root process (18). For short time intervals the unconditional distribution of z is approximately Gamma with density

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

and parameters $a, b > 0$. This conclusion relies on a similar statement by Cox Ingersoll and Ross (1985) for their continuous-time analog and Sun's (1992) demonstration that the

discrete time model converges to Cox-Ingersoll-Ross. The mean and variance for a Gamma random variable are ab and $ab^2\Gamma$ which defines the parameters as

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

In the continuous-time limit $1-\varphi^2 \to 2\kappa = 2(1-\varphi)\Gamma$ so $a \to 2(1-\varphi)\theta/\sigma^2\Gamma$ the Feller ratio in inequality (26). We compute other moments using the moment generating function:

$$\phi(s) = (1-bs)^{-a},$$

for $s \leq 1/b$. Indicators of the first four moments are

$$\begin{aligned} \mu_1(\text{mean}) &= \theta \\ \mu_2(\text{variance}) &= \frac{\theta\sigma^2}{(1-\varphi^2)} \\ \gamma_1(\text{skewness}) &= \frac{2}{a^{1/2}} = 2\left(\frac{\sigma^2}{(1-\varphi^2)\theta}\right)^{1/2} \\ \gamma_2(\text{kurtosis}) &= \frac{6}{a} - 3 = \frac{6\sigma^2}{(1-\varphi^2)\theta} - 3. \end{aligned}$$

In continous time the Feller condition (19) is equivalent to $a \geq 1$. Thus parameter values implying small a violate the Feller condition and generate large values of the skewness and kurtosis measures γ_1 and γ_2.

References

Ahn, Dong-Hyun, 1995, "Common factors and local factors: Implications for term structures and exchange rates," unpublished manuscript, New York University, October.

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.

Backus, David, and Stanley Zin, 1994, "Reverse engineering the yield curve," unpublished manuscript, Stern School of Business, New York University, March.

Bakshi, Gurdip, and Zhiwu Chen, 1995, "Equilibrium valuation of foreign exchange claims," unpublished manuscript, University of New Orleans, July

Bansal, Ravi, 1991, "Can nonseparabilities explain exchange rate movements and risk premia?" unpublished manuscript, Carnegie Mellon University.

Bansal, Ravi, 1995, "The dynamics of multiple term structures and exchange rate movements," unpublished manuscript, Fuqua School of Business, Duke University, July

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, 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, 1992, "The implications of first-order risk aversion for asset market risk premiums," unpublished manuscript, Northwestern University, December.

Boessaerts, 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.

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, 1993, "A yield-factor model of interest rates," unpublished manuscript, Stanford University, September.

Duffie, Darrell, and Kenneth Singleton, 1995, "An econometric model of the term structure of interest rate swap yields," unpublished manuscript, Stanford University, September.

Engel, Charles, 1995, "The forward discount anomaly and the risk premium: A survey of recent evidence," NBER Working Paper No. 5312, October; forthcoming, *Journal of Empirical Finance*.

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, 1994, "Fixes: Of the forward discount puzzle," NBER Working Paper No. 4928, November.

Frachot, Antoine, 1994, "A reexamination of the uncovered interest rate parity hypothesis," unpublished manuscript, Banque de France, January.

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, 1995, "An examination of uncovered interest rate parity in segmented international commodity markets," unpublished manuscript, University of British Columbia, April.

Kritzman, Mark, 1994, "What practitioners need to know... About higher moments," *Financial Analysts Journal 50* (September/October), 10-17.

Macklem, Tiff, 1991, "Forward exchange rates and risk premiums in artificial economies," *Journal of International Money and Finance 10*, 365-391.

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 Dev	Skewness	Kurtosis	Autocorr
		A. Depreciation Rate$s_{t+1} - s_t$			
British Pound	−0.0017	0.0342*	−0.187	2.075	0.084
Canadian Dollar	−0.0015	0.0122*	−0.343	0.636	0.057
French Franc	−0.0005	0.0328*	−0.357	1.198	−0.002
German Mark	0.0021	0.0340*	−0.289	0.901	−0.015
Italian Lira	−0.0038	0.0334*	−0.712	1.961	0.049
Japanese Yen	0.0044	0.0324*	0.403	0.622	0.067
		B. One-Month Interest Rater_t			
American Dollar	0.0069*	0.0030*	0.996	0.884	0.957*
British Pound	0.0093*	0.0027*	0.045	−0.077	0.915*
Canadian Dollar	0.0081*	0.0028*	0.822	1.145	0.965*
French Franc	0.0091*	0.0035*	2.380	7.906	0.755*
German Mark	0.0053*	0.0020*	0.652	−0.235	0.969*
Italian Lira	0.0122*	0.0045*	1.732	4.140	0.743*
Japanese Yen	0.0046*	0.0020*	0.564	2.095	0.914*
		C. Forward Premium$f_t - s_t = r_t - r_t^*$			
British Pound	−0.0024*	0.0027*	−0.053	1.169	0.900*
Canadian Dollar	−0.0014*	0.0014*	0.018	0.480	0.842*
French Franc	−0.0023*	0.0032*	−0.670	2.546	0.660*
German Mark	0.0017*	0.0029*	−0.573	0.088	0.953*
Italian Lira	−0.0056*	0.0045*	−2.117	6.384	0.724*
Japanese Yen	0.0021*	0.0029*	0.298	0.303	0.888*

Entries are sample moments of depreciation rates$s_{t+1} - s_t$ one-month eurocurrency interest ratesr_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, Skewness an estimate of the skewness measure γ_1, Kurtosis an estimate of the kurtosis measure γ_2, and Autocorr the first autocorrelation. The skewness and kurtosis measures are defined, specifically, in terms of central moments μ_j: $\gamma_1 = \mu_3/\mu_2^{3/2}$ and $\gamma_2 = \mu_4/\mu_2^2 - 3$. Both are zero for normal random variables. Our estimates replace population moments with sample moments.

Table 2

Forward Premium Regressions

Currency	α_1	α_2	Std Er	R^2
British Pound	−0.0062	−1.840	0.0339	0.0213
	(0.0027)	(0.847)		
Canadian Dollar	−0.0036	−1.575	0.0120	0.0341
	(0.0009)	(0.460)		
French Franc	−0.0021	−0.674	0.0328	0.0042
	(0.0031)	(0.827)		
German Mark	0.0033	−0.743	0.0340	0.0041
	(0.0025)	(0.805)		
Italian Lira	−0.0042	−0.073	0.0335	0.0001
	(0.0039)	(0.453)		
Japanese Yen	0.0080	−1.711	0.0320	0.0230
	(0.0024)	(0.643)		

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

$$s_{t+1} - s_t = \alpha_1 + \alpha_2(f_t - s_t) + \text{residual},$$

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 and Std Er is the estimated standard deviation of the residual.

Figure 1
Mean Depreciation Rates and Forward Premiums