Risk Premia in Forward Foreign Exchange Markets: A Comparison of Signal Extraction and Regression Methods

Prasad V. Bidarkota

Department of Economics, Florida International University

Abstract: We investigate time varying risk premia in forward dollar/pound monthly exchange rates over the last two decades. We study this issue using a signal plus noise model and separately using regression techniques. Our models account for time varying volatility and non-normalities in the observed series. Our signal plus noise model fails to isolate a statistically significant risk premium component whereas our regression model does. We attribute the discrepancy in the results from the two methods to the low power of the signal plus noise model in discriminating between a time varying risk premium component and a serially uncorrelated spot exchange rate expectational error. An important reason for the low power of the signal plus noise model is its failure to use information on current period forward rates in extracting the risk premium.

Key phrases: spot foreign exchange rates; forward foreign exchange rates; time-varying risk premium; signal extraction; non-normality; volatility persistence

JEL Codes: F31, C5, G12

August 19, 2004

Address for correspondence: P.V.Bidarkota, Department of Economics, University Park DM 320A, Florida International University, Miami, FL 33199, USA; Tel: +1-305-348-6362; Fax: +1-305-348-1524
E-mail address: bidarkot@fiu.edu

Introduction

The possible existence of risk premia in forward foreign exchange rates has been extensively investigated in the literature. Lewis (1995) and Engel (1996) provide surveys of this literature. That forward rates do not provide conditionally unbiased forecasts of future spot exchange rates has been firmly established in a number of studies. Under rational expectations, this lack of unbiasedness implies the existence of risk premia in forward foreign exchange rates.

Several studies have attempted to measure the size of the risk premium and to characterize its time series properties. Several alternative approaches have been tried in this regard, including regression techniques (see, for instance, Fama 1984, and Lewis 1995), vector autoregressions (VARs) (Canova and Ito 1991), signal extraction methods (Wolff 1987, 2000, Cheung 1993, Hai et al. 1997, Bhar et al. 2002), and survey based methods (Froot and Frankel 1989).

Regression-based approaches involve regressing the ex post forward bias (or alternatively, the change in the spot exchange rate) on variables available in the information set, such as the forward premium. The choice of the explanatory variables is often arbitrary (Hansen and Hodrick 1980).

Signal extraction methods obviate the need to specify explanatory variables. Wolff (1987) provides early estimates of the risk premium using a univariate version within this framework (see also Nijman et al. 1993 for a clarification). Signal extraction based on a bivariate model has been attempted in Hai at al. (1997). In all these models, after obtaining maximum likelihood

estimates of unknown parameters, full sample smoothed estimates of the state vector (see, for instance, Harvey 1989, p. 149) are generated that provide estimates of the unobservable risk premium.

Both regression and signal plus noise models adopt a homoskedastic Gaussian framework. However, several studies have documented that spot exchange rates are non-Gaussian (Booth and Glassman, 1987, and Tucker and Pond, 1988) and, as summarized by So (1987), so are forward foreign exchange rates. Time varying volatility in these rates has also been widely documented (see Frankel and Rose 1995 for a survey article on empirical research on nominal exchange rates).

Failure to take into account any potential non-normalities and conditional heteroskedasticity results in estimation inefficiencies. Accurate and precise estimation and characterization of the time series properties of the risk premia are important. This is because intertemporal equilibrium models seeking to explain the behavior of forward foreign exchange rates are judged based on whether or not they can account for these time series properties (see, for instance, Backus et al. 1993 for such an exercise, and Engel 1996 for a survey on such efforts).

In this study, we investigate the possible presence of risk premia in monthly forward dollar/pound exchange rates for the past two decades using both regression and signal extraction methods, taking into account any non-normalities and volatility persistence that may exist. Our estimation methods are more

efficient, precisely because they take these features of the data into account. We compare the statistical outcomes from the two techniques, and explore some reasons why the two methods draw differing conclusions on the presence of a risk premium component.

This paper is organized as follows. Section 2 lays out the signal extraction framework to identify risk premia in forward exchange rates. Section 3 provides summary statistics on the data and presents empirical results of estimating the signal extraction model. Section 4 describes, and presents empirical results on, the main hypotheses of interest regarding the nature and existence of the risk premium. Section 5 examines the risk premium issue using regression-based methods. Section 6 provides a discussion of why the empirical restuls from the two methods differ. Section 7 concludes with some observations derived from our analysis.

2. A Signal Plus Noise Model for the Risk Premium

In this section, we set out the signal plus noise model for the risk premium. We lay out the model and discuss its key features in section 2.1, and in section 2.2 discuss some issues that arise in its estimation.

2.1 Signal Plus Noise Model

Let F_t^{t+1} denote the forward foreign exchange rate observed at time t for currency to be delivered at time $t+1$. Let S_t denote the spot exchange rate observed at time t and let $E_t(.)$ represent the mathematical expectation conditional on all the information available at time t. Let lowercase letters denote the natural logarithms of these variables. Then, we have:

$$f_t^{t+1} = E_t(s_{t+1}) + p_t \tag{1}$$

where p_t is interpreted as an unobservable risk premium.

Subtracting s_{t+1} from both sides of Equation (1), we get:

$$f_t^{t+1} - s_{t+1} = E_t(s_{t+1}) - s_{t+1} + p_t \tag{2}$$

which can be rewritten as:

$$f_t^{t+1} - s_{t+1} = p_t + v_{t+1} \tag{3a}$$

or, defining the ex post forward bias $y_{t+1} \equiv f_t^{t+1} - s_{t+1}$, as:

$$y_{t+1} = p_t + v_{t+1} \tag{3b}$$

where $v_{t+1} \equiv E_t(s_{t+1}) - s_{t+1}$ is a serially uncorrelated white noise error term reflecting new information about s_{t+1} that arrives between time t and $t+1$.

Equation (3a) or (3b) can be viewed as the observation equation of a state space (or unobserved components) model where p_t is the signal of interest that is

only observed contaminated with noise ν_{t+1}. Our interest here is in extracting p_t from its noisy observable indicator $f_t^{t+1} - s_{t+1}$.

In order to make signal extraction operational, we need to specify a law of motion for the state variable p_t (the state transition equation). There is some evidence in the literature that risk premia exhibit persistence over time (see Canova and Ito 1991 and Engel 1996). Therefore, following Wolff (1987) and Nijman et al. (1993), we choose a simple first-order autoregression to characterize the dynamics of p_t:

$$(p_t - \mu) = \phi(p_{t-1} - \mu) + \eta_t \tag{4}$$

Equations (3b) and (4) together constitute our state space model.

We need to specify the distribution of the errors ν_{t+1} and η_t in order to complete the description of the state space model. Wolff (1987) assumes homoskedastic Gaussian distributions for these errors. Hai et al. (1997) also assume homoskedastic Gaussian errors for their bivariate model for spot and forward rates that features a common unobserved component. However, there is evidence of volatility clustering and fat tails in the distribution of spot exchange rates (Booth and Glassman, 1987, and Tucker and Pond, 1988), forward rates (see the references in So 1987), and risk premia (Canova and Ito 1991, and Engel 1996). Our specification of the signal plus noise model is therefore designed to reflect these twin features.

6

Accordingly, we model $v_{t+1} \equiv c_t z_{1t+1}$ where $z_{1t+1} \sim$ iid $S_\alpha(0,1)$. A random variable X is said to have a symmetric stable distribution $S_\alpha(0,c)$, if its log-characteristic function can be expressed as:

$$\ln E \exp(iXt) = i\delta t - |ct|^\alpha. \tag{5}$$

The parameters $c > 0$ and $\delta \in (-\infty, \infty)$ are measures of scale and location, respectively, and $\alpha \in (0,2]$ is the characteristic exponent governing the tail behavior, with a smaller value of α indicating thicker tails. The normal distribution belongs to the symmetric stable family with $\alpha = 2$, and is the only member with finite variance, equal to $2c^2$. Appendix A provides additional details on stable distributions.

The term c_t captures volatility clustering. It is posited as following a GARCH(1,1)-like process:

$$c_t^\alpha = \omega + \beta c_{t-1}^\alpha + \delta | y_t - E(y_t | y_1, y_2, ..., y_{t-1}) |^\alpha \tag{6}$$

with the restrictions $\omega > 0, \beta \geq 0,$ and $\delta \geq 0$. When the errors are normal (i.e. when $\alpha = 2$ is imposed), this model for volatility persistence reduces to the familiar GARCH-normal process.

The state error driving the risk premium is modeled as $\eta_t \equiv c_\eta c_t z_{2t}$ where $z_{2t} \sim$ iid $S_\alpha(0,1)$ and is completely independent of z_{1t+1} at all leads and lags. Here, $c_\eta \geq 0$ is the signal to noise scale ratio.

To summarize, our signal plus noise model for extracting risk premium in forward rates is the following:

$$y_{t+1} = p_t + v_{t+1}, \qquad v_{t+1} \sim c_t z_{1t+1}, \quad z_{1t+1} \sim \text{iid } S_\alpha(0,1) \qquad (7a)$$

$$(p_t - \mu) = \phi(p_{t-1} - \mu) + \eta_t \qquad \eta_t \sim c_\eta c_t z_{2t}, \quad z_{2t} \sim \text{iid } S_\alpha(0,1) \qquad (7b)$$

$$c_t^\alpha = \omega + \beta c_{t-1}^\alpha + \delta \mid y_t - E(y_t \mid y_1, y_2, ..., y_{t-1}) \mid^\alpha \qquad (7c)$$

We shall refer to this most general model described in Equations (7) as Model 1.

2.2 Estimation Issues

The conditionally non-normal nature of the state space model in Equations (7) creates complications in estimation, even without the presence of conditional heteroskedasticity. This is because the Kalman filter is no longer optimal due to the non-Gaussian nature of the shocks.

However, the general recursive filtering algorithm due to Sorenson and Alspach (1971) provides the optimal filtering and predictive densities under any given distributions for the errors, and a formula for computing the likelihood function. Appendix B gives these formulae. The recursive equations for computing the filtering and predictive densities are given in the form of integrals, whose closed-form analytical expressions are generally intractable, except in very special cases. In this paper, we numerically evaluate these integrals. Details on the numerical implementation procedure adopted are given in Appendix C.

The stable distribution and density may be evaluated by using Zolotarev's (1986, p.74,78) proper integral representations, or by taking the inverse Fourier transform of the characteristic function. McCulloch (1996a) has developed a fast numerical approximation to the stable distribution and density that has an expected relative density precision of 10^{-6} for $\alpha \in [0.84,2]$. We therefore restrict ourselves in this paper to stable distributions with α in this range for computational convenience.

There is some empirical evidence of skewness in risk premia (Canova and Ito 1991). Although asymmetric stable distributions exist and are well-defined, the fast numerical approximation to the stable distribution and density functions developed by McCulloch (1996a) works only for the symmetric stable distributions. Hence, we restrict ourselves to these symmetric distributions in this paper.

Lumsdaine (1996) shows that the effect of initial values in the GARCH volatility process on the properties of the parameter estimators in GARCH(1,1) and IGARCH(1,1) models is asymptotically negligible. Diebold and Lopez (1995) suggest setting the initial conditional variance (equal to $2c_0^2$, when it exists) equal to the sample variance at the first iteration and at subsequent iterations to the sample variance from a simulated realization with the estimated parameters (from the previous iteration). Engle and Bollerslev (1986) suggest initializing the GARCH process using estimates of c_0 based on sample values. Here, we set the

9

value of c_0 equal to its unconditional value obtained from the volatility process in Equation (7c).

3. Empirical Results

3.1 Summary Statistics

We work with monthly US dollar / British pound exchange rates obtained from DataStream.[1] One month forward and subsequently observed spot rates span the period November 1983 through June 2004.

Figure 1 plots the ex post forward bias $f_t^{t+1} - s_{t+1}$ that is composed of a risk premium and an expectational error, as given in Equation (3a). Summary statistics indicate a mean bias of -0.30 percent per month that is not statistically significant at the 0.05 level and a variance of 0.10 percent. The series has the skewness coefficient of 0.29 (p-value of 0.03) and kurtosis of 5.53 (p-value for $kurtosis = 3$ of 2.5e-16). The Jarque-Bera test strongly rejects normality (p-value of 9.4e-16).

3.2 Estimation Results

The maximum likelihood (ML) estimates of Model 1 are presented in the first row of Table 1. The results indicate an estimated mean risk premium μ of -

---

[1] Thanks are due to K.M. Kiani for assistance in acquiring this data series.

0.25 percent per month. The ARCH parameter δ is estimated to be 0.03, and the volatility persistence parameter β is 0.93. The AR coefficient of the risk premium ϕ is fairly persistent at an estimate of 0.53, although its scale is only 0.31 times the scale of the expectational error in the spot exchange rate (this is the estimate of the signal-to-noise scale ratio c_η). The characteristic exponent α is estimated to be 1.89.

A plot of the mean estimate of the risk premium obtained from the filter density, $E\left(p_t \mid y_1, y_2, ..., y_t\right)$, appears in Figures 2 and 3, along with the ex post forward bias. The figures show the risk premium to be generally small, usually less than one percent in magnitude. An exception occurs in December 1992 when the risk premium exceeds 5 percent. The premium switches signs often, taking positive and negative values at various times. As evident in Figure 2, its variability is small compared to the variability of the ex post forward bias. This is also reflected in the small signal-to-noise scale ratio c_η of 0.31.

4. Hypotheses Tests

In this section, we describe in detail several hypotheses of interest concerning the risk premium. Restricted models under the null hypotheses are set up in each instance. Empirical estimates of the restricted models, and results of hypotheses tests, are reported and discussed.

4.1 Is The Risk Premium Constant?

In the first instance, we ask whether the risk premium is constant rather than time varying. To test the null hypothesis of a constant risk premium, we consider the following restricted version of the general Model 1 given in Equations (7):

$$y_{t+1} = \mu + v_{t+1}, \quad v_{t+1} \sim c_t z_{1t+1}, \quad z_{1t+1} \sim \text{iid } S_\alpha(0,1) \qquad (8a)$$

$$c_t^\alpha = \omega + \beta c_{t-1}^\alpha + \delta \, | \, y_t - E(y_t \, | \, y_1, y_2, ..., y_{t-1}) \, |^\alpha \qquad (8b)$$

Under this null hypothesis, the ex post forward bias is simply equal to a constant risk premium μ plus the expectational error in the conditional forecast of the spot exchange rate v_{t+1}. In what follows, we shall refer to this model with a constant risk premium described in Equations (8) as Model 2.

Maximum likelihood estimates of Model 2 are given in the second row of Table 1. Estimates of common parameters are very similar to those obtained with the time-varying risk premium Model 1. The only exception is the estimate of the volatility parameter ω which is now four times larger. This is understandable since all the variation in the ex post forward bias is now solely attributed to the expectational errors, rather than to a combination of expectational errors and time-varying risk premia.

The constant risk premium model imposes the two restrictions $\phi = c_\eta = 0$ on the time varying risk premium model. A test of the validity of these restrictions can be conducted with a likelihood ratio (LR) test. However, the standard LR test

is not applicable in this case. The reason is because, under the null hypothesis, the value of c_η lies on the boundary of admissible values for it. The derivation of the asymptotic χ^2 distribution of the LR statistic requires that the likelihood function be approximately quadratic in the region in which the null hypothesis and the global optima lie. This is clearly violated at the boundary. Therefore, standard asymptotic distribution theory does not go through.

Since estimation of the alternative Model 1 in our case is computationally very intensive, we generate small sample critical values for this test by Monte Carlo simulations from Gaussian homoskedastic versions of the null and alternative models. Homoskedastic models are discussed later in subsection 4.3 and Gaussian versions of all models in subsection 4.4.

The LR test statistic for the null hypotheses $\phi = c_\eta = 0$ (comparing Models 1 and 2) is 1.61. Critical values derived from the Monte Carlo simulation are 3.23 and 2.36 at the 0.05 and 0.10 significance levels, respectively. Thus, we fail to reject constancy of the risk premium even at the 0.10 significance level.

4.2 Is There a (Constant) Risk Premium?

Given the evidence against time variation in the risk premium, we go on to ask whether a risk premium actually exists in the forward foreign exchange markets. To test the null hypothesis of no risk premium, we consider the

following restricted version of the constant risk premium Model 2 given in Equations (8):

$$y_{t+1} = v_{t+1}, \quad v_{t+1} \sim c_t z_{1t+1}, \qquad z_{1t+1} \sim \text{iid } S_\alpha(0,1) \qquad (9a)$$

$$c_t^\alpha = \omega + \beta c_{t-1}^\alpha + \delta \mid y_t \mid^\alpha \qquad (9b)$$

This restricted model is obtained by setting $\mu = 0$ in Model 2. Under this null hypothesis, the ex post forward bias is simply equal to the expectational error in the conditional forecast of the spot exchange rate v_{t+1}. Henceforth, the model in Equations (9) will be referred to as Model 3.

ML estimates of Model 3 are presented in the third row of Table 1. Most parameter estimates are similar to the corresponding estimates obtained for Model 2. The LR test statistic for $\mu = 0$ is 1.958, with a p-value of 0.162 from the asymptotic χ_1^2 distribution. Thus, in the dollar/pound forward exchange market there does not appear to be any statistically significant risk premium.

4.3 What Happens in a Homoskedastic Setting?

A test for lack of volatility clustering, or equivalently a test for homoskedasticity, can be formulated as a test of $\beta = \delta = 0$. In this case, the model under the null hypothesis is obtained by setting $\beta = \delta = 0$ in the alternative Model 3.

This yields Model 4, which can be written as:

14

$$y_{t+1} = \nu_{t+1}, \quad \nu_{t+1} \sim c z_{1t+1}, \quad z_{1t+1} \sim \text{iid } S_\alpha(0,1) \tag{10}$$

ML estimates of the homoskedastic Model 4 are reported in the fourth row of Table 1. In this case, the constant scale parameter c is estimated to be 0.020. The LR test statistic for no volatility clustering in the ex post forward bias (test for $\beta = \delta = 0$) works out to be 19.393, with a p-value from the asymptotic χ^2_2 distribution of 6.1e-5. Thus, we overwhelmingly reject homoskedasticity in favor of volatility clustering.

Figure 4 plots the scales c_t obtained from Model 3. The figure clearly shows time-varying volatility, with the estimated scales ranging from 1.4 percent to 4.4 percent per month. Volatility shows two prominent spikes, one in June 1986 and a more prominent one in January 1993. The latter episode coincides with the surge in the risk premium in December 1992 noted earlier in section 3.2. After the 1993 spike, volatility falls to a level distinctly lower compared to the earlier period and remains low through the end of the sample.

In order to assess the consequences of ignoring conditional heteroskedasticity on the inferences drawn regarding the presence or absence of time-varying or constant risk premia in the forward foreign exchange rates, we consider homoskedastic versions of the general time-varying risk premium Model 1 and the constant risk premium Model 2.

A homoskedastic time-varying risk premium Model 5 takes the form:

$$y_{t+1} = p_t + \nu_{t+1}, \quad \nu_{t+1} \sim c z_{1t+1}, \quad z_{1t+1} \sim \text{iid } S_\alpha(0,1) \tag{11a}$$

15

$$(p_t - \mu) = \phi(p_{t-1} - \mu) + \eta_t \qquad \eta_t \sim c_\eta c z_{2t}, \quad z_{2t} \sim \text{iid } S_\alpha(0,1) \qquad (11b)$$

A homoskedastic constant risk premium Model 6 takes the form:

$$y_{t+1} = \mu + \nu_{t+1}, \quad \nu_{t+1} \sim c z_{1t+1}, \quad z_{1t+1} \sim \text{iid } S_\alpha(0,1) \qquad (12)$$

ML estimates of Models 5 and 6 are reported in the last two rows of Table 1. The LR test statistic for a constant risk premium in this homoskedastic setting, i.e. a test for the restriction $\phi = c_\eta = 0$ (comparing Model 5 versus Model 6) is 2.63, higher than in the conditionally heteroskedastic case. Using small sample critical values reported in the last paragraph of section 4.1, we reject constant risk premium in favor of time varying risk premium at the 0.10 significance level but fail to reject at the 0.05 level. Thus, ignoring conditional heteroskedasticity in the ex post forward bias could potentially lead to a false statistical inference in favor of time-varying risk premia in the forward dollar/pound exchange rates.

In order to assess the consequences of ignoring conditional heteroskedasticity on the inference on the presence of a constant risk premium, we can compare Models 6 and 4. The LR test statistic for the restriction $\mu = 0$ is now 3.865, with a p-value of 0.049 derived from the χ_1^2 distribution. Thus, once again it appears that ignoring volatility clustering leads us to infer incorrectly that there is a (time invariant) risk premium in the forward dollar/pound exchange rates.

4.4. What Happens in a Gaussian Setting?

Given the conclusions of the previous subsections that there is no risk premium in the dollar/pound forward rates (reached using conditionally heteroskedastic models) and that volatility clustering occurs in the ex post forward bias, we consider a Gaussian version of Model 3 in order to test for normality. The Gaussian null model is obtained by setting $\alpha = 2$ in Model 3

$$y_{t+1} = v_{t+1}, \quad v_{t+1} \sim \sqrt{2}c_t z_{1t+1}, \qquad z_{1t+1} \sim \text{iid } N(0,1) \qquad (13a)$$

$$c_t^2 = \omega + \beta c_{t-1}^2 + \delta \mid y_t \mid^2 \qquad (13b)$$

ML estimates of this model are presented in the third row of Table 2. Parameter estimates of the volatility process are very similar to those obtained in the stable case. Volatility persistence is a little lower (β estimate is 0.724 compared to 0.812 in the stable case) and the ARCH parameter is a little higher (δ estimate is 0.115 compared to 0.064 in the stable case). Test for normality can be based on testing for $\alpha = 2$. However, the LR test statistic has a non-standard distribution, since the null hypothesis lies on the boundary of admissible values for α, and, hence, the standard regularity conditions are not satisfied. The small-sample critical values for such a test have been tabulated in McCulloch (1997). The LR test statistic (comparing stable and Gaussian versions of Model 3) turns out to be 5.378 and the null hypothesis is rejected at better than the 0.01 significance level using critical value from McCulloch (1997).

We go on to ask what the consequences of ignoring non-normality would be on the inferences regarding the risk premium drawn from a signal plus noise

model. A homoskedastic version of such a Gaussian setup is employed by Wolff (1987).

We report results from estimating conditionally Gaussian versions of all the models discussed in section 4 in the remaining rows of Table 4. Figure 5 plots the estimated risk premia obtained with the Gaussian version of Model 1. As the figure shows, estimated premia are always negative in this setup. Compared with the estimates obtained from the stable Model 1 (plotted in Figure 3), we find that these estimates are much smaller in magnitude. Two further differences can be seen in the two figures. In June 1985, when there is a large dip in the ex post forward bias lasting for one period, the Gaussian model gives a large rise in the premium whereas the stable model gives an unremarkable (in magnitude) estimate. In January 1993, when there is a big jump in the ex post forward bias lasting for two periods, the Gaussian model gives a large fall in the premium whereas the stable model gives a huge increase. Thus, the Gaussian model attributes any large movements in the ex post bias to movements in risk premia (in the opposite direction) whereas the stable model infers large movements in risk premia only when there are sustained big changes in the ex post bias (and in the same direction). This contrasting behavior of Gaussian and stable state space models is further illustrated in a different context in Bidarkota and McCulloch (1998).

Figure 6 plots the estimated scales from the Gaussian version of Model 3. Once again, the figure shows highly non-constant and persistent volatility. Compared to the scales from the stable model in Figure 4, these are generally larger.

All the statistical inferences on the hypotheses of interest discussed in section 4 remain qualitatively unchanged in the Gaussian framework, with two exceptions. These exceptions have to do with the hypotheses of a time varying risk premium component (as opposed to a constant risk premium) and constant risk premium component (as opposed to no risk premium) in forward rates. First, in the Gaussian homoskedastic case, we fail to reject constant risk premium even at the 0.10 level (comparing Models 5 and 6). Second, contrary to the results under non-normality, we fail to reject the absence of a constant risk premium (comparing Models 4 and 6). Thus, our statistical inferences on these two hypotheses in a homoskedastic Gaussian framework are identical to those obtained in a conditionally heteroskedastic stable framework (cf. with the inferences in subsections 4.1 and 4.2).

4.5. Discussion

Our results on the lack of significant risk premium components in forward dollar/pound exchange rates using the signal plus noise model within a conditionally heteroskedastic non-normal setup are contrary to the findings in

19

other studies such as Wolff (1987), who uses a similar methodology in a homoskedastic Gaussian setting.

To help understand our negative results, we look at the autocorrelations of the ex post forward bias $y_{t+1} \equiv f_t^{t+1} - s_{t+1}$. All the autocorrelation and partial autocorrelation coefficients are less than 0.1 in magnitude and statistically insignificant at the 0.05 significance level.

However, it needs to be emphasized that lack of strong autocorrelations in y_{t+1} need not signal the absence of any time varying risk premium components. It is possible that the predictable component is obscured by a large white noise expectational error v_{t+1}. For instance, with the homoskedastic Gaussian Model 5, one can easily show that the first order autocorrelation coefficient is given by $\phi / \left\{ 1 + \left(1 - \phi^2 \right) / \sigma_\eta^2 \right\}$. Even with a ϕ as large as 0.9, this coefficient is only 0.05 when c_η is 0.1. This point is also emphasized by Fama (1984).

With the homoskedastic Gaussian Model 5 estimates, variance of the risk premium works out to be 4.09×10^{-4}, of which an overwhelming portion turns out to be the variance of noise in it (variance of η_t is 3.96×10^{-4}). Furthermore, variance of the expectational error v_{t+1} is 5.80×10^{-4}, which is larger than the variance of the risk premium. These variance decompositions are qualitatively very different from those reported in Wolff (1987), who finds the variance of the

risk premium to be larger than the variance of the expectational error. Thus, this provides further clues as to why our results regarding risk premia might be so different.

Most significantly, statistical inferences in Wolff (1987) are based on asymptotic χ_1^2 critical values for the LR test for the significance of the autoregressive coefficient in the risk premium dynamics, as in our Gaussian version of Equation (11b). However, as emphasized in section 4.1 earlier, these χ_1^2 critical values are invalid.

The magnitude of the LR test statistic in Wolff (1987) is quite large (9.857 for the dollar/pound exchange rate). Therefore, even with Monte Carlo critical values, it is quite likely that his LR test would in fact still reject no time varying risk premium. Therefore, the most likely explanation for why our results differ seems to be the different sample period used here.

5. Regression Model for Risk Premium

We now turn to an analysis of risk premia in forward rates using regression methods, as in Fama (1984) and several other studies (see Lewis 1995 for a survey). A typical practice in these types of studies is to run a regression of the ex post forward bias on the forward premium:

$$f_t^{t+1} - s_{t+1} = a + b\left(f_t^{t+1} - s_t\right) + u_{t+1} \qquad (14)$$

where u_{t+1} is a regression error. Absence of a risk premium component implies $a = b = 0$. A constant risk premium (as opposed to time varying premium) implies $a \neq 0$ and $b = 0$.

While most studies typically assume $u_{t+1} \sim \text{iid } N(0,\sigma^2)$, given the evidence on conditional heteroskedasticity and non-normality in the previous section, we entertain these possibilities in the regression residual above. Thus, we assume $u_{t+1} \equiv c_t z_{t+1}$ where $z_{t+1} \sim \text{iid } S_\alpha(0,1)$. Here, the time varying volatility c_t follows the process:

$$c_t^\alpha = \omega + \beta c_{t-1}^\alpha + \delta \, | f_{t-1}^t - s_t - a - b \left(f_{t-1}^t - s_{t-1} \right) |^\alpha \qquad (15)$$

Thus, our complete regression model is as follows:

$$f_t^{t+1} - s_{t+1} = a + b \left(f_t^{t+1} - s_t \right) + u_{t+1}, \; u_{t+1} \sim c_t z_{t+1}, \; z_{t+1} \sim \text{iid } S_\alpha(0,1) \qquad (16a)$$

$$c_t^\alpha = \omega + \beta c_{t-1}^\alpha + \delta \, | f_{t-1}^t - s_t - a - b \left(f_{t-1}^t - s_{t-1} \right) |^\alpha \qquad (16b)$$

Using our naming convention from the signal plus noise models, we call this the stable regression Model 1. We also consider three restricted versions of this regression model. A homoskedastic version of Model 1 is termed stable regression Model 5. Gaussian versions of these two models (with the restriction $\alpha = 2$ imposed) are referred to as Gaussian regression Model 1 and Model 5, respectively.

Maximum likelihood estimates of all these models are presented in Table 3. Estimate of the intercept is 0.335 percent per month with a standard error of 0.209 with the stable Model 1 (p-value of 0.109). Thus, it is statistically significant (barely) at the 0.11 significance level. However, the regression slope coefficient is estimated at 4.015 with a standard error of just 0.845, which is strongly statistically significant. Similar inferences follow with all the other model estimates reported in Table 3. In every case, the slope coefficient is strongly significant, although its point estimates from the Gaussian models are substantially smaller.

Figures 7 and 8 plot the time series of risk premia estimated with the stable regression Model 1 (fitted line in the regression). Comparison with plots of estimated risk premia with the stable signal plus noise Model 1 in figures 2 and 3 shows that the two estimates differ substantially. Risk premia from the regression model show a great deal more persistence (as evident in Figure 8 versus Figure 3). Premia take either negative (or positive) values for prolonged successive periods of time. Unlike the estimate from the signal plus noise model, there is no sharp positive spike in the premium in December 1992. In fact, the premium for that period is negative.

LR tests for homoskdasticity and normality are easily rejected at the usual significance levels. Estimated scales from the regression Model 1 are plotted in

Figure 9. Once again the figure shows large persistence in volatility. Once again, there is evidence of substantial reduction in volatility after its peak in 1993.

Overall, our regression results confirm the findings in several prior studies of a significant time varying risk premium in the dollar/pound forward exchange rates.

6. Explaining the Discrepancy between Signal Plus Noise Models and Regression Methods

Our results on the lack of significant risk premium components in forward dollar/pound exchange rates using the signal plus noise model are contrary to our findings using the regression approach.

To understand these results, we need to explain why the signal plus noise model is not very powerful at discriminating between a time varying risk premium component and noise. One likely reason for the low power in the signal plus noise models lies in its inherently univariate framework. The model only uses information on the ex post forward bias $f_t^{t+1} - s_{t+1}$. On the other hand the regression model uses information on the forward premium $f_t^{t+1} - s_t$ as well.

A second important reason has to do with the information set available at time t in the context of the two models. While the regression model incorporates at time t the available information on forward prices f_t^{t+1} in the explanatory

variable on the right hand side, the specific version of the signal plus noise model used in this paper (which was motivated by Wolff 1987) does not. In this model, information on the forward price f_t^{t+1} is in effect used only at time $t+1$ in the form of the ex post forward bias.

An examination of the maximized log-likelihood values across all models indicates that the stable conditionally heteroskedastic regression Model 1 dominates all other models. Although a comparison is strictly invalid because of the non-nested nature of the regression and signal plus noise models, the higher maximized log-likelihood values do suggest superior fit of the regression model to the ex post forward bias data. This is also revealed in a plot of the estimates of the risk premia in Figure 7 as compared to Figure 2. Estimates of the risk premia in Figure 2 appear closer to a white noise process than those in Figure 7.

7. Conclusions

In this paper we investigate the possible presence of time varying risk premia in forward dollar/pound monthly exchange rates for the period 1983:11 through 2004:6. We study this issue using two different methodologies. One is the univariate signal plus noise model, used in Wolff (1987, 2000), Nijman et al. (1993), and Cheung (1993). The other is a simple regression of the ex post forward bias on the forward premium. We improve on the previous studies by

explicitly taking into account volatility clustering and non-normalities documented in earlier studies on exchange rates.

Our signal plus noise model fails to reveal a statistically significant risk premium component in forward rates, in contrast to the results in Wolff (1987). Unlike this study, our inference is based on small sample critical values of the likelihood ratio test statistic generated by Monte Carlo simulations. Our regression method finds a statistically significant risk premium component, in accord with results from several studies on this issue.

We conclude that the univariate signal plus noise model is not very powerful at discriminating between a persistent risk premium component and a serially uncorrelated expectational error in spot exchange rates. Apart from its univariate nature, another important reason for lack of power in the signal plus noise model is the fact that the model, as formulated, does not use information on current period observed forward rates in extracting unobservable risk premium components.

REFERENCES

Backus, D.K., A.W. Gregory, and C.I. Telmer, 1993, 'Accounting for forward rates in markets for forward currency,' The Journal of Finance, Vol.XLVIII, No.5, 1887-1908.

Bhar, R., C. Chiarella, and T.M. Pham, 2002, 'Modeling the currency forward risk premium: Theory and evidence,' Unpublished manuscript, University of Technology (Sydney).

Bidarkota, P.V. and J.H. McCulloch, 1998, 'Optimal univariate inflation forecasting with symmetric stable shocks,' Journal of Applied Econometrics, 13, 659-670.

Boothe, P. and D. Glassman, 1987, 'The statistical distribution of exchange rates: Empirical evidence and economic implications,' Journal of International Economics, 22, 297-319.

Canova, F. and T. Ito, 1991, 'The time series properties of the risk premium in the yen/dollar exchange market,' Journal of Applied Econometrics, 6, 125-142.

Cheung, Y-W., 1993, 'Exchange rate risk premiums,' Journal of International Money and Finance, 12, 182-194.

Diebold, F.X. and J.A. Lopez, 1995, Modeling volatility dynamics, in K. Hoover (ed.), Macroeconometrics: Developments, Tensions and Prospects, Boston: Kluwer Academic Press, 427-72.

Engel, C., 1996, 'The forward discount anomaly and the risk premium: A survey of recent evidence,' Journal of Empirical Finance, 3, 123-192.

Engle, R.F. and T. Bollerslev, 1986, 'Modeling the persistence of conditional variances,' Econometric Reviews 5, 1-50.

Fama, E.F., 1984, 'Forward and spot exchange rates,' Journal of Monetary Economics, 14, 319-338.

Frankel, J.A. and A.K. Rose, 1997, Empirical Research on Nominal Exchange Rates, in Grossman, G.M., Rogoff, K., eds., Handbook of International Economics 3 (North-Holland) 1689-1729.

Froot, K.A. and J.A. Frankel, 1989, 'Forward discount bias: Is it an exchange risk premium?' The Quarterly Journal of Economics, 104, 139-161.

Hai, W., N.C. Mark, and Y. Wu, 1997, 'Understanding spot and forward exchange rate regressions,' Journal of Applied Econometrics, Vol.12, 715-734.

Hansen, L.P. and R.J. Hodrick, 1980, 'Forward exchange rates as optimal predictors of future spot rates: An econometric analysis,' Journal of Political Economy, Vol.88, No.5, 829-853.

Harvey, A.C, 1992, Forecasting, Structural Time Series Models and the Kalman Filter, Cambridge University Press, Cambridge, UK.

Kitagawa, G., 1987, 'Non-Gaussian state space modeling of nonstationary time series,' Journal of the American Statistical Association, Vol.82, No.400, 1032-63.

Lewis, K.K., 1995, Puzzles in international financial markets, in Grossman, G.M., Rogoff, K., eds., Handbook of International Economics 3 (North-Holland) 1913-1971.

Lumsdaine, R., 1996, 'Consistency and asymptotic normality of the quasi-maximum likelihood estimator in IGARCH(1,1) and covariance stationary GARCH(1,1) models,' Econometrica, Vol.64, No.3, 575-96.

McCulloch, J.H., 1997, 'Measuring tail thickness in order to estimate the stable index α: A critique,' Journal of Business and Economic Statistics, Vol.15, No.1, 74-81.

_____, 1996a, 'Numerical approximation of the symmetric stable distribution and density,' in R. Adler, R. Feldman, and M.S. Taqqu (eds.), A Practical Guide to Heavy Tails: Statistical Techniques for Analyzing Heavy Tailed Distributions, Boston: Birkhauser.

_____, 1996b. Financial applications of stable distributions, in: Maddala, G.S., Rao, C.R., eds., Handbook of Statistics, Vol.14 (Elsevier, Amsterdam) 393-425.

Nijman, T.E., F.C. Palm, and C.C.P. Wolff, 1993, 'Premia in forward foreign exchange as unobserved components: A note,' Journal of Business & Economic Statistics, Vol.11, No.3, 361-365.

So, J.C., 1987, 'The sub-Gaussian distribution of currency futures: stable Paretian or nonstationary?' The Review of Economics and Statistics, Vol. , No. , 100-107.

Sorenson, H.W. and D.L. Alspach, 1971, 'Recursive Bayesian estimation using Gaussian sums,' Automatica 7, 465-79.

Tucker, A.L. and L. Pond, 1988, 'The probability distribution of foreign exchange price changes: Tests of candidate processes," The Review of Economics and Statistics, Vol. , No. , 638-647.

Wolff, **C.C.P.**,2000, 'Measuring the forward foreign exchange risk premium: multicountry evidence from unobserved components models,' Journal of International Financial Markets, Institutions, and Money, 10, 1-8.

_____, 1987, 'Forward foreign exchange rates, expected spot rates, and premia: A signal-extraction approach,' The Journal of Finance, Vol.XLII, No.2, 395-406.

Zolotarev, V.M., 1986, One dimensional stable laws, American Mathematical Society. (Translation of *Odnomernye Ustoichivye Raspredeleniia* (NAUKA, Moscow, 1983).)

Appendix A. Stable Distributions and Their Properties

This section draws heavily from McCulloch (1996b). Stable distributions $S(x;\alpha,\beta,c,\delta)$ are determined by four parameters. The location parameter $\delta \in (-\infty,\infty)$ shifts the distribution to the left or right, while the scale parameter $c \in (0,\infty)$ expands or contracts it about δ, so that

$$S(x;\alpha,\beta,c,\delta) = S((x-\delta)/c;\alpha,\beta,1,0). \tag{A1}$$

The standard stable distribution function has $c = 1$ and $\delta = 0$. If a random variable X has a stable distribution, it is represented as $X \sim S(\alpha,\beta,c,\delta)$.

The characteristic exponent $\alpha \in (0,2]$ governs the tail behavior, and therefore the degree of leptokurtosis. When $\alpha = 2$, the normal distribution results, with variance $2c^2$. For $\alpha < 2$, the variance is infinite. When $\alpha > 1$, $E(X) = \delta$; but if $\alpha \leq 1$, the mean is undefined.

The skewness parameter $\beta \in [-1,1]$ is defined such that $\beta > 0$ indicates positive skewness. If $\beta = 0$, the distribution is symmetric stable. As $\alpha \uparrow 2$, β loses its effect and becomes unidentified.

Stable distributions are defined most concisely in terms of their log-characteristic functions:

$$\ln E \exp(iXt) = i\delta t + \psi_{\alpha,\beta}(ct) \tag{A2}$$

where

$$\psi_{\alpha,\beta}(t) = \begin{cases} -|t|^\alpha \, (1 - i\beta\,\text{sign}(t)\tan(\pi\alpha/2)) & \text{for}\, \alpha \neq 1 \\ -|t|\,(1 + i\beta(2/\pi)\text{sign}(t)\ln|t|) & \text{for}\, \alpha = 1 \end{cases} \tag{A3}$$

is the log-characteristic function for $S(\alpha,\beta,1,0)$.

When $\alpha < 2$, stable distributions have tails that behave asymptotically like $x^{-\alpha}$ and give the stable distributions infinite absolute population moments of order greater than or equal to α.

Let $X \sim S(\alpha,\beta,c,\delta)$ and a be any real constant. Then (A2) implies:

$$aX \sim S(\alpha,\text{sign}(a)\beta,|a|c,a\delta) .\qquad\qquad (A4)$$

Let $X_1 \sim (\alpha,\beta_1,c_1,\delta_1)$ and $X_2 \sim (\alpha,\beta_2,c_2,\delta_2)$ be independent drawings from stable distributions with a common α. Then $Y = X_1 + X_2 \sim S(\alpha,\beta,c,\delta)$, where

$$c^\alpha = c_1^\alpha + c_2^\alpha \qquad\qquad (A5)$$

$$\beta = (\beta_1 c_1^\alpha + \beta_2 c_2^\alpha) / c^\alpha \qquad\qquad (A6)$$

$$\delta = \begin{cases} \delta_1 + \delta_2 & \text{for } \alpha \neq 1 \\ \delta_1 + \delta_2 + 2(\beta c \ln(c) - \beta_1 c_1 \ln(c_1) - \beta_2 c_2 \ln(c_2)) / \pi & \text{for } \alpha = 1. \end{cases} \qquad (A7)$$

When $\beta_1 = \beta_2$, β equals their common value, so that Y has the same shaped distribution as X_1 and X_2. This is the "stability" property of stable distributions that leads directly to their role in the central limit theorem, and makes them particularly useful in financial portfolio theory. When $\beta_1 \neq \beta_2$, β lies between β_1 and β_2.

For $\alpha < 2$ and $\beta > -1$, the long upper Paretian tail of $X \sim S(\alpha,\beta,c,\delta)$ makes Ee^X infinite. However, when $\beta = -1$,

$$\ln \mathrm{E}e^{X} = \begin{cases} \delta - c^{\alpha}\sec(\pi\alpha/2), & \alpha \neq 1 \\ \delta + (2c/\pi)\ln c, & \alpha = 1 \end{cases} \qquad (A8)$$

This formula greatly facilitates asset pricing under log-stable uncertainty.

See also Zolotarev (1986, p.112) and McCulloch (1996b).

Appendix B. Sorenson-Alspach Filtering Equations

Let $y_t, t = 1,...,T$, be an observed time series and x_t an unobserved state variable, stochastically determining y_t. Denote $Y_t = \{y_1,...,y_t\}$. The recursive formulae for obtaining one-step ahead prediction and filtering densities, due to Sorenson and Alspach (1971), are as follows:

$$p(x_t|Y_{t-1}) = \int_{-\infty}^{\infty} p(x_t|x_{t-1})p(x_{t-1}|Y_{t-1})dx_{t-1}, \qquad (B1)$$

$$p(x_t|Y_t) = p(y_t|x_t)p(x_t|Y_{t-1})/p(y_t|Y_{t-1}), \qquad (B2)$$

$$p(y_t|Y_{t-1}) = \int_{-\infty}^{\infty} p(y_t|x_t)p(x_t|Y_{t-1})dx_t. \qquad (B3)$$

Finally, the log-likelihood function is given by:

$$\log p(y_1,...,y_T) = \sum_{t=1}^{T} \log p(y_t|Y_{t-1}). \qquad (B4)$$

These formulae have been applied to non-Gaussian data and extended to include a smoother formula by Kitagawa (1987). When shocks are normal ($\alpha = 2$ in our models), this filter collapses to the Kalman filter.

In the model given in Equations (7) in the main text, y_{t+1} is the observed series,

$$p(y_{t+1} \mid p_t) = s_\alpha(y_{t+1} - p_t; 0, c_t)$$

and
$$p(p_t \mid p_{t-1}) = s_\alpha(p_t - \mu - \phi(p_{t-1} - \mu); 0, c_\eta c_t),$$

where $s_\alpha(x; \delta, c)$ is the symmetric stable density. The filter is initialized by the unconditional distribution of the state variable since the process for risk premium is strictly stationary, i.e.

$$p(p_0 \mid Y_0) = s_\alpha(p_0; \mu, c_{p_0}), \ c_{p_0} = c_\eta c_0 / (1 - \phi^\alpha)^{1/\alpha}.$$

where c_0 is the unconditional mean of c_t which evolves according to the volatility process given in Equation (7c). Starting points for the hyperparameter estimation are obtained from the Kalman filter under normality.

Appendix C. Numerical Implementation of Filtering Equations

The Sorenson-Alspach filter and predictive densities were evaluated at a grid of 100 points equally spaced on a truncated portion of the real line. The left truncation point was chosen to lie 4 standard deviations (of the ε shock as measured by a preliminary Kalman filter) below the minimum observed excess return and the right truncation point 4 standard deviations above the maximum observed return. The likelihood and the predictive density integrals (Equations (B3) and (B1) resp.) were evaluated numerically by a piecewise cubic quadrature technique, as follows: Integration between any two interior nodes was performed by fitting a piecewise cubic function through the four nearest nodes and approximating the required area under the integrand between those nodes by the

area under the cubic. The outermost intervals employ the same cubics as the adjacent intervals. For equispaced nodes, 8 or more in number, this quadrature procedure yields the weights 8/24, 31/24, 20/24, 25/24, 1, 1, ..., 1, 25/24, 20/24, 31/24, 8/24 for the ordinates. The numerically computed predictive density was normalized in order to ensure that it integrated to unity. The piecewise linear interpolation and the trapezoidal rule for integration suggested by Kitagawa (1987) was not employed. Hodges and Hale (1993) propose an integration by parts procedure to speed up the Kitagawa procedure, but this was not employed either.

The accuracy of our numerical quadrature can be gauged by a comparison of the maximized log-likelihood value for Model 2 in Equations (4) obtained from our numerical integration with α restricted to be 2, with that obtained from the Kalman filter (which is optimal in this Gaussian case), for given values of the other hyperparameters. We verified that, with 100 nodes, our numerical approximation gives log-likelihood values accurate to one decimal place at the estimated hyperparameters of the Gaussian Model 2. In light of this our numerical integration appears to be sufficiently accurate for drawing valid inferences from data. Calculations were carried out in GAUSS on a Pentium personal computer.

Table 1: Signal Plus Noise Stable Model Estimates

$$y_{t+1} = p_t + v_{t+1}, \qquad v_{t+1} \sim c_t z_{1t+1}, \quad z_{1t+1} \sim \text{iid } S_\alpha(0,1) \tag{7a}$$

$$(p_t - \mu) = \phi(p_{t-1} - \mu) + \eta_t \qquad \eta_t \sim c_\eta c_t z_{2t}, \quad z_{2t} \sim \text{iid } S_\alpha(0,1) \tag{7b}$$

$$c_t^\alpha = \omega + \beta c_{t-1}^\alpha + \delta \, | \, y_t - E(y_t \, | \, y_1, y_2, ..., y_{t-1}) \, |^\alpha \tag{7c}$$

The most general model is Model 1 given in Equations (7) above. We get Model 2 by setting $\phi = c_\eta = 0$ in Model 1.

Imposing $\mu = 0$ on Model 2 gives Model 3. Setting $\beta = \delta = 0$ in Model 3 yields Model 4. Restricting $\beta = \delta = 0$ in Model 1

gives Model 5. Finally, imposing $\phi = c_\eta = 0$ on Model 5 gives Model 6. Hessian-based standard errors are reported in

parentheses beneath the parameter estimates.

Stable Models	α	μ $\times 10^2$	ω $\times 10^5$	β	δ	c $\times 10^2$	c_η	ϕ	LogL
Model 1	1.893	-0.252	1.130	0.928	0.026		0.309	0.527	522.213
	(0.090)	(0.198)	(1.249)	(0.050)	(0.018)		(0.211)	(0.213)	
Model 2	1.907	-0.244	4.649	0.773	0.074				521.408
	(0.092)	(0.173)	(3.714)	(0.123)	(0.043)				
Model 3	1.911		3.525	0.812	0.064				520.429
	(0.055)		(3.371)	(0.125)	(0.042)				
Model 4	1.863					1.995			510.733
	(0.059)					(0.104)			
Model 5	1.873	-0.377				1.808	0.349	0.512	513.981
	(0.094)	(0.213)				(0.159)	(0.188)	(0.156)	
Model 6	1.859	-0.364				1.971			512.666
	(0.067)	(0.184)				(0.103)			

Table 2: Signal Plus Noise Gaussian Model Estimates

$$y_{t+1} = p_t + v_{t+1}, \qquad v_{t+1} \sim \sqrt{2}c_t z_{1t+1}, \qquad z_{1t+1} \sim \text{iid } N(0,1)$$

$$(p_t - \mu) = \phi(p_{t-1} - \mu) + \eta_t \qquad \eta_t \sim \sqrt{2}c_\eta c_t z_{2t}, \qquad z_{2t} \sim \text{iid } N(0,1)$$

$$c_t^2 = \omega + \beta c_{t-1}^2 + \delta \, | \, y_t - E(y_t \, | \, y_1, y_2, ..., y_{t-1}) \, |^2$$

The most general Gaussian model is Model 1 given in the Equations above. We get Model 2 by setting $\phi = c_\eta = 0$ in Model 1. Imposing $\mu = 0$ on Model 2 gives Model 3. Setting $\beta = \delta = 0$ in Model 3 yields Model 4. Restricting $\beta = \delta = 0$ in Model 1 gives Model 5. Finally, imposing $\phi = c_\eta = 0$ on Model 5 gives Model 6. Hessian-based standard errors are reported in parentheses beneath the parameter estimates.

Gaussian Models	μ $\times 10^2$	ω $\times 10^5$	β	δ	c $\times 10^2$	c_η	ϕ	LogL
Model 1	-0.248	3.675	0.692	0.136		0.256	-0.140	519.012
	(0.155)	(3.128)	(0.119)	(0.079)		(0.939)	(0.517)	
Model 2	-0.245	3.936	0.691	0.128				518.957
	(0.156)	(2.706)	(0.120)	(0.052)				
Model 3		3.400	0.724	0.115				517.740
		(2.534)	(0.116)	(0.050)				
Model 4					2.237			502.557
					(0.101)			
Model 5	-0.308				1.705	0.827	0.172	504.349
	(0.217)				(3.152)	(3.825)	(0.879)	
Model 6	-0.297				2.227			503.651
	(0.200)				(0.100)			

37

$$f_t^{t+1} - s_{t+1} = a + b\left(f_t^{t+1} - s_t\right) + u_{t+1}, \ u_{t+1} \sim c_t z_{t+1}, \ z_{t+1} \sim \text{iid } S_\alpha(0,1) \qquad (16a)$$

$$c_t^\alpha = \omega + \beta c_{t-1}^\alpha + \delta \mid f_{t-1}^t - s_t - a - b\left(f_{t-1}^t - s_{t-1}\right) \mid^\alpha \qquad (16b)$$

The most general model is Model 1 given in Equations (16) above. We get Model 5 by setting $\beta = \delta = 0$ in Model 1. Setting $\alpha = 2$ gives Gaussian versions of these models. Hessian-based standard errors are reported in parentheses beneath the parameter estimates.

Regression Models	α	a	b	ω	β	δ	c	LogL
		$\times 10^2$		$\times 10^5$			$\times 10^2$	
Stable Models								
Model 1	1.907	0.335	4.015	9.774	0.628	0.096		531.107
	(0.075)	(0.209)	(0.845)	(5.697)	(0.155)	(0.046)		
Model 5	1.827	0.426	4.055				1.849	524.166
	(0.061)	(0.235)	(0.810)				(0.099)	
Gaussian Models								
Model 1	2	0.216	2.656	6.984	0.603	0.141		523.954
	(restricted)	(0.221)	(0.828)	(4.482)	(0.157)	(0.059)		
Model 5	2	0.429	3.434				2.162	510.981
	(restricted)	(0.270)	(0.883)				(0.097)	

Fig.1 Forward-Spot Exchange Rate Differentials

$$y_{t+1} \equiv f_t^{t+1} - s_{t+1}$$

ln(dollars per pound)

year
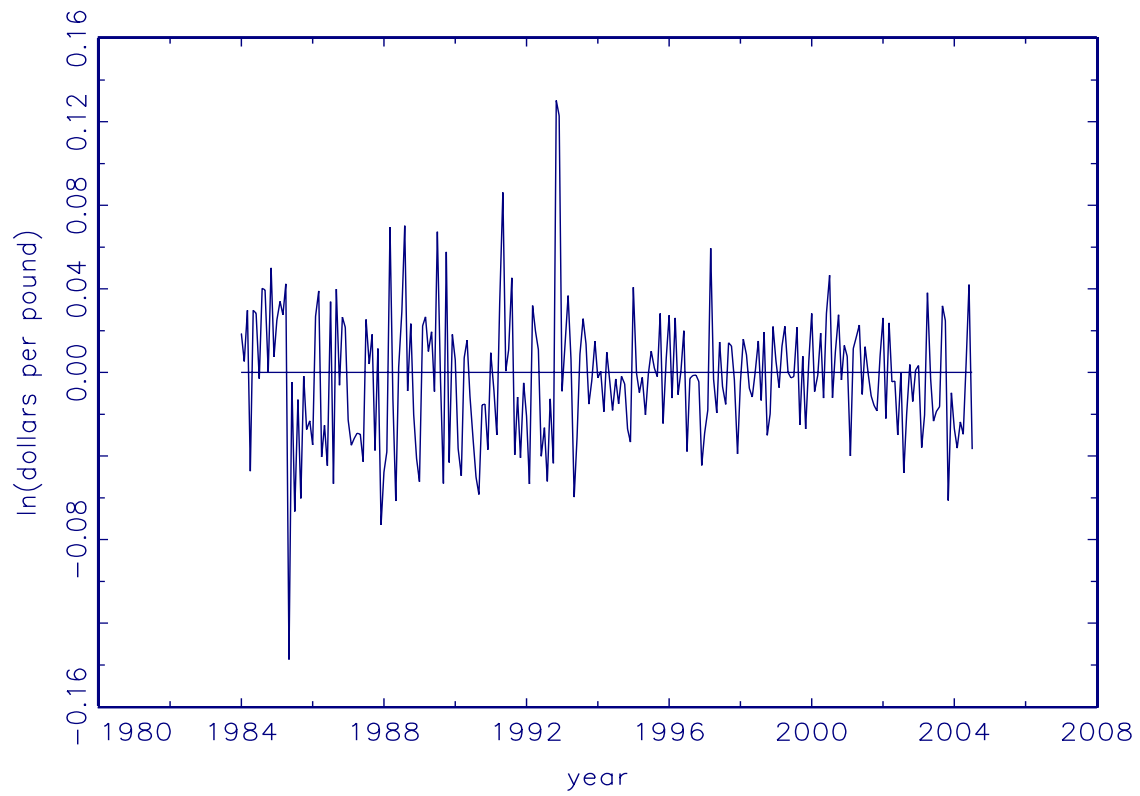
Fig.2 Estimates of Risk Premium from stable Model 1
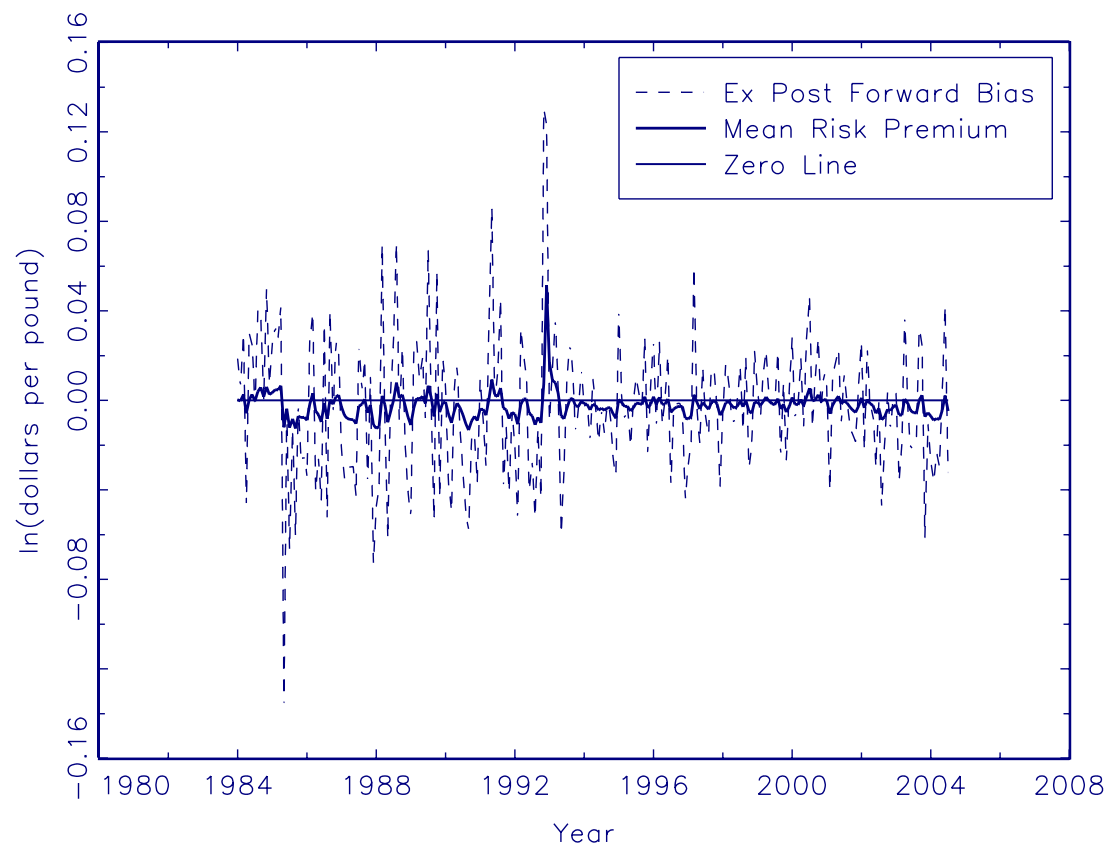
$$E\left(p_t \mid y_1, y_2, ..., y_t\right)$$

Fig.3 Estimates of Risk Premium from stable Model 1
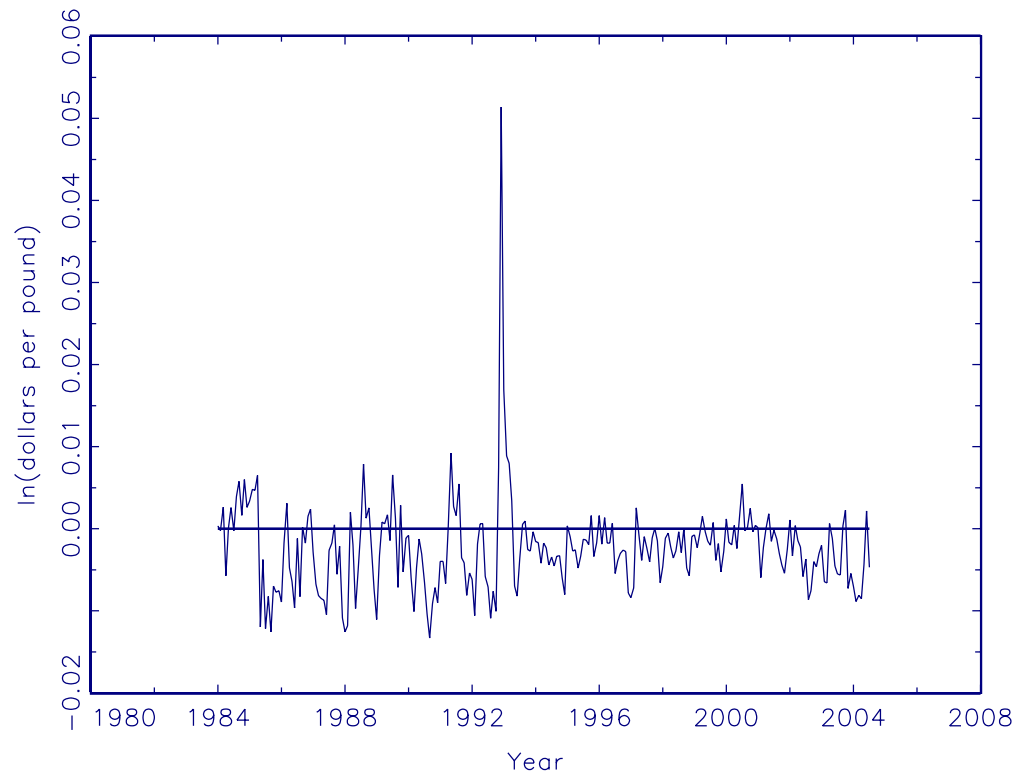
$$E\left(p_t \mid y_1, y_2, ..., y_t\right)$$

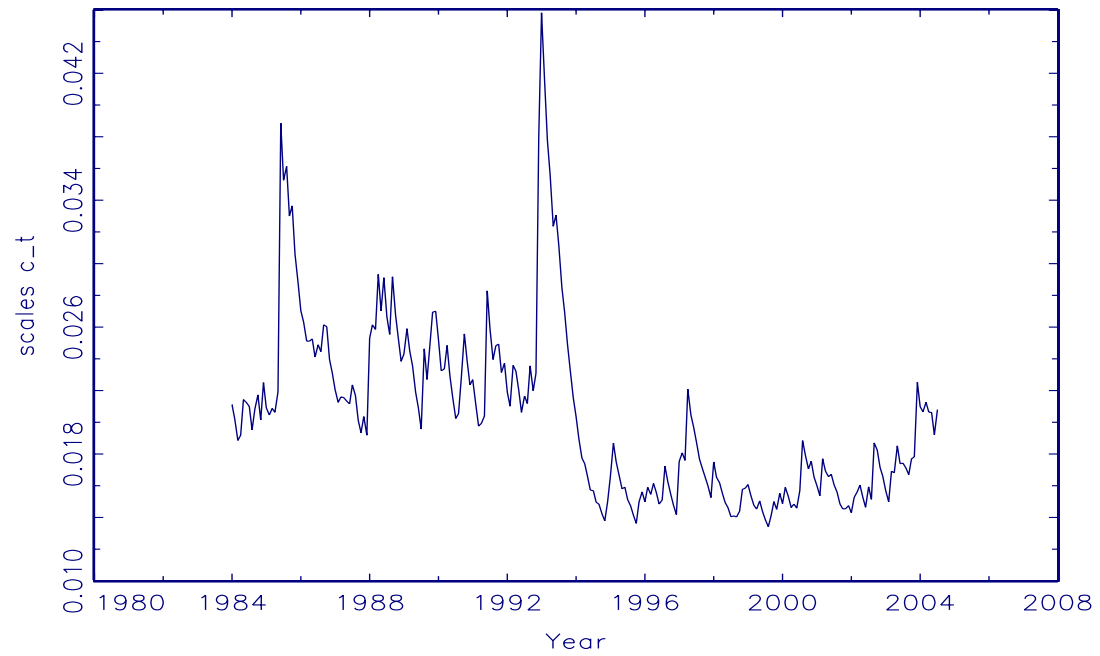Fig.4 Estimates of Conditional Scales c_t from stable Model 3

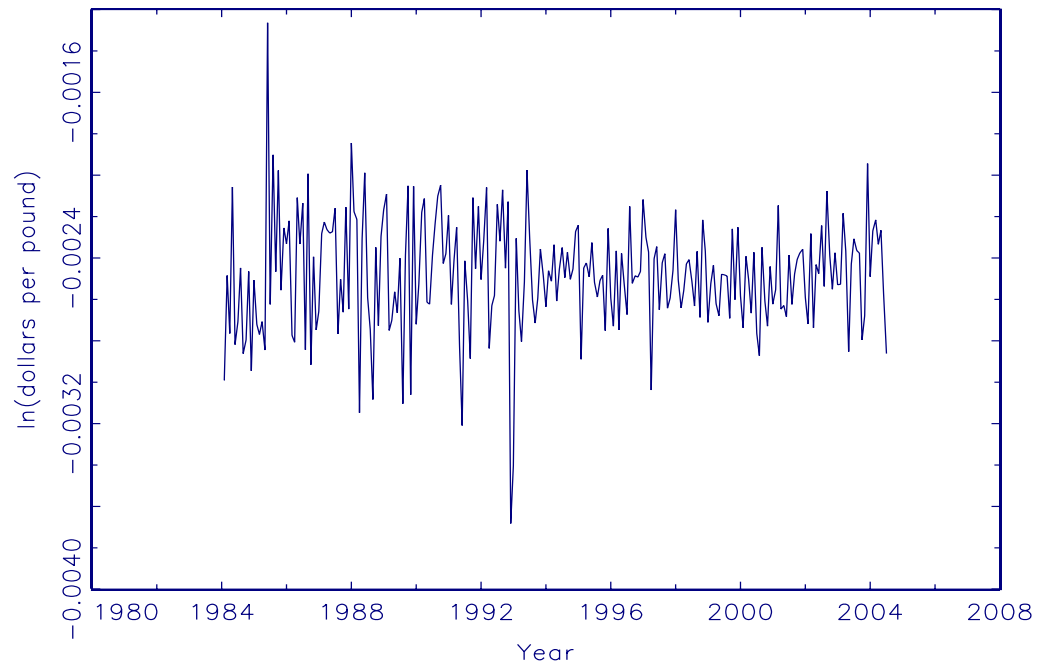Fig.5 Estimates of Risk Premium from Gaussian Model 1
$$E\left(p_t \mid y_1, y_2, ..., y_t\right)$$

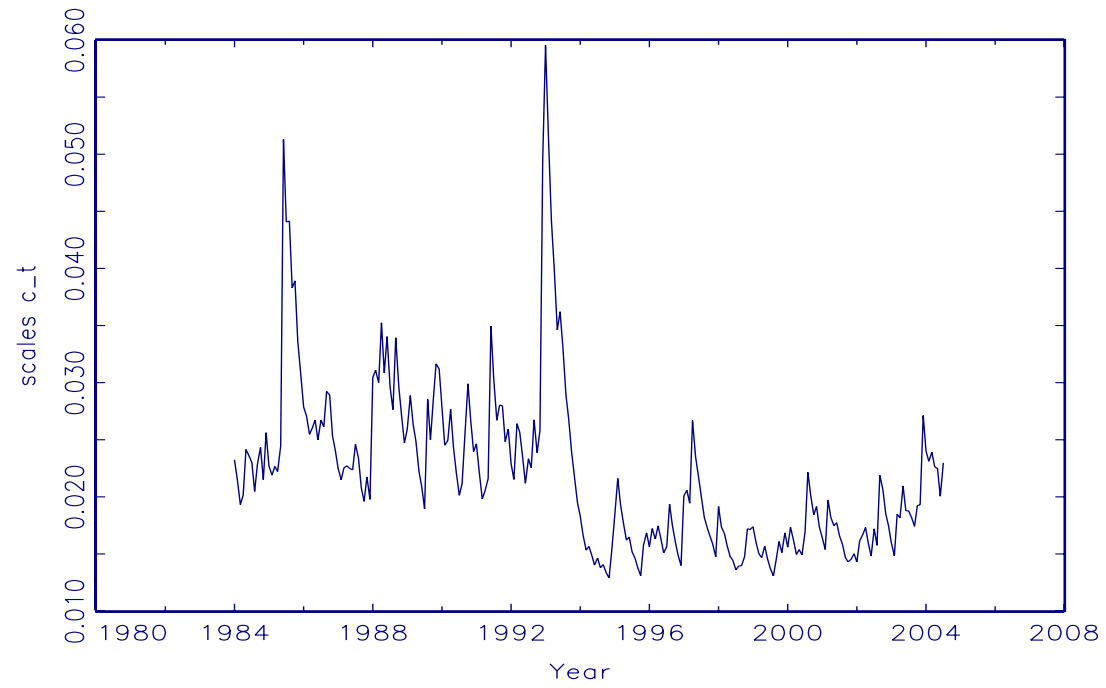Fig.6 Estimates of Conditional Scales c_t from Gaussian Model 3

45

Fig.7 Estimates of Risk Premium from Regression Model 1
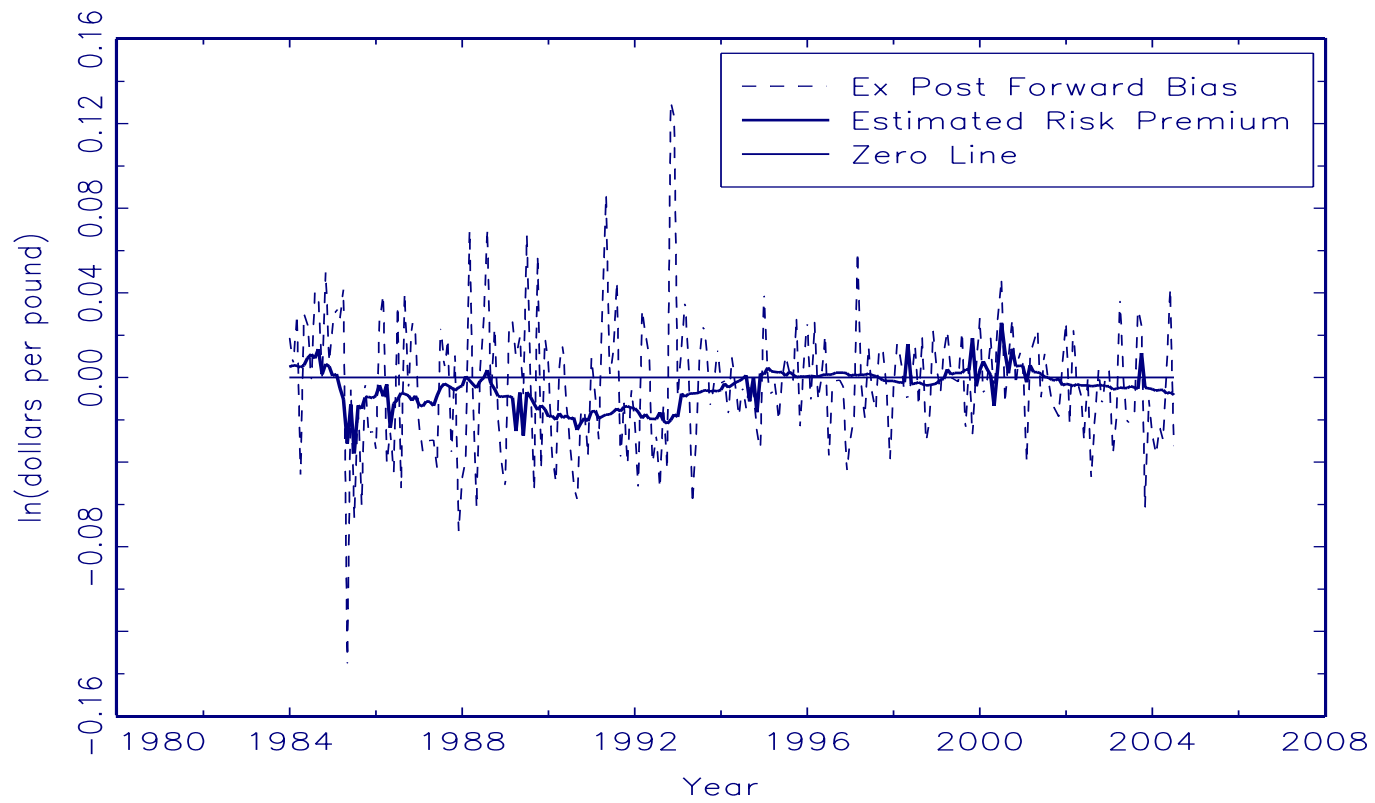
$$E\left(p_t \mid f_t^{t+1} - s_t\right)$$

Fig.8 Estimates of Risk Premium from Regression Model 1
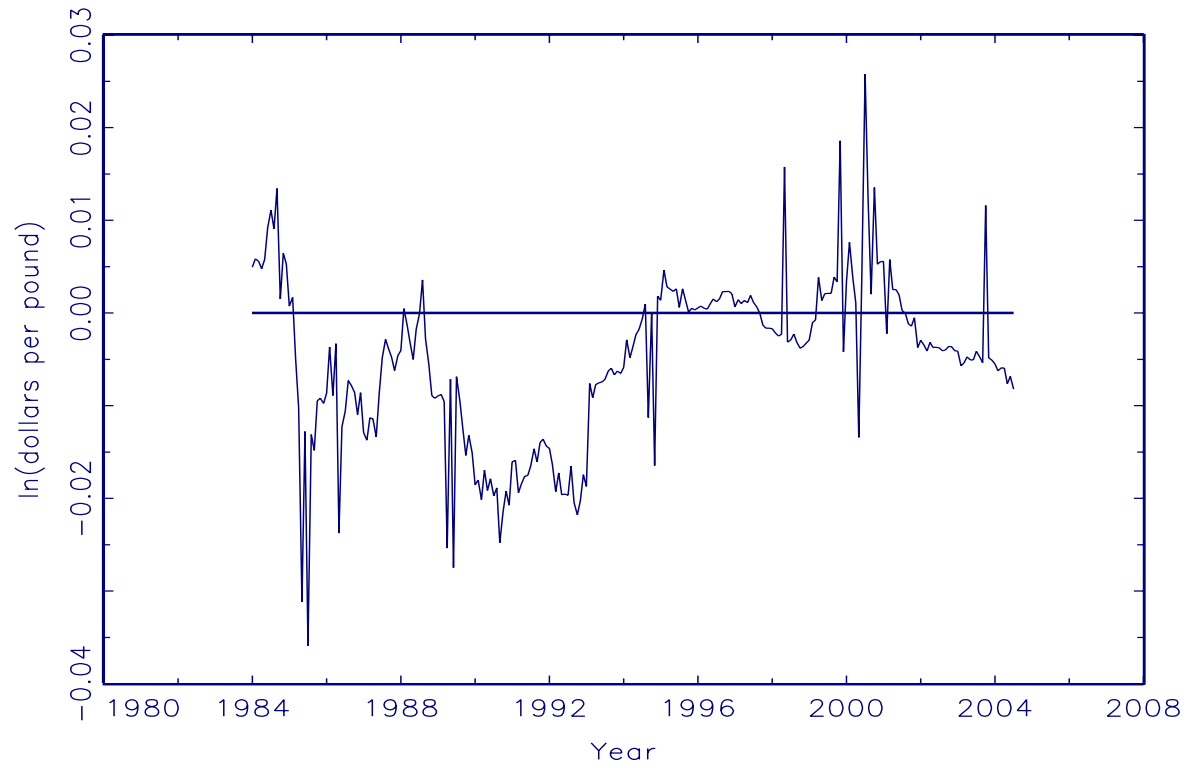$$E\left(p_t \mid f_t^{t+1} - s_t\right)$$

Fig.9 Estimates of Conditional Scales c_t from Regression Model 1