Tests of the risk premium on foreign currency futures implied by the intertemporal asset pricing theory

RAMAZAN GENCAY
Department of Economics, University of Windsor, Windsor, Ontario, Canada, N9B 3P4

In earlier literature on futures exchange rates, some evidence against the martingale hypothesis is not accompanied by an explicit alternative hypothesis derived from the asset pricing theory. This paper utilizes the asset pricing theory to show that the rate of growth of futures price is a function of the rate of growth of the per capita consumption, the logarithm of the gross risk-free rate of return and the time varying risk premium. The form of risk premium implied by the asset pricing theory is examined and its statistical significance is tested with the British Pound and German Deutschmark futures. Consequently, the specification of the test equation implied by the asset pricing theory is tested against ad hoc test equations of the martingale hypothesis. For both currency futures, there is strong evidence against the ad hoc test equations of the martingale hypothesis.

I. INTRODUCTION

The testing of the existence of a risk premium in futures asset prices has been a major focus of attention in empirical financial economics. Samuelson (1965) originally argued that futures prices in an efficient market would be a martingale. A martingale hypothesis postulates

\[ f_t = E_t(f_{t+1}) \] (1)

where \( f_t \) is today's futures price, \( f_{t+1} \) is tomorrow's futures price and \( E_t \) is the expectation conditioned on the information available at time \( t \), \( I_t \). The martingale hypothesis states that today's futures price is the best guess for predicting tomorrow's futures price, given the information available at time \( t \).

On the other hand, the asset pricing paradigm predicts that the martingale hypothesis does not hold due to the existence of a risk premium which separates today's price from its expected price tomorrow:

\[ f_t = E_t(f_{t+1}) + R(z_{t+1} | I_t) \] (2)

where \( R \) is a risk premium function and \( z_{t+1} \) is a vector of relevant economic and financial variables at \( t + 1 \). Under rational expectations, \( f_{t+1} = E_t(f_{t+1}) + e_{t+1} \) and \( e_{t+1} \) is the forecast error which is orthogonal to information set at \( t \).

Hence Equation 2 yields

\[ f_{t+1} - f_t = -R(z_{t+1} | I_t) + e_{t+1} \] (3)

Therefore the two hypotheses are:

\[ H_0: R(z_{t+1} | I_t) = 0 \]
\[ H_1: R(z_{t+1} | I_t) \neq 0. \]

Evidence against the martingale hypothesis with currency futures has developed along two different paths. The first path is followed by McCurdy and Morgan (1987, 1988), Hodrick and Srivastava (1987) and Taylor (1986). McCurdy and Morgan (1987) investigated the martingale hypothesis for futures prices for five currencies with daily and weekly data. Although they found evidence against the martingale hypothesis for all currencies with the daily data, the martingale hypothesis was rejected for only Deutschmark with weekly data. McCurdy and Morgan (1988) examined Deutschmark futures price data in a different time period. They rejected the martingale hypothesis with daily data, but retained with weekly data. Hodrick and Srivastava (1987) rejected the martingale hypothesis with daily exchange rate futures but reported that the rejection is much weaker with
monthly data. The shortcoming of the testing procedure\(^1\) employed in these papers is that the asset pricing theory (e.g. Equation 2) is not utilized to derive a test equation for the alternative hypothesis. One of the main reasons is that some of the arguments of function \(R(z, e|I_e)\) are unobservable. The shortcoming of this approach is that the testing does not indicate the potential sources of rejection as it does not utilize a specific equilibrium model as an alternative hypothesis. Without such a link it is not possible to empirically distinguish between an inefficient market and a time-varying risk premium.

The second path is to construct such a link by deriving a test equation from an equilibrium model.\(^2\) Extensions of intertemporal asset pricing models to foreign currencies are done by Hodrick (1981), Stulz (1981), Lucas (1982) and Hodrick and Srivastava (1984). Domowitz and Hakko (1985) utilize the model of Lucas (1982) which is extended by Hodrick and Srivastava (1984) to derive a test equation with time-varying risk premium that is a function of the conditional variance of domestic and foreign currency. Their results indicate some evidence of a non-zero risk premium for monthly currencies. In two recent papers McCurdy and Morgan (1992a, 1992b) utilize the rational expectations general equilibrium model of Richard and Sundaresan (1981) to derive test equations for foreign exchange futures. With nonexpected utility explanations of prices and under the assumption that the first differences of the logs of consumption and wealth are jointly normally distributed, McCurdy and Morgan (1992a) express the expected logarithm of real total return of an asset as a linear function of the covariances with the logarithm of real total return of wealth and the change in the logarithm of consumption. By constructing a benchmark portfolio for the consumption component and using a world index for the wealth component, they design a test equation which is linear in mean and estimate it by autoregressive conditional heteroscedasticity\(^3\) (ARCH) of Engle (1982). They find no evidence of risk premia when the price of covariance risk is constant, but when the price of the covariance risk is replaced by the conditional expectation of excess returns from the benchmark portfolios and their conditional variances, they obtain evidence of risk premia. McCurdy and Morgan (1992b) measure the covariance risk relative to excess returns from a broadly diversified international portfolio of equities and find strong evidence of risk premia with five weekly currencies.

Hodrick and Srivastava (1987) extended the two-country intertemporal asset pricing model of Lucas (1982) to foreign currency futures. Their development, however, does not discuss how money is formally introduced into the model in order that it will not be an asset whose rate of return is dominated. In the present work, a Svensson’s (1985) style cash-in-advance economy with currency futures markets is constructed which is an extension of Richard and Sundaresan’s (1981) intertemporal asset pricing model with money introduced as a cash-in-advance constraint. It is shown that the rate of growth in a futures price is a function of the rate of growth of the per capita consumption and a time varying risk premium. The form of the time varying risk premium is not parametrically constrained by the theory. The time varying risk premium function is estimated by non-parametric kernel estimation as suggested by Pagan and Ullah (1988). Non-parametric estimation techniques have been successful in applications to models with risk terms. Pagan and Hong (1989) estimated an equity premium reflected in the variance of stock returns by non-parametric kernel and flexible Fourier techniques and found these techniques preferable to parametric ones. Pagan and Schwert (1990) applied these techniques to alternative models of conditional stock volatility. Diebold and Nason (1990) estimated the conditional mean functions of ten major nominal dollar spot rates by nearest-neighbour technique, but did not find any out-of-sample improvement over a simple random walk model. McCurdy and Stengos (1992) obtain similar results by showing that the parametric estimator of the risk premia has more out-of-sample forecastability\(^4\) than does the kernel estimator. Pagan and Ullah (1988) provide an integrated approach to the estimation of models with risk terms.

Section II is on the theoretical model leading to the test equation implied by the asset pricing theory. Section III details the modelling strategy and the empirical findings. Conclusion follows thereafter.

II. THE MODEL

A representative firm in a closed economy produces a stochastic nonstorable output \(\{y_t\}\). The random supply of

---

\(^1\)This testing procedure mainly involves construction of an ad hoc alternative hypothesis from the null model. The test results indicate whether the martingale hypothesis holds or does not hold for this particular alternative hypothesis. However, if an ad hoc test equation leads to the rejection of the null hypothesis, the researcher cannot assess what this means for the risk premium implied by an asset pricing theory. Consequently, the rejection of the null hypothesis with an ad hoc test equation does not necessarily imply that the risk premium implied by an asset pricing theory is statistically significant or vice versa.

\(^2\)The first application of a single period capital asset pricing model (CAPM) to commodity futures was by Dusak (1973), who found no evidence of risk premium. Black (1976) derived testable alternative specifications of futures commodity contracts with single period CAPM.

\(^3\)An excellent survey of the application of the ARCH models to finance is presented in Bollerslev, Chou and Kroner (1992).

\(^4\)This evidence supports the findings of Hsieh (1989) that the temporal dependence in exchange rates is not in the conditional mean but in the conditional variance.
Risk premium on foreign currency futures

The cost of \( n_{t-1} r_{t-1} \), the number of futures contracts, is zero. When the agent initiates \( n_{t-1} r_{t-1} \) dollars of futures contracts he also buys \( n_{t-1} f_{t-1} \) dollar amount of one period discount bonds. \( f_{t-1} \) is the time \( t-1 \) price of a futures contract for delivery at time \( t \), and is in units of money. Therefore, the consumer's cash flow due to the involvement in the futures market is given as

\[
\text{Cash Flow at } t-1 = -n_{t-1} f_{t-1} + n_{t-1} r_{t-1} (f_{t} - f_{t-1}) = n_{t-1} r_{t-1} f_{t}
\]

where \( n_{t-1} r_{t-1} (f_{t} - f_{t-1}) \) is the futures payoff received or paid at time \( t \) and \( f_{t-1} n_{t-1} r_{t-1} \) is the gross nominal return on discount bond holdings received at time \( t \). The no-arbitrage condition is that the present value of cash flow at time \( t \) is equal to the cash flow at \( t-1 \), \( f_{t-1} = PV(r_{t-1} f_{t}) \) where \( PV \) is the present value operator. This no-arbitrage condition is Proposition 2 of Cox, Ingersoll and Ross (1981). To derive a closed form solution for the present value operator we incorporate this condition into the budget constraint of the consumer. It is assumed that futures payoff \( n_{t-1} r_{t-1} (f_{t} - f_{t-1}) \) and the gross nominal return on bond holdings \( f_{t-1} n_{t-1} r_{t-1} \) are received after the goods market is closed. Hence, the timing of transactions in goods and asset markets is as in Svensson (1983).

The payoffs received from the holdings of futures contracts and the holdings of discount bonds, together with the leftover money balances from the purchase of the consumption good, \( (\pi m_t - c_t) \), are used to buy bonds, initiate futures contracts and determine next period cash holdings. The consumer's budget constraint can then be written as

\[
\pi m_{t+1} + \pi f_{t} n_t + \pi b_t \\
\leq (\pi m_t - c_t) + (w_t - 1) \pi m_t + \pi b_{t-1} r_{t-1} \\
+ \pi r_{t-1} n_{t-1} (f_t - f_{t-1}) + \pi f_{t-1} n_{t-1} r_{t-1} + y_t
\]

By rearrangement we obtain

\[
\pi m_{t+1} + \pi f_{t} n_t + \pi b_t \\
\leq (\pi m_t - c_t) + (w_t - 1) \pi m_t + \pi b_{t-1} r_{t-1} \\
+ \pi r_{t-1} n_{t-1} f_t + y_t
\]

where \( m_{t+1} \) is the money holdings to be carried into period \( t+1 \). Each period, the market clearing conditions for goods and asset markets are

\[
c_t = y_t, \quad m_t = m_t^0, \quad m_{t+1} = m_{t+1}^0 = w_t m_t, \quad r_t = 1, \quad n_t = 0.
\]

The representative agent maximizes Equation 5 subject to Equation 6 and Equation 8. I choose to solve the problem

The discussion here covers all financial futures with nominal payoffs such as treasury bills.
by forming a Lagrangean,
\[ J = E_0 \sum_{i=0}^{\infty} \beta^i u(c_i) + \lambda_i [(\pi_i m_i - c_i) \\
+ (w_{i-1}) \pi_i m_i^2 + \pi_i h_{i-1} r_{i-1} + \pi_i r_{i-1} m_i f_t \\
+ y_i - \pi_i m_{i+1} - \pi_i f_t m_i - \pi_i h_i] + \mu_i [(\pi_i m_i - c_i)] \] (9)

The first-order necessary conditions are:
\[ c_i : u'(c_i) = \lambda_i + \mu_i \] (10)
\[ b_i : \pi_i \lambda_i = \beta E_i \pi_{i+1} \lambda_{i+1} + \mu_{i+1} \] (11)
\[ m_{i+1} : \pi_i \lambda_i = \beta E_i \pi_{i+1} (\lambda_{i+1} + \mu_{i+1}) \] (12)
\[ n_i : \pi_i \lambda_i f_t = \beta E_i \pi_{i+1} \lambda_{i+1} f_{i+1} + \mu_{i+1} \] (13)
\[ c_i \leq \pi_i m_i, \quad \mu_i \geq 0, \quad \mu_i (\pi_i m_i - c_i) = 0 \] (14)

Equation 10 states that the marginal utility of real wealth and liquidity services of real balances are equal to the marginal utility of consumption. The equilibrium money price of a one-period discount bond is obtained from Equation 11 by rearrangement
\[ r_t = \beta E_t \left( \frac{\lambda_{i+1} \pi_{i+1}}{\lambda_i \pi_i} \right) \] (15)

Equation 12 denotes that when $\mu$ is binding, the marginal utility of nominal wealth has to be equal to the expected discounted nominal value of next period’s consumption. Equation 13 states that the foregone marginal utility of nominal wealth from the purchase of $f_t$ dollars is equal to the expected marginal discounted benefit of nominal wealth of payoff $f_{i+1} r_t$. The futures pricing function obtained from Equation 13 is
\[ f_t = \beta E_t \left( \frac{\lambda_{i+1} \pi_{i+1}}{\lambda_i \pi_i} \right) f_{i+1} r_t. \] (16)

Equation 14 denotes the complementary slackness condition. I will assume that liquidity constraint binds so that $\mu$ is always strictly positive. Under this assumption the present value operator in Equation 15 is not simply the ratio of marginal utilities of consumption in successive periods. Define
\[ k_{i+1} = \beta \left( \frac{\lambda_{i+1} \pi_{i+1}}{\lambda_i \pi_i} \right) r_t, \quad l_{i+1} = \frac{f_{i+1} r_t}{f_t} \]

so that Equation 16 can be written as
\[ 1 = E_t k_{i+1} l_{i+1} \] (17)

We can now state and prove our two main propositions.

**Proposition 1**
Assume that $k_{i+1} l_{i+1} \neq 0$ with probability one. Let
\[ e_{i+1} = \frac{E_t k_{i+1} l_{i+1}}{k_{i+1} l_{i+1}} \]
be a stochastic process and $l_t$ be the smallest $\sigma$-algebra such that $k_t, l_t, e_t, k_{t-1}, l_{t-1}, e_{t-1}, \ldots$ are all measurable. Then for all integrable functions $g$,
\[ g(E_t k_{i+1} l_{i+1}) = E_t g(k_{i+1} l_{i+1} e_{i+1}) \]

**Proof:**
\[ E_t g(k_{i+1} l_{i+1} e_{i+1}) = E_t g(E_t k_{i+1} l_{i+1}) = g(E_t k_{i+1} l_{i+1}) \]

**Proposition 2**
A second-order Taylor expansion of $\log(E_t k_{i+1} l_{i+1})$ around the unconditional mean of $e_{i+1}, \bar{e} = E(e_{i+1})$, is
\[ \log(f_{i+1}) - \log(f_t) = \gamma_0 + \gamma_1 (\log(\lambda_{i+1} \pi_{i+1}) - \log(\lambda_i \pi_i)) \]
\[ + \gamma_2 \log(r_t) + \gamma_3 \sigma_t^2 + v_{i+1} \] (18)

where $\gamma_0 = \frac{\mu-1}{\beta} + \frac{1}{2}(1 - \beta)^2 - \log(\beta) - \log(\bar{e}), \quad \gamma_1 = \gamma_2 = -1, \quad \gamma_3 = \frac{1}{2} \text{ and } v_{i+1} \sim (0, \sigma_t^2).$

**Proof:** See appendix.

Two remarks can be made from Proposition 2. First, an increase in the rate of growth of the marginal utility of real wealth and nominal gross rate of return lead to an equal decrease in the rate of growth of futures prices. Second, the higher the conditional variance of the innovations, the higher will be the increase in the rate of growth of futures prices.

### III. MODELLING STRATEGY AND EMPIRICAL FINDINGS

By specifying $u(c) = \log(c)$ and assuming that the marginal utility of liquidity services for real balances, $\mu = \mu$ is

---

6Svensson (1985) gives sufficient conditions for the existence of a unique stochastic stationary rational expectations solution with finite value for the objective function.

7Equations 15 and 16 together lead to the Theorem 1 of Richard and Sundaresan (1981) which they derived in a rational expectations general equilibrium model in continuous time of a multigood, identical consumer economy with constant stochastic returns to scale production.

8From Equation 17 $k_{i+1} l_{i+1} = 1 + w_{i+1}$ where $w_{i+1}$ is the forecast error and is orthogonal to the information set at time $t$. Then
\[ e_{i+1} = \frac{1}{1 + w_{i+1}}. \]
Risk premium on foreign currency futures

constant\(^9\) for all \(t\), Equation 18 is written as

\[
\log(f_{t+1}) - \log(f_t) = \gamma_0 + \gamma_1 [\log(p_{t+1}c_{t+1}) - \log(p_c)] + \gamma_2 \log(r_t) + \gamma_3 \sigma_t^2 + \epsilon_{t+1}
\]

(19)

where \(\gamma_1 = 1; \gamma_2 = -1; \gamma_3 = \frac{1}{2}(p_c)\) is nominal consumption expenditures at \(t\) and \(\sigma_t^2\) is the time-varying risk premium. The risk premium function is represented by \(E_t\phi_t^2 = \sigma_t^2\) and \(\phi_t^2 = (\psi_t - E_t\psi_t)^2\) where \(\psi_t = \log(f_{t+1}) - \log(f_t)\).

Notice that the change in the purchasing power of money has a direct effect on futures prices via nominal consumption expenditures. Suppose that per capita real consumption in two successive periods is the same, but the purchasing power of the dollar is increased, i.e. \(\pi_{t+1} > \pi_t\). This means that the agent will spend less dollars tomorrow to buy today's consumption level. From Equation 19, this will lead to a decrease in the rate of growth of futures prices.

The estimation problem is that no direct observations on \(\sigma_t^2\) exist, and the theory does not describe how \(\sigma_t^2\) varies with the information at \(t\). If it were possible to construct a series \(\phi_t^2\) which had the strong property,\(^10\) then \(\sigma_t^2\) could be replaced by \(\phi_t^2\) and Equation 19 could be estimated by ordinary least squares (OLS). However, if the constructed series \(\phi_t^2\) has the weak property,\(^11\) then OLS will be inconsistent due to the regressors \(\sigma_t^2\) being observed with error. Pagan and Ullah (1988) propose a solution to this problem by estimating Equation 19 not by OLS, but by instrumental variables (IV) with instruments constructed from \(I_t\). Of course instrument construction has to be done carefully, since the instruments must be as highly correlated with \(\phi_t^2\) as possible. An important aspect of the IV strategy is that it only requires \(m_t = E_t\psi_t\) to be estimated consistently and some estimate of \(\sigma_t^2\).

To construct our proxy series \(\phi_t^2\), we need to know how \(\sigma_t^2\) varies with the information available at \(t\), and what the mapping is between this information and \(\sigma_t^2\). Pagan and Ullah (1988) show that the consistent estimation of the parameter of the model require that \(I_t\) be known, so the choice of the information set has to be wide enough to encompass \(I_t\). In this paper, we assume that econometrician's information set incorporates \(I_t\).

\(m_t\) will normally be a nonlinear function of \(I_t\). A general dependence of \(m_t\) on \(I_t\) can be carried out by non-parametric estimation methods such as the non-parametric kernel regression. With such an estimate of \(m_t\), \(\phi_t^2\) can be constructed by \(\phi_t^2 = (\psi_t - m_t)^2\).

Thus the model which has to be estimated by IV is

\[
\log(f_{t+1}) - \log(f_t) = \gamma_0 + \gamma_1 [\log(p_{t+1}c_{t+1}) - \log(p_c)] + \gamma_2 \log(r_t) + \gamma_3 \phi_t^2 + \gamma_5 (\phi_t^2 - \phi_t^2) + \epsilon_{t+1}
\]

(20)

with constant, \(\log(p_{t+1}c_{t+1}) - \log(p_c)\), \(\log(r_t)\) and \(\phi_t^2\) as instruments.

Non-parametric kernel estimation

Consider an economic model

\[
y = M(x) + u
\]

(21)

where \(M(x)\) is the conditional mean of a random variable \(y\) and \(x\) is a \(q\times1\) vector of conditioning variables. In parametric estimation \(M(x)\) is known, but in non-parametric regression it is an unspecified functional form. The usual kernel estimator \(M(x) = \hat{E}(y|x)\) is written as

\[
\hat{M}(x) = \frac{\hat{s}(x)}{\hat{f}(x)}
\]

(22)

where

\[
\hat{s}(x) = \frac{1}{n[\Omega^{-1}]_{j,j}} \sum^n_{i=1} y_i K_{ij}; \quad \hat{f}(x) = \frac{1}{n[\Omega^{-1}]_{j,j}} \sum^n_{i=1} K_{ij}
\]

(23)

\(K_j = K(\Omega^{-1}(x - x_i))\); \(\Omega = \text{diag}(a_1, a_2, \ldots, a_q)\); \(a_i, i = 1, 2, \ldots, q\) are the smoothing parameters or window widths; \(n\) is the sample size and \(j\)‘s indicate individual observations. Since it is possible to choose the function \(K\) so that it is continuous, the resulting kernel estimator of the density function will also be continuous. In the present paper, the kernel is chosen to be the standard multivariate normal density function. Robinson (1983) derives the asymptotic distribution of the non-parametric kernel estimator of the joint density and of the conditional mean within the time series context.

An important consideration in the literature is the choice of the bandwidth parameter \(a\). Too large a value of \(a\) induces bias and too small a value induces imprecise estimates. Robinson (1983) and Ullah (1988), among others, summarize the conditions that the kernel function and the bandwidth parameter \(a\) will have to satisfy to obtain the asymptotic properties of the regression function estimator. Ullah (1988) suggests setting the window width \(a_t\) to

\[
a_t = s_t n^{-1/(4+\epsilon)}
\]

(24)

\(^9\)In Equation 10 if \(\mu_t = 0\) then \(\lambda_t = u(c_t)\) which implies that the marginal utility of wealth is equal to the marginal utility of consumption. Since the utility function is arbitrary up to any monotonic transformation, marginal utility of wealth is also arbitrary relative to this transformation. Suppose that \(u(c) = \log(c)\) and a monotonic transformation of \(u\) is \(\varphi(u) = u + \mu\varepsilon\) so that \(\varphi(u) > 0\). Then \(\varphi(c) = u(c) + \mu\). By replacing \(u(c)\) by \(\varphi(c)\) in \(\lambda_t = u(c_t)\) we get \(\lambda_t = \varphi(c_t) + \mu\). Therefore, we will assume that \(\mu_t = \mu\) and use \(\lambda_t = \varphi(c_t)\) instead of Equation 10.

\(^10\)A series is said to possess the strong property if \(\phi_t^2(N) \rightarrow \sigma_t^2\) for all \(N \rightarrow \infty\) where \(N\) is not dependent on \(t\).

\(^11\)A series is said to possess the weak property if \(E_t \phi_t^2 = \sigma_t^2\).
where $i = 1, 2, \ldots, q$ and $s_i$ denotes the standard deviation of $x_i$. The asymptotic variance of a kernel regression estimate $M(x)$ is given by

$$V(x) = \frac{1}{n|\Omega|} \frac{\text{Var}(y/x)}{f(x)} \int K^2$$

from which it is straightforward to construct standard errors.

**Data description**

The daily British Pound and German Deutschemark futures exchange rate series are obtained from the Center of Research for Futures Markets of the University of Chicago covering the period July 1972 to December 1986 and Reuters for January 1987 to December 1988. The monthly series are then constructed by choosing the mid-month futures price corresponding to the shortest maturity contracts which are delivered on the third Wednesday of March, June, September and December. The mid-month refers to the 15th day of every month and 14th day of February. If the 15th day is on a weekend, the previous Friday’s price is chosen. The mid-month futures are chosen to match the futures prices with the monthly average consumption. The monthly US nominal consumption expenditure series for nondurables, and the monthly US population series are compiled from the US Department of Commerce, Survey of Current Business. To obtain the per capita consumption, consumption series are normalized by the population series. The nominal gross risk-free rate of return series is constructed from the interest rate on US30-Day Commercial Paper and obtained from CANSIM matrix 2545.

**Empirical findings**

Choosing three lags of $\psi = \log(f_{i+1}) - \log(f_i)$ as the conditioning elements, the conditional mean $m_t$ is estimated by non-parametric kernel by regressing $\psi$ on $\psi_{t-1}$, $\psi_{t-2}$ and $\psi_{t-3}$. Table 1 gives the first ten autocorrelation coefficients of the squared non-parametric residuals along with their $t$-statistics. There is no evidence of ARCH effects observed which makes the further lags of $\psi$, in the conditioning set unnecessary. The proxy $\phi_t^{2*}$ for $\sigma_t^2$ is generated by $\phi_t^{2*} = (\psi_t - m_t)^2$ where $m_t$ is the non-parametric conditional mean estimate. An estimate of $\sigma_t^2$, $\delta_t^2$ is needed as an instrument for $\sigma_t^2$. $\delta_t^2$ is calculated non-parametrically by calculating $\hat{\delta}_t^2 = E(\phi_t^2)$. IV estimation indicates the presence of a first order autocorrelation in British Pound futures. Accordingly, the lagged rate of growth of British Pound is included as a regressor in Equation 20. Diagnostic tests applied after IV estimation give test statistic values of 0.331 and 0.572 for first order autocorrelation and 2.456 and 7.976 for heteroscedasticity. Thus there is some evidence of heteroscedasticity in the Deutschemark residuals. The results for the IV procedure with Equation 20 are presented in Table 2 with a White (1980) adjustment to the calculated $t$-statistics for the Deutschemark futures.

For both currencies it is clear that the martingale hypothesis is rejected. The risk premium for the British Pound futures and the intercept and the risk-free rate for the Deutschemark futures are statistically significant at the 5% level. This raises the issue about which model specification characterizes the data better, an ad hoc test equation or the test equation derived from the asset pricing model. This can easily be demonstrated by a non-nested testing methodology such as the J test of Davidson and MacKinnon (1981). Two different specifications for the ad hoc test equation are chosen. The first specification is

$$\log(f_{i+1}) - \log(f_i) = a_0 + a_2[\log(f_{i-1}) - \log(f_{i-2})] + \epsilon_i$$

(26)

The parameter estimates of this model are presented in Table 3. Equation 26 is very similar to the main test equation of McCurdy and Morgan (1988), which they estimate

<table>
<thead>
<tr>
<th>Table 1. Autocorrelation function of squared non-parametric residuals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lag</td>
</tr>
<tr>
<td>------</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>5</td>
</tr>
<tr>
<td>6</td>
</tr>
</tbody>
</table>

The estimates of the autocorrelation coefficients (a.c.f.) are obtained by regressing $\delta_t^2$ against a constant and $\delta_{i,j}^2 (j = 1, 2, \ldots, 10)$ one at a time. $t$ is the absolute value of a $t$ statistic.

13For British Pound the estimated model is

$$\log(f_{i+1}) - \log(f_i) = \gamma_0 + \gamma_1[\log(p_{t+1}c_{t+1}) - \log(p_{t}c_{t})] + \gamma_2[\log(f_{t-1}) - \log(f_{t-2})] + \gamma_3[\log(r_t) + \gamma_4\phi_t^{2*}]$$

with constant, $\log(p_{t+1}c_{t+1}) - \log(p_{t}c_{t})$, $\log(r_t)$, and $\delta_t^2$ as instruments.

14The calculated test statistics are the Breusch (1979) and Godfrey (1978) procedure and should be referred to as $\chi^2(1)$.

15This is the Koenker (1981) procedure and should be referred to as $\chi^2(3)$. 

R. Gencay
Risk premium on foreign currency futures

Table 2. Instrumental variables estimation of
\[ \log(f_{t+1}) - \log(f_t) = \gamma_0 + \gamma_1(\log(p_{t+1,c_t+1}) - \log(p_{c_t})) + \gamma_2(\log(f_t) - \log(f_{t-1})) + \gamma_3 \log(r_t) + \gamma_4 \phi_t^2 + \gamma_5(\phi_t^2 - \phi_{t-1}^2) + \epsilon_{t+1} \]

<table>
<thead>
<tr>
<th>British Pound</th>
<th>German Deutschmark</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma_0$</td>
<td>0.0035 (0.4455)</td>
</tr>
<tr>
<td>$\gamma_1$</td>
<td>-0.2709 (0.8642)</td>
</tr>
<tr>
<td>$\gamma_2$</td>
<td>0.2505 (1.0633)</td>
</tr>
<tr>
<td>$\gamma_3$</td>
<td>-0.1236 (1.5463)</td>
</tr>
<tr>
<td>$\gamma_4$</td>
<td>8.2097 (3.0223)</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.0843</td>
</tr>
</tbody>
</table>

The numbers in the first parentheses are the absolute values of the t-statistics and the numbers in the second parentheses are heteroscedastic consistent versions. For the British Pound, the constant, \( \log(p_{t+1,c_t+1}) - \log(p_{c_t}) \), \( \log(p_{c_t}) - \log(p_{t-1,c_t-1}) \), \( \log(r_t) \), and \( \phi_t^2 \) are used as instruments. For the Deutschmark, the constant, \( \log(p_{t+1,c_t+1}) - \log(p_{c_t}) \), \( \log(r_t) \), and \( \phi_t^2 \) are used as instruments.

Table 3. Ordinary least squares estimation of
\[ \log(f_{t+1}) - \log(f_t) = \alpha_0 + \alpha_1(\log(f_t) - \log(f_{t-1})) + \alpha_2(\log(f_{t-1}) - \log(f_{t-2})) + \epsilon_t \]

<table>
<thead>
<tr>
<th>British Pound</th>
<th>German Deutschmark</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td>Model 2</td>
</tr>
<tr>
<td>$\alpha_0$</td>
<td>-0.0015 (0.6592)</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>0.1527 (2.1137)</td>
</tr>
<tr>
<td>$\alpha_2$</td>
<td>-0.0156 (0.2151)</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.0002</td>
</tr>
</tbody>
</table>

The absolute values of t-statistics are in parentheses.

by ARCH (see Engle, 1982; Bollerslev, 1986; and Engle and Bollerslev, 1986). The presence of an ARCH(1) and ARCH(2) process in the residuals of Equation 26 is tested by using monthly data. The test statistics obtained were 0.032 and 0.239 for the British Pound and 0.116 and 0.118 for the Deutschmark which are statistically insignificant.\(^\text{15}\) Assuming that Equation 26 is specified under $H_0$, the artificial compound model is written by

\[ \log(f_{t+1}) - \log(f_t) = \alpha_0 + \alpha_2[\log(f_{t-1}) - \log(f_{t-2})] + \alpha_3 \Psi_t + \epsilon_t \]

(27)

where variable $\Psi_t$ in Equation 27 refers to the fit obtained from the IV estimation\(^\text{16}\) of Equation 20. The logic underlying the $J$ test is that, because the hypotheses are non-nested, the truth of $H_0$ implies that $H_1$ is false. If $H_0$ is the true model then the true value of $\delta_3$ in Equation 27 is zero. Thus, $H_0$ is accepted or rejected based on the significance of $\delta_3$: if $\delta_3$ is insignificant, $H_0$ is accepted, while if $\delta_3$ is significant, $H_0$ is rejected. Davidson and MacKinnon (1981) show that a standard $t$ test can be used to determine if $\delta_3$ is significantly different from zero. A $t$ test on the coefficient of $\Psi_t$ is presented\(^\text{17}\) in the first row of Table 4(a). The results indicate that for both currencies the ad hoc model is rejected in favour of the asset pricing model. The test of the asset pricing model against the ad hoc model leads to retaining the asset pricing model. Therefore, the asset pricing model comes out as a better specification from the non-nested testing.

The other specification as the ad hoc test equation\(^\text{18}\) is

\[ \log(f_{t+1}) - \log(f_t) = \alpha_0 + \alpha_1[\log(f_t) - \log(f_{t-1})] + \alpha_2[\log(f_{t-1}) - \log(f_{t-2})] + \epsilon_t \]

(28)

The calculated $J$ test results are presented in the second row of Table 4(a). The evidence reveals that the test equation derived from the intertemporal asset pricing theory is a better specification of the German Deutschmark and British Pound futures.

The data period covers approximately 16 years. The length of the sample horizon may raise the issue of the sensitivity of the results over shorter periods. This is done by splitting the data on three five year sub-periods. For each sub-period the $J$ test is calculated and the results are presented in Table 4(b), (c) and (d). The results here confirm the

---

\(^{15}\)This is consistent with Diebold (1988) and Baur and Bollerslev (1989) that ARCH effects tend to weaken with less frequently sampled data.

\(^{16}\)The test statistic for an ARCH(1) process was obtained as is $nR^2$ from the artificial regression $I^2 = \alpha_0 + \alpha_1 I^2_{t-1} + \cdots + \alpha_n I^2_{t-n} + \epsilon_t$, where $I^2_t$ are the squared residuals from the AR fit. The test statistic is distributed as $\chi^2$.

\(^{17}\)When Equation 20 is the null model then the modified version of the $J$ test in MacKinnon, White and Davidson (1983) for the two-stage least squares regression is used.

\(^{18}\)The $J$ test is not the only non-nested test; others include the Cox (1961) test and the Fisher–McAleer (1981) test. In an evaluation of several non-nested tests, Godfrey and Pesaran (1983) find that in small samples the $J$ test has a tendency to reject a true null model too often. This is particular so when the true model fits the data so poorly and when the number of regressors in the false model exceeds that in the true model. The Fisher–McAleer test was also tried. The results are very comparable to the $J$ test results.

\(^{19}\)The parameter estimates of these equations are presented in Table 3.
Table 4. Nonnested comparison of the asset pricing model (AM) with an ad hoc model (HM)

<table>
<thead>
<tr>
<th></th>
<th>British Pound (AM versus HM)</th>
<th>German Deutschemark (AM versus HM)</th>
<th>German Deutschemark (HM versus AM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) Sample period: July 1972–December 1988</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 1</td>
<td>0.3086 (0.2680)</td>
<td>0.0438 (0.0435)</td>
<td>4.3747 (2.7529)</td>
</tr>
<tr>
<td></td>
<td>5.9205 (3.7424)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 2</td>
<td>0.3082 (0.2675)</td>
<td>0.5756 (0.5562)</td>
<td>4.2941 (2.6602)</td>
</tr>
<tr>
<td></td>
<td>5.4533 (3.1341)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(b) Sample period: July 1972–June 1977</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 1</td>
<td>0.8562 (0.8213)</td>
<td>0.2145 (0.1819)</td>
<td>3.0631 (2.8345)</td>
</tr>
<tr>
<td></td>
<td>3.8452 (3.6231)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 2</td>
<td>0.8412 (0.8311)</td>
<td>0.3451 (0.3271)</td>
<td>3.0945 (3.0412)</td>
</tr>
<tr>
<td></td>
<td>3.5412 (3.4567)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(c) Sample period: July 1972–June 1982</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 1</td>
<td>0.3804 (0.3446)</td>
<td>0.6539 (0.5921)</td>
<td>3.5847 (3.3100)</td>
</tr>
<tr>
<td></td>
<td>3.2131 (3.1192)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 2</td>
<td>0.3847 (0.3404)</td>
<td>0.5834 (0.5380)</td>
<td>3.4178 (3.4178)</td>
</tr>
<tr>
<td></td>
<td>3.3256 (3.0123)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(d) Sample period: July 1982–June 1988</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 1</td>
<td>0.4512 (0.4449)</td>
<td>0.3214 (0.2903)</td>
<td>3.3154 (3.2234)</td>
</tr>
<tr>
<td></td>
<td>3.2657 (3.2145)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 2</td>
<td>0.4783 (0.4312)</td>
<td>0.5834 (0.5312)</td>
<td>3.4178 (3.2178)</td>
</tr>
<tr>
<td></td>
<td>3.3256 (3.1567)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

AM and HM refer to the test equations derived from the asset pricing theory and the ad hoc testing methods, respectively. The numbers in parentheses are the heteroscedasticity consistent versions of the J tests. Models 1 and 2 are:

1. \[ \log(f_{t+1}) - \log(f_t) = \alpha_0 + \alpha_2(\log(f_{t-1}) - \log(f_{t-2})) + \epsilon_t, \]
2. \[ \log(f_{t+1}) - \log(f_t) = \alpha_0 + \alpha_1(\log(f_t) - \log(f_{t-1})) + \alpha_2(\log(f_{t-1}) - \log(f_{t-2})) + \epsilon_t, \]

earlier results obtained from the full sample that the ad hoc test equations are rejected in favour of the test equation of the asset pricing theory.

IV. CONCLUSIONS

Some evidence in earlier literature with foreign currency futures rejected the martingale hypothesis without an explicit model of an alternative hypothesis derived from the asset pricing theory (e.g. Equation 2). One of the main reasons is that some of the arguments of function \( R_{\epsilon, I_f} \) are unobservable. The limitation of this approach is that the testing does not indicate the potential sources of rejection, as it does not utilize a specific equilibrium model as an alternative hypothesis. Without such a link, it is not possible to empirically distinguish between an inefficient market and a time-varying risk premium.

This paper has utilized the intertemporal asset pricing theory to show that the rate of growth in a futures price is a function of the rate of growth of the per capita consumption, logarithm of the nominal gross risk-free rate of return and a time varying risk premium. The form of risk premium implied by the asset pricing theory is examined and its statistical significance is tested with British Pound and German Deutschemark futures. The evidence strongly rejects the ad hoc test equations of the martingale hypothesis in favour of the test equation derived from the asset pricing theory.

The results of this paper indicate that ad hoc test equations of the martingale hypothesis are misspecified. One source of the misspecification in ad hoc test equations is the omission of the consumption growth from the test equation. The omission of the consumption data is justified on the grounds that it is not available in higher frequencies and it is noisy. Recently, Campbell (1993) suggests a loglinear approximation to the budget constraint to substitute consumption out of a standard intertemporal asset pricing model. Campbell's results show that the innovation to log consumption equals the innovation to the log return on the market portfolio and the revisions in the discounted value of future market returns. It is possible to extend Campbell...
Risk premium on foreign currency futures

(1993) results to the pricing of the futures exchange rates such that innovation in log consumption is expressed as a function of the innovations in the market return. Then, the test equation of the futures asset pricing theory would provide evidence to distinguish between an inefficient market and a time varying risk premium.

ACKNOWLEDGEMENTS

I am grateful to Dee Dechert, Ieuuan Morgan and an anonymous referee for detailed comments on an earlier version and Ieuuan Morgan and Tom McCurdy for providing the futures exchange rate data. An earlier version of this paper is circulated under the title of 'On the Mapping of the Risk Premium in Foreign Exchange Futures'. I would like to thank the Social Sciences and Humanities Research Council of Canada and the Natural Sciences and Engineering Council of Canada for financial support.

REFERENCES


APPENDIX

Proposition 2

A second-order Taylor expansion of \( \log(E_i k_{i+1} l_{i+1}) \) around the unconditional mean of \( e_{i+1} = (e_{i+1}) \), is

\[
\log(f_{i+1}) - \log(f_i) = \gamma_0 + \gamma_1 \left[ \log(\lambda_{i+1} \pi_{i+1}) - \log(\lambda_i \pi_i) \right] + \gamma_2 \log(r_i) + \gamma_3 \sigma_i^2 + \nu_{i+1}
\]

where \( \gamma_0 = \frac{\sigma_i^2}{2 \sigma_i^2} + \left(\frac{3}{2} - 1\right)^2 - \log(\beta) - \log(\bar{\varepsilon}) \), \( \gamma_1 = \gamma_2 = -1 \), \( \gamma_3 = \frac{1}{2} \) and \( \nu_{i+1} \sim (0, \sigma_{e_i}^2) \).

Proof:

By applying Proposition 1 and choosing \( g(x) = \log(x) \)

Equation 17 is written as

\[
0 = \log(E_i k_{i+1} l_{i+1}) = E_i \log(k_{i+1} l_{i+1} e_{i+1})
\]

Consider a linear expansion of \( E_i \log(k_{i+1} l_{i+1} e_{i+1}) \) around the unconditional mean of \( e_{i+1} = (e_{i+1}) \),

\[
E_i \log(k_{i+1} l_{i+1} e_{i+1}) = E_i \log(k_{i+1} l_{i+1} (e_{i+1} - \bar{e})) + \frac{1}{\bar{e}} E_i (k_{i+1} l_{i+1} (e_{i+1} - \bar{e}))
\]

\[
- \frac{1}{2 \bar{e}} E_i (k_{i+1} l_{i+1} (e_{i+1} - \bar{e}))^2.
\]

From Proposition 1, \( E_i k_{i+1} l_{i+1} e_{i+1} = E_i k_{i+1} l_{i+1} \) and from Equation 17, \( E_i k_{i+1} l_{i+1} = 1 \), hence

\[
- \frac{1}{\bar{e}} E_i (k_{i+1} l_{i+1} (e_{i+1} - \bar{e})) = (1 - \bar{e}) / \bar{e},
\]

and

\[
E_i \log(k_{i+1} l_{i+1} \bar{e}) = E_i \log(k_{i+1} l_{i+1}) + \log(\bar{e}).
\]

Also, notice that the third component of Equation 29 is written as

\[
- \frac{1}{2 \bar{e}} E_i (k_{i+1} l_{i+1} (e_{i+1} - \bar{e}))^2
\]

\[
- \frac{1}{2 \bar{e}} \{E_i (1 - k_{i+1} l_{i+1} \bar{e})^2 \}
\]

since \( k_{i+1} l_{i+1} e_{i+1} = E_i k_{i+1} l_{i+1} \). By factoring out \( \bar{e} \) and adding and subtracting one, we obtain

\[
- \frac{1}{2} \left\{ E_i \left(1 - k_{i+1} l_{i+1} + \frac{1}{\bar{e}} - 1 \right)^2 \right\}
\]

By defining \( w_{i+1} = k_{i+1} l_{i+1} - 1 \), we get

\[
- \frac{1}{2} \left\{ E_i \left(1 - w_{i+1} + \frac{1}{\bar{e}} - 1 \right)^2 \right\}
\]

\[
- \frac{1}{2} \left\{ E_i w_{i+1}^2 + \left(\frac{1}{\bar{e}} - 1 \right)^2 \right\}
\]

(32)

where \( E_i w_{i+1} = 0 \). By combining Equations 30, 31 and 32, we obtain

\[
0 = E_i \log(k_{i+1} l_{i+1}) + \log(\bar{e}) + \frac{1 - \bar{e}}{\bar{e}}
\]

\[
- \frac{1}{2} \left\{ \sigma_i^2 + \left(\frac{1}{\bar{e}} - 1 \right)^2 \right\}
\]

(33)

where \( w_{i+1} \sim (0, \sigma_i^2) \). \( E_i \log(k_{i+1} l_{i+1}) \) is written as

\[
E_i \log(k_{i+1} l_{i+1}) = \log(\beta) + \log(r_i) + \log(\lambda_{i+1} \pi_{i+1})
\]

\[
- \log(\lambda_i \pi_i) + \log(f_{i+1})
\]

(34)

By rearranging terms in Equations 33 and 34 we obtain

\[
\log(f_{i+1}) - \log(f_i) = \gamma_0 + \gamma_1 \left[ \log(\lambda_{i+1} \pi_{i+1}) - \log(\lambda_i \pi_i) \right] + \gamma_2 \log(r_i) + \gamma_3 \sigma_i^2 + \nu_{i+1}
\]

where \( \gamma_0 = \frac{\sigma_i^2}{2 \sigma_i^2} + \left(\frac{3}{2} - 1\right)^2 - \log(\beta) - \log(\bar{e}) \), \( \gamma_1 = \gamma_2 = -1 \), \( \gamma_3 = \frac{1}{2} \) and \( \nu_{i+1} \sim (0, \sigma_{e_i}^2) \).