NBER WORKING PAPER SERIES

EXPECTATIONS HYPOTHESES TESTS

Geert Bekaert
Robert J. Hodrick

Working Paper 7609
http://www.nber.org/papers/w7609

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 2000

We are grateful to David Ng and Paul Sengmueller of Columbia University and especially Joe Chen of Stanford University for research assistance. Geert Bekaert gratefully acknowledges financial support from the National Science Foundation. The views expressed herein are those of the authors and are not necessarily those of the National Bureau of Economic Research.

Expectations Hypotheses Tests
Geert Bekaert and Robert J. Hodrick
NBER Working Paper No. 7609
March 2000
JEL No. E4, F3, C5

ABSTRACT

We investigate the Expectations Hypotheses of the term structure of interest rates and of the foreign exchange market using vector autoregressive methods for the U.S. dollar, Deutsche mark, and British pound interest rates and exchange rates. In addition to standard Wald tests, we formulate Lagrange Multiplier and Distance Metric tests which require estimation under the non-linear constraints of the null hypotheses. Estimation under the null is achieved by iterating on approximate solutions that require only matrix inversions. We use a bias-corrected, constrained vector autoregression as a data generating process and construct extensive Monte Carlo simulations of the various test statistics under the null hypotheses. Wald tests suffer from severe size distortions and use of the asymptotic critical values results in gross over-rejection of the null. The Lagrange Multiplier tests slightly under-reject the null, and the Distance Metric tests over-reject. Use of the small sample distributions of the different tests leads to a common interpretation of the validity of the Expectations Hypotheses. The evidence against the Expectations Hypotheses for these interest rates and exchange rates is much less strong than under asymptotic inference.

Geert Bekaert
Columbia Business School
3022 Broadway
New York, NY 10027
NBER and Stanford University
gb241@columbia.edu

Robert J. Hodrick
Columbia Business School
3022 Broadway
New York, NY 10027
NBER
rh169@columbia.edu

According to the Expectations Hypothesis, information in current interest rates provides the conditional expectation of future asset prices. The Expectations Hypothesis of the term structure of interest rates (EH-TS) states that the current term spread between a long-term interest rate and a short-term interest rate is the expected value of a weighted average of the expected future changes in the short-term interest rate. This theory, popularized in the writings of Fisher (1930), Keynes (1930), and Hicks (1953), continues to be a way that many economists think about the determination of long-term interest rates. The Expectations Hypothesis in the foreign exchange market (EH-FX) states that the interest-rate differential between two currencies is the conditional expected value of the rate of depreciation of the high interest-rate currency relative to the low interest-rate currency. Again, Fisher (1930) and Keynes (1930) discussed this hypothesis. Because of covered interest arbitrage, the interest differential equals the forward premium, which is the percentage difference between the forward exchange rate and the spot rate. Hence, the EH-FX is equivalent to the Unbiasedness Hypothesis, which is the proposition that the logarithm of the forward exchange rate is an unbiased predictor of the logarithm of the future spot rate. Many economists also currently view the EH-FX as the way that forward exchange rates are determined.

These Expectations Hypotheses (EHs) continue to have adherents because most modern asset pricing theories imply either that expected future interest rates and exchange rates are related to current interest rates directly through the EHs or with the addition to the EHs of risk premiums. If these risk premiums are constant, the EHs can be said to hold because the temporal variation in expected future asset prices drives the variability in current interest rates. If the risk premiums are variable, the EHs will not hold, but the literature has had surprisingly little success generating risk premiums that explain the empirical evidence.

Empirical tests of the EHs are too numerous to enumerate. For the EH-FX, the statistical evidence surveyed by Hodrick (1987), Bekaert and Hodrick (1993), and Engel (1996) strongly rejects the hypothesis. In particular, high interest rate currencies do not depreciate as much as is predicted by the theory. For the EH-TS, the evidence is more mixed. The EH-TS is often strongly rejected with U.S. dollar (USD) interest rates, but for the currencies of a number of other countries, standard tests often fail to reject.[2]

[2] Campbell and Shiller (1991) and Bekaert, Hodrick, and Marshall (1997) examine the USD evidence. For other currencies see Hardouvelis (1994), Gerlach and Smets (1997), Dahlquist and Jonsson (1995), and Bekaert, Hodrick, and Marshall (1999).

There are three main potential reasons for the rejection of the EHs. First, the EHs are based on the assumption of rational expectations and unlimited arbitrage. It may be that irrational investors make systematic forecast errors, and the ability of rational investors to profit from this situation is limited by their risk aversion. Second, the presence of time-varying risk premiums means that standard tests of the EHs omit the variables capturing the risk premium. If these variables are correlated with interest rates, the estimated coefficients would be pulled away from those implied by the EHs. Third, the tests themselves may lead to false rejections because of their poor properties in finite samples, which can be caused by highly persistent variables, peso problems, or learning.[3] Recently, Bekaert, Hodrick and Marshall (1997, 1999) and Valkanov (1998) have analyzed the poor finite sample behavior of EH-TS tests, and Baillie and Bollerslev (1998), Maynard and Phillips (1998), and Roll and Yan (1998) have argued that poor small-sample behavior may explain the results of EH-FX tests. These papers note that if standard tests are poorly behaved in small samples, inference based on standard asymptotic distribution theory is distorted, and alternative methods of inference are necessary.

In this paper, we re-consider the EHs in a vector autoregressive (VAR) framework. Apart from standard Wald tests, we also investigate Lagrange Multiplier (LM) and Distance Metric (DM) tests that require imposition of the null hypothesis in the estimation. Because the restrictions of the EH-TS are highly non-linear, estimating under these restrictions is generally a non-trivial exercise. We develop an easy-to-implement procedure that extends the suggested estimator of Newey and McFadden (1994) and that works quite well. Once we have estimated the VAR subject to the constraints of the EHs, we can use this system as a data generating mechanism to investigate the small-sample properties of the various tests. We find that the Wald test, the test predominantly used in the literature, has by far the worst small-sample properties. Conducting inference with Wald tests would therefore often be very mis-leading, since the sizes of the tests are quite poor. The DM tests also overreject, but less strongly than the Wald tests. The LM tests, on the other hand, are slightly conservative. Overall, the LM tests perform the best. When reconsidering the evidence on the EHs for the USD, the Deutsche mark (DEM), and the British pound (GBP), we find that inference with the small-sample distributions considerably weakens the case against the EHs.

[3] An early reference to the small-sample problem is Mankiw and Shapiro (1986). Evans (1996) surveys the peso-problem literature, and Lewis (1989) is an early example of the role of learning.

The paper is organized as follows. Section I examines the conditions under which the EHs arise in a no-arbitrage framework. Section II details several econometric approaches to testing the EHs, developing both standard regression tests and the more novel VAR-based tests we propose. Section III briefly describes the data on interest rates and exchange rates. Section IV examines the small-sample properties of the various tests using bootstrap and Monte Carlo analysis. Section V applies the tests to the data. The conclusions summarize our findings and reflect on the usefulness of our test procedures and the technique for imposing non-linear constraints in other settings.

I. The Expectations Hypotheses

By the EH-TS for a particular currency j, we mean that the continuously compounded zero-coupon n-period interest rate, $i^j_{t,n}$, equals the average of the current and expected future short interest rates plus a maturity specific constant:

$$i^j_{t,n} = \frac{1}{n}\sum_{h=0}^{n-1} E_t(i^j_{t+h}) + \alpha^j_n, \tag{1}$$

where we drop the maturity subscript for one period rates.

By the EH-FX, we mean the proposition that the conditional expectation of the continuously compounded rate of appreciation of currency j relative to currency k equals the differential between the continuously compounded interest rates for the two currencies plus a constant. Let S_t denote the currency-k price of currency j. Then, with lower-case letters indicating either natural logarithms of upper-case counterparts or continuously compounded interest rates, the EH-FX is

$$E_t(s_{t+n} - s_t) = \alpha^{k,j}_n + n(i^k_{t,n} - i^j_{t,n}). \tag{2}$$

It is straightforward to demonstrate that these expectation hypotheses are consistent with a class of modern financial models in which assets are priced by no arbitrage restrictions. In economies that do not admit arbitrage, any return denominated in currency j, R^j_{t+1}, satisfies

$$E_t(M^j_{t+1}R^j_{t+1}) = 1, \tag{3}$$

where M^j_{t+1} denotes the currency-j pricing kernel. When the returns and the pricing kernels are log-normally distributed, equation (3) implies the following:

$$E_t(m^j_{t+1}) + 0.5V_t(m^j_{t+1}) + E_t(r^j_{t+1}) + 0.5V_t(r^j_{t+1}) + C_t(m^j_{t+1}, r^j_{t+1}) = 0, \tag{4}$$

where the conditional variance and covariance are denoted $V_t(.)$ and $C_t(.)$, respectively. Because the rate of return associated with the continuously compounded one-period interest rate is in the time t information set, equation (4) implies that

$$i_t^j = -\left[E_t(m_{t+1}^j) + 0.5V_t(m_{t+1}^j)\right]. \tag{5}$$

To derive the implications for the term structure of interest rates, consider the continuously compounded, one-period rate of return on an n-period bond, $rb_{t+1,n}^j \equiv i_{t,n}^j + (n-1)(i_{t,n}^j - i_{t+1,n-1}^j)$. Using equations (4) and (5) we find

$$E_t(rb_{t+1,n}^j) - i_t^j = -\left[0.5V_t(rb_{t+1,n}^j) + C_t(m_{t+1}^j, rb_{t+1,n}^j)\right]. \tag{6}$$

The right-hand side of equation (6) is a constant for any bond-pricing model, such as Vasicek's (1977), in which the logarithmic pricing kernel is conditionally homoskedastic. Let this constant be denoted c_n^j. By using the definition of the rate of return on the bond and the relation between logarithmic bond prices and yields to maturity, equation (6) implies

$$ni_{t,n}^j = i_t^j + E_t\left[(n-1)i_{t+1,n-1}^j\right] + c_n^j. \tag{7}$$

Recursive application of equation (7) and use of the law of iterated expectations implies equation (1) with $\alpha_n^j \equiv \sum_{h=2}^n c_h^j$.

Note that any currency-j return can be converted into a currency-k return by multiplying by S_{t+1}/S_t, which recognizes that one must first purchase one unit of currency j with currency k and then resell the currency j return for currency k. Hence, if markets are complete, and by using equation (3) for each currency, we find that the difference of the logarithms of the pricing kernels equals the rate of appreciation of currency j relative to currency k:

$$m_{t+1}^j - m_{t+1}^k = s_{t+1} - s_t. \tag{8}$$

We can derive the implications for the EH-FX by taking the conditional expectation of equation (8) and substituting from equation (5) evaluated for each of the currencies:

$$E_t(s_{t+1} - s_t) = (i_t^k - i_t^j) + 0.5V_t(m_{t+1}^k) - 0.5V_t(m_{t+1}^j). \tag{9}$$

As with the term structure, the EH-FX is true in economies with conditionally homoskedastic logarithmic pricing kernels. It is possible to derive general expressions for the term premiums and foreign exchange premiums in terms of the conditional moments of the logarithm of the pricing kernel under much weaker conditions than log-normality. The Appendix demonstrates that the conditions for the EHs to hold are constancy of all second and higher order conditional moments of the log pricing kernel.

The logic that leads to equation (9) can also be used to verify that equation (2) holds for the n-period maturity. Note that $s_{t+1}-s_t+i_t^k-i_t^j$ is a one-period excess rate of return that also satisfies equation (6). After substituting into equation (6), the right-hand side is then the one-period foreign exchange risk premium previously indicated by $\alpha_1^{k,j}$. When both the one-period EH-FX holds and the n-period EH-TS holds in both currencies, the n-period foreign exchange premium is constant and equals

$$\alpha_n^{k,j} = E_t\left[(s_{t+n}-s_t) - n(i_{t,n}^k - i_{t,n}^j)\right] = n\alpha_1^{k,j} - n(\alpha_n^k - \alpha_n^j). \tag{10}$$

An investment of a unit of currency k in the n-period currency-j bond earns the currency-j term premium and n times the one-period foreign exchange premium. The opportunity cost is the currency-k term premium.

II.Econometric Procedures

This section develops several alternative econometric approaches to testing the EHs derived in equations (1) and (2). We begin with traditional single-equation specifications and then consider tests based on unconstrained and constrained VARs. Since the validity of the asymptotic distributions of the various test statistics is questionable in the sample sizes we have available, we do not present any estimation results until we have developed all of the statistics and explained how we will assess their finite sample properties.

The derivation of the asymptotic properties of the test statistics relies on Hansen's (1982) Generalized Method of Moments (GMM), which uses orthogonality conditions defined by the theory to develop tests. The orthogonality conditions are based on the assumption of rational expectations, which implies that the realization of a random variable is equal to its conditional expectation plus an error term that is orthogonal to the information set used to form the expectation. To represent a vector of orthogonality conditions specified by the expectation theories, let y_t be a vector of data in the time t information set, and let x_{t-1} be a vector of instruments that are in the time t-1 information set. Let $h(y_t, x_{t-1}, \theta)$ be a vector-valued function of the data and the parameters to be estimated, θ, with the property that

$$E_t\left[h(y_t, x_{t-1}, \theta)\right] = 0 \tag{11}$$

when the null hypothesis is true and the function is evaluated at the parameter θ_0. Let the vector

η_t be an error process defined by the rational expectations assumption applied to equation (11), and define the vector $z_t \equiv (y_t', x_{t-1}')'$ and the vector-valued function of the data and the parameters, $g(z_t, \theta) \equiv \eta_t \otimes x_{t-1}$. Then, the unconditional orthogonality conditions used in a GMM estimation are

$$E[g(z_t, \theta)] = 0. \tag{12}$$

Estimation uses the corresponding sample moment conditions for a sample of size T:

$$g_T(\theta) \equiv \frac{1}{T} \sum_{t=1}^{T} g(z_t, \theta). \tag{13}$$

The parameters are estimated by minimizing the GMM criterion function which is a quadratic form in the sample orthogonality conditions using a weighting matrix, W:

$$J_T(\theta) \equiv g_T(\theta)' W g_T(\theta). \tag{14}$$

Hansen (1982) demonstrates that the optimal weighting matrix is a consistent estimate of the inverse of

$$\Omega \equiv \sum_{k=-\infty}^{k=\infty} E\left[g(z_t, \theta) g(z_{t-k}, \theta)'\right]. \tag{15}$$

Let the gradient of the sample orthogonality conditions be $G_T(\theta) \equiv \nabla_\theta g_T(\theta)$, and let Ω_T represent a consistent estimate of Ω. When the weighting matrix is chosen optimally as Ω_T^{-1}, the GMM asymptotic distribution theory implies that

$$\sqrt{T}(\widehat{\theta} - \theta_0) \rightarrow N[0, (G_T' \Omega_T^{-1} G_T)^{-1}] \tag{16}$$

where $\widehat{\theta}$ denotes the parameter estimate and the symbol $\rightarrow$ denotes convergence in distribution. The standard errors implicit in equation (16) are autocorrelation and heteroskedasticity consistent.

Regression Tests

It is straightforward to derive ordinary least squares (OLS) regression tests of the various expectation hypotheses. Under rational expectations, equation (2) evaluated for $n = 1$ becomes

$$s_{t+1} - s_t = \alpha_1^{k,j} + \beta_1^{k,j}(i_t^k - i_t^j) + \epsilon_{t+1} \tag{17}$$

where ϵ_{t+1} is the rational expectations error term and the null hypothesis is that the slope coefficient equals one. A GMM estimation based on the orthogonality of the error term to a constant and the interest differential reduces to OLS estimation of equation (17), and setting $k = 0$ in equation (15) produces heteroskedasticity consistent standard errors.

Campbell and Shiller (1991) propose two distinct regression tests of the EH-TS based on equation (1). The first specification test can be derived directly from equation (1) under the assumption of rational expectations:

$$\frac{1}{n}\sum_{h=0}^{n-1}(i^j_{t+h}) - i^j_t = \alpha^j_n + \beta^j_n(i^j_{t,n} - i^j_t) + \nu_{t+n-1}. \tag{18}$$

The null hypothesis is again that the slope coefficient equals one, and the estimation uses the fact that the error term, ν_{t+n-1}, is orthogonal to a constant and the term spread at time t. While OLS provides the parameter estimates, appropriate GMM standard errors must allow for the serial correlation of the errors induced by overlapping observations by setting $k = n-1$ in equation (15).

The second specification test of Campbell and Shiller (1991) is derived by rearranging equation (7) and using rational expectations:

$$i^j_{t+1,n-1} - i^j_{t,n} = \delta^j_n + \gamma^j_n \frac{1}{(n-1)}(i^j_{t,n} - i^j_t) + \xi_{t+1}. \tag{19}$$

The OLS specification uses the orthogonality of the error term, ξ_{t+1}, to a constant and the adjusted term spread, and the null hypothesis is again that the slope coefficient equals one. Standard errors can be constructed by setting $k = 0$ in equation (15). When only constant maturities are available, this specification test is often performed with $i^j_{t+1,n}$ on the left-hand side instead of $i^j_{t+1,n-1}$. Bekaert, Hodrick, and Marshall (1997) note that this change of variables leads to an upward bias in the prediction of the slope coefficient such that values greater than one are expected under the null hypothesis, even asymptotically.

Tests from Unconstrained Vector Autoregressions

It is also possible to develop GMM-based tests of the expectation hypotheses using the orthogonality conditions of a VAR. With a VAR, one can test the theory directly as well as calculate implied slope coefficients that are analogous to the directly estimated OLS slope coefficients discussed above. Below, we examine VARs that involve a two-country framework using data from three developed economies to investigate the various EHs. For convenience of presentation, we number the currencies and use standard currency abbreviations in the following way: one for the USD, two for the DEM, and three for the GBP. Thus, i^1_t is the USD short interest rate, and sp^2_t is the spread between the DEM long interest rate and the DEM short interest rate. Since all rates of change of exchange rates are expressed versus the USD, Δs^j_t is the rate of appreciation of the USD relative to currency j, for $j = 2, 3$. The variables in the VAR are the rate of appreciation of the USD relative to a currency j, the USD interest rate, the currency-j interest rate, the USD spread, and the currency-j spread. To develop the econometric model, stack the five variables into the vector $y_t \equiv (\Delta s_t^j, i_t^1, i_t^j, sp_t^1, sp_t^j)'$. Then, let a K-th order VAR represent the demeaned data generating process for y_t:

$$y_{t+1} = \sum_{k=1}^{K} B_k y_{t-k+1} + \eta_{t+1} \tag{20}$$

where the parameters B_k represent five-dimensional square matrixes of coefficients, and η_{t+1} is the vector of innovations that is orthogonal to the time t information set. The first-order companion form of the VAR can be represented using the vector $x_t \equiv (y_t', y_{t-1}', ...y_{t-K+1}')'$:

$$x_{t+1} = \Theta x_t + \xi_{t+1}. \tag{21}$$

The parameter matrix, Θ, is a 5K-dimensional square matrix with the B_k matrixes stacked horizontally in the first five rows, a 5(K-1) identity matrix beneath these parameters on the left, and zeroes elsewhere. The innovation vector, $\xi_{t+1} \equiv (\eta_{t+1}', 0...0)$, has variance matrix Σ. With this specification there are (25K) parameters in θ_0.

We use the VAR parameters and the asymptotic distribution in equation (16) to generate test statistics that are based on implied counterparts of the OLS slope coefficients. We can also develop tests of the full restrictions of the EHs in the VAR framework. To derive these tests we need to consider the implications of the EHs for the coefficients of the VAR.

Although the EHs are based on the full information set of economic agents, as long as that information set includes the information on the right-hand sides of the VAR equations, the law of iterated expectations implies that we can use the VAR to test the theories. From the companion form of the VAR in equation (21), we know that forecasts of x_{t+h}, based on the information in the VAR at time t, may be generated as

$$E_t^x(x_{t+h}) = \Theta^h x_t, \tag{22}$$

where the expectation is with respect to the information set of the VAR. The EHs consequently imply highly non-linear sets of restrictions on the parameters. To derive the constraints on the parameters, define the indicator vectors, e_j, which have dimension 5K, a one in the j-th position, and zeroes elsewhere. The vector of restrictions from the EH-FX for one-period interest rates may be written as

$$e_1'\Theta = (e_3' - e_2'). \tag{23}$$

Next, consider the derivation of the restrictions of the EH-TS for each currency. For the USD interest rates, equation (1) implies the following restrictions on the underlying parameters of the VAR:

$$e_4' = e_2' \left[(1/n)(I - \Theta^n)(I - \Theta)^{-1} - I \right]. \tag{24}$$

The analogous restrictions associated with the EH-TS of the foreign currency interest rates are

$$e_5' = e_3' \left[(1/n)(I - \Theta^n)(I - \Theta)^{-1} - I \right]. \tag{25}$$

The representations of the EHs in equations (23)-(25) allow estimation of implied slope coefficients that are analogous to the directly estimated OLS coefficients. For example, the implied slope coefficient from the VAR that is analogous to the slope coefficient in equation (17) is

$$\beta_{k,j} = \frac{e_1' \Theta \Psi (e_3 - e_2)}{(e_3 - e_2)' \Psi (e_3 - e_2)} \tag{26}$$

where Ψ is the unconditional variance of x_t, which is computed from $vec(\Psi) = (I - \Theta \otimes \Theta')^{-1} vec(\Sigma)$. The numerator of equation (26) is the covariance between the expected future rate of appreciation and the interest differential, while the denominator is the variance of the interest differential. Similarly, the implied slope coefficient for the USD EH-TS analogous to equation (18) is the covariance between the average of the expected future interest rates and the current interest rate spread divided by the variance of the current spread:

$$\beta_n^{USD} = \frac{e_2' [(1/n)(I - \Theta^n)(I - \Theta)^{-1} - I] \Psi e_4}{e_4{}' \Psi e_4}. \tag{27}$$

The implied OLS coefficient corresponding to equation (19) for the USD which uses the substitution of the n-period rate for the n-1 period rate is

$$\gamma_n^{USD} = \frac{(e_4 + e_2)' (\Theta - I) \Psi e_4 (n-1)}{e_4{}' \Psi e_4}. \tag{28}$$

To develop Wald tests of the three expectation hypotheses, let the null hypotheses in equations (23)-(25) be summarized by

$$H_0 : a(\theta_0) = 0, \tag{29}$$

where $a(\theta_0)$ is a $15K$-dimensional vector that is non-linear in the underlying parameters. Let the sample counterpart of this vector be $a_T(\theta)$, let the gradient of the constraints with respect to the parameters be $A_T \equiv \nabla_\theta a_T(\theta)$, and let $B_T \equiv G_T' \Omega_T^{-1} G_T$. Then, it follows from a Taylor's Series approximation that

$$\sqrt{T} a_T(\widehat{\theta}) \rightarrow N(0, A_T B_T^{-1} A_T'). \tag{30}$$

A Wald test of the null hypothesis asks how close are the constraints to being satisfied at the unconstrained parameter values. The test statistic follows from the asymptotic distribution in equation (30):

$$Ta_T(\widehat{\theta})'(A_T B_T^{-1} A_T')^{-1} a_T(\widehat{\theta}) \rightarrow \chi^2(15K). \tag{31}$$

Estimation under the null hypothesis

Both Distance Metric statistics, which are based on intuition from maximum likelihood, and Lagrange Multiplier statistics require estimation of the parameters subject to the highly non-linear constraints of equation (29), which is quite difficult. One approach to constrained estimation follows Melino (1983), who corrected an error in Sargent's (1979) maximum likelihood estimation of the EH-TS. Melino (1983) recognizes that the EH-TS imposes significant restrictions on the eigenvectors of Θ.

To understand these restrictions, consider a first-order VAR in which the five eigenvalues of Θ are distinct. In this case, we can do an eigenvalue decomposition:

$$\Theta = P\Lambda P^{-1}, \tag{32}$$

where Λ is the diagonal matrix of eigenvalues and P is the matrix with the corresponding eigenvectors in its columns. Now, to derive the restrictions of the EHs, substitute from equation (32) into equation (23) and multiply from the right-hand side by P implies

$$e_1' P\Lambda = (e_3' - e_2')P. \tag{33}$$

Let the diagonal elements of Λ be λ_j, and let the rows of P be P_i, with distinct elements P_{ij}. Since P_1 can be normalized to a row vector of ones, this constraint implies

$$P_{3j} = P_{2j} + \lambda_j. \tag{34}$$

By substituting equation (32) into equations (24) and (25) and simplifying, we find

$$e_4' P = e_2' P[(1/n)(I - \Lambda^n)(I - \Lambda)^{-1} - I] \tag{35}$$

and

$$e_5' P = e_3' P[(1/n)(I - \Lambda^n)(I - \Lambda)^{-1} - I]. \tag{36}$$

The restrictions in equations (34)-(36) imply that the ten free parameters of the constrained estimation of a first-order VAR are the five eigenvalues and the five parameters of the second row of the eigenvectors. All other parameters are functions of these fundamental parameters. Since the eigenvalues can be complex conjugates, direct estimation of the constrained system is quite complicated because the search must be conducted over potentially complex numbers.

To estimate the parameters, θ, subject to the constraints in equation (29), we instead follow an indirect route that extends the estimator proposed by Newey and McFadden (1994). Define the Lagrangian for the constrained GMM maximization problem to be.

$$L(\theta, \gamma) = -(1/2) g_T'(\theta) \Omega_T^{-1} g_T(\theta) - a_T(\theta)' \gamma \tag{37}$$

where γ is a vector of Lagrange multipliers. Let an overbar denote estimates subject to the constraints. Then, the first-order conditions for this problem can be written as

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -G_T' \Omega_T^{-1} \sqrt{T} g_T(\overline{\theta}) - A_T' \sqrt{T} \overline{\gamma} \\ -\sqrt{T} a_T(\overline{\theta}) \end{bmatrix}. \tag{38}$$

While equation (38) is non-linear in the parameters, we can derive an approximate asymptotic solution using the law of large numbers and a Taylor's Series expansion. Recognize that

$$\sqrt{T} g_T(\theta_0) \rightarrow N(0, \Omega), \tag{39}$$

$$\sqrt{T} g_T(\overline{\theta}) \approx \sqrt{T} g_T(\theta_0) + G_T \sqrt{T}(\overline{\theta} - \theta_0), \tag{40}$$

and

$$\sqrt{T} a_T(\overline{\theta}) \approx \sqrt{T} a_T(\theta_0) + A_T \sqrt{T}(\overline{\theta} - \theta_0). \tag{41}$$

Under the null hypothesis, $a_T(\theta_0) = 0$. Hence, when we substitute from equations (40) and (41) into the first-order conditions, we find

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -G_T' \Omega_T^{-1} \sqrt{T} g_T(\theta_0) \\ 0 \end{bmatrix} - \begin{bmatrix} B_T & A_T' \\ A_T & 0 \end{bmatrix} \begin{bmatrix} \sqrt{T}(\overline{\theta} - \theta_0) \\ \sqrt{T} \overline{\gamma} \end{bmatrix}. \tag{42}$$

The formula for a partitioned inverse implies that

$$\begin{bmatrix} B_T & A_T' \\ A_T & 0 \end{bmatrix}^{-1} = \begin{bmatrix} B_T^{-1/2} M_T B_T^{-1/2} & B_T^{-1} A_T' (A_T B_T^{-1} A_T')^{-1} \\ (A_T B_T^{-1} A_T')^{-1} A_T B_T^{-1} & -(A_T B_T^{-1} A_T')^{-1} \end{bmatrix} \tag{43}$$

where $M_T \equiv I - B_T^{-1/2} A_T' (A_T B_T^{-1} A_T')^{-1} A_T B_T^{-1/2}$ is an idempotent matrix. Thus, the asymptotic distribution for the constrained estimator and the Lagrange multiplier is

$$\sqrt{T}(\overline{\theta} - \theta_0) \rightarrow N(0, B_T^{-1/2} M_T B_T^{-1/2}) \tag{44}$$

$$\sqrt{T}\overline{\gamma} \to N[0, (A_T B_T^{-1} A_T')^{-1}]. \tag{45}$$

Although direct maximization of the Lagrangian in equation (38) is feasible, it is often computationally difficult. We instead extend the approach suggested in Newey and McFadden (1994) who demonstrate how to derive a constrained consistent estimator starting from an initial unconstrained consistent estimator and using only matrix algebra. Let $\widetilde{\theta}$ represent an initial consistent unconstrained estimate. Then, we have

$$g_T(\overline{\theta}) \approx g_T(\widetilde{\theta}) + G_T(\overline{\theta} - \widetilde{\theta}) \tag{46}$$

$$a_T(\overline{\theta}) \approx a_T(\widetilde{\theta}) + A_T(\widetilde{\theta} - \overline{\theta}). \tag{47}$$

After substituting into the first-order conditions and solving, we find

$$\overline{\theta} \approx \widetilde{\theta} - B_T^{-1/2} M_T B_T^{-1/2} G_T' \Omega_T^{-1} g_T(\widetilde{\theta}) - B_T^{-1} A_T' (A_T B_T^{-1} A_T')^{-1} a_T(\widetilde{\theta}) \tag{48}$$

$$\overline{\gamma} \approx -(A_T B_T^{-1} A_T')^{-1} A_T B_T^{-1} G_T' \Omega_T^{-1} g_T(\widetilde{\theta}) + (A_T B_T^{-1} A_T')^{-1} a_T(\widetilde{\theta}). \tag{49}$$

While Newey and McFadden (1994) note that the estimators in equations (48) and (49) are consistent, they do not satisfy the constrained optimization problem exactly. In constructing our constrained estimates, we iterated on equations (48) and (49), substituting the first constrained estimate for the initial consistent unconstrained estimate to derive a second constrained estimate, and so forth. We stopped the iterative process when the resulting constrained estimate satisfied the constraints, i.e. when $a_T(\overline{\theta}) = 0$.

The values of the Lagrange multipliers are not zero at the constrained parameter estimates when imposition of the constraints significantly affects the value of the objective function. An LM test asks whether we can reject the hypothesis that the multipliers are jointly zero. From equation (45), the LM test for a K-th-order system is

$$T\overline{\gamma}'(A_T B_T^{-1} A_T')\overline{\gamma} \to \chi^2(15K). \tag{50}$$

A GMM-based distance metric (DM) test, analogous to a likelihood ratio test, can also be developed. Typically, this test is constructed as the sample size times the difference between the GMM objective function evaluated at a constrained estimate and the GMM objective function evaluated at the unconstrained estimate using the same weighting matrix in each estimation. Since our unconstrained problem is just identified, the value of the GMM objective function is zero in this case. Hence, the DM test for a K-th-order system is

$$T g_T(\overline{\theta})' \Omega_T^{-1} g_T(\overline{\theta}) \to \chi^2(15K). \tag{51}$$

III. The Data

Table 1 presents some summary statistics for the variables. All variables are measured in percentage points per annum. Monthly rates of appreciation are annualized by multiplying by 1200. While this transformation does not affect the interpretation of the mean returns, the annualized standard deviation is not the standard deviation associated with an annual holding period. The sample period is January 1975 to July 1997. The exchange rates and Eurocurrency interest rates are from Datastream. The dollar-based exchange rates are calculated from the quoted sterling exchange rates which are closing middle rates provided by Reuters.

Notice that the rates of appreciation are quite volatile and have very small autocorrelations. The one-month interest rates are all highly autocorrelated, and the spreads between twelve-month rates and one-month rates are persistent but not as highly autocorrelated as the short rates. Use of interest rates and spreads as predictors of the rates of appreciation is consistent with the idea that predictable changes in asset prices are small relative to their unpredictable changes.

IV. Econometric Analysis of Test Statistics

The goal of this section is two-fold. We first integrate our analysis with the recent evidence on the small-sample characteristics of standard regression tests of the EH-TS and the EH-FX. Various authors, including Bekaert, Hodrick and Marshall (1997, 1999), Schotman (1996), and Valkanov (1998) have demonstrated that the standard regression tests of the EH-TS are ill-behaved in small samples under a variety of data generating processes (DGPs). In particular, small-sample biases arise for essentially the same reason that was first discussed by Kendall (1954) in the context of estimation of the parameters of autoregressive processes. The regressors are serially correlated lagged dependent variables. Although the parameter estimates are consistent, the absence of strict exogeneity of the regressors implies bias in small samples.[4] In EH-TS tests, the regression coefficients are upwardly biased and their small-sample distributions are very dispersed. Tauchen (1985) and Baillie and Bollerslev (1998) have also shown that EH-FX regressions suffer from a similar problem. Unfortunately, research about the small-sample problems of doing inference about the validity of the EHs does not arrive at a common conclusion.

[4]Stambaugh (1999) provides a recent Bayesian treatment of these issues.

Our VAR model imposes the three EHs while matching the time-series properties of the data. Hence, we derive the small-sample distributions of the regression coefficients under the null hypothesis within a model that accommodates realistic persistence in both the foreign and local interest rates and Granger-causality of interest rates both by spreads and exchange-rate changes. Moreover, we compare the distributions of the standard regression coefficients with the distributions of the slope coefficients implied by the VAR. If the VAR adequately captures the dynamics of the data, we obtain slightly more efficient estimates in some instances. For example, the long-run (12 month) unbiasedness test and the test of Equation (18) lead to the loss of data, which is not the case in the VAR.

A second goal of this section is to examine whether alternatives to the simple Wald test have superior small-sample properties. By imposing the non-linear constraints on the VAR dynamics, we are also able to examine the relative size and power properties of the Wald, LM and DM tests described in Section III.[5] Given the well-known problems with Wald tests in general (as discussed in Burnside and Eichenbaum (1996) for example), it may well be that these other tests have superior small-sample properties.

Alternative Data Generating Processes

We use two DGPs in the Monte Carlo analysis. Both start from an unconstrained five variable VAR. In principle, we could then apply the iterative scheme described in Section III to find the VAR parameters that impose the null. However, as Bekaert, Hodrick, and Marshall (1997) note, the estimated VAR parameters are biased in small samples. Hence, these parameters do not constitute a relevant starting point.

The bias-correction we implement proceeds as follows. We use the estimated unconstrained VAR parameters to generate 100,000 artificial data sets of 269 observations using an i.i.d. bootstrap of the residuals. We re-estimate the VAR parameters from these replications of the initial data. The bias in the estimated parameters is estimated by the difference between the known parameters of the DGP and the means of the Monte Carlo distributions based on the 100,000 replications. We then bias-correct the original estimates by adding these biases to the original unconstrained estimates. This yields a bias-corrected set of unconstrained parameters, μ^u and A^u, which are also used in

[5]Ligeralde (1997) examines the small-sample performance of various methods of constructing Wald tests. The differential performance across alternatives is mostly due to how one deals with the serial correlation induced by the overlapping error structure in the data. In our VAR setting however, this overlapping data problem does not arise.

simulations to represent an alternative hypothesis in which there are violations of the EHs. To determine bias-corrected parameters that satisfy the null hypothesis, we use μ^u and A^u to simulate a very long series (70,000 observations plus 1,000 starting values that are discarded), which is then subjected to the iterative estimation scheme described in Section III. These parameters are our bias-corrected constrained parameters, μ^c and A^c

In all cases we use a first-order VAR as that is the order chosen by the Schwarz Criterion. Table A1 in the Appendix reports these test statistics in Panel A along with Cumby-Huizinga (1992) l-tests for residual serial correlation in Panel B. Only for the residuals of the USD and DEM spreads in that VAR do we find any evidence inconsistent with the first-order model. The three panels of Table A2 report the unconstrained parameter estimates with their bias-corrected counterparts for the three VARs. Table A3 reports the estimates of the VAR coefficients that are constrained to satisfy the EHs.

In all of our experiments, we use the constrained coefficients that are estimated from simultaneously imposing the EHs. For the first DGP, we bootstrap the original residuals from the unconstrained VAR, and reconstruct constrained and unconstrained data, using μ^c and A^c, and μ^u and A^u, respectively. Whereas the sample size for each experiment is 269, each experiment generates an initial 1,000 observations that are discarded. We also check the validity of the computer code by letting the sample size become very large and verifying convergence to the asymptotic distributions. We conduct this bootstrap procedure for both the DEM-USD and the GBP-USD VARs.

Although the bootstrap procedure captures skewness and leptokurtosis in the data, it is potentially unrealistic because it destroys higher-order dependence in the residuals (for example, volatility clustering).[6] To accommodate temporal heteroskedasticity and its potential effects on small-sample distributions, we also use a Monte Carlo experiment based on a parameterized model of the residuals. We use the same conditional mean coefficient matrices as in the bootstrap DGP, but we draw the error terms according to a multivariate GARCH model.

The GARCH model is similar to the factor GARCH models of Engle, Ng and Rothschild (1990), Bekaert and Harvey (1997), and Bekaert, Hodrick, and Marshall (1997). We model the innovation

[6]We experimented with stationary bootstrap methods, as in Politis and Romano (1994) and Politis, Romano, and Wolf (1997), which allow for dependence, but they do not seem well-suited for problems where the data are highly persistent but residuals ought to be uncorrelated.

vector, η_t, as a factor structure with the innovations of the short rates in the two countries as the factors. Thus,

$$\eta_t = F e_t \tag{52}$$

with

$$F = \begin{bmatrix} 1 & f_{12} & f_{13} & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & f_{32} & 1 & 0 & 0 \\ 0 & f_{42} & f_{43} & 1 & 0 \\ 0 & f_{52} & f_{53} & 0 & 1 \end{bmatrix}. \tag{53}$$

Note that the innovation in the USD interest rate affects the innovation in the foreign interest rate, but the foreign interest rate shock does not affect the USD interest rate innovation. In effect, f_{32} determines the correlation between the two fundamental shocks to the system. In equation (52), the vector e_t represents the idiosyncratic innovations. Hence, $E_{t-1}\left[e_t e_t'\right] = H_t$, where H_t is a diagonal matrix. As a result, the conditional covariance matrix of the innovations, η_t,which is denoted Σ_t, can be written as $\Sigma_t = F H_t F'$. We assume that elements in H_t corresponding to the two factors and the conditional exchange rate variance follow a GARCH(1,1) process (see Bollerslev (1986)). For the conditional variances of the interest rates, we augment the model to allow the conditional variance to depend on the past interest rate as in the univariate model of Gray (1996). Thus, the model for the conditional variances can be written as follows:

$$h_t^j = \beta_j h_{t-1}^j + \alpha_j (e_{t-1}^j)^2 + \omega_j (i_{t-1}^j), \qquad j = 2, 3. \tag{54}$$

The modification to the usual GARCH model accommodates the dramatic shift in short-rate volatility during the monetary targeting period of 1979-1982. In this model, the conditional variances of the twelve-month term spreads and the rate of change of the exchange rate have three components: a component linear in the conditional variance of the USD short rate, a component linear in the conditional variance of the foreign short rate and an idiosyncratic component. Compared to other multivariate GARCH models, the model is very parsimonious with only 18 parameters. This parsimony is achieved by restricting the covariance matrix to depend only on the conditional variances of the two short rates.

To estimate the model in equations (52) to (54), we exploit the block-diagonal nature of the information matrix and estimate the multivariate GARCH model from the VAR residuals, using quasi-maximum likelihood. Hence, we assume normal innovations to construct the likelihood function although the true distribution of the innovations may not be normal. White (1982) and Bollerslev and Wooldridge (1992) show that the resulting estimator is consistent and asymptotically normal.

Tables A4 and A5 contain the estimation results for the GARCH models for the DEM and USD rates and the GBP and USD rates, respectively. We first discuss the DEM-USD system. The conditional variances of the USD and DEM short rates are moderately persistent with large ARCH coefficients. There is some remaining time-variation in the idiosyncratic component of the conditional variance of the exchange rate, but it shows little persistence. The exchange rate shows small, but statistically significantly positive factor loadings with respect to both the USD and the DEM interest rates.[7] The USD term spread residual is negatively correlated with the USD short-rate shock, as is expected, and it only weakly depends on the DEM rate. The DEM spread residual is also strongly negatively correlated with the DEM short rate, but it is correlated positively with the USD short rate. This does not necessarily imply that unexpected increases in the USD short rate steepen the German yield curve, since the USD short rate is positively related to the DEM short rate, and increases in the DEM short rate increases flatten the yield curve.

Table A5 reports the GBP-USD system. The estimates are in many ways qualitatively similar to the DEM-USD system although the conditional variances of both the USD and GBP short rates show more persistence. We again find positive exchange-rate factor loadings with respect to both the USD and GBP interest rate shocks, but the GBP interest rate effect is statistically insignificant. The factor loadings for the spreads also have the same signs as in the DEM-USD system. The covariance between the USD and GBP interest rate shocks is much lower than the comparable one between the USD and DEM interest rates. The USD-GBP system does somewhat under-predict the unconditional variances of both the USD and the GBP interest rates.

As mentioned above, the innovations in the Monte Carlo experiments are drawn either from the bootstrap procedure or the GARCH models, and the DGP satisfies the null of the EHs using the bias-corrected, constrained VAR parameters. The bias-corrected, unconstrained VARs serve as natural alternative models.

Properties of Test Statistics in Finite Samples

[7] In traditional theories of exchange rate determination, the correlation of exchange rate innovations with interest rate innovations depends on whether the shock causing interest rates to move reflects a change in expected inflation or in the expected real rate. The latter case predicts a positive correlation for the USD and a negative correlation for the DEM. That is, if the USD (DEM) short rate unexpectedly rises, the dollar (mark) ought to appreciate.

From the DGPs described above, we simulate 25,000 artificial samples of 269 observations. We focus on two sets of results. First, we investigate the small-sample distributions of the various regression coefficients in the standard regressions used to test the EHs. Second, we examine the performance of the three test statistics (Wald, LM and DM) in terms of size and power against the alternative hypothesis.

Tables 2 and 3 present some relevant characteristics of the small-sample distributions of the slope coefficients in the various regression tests under the two different data generating processes, the bootstrap in Table 2 and the GARCH model in Table 3. We report only the left-hand tail area quantiles because the sample parameter estimates are all less than the null value. We consider both OLS regression coefficients and regression coefficients implied by the VAR parameters.

A comparison of the means and the medians for the distributions of the OLS coefficients indicates that they show little asymmetry. The biases, defined as the deviations of the mean values of the empirical distributions from the values under the null hypothesis, are rather small for the EH-FX tests and the EH-TS tests based on equation (18). The biases are considerably larger for the tests based on equation (19), but most of this bias is due to the maturity mis-match between the twelve-month interest rate used in the test and the eleven-month interest rate that should be used. Hence, the bias largely remains present even in samples of 50,000. The biases in the EH-TS tests are consistent with the results in Bekaert, Hodrick, and Marshall (1997) where biases only become quite substantial for longer maturities.

Now consider the dispersion of the slope coefficients. The standard deviations of the empirical distributions in Panels A and B of Table 2 are larger than their corresponding values in Table 3 except for those associated with the DEM term structure. This reflects the inability of the GARCH models to match the fat tails in the data. The standard deviations of the FX slopes for the DEM in Panel A are much larger than the standard deviations of the term structure slopes, but they are not noticeably larger than the asymptotic standard errors except for the regression at the twelve-month horizon. This is true in Panel B for the GBP as well except the standard deviations of the FX slopes are now smaller than the asymptotic values. Notice also that the left tails of the distributions of the FX tests in Table 2 include substantially negative values. The slope coefficients from equation (19) also show much more dispersion than those from equation (18). Note that these results are similar across the two currencies.

The small-sample distributions for the implied regression coefficients from the VARs are quite similar to the distributions of the OLS regression coefficients. Overall, the biases are slightly smaller, but with a few exceptions for the GBP, the quantiles are remarkably alike across the two sets of coefficients. This indicates that the VAR generally provides a good description of the relevant dynamics of the data. Note though that the dispersion of the small-sample distributions is sometimes larger for the EH-TS tests because there are a few extreme observations.[8]

Tables 4, 5, and 6 focus on the small-sample properties of the various test statistics.[9] Table 4 considers properties of the small-sample distribution from the bootstrap DGP. We consider first the EH-FX tests individually for one-month and twelve-month horizons and jointly for both horizons. We then consider the EH-TS tests for the USD and for the DEM in Panel A and for the USD and the GBP in Panel B. Finally, we consider joint tests of all three EHs. As noted above, with a first-order VAR, each individual test imposes five restrictions on the VAR parameters. Hence, the appropriate asymptotic distributions for comparison purposes are the $\chi^2(5)$ for the individual tests, $\chi^2(10)$ for the joint FX test, and $\chi^2(15)$ for the simultaneous test of all EHs.

Panel A of Table 4 reveals that the means of the small-sample distributions are slightly higher than the corresponding chi-square means, in all but one case (the joint LM test). The upward bias is most severe for the Wald tests and very small for the LM tests. A similar relation holds for the dispersions of the test statistics. The DM tests and especially the Wald tests show much more dispersion than their corresponding asymptotic distributions. The distributions of the Wald tests are significantly shifted to the right, as are the distributions of the DM tests, but less dramatically so. The LM tests actually show slightly less dispersion than the corresponding chi-square distributions. Given these findings, it is not surprising that the empirical critical values do not correspond with the asymptotic ones. The LM tests slightly under-reject at the asymptotic critical value in some cases, but in general their small-sample distributions are far closer to the asymptotic distributions than those of the other test statistics. The distortions of the Wald test appear worst for the EH-TS tests. The distortions of the Wald test also worsen considerably when the number of restrictions increases. For example, whereas the 99% quantile for a $\chi^2(15)$ is 30.58, the 99% value in the

[8] Occasionally, the implied coefficients show rather extreme standard deviations which can be traced to outliers caused by VAR nonstationarity. The removal of one outlier typically suffices to bring the standard deviation back in line with the other results.

[9] These test statistics are asymptotically pivotal because their limiting distributions do not depend on any unknown parameters in contrast to the distributions of the regression coefficients. Statisticians argue that examining asymptotically pivotal statistics improves finite-sample inference. See Berkowitz and Kilian (1996), for example.

small-sample distribution of the Wald test for all restrictions in the DEM-USD system is 63.66.

Panel B of Table 4 reports the same characteristics for the GBP system. All the observations made above remain valid, but we now record a few more instances in which the empirical mean of the LM tests is slightly below the asymptotic mean. In general, the closeness of the results between the two panels is extremely encouraging. For example, the means of the small-sample distributions for the EH-FX tests are at most 0.10 apart across the two tables. For the individual EH-TS tests we have four sets of results (the USD twice, the DEM and the GBP). Across these four sets of results, the 95% quantiles vary between 17.32 and 19.66 for the Wald test, between 10.79 and 11.24 for the LM test (the corresponding chi-square value is 11.07), and between 14.60 and 15.21 for the DM test. This is a clear illustration of the remarkable robustness across currencies of our distributions, and it nicely illustrates the relative qualities of the test statistics.

Table 5 repeats all of these results for the GARCH DGP. All of the results remain robust. To illustrate, let us focus on the joint tests, since they feature the largest distortions. The means of the Wald tests are 27.47 in the DEM-USD system and 28.22 in the GBP-USD system, and the 95% quantiles are 54.27 and 56.33, respectively. The distortions here are somewhat larger than for the bootstrap results where the 95% quantiles are 47.46 and 52.50, respectively. Compared to the 95% critical value of a $\chi^2(15)$ of 25, the size distortions are considerable. There is also a significant rightward shift for the DM test. Its mean is 18.25 in the DEM-USD system and 17.76 in the GBP-USD system. The 95% quantiles are 28.50 and 27.17, respectively. Apart from showing a much smaller distortion relative to the Wald test, the small-sample distribution of the DM test is also more alike across currencies and DGPs. The 95% quantiles in the bootstrap case are 27.98 for the DEM-USD system and 31.16 for the GBP-USD case. The LM test is again the best-behaved. The mean of its distribution is 15.42 for the DEM-USD system and 14.78 for the GBP-USD system, which is very close to the mean of the $\chi^2(15)$. Since the empirical distributions of the LM tests have smaller variances than the asymptotic distributions, it is not surprising that the 95% quantiles are lower than the corresponding value of 25 for a $\chi^2(15)$. The 95% quantiles are 23.27 in the DEM-USD system and 21.95 in the GBP-USD system. The under-coverage of the LM test is worst for the joint test. Since the 95% critical values in the bootstrap case were 22.15 for the DEM-USD system and 21.57 for the GBP-USD case, this test also shows remarkable robustness across currencies and DGPs.

In Table 6 we focus on the empirical size and the empirical power of the various tests at the nominal 5% significance level. The empirical size of a test is the percent of the Monte Carlo experiments conducted under the null hypothesis in which the test statistic exceeds the asymptotic critical value associated with a 5% type one error. These values are reported in Panel A. The empirical power of a test is the percent of Monte Carlo experiments conducted under the alternative hypothesis in which the test statistic exceeds the empirical critical value. These critical values are reported in Tables 4 and 5. Panel B of Table 6 reports the values for the empirical powers of the tests where the alternative hypothesis is the unconstrained VAR.

Whereas all tests show size distortions, the Wald test has by far the worst size properties. Its empirical size for a 5% nominal test is at least 10.50 %. The empirical size is considerably worse for the EH-TS tests reaching 26.2% for the USD test in the DEM-USD GARCH DGP. For the joint test of the EHs, the empirical sizes of the Wald tests vary between 44.6% and 50.6%. Since this test has been the one used most in empirical work, these findings may potentially change inference regarding the validity of the EHs. The DM tests also have size distortions for the EH-FX tests with a largest empirical size of 15.4%, but the sizes of the DM tests are smaller than those of the corresponding Wald tests, except in one case. The empirical sizes of the LM tests for a 5% nominal size vary between 0.7% and 7.9%. In the majority of the cases, the sizes of the LM tests are smaller than 5%, and in virtually half of the cases the empirical size is within 1% of the nominal size.

To assess the power of the tests, we use the unconstrained VAR as the alternative hypothesis. We find that the power of the tests depends critically on which of the null hypotheses is tested and to some extent on the DGP. For the DEM-USD system, the EH-TS tests are more powerful than the EH-FX tests. Note that the information set considered for the EH-FX test is larger than what is typically considered in regression tests, where changes in foreign exchange rates are regressed on an interest differential. Here the coefficients on the interest rates are allowed to be different in absolute magnitude and the spreads are allowed to predict changes in exchange rates. The power of the FX tests hovers around 55%. For the test of the EH-TS in the DEM, the power is very high for the bootstrap DGP (in excess of 95%) and between 74.3% and 85.4% for the GARCH system. For the USD term structure, the roles are reversed, with the GARCH system yielding more powerful tests, generally in excess of 95%. Nevertheless, even for the bootstrap system, power is still in excess of 80%. For the joint EH-FX test, power is slightly in excess of 50% for the LM and DM tests, but drops to 45.1% for the Wald test in the bootstrap DGP and to 12.5% for the Wald test in the GARCH DGP.

For the GBP system, power is always higher for the EH-FX tests than in the DEM-USD system. We find that empirical power in these test varies between 56.5% and 99.4%. EH-TS tests for the USD and the GBP are less powerful with the exception of the GBP tests for the GARCH system in which they exceed 94% in every case.

For the joint tests of the EHs, there is uniformly high power. For the GARCH system, the LM and DM tests have power over 99%, while the power is generally smaller for the bootstrap DGP, it never falls below 91.8%.

We also checked to see that all tests are consistent in that power goes to one when the sample is increased. Simulations of samples with 50,000 observations reveal powers very close to 1.00 for all tests. For our small samples of 269 observations, it is important to assess which tests are most powerful. Of course, we already know that the LM test has superior size properties and should be the preferred test, if it has comparable power to the other tests. Across the two DGPs, the two currencies, and the various tests, we can make a total of 20 power comparisons. In 17 cases, the DM test is most powerful and comes in second in the three other cases. The LM test is never the most powerful test, but comes in second in 13 cases. Moreover, whereas the Wald test is sometimes more powerful than the LM test, when it is not, its power is substantially below that of the other tests.

Taken together, our results strongly suggest avoiding use of the Wald test. The DM test has reasonable size properties, but its use would lead to over-rejection of the null hypothesis. It is also quite powerful when applied correctly. The LM test is by far the best test. It has very good size properties, and it has good power. In some cases, it may turn out to be a slightly conservative test, which fails to reject the null when it is false. Ironically, the LM test is arguably the least used of all in applied work. Having established the small-sample properties of the various test statistics allows us to revisit the evidence on the EHs in the data.

V. Statistical Analysis of the Data

This section evaluates the validity of the EHs using the small-sample distributions developed above. Two types of evidence are interpreted. First, we consider the regression evidence corresponding to equations (17), (18), and (19) for slope coefficients from ordinary least squares (OLS) regressions and the corresponding implied coefficients (IOLS) from the VAR. Then, we consider the test statistics from the VARs.

The Regression Evidence

Consider first the results in Table 2. For the DEM/USD rate, the slope coefficients corresponding to equation (17) of -0.527 (OLS) and -0.498 (IOLS) fall between the 2.5% and 5% quantiles of the empirical distributions. After allowing for a two-sided test, this evidence is consistent with the large-sample inference one would do based on asymptotic standard errors of 0.923 (OLS) and 0.979 (IOLS) which produces p-values of .1 for the null that the slope coefficient is one. Similarly, for the GBP/USD rate, the coefficient estimates of -1.654 (OLS) and -1.662 (IOLS) are well below the 0.5% quantile of the empirical distributions. Hence, the small-sample inference supports the asymptotic inference that rejects the null at smaller than a 1% marginal level of significance. The evidence for both of these rates at the twelve-month horizon is not quite as strong.

The situation for the term structure is in many ways the reverse of the above. We now reject the EH-TS for the USD at the 1% level for equation (18) and at the 5% level for equation (19). We reject very strongly in the DEM term structure, but we do not reject at all in the GBP term structure.

Similar inference can be drawn from the distributions in Table 3 because the GARCH model generally produces less dispersion in the slope coefficients. The p-values of the FX tests would actually be smaller than the asymptotic p-values.

The VAR tests

Analysis of Tables 4 and 5 indicates that use of empirical critical values generally weakens the evidence against the EHs for the Wald and DM tests. Consider a researcher who conducts inference using Wald tests and their asymptotic critical values, which is undoubtedly the most common approach in the literature. Such a researcher would conclude that there is strong evidence against the EH-FX for the GBP/USD but not in the DEM/USD, that the EH-TS is rejected for the USD and the DEM but not for the GBP. Notice, though, that a joint test of all the EHs would reveal very strong evidence against the hypotheses in both currency markets and all three term structures.

When empirical critical values are used, the evidence against the EHs weakens considerably. In fact, all tests fail to reject at the 1% marginal level of significance, and most of the joint evidence yields (marginal) 5% rejections. Given the LM test's superior size properties, a researcher using such a test, even with the asymptotic critical values, would typically reach the right conclusion.

Note that using the appropriate empirical critical values for the different tests generally leads to a more common interpretation of the data across the tests than is afforded from the asymptotic distributions. For example, in the joint test of all the EHs's in the DEM-USD system, which is reported in Table 4, the Wald, DM, and LM test statistics are 47.76, 27.94, and 21.34, respectively. Since the 5% critical value of a $\chi^2(15)$ is 25, asymptotic inference is quite different depending on the statistic chosen. The results in Table 4 indicate that each of the statistics is quite close to the 95% quantiles of the empirical distributions, which are 47.46, 27.98, and 22.15, respectively.

VI. Conclusions

The goal of this paper is to evaluate the Expectations Hypotheses of the term structure of interest rates and of the foreign exchange market using alternative statistical techniques and extensive Monte Carlo methods. We find no evidence against the EH-FX for the DEM/USD foreign exchange market, but we marginally reject it for the GBP/USD market at either the 5% or 10% marginal level of significance depending on the test statistic. The lack of strong evidence against the EH-FX for these major currencies is consistent with the findings of Huisman, Koedijk, Kool, and Nissen (1998) and Bansal and Dahlquist (1999) who use panel data techniques with fifteen and twenty-eight countries, respectively.

For the EH-TS, the evidence is more mixed. There is no evidence at all against the EH-TS for the GBP, weak evidence against the EH-TS for the USD (at most 5% rejections) and somewhat stronger evidence against the EH-TS for the DEM, where the DM test rejects at the 1% level for both DGPs. However, the other tests reject at the 5% or 10% level depending on the DGP. The joint tests of the EHs never reject at the 1% level and the strongest evidence against the joint hypotheses occurs in the GBP-USD bootstrap system, where the Wald and DM tests reject at the 5% level and the LM test at the 10% level.

These rejections are much less dramatic than the asymptotic distributions imply. In general, we find severe size distortions in the Wald tests and to a lesser extent in Distance Metric tests. The test with the best performance for our sample size is the Lagrange Multiplier test. While estimation of VARs subject to highly non-linear restrictions is often technically demanding, we find that iterating on the approximate solution of Newey and McFadden (1994) easily converged to estimators that satisfied the constraints.

This technique is not only useful in formulating alternative test statistics to the usual Wald tests, it also delivers the dynamics of the data under the null hypothesis. This allows the straightforward development of Monte Carlo experiments to analyze the small-sample distributions of test statistics. There are also many environments in which contrasting constrained with unconstrained dynamics can yield useful insights. As one example, consider the effect of monetary policy on the aggregate economy. Policy analysis, such as Bernanke, Gertler and Watson (1997), often uses VARs to trace out these effects. If some of the effects occur through changes in long rates, it may be instructive to compare the predictions of models estimated under the EH with unconstrained VAR dynamics, especially since the EH is a working hypothesis of many policy makers.

While the distortions in the test statistics provide a partial rehabilitation to the EHs, it remains inconsistent with the data. Moreover, our results cannot be generalized to other currencies. There are several possible ways to go in explaining the findings. First, it is unlikely that the EHs are literally true because of the requirement that risk premiums are constant. Indeed, Bekaert, Hodrick, and Marshall (1999) find that allowing for a small amount of variation in term premiums in the bond market improves the ability of the EH-TS test statistics to match the data. Second, although we allow for a rich data generating process, it may be that the real world is more complicated than this and that peso problems may consequently plague the statistical analysis. Once again, Baillie and Bollerslev (1998), Bekaert, Hodrick and Marshall (1999), and others have experimented with alternative DGPs that may provide richer and more realistic environments than our constrained VARs.

References

Baillie, R. T. and T. Bollerslev, 1998, "The Forward Premium Anomaly Is Not as Bad as You Think," manuscript, University of Virginia.

Bansal, R. and M. Dahlquist, 1999, "The Forward Premium Puzzle: Different Tales from Developed and Emerging Economies," manuscript, Duke University.

Bekaert, G. and C. R. Harvey, 1997, "Emerging Equity Market Volatility," *Journal of Financial Economics* 43, 29-77.

Bekaert, G. and R. J. Hodrick, 1993, "On Biases in the Measurement of Foreign Exchange Risk Premiums," *Journal of International Money and Finance* 12, 115-138.

Bekaert, G., R. J. Hodrick, and D. Marshall, 1997, "On Biases in Tests of the Expectations Hypothesis of the Term Structure of Interest Rates," *Journal of Financial Economics* 44, 309-348.

Bekaert, G., R. J. Hodrick, and D. Marshall, 1999, "Peso Problem' Explanations for Term Structure Anomalies," revision of NBER Working Paper No. 6147.

Berkowitz, J. and L. Kilian, 1996, "Recent Developments in Bootstrapping Time Series," manuscript, Federal Reserve Board.

Bernanke, B. S., M. Gertler, and M. Watson, 1997, "Systematic Monetary Policy and the Effects of Oil Price Shocks," *Brookings Papers on Economic Activity* 1, 91-142.

Bollerslev, T., 1986, "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics* 31, 307-328.

Bollerslev, T. and J. M. Wooldridge, 1992, "Quasi-Maximum Likelihood Estimation and Inference in Dynamic Models with Time Varying Covariances," *Econometric Reviews* 11, 143-172.

Burnside, C. and M. Eichenbaum, 1996, "Small-Sample Properties of GMM-Based Wald Tests," *Journal of Business and Economic Statistics* 14, 294-308.

Campbell, J. Y., and R. J. Shiller, 1991, "Yield Spreads and Interest Rate Movements: A Bird's Eye View," *Review of Economic Studies* 58, 495-514.

Cumby, R. E. and J. Huizinga, 1992, "Testing the Autocorrelation Structure of Disturbances in Ordinary Least Squares and Instrumental Variables Regressions," *Econometrica* 60, 185-196.

Dahlquist, M. and G. Jonsson, 1995, "The Information in Swedish Short-maturity Forward Rates," *European Economic Review* 39, 1115-1131.

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., V. M. Ng, and M. Rothschild, 1990, "Asset Pricing with Factor-ARCH Structure: Empirical Estimation for Treasury Bills," *Journal of Econometrics* 45, 213-237.

Evans, M. D. D., 1996, "Peso Problems: Their Theoretical and Empirical Implications," in G.S. Maddala and C.R. Rao, eds., *Handbook of Statistics, Vol. 14*, Elsevier Science, 613-646.

Fisher, I., 1930, *The Theory of Interest*, New York: Macmillan.

Gerlach, S. and F. Smets, 1997, "The Term Structure of Euro-Rates: Some Evidence In Support of the Expectations Hypothesis," *Journal of International Money and Finance* 16, 305-321.

Gray, S. F., 1996, "Modeling The Conditional Distribution Of Interest Rates As A Regime-switching Process," *Journal of Financial Economics* 42, 27-62.

Hansen, L. P., 1982, "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica* 50, 1029-1054.

Hardouvelis, G. A., 1994, "The Term Structure Spread and Future Changes in Long and Short Rates in the G7 Countries: Is There a Puzzle?" *Journal of Monetary Economics* 33, 255-284.

Hicks, 1953, *Value and Capital*, London: Oxford University Press.

Hodrick, R. J., 1987, *The Empirical Evidence on the Efficiency of Forward and Futures Foreign Exchange Markets*. Harwood Academic Publishers, Chur, Switzerland.

Huisman, R., K. Koedijk, C. Kool, and F. Nissen, 1998, "Extreme Support for Uncovered Interest Parity," *Journal of International Money and Finance17*, 211-228.

Kendall, M. G., 1954, "Note on the Bias in the Estimation of Autocorrelation," *Biometrika* 41, 403-404.

Keynes, J. M., 1930, *A Treatise on Money*. London: Macmillan.

Lewis, K. K., 1989, "Changing Beliefs and Systematic Rational Forecast Errors with Evidence from Foreign Exchange" *American Economic Review* 79, 621-636.

Ligeralde, A. V., 1997, "Covariance Matrix Estimators And Tests Of Market Efficiency," *Journal of International Money and Finance* 16, 323-343.

Mankiw, N. G. and M. Shapiro, 1986, "Do We Reject Too Often: Small Sample Properties of Rational Expectations Models," *Economics Letters* 20, 139-145.

Maynard, A. and P. C. B. Phillips, 1998, "Rethinking an Old Empirical Puzzle: Econometric Evidence on the Forward Discoun Anomaly," manuscript, Yale University.

Melino, A., 1983, "Essays on Estimation and Inference in Linear Rational Expectations Models," Ph.D. dissertation, Harvard University, Cambridge, MA.

Newey, W. K. and D. L. McFadden, 1994, "Large Sample Estimation and Hypothesis Testing," in R. F. Engle and D. L. McFadden, eds., *Handbook of Econometrics: Volume 4.* Elsevier Science: Amsterdam, The Netherlands, 2111-2245.

Politis, D. N. and J. P. Romano, 1994, "The Stationary Bootstrap," *Journal of the American Statistical Association* 89, 1303-1313.

Politis, D. N., J. P. Romano, and M. Wolf, 1997, "Subsampling for Heteroskedastic Time Series," *Journal of Econometrics* 81, 281-317.

Roll, R. and S. Yan, 1998, "An Explanation of the Forward Premium Puzzle'," UCLA CIBER Working Paper No. 98-13.

Sargent, T. J., 1979, "A Note on Maximum Likelihood Estimation of the Rational Expectations Model of the Term Structure," *Journal of Monetary Economics* 5, 133-144.

Schotman , P., 1996, "Small Sample Properties of the Regression Test of the Expectations Hypothesis of the Term Structure," *Economics Letters* 57, 129-134.

Stambaugh, R., 1999, "Predictive Regressions," *Journal of Financial Economics* 54, 375-421.

Tauchen, G., 1985, "The Small-Sample Bias of Tests for a Risk Premium in Forward Exchange Rates," manuscript, Duke University.

Valkanov, R., 1998, "The Term Structure with Highly Persistent Interest Rates," manuscript, Princeton University.

Vasicek, O., 1977, "An Equilibrium Characterization of the Term Structure," *Journal of Financial Economics* 5, 177-188.

White, H., 1982, "Maximum Likelihood Estimation Of Misspecified Models," *Econometrica* 50, 1-26.

Appendix

This Appendix examines the implications of economies that do not admit arbitrage for the expectations hypothesis of the term structure of interest rates. From equation (3) of the paper, the n-period interest rate can be written as follows:

$$i_{t,n} = -\frac{1}{n}\log\left[E_t[\exp(m_{t+n,n})]\right] \tag{55}$$

where the log of the n-period pricing kernel is $m_{t+n,n} \equiv \sum_{i=1}^{n} m_{t+i}$. A Taylor's Series expansion of $exp(m_{t+n,n})$ around the mean yields the following expression:

$$\exp\left(m_{t+n,n}\right) = \exp[E_t\left(m_{t+n,n}\right)]\left[1+\sum_{p=1}^{\infty}\frac{[m_{t+n,n}-E_t(m_{t+n,n})]^p}{p!}\right]. \tag{56}$$

Therefore,

$$E_t[\exp\left(m_{t+n,n}\right)] = \exp[E_t\left(m_{t+n,n}\right)]\left[1+\sum_{p=2}^{\infty}\frac{\nu_{t,n}(p)}{p!}\right] \tag{57}$$

where $\nu_{t,n}(p)$ is the p-th conditional central moment of $m_{t+n,n}$.

By applying equation (57) repeatedly for $n = 1$, and replacing interest rates by conditional expectations of pricing kernels as in equation (55), we derive the following general term premium:

$$\alpha_{t,n} = \frac{1}{n}\left[E_t\left(\sum_{i=1}^{n}\log\left[1+\sum_{p=2}^{\infty}\frac{\nu_{t+i-1,1}(p)}{p!}\right]\right) - \log\left(1+\sum_{p=2}^{\infty}\frac{\nu_{t,n}(p)}{p!}\right)\right]. \tag{58}$$

To compute the foreign exchange risk premium, use the complete markets assumption to express exchange rate changes in terms of conditional pricing kernels and then use the results in equations (55) through (57) to obtain:

$$\begin{aligned} E_t\left(\Delta s_{t+1}\right) &= E_t\left(m_{t+1}^{j}\right) - E_t\left(m_{t+1}^{k}\right) = \\ &= \left(i_t^k - i_t^j\right) + \log\left[1+\sum_{p=2}^{\infty}\frac{\nu_{t,1}^{k}(p)}{p!}\right] - \log\left[1+\sum_{p=2}^{\infty}\frac{\nu_{t,1}^{j}(p)}{p!}\right] \end{aligned} \tag{59}$$

Under log-normality, the last two terms reduce to half the difference of the conditional variances of the two log pricing kernels, as in equation (9) in the paper.

Table 1: Summary Statistics

The sample contains 270 monthly observations from January 1975 to July 1997. The currencies are numbered 1 for the USD, 2 for the DEM, 3 for the GBP. The continously compounded rates of appreciation of the USD versus currency j are denoted Δs_t^j . The short-term interest rate for currency j is i_t^j . The spread between the twelve-month rate and the one-month rate for currency j is sp_t^j .

Variable	Mean	Stand. Dev.	Minimum	Maximum	Autocorrelations 1	2	3
Δs_t^2	-1.084	40.617	-120.068	132.234	-0.019	0.105	0.03
Δs_t^3	1.661	39.987	-163.297	157.455	0.083	0.036	-0.013
i_t^1	7.943	3.52	3.059	20.081	0.968	0.932	0.899
i_t^2	6.011	2.445	2.248	14.907	0.975	0.96	0.945
i_t^3	10.652	3.373	4.647	20.204	0.957	0.92	0.88
sp_t^1	0.164	0.925	-4.882	1.823	0.817	0.673	0.554
sp_t^2	0.071	0.683	-3.24	2.666	0.805	0.722	0.666
sp_t^3	-0.429	1.106	-4.273	1.898	0.817	0.707	0.638

Table 2: Empirical Distributions of the Regression Coefficients

under the EH Null with Bootstrap Innovations

The Table provides summary statistics for the empirical distributions generated from a constrained VAR with a bootstrap of the residuals. The summary statistics are the Mean, Median, Standard Deviation (Std. Dev.) and the 0.5%, 2.5%, and 5% quantiles. The statistics are the slope coefficients in regression tests. An R indicates the direct regression, and an I indicates an implied regression from a VAR. FX1 and FX12 are the one-month and twelve-month EH-FX tests as in equation (17). CUR1 and CUR2 refer to EH-TS tests as in equations (18) and (19), where CUR signifies either USD, DEM, or GBP interest rates. The point estimate is Sample Stat., and the asymptotic standard error is Asymp. S.E.

Panel A: DEM-USD VAR								
Slope	Mean	Median	Std. Dev.	0.5%	2.5%	5%	Sample Stat.	Asymp. s.e.
FX1-R	1.009	0.991	0.926	-1.503	-0.793	-0.470	-0.527	0.923
FX1-I	1.011	0.997	0.955	-1.543	-0.808	-0.490	-0.498	0.979
FX12-R	0.924	0.922	1.022	-1.979	-1.103	-0.737	-0.273	0.467
FX12-I	0.941	0.938	0.939	-1.447	-0.806	-0.500	-0.729	0.982
USD1-R	1.073	1.070	0.194	0.591	0.699	0.756	0.466	0.142
USD1-I	1.025	1.031	0.440	0.590	0.729	0.783	0.664	0.280
USD2-R	1.505	1.493	0.516	0.183	0.510	0.682	0.237	0.929
USD2-I	1.488	1.479	0.782	0.155	0.505	0.677	0.271	0.777
DEM1-R	1.028	1.028	0.143	0.658	0.748	0.793	0.558	0.151
DEM1-I	0.990	0.993	0.193	0.673	0.762	0.803	0.669	0.148
DEM2-R	1.507	1.513	0.361	0.488	0.774	0.903	0.146	0.368
DEM2-I	1.495	1.504	0.415	0.480	0.750	0.886	0.193	0.339

Panel B: GBP-USD VAR								
Slope	Mean	Median	Std. Dev.	0.5%	2.5%	5%	Sample Stat.	Asymp. s.e.
FX1-R	1.052	1.045	0.673	-0.778	-0.261	-0.035	-1.654	0.911
FX1-I	1.050	1.045	0.680	-0.801	-0.277	-0.048	-1.662	0.936
FX12-R	0.985	0.990	0.741	-1.107	-0.507	-0.231	-0.867	0.651
FX12-I	0.986	0.992	0.644	-0.784	-0.293	-0.081	-1.341	0.942
USD1-R	1.059	1.057	0.197	0.556	0.679	0.737	0.466	0.142
USD1-I	1.018	1.010	1.567	0.557	0.710	0.766	0.624	0.348
USD2-R	1.512	1.502	0.509	0.163	0.530	0.698	0.237	0.929
USD2-I	1.500	1.489	0.954	0.124	0.503	0.672	0.255	0.707
GBP1-R	1.054	1.052	0.174	0.608	0.720	0.774	0.839	0.176
GBP1-I	1.007	1.009	0.289	0.622	0.739	0.786	0.817	0.241
GBP2-R	1.546	1.536	0.425	0.462	0.730	0.859	0.938	0.483
GBP2-I	1.535	1.525	0.580	0.415	0.705	0.840	0.928	0.467

Table 3: Empirical Distributions of the Regression Coefficients

under the EH Null with GARCH Innovations

The Table provides summary statistics for the empirical distributions generated from a constrained VAR with a GARCH model of the residuals. The summary statistics are the Mean, Median, Standard Deviation (Std. Dev.) and the 0.5%, 2.5%, and 5% quantiles. The statistics are the slope coefficients in regression tests. An R indicates the direct regression, and an I indicates an implied regression from a VAR. FX1 and FX12 are the one-month and twelve-month EH-FX tests in the foreign exchange market as in equation (17). CUR1 and CUR2 refer to EH-TS tests as in equations (18) and (19), where CUR signifies either USD, DEM, or GBP interest rates. The point estimate is Sample Stat., and the asymptotic standard error is Asymp. S.E.

Panel A: DEM-USD VAR

Slope	Mean	Median	Std. Dev.	0.5%	2.5%	5%	Sample Stat.	Asymp. s.e.
FX1-R	0.998	0.998	0.226	0.359	0.545	0.631	-0.527	0.923
FX1-I	0.998	0.997	0.229	0.343	0.539	0.626	-0.498	0.979
FX12-R	0.962	0.971	0.259	0.173	0.416	0.522	-0.273	0.467
FX12-I	0.965	0.972	0.246	0.224	0.445	0.544	-0.729	0.982
USD1-R	1.053	1.058	0.131	0.703	0.790	0.833	0.466	0.142
USD1-I	1.047	1.068	1.991	0.679	0.807	0.855	0.664	0.280
USD2-R	1.421	1.398	0.343	0.633	0.807	0.898	0.237	0.929
USD2-I	1.405	1.397	2.440	0.585	0.796	0.888	0.271	0.777
DEM1-R	1.116	1.114	0.242	0.504	0.647	0.719	0.558	0.151
DEM1-I	1.066	1.085	1.531	0.416	0.680	0.757	0.669	0.148
DEM2-R	1.626	1.603	0.658	0.029	0.391	0.585	0.146	0.368
DEM2-I	1.589	1.589	2.548	-0.058	0.361	0.558	0.193	0.339

Panel B: GBP-USD VAR								
Slope	Mean	Median	Std. Dev.	0.5%	2.5%	5%	Sample Stat.	Asymp. s.e.
FX1-R	1.002	1.003	0.288	0.227	0.430	0.530	-1.654	0.911
FX1-I	1.002	1.003	0.290	0.228	0.429	0.527	-1.662	0.936
FX12-R	0.898	0.910	0.546	-0.645	-0.216	-0.018	-0.867	0.651
FX12-I	0.914	0.920	0.458	-0.320	0.005	0.151	-1.341	0.942
USD1-R	1.071	1.072	0.180	0.597	0.715	0.772	0.466	0.142
USD1-I	1.024	1.030	0.236	0.621	0.736	0.790	0.624	0.348
USD2-R	1.579	1.554	0.454	0.521	0.761	0.877	0.237	0.929
USD2-I	1.573	1.549	0.553	0.507	0.753	0.872	0.255	0.707
GBP1-R	1.021	1.024	0.083	0.798	0.851	0.879	0.839	0.176
GBP1-I	1.010	1.015	0.135	0.828	0.876	0.902	0.817	0.241
GBP2-R	1.344	1.336	0.164	0.966	1.046	1.089	0.938	0.483
GBP2-I	1.342	1.334	0.232	0.959	1.044	1.087	0.928	0.467

Table 4: Empirical Distributions of Wald, Lagrange Multiplier and Distance Metric Tests

under the EH Null with Bootstrap Innovations

The Table provides summary statistics for the empirical distributions generated from a constrained VAR with a bootstrap of the residuals. The summary statistics are the Mean, Median, Standard Deviation (Std. Dev.) and the 90%, 95%, and 99% quantiles. The statistics are the Wald (W), Lagrange Multiplier (LM), and Distance Metric (DM) tests. FX1 and FX12 are the one-month and twelve-month EH-FX tests. The asymptotic distribution is a $P^2(5)$. FX1-12 examines the one-month and twelve-month joint EH-FX test. The asymptotic distribution is a $P^2(10)$. The EH-TS tests are labelled by currency. The asymptotic distribution is a $P^2(5)$. Joint-EH is a simultaneous test of the restrictions of the EH-FX and the EH-TS in each currency. The asymptotic distribution is a $P^2(15)$. The currencies are the USD, the DEM, and the GBP. The sample statistic is Sample Stat., and its asymptotic p-value is Asymp. p-value.

Panel A: DEM-USD VAR								
	Mean	Median	Std. Dev.	90%	95%	99%	Sample Stat.	Asymp p-value
$P^2(5)$	5.00	4.35	3.16	9.24	11.07	15.09		
FX1 W	6.23	5.30	4.23	11.67	14.35	20.18	6.953	0.224
FX1 LM	5.08	4.59	2.91	8.98	10.61	13.73	5.666	0.340
FX1 DM	6.19	5.32	4.06	11.67	14.16	19.40	8.770	0.119
FX12 W	6.20	5.23	4.35	11.66	14.37	21.57	6.740	0.241
FX12 LM	5.07	4.57	2.91	8.97	10.60	13.81	6.093	0.297
FX12 DM	6.13	5.28	4.01	11.51	13.95	18.92	9.993	0.079
$P^2(10)$	10.00	9.34	4.47	15.99	18.31	23.21		
FX1-12 W	11.09	9.09	9.99	18.57	23.93	41.60	8.063	0.623
FX1-12 LM	10.38	10.02	3.82	15.51	17.16	20.55	9.133	0.520
FX1-12 DM	12.82	12.11	5.51	20.14	22.53	27.69	12.110	0.278
$P^2(5)$	5.00	4.35	3.16	9.24	11.07	15.09		
USD W	7.98	6.55	5.91	15.17	18.91	29.20	14.898	0.011
USD LM	5.65	5.20	2.98	9.72	11.24	14.00	12.627	0.027
USD DM	6.80	5.99	4.10	12.31	14.60	19.52	18.134	0.003
DEM W	7.28	6.01	5.29	13.83	17.32	26.24	19.890	0.001
DEM LM	5.37	4.93	2.92	9.33	10.79	13.80	10.971	0.052
DEM DM	6.76	5.88	4.51	12.48	14.89	20.20	24.498	0.000
$P^2(15)$	15.00	14.34	5.48	22.31	25.00	30.58		
Joint-EH W	25.63	23.54	11.76	40.56	47.46	63.66	47.758	0.000

Joint-EH LM	14.96	14.76	4.16	20.41	22.15	25.35	21.336	0.126
Joint-EH DM	18.19	17.77	5.91	25.62	27.98	32.97	27.937	0.022

Panel B: GBP-USD VAR

	Mean	Median	Std. Dev.	90%	95%	99%	Sample Stat.	Asymp p-value
$P^2(5)$	5.00	4.35	3.16	9.24	11.07	15.09		
FX1 W	6.30	5.38	4.21	11.79	14.35	20.58	17.664	0.003
FX1 LM	4.98	4.55	2.76	8.70	10.15	13.00	9.340	0.096
FX1 DM	6.18	5.37	3.94	11.45	13.76	18.71	15.678	0.005
FX12 W	6.24	5.26	4.29	11.75	14.39	20.80	16.622	0.005
FX12 LM	4.98	4.56	2.77	8.74	10.09	13.09	9.753	0.083
FX12 DM	6.15	5.35	3.92	11.39	13.62	18.58	15.050	0.010
$P^2(10)$	10.00	9.34	4.47	15.99	18.31	23.21		
FX1-12 W	13.08	10.48	13.64	21.74	28.16	55.84	23.728	0.008
FX1-12 LM	10.34	10.04	3.70	15.30	16.93	19.89	14.962	0.133
FX1-12 DM	12.89	12.33	5.24	19.99	22.43	27.07	26.197	0.003
$P^2(5)$	5.00	4.35	3.16	9.24	11.07	15.09		
USD W	8.00	6.61	5.85	15.07	18.87	28.51	13.247	0.021
USD LM	5.61	5.20	2.89	9.53	10.97	13.58	7.813	0.167
USD DM	6.95	6.14	4.15	12.57	14.87	19.86	13.154	0.022
GBP W	7.96	6.38	6.21	15.40	19.66	30.91	4.666	0.458
GBP LM	5.54	5.08	3.03	9.67	11.23	14.38	4.086	0.537
GBP DM	6.86	5.99	4.32	12.71	15.21	20.53	6.032	0.303
$P^2(15)$	15.00	14.34	5.48	22.31	25.00	30.58		
Joint-EH W	27.65	25.15	13.22	44.32	52.50	71.86	55.511	0.000
Joint-EH LM	14.81	14.64	3.93	20.01	21.57	24.43	20.109	0.168
Joint-EH DM	18.04	17.72	5.30	25.18	27.34	31.16	27.937	0.022

Table 5: Empirical Distributions of Wald, Lagrange Multiplier and Distance Metric Tests

under the EH Null with GARCH Innovations

The Table provides summary statistics for the empirical distributions generated from a constrained VAR with a GARCH model of the residuals. The summary statistics are the Mean, Median, Standard Deviation (Std. Dev.) and the 90%, 95%, and 99% quantiles. The statistics are the Wald (W), Lagrange Multiplier (LM), and Distance Metric (DM) tests. FX1 and FX12 are the one-month and twelve-month EH-FX tests. The asymptotic distribution is a $\chi^2(5)$. FX1-12 examines the one-month and twelve-month joint EH-FX test. The asymptotic distribution is a $\chi^2(10)$. The EH-TS tests are labelled by currency. The asymptotic distribution is a $\chi^2(5)$. Joint-EH is a simultaneous test of the restrictions of the EH-FX and the EH-TS in each currency. The asymptotic distribution is a $\chi^2(15)$. The currencies are the USD, the DEM, and the GBP. The point estimate is Sample Stat., and its asymptotic p-value is Asymp. p-value.

Panel A: DEM-USD VAR								
	Mean	Median	Std. Dev.	90%	95%	99%	Sample Stat.	Asymp. p-value
$\chi^2(5)$	5.00	4.35	3.16	9.24	11.07	15.09		
FX1 W	6.05	5.16	4.04	11.47	13.87	19.32	6.953	0.224
FX1 LM	5.02	4.52	2.88	8.98	10.54	13.68	5.666	0.340
FX1 DM	5.94	5.12	3.83	11.11	13.39	18.25	8.770	0.119
FX12 W	6.14	5.10	4.34	11.75	14.47	21.09	6.740	0.241
FX12 LM	5.06	4.55	2.92	9.08	10.59	13.84	6.093	0.297
FX12 DM	6.00	5.17	3.89	11.32	13.60	18.40	9.993	0.079
$\chi^2(10)$	10.00	9.34	4.47	15.99	18.31	23.21		
FX1-12 W	15.87	12.36	15.81	27.75	37.03	68.33	8.063	0.623
FX1-12 LM	10.34	9.96	3.96	15.66	17.45	21.12	9.133	0.520
FX1-12 DM	12.33	11.67	5.25	19.44	21.96	27.15	12.110	0.278
$\chi^2(5)$	5.00	4.35	3.16	9.24	11.07	15.09		
USD W	9.05	6.79	8.26	17.93	23.98	40.39	14.898	0.011
USD LM	5.91	5.36	3.30	10.39	12.13	15.73	12.627	0.027
USD DM	7.10	6.17	4.50	13.13	15.73	21.50	18.134	0.003
DEM W	8.71	6.71	7.40	17.29	22.22	36.63	19.890	0.001
DEM LM	5.93	5.37	3.30	10.47	12.16	15.64	10.971	0.052
DEM DM	7.16	6.20	4.50	13.27	15.90	21.52	24.498	0.000
$\chi^2(15)$	15.00	14.34	5.48	22.31	25.00	30.58		
Joint-EH W	27.47	24.29	14.68	45.29	54.27	79.32	47.758	0.000

Joint-EH LM	15.42	15.17	4.45	21.36	23.24	26.62	21.336	0.126
Joint-EH DM	18.25	17.82	5.77	25.99	28.50	33.02	27.937	0.022

Panel B: GBP-USD VAR

	Mean	Median	Std. Dev.	90%	95%	99%	Sample Stat.	Asymp p-value
$P^2(5)$	5.00	4.35	3.16	9.24	11.07	15.09		
FX1 W	6.25	5.30	4.22	11.78	14.32	20.31	17.664	0.003
FX1 LM	4.97	4.51	2.79	8.72	10.26	13.21	9.340	0.096
FX1 DM	6.05	5.24	3.86	11.27	13.54	18.42	15.678	0.005
FX12 W	6.42	5.30	4.60	12.31	15.29	22.57	16.622	0.005
FX12 LM	5.06	4.61	2.84	8.92	10.43	13.59	9.753	0.083
FX12 DM	6.20	5.37	3.97	11.55	13.86	18.95	15.050	0.010
$P^2(10)$	10.00	9.34	4.47	15.99	18.31	23.21		
FX1-12 W	16.52	13.20	14.77	28.52	37.33	70.07	23.728	0.008
FX1-12 LM	10.25	9.91	3.78	15.28	17.03	20.19	14.962	0.133
FX1-12 DM	12.57	11.96	5.17	19.61	22.04	26.68	26.197	0.003
$P^2(5)$	5.00	4.35	3.16	9.24	11.07	15.09		
USD W	8.23	6.27	7.17	16.34	21.23	36.32	13.247	0.021
USD LM	5.52	5.01	3.10	9.73	11.41	14.66	7.813	0.167
USD DM	6.73	5.86	4.28	12.50	15.05	20.40	13.154	0.022
GBP W	7.89	5.97	6.87	15.89	20.67	33.70	4.666	0.458
GBP LM	5.41	4.85	3.13	9.69	11.39	14.82	4.086	0.537
GBP DM	6.65	5.68	4.35	12.55	15.18	20.56	6.032	0.303
$P^2(15)$	15.00	14.34	5.48	22.31	25.00	30.58		
Joint-EH W	28.22	24.95	15.04	46.73	56.33	80.91	55.511	0.000
Joint-EH LM	14.78	14.59	4.13	20.26	21.95	25.08	20.109	0.168
Joint-EH DM	17.76	17.45	5.41	25.01	27.17	31.21	27.937	0.022

Table 6: Empirical Size and Power of

Wald, Lagrange Multiplier and Distance Metric Tests

The Table provides empirical sizes and powers from the empirical distributions of various test statistics. The empirical size is the percent of the Monte Carlo experiments generated when the null hypothesis is true in which the test statistic exceeds the 5% asymptotic critical value. The power of the test is the percent of the empirical distribution generated when the alternative hypothesis is true that exceeds the 5% critical value of the empirical distribution generated when the null hypothesis is true. The statistics are the Wald (W), Lagrange Multiplier (LM), and Distance Metric (DM) tests. The symbol B signifies the bootstrap DGP distributions, and the symbol G signifies the GARCH model DGP distributions. FX-1 and FX-12 are the one-month and twelve-month EH-FX tests. Joint FX is the simultaneous test of both horizons. The EH-TS tests are labeled by currency. Joint is a simultaneous test of the restrictions of the EH-FX and the two EH-TS.

	DEM-USD						GBP-USD					
	FX-1	FX-12	Joint FX	EH-USD	EH-DEM	Joint	FX-1	FX-12	Joint FX	EH-USD	EH-GBP	Joint
Panel A: Empirical Size												
W-B	11.8	11.5	10.5	21.9	17.8	44.6	12.1	12.0	15.6	21.8	22.0	50.6
LM-B	4.1	4.0	3.0	5.4	4.4	1.2	3.2	3.1	2.5	4.8	5.3	0.7
DM-B	11.6	11.3	15.2	14.3	14.7	11.9	11.2	11.1	15.4	15.1	15.4	10.6
W-G	11.1	11.9	25.9	26.2	24.9	47.7	12.2	13.4	28.5	22.6	21.2	49.9
LM-G	3.9	4.0	3.5	7.7	7.9	2.3	3.3	3.6	2.9	5.8	5.7	1.1
DM-G	10.2	10.8	13.2	16.8	17.4	12.8	10.6	11.7	13.7	14.5	14.7	10.0
Panel B: Empirical Power												
W-B	55.8	49.7	45.1	80.1	95.3	95.8	77.8	72.7	60.2	69.1	45.6	91.8
LM-B	54.1	51.1	52.0	84.4	95.5	96.1	64.0	70.9	79.5	72.5	47.6	91.8
DM-B	55.2	52.8	54.1	88.2	97.7	98.9	72.0	78.1	85.5	76.1	49.5	95.3
W-G	57.8	54.8	12.5	95.6	74.3	98.6	67.8	61.9	31.2	95.3	94.1	99.8
LM-G	56.2	57.6	53.0	97.9	83.7	99.0	56.5	59.8	99.1	96.9	95.4	99.9
DM-G	58.0	58.7	54.2	98.4	85.4	99.5	61.8	66.8	99.4	97.4	95.7	99.9

Table A1: VAR Order Tests

Panel A: Schwarz Criteria				
VAR	Lag 1	Lag 2	Lag 3	Lag 4
DEM-USD	1.677	1.805	2.062	2.414
GBP-USD	3.87	4.176	4.571	4.823

Panel B: Cumby-Huizinga l-tests p-values			
	Lag 1	Lags 1-3	Lags 1-6
VAR DEM-USD			
Δs_t^2	0.798	0.459	0.625
i_t^1	0.321	0.536	0.187
i_t^2	0.703	0.868	0.31
sp_t^1	0.049	0.228	0.188
sp_t^2	0.024	0.082	0.107
VAR GBP-USD			
Δs_t^3	0.443	0.637	0.377
i_t^1	0.306	0.598	0.217
i_t^3	0.805	0.987	0.148
sp_t^1	0.687	0.766	0.799
sp_t^3	0.358	0.726	0.811

Table A2: Unconstrained VAR Dynamics with OLS and Bias-Corrected Coefficients

Panel A: DEM-USD	Coef. Δs_{t-1}^2 Bias-Corrected (s.e.)	Coef. i_{t-1}^1 Bias-Corrected (s.e.)	Coef. i_{t-1}^2 Bias-Corrected (s.e.)	Coef. sp_t^1 Bias-Corrected (s.e.)	Coef. sp_t^2 Bias-Corrected (s.e.)
Δs_t^2	-0.026	1.706	-0.518	3.812	-6.431
	-0.011	1.667	-0.020	3.837	-5.936
	(0.075)	(1.277)	(1.692)	(5.629)	(4.459)
i_t^1	0.000	1.030	-0.035	0.287	-0.175
	0.000	1.039	-0.020	0.273	-0.165
	(0.001)	(0.036)	(0.024)	(0.164)	(0.123)
i_t^2	0.002	-0.021	1.001	-0.209	0.312
	0.002	-0.023	1.012	-0.207	0.303
	(0.001)	(0.011)	(0.012)	(0.071)	(0.073)
sp_t^1	0.000	-0.036	-0.015	0.699	0.019
	0.000	-0.033	-0.018	0.725	0.023
	(0.001)	(0.018)	(0.015)	(0.066)	(0.056)
sp_t^2	-0.001	0.032	-0.038	0.152	0.655
	-0.001	0.029	-0.033	0.147	0.684
	(0.001)	(0.011)	(0.012)	(0.047)	(0.063)

Panel B: GBP-USD	Coef. Δs_{t-1}^3 Bias-Corrected (s.e.)	Coef. i_{t-1}^1 Bias-Corrected (s.e.)	Coef. i_{t-1}^3 Bias-Corrected (s.e.)	Coef. sp_t^1 Bias-Corrected (s.e.)	Coef. sp_t^3 Bias-Corrected (s.e.)
Δs_t^3	0.060	1.524	-0.390	1.405	4.709
	0.077	1.555	-0.273	1.486	4.625
	(0.080)	(1.441)	(1.151)	(4.288)	(3.498)
i_t^1	-0.002	1.023	-0.019	0.272	-0.072
	-0.002	1.042	-0.019	0.266	-0.075
	(0.001)	(0.038)	(0.023)	(0.150)	(0.103)
i_t^3	0.001	0.025	1.022	-0.078	0.370
	0.001	0.017	1.035	-0.086	0.361
	(0.001)	(0.034)	(0.041)	(0.085)	(0.106)
sp_t^1	0.001	-0.042	0.010	0.693	0.060
	0.001	-0.042	0.013	0.718	0.064
	(0.001)	(0.020)	(0.013)	(0.059)	(0.047)
sp_t^3	-0.001	0.027	-0.081	0.184	0.602
	-0.001	0.028	-0.075	0.183	0.630
	(0.001)	(0.018)	(0.022)	(0.065)	(0.061)

Table A3: Constrained VAR Dynamics with OLS and Bias-Corrected Coefficients

Panel A: DEM-USD					
	Coef. Δs_{t-1}^2 Bias-Corrected (s.e.)	Coef. i_{t-1}^1 Bias-Corrected (s.e.)	Coef. i_{t-1}^2 Bias-Corrected (s.e.)	Coef. sp_t^1 Bias-Corrected (s.e.)	Coef. sp_t^2 Bias-Corrected (s.e.)
Δs_t^2	0.000 0.000 (0.076)	-1.000 -1.000 (1.357)	1.000 1.000 (1.789)	0.000 0.000 (5.986)	0.000 0.000 (4.618)
i_t^1	-0.001 0.000 (0.001)	1.034 1.022 (0.038)	0.007 0.020 (0.028)	0.475 0.401 (0.179)	-0.021 0.018 (0.130)
i_t^2	0.000 0.001 (0.001)	-0.031 -0.021 (0.011)	1.014 1.109 (0.012)	-0.152 -0.139 (0.075)	0.393 0.414 (0.072)
sp_t^1	0.001 0.000 (0.001)	-0.042 -0.029 (0.019)	-0.009 -0.025 (0.017)	0.626 0.707 (0.068)	0.022 -0.025 (0.059)
sp_t^2	0.000 -0.001 (0.001)	0.037 0.025 (0.011)	-0.019 -0.027 (0.011)	0.175 0.160 (0.051)	0.720 0.692 (0.063)
GBP-USD	Coef. Δs_{t-1}^3 Bias-Corrected (s.e.)	Coef. i_{t-1}^1 Bias-Corrected (s.e.)	Coef. i_{t-1}^3 Bias-Corrected (s.e.)	Coef. sp_t^1 Bias-Corrected (s.e.)	Coef. sp_t^3 Bias-Corrected (s.e.)
Δs_t^3	0.000 0.000 (0.077)	-1.000 -1.000 (1.664)	1.000 1.000 (1.370)	0.000 0.000 (5.327)	0.000 0.000 (4.026)
i_t^1	-0.001 -0.001 (0.001)	1.026 1.026 (0.039)	-0.008 -0.006 (0.025)	0.413 0.397 (0.158)	-0.065 -0.039 (0.108)
i_t^3	0.001 0.000 (0.001)	-0.024 -0.032 (0.034)	1.066 1.064 (0.041)	-0.176 -0.202 (0.091)	0.486 0.478 (0.111)
sp_t^1	0.001 0.001 (0.001)	-0.033 -0.034 (0.019)	0.009 0.007 (0.013)	0.698 0.714 (0.059)	0.072 0.044 (0.047)
sp_t^3	-0.001 0.000 (0.001)	0.027 0.037 (0.018)	-0.084 -0.081 (0.022)	0.199 0.230 (0.066)	0.609 0.617 (0.061)

Table A4: DEM-USD GARCH Model

	Δs_t^2	i_t^1	i_t^2	sp_t^1	sp_t^2
Δs_t^2	1.0000	0.0019 (0.0006)	0.0014 (0.0004)	0.0000	0.0000
i_t^1	0.0000	1.0000	0.0000	0.0000	0.0000
i_t^2	0.0000	0.4689 (0.0672)	1.0000	0.0000	0.0000
sp_t^1	0.0000	-0.6217 (0.0919)	-0.0377 (0.0494)	1.0000	0.0000
sp_t^2	0.0000	0.1882 (0.0847)	-0.6053 (0.0675)	0.0000	1.0000
	T_j (s.e.)	$\$_j$ (s.e.)	$''_j$ (s.e.)		
$h_{1,t}$	0.1461 (0.0155)	0.00002 (0.00173)	0.0980 (0.0691)		
$h_{2,t}$	0.00006 (0.00002)	0.5355 (0.1182)	0.2222 (0.0761)		
$h_{3,t}$	0.00004 (0.00001)	0.4642 (0.0881)	0.2777 (0.0890)		
$h_{4,t}$	0.00003 (0.000002)	0.0	0.0		
$h_{5,t}$	0.000014 (0.000001)	0.0	0.0		

Table A5: GBP-USD GARCH Model

	Δs_t^3	i_t^1	i_t^3	sp_t^1	sp_t^3
Δs_t^3	1.0000	0.0013 (0.0006)	0.00007 (0.00087)	0.0000	0.0000
i_t^1	0.0000	1.0000	0.0000	0.0000	0.0000
i_t^3	0.0000	0.1289 (0.0376)	1.0000	0.0000	0.0000
sp_t^1	0.0000	-0.5175 (0.0867)	-0.0475 (0.0864)	1.0000	0.0000
sp_t^3	0.0000	0.1302 (0.0514)	-0.8097 (0.0795)	0.0000	1.0000
	T_j (s.e.)	$\$_j$ (s.e.)	$"_j$ (s.e.)		
$h_{1,t}$	0.1243 (0.0143)	0.000000 (0.00023)	0.2004 (0.0914)		
$h_{2,t}$	0.000019 (0.000021)	0.6877 (0.1185)	0.2215 (0.0727)		
$h_{3,t}$	0.000033 (0.000014)	0.7657 (0.0662)	0.0991 (0.0355)		
$h_{4,t}$	0.000027 (0.000002)	0.0	0.0		
$h_{5,t}$	0.000035 (0.000003)	0.0	0.0		