RESEARCH SEMINAR IN INTERNATIONAL ECONOMICS

School of Public Policy
University of Michigan
Ann Arbor, Michigan 48109-1220

Discussion Paper No. 401

Exchange Rates and Monetary Fundamentals:

What Do We Learn From Long-Horizon Regressions?

Lutz Kilian
University of Michigan

May 28, 1997

Recent RSIE Discussion Papers are available on the World Wide Web
at: http://www.spp.umich.edu/rsie/workingpapers/wp.html

Exchange Rates and Monetary Fundamentals:

What Do We Learn From Long-Horizon Regressions?

Lutz Kilian

University of Michigan

First Draft: December 1996
This Version: May 28, 1997

Abstract: Long-horizon regression tests are widely used in empirical finance, despite evidence of severe size distortions. I propose a new bootstrap method for small-sample inference in long-horizon regressions. A Monte Carlo study shows that this bootstrap test greatly reduces the size distortions of conventional long-horizon regression tests. I also find that long-horizon regression tests do not have power advantages against economically plausible alternatives. The apparent lack of higher power at long horizons suggests that previous findings of increasing long-horizon predictability are more likely due to size distortions than to power gains. I illustrate the use of the bootstrap method by analyzing whether monetary fundamentals help predict changes in four major exchange rates. In contrast to earlier studies, I find only weak evidence of exchange rate predictability and no evidence of increasing long-horizon predictability. Many of the differences in results can be traced to the implementation of the test.

JEL Codes: C22, C32, C52, C53, F31, F47
Key Words: empirical finance, long-horizon regression tests, bootstrap, size, power, net present value model, monetary exchange rate model, forecasting.

Acknowledgments: The comments of Jeremy Berkowitz, Robert Barsky, Ufuk Demiroglu, Frank Diebold, Phil Howrey, Lorenzo Giorgianni, Shinichi Sakata, and Matthew Shapiro helped improve the paper. I thank Jeremy Berkowitz, Lorenzo Giorgianni, and Nelson Mark for providing ready access to their data and programs.

Correspondence to: Department of Economics, University of Michigan, Ann Arbor, MI 48109-1220. Fax: (313) 764-2769. Phone: (313) 764-2320. Email: lkilian@econ.lsa.umich.edu

1. Introduction

Long-horizon regression tests are widely used in empirical finance as tests of market efficiency. They have been used for example in exchange rate prediction (e.g., Mark (1995), Chinn and Meese (1995)) in the analysis of dividend yields and expected stock returns (e.g., Fama and French (1988), Campbell and Shiller (1988)) and in studies of the term structure of interest rates (e.g., Fama and Bliss (1987), Cutler, Poterba, and Summers (1991)).[1] In the absence of market efficiency, deviations of asset prices from their long-run equilibrium value should help predict cumulative future asset returns. Regression tests of this hypothesis typically find strong evidence of predictability at long forecast horizons, but cannot reject the null of unpredictable asset returns at short forecast horizons. This finding is often interpreted as evidence of increasing power at higher forecast horizons. However, there exists a large body of literature which questions this interpretation of long-horizon regression test results. For example, Mankiw, Romer, and Shapiro (1991), Hodrick (1992), Nelson and Kim (1993), Bollerslev and Hodrick (1995), and Berkowitz and Giorgianni (1997) have documented that conventional long-horizon regression tests are biased in favor of finding predictability. Severe size distortions may arise from spurious regression fits and from small-sample bias in the estimates of regression coefficients and asymptotic standard errors. Previous attempts to mitigate these size distortions have only been partially successful. In this paper, I propose a new bootstrap method for small-sample inference in long-horizon regressions. I present Monte Carlo evidence which shows that this bootstrap test is indeed fairly accurate in realistic situations. It greatly reduces the size distortions of conventional long-horizon regression tests and is immune from problems of spurious regression fit.

[1] Campbell and Shiller (1988) and Mankiw, Romer, and Shapiro (1991) discuss the close relationship between long-horizon regression tests and volatility tests.

To illustrate the use of this bootstrap method I reexamine the question of whether monetary fundamentals help predict changes in major exchange rates. In recent years, this question has received considerable interest in the international finance literature. I consider both in-sample and out-of-sample test statistics for an extended data set based on Mark (1995). In contrast to Mark (1995) and Chinn and Meese (1995), I find that there is only weak evidence that monetary fundamentals help predict exchange rates and no evidence of increasing long-horizon predictability. I show that many of the differences in results can be traced to the implementation of the test.

While no one believes that the monetary exchange rate model holds period by period, many economists would consider it a reasonable description of the long run. This makes the absence of a pattern of increasing long-horizon predictability in the data surprising. The model predicts that at least in the long run the exchange rate must revert to its equilibrium value. As a result, current deviations from the long-run equilibrium value of the exchange rate contain useful information about future changes of the exchange rate, especially at long forecast horizons. This suggests that the power of long-horizon regression tests ought to increase with the forecast horizon. However, empirical evidence supporting that conjecture has apparently never been presented. This paper makes an effort to fill that gap. The stable and fairly accurate size of the bootstrap test makes it straightforward to evaluate the power of the bootstrap test against economically plausible alternatives. I conduct a Monte Carlo experiment based on data generating processes calibrated to the data used in the empirical study.

The natural alternative against which to test the null hypothesis of no predictability is the vector error correction model implied by the underlying rational expectations model. In addition, I study the power of the bootstrap tests against models in which asset prices contain a highly

serially correlated transitory noise component. This noise component can be thought of as fads in investors' behavior or short-term speculative dynamics. The intuition is that serially correlated noise may obscure the tendency of the economy to revert to equilibrium in the short-run; in the long run, however, investor fads will die out and the true pattern of mean reversion in the exchange rate market will be revealed. Thus, investor fads may provide an alternative rationale for the belief that regression tests have higher power at long horizons. As I will show, the results of these two Monte Carlo studies suggest the need for a reevaluation of the long-horizon regression test methodology with important implications for many areas of empirical finance.

The remainder of the paper is organized as follows. For expository purposes, I present the bootstrap test in the context of the long-run monetary model of exchange rate determination. The discussion can be modified easily to apply to other net present value models (see Campbell and Shiller (1987)). The structure of the model is also identical to the permanent income theory of consumption under rational expectations (e.g., Campbell (1987)). Section 2 contains some useful statistical relationships based on the monetary exchange rate model which underlies the long-horizon regressions in Mark (1995) and Chinn and Meese (1995). In section 3, these relationships are used to motivate the bootstrap algorithm. Section 3 also compares this bootstrap test to earlier bootstrap tests for long-horizon regressions due to Campbell (1993) and Mark (1995). Section 4 presents the empirical findings and explains the differences to Mark's results. Section 5 examines the sensitivity of the results to alternative assumptions about the data generating process. Section 6 analyzes the size and power of the bootstrap long-horizon regression test, and section 7 concludes.

2. The Monetary Model in Vector Error Correction Representation

In the standard long-run monetary model of exchange rate determination it is assumed that purchasing power parity and uncovered interest parity hold. Demand for log real balances is static and linearly related to log real income and the nominal interest rate. Denote the money demand income elasticity by λ and the money-demand interest rate semi-elasticity by ϕ. In the empirical part, λ will be set to 1 following Mark (1995). Further let $\delta \equiv \phi/(1+\phi)$. In the absence of speculative bubbles, the model implies that the log exchange rate for two identical countries is determined by:

$$(1) \qquad e_t = (1-\delta)\ E_t\left(\sum_{j=0}^{\infty} \delta^j\ f_{t+j}\right)$$

where $f_t \equiv (m_t - m_t^*) - \lambda\ (y_t - y_t^*)$ and * denotes the foreign country. Subtracting f_t from both sides and rearranging yields:

$$(2) \qquad e_t - f_t = E_t\left(\sum_{j=1}^{\infty} \delta^j\ \Delta f_{t+j}\right)$$

Provided that f_t is a serially correlated stationary process in first differences, (1) implies that e_t ~ I(1) and by (2) e_t - f_t ~ I(0). Thus, e_t and f_t are cointegrated with cointegrating vector C′ = [1, -1], and f_t may be interpreted as the long-run equilibrium value of the spot exchange rate. The implied joint time series process for e_t and f_t may be represented as a bivariate vector autoregression (VAR) for $y_t = (e_t, f_t)'$:

$$(3) \qquad y_t = \nu + \Phi_1\ y_{t-1} + \ldots + \Phi_p\ y_{t-p} + u_t$$

where u_t is assumed to be iid white noise with vector mean zero and nonsingular covariance matrix Σ_u = E($u_t\ u_t$') and ν is the intercept. Let $z_t \equiv e_t$ - f_t denote the deviation of the spot

exchange rate from its fundamental value. As noted by Berkowitz and Giorgianni (1997), the level VAR may then be rewritten in vector error correction (VEC) form as:

$$\Delta y_t = \nu + \xi_0 y_{t-1} + \xi_1 \Delta y_{t-1} + \ldots + \xi_{p-1} \Delta y_{t-p+1} + u_t \tag{4}$$

where $\xi_0 = -HC'$ is a (2x2) matrix with rank r = 1 and H = $(h_1, h_2)'$ and C = $(c_1, c_2)'$ are (2x1) vectors. Given C′ = [1, -1] we can write:

$$\xi_0 y_{t-1} = -H(C' y_{t-1}) = -H(e_{t-1} - f_{t-1}) = -H z_{t-1}$$

Substituting into (4) we obtain the VEC model:

$$\Delta y_t = \nu - H z_{t-1} + \xi_1 \Delta y_{t-1} + \ldots + \xi_{p-1} \Delta y_{t-p+1} + u_t \tag{4′}$$

which may also be written as:

$$\begin{aligned} e_t &= \nu_e + e_{t-1} - h_1 z_{t-1} + \xi_1^{11} \Delta e_{t-1} + \xi_1^{12} \Delta f_{t-1} + \ldots + \xi_{p-1}^{11} \Delta e_{t-p+1} + \xi_{p-1}^{12} \Delta f_{t-p+1} + u_{1t} \\ f_t &= \nu_f + f_{t-1} - h_2 z_{t-1} + \xi_1^{21} \Delta e_{t-1} + \xi_1^{22} \Delta f_{t-1} + \ldots + \xi_{p-1}^{21} \Delta e_{t-p+1} + \xi_{p-1}^{22} \Delta f_{t-p+1} + u_{2t} \end{aligned} \tag{4″}$$

Subtracting the second from the first equation in (4″) gives:

$$z_t = (\nu_e - \nu_f) + \rho z_{t-1} + \tilde{u}_t \tag{5}$$

where $\rho = 1 - h_1 + h_2$ and the remainder term $\tilde{u}_t$ will in general be serially correlated. From $z_t \equiv e_t - f_t \sim$ I(0) it follows that $|\rho| < 1$ and hence $h_1 > h_2$. For $h_1 = 0$ this implies the restriction $-2 < h_2 < 0$.

3. Bootstrapping Long-Horizon Regression Tests

Numerous econometric studies have found that the random walk model provides more accurate forecasts than other models of the exchange rate (e.g., Meese and Rogoff (1983, 1988), Diebold and Nason (1990)). Thus, the random walk model is a natural benchmark in judging forecast performance. The monetary model of section 2 suggests that regressions of the form:

(6) $$e_{t+k} - e_t = a_k + b_k\, z_t + \varepsilon_{t+k}, \qquad k = 1, 4, 8, 12, 16$$

may improve forecast accuracy relative to the random walk forecast:

(7) $$e_{t+k} - e_t = d_k + \varepsilon_{t+k}, \qquad k = 1, 4, 8, 12, 16$$

by exploiting the mean reversion of z_t. This conjecture can be tested as $H_0\colon\ b_k = 0$ vs. $H_1\colon$ $b_k < 0$ for a given forecast horizon *k*, or jointly for all forecast horizons as $H_0\colon\ b_k = 0\ \forall\, k$ vs. $H_1\colon\ b_k < 0$ for some *k*. In essence, this is a standard Granger noncausality test for z_t in (6) based on the full sample. Alternatively, the out-of-sample prediction mean-squared error of models (6) and (7) based on a sequence of rolling forecasts may be evaluated using Theil's U-statistic or the DM statistic of Diebold and Mariano (1995). A formal test compares the null of equal forecast accuracy against the one-sided alternative that forecasts from (6) are more accurate than forecasts from (7). It is well known that asymptotic critical values for these test statistics are seriously biased in small samples. To mitigate these size distortions critical values may be calculated based on the bootstrap approximation of the finite sample distribution of the test statistic under the null hypothesis of no exchange rate predictability in the cointegrated model (4″) or some equivalent representation of the data generating process. Unlike asymptotic or exact finite sample critical values, bootstrap critical values based on the percentiles of the bootstrap distribution automatically adjust for the increase in the dispersion of the finite-sample distribution of the test statistic that occurs in near-spurious regressions as the sample size grows. As a result, bootstrap inference is immune from the near-spurious regression problem discussed in Berkowitz and Giorgianni (1997). However, special care must be taken to ensure the validity of the bootstrap model under the null.[2]

[2] For a recent review of the bootstrap testing methodology in time series models see Li and Maddala (1996).

3. 1. Bootstrapping Long-Horizon Regression Tests under the Null Hypothesis

A valid bootstrap algorithm may be readily constructed from representation (4′′). Recall that under $H_0: h_1 = 0$ and $h_2 < 0$, so e_t and f_t are cointegrated with vector C′ = [1, -1]. This suggests bootstrapping the VEC model:

$$
\begin{aligned}
\Delta e_t &= \nu_e + u_{1t} \\
\Delta f_t &= \nu_f - h_2\, z_{t-1} + \sum\nolimits_{j=1}^{p-1} \xi_j^{21}\, \Delta e_{t-j} + \sum\nolimits_{j=1}^{p-1} \xi_j^{22}\, \Delta f_{t-j} + u_{2t}
\end{aligned}
\tag{8}
$$

subject to the constraint that $-2 < h_2 < 0$ where p has been determined under H_0 by a suitable lag order selection criterion such as the SIC. This restricted model by construction has the same iid innovations as model (4).[3] Under the null hypothesis of no exchange rate predictability, it is known that $h_1 = 0$ which imposes the restriction $-2 < h_2 < 0$ for cointegration to exist. This condition must be imposed in estimating (8) to ensure that the bootstrap critical values generated from the estimated model are the critical values of the test statistic under H_0. It is also necessary to ensure the stationarity of the bootstrap data generating process for z_t in small samples. Estimation of (8) thus requires the use of constrained estimated generalized least squares (EGLS) with all coefficients but ν_e set equal to zero in the first equation and $(-h_2)$ constrained to be positive, but smaller than 2, in the second equation.[4]

The asymptotic validity of this bootstrap procedure follows from the standard assumptions in Bose (1988) after observing that the VEC model in (A2) may be equivalently represented as a VAR in Δe_t and z_t. Under the null hypothesis, the restricted EGLS estimator

[3] The assumption of iid innovations is not a concern for the quarterly data used in this paper. Note that in applications using monthly or weekly financial data with time varying volatility the procedure may be suitably modified by resampling blocks of residuals.

[4] In practice, h_2 may be constrained to some negative number arbitrarily close to zero. Under the null hypothesis, this constraint will not be binding asymptotically, so the asymptotic validity of the bootstrap procedure is not affected, regardless of the precise value of the constraint. EGLS estimation was implemented using an adaptation of the algorithm described in Lütkepohl (1991, pp. 168).

asymptotically converges to the standard LS estimator considered by Bose. Note that the discontinuity in the asymptotic distribution discussed in Basawa et al. (1991) does not arise in this model, because the cointegrating vector has been imposed in the vector error correction model.

Additional restrictions on the bootstrap data generating process may arise in special cases. For example, the null hypothesis that the exchange rate is known to follow a random walk without drift implies the restrictions $\nu_e = 0$ in (8) and $d_k = 0$ in the forecast model (7).[5] However, such an assumption is tenuous at best, and may result in spurious inference. Under the less restrictive assumption that the exchange rate follows a random walk, possibly with drift, ν_e and d_k must remain unrestricted.[6] Section 4.3. will illustrate what difference this seemingly innocuous assumption may make. Moreover, it is important that *both* drift terms remain unrestricted. Allowing for a possible drift in the exchange rate in specifying the bootstrap replica of the population process ($\nu_e \neq 0$), while ignoring this same drift in the no-change forecast of the exchange rate ($d_k = 0$) violates the requirement that the bootstrap model must be consistent with the model under the null hypothesis. The resulting bootstrap critical values will not accurately reflect the null hypothesis. It can be shown that the omission of d_k may lead to sizable changes in the *p*-values of the out-of-sample statistics.

3. 2. Comparison with Earlier Bootstrap Long-Horizon Regression Tests

The bootstrap test proposed in this paper is not the first attempt at bootstrapping long-

[5] Note that under the null hypothesis of a random walk without drift the intercept in (6) will be zero as well. An intercept must be included, however, because under the alternative hypothesis z_t enters with possibly nonzero mean.
[6] For example, Diebold, Gardeazabal, and Yilmaz (1994, p. 732) argue for including a drift, unless there is irrefutable evidence to the contrary.

horizon regression tests. The most sophisticated application to date is Mark (1995). His method is somewhat different, and it will be useful to clearly delineate the differences and the common features. Mark postulates the following model where the null hypothesis that the exchange rate follow a random walk has been imposed:

$$
\begin{aligned}
e_t - e_{t-1} &= a_0 + \varepsilon_{1t} \\
z_t &= b_0 + \sum_{j=1}^{J} b_j \; z_{t-j} \; + \varepsilon_{2t}
\end{aligned}
\tag{9}
$$

The innovations $\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t})'$ are iid and $\Sigma_\varepsilon = \mathrm{E}(\varepsilon_t \, \varepsilon_t')$. He estimates each equation of this model by OLS and generates bootstrap data conditional on the fitted values, possibly after correcting for bias in the second equation.

A simple example will illustrate how this bootstrap procedure relates to the bootstrap procedure described earlier. Suppose that the exchange rate follows a random walk and e_t and f_t are cointegrated such that $z_t \sim \mathrm{I}(0)$. For simplicity further suppose that there is just one lagged difference in model (4″). Then under H_0:

$$
\begin{aligned}
e_t &= \nu_e + e_{t-1} + u_{1t} \\
f_t &= \nu_f + f_{t-1} - h_2 (e_{t-1} - f_{t-1}) + \xi_1^{21} (e_{t-1} - e_{t-2}) + \xi_1^{22} (f_{t-1} - f_{t-2}) + u_{2t}
\end{aligned}
\tag{10}
$$

This vector error correction model may be expressed as a subset VAR in Δe_t and z_t. Pre-multiplying (10) by a conformable identity matrix whose second row has been replaced by C′ yields an equivalent representation based on Campbell (1987):

$$
\begin{aligned}
\Delta e_t &= \nu_e + u_{1t} \\
z_t &= \nu_e - \nu_f + (1 + h_2 + \xi_1^{22}) z_{t-1} - \xi_1^{22} z_{t-2} - (\xi_1^{21} + \xi_1^{22}) \Delta e_{t-1} + u_{1t} - u_{2t}
\end{aligned}
\tag{10′}
$$

Note that without the restrictions under the null hypothesis this model would be identical to the VAR models used in Campbell and Shiller (1987, 1988). By substituting for the lagged

Δe_t in the second equation of (10′), one may express the system in terms of the two marginal time series processes for Δe_t and z_t:

$$(10'') \quad \begin{aligned} \Delta e_t &= \nu_e + u_{1t} \\ z_t &= (1+\xi_1^{21}+\xi_1^{22})\,\nu_e - \nu_f + (1+h_2+\xi_1^{22})\,z_{t-1} - \xi_1^{22}\,z_{t-2} + u_{1t} - (\xi_1^{21}+\xi_1^{22})\,u_{1t-1} - u_{2t} \end{aligned}$$

The second equation of this system is the sum of a white noise process u_{2t} and an ARMA(2,1) process in z_t and u_{1t}. Engel (1984) proves that the sum of two possibly correlated ARMA processes will remain an ARMA process. This suggests approximating the ARMA process for z_t in (10′′) by a suitable higher order AR process, which results in Mark's model (9).

Provided that the estimated process for z_t is stationary, the bootstrap critical values from model (9) will thus be asymptotically equivalent to those from model (8). However, in small samples, they will tend to be biased, given the approximation error for $p > 1$ and given the fact that the model is estimated by equation-by-equation least-squares methods rather than EGLS. The bootstrap method used by Mark also does not formally impose the stationarity of z_t in resampling. This stationarity condition is the equivalent of the restrictions on h_2 in section 3. Moreover, estimates of (9) do not impose the restriction that $h_1 = 0$ under H_0. Put differently, unlike the bootstrap model in section 3, specification (9) is the same whether $h_1 = 0$ or not. This means that the bootstrap critical values would be biased, even if (9) were estimated by EGLS and if there were no approximation error.

In related work, Campbell (1993) considers a special case of the bootstrap algorithm in Mark (1995). In his model $J = p = 1$, and there is no lag order uncertainty. Both authors also consider corrections for bias in the least-squares coefficient estimates in the second equation of (9). Note that the OLS small-sample bias adjustment proposed in Campbell (1993) ignores the fact that the AR(1) is embedded in a regression system. Mark also reports results after bias

corrections, but his bias-adjustment procedure is unconventional. However, Campbell's procedure can be easily generalized to AR(p) models. An appropriate closed form solution for the first-order mean bias of the OLS estimator has been proposed by Shaman and Stine (1988). Similar bias corrections could also be employed for the EGLS estimate. For the method proposed in section 3, these bias corrections had little effect, given the proximity of the point estimate to the nonstationary region.

4. Empirical Results

The data set for this paper has been constructed from *OECD Main Economic Indicators* data for 1973:II-1994:IV. All data have been transformed exactly as described in Mark (1995). The data set includes the U.S. dollar exchange rates of the Canadian Dollar, the German Mark, the Japanese Yen, and the Swiss Franc. In the remainder of this section, the bootstrap algorithm described in section 3 will be put into practice. It will be useful to begin with some preliminary analysis of the unrestricted vector error correction model and the cointegration properties of the data.

4.1. Cointegration Tests

An indirect test of the theoretical underpinnings of long-horizon regressions may be conducted by testing the assumption of cointegration. Clearly, in the absence of cointegration there is no theoretical reason for regressing future changes in the exchange rate on the deviation of the spot exchange rate from its fundamental value. Standard tests of cointegration require the exchange rate and the fundamental to be individually I(1). This assumption will be relaxed later. Table 1 presents evidence based on the Horvath-Watson Wald test of the null of no cointegration

between e_t and f_t against the cointegrating vector implied by the monetary model.[7] All models include an intercept, as the error correction term z_t under the alternative is not mean zero in general. At the 10 percent level, the null of no cointegration cannot be rejected for any country. Thus, we cannot rule out that the data were generated from a difference stationary process. Both a difference stationary VAR and a VEC model are potentially consistent with the data.

A more informative approach focuses on the probability of making a type I error and a type II error. Adapting an idea of Rudebusch (1993), the marginal significance level of the Horvath-Watson test may be simulated based on the best-fitting model under H_0. Then the power of the test at the marginal significance level may be simulated based on the best-fitting model under H_1.[8] Table 1 displays the type-I error and type-II error for a rejection of H_0 implied by this bootstrap procedure. For Switzerland the type-II error clearly exceeds the type-I error, suggesting that the observed test statistic is more likely to have been generated under H_1 than under H_0, despite the nonrejection of H_0. For Canada, the test statistic is about equally likely to have been generated under H_1 or H_0. For Japan and Germany, the type-I error far outweighs the type-II error. This evidence does not resolve the controversy over whether there is cointegration in the data or not, but it establishes two points: First, if there is cointegration, the mean reversion of the error correction term is typically very slow. This would explain why exchange rates often appear unpredictable. Second, if the monetary model of section 2 holds, one would expect to find the strongest evidence of exchange rate predictability in the Swiss data, followed by the Canadian data. In contrast, for Germany and Japan there is very little evidence

[7] As noted by Berkowitz and Giorgianni (1997) the Horvath-Watson test is more powerful than the standard cointegration tests reported for example in Chinn and Meese (1995) and Mark (1995).

[8] Under both the null and the alternative the best-fitting model was selected based on the SIC allowing for lag orders between 0 and 4. The qualitative results in Table 1 are robust to alternative assumptions.

of mean reversion in the error correction term, and fundamentals would not be expected to improve forecast accuracy.

4. 2. Bootstrap Inference: Granger Causality

Estimates of the bootstrap data generating process are obtained as outlined in section 3. The constraint on h_2 is binding only for Germany. With the lag order constrained to lie between 0 and 4, the SIC selects two augmented lags for Canada, zero lags for Germany, and one lag each for Japan and Switzerland. Since the Bera-Jarque test rejects the null of Gaussian innovations for Canada at the 95 percent level, all bootstrap inference in this paper will be based on nonparametric resampling of the residuals.[9] The number of bootstrap replications is 2,000. Table 2 reports the bootstrap p-values under the null hypothesis that the exchange rate follows a random walk without drift. The restrictive assumption of zero drift will be relaxed later. Table 2 shows *p*-values for a number of key statistics.[10] t_{20} and t_A are the *t*-statistics for the slope coefficient in the long-horizon regression, with the subscript indicating whether the robust standard error is calculated based on a fixed truncation lag of 20 or Andrews' procedure. DM_{20} and DM_A refer to the corresponding Diebold-Mariano statistics and U to Theil's U-statistic.

It is instructive to first focus on the columns for t_{20} and t_A. For k = 1, only the t_{20} - and

[9] To check the sensitivity of bootstrap inference to small sample bias in the coefficient estimates in (10), a procedure analogous to that in Kilian (1997) was employed. The differences in results were so minor that no results will be reported.

[10] No slope coefficients are reported. As pointed out in Berkowitz and Giorgianni (1997), under the alternative hypothesis the slope coefficients will increase with the forecast horizon by construction, so that evidence of increasing slopes does not imply increased long-horizon predictability. This observation applies whether or not the slope coefficients are bias-adjusted. Similarly, statistical or visual measures of in-sample fit alone cannot be regarded as informative. Instead, results will be presented for marginal significance levels only. The use of bootstrap p-values also avoids the problem of spurious fits discussed in Berkowitz and Giorgianni (1997).

t_A-test statistics for Canada are significant at the 10 percent level. For Switzerland the corresponding *p*-values slightly exceed 10 percent, and for Japan and Germany they are as high as 42 to 77 percent. This result is roughly consistent with the evidence based on the Horvath-Watson test in section 4.1. A basic premise in the long-horizon regression test literature is that the power of tests of unpredictability can be expected to increase with the time horizon. Even if there is little evidence of exchange rate predictability for $k = 1$, as in Table 2, it is conjectured, raising the time horizon will reveal stronger evidence. There is little support for that conjecture in Table 2. With the possible exception of Switzerland, *p*-values do not decline substantially with rising *k*. Moreover, a proper test of the null of no exchange rate predictability recognizes the interdependence of the *t*-statistics across forecast horizons. As Berkowitz and Giorgianni (1997) show, $b_k = b_1 \sum_{i=0}^{k-1} \rho^i$, so the individual *t*-statistics must not be viewed as independent tests. To circumvent the problem of the dependency of the b_k coefficients, Mark suggests bootstrapping the distribution of the infimum of the *t*-statistics across the five time horizons of interest rather than the individual *t*-statistics for a given *k*. In Table 1, this statistic is labeled the joint *t*-test statistic. Table 1 shows that except for Switzerland none of the joint *t*-tests are significant at the 10 percent level. The in-sample *t*-statistics thus provide some prima facie evidence of exchange rate predictability for Canada and Switzerland, but no evidence for Germany and Japan.

4.3. Bootstrap Inference: Out-of-sample Forecast Accuracy

Turning to the out-of-sample evidence in Table 2, there seems to be strong evidence of out-of-sample predictability at horizon $k = 1$ for Canada and Japan, according to all three statistics, and some evidence for Switzerland according to the joint *U*-test statistic (defined as the

infimum of *U*-test statistics across *k*). There is no evidence of predictability at any horizon for Germany. Given the results of the cointegration test in section 4.1., these results are disturbing. While it may seem intuitive to find strong exchange rate predictability in the Canadian and Swiss data, the high persistence of the Japanese error correction term makes such a finding economically implausible. Moreover, the evidence for Japan appears stronger than for Switzerland, which seems incompatible with the results of the cointegration test.

This puzzle may be resolved by keeping in mind that Table 2 does *not* establish that economic fundamentals are responsible for the improved forecast accuracy; rather they measure the *joint* contribution of the drift term *and* the error correction term in the long-horizon regression forecast. Note that the out-of-sample statistics used in Mark (1995) and in Tables 2 and 3 compare the long-horizon regression forecast

$$(6) \qquad e_{t+k} - e_t = a_k + b_k\, z_t + \varepsilon_{t+k}\,, \qquad k = 1, 4, 8, 12, 16$$

with the forecast based on the driftless random walk:

$$(7') \qquad e_{t+k} - e_t = \varepsilon_{t+k}\,, \qquad k = 1, 4, 8, 12, 16\,.$$

Thus, the superior out-of-sample accuracy of (6) may be due to the fact that regression (6) picks up an apparent drift in the exchange rate over the sample period or due to the inclusion of the error correction term. The reason for the improved forecast performance is not identified. This makes it impossible to interpret a significant improvement in forecast accuracy as evidence in favor of monetary exchange rate models. The out-of-sample statistics may either overstate or understate the true contribution of the fundamental by lumping its effect together with that of the drift term.

To isolate the *marginal* contribution of z_t, one has to allow for a drift in the random walk forecast as in (7). To control for the drift term, I recalculated the results in Table 2 and

compared them to the forecast of a random walk allowing for drift. This leads to a striking change in the out-of-sample statistics, while leaving the in-sample results unchanged. Table 3 summarizes the findings. Not only do all one-step ahead *DM*-statistics and *U*-statistics for Switzerland now turn significant, but Switzerland and Canada are the only countries with significant out-of-sample statistics. *P*-values for Germany and Japan range from 27 to 54 percent. This result is exactly what one would have expected based on the cointegration test in section 4.1. It is also broadly consistent with the in-sample evidence in section 4.2. Moreover, there is no pattern of increasing significance with rising forecast horizon for any of the out-of-sample statistics. As in Table 2, it is useful to compare the one-step-ahead forecast test results to results for the joint *DM* test statistics (based on the supremum of the test statistic across *k*) and the joint *U*-test. With the exception of the *p*-value of the Canadian *U*-statistic (which is 13.7 percent), all *p*-values for Canada and Switzerland are significant at the 10 percent level, consistent with the results for $k = 1$. The *p*-values for Germany or Japan range from 34 to 39 percent. These out-of-sample test results demonstrate that fundamentals, after controlling for the possible presence of drift terms, may indeed improve the accuracy of real-time forecasts of the Swiss and Canadian exchange rate, but do not help in forecasting the DM and the Yen exchange rate.[11]

4. 4. Reconciling the Results with Earlier Findings

The preceding results are not directly comparable to Mark (1995) because they are based on an extended data set. Table 4 therefore shows the corresponding results for Mark's (1995)

[11] It may appear that an alternative test of the marginal contribution of z_t would be to estimate the long-horizon regression excluding the intercept. However, that regression would not be valid under the alternative hypothesis, unless it were known with certainty that f_t does not have a drift.

bootstrap procedure based on model (9).[12] The procedure differs slightly from Mark in that the lag order J of the process for z_t is selected by the AIC, given an upper bound of eight lags. The AIC selects $J = 7$ for Canada, $J = 5$ for Germany, $J = 1$ for Japan, and $J = 5$ for Switzerland. The roots ρ of the estimated processes are 0.9759, 0.8877, 0.9624, and 0.8871, respectively. These roots are consistently lower than the estimates implied by the restricted VEC model, which is suggestive of OLS small-sample bias.

Table 4a shows that extending the sample alone suffices to reverse many of Mark's (1995) results. There is only weak evidence of in-sample predictability for Canada and Switzerland and none for Japan and Germany. Except for Switzerland there is no significant evidence of out-of-sample predictability. While *p*-values often decline with rising forecast horizon, this tendency is by no means universal or as pronounced as in Mark (1995). This striking reversal suggests that much of the earlier evidence of exchange-rate predictability may simply have been an artifact of the small sample.

This does not necessarily mean, however, that the new results are more reliable. There are important differences to the results for the restricted VEC procedure in Tables 2 and 3. For example, the *p*-values for the in-sample statistics may be up to 0.30 lower or up to 0.09 higher than in Tables 2 and 3. These differences are an indication of the bias in the bootstrap critical values that arises because of approximation error and because many restrictions under the null hypothesis have not been imposed in the estimation of model (9). Moreover, while the in-sample results in Table 4a appear broadly similar to Tables 2 and 3, the out-of-sample results are highly sensitive to bias corrections.

[12] Alternatively, the results could have been compared for the original data set of Mark (1995). However, the results for the extended data set are of greater substantive interest and more reliable statistically.

Table 4b shows the corresponding results after suitable bias corrections of the autoregressive coefficients. The OLS bias estimates are based on Shaman and Stine (1988). Bias adjustments push the roots ρ of the estimated processes up to 0.9866, 0.9656, 0.9999, and 0.8930, respectively, closer to the estimates of ρ under the restricted VEC model. The substantial change in the estimates of ρ for Japan and Germany suggests that the bootstrap data generating process without bias corrections, on which the *p*-values in Table 4a conditioned, is likely to be quite misleading.

Correcting for bias leads to another striking reversal of the results. Note that after bias correction there is strong evidence of out-of-sample predictability for Japan, similar to the results in Table 2 for the model without drift. This result is quite different both from Table 4a and from the original results in Mark (1995), but it is exactly what one would have expected from a bootstrap model designed to compare the long-horizon regression forecast to that of a random walk without drift. Results for Switzerland and Germany are also similar to Table 2, but Mark's procedure does not detect any out-of-sample predictability for Canada. Some of the remaining differences in out-of-sample predictability are likely to be due to the inconsistent treatment of the drift term in the bootstrap data generating process and the random walk forecast model. Note that Tables 4a and 4b compare the out-of-sample accuracy of forecasts from long-horizon regressions to a simple no-change forecast of the exchange rate. The bootstrap model (9) implies a contradictory set of beliefs in which the researcher allows for a possible drift in the exchange rate in specifying the bootstrap replica of the population process ($\nu_e \neq 0$), but insists on ignoring this drift when using the no-change forecast of the exchange rate ($d_k = 0$). In addition, important differences also remain in-sample, as evidenced by the substantially lower *p*-values for

Japan in Table 4b compared to Tables 2 or 3. These differences (often in excess of 0.50) again are a consequence of the failure to impose all restrictions implied by the null hypothesis.

Finally, neither Table 4a nor Table 4b appear consistent with the evidence of cointegration based on the Horvath-Watson test. There is no reason to expect strong mean reversion in the Japanese exchange rate, for example, as Table 4b would suggest. As shown earlier, this counterintuitive result appears to be an artifact of the questionable assumption that the exchange rate follows a random walk without drift. The out-of-sample tests in Table 4 do not test the hypothesis that economic fundamentals improve the accuracy of exchange rate forecasts; rather they measure the *joint* contribution of the drift term *and* the error correction term in the long-horizon regression forecast. As a result, it is not possible to interpret a significant improvement in forecast accuracy as evidence in favor of monetary exchange rate models. In contrast, the results for the restricted VEC model in Table 3 are consistent with the cointegration evidence and measure the *marginal* contribution of the error correction term to forecast accuracy.

5. Sensitivity Analysis

The evidence presented so far has exploited the notion of cointegration between macroeconomic fundamentals and the exchange rate. Assuming that fundamentals are I(1), the cointegration result follows from the theoretical model of section 2. However, it is quite possible that the model is wrong, in which case cointegration may not hold even if the fundamental is I(1).[13] As Table 1 suggests, the cointegration assumption cannot be taken for granted. It is therefore important to assess the sensitivity of the findings in Table 3. Table 5 presents the

[13] While the error correction term in models without cointegration is I(1) by construction, under the null hypothesis of no exchange rate predictability the slope coefficient of the long-horizon regression is zero and hence the critical values are well-defined.

corresponding results under the assumption that the true vector process is difference stationary with the exchange rate further restricted to be a random walk. Estimation is by constrained EGLS. The results for Germany and Japan are unchanged. This finding is not surprising, given the continuity of the finite-sample distribution. With estimates of the root ρ of the error correction term z_t very close to 1 under the null hypothesis, there is little to distinguish a difference stationary process from a VEC model, especially for Japan and Germany. For Canada, the two *t*-statistics for $k = 1$ turn insignificant at the 10 percent level (with *p*-values of 11.6 and 13.7 percent), but the out-of-sample results are not affected. For Switzerland, only one of the three out-of-sample statistics for $k = 1$ remains significant at the 10 percent level (the other two having *p*-values of 13 and 12.5 percent). Similarly, only one of the three joint out-of-sample tests remains significant at the 10 percent level. However, the two joint *DM*-statistics are only barely insignificant with *p*-values of 10.9 and 10.3 percent. The in-sample evidence weakens especially for $k = 1$, dropping to 28.4 and 22.6 percent, but both joint *t*-tests remain highly significant. Overall, dropping the cointegration assumption weakens the evidence of exchange rate predictability, but cannot reverse it.[14]

A further possibility is that, in addition to the model being false, the fundamental is not I(1). So far this I(1) assumption has been taken for granted. Table 6 applies the Rudebusch (1993) bootstrap procedure to the augmented Dickey-Fuller (ADF) test of the null of a unit root in f_t. Panel (a) shows the type I and type II errors for the alternative hypothesis of a level stationary process with intercept, but excluding a time trend. The bootstrap marginal significance level in no case allows the rejection of H_0. However, for Canada the type II error

[14] Virtually identical results hold if the fundamental is assumed not to respond to past values of the exchange rate, except that the *p*-values of the Canadian one-step-ahead *t*-statistics improve to 8.8 and 10.3 percent.

far exceeds the type I error suggesting that the observed ADF test statistics is more likely to have been generated under the alternative hypothesis. If true, this test result would imply that z_t is I(1) which would invalidate the bootstrap *p*-values for Canada in Tables 2 and 3. Similarly, panel (b) of Table 6 presents the type I and type II error for the ADF test of the null of a unit root in fundamentals against the alternative of a trend stationary process. For all countries but Switzerland the marginal significance level is too high to reject H_0. For Switzerland, the test rejects the null hypothesis of a unit root in the fundamental at the 10 percent level. This rejection coincides with a very high type II error, and casts further doubt on the findings in Tables 2 and 3.

To verify the sensitivity of the conclusions in section 4, Table 7 presents additional evidence for Switzerland and Canada based on the test results in Table 6. The critical values implied by these data generating processes may be interpreted as the critical values under the joint null of *no* cointegration and no exchange rate predictability. The bootstrap data generating process for Canada is of the form:

$$\begin{aligned} \Delta e_t &= \nu_e + u_{1t} \\ f_t &= \nu_f + \sum_{j=1}^{p} \Phi_j^{21} f_{t-j} + u_{2t} \end{aligned} \tag{11}$$

whereas the process for Switzerland includes an additional log-linear time trend in the equation for f_t. Table 7 shows that the results of Table 3 are essentially unchanged.[15]

6. The Size and Power of Long-Horizon Regression Tests

To assess the reliability of the proposed bootstrap procedure and to help explain the results of section 4, I conduct a Monte Carlo study of the size and power of the bootstrap test. It is widely presumed that long-horizon tests have distinct power advantages over standard tests.

[15] Virtually identical results are obtained if additional lagged differences of the exchange rate are included in (11).

This belief is central to the long-horizon regression test literature. The idea of increasing long-horizon predictability seems hard to reconcile with the fairly stable pattern of p-values for Germany, Japan, and Switzerland in Table 3. For Canada there even is a tendency for the p-values to *increase* with the time horizon, suggesting that long-horizon tests may in fact have lower power. What is even more curious is the absence of analytic or simulation results substantiating the claims of higher power at long forecast horizons. The aim of this section is to examine the basis of these claims for some realistic data generating processes. Before studying the power properties of the bootstrap tests of section 3, it is useful to examine their accuracy in small samples.

Table 8 (a) shows the effective size of the nominal 10 percent test based on bootstrap p-values. All results are based on 500 trials with 2,000 bootstrap replications each. The approximate Monte Carlo simulation error is 0.0134. The data generating process is based on the restricted VEC model under the null hypothesis that the exchange rate follows a random walk (possibly with drift) and the exchange rate and the fundamental are cointegrated. Separate data generating processes (DGPs) are estimated for each country. The lag orders are based on the SIC as in all previous applications. For each trial, the bootstrap procedure of Table 3 is used to calculate the p-value, and the rejection rates are tabulated. Table 8 (a) suggests that the bootstrap tests are fairly accurate, especially for the out-of-sample statistics. Moreover, the size is fairly constant across forecast horizons. The differences in the accuracy of the one-step ahead forecast horizon test and the joint test are statistically significant in only two of twenty cases. For the Canadian DGP, the size of the t_{20}-test slightly increases, and for the German DGP the size of the t_A-test falls slightly. Overall, there is strong evidence that any systematic differences between test results for short and long horizon tests must be due to differences in power.

Table 8 (b) shows the power of the nominal 10 percent test against the alternative of an unrestricted VEC model. The DGP is based on the best-fitting unrestricted VEC model estimates for Canada, Germany, Japan, and Switzerland. The roots ρ of the error correction term are 0.9521, 0.9677, 0.9624 and 0.9075, respectively. The simulation results suggest that there are no power advantages to long-horizon regression tests.[16] For the Japanese and German DGP, formal significance tests for the difference in power between the one-step ahead test and the joint test cannot reject the null hypothesis that power remains constant. For Switzerland, there is evidence of significant declines in power with rising forecast horizon. For Canada, the results are mixed, but only in the case of the U-statistic power increases significantly. In no case, the increase in power appears economically significant.[17]

The absence of increasing long-horizon power against one alternative does not rule out that long-horizon tests may have higher power against other alternatives. For example, it is often suggested that the predictability of the exchange rate may be obscured by speculative dynamics or noise trading following fads.[18] The transitory noise is expected to subside at longer time horizons revealing that the exchange rate is fundamentally driven by the monetary model. As a result, long-horizon tests can be expected to have higher power. This notion may be formalized by modeling the exchange rate as the sum of the exchange rate implied by the unrestricted VEC model (4) and a serially correlated noise component $e_t = e_t^* + n_t$. For simplicity the noise

[16] The power results for the *t*-statistics are consistent with size-adjusted power results in Demiroglu and Shapiro (1997) for a simpler model. Campbell (1993) reaches similar conclusions for unweighted long-horizon regression *t*-tests, using a somewhat different methodology. He also reports more favorable Monte Carlo simulation results for a weighted long-horizon regression bootstrap test. As far as I know, there does not exist any related work on the size and power of the out-of-sample tests.

[17] The power results remain essentially unchanged after accounting for the slight, but statistically significant changes in size pointed out earlier. For Japan, the size-adjusted power of the *DM(20)* test is likely to be somewhat higher, and for Canada the size-adjusted power of the *t(20)*-test is likely to fall.

[18] Cutler, Poterba, and Summers (1991) document evidence of speculative dynamics in many asset returns. Such dynamics may arise as the result of feedback trading (e.g., Cutler, Poterba, and Summers (1990)).

component is assumed to follow an AR(1) process $n_t = \alpha\ n_{t-1} + \varepsilon_{nt}$ with a half-life of approximately two years (corresponding to $\alpha = 0.92$) and a variance calibrated to match 75 percent of the variance of the first-differenced exchange rate. Less extreme assumptions would be equally plausible, but are unlikely to generate important transitory movements or to dramatically alter the power of the test.[19] The experiment deliberately tilts the playing field in favor of finding higher power. Contrary to the conjecture, Table 9 shows that adding a fad component to the Swiss and to the Canadian exchange rate does not fundamentally alter the power of the tests. With the exception of the U-statistic for Canada there is no evidence of increasing power, but some significant evidence of declining power with rising forecast horizon.

The Monte Carlo evidence of stable size and typically constant or declining power with rising forecast horizon casts doubt on the use of long-horizon regression tests in the literature. There is no evidence that such tests perform systematically better than standard tests based on one-step ahead forecasts, but significant evidence that they may perform much worse. The evidence of lower power is quite intuitive, given the shortening of the effective sample, as the forecast horizon increases. The Monte Carlo study also confirms the reliability of the bootstrap methodology proposed in section 3, and it explains the absence of a pattern of increased long-horizon predictability in Table 3.

7. Concluding Remarks

Long-horizon regression tests are widely used in empirical finance as tests of market efficiency. In the absence of market efficiency, deviations of asset prices from their long-run

[19] The model of the transitory components is similar to Summers (1986) and Poterba and Summers (1988). Given the low power of standard tests, it is difficult to directly identify and estimate the transitory component. However, if the power results were sensitive to serially correlated transitory noise of general form, one would expect the stylized Monte Carlo experiment in Table 8 to reveal this fact.

equilibrium value should help predict cumulative future asset returns. Regression tests of this hypothesis typically find strong evidence that economic fundamentals help predict asset returns at long forecast horizons, but not at short horizons. The interpretation of these results, however is far from clear. Numerous studies have documented severe size distortions of long-horizon regression tests. In this paper, I proposed a new bootstrap method for small-sample inference in long-horizon regressions. I presented simulation evidence that this bootstrap method greatly reduces the size distortions of conventional long-horizon regression tests in realistic situations. The remaining size distortions are typically small, and the size of the test appears stable across forecast horizons.

I illustrated the use of this bootstrap method by analyzing the long-horizon predictability based on monetary fundamentals of four major exchange rates. In recent years, this question has received considerable interest in the international finance literature (e.g., Chinn and Meese (1995), Mark (1995), Chen and Mark (1996)). My results differed in important ways from the earlier literature. I showed that many of the differences in results can be traced to the implementation of the long-horizon regression test. Two substantive results stand out: First, unlike earlier studies, I found only weak evidence that fundamentals help predict the Swiss Franc and the Canadian dollar rate, but no evidence for Germany and Japan. This finding is consistent with evidence based on the Horvath and Watson (1995) test of the null of no cointegration. It also appears remarkably robust to whether or not cointegration is assumed under the null hypothesis of no exchange rate predictability.

Second, in contrast to the earlier literature, I found no evidence of patterns of increasing long-horizon predictability in exchange rates. The latter finding may seem surprising, given the fundamental premise of the long-horizon regression test literature that power improves at long

forecast horizons. It is precisely at these long horizons that we would expect the exchange rate to be predictable based on the monetary model. However, it has not been demonstrated to date that in realistic situations power actually increases as the time horizon grows. This paper made an effort to fill that gap. I fit several monetary exchange rate models to the data used in the empirical study. Based on these data generating processes, I investigated the power of the bootstrap long-horizon regression test by Monte Carlo simulation for each of the four currencies. The stable and fairly accurate size of the test makes it straightforward to evaluate the power of the bootstrap test against economically plausible alternatives.

The natural alternative against which to test the null hypothesis of no predictability is the vector error correction model implied by the underlying net present value model. I found that there is no evidence that long-horizon tests are systematically more powerful than standard tests if the net present value model is true. This finding is consistent with the pattern of stable or increasing bootstrap *p*-values found in the data. In fact, in many cases the power of long-horizon tests declines with the forecast horizon. The evidence of lower power is quite intuitive in out-of-sample forecasts, given the shortening of the effective sample, as the forecast horizon increases.

I also considered a second economically plausible alternative allowing for short-term speculative dynamics in the exchange rate. Fama and French (1988) and Poterba and Summers (1988), among others, have argued that forecasting ability should be easier to detect at long horizons if asset prices contain a highly serially correlated transitory noise component. This noise component can be thought of as fads in investors' behavior. I presented additional Monte Carlo evidence for such a model calibrated to actual data. I again found that power does not improve with higher forecast horizons and may in fact decline.

The persistent lack of evidence of higher power at long horizons suggests that previous findings of increasing long-horizon predictability are more likely due to size distortions than to power gains. The notion that long-horizon tests enjoy power advantages is central to the long-horizon regression literature. The results of this paper may therefore come as a surprise, but they reinforce and extend similar results for the size-adjusted power of long-horizon regression tests in Campbell (1993) and Demiroglu and Shapiro (1997). They also are consistent with Monte Carlo evidence of the exact finite sample distributions of the test statistic in Bollerslev and Hodrick (1995, p. 434).

It may be tempting to conclude that, if there are no power gains, we might as well avoid the statistical complications of long-horizon regressions and rely on more conventional tests of predictability. For example, Campbell and Shiller (1987, 1988) have used predictability tests based on cointegrated VAR models which are free from the complications of econometric inference in long-horizon regressions. One drawback of the Campbell-Shiller VAR approach is that it uses the full sample in calculating *ex-post* measures of predictability. However, out-of-forecasts could be constructed from the same model using rolling or recursive regressions. While the Campbell-Shiller model is not designed to be bootstrapped under the null of no predictability, the bootstrap methodology proposed in this paper could easily be modified to generate out-of-sample forecast statistics based on the restricted VEC model. This model is fully consistent with the Campbell-Shiller VAR model under the null hypothesis of unpredictable asset returns. Future research will have to systematically explore these approaches and compare them with the long-horizon regression test.

At this point it would be premature to completely discard the idea of long-horizon regression tests. However, the evidence presented in this paper clearly shifts the burden of proof

to the advocates of long-horizon regression tests. What needs to be demonstrated is that there are other economically plausible alternatives against which long-horizon regression tests have power advantages. It would also be useful to investigate in more detail the power of weighted long-horizon regressions, building on preliminary results in Campbell (1993). Campbell's results suggest that in some cases regression tests based on weighted cumulative forecasts may have power advantages at longer horizons.

References:

Basawa, I.V., A.K. Mallik, W.P. Cormick, J.H. Reeves, and R.L. Taylor (1991), "Bootstrapping Unstable First-Order Autoregressive Processes," *Annals of Statistics* 19 (2), 1098-1101.

Berkowitz, Jeremy, and Lorenzo Giorgianni (1997), "Long-Horizon Exchange Rate Predictability?," Working Paper No. 97/6, International Monetary Fund.

Bollerslev, Tim, and Robert J. Hodrick (1995), "Financial Market Efficiency Tests," in: H. Pesaran and M. Wickens (eds.), *Handbook of Applied Econometrics*, Cambridge, MA: Blackwell Publishers.

Bose, Arup (1988), "Edgeworth Correction by Bootstrap in Autoregressions," *Annals of Statistics* 16 (4), 1709-1722.

Campbell, John Y. (1993), "Why Long-Horizons: A Study of Power Against Persistent Alternatives," NBER Technical Working Paper No. 142.

Campbell, John Y. (1987), "Does Saving Anticipate Declining Labor Income? An Alternative Test of the Permanent Income Hypothesis," *Econometrica*, 55 (6), 1249-1273.

Campbell, John Y., and Robert J. Shiller (1987), "Cointegration and Tests of Present Value Models," *Journal of Political Economy*, 95 (5), 1062-1088.

Campbell, John Y., and Robert J. Shiller (1988), "Stock Prices, Earnings, and Expected Dividends," *Journal of Finance*, 43, 661-676.

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

Chen, Jian, and Nelson C. Mark (1996), "Alternative Long-Horizon Exchange Rate Predictors," Working Paper No. 96-05, Department of Economics, Ohio State University.

Chinn, Menzie D., and Richard A. Meese (1995), "Banking on Currency Forecasts: How Predictable is Change in Money?," *Journal of International Economics*, 38, 161-178.

Cutler, David M., James M. Poterba, and Lawrence H. Summers (1991), "Speculative Dynamics," *Review of Economic Studies*, 58 (3), 529-546.

Cutler, David M., James M. Poterba, and Lawrence H. Summers (1990), "Speculative Dynamics and the Role of Feedback Traders," *American Economic Review*, 80 (2), 63-68.

Demiroglu, Ufuk, and Matthew Shapiro (1997), "Predictability in Financial Markets and the Power of Long-Horizon Tests," manuscript, Department of Economics, University of Michigan.

Diebold, Francis X., and Roberto Mariano (1995), "Comparing Predictive Accuracy," *Journal of Business and Economic Statistics*, 13, 253-262.

Diebold, Francis X., Javier Gardeazabal, and Kamil Yilmaz (1994), "On Cointegration and Exchange Rate Dynamics," *Journal of Finance*, 49 (2), 727-735.

Diebold, Francis X., and James A. Nason (1990), "Nonparametric Exchange Rate Prediction?," *Journal of International Economics*, 28 (3-4), 315-332.

Engel, E. M. R. A. (1984), "A Unified Approach to the Study of Sums, Products, Time-Aggregation and Other Functions of ARMA Processes," *Journal of Time Series Analysis*, 5 (3), 159-171.

Fama, Eugene F., and Robert R. Bliss (1987), "The Information in Long-Maturity Forward Rates," *American Economic Review*, 77 (4), 680-692.

Fama, Eugene F., and Kenneth R. French (1988), "Dividend Yields and Expected Stock Returns," *Journal of Financial Economics*, 22, 3-25.

Hodrick, Robert J. (1992), "Dividend Yields and Expected Stock Returns: Alternative Procedures for Inference and Measurement," *Review of Financial Studies*, 5 (3), 357-386.

Horvath, Michael T.K., and Mark W. Watson (1995), "Testing for Cointegration when some of the Cointegrating Vectors are Prespecified," *Econometric Theory*, 11, 984-1014.

Kilian, Lutz (1997), "Small-Sample Confidence Intervals for Impulse Response Functions," forthcoming: *Review of Economics and Statistics*.

Li, Hongyi, and G.S. Maddala (1996), "Bootstrapping Time Series Models," *Econometric Reviews*, 15 (2), 115-158.

Lütkepohl, Helmut (1991), *An Introduction to Multiple Time Series Analysis*, Springer-Verlag: New York.

Mankiw, N. Gregory, David Romer, and Matthew D. Shapiro (1991), "Stock Market Forecastability and Volatility: A Statistical Appraisal," *Review of Economic Studies*, 58, 455-477.

Mark, Nelson C. (1995), "Exchange Rates and Fundamentals: Evidence on Long-Horizon Predictability," *American Economic Review*, 85 (1), 201-218.

Meese, Richard, and Kenneth Rogoff (1983), "Empirical Exchange Rate Models of the 1970s: Do They Fit Out of Sample?," *Journal of International Economics*, 14 (1-2), 3-24.

Meese, Richard, and Kenneth Rogoff (1988), "Was It Real? The Exchange-Rate Interest Differential Relation Over the Modern Floating-Rate Period," *Journal of Finance*, 4 (4), 933-948.

Ng, Serena, and Pierre Perron (1995), "Unit Root Tests in ARMA Models with Data-Dependent Methods for the Selection of the Truncation Lag," *Journal of the American Statistical Association*, 90, 268-281.

Poterba, James H., and Lawrence H. Summers (1988), "Mean Reversion in Stock Prices: Evidence and Implications," *Journal of Financial Economics*, 22 (1), 27-59.

Rudebusch, Glenn D. (1993), "The Uncertain Unit Root in Real GNP," *American Economic Review*, 83 (1), 264-272.

Shaman, Paul, and Robert A. Stine (1988), "The Bias of Autoregressive Coefficient Estimators," *Journal of the American Statistical Association*, 83, 842-848.

Summers, Lawrence H. (1986), "Does the Stock Market Rationally Reflect Fundamental Values?," *Journal of Finance*, 41 (3), 591-601.

Table 1: Probabilities of the observed Horvath-Watson test statistic for the null of no cointegration having been generated under the null hypothesis and the alternative hypothesis

	Probability of Type I Error (percent)	Probability of Type II Error (percent)
Canadian Dollar	25.80	27.85
German Mark	78.75	5.40
Japanese Yen	68.65	7.55
Swiss Franc	13.55	34.75

NOTES: The type-I error probability is the marginal significance level based on simulations of the test statistic under the best-fitting difference stationary model. Given this type-I error, the type-II error probability is based on the distribution of the test statistic under the best-fitting vector error correction model. Both models were selected using the SIC and allowing for up to 4 augmented lags. The results are based on the extended data set for 1973:II-1994:IV. Similar results are obtained for Mark's original data set and using alternative lag order selection procedures. All distributions are based on 2,000 replications.

Table 2: Bootstrap p-values from data generating process restricted under the joint null hypothesis of cointegration and a driftless random walk in the exchange rate

					Relative to driftless random walk		
	ρ	k	t(20)	t(A)	U	DM(20)	DM(A)
Canadian Dollar	0.9732	1	**0.054**	**0.065**	**0.045**	**0.065**	**0.065**
		4	0.132	0.132	0.147	0.139	0.139
		8	0.172	0.159	0.139	0.145	0.144
		12	0.286	0.305	0.341	0.247	0.296
		16	0.392	0.415	0.539	0.469	0.603
		joint	0.250	0.305	0.257	0.214	0.228
German Mark	1.0000	1	0.420	0.599	0.348	0.470	0.500
		4	0.486	0.535	0.329	0.675	0.556
		8	0.520	0.491	0.334	0.316	0.313
		12	0.473	0.458	0.268	0.249	0.250
		16	0.428	0.460	0.206	0.231	0.229
		joint	0.498	0.521	0.259	0.348	0.343
Japanese Yen	0.9998	1	0.771	0.669	**0.080**	**0.042**	**0.038**
		4	0.738	0.772	**0.079**	**0.071**	**0.061**
		8	0.620	0.607	**0.072**	**0.083**	**0.068**
		12	0.497	0.502	**0.099**	0.142	0.128
		16	0.557	0.570	0.112	0.173	0.165
		joint	0.617	0.613	0.141	0.203	0.183
Swiss Franc	0.9850	1	0.150	0.107	0.107	0.123	0.115
		4	0.132	0.102	0.117	0.130	0.131
		8	**0.095**	**0.085**	0.113	0.137	0.134
		12	**0.036**	**0.034**	**0.046**	**0.092**	**0.091**
		16	**0.001**	**0.003**	**0.003**	**0.073**	**0.072**
		joint	**0.002**	**0.006**	**0.003**	0.111	0.103

NOTES: The table presents the marginal significance levels of the *t*-statistic of the slope coefficient in the long-horizon regression, of Theil's U-statistic and of the Diebold-Mariano statistic of equal out-of-sample forecast accuracy. Results are shown for alternative forecast horizons *k*. The entry *joint* refers to the distribution of the *sup* or the *inf* of the test statistic across *k*. The p-values for the one-sided test are generated under the null hypothesis based on nonparametric i.i.d. resampling of the residuals in the best-fitting restricted vector error correction model. ρ is the implied root of the regressor z_t under H_0. Boldface indicates significance at the 10 percent level.

Table 3: Bootstrap p-values from data generating process restricted under the joint null hypothesis of cointegration and a random walk with drift in the exchange rate

					Relative to random walk with drift		
	ρ	k	t(20)	t(A)	U	DM(20)	DM(A)
Canadian Dollar	0.9731	1	**0.054**	**0.065**	**0.048**	**0.055**	**0.065**
		4	0.131	0.131	**0.055**	**0.050**	**0.066**
		8	0.172	0.160	**0.052**	**0.022**	**0.021**
		12	0.287	0.304	0.129	0.116	0.104
		16	0.391	0.414	0.250	0.345	0.421
		joint	0.248	0.304	0.137	**0.069**	**0.066**
German Mark	1.0000	1	0.346	0.492	0.352	0.537	0.375
		4	0.395	0.442	0.309	0.407	0.365
		8	0.402	0.372	0.308	0.303	0.300
		12	0.359	0.350	0.272	0.274	0.274
		16	0.344	0.356	0.281	0.299	0.297
		joint	0.385	0.393	0.348	0.394	0.391
Japanese Yen	0.9997	1	0.629	0.520	0.269	0.360	0.309
		4	0.589	0.573	0.276	0.290	0.291
		8	0.473	0.461	0.217	0.221	0.217
		12	0.362	0.361	0.276	0.277	0.277
		16	0.403	0.406	0.389	0.346	0.352
		joint	0.463	0.455	0.336	0.360	0.352
Swiss Franc	0.9850	1	0.150	0.107	**0.061**	**0.081**	**0.049**
		4	0.132	0.102	**0.091**	0.102	**0.093**
		8	**0.095**	**0.085**	**0.089**	0.112	0.108
		12	**0.035**	**0.034**	**0.032**	**0.064**	**0.068**
		16	**0.001**	**0.003**	**0.007**	**0.049**	**0.037**
		joint	**0.002**	**0.006**	**0.011**	**0.094**	**0.066**

NOTES: The table presents the marginal significance levels of the *t*-statistic of the slope coefficient in the long-horizon regression, of Theil's U-statistic and of the Diebold-Mariano statistic of equal out-of-sample forecast accuracy. Results are shown for alternative forecast horizons *k*. The entry *joint* refers to the distribution of the *sup* or the *inf* of the test statistic across *k*. The p-values for the one-sided test are generated under the null hypothesis based on nonparametric i.i.d. resampling of the residuals in the best-fitting restricted vector error correction model. ρ is the implied root of the regressor z_t under H_0. Boldface indicates significance at the 10 percent level.

Table 4a: Bootstrap p-values from data generating process under the null hypothesis of a random walk in the exchange rate as outlined in Mark (1995)

					Relative to random walk with drift		
	ρ	k	t(20)	t(A)	U	DM(20)	DM(A)
Canadian Dollar	0.9759	1	**0.058**	**0.073**	0.224	0.240	0.246
		4	0.115	0.117	0.380	0.373	0.373
		8	0.140	0.128	0.338	0.346	0.344
		12	0.225	0.243	0.545	0.469	0.521
		16	0.318	0.340	0.712	0.672	0.768
		joint	0.201	0.249	0.481	0.428	0.447
German Mark	0.8877	1	0.200	0.303	0.351	0.506	0.532
		4	0.247	0.280	0.336	0.692	0.568
		8	0.269	0.247	0.322	0.316	0.309
		12	0.240	0.233	0.264	0.247	0.248
		16	0.193	0.221	0.204	0.225	0.222
		joint	0.271	0.285	0.254	0.336	0.328
Japanese Yen	0.9624	1	0.603	0.471	0.290	0.155	0.148
		4	0.580	0.555	0.287	0.198	0.179
		8	0.477	0.468	0.254	0.210	0.201
		12	0.389	0.419	0.269	0.301	0.296
		16	0.445	0.488	0.311	0.365	0.361
		joint	0.524	0.545	0.336	0.359	0.349
Swiss Franc	0.8871	1	**0.091**	**0.059**	0.173	0.189	0.176
		4	**0.087**	**0.066**	0.186	0.197	0.199
		8	**0.041**	**0.037**	0.172	0.195	0.190
		12	**0.008**	**0.007**	**0.068**	0.127	0.136
		16	**0.000**	**0.000**	**0.009**	**0.095**	**0.096**
		joint	**0.000**	**0.000**	**0.009**	0.132	0.124

NOTES: The table presents the marginal significance levels of the *t*-statistic of the slope coefficient in the long-horizon regression, of Theil's U-statistic and of the Diebold-Mariano statistic of equal out-of-sample forecast accuracy. Results are shown for alternative forecast horizons *k*. The entry *joint* refers to the distribution of the *sup* or the *inf* of the test statistic across *k*. All results are based on OLS estimation of model (9). The p-values for the one-sided test are generated based on nonparametric i.i.d. resampling of the residuals of the best-fitting model selected by the AIC. ρ is the largest root of the autoregressive process for z_t. Boldface indicates significance at the 10 percent level.

Table 4b: Bootstrap p-values from bias-corrected data generating process under the null hypothesis of a random walk in the exchange rate as outlined in Mark (1995)

					Relative to random walk with drift		
	ρ	k	t(20)	t(A)	U	DM(20)	DM(A)
Canadian Dollar	0.9866	1	**0.054**	**0.071**	0.225	0.245	0.249
		4	0.106	0.106	0.387	0.380	0.379
		8	0.125	0.115	0.344	0.347	0.347
		12	0.212	0.232	0.553	0.486	0.535
		16	0.294	0.313	0.715	0.673	0.777
		joint	0.183	0.227	0.501	0.437	0.456
German Mark	0.9656	1	0.301	0.442	0.339	0.502	0.528
		4	0.338	0.379	0.317	0.686	0.560
		8	0.365	0.330	0.305	0.311	0.307
		12	0.331	0.316	0.259	0.238	0.242
		16	0.279	0.302	0.199	0.222	0.220
		joint	0.358	0.368	0.255	0.330	0.324
Japanese Yen	0.9999	1	0.237	0.145	0.140	**0.023**	**0.036**
		4	0.215	0.185	0.126	**0.055**	**0.051**
		8	0.136	0.117	0.101	**0.069**	**0.061**
		12	**0.084**	**0.073**	**0.088**	0.101	**0.095**
		16	0.108	0.101	**0.089**	0.108	0.106
		joint	0.117	0.107	0.118	0.120	0.118
Swiss Franc	0.8930	1	0.103	**0.062**	0.169	0.189	0.176
		4	**0.091**	**0.066**	0.182	0.188	0.193
		8	**0.055**	**0.045**	0.161	0.183	0.178
		12	**0.015**	**0.011**	**0.067**	0.117	0.123
		16	**0.000**	**0.000**	**0.010**	**0.095**	**0.089**
		joint	**0.001**	**0.001**	**0.010**	0.132	0.121

NOTES: The table presents the marginal significance levels of the *t*-statistic of the slope coefficient in the long-horizon regression, of Theil's U-statistic and of the Diebold-Mariano statistic of equal out-of-sample forecast accuracy. Results are shown for alternative forecast horizons *k*. The entry *joint* refers to the distribution of the *sup* or the *inf* of the test statistic across *k*. All results are based on bias-corrected OLS estimates of model (9). The p-values for the one-sided test are generated based on nonparametric i.i.d. resampling of the residuals of the best-fitting model selected by the AIC. ρ is the largest root of the autoregressive process for z_t. Boldface indicates significance at the 10 percent level.

Table 5: Bootstrap p-values from difference-stationary VAR process for the exchange rate and the fundamental with the exchange rate restricted to be a random walk with drift under the null hypothesis

				Relative to random walk with drift		
	k	t(20)	t(A)	U	DM(20)	DM(A)
Canadian Dollar	1	0.116	0.137	**0.065**	**0.055**	**0.064**
	4	0.236	0.229	**0.083**	**0.052**	**0.074**
	8	0.277	0.253	**0.077**	**0.022**	**0.024**
	12	0.415	0.436	0.164	0.129	0.122
	16	0.530	0.546	0.255	0.359	0.418
	joint	0.338	0.392	0.174	**0.070**	**0.072**
German Mark	1	0.346	0.492	0.352	0.537	0.375
	4	0.395	0.442	0.309	0.407	0.365
	8	0.402	0.372	0.308	0.303	0.300
	12	0.359	0.350	0.272	0.274	0.274
	16	0.344	0.356	0.281	0.299	0.297
	joint	0.385	0.393	0.348	0.394	0.391
Japanese Yen	1	0.686	0.581	0.310	0.398	0.351
	4	0.657	0.641	0.322	0.333	0.329
	8	0.543	0.530	0.259	0.254	0.250
	12	0.417	0.423	0.314	0.314	0.314
	16	0.477	0.486	0.445	0.390	0.395
	joint	0.535	0.534	0.399	0.417	0.414
Swiss Franc	1	0.284	0.226	0.130	0.125	**0.075**
	4	0.244	0.187	0.171	0.144	0.134
	8	0.167	0.143	0.167	0.168	0.171
	12	**0.059**	**0.056**	**0.090**	0.100	0.120
	16	**0.007**	**0.014**	**0.031**	**0.067**	**0.072**
	joint	**0.010**	**0.022**	**0.036**	0.109	0.103

NOTES: The table presents the marginal significance levels of the *t*-statistic of the slope coefficient in the long-horizon regression, of Theil's U-statistic and of the Diebold-Mariano statistic of equal out-of-sample forecast accuracy. Results are shown for alternative forecast horizons *k*. The entry *joint* refers to the distribution of the *sup* or the *inf* of the test statistic across *k*. The p-values for the one-sided test are generated under the null hypothesis based on nonparametric i.i.d. resampling of the residuals in the best-fitting vector autoregressive model. All estimated models are stationary. Boldface indicates significance at the 10 percent level.

Table 6: Probabilities of the observed augmented Dickey-Fuller test statistic for the fundamental having been generated under the null hypothesis and the alternative hypothesis

(a) Testing the null hypothesis of a unit root in the fundamental against the alternative of a stationary autoregressive process with intercept		
Currency	Probability of Type I Error (percent)	Probability of Type II Error (percent)
Canadian Dollar	18.75	59.55
German Mark	38.05	18.05
Japanese Yen	61.50	13.25
Swiss Franc	76.70	10.75

(b) Testing the null hypothesis of a unit root in the fundamental against the alternative of a stationary autoregressive process with intercept and deterministic linear time trend		
Currency	Probability of Type I Error (percent)	Probability of Type II Error (percent)
Canadian Dollar	41.75	35.60
German Mark	74.20	9.85
Japanese Yen	51.20	29.55
Swiss Franc	5.40	65.50

NOTES: The type-I error probability is the marginal significance level based on simulations of the test statistic under the best-fitting difference stationary model. Given this type-I error, the type-II error probability is based on the distribution of the test statistic under the best-fitting level-stationary model. Both models were selected using the sequential t-test of Ng and Perron (1995) allowing for up to 8 augmented lags. The results are based on the extended data set for 1973:II-1994:IV. All distributions are based on 2,000 replications. The qualitative results are robust to bias-corrections of the bootstrap data generating process.

Table 7: (a) Bootstrap p-values from data generating process restricted under the joint null hypothesis of no cointegration, a stationary AR process for the fundamental and a random walk with drift in the exchange rate

				Relative to random walk with drift		
	k	t(20)	t(A)	U	DM(20)	DM(A)
Canadian Dollar	1	0.114	0.130	**0.060**	**0.059**	**0.064**
	4	0.219	0.219	**0.076**	**0.048**	**0.065**
	8	0.257	0.240	**0.064**	**0.023**	**0.021**
	12	0.425	0.442	0.140	0.115	0.104
	16	0.526	0.547	0.254	0.357	0.421
	joint	0.349	0.422	0.159	**0.076**	**0.072**

(b) Bootstrap p-values from data generating process restricted under the joint null hypothesis of no cointegration, a trend-stationary AR process for the fundamental and a random walk with drift in the exchange rate

				Relative to random walk with drift		
	k	t(20)	t(A)	U	DM(20)	DM(A)
Swiss Franc	1	0.357	0.267	0.145	0.147	**0.089**
	4	0.315	0.241	0.181	0.164	0.152
	8	0.214	0.194	0.193	0.188	0.190
	12	**0.081**	**0.082**	**0.096**	0.119	0.140
	16	**0.004**	**0.015**	**0.028**	**0.088**	**0.087**
	joint	**0.005**	**0.024**	**0.033**	0.135	0.116

NOTES: The tables present the marginal significance levels of the *t*-statistic of the slope coefficient in the long-horizon regression, of Theil's U-statistic and of the Diebold-Mariano statistic of equal out-of-sample forecast accuracy. Results are shown for alternative forecast horizons *k*. The entry *joint* refers to the distribution of the *sup* or the *inf* of the test statistic across *k*. The p-values for the one-sided test are generated under the null hypothesis based on nonparametric i.i.d. resampling of the residuals in the best-fitting vector autoregressive model. All estimated models are stationary. Boldface indicates significance at the 10 percent level.

Table 8: Size and Power of Nominal 10 Percent Long-Horizon Regression Tests

		(a) Size					(b) Power				
				Relative to random walk with drift					Relative to random walk with drift		
Currency	k	t(20)	t(A)	U	DM(20)	DM(A)	t(20)	t(A)	U	DM(20)	DM(A)
Canadian Dollar	1	0.054	0.056	0.068	0.084	0.076	0.370	0.342	0.280	0.282	0.276
	4	0.052	0.060	0.086	0.092	0.090	0.418	0.408	0.324	0.282	0.280
	8	0.064	0.060	0.080	0.096	0.100	0.410	0.400	0.324	0.248	0.254
	12	0.080	0.066	0.092	0.084	0.084	0.372	0.370	0.330	0.266	0.278
	16	0.090	0.084	0.072	0.080	0.082	0.344	0.336	0.312	0.214	0.218
	joint	0.082	0.076	0.068	0.084	0.082	0.366	0.366	0.338	0.240	0.244
German Mark	1	0.102	0.106	0.118	0.126	0.110	0.176	0.194	0.192	0.172	0.160
	4	0.092	0.098	0.126	0.108	0.106	0.172	0.184	0.186	0.182	0.162
	8	0.084	0.086	0.122	0.108	0.106	0.190	0.194	0.198	0.174	0.164
	12	0.084	0.088	0.118	0.104	0.104	0.192	0.202	0.198	0.178	0.184
	16	0.086	0.092	0.110	0.104	0.108	0.180	0.198	0.174	0.168	0.172
	joint	0.090	0.078	0.110	0.096	0.106	0.196	0.204	0.192	0.186	0.178
Japanese Yen	1	0.090	0.078	0.102	0.100	0.102	0.178	0.176	0.168	0.166	0.160
	4	0.072	0.084	0.104	0.096	0.098	0.160	0.168	0.176	0.170	0.164
	8	0.058	0.064	0.104	0.082	0.092	0.158	0.152	0.162	0.152	0.146
	12	0.076	0.068	0.110	0.094	0.100	0.172	0.154	0.172	0.154	0.154
	16	0.084	0.080	0.104	0.086	0.090	0.164	0.160	0.172	0.156	0.150
	joint	0.074	0.060	0.102	0.086	0.104	0.168	0.154	0.178	0.176	0.148
Swiss Franc	1	0.078	0.084	0.104	0.114	0.114	0.394	0.426	0.364	0.322	0.318
	4	0.088	0.084	0.102	0.106	0.096	0.436	0.440	0.374	0.336	0.320
	8	0.084	0.082	0.088	0.086	0.080	0.370	0.358	0.334	0.296	0.292
	12	0.092	0.090	0.088	0.092	0.086	0.312	0.276	0.304	0.256	0.266
	16	0.094	0.078	0.096	0.086	0.092	0.224	0.206	0.250	0.226	0.226
	joint	0.080	0.080	0.096	0.096	0.110	0.318	0.276	0.292	0.276	0.268

NOTES: The table presents the rejection rates of the one-sided nominal 10 percent test based on bootstrap *p*-values. The data generating processes are based on the best-fitting VEC model under H_0 and under H_1 for each country. Results are based on 500 Monte Carlo trials with 2,000 bootstrap replications in each trial. The sample size is T=85. The Monte Carlo simulation error of the size estimate is approximately 0.0134.

Table 9: Power of Nominal 10 Percent Long-Horizon Regression Tests Against VEC Model with Additional Speculative Fads in the Exchange Rate

		Power				
				Relative to random walk with drift		
Currency	k	t(20)	t(A)	U	DM(20)	DM(A)
Canadian Dollar	1	0.382	0.370	0.268	0.260	0.248
	4	0.432	0.460	0.350	0.310	0.314
	8	0.420	0.414	0.344	0.272	0.282
	12	0.382	0.368	0.332	0.272	0.262
	16	0.318	0.306	0.310	0.232	0.234
	joint	0.362	0.340	0.334	0.250	0.254
Swiss Franc	1	0.428	0.492	0.384	0.332	0.332
	4	0.438	0.452	0.402	0.344	0.334
	8	0.366	0.360	0.358	0.288	0.282
	12	0.304	0.284	0.310	0.278	0.280
	16	0.246	0.228	0.230	0.246	0.236
	joint	0.326	0.280	0.294	0.308	0.300

NOTES: The table presents the rejection rates of the one-sided nominal 10 percent test based on bootstrap *p*-values. The data generating process is an unobserved components model. The exchange rate is modeled as the sum a long-run component based on the monetary model of exchange rate determination and a transitory speculative component with a half-life of two years. Details are described in the text. Results are based on 500 Monte Carlo trials with 2,000 bootstrap replications in each trial. The sample size is T=85.