Does the Absence of Cointegration Explain the Typical Findings in Long Horizon Regressions?

Robert-Paul Berben*
Department of Applied Economics, University of Nijmegen

Dick van Dijk†
Tinbergen Institute, Erasmus University Rotterdam

April 6, 1998

Abstract

One of the stylized facts in financial and international economics is that of increasing predictability of variables such as exchange rates and stock returns at longer horizons. This fact is based upon applications of long horizon regressions, from which the typical findings are that the point estimates of the regression parameter, the associated t-statistic, and the regression R^2 all tend to increase as the horizon increases. Such long horizon regression analyses implicitly assume the existence of cointegration between the variables involved. In this paper, we investigate the consequences of dropping this assumption. In particular, we look upon the long horizon regression as a conditional error-correction model and interpret the test for long horizon predictability as a single equation test for cointegration. We derive the asymptotic distributions of the estimator of the regression parameter and its t-statistic for arbitrary horizons, under the null hypothesis of no cointegration. It is shown that these distributions provide an alternative explanation for at least part of the typical findings. Furthermore, the distributions are used to derive a Phillips-Perron type correction to the ordinary least-squares t-statistic in order to endow it with a stable size for given, arbitrary, horizon. A local asymptotic power analysis reveals that the power of long horizon regression tests does not increase with the horizon. Exchange rate data are used to demonstrate the empirical relevance of our theoretical results.

Keywords: long horizon regression, error-correction model, cointegration, local power analysis, monetary exchange rate model, exchange rate forecasting

JEL classification: C22, C51, F31, F47

*Department of Applied Economics, University of Nijmegen, P.O. Box 9108, NL-6500 HK, Nijmegen, The Netherlands, email: r.berben@bw.kun.nl

†Tinbergen Institute, Erasmus University of Rotterdam, P.O. Box 1738, NL-3000 DR, Rotterdam, The Netherlands, email: djvandijk@few.eur.nl

We thank Eelke de Jong and Philip Hans Franses for their helpful comments and suggestions. Responsibility for any remaining errors is entirely ours. Berben acknowledges support from the Economics Research Foundation (ESR), which is part of the Netherlands Organization for Scientific Research (NWO).

1 Introduction

One of the stylized facts in financial and international economics is that variables, such as exchange rates and stock prices, which are not predictable at short horizons of up to a year, say, contain a significant predictable component at longer horizons. The statistical evidence upon which this claim is built is obtained from long horizon regressions, relating the change in the economic variable of interest k periods ahead to its current deviation from an equilibrium or fundamental value, which usually is provided by economic theory[1]. The typical findings from applying such regressions are that the point estimates of the regression parameter, the associated t-statistic to test its significance, and the regression R^2 all increase (in absolute value) as the horizon k increases. Long horizon regressions have been used to analyse the term structure of interest rates (Fama and Bliss, 1987, Fama, 1990, Campbell and Shiller, 1991), stock returns and dividend yields (Campbell and Shiller, 1988, Fama and French, 1988), nominal exchange rates and fundamentals (Mark, 1995, Chinn and Meese, 1995), and real exchange rates (Mark and Choi, 1997), among others.

Economic explanations for the apparent long-horizon predictability tend to be taken from the literature on the micro-structure of financial markets (Cutler *et al.*, 1990, De Long *et al.*, 1990). The change in an economic variable is thought to be largely determined by fads in investors' behaviour in the short run, while in the long run economic fundamentals play a more important role. For example, Mark (1995, p.215) suggests that "while short horizon changes tend to be dominated by noise, the noise is apparently averaged out over time, thus revealing systematic exchange-rate movements that are determined by economic fundamentals", see also Fama and French (1988) and Poterba and Summers (1988) for similar arguments. Frequently, this line of reasoning is reversed to argue that the power of tests of predictability increases as the horizon increases.

Attempts to explain the apparent success of the long horizon regression methodology with statistical arguments have centered on the small sample bias in both the regression coefficient and its standard error, which are caused by endogeneity of the regressor and

[1] The term 'long horizon regressions' has also been used to characterize regressions in which the change in one economic variable k periods ahead is related to (the change in) another, predetermined, macroeconomic variable (see Mishkin, 1990*a*,*b*, and Boudoukh and Richardson, 1993, among others). The statistical intricacies of such regressions are analysed in Richardson and Stock (1989), Boudoukh and Richardson (1994), and Campbell *et al.* (1997, Sec. 7.2).

the use of overlapping observations. If these are not properly accounted for, long horizon regressions are biased towards finding predictability at longer horizons, see Hodrick (1992), Nelson and Kim (1993), Goetzmann and Jorion (1993), Bollerslev and Hodrick (1995), and Kilian (1997), among others.

Most of this discussion around long horizon regressions has implicitly assumed the existence of a long-run relationship between the variable of interest and its fundamental. Put differently, since the variables involved usually are (assumed to be) nonstationary, it is assumed that they are cointegrated. If this is indeed the case, the test for long-horizon predictability is in fact a test for weak exogeneity of the economic variable. The relevant asymptotic distribution then simply is Gaussian, and the observed systematic deviations in the empirical results from long horizon regressions have to be caused by small sample effects.

In many applications the existence of cointegration is in fact a heroic assumption to make. Many papers start with a cointegration analysis of the variables involved, and in many cases no or very little evidence in favor of cointegration is found. If cointegration does not hold, the test for long horizon predictability itself becomes a test for cointegration between the economic variable and its fundamental. The relevant asymptotic distribution then no longer is Gaussian, but rather a mixture of a Gaussian and a Dickey-Fuller distribution, cf. Kremers *et al.* (1992) and Zivot (1996).

In the present paper we aim to investigate whether this alternative asymptotic distribution theory can help explain the empirical results from long horizon regressions. In particular, we derive the asymptotic distributions of the slope parameter and the corresponding t-statistic for general horizon k under the absence of cointegration. After applying a Phillips-Perron type correction, the asymptotic distribution of the test for horizon equal to one period turns out to be equal to Hansen's (1995) covariate augmented Dickey-Fuller distribution. Furthermore, the distribution of the regression coefficient is 'blown up' proportionally to the horizon k. The distribution for the naive least-squares based t-statistic is rescaled with $\sqrt{k}$, while the distribution of the t-statistic which correctly accounts for the presence of serial correlation is invariant across horizons. However, a Monte Carlo study shows that in small samples the latter t-statistic with a Phillips-Perron type correction appears to suffer from substantial size distortions.

In addition, we analyse the asymptotic power of long horizon regression tests against local alternatives and demonstrate that, for a given local alternative, local power is independent of the horizon. Thus, from an asymptotic point of view there is no reason for using long horizon regression tests.

The remainder of this paper is organised as follows. We start in Section 2 with an empirical illustration to introduce the concept of long horizon regressions and to demonstrate the characteristic findings of such regressions. In Section 3 the asymptotic distributions of the statistics under the null hypothesis of no cointegration are derived. In Section 4 we analyse the power of long horizon regression tests against local, near-cointegrated, alternatives. In Section 5, the empirical illustration is revisited and re-analyzed in the light of the theoretical findings. Finally, Section 6 concludes. Proofs of the Theorems are gathered in the Appendix.

2 Empirical illustration: exchange rates and economic fundamentals

Long horizon regressions relate the change in an economic variable y_t k periods ahead to its current deviation from a presumed equilibrium value x_t, i.e.,

$$\Delta_k y_{t+k} = \beta_k (y_t - x_t) + e_{t+k,k}, \qquad t = 1, 2, \ldots, T-k, \tag{1}$$

where $\Delta_k y_{t+k} \equiv y_{t+k} - y_t$ for all $k > 0$ and T is the sample size. The properties of the error term $e_{t+k,k}$ will be discussed at length in Section 3. The predictability of y_t at horizon k is assessed by testing the null hypothesis $H_0 : \beta_k = 0$ against the alternative $H_1 : \beta_k < 0$. When long horizon regressions of the form (1) are applied to economic data, the outcome invariably appears to be that the estimated regression coefficient $\hat{\beta}_k$, the t-statistic $t(\hat{\beta}_k)$, and the regression R^2, denoted as R_k^2, all increase in absolute value as the horizon k increases. This usually is taken as evidence that x_t helps predicting long-run movements in y_t better than short-run changes. In this section we illustrate these typical findings by means of an example involving the relationship between changes in the nominal exchange rate and an economic fundamental, cf. Mark (1995) and Chinn and Meese (1995).

The basic monetary exchange rate model with flexible prices[2] assumes that purchasing power parity and uncovered interest rate parity hold, and that log money demand is

[2]See Frenkel (1976), Mussa (1976), and Bilson (1978).

static and depends linearly on real income and the nominal interest rate. Using rational expectations the model implies that for two identical countries the log exchange rate s_t, expressed as the number of units of the domestic currency per unit of foreign currency, is determined as

$$s_t = \frac{1}{1+\omega} \sum_{i=0}^{\infty} \left(\frac{\omega}{1+\omega} \right)^i E_t(f_{t+i}), \tag{2}$$

where the economic fundamental f_t is given by

$$f_t = (m_t - m_t^*) - \lambda(y_t - y_t^*), \tag{3}$$

where m_t denotes the log of the domestic money stock and y_t the log of domestic real income. An asterisk indicates a foreign variable. The parameters ω and λ denote the common semi-interest elasticity and common income elasticity of money demand, respectively. Subtracting f_t from both sides of (2) yields, after some rearrangements,

$$s_t - f_t = \sum_{i=1}^{\infty} \left(\frac{\omega}{1+\omega} \right)^i E_t(\Delta f_{t+i}). \tag{4}$$

Under the assumption that f_t is stationary in first differences, it is seen from (2) that s_t is nonstationary, and from (4) it follows that $s_t - f_t$ is stationary. Hence, the log exchange rate and the fundamental are cointegrated with cointegrating vector equal to $(1, -1)$. Thus, the exchange rate might be expected to react to deviations from its fundamental value, according to an error-correction mechanism

$$\Delta s_{t+1} = \beta(s_t - f_t) + e_{t+1}. \tag{5}$$

If (5) is estimated using post-Bretton Woods data, the hypothesis that $\beta = 0$ typically can not be rejected, see Meese (1990) and Frankel and Rose (1995) and the references cited therein. This is usually interpreted as evidence for the fact that in the short run, changes in the exchange rate are determined by factors other than the economic fundamentals in f_t. However, at the same time it is believed that the exchange rate can not wander around indefinitely. In the long run, it should ultimately revert to its equilibrium value. Hence, the change in the exchange rate measured over longer horizons should be, at least partly, predictable from its current deviation from equilibrium. This is the basic motivation given by Mark (1995) and Chinn and Meese (1995) for considering long horizon regressions for exchange rate data, i.e., (1) with $y_t = s_t$ and $x_t = f_t$.

We replicate the analysis of Mark (1995)[3] on an extended sample period. Quarterly observations covering 1973Q1-1997Q3 on nominal money supply M1, real gross domestic product [GDP] and end-of-period exchange rates *vis-à-vis* the US dollar for Germany, Canada, Japan and Switzerland are obtained from the OECD *Main Economic Indicators*. The real GDP data are seasonally adjusted, while the money supply data is not. We apply the same procedure as Mark (1995) to eliminate the seasonality in money supply by substituting all observations with the average money supply during the current and previous three quarters. The effective sample size then is equal to 94 observations[4]. To construct the fundamental f_t as given in (3), we follow Mark (1995) and set $\lambda = 1$. A constant α_k is included in the regression (1) as well.

Table 1 reports some key results from estimating the long horizon regression (1) for $k = 1, 4, 8, 12$, and 16, i.e., for horizons varying between one quarter and four years. Several different t-statistics for testing the significance of β_k are reported. Under the null hypothesis, the errors $e_{t,k}$ in the regression for $k > 1$ are autocorrelated by construction, which causes the usual least squares estimate of the residual variance to be biased downwards. This in turn inflates the t-statistic associated with β_k, denoted as $t_{LS}(\hat{\beta}_k)$. To account for this serial correlation, we estimate the residual variance[5] by making use of a nonparametric estimator, popularized by Newey and West (1987), i.e.,

$$\hat{\sigma}^2_{e_{t,k}} = \frac{1}{T}\sum_{t=k}^{T} e^2_{t,k} + 2\sum_{j=1}^{l} w(j,l)\frac{1}{T}\sum_{t=k+j}^{T} e_{t,k}e_{t-j,k}, \tag{6}$$

where $w(j,l)$ is a kernel function which assigns weights to the covariances, and l is the truncation lag. Here we use the Bartlett kernel[6], i.e., $w(j,l) = 1 - j/(l+1)$. The rule-of-thumb of Andrews (1991) is the most popular method for choosing the truncation lag l, while l equal to some large fixed value or $l = k-1$ are also used. Here we report results

[3] Chinn and Meese (1995) consider alternative exchange rate models, which entail incorporating additional terms in the fundamental f_t, such as nominal interest rates, the inflation rate, and real trade balances. Results obtained from these models are comparable to those presented here.

[4] For Switzerland, the GDP series is available only up to 1996Q2. Hence, for this country the effective sample consists of 89 observations.

[5] Notice that we deviate from standard practice here, which is to estimate the entire covariance matrix of (α_k, β_k) by means of a serial correlation robust estimator. Since we interpret the test of long-horizon predictability as a test for cointegration, the regressor $s_t - f_t$ is nonstationary under the null hypothesis, which invalidates the use of such estimators.

[6] Uniform weights for the covariances, (i.e. $w(j,l) = 1$ for all $j = 1,\ldots,l$ in (6)), cf. Hansen and Hodrick (1980), are also commonly applied, see Fama and French (1988), among others. To save space, we do not report results involving this estimator here, as they are very similar to the ones obtained with the Newey-West estimator. These results are available upon request.

for $l = 20$ and l set according to Andrews' rule, cf. Mark (1995).

- insert Table 1 about here -

From Table 1 it is seen that for all currencies $\hat{\beta}_k$, the least-squares based t-statistic $t_{LS}(\hat{\beta}_k)$, the Newey-West based t statistics $t_A(\hat{\beta}_k)$ and $t_{20}(\hat{\beta}_k)$, and the regression R^2 all increase in absolute value as the horizon k increases. Although the increase is not as convincing as in Mark (1995)[7], we might still be tempted to conclude that "the improved fit as k increases suggests that the noise that dominates quarter-to-quarter changes in s_t averages out over long horizons"(Mark, 1995, p.210). In Section 5 we discuss whether this conclusion is valid or not by assessing the significance of the various t-statistics. In the next two sections, we first turn our attention to the asymptotic distributions of $\hat{\beta}_k$ and $t(\hat{\beta}_k)$ under the null hypothesis of no cointegration and under local alternatives. These two sections can be skipped by readers who are not interested in the technical details. A brief summary of the main results is given at the beginning of Section 5.

3 Asymptotic distributions in long horizon regressions

If long horizon regressions are applied under the (implicit) assumption of cointegration, the test for predictability is in fact a test for weak exogeneity of the economic variable. The relevant asymptotic distribution in that case is Gaussian. If the fundamental is weakly exogenous however, the test for predictability becomes a test for cointegration, which changes the relevant asymptotic distribution theory. The aim of this section is twofold. First, we explore whether this alternative asymptotic distribution theory can provide an alternative explanation for the observed increase in the absolute values of the $\hat{\beta}_k$'s and their t-statistics with k. In particular, we derive the asymptotic distributions of these statistics under the null of no cointegration and inspect the dependence of these distributions on the horizon k. Second, we assess whether these asymptotic distributions provide a reliable base for inference in small samples. Towards this end, we conduct a limited Monte Carlo study.

[7]The sample period considered by Mark (1995) runs to 1991Q4. The difference in the results appears to be caused mainly by a decline in the US money supply since 1994, which is not matched by an appropriate change in other variables entering the fundamental or in the exchange rates.

This section is divided into two parts. In the first subsection we focus on the long horizon regression with horizon equal to one, i.e., a basic error-correction model. In the second subsection we generalize the results on the asymptotic distributions to arbitrary horizons, which allows us to examine the relationship between the statistics for different horizons.

3.1 Short horizon regression

Taking the horizon k in (1) equal to one results in the single equation conditional error-correction model [CECM]

$$\Delta y_{t+1} = \beta(y_t - x_t) + e_{t+1}, \tag{7}$$

where the subscripts 1 on β and e_t have been omitted to economize on notation. See Boswijk (1992) for a thorough introduction to CECMs. It will prove useful later on to rewrite the CECM (7) as

$$\Delta z_{t+1} = \beta z_t + v_{t+1}, \tag{8}$$

where $z_t \equiv y_t - x_t$ and $v_t = e_t - \Delta x_t$, cf. Zivot (1996).

The test of predictability of y_t is expressed as $H_0 : \beta = 0$ versus $H_1 : \beta < 0$. It follows that, if x_t is assumed to be weakly exogenous, this is in fact a test for cointegration between x_t and y_t with a pre-specified cointegrating vector equal to (1,-1). Kremers *et al.* (1992) suggest to use the t-statistic $t(\hat{\beta})$ to test for cointegration in (7), see also Boswijk (1994) for an extensive discussion on testing for cointegration in CECMs. The asymptotic distribution Kremers *et al.* (1992) derive depends on a nuisance parameter which can take on any positive value and, therefore, is of little value in practice. Zivot (1996) provides an alternative representation of the asymptotic distribution by exploiting the link between testing for cointegration in (7) with the covariate augmented Dickey-Fuller [CADF] test for a unit root developed by Hansen (1995). In this way, the asymptotic distribution, which is a mixture of a Dickey-Fuller unit root distribution and a standard normal distribution, depends on a nuisance parameter which takes on values in the unit interval and can be estimated consistently.

Hansen (1995) and Zivot (1996) allow an arbitrary number of lagged values of both Δy_t and Δx_t to enter (7), which justifies their assumption that e_t is neither autocorrelated nor correlated with Δx_t. However, as the short run dynamics are not modelled explicitly in the long horizon regression (1), we are forced to assume that these are captured by the error terms e_t and v_t. Hence our analysis is based on the approach by Phillips and Perron (1988) by imposing only certain mixing conditions on $(\Delta x_t, e_t)$. In particular, we invoke the following assumptions, cf. Hansen (1995) and Herrndorf (1984).

Assumption 1: For some $p > r > 2$,

1. $(\Delta x_t, e_t)$ is covariance stationary and strong mixing with mixing coefficients α_m, which satisfy $\sum_{m=1}^{\infty} \alpha_m^{1/r-1/p} < \infty$;

2. $\sup_t E(|\Delta x_t|^p + |e_t|^p) < \infty$.

These assumptions have become conventional in the time series literature (see Phillips, 1987). They appear to be suitable for financial data (exchange rates, interest rates, stock returns) as they allow for a fair amount of heterogeneity and impose only mild restrictions on moments and temporal dependence.

Define the long run covariance matrix as

$$\Omega = \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \sum_{l=-\infty}^{\infty} E \begin{pmatrix} v_t \\ e_t \end{pmatrix} (v_{t-l} \;\; e_{t-l}) = \begin{pmatrix} \omega_v^2 & \omega_{ve} \\ \omega_{ve} & \omega_e^2 \end{pmatrix}, \tag{9}$$

which can be decomposed as $\Omega = \Sigma + \Lambda + \Lambda'$, where

$$\Sigma = \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} E \begin{pmatrix} v_t \\ e_t \end{pmatrix} (v_t \;\; e_t) = \begin{pmatrix} \sigma_v^2 & \sigma_{ve} \\ \sigma_{ve} & \sigma_e^2 \end{pmatrix}, \tag{10}$$

$$\Lambda = \lim_{T\to\infty} \frac{1}{T} \sum_{t=2}^{T} \sum_{l=1}^{t-1} E \begin{pmatrix} v_{t-l} \\ e_{t-l} \end{pmatrix} (v_t \;\; e_t). \tag{11}$$

In addition define the long-run correlations

$$\rho^2 = \omega_{ve}^2 / \omega_e^2 \omega_v^2, \tag{12}$$

$$R^2 = \omega_e^2 / \omega_v^2. \tag{13}$$

Hence, ρ^2 is the squared long-run correlation between e_t and v_t, while R^2 indicates the proportion of the long-run variance of v_t explained by e_t. Note that, if y_t and x_t are uncorrelated random walks, ρ^2 is equal to R^2.

It follows from the above that

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{\lfloor rT \rfloor} \begin{pmatrix} v_t \\ e_t \end{pmatrix} \Rightarrow \begin{pmatrix} \omega_v W_v(r) \\ \omega_e(\rho W_v(r) + \sqrt{1-\rho^2} W_{e\cdot v}(r)) \end{pmatrix}, \tag{14}$$

where $W_v(r)$ and $W_{e\cdot v}(r)$ are two independent Wiener processes, $\lfloor\cdot\rfloor$ denotes integer part, and $\Rightarrow$ denotes weak convergence. Throughout the rest of the paper, we frequently use the shorthand notation $\int_0^1 W_v^2$ to denote the integral with respect to Lebesque measure, i.e., $\int_0^1 W_v(r)^2 dr$. The next theorem discusses the asymptotic behaviour of the least squares estimator of β in (7) and the least-squares based and Newey-West type t-statistics for testing the null hypothesis $H_0 : \beta = 0$.

Theorem 1 (Asymptotic behaviour of $\hat{\beta}$ and $t(\hat{\beta})$) *Under the null hypothesis $\beta = 0$ and assuming that x_t is weakly exogenous,*

$$\begin{aligned}
T(\hat{\beta}-\beta) &\Rightarrow R\left(\rho\frac{\int_0^1 W_v dW_v}{\int_0^1 W_v^2} + \sqrt{1-\rho^2}\frac{\int_0^1 W_v dW_{e\cdot v}}{\int_0^1 W_v^2}\right) + \frac{1}{\omega_v^2}\frac{\Lambda_{ve}}{\int_0^1 W_v^2}, &(15)\\
t_{LS}(\hat{\beta}) &\Rightarrow \frac{\omega_e}{\sigma_e}\left(\rho\frac{\int_0^1 W_v dW_v}{(\int_0^1 W_v^2)^{1/2}} + \sqrt{1-\rho^2}\mathcal{N}(0,1)\right) + \frac{1}{\omega_v\sigma_e}\frac{\Lambda_{ve}}{(\int_0^1 W_v^2)^{1/2}}, &(16)\\
t_{NW}(\hat{\beta}) &\Rightarrow \rho\frac{\int_0^1 W_v dW_v}{(\int_0^1 W_v^2)^{1/2}} + \sqrt{1-\rho^2}\mathcal{N}(0,1) + \frac{1}{\omega_v\omega_e}\frac{\Lambda_{ve}}{(\int_0^1 W_v^2)^{1/2}}, &(17)
\end{aligned}$$

where Λ_{ve} is the upper right element of Λ.

The asymptotic distribution for the t-statistics is a linear combination of the standard normal and the Dickey-Fuller unit root distribution. The relative importance of these depends on the value of ρ: if $\rho = 1$, the distribution collapses to the Dickey-Fuller distribution, while the distribution approaches the standard normal as ρ tends to zero. These distributions differ from Hansen's (1995) CADF distribution only with respect to the last term and a scaling factor ω_e/σ_e in the case of the least-squares based t-statistic. The formulae closely resemble the formulae by Phillips and Perron (1988) for the case of univariate unit root tests. All covariance matrices defined above can be consistently estimated by Newey-West type estimators of the type given in (6), which also allows to obtain a consistent estimate of the nuisance parameter ρ. Hence, it is possible to apply a correction to the conventional t-statistic along the same lines as Phillips and Perron (1988) in order to be able to use the critical values of Hansen. This is shown in Corollary 1.

Corollary 1 (Asymptotic behaviour modified $t(\hat{\beta})$)

$$\begin{aligned}
\frac{\hat{\sigma}_e}{\hat{\omega}_e}t_{LS}(\hat{\beta}) - \frac{\hat{\Lambda}_{ve}}{\hat{\omega}_e(T^{-2}\sum_{t=1}^{T-1}(y_t-x_t)^2)^{1/2}} &\Rightarrow \rho\frac{\int_0^1 W_v dW_v}{(\int_0^1 W_v^2)^{1/2}} + \sqrt{1-\rho^2}\mathcal{N}(0,1),\\
t_{NW}(\hat{\beta}) - \frac{\hat{\Lambda}_{ve}}{\hat{\omega}_e(T^{-2}\sum_{t=1}^{T-1}(y_t-x_t)^2)^{1/2}} &\Rightarrow \rho\frac{\int_0^1 W_v dW_v}{(\int_0^1 W_v^2)^{1/2}} + \sqrt{1-\rho^2}\mathcal{N}(0,1).
\end{aligned}$$

Finally it should be remarked that deterministic terms, like a constant α in the CECM can be incorporated in the same fashion as in Hansen (1995) and Zivot (1996).

3.2 Broadening the Horizons

Under the null hypothesis $\beta = 0$, the long horizon regression (1) can be rewritten as

$$\Delta_k z_{t+k} = \beta_k z_t + v_{t+k} + \ldots + v_{t+1}, \tag{18}$$

where, as before, $z_t \equiv y_t - x_t$ and $v_t = e_t - \Delta x_t$. This result is used to proof the following theorem, which gives the asymptotic distributions for the long horizon regression statistics for general horizon k under the null hypothesis of no cointegration.

Theorem 2 (Asymptotic behaviour of $\hat{\beta}_k$ and $t(\hat{\beta}_k)$) *Under the null hypothesis $\beta = 0$, and assuming that x_t is weakly exogenous,*

$$\begin{aligned} T(\hat{\beta}_k - \beta_k) \quad &\Rightarrow \quad kR\left(\rho\frac{\int_0^1 W_v dW_v}{\int_0^1 W_v^2} + \sqrt{1-\rho^2}\frac{\int_0^1 W_v dW_{e\cdot v}}{\int_0^1 W_v^2}\right) + \\ &\qquad k\frac{1}{\omega_v^2}\left(\frac{\Lambda_{ve}}{\int_0^1 W_v^2} - \frac{\Lambda_{ve,k-1}}{\int_0^1 W_v^2}\right), \end{aligned} \tag{19}$$

$$\begin{aligned} t_{LS}(\hat{\beta}_k) \quad &\Rightarrow \quad \frac{k\omega_e}{\sqrt{k\sigma_e^2 + 2\sum_{j=1}^{k-1}\Lambda_{ee,j}}}\left(\rho\frac{\int_0^1 W_v dW_v}{(\int_0^1 W_v^2)^{1/2}} + \sqrt{1-\rho^2}\mathcal{N}(0,1)\right) + \\ &\qquad \frac{k}{\sqrt{k\sigma_e^2 + 2\sum_{j=1}^{k-1}\Lambda_{ee,j}}}\frac{1}{\omega_v}\left(\frac{\Lambda_{ve}}{(\int_0^1 W_v^2)^{1/2}} - \frac{\Lambda_{ve,k-1}}{(\int_0^1 W_v^2)^{1/2}}\right), \end{aligned} \tag{20}$$

$$\begin{aligned} t_{NW}(\hat{\beta}_k) \quad &\Rightarrow \quad \left(\rho\frac{\int_0^1 W_v dW_v}{(\int_0^1 W_v^2)^{1/2}} + \sqrt{1-\rho^2}\mathcal{N}(0,1)\right) + \\ &\qquad \frac{1}{\omega_v\omega_e}\left(\frac{\Lambda_{ve}}{(\int_0^1 W_v^2)^{1/2}} - \frac{\Lambda_{ve,k-1}}{(\int_0^1 W_v^2)^{1/2}}\right), \end{aligned} \tag{21}$$

where

$$\Lambda_{ve,k-1} \quad = \quad \lim_{T\to\infty} T^{-1}\sum_{t=k+1}^{T}\sum_{i=1}^{k-1} Ev_{t-i}e_t, \tag{22}$$

$$\Lambda_{ee,k-1} \quad = \quad \lim_{T\to\infty} T^{-1}\sum_{t=k+1}^{T}\sum_{i=1}^{k-1} Ee_{t-i}e_t. \tag{23}$$

Corollary 2 *If e_t is not autocorrelated then*

$$t_{LS}(\hat{\beta}_k) \Rightarrow \sqrt{k}\left(\rho\frac{\int_0^1 W_v dW_v}{(\int_0^1 W_v^2)^{1/2}} + \sqrt{1-\rho^2}\mathcal{N}(0,1)\right) + \sqrt{k}\frac{1}{\omega_v\omega_e}\left(\frac{\Lambda_{ve}}{(\int_0^1 W_v^2)^{1/2}} - \frac{\Lambda_{ve,k-1}}{(\int_0^1 W_v^2)^{1/2}}\right).$$

From (19) it is seen that the asymptotic distribution of the regression coefficient $\hat{\beta}_k$ is k times the distribution of $\hat{\beta}_1$, apart from the last correction term. Similarly, the distribution of the least squares based t-statistic depends on the horizon k. Corollary 2 demonstrates that if e_t is uncorrelated, the scaling factor is $\sqrt{k}$. Notice that these asymptotic distributions explain the increase in the absolute value of estimates of β_k and $t_{LS}(\hat{\beta}_k)$ with k. Given that the Dickey-Fuller distribution is skewed to the right, it is not surprising to find a negative estimate for β_1 and $t(\hat{\beta}_1)$ on average. The dependence of the asymptotic distributions of $\hat{\beta}_k$ and $t_{LS}(\hat{\beta}_k)$ on k ensure that these statistics become more negative as the horizon increases. The asymptotic results presented here corroborate simulation evidence presented in Berkowitz and Giorgianni (1997) showing that even if cointegration does not hold the means of $\hat{\beta}_k$ and $t_{LS}(\hat{\beta}_k)$ become more negative as k increases.

If Newey-West type covariance matrix estimators are used, the asymptotic distribution of the t-statistic still depends on k through the last correlation term in (21). The effect of this correction term is much smaller however than the scaling factors involved in the distributions for the parameter estimate and the least-squared based t-statistic. In the case of uncorrelated errors, this correction term disappears and the asymptotic distribution of the t-statistic is independent of k. Hence, for Newey-West based t-statistics the asymptotic distributions can not clearly explain the typical findings. The Monte Carlo results in Berkowitz and Giorgianni (1997) however suggest that the small-sample distributions of these statistics does shift to the left as k increases.

Corollary 3 shows how to modify $t_{LS}(\hat{\beta}_k)$ and $t_{NW}(\hat{\beta}_k)$ as k increases. Due to these modifications, it is possible to use Hansen's asymptotic CADF distribution, as tabulated in his Table 1, for arbitrary k.

Corollary 3 (Asymptotic behaviour of modified $t(\hat{\beta}_k)$)

$$\frac{\sqrt{k\hat{\sigma}_e^2 + 2\sum_{j=1}^{k-1}\hat{\Lambda}_{ee,j}}}{k\hat{\omega}_e} t_{LS}(\hat{\beta}_k) - \frac{\hat{\Lambda}_{ve} - \hat{\Lambda}_{ve,k-1}}{\hat{\omega}_e(T^{-2}\sum_{t=1}^{T-k}(y_t - x_t)^2)^{1/2}} \Rightarrow \rho \frac{\int_0^1 W_v dW_v}{(\int_0^1 W_v^2)^{1/2}} + \sqrt{1-\rho^2}\mathcal{N}(0,1),$$

$$t_{NW}(\hat{\beta}_k) - \frac{\hat{\Lambda}_{ve} - \hat{\Lambda}_{ve,k-1}}{\hat{\omega}_e(T^{-2}\sum_{t=1}^{T-k}(y_t - x_t)^2)^{1/2}} \Rightarrow \rho \frac{\int_0^1 W_v dW_v}{(\int_0^1 W_v^2)^{1/2}} + \sqrt{1-\rho^2}\mathcal{N}(0,1).$$

In order to examine the relationship between the different R^2's of the long horizon regression, we return to equation (1). For arbitrary k and fixed T it follows that

$$R_k^2 = 1 - \frac{\sum_{t=1}^{T-k}(\hat{e}_{t+k,k})^2}{\sum_{t=1}^{T-k}(\hat{e}_{t+k,k})^2 + \hat{\beta}_k^2 \sum_{t=1}^{T-k}(y_t - x_t)^2}. \tag{24}$$

Under the null hypothesis of no cointegration, R_k^2 converges to zero for all horizons. However, if k increases, R_k^2 tends to go to zero more slowly, which can be seen from

$$R_k^2 \Rightarrow 1 - \frac{k\sigma_e^2 + 2\sum_{j=1}^{k-1} \Lambda_{ee,j}}{k\sigma_e^2 + 2\sum_{j=1}^{k-1} \Lambda_{ee,j} + (k/T)T^2\hat{\beta}^2 \int_0^1 W_v^2}. \tag{25}$$

We perform a limited Monte Carlo experiments to investigate whether the asymptotic distributions derived above are suited for inference in small samples. We generate y_t and x_t as two possibly correlated random walks, $\Delta y_t = e_t$ and $\Delta x_t = u_t$, where the innovations of the random walks are generated according to

$$\begin{pmatrix} e_t \\ u_t \end{pmatrix} = \begin{pmatrix} \alpha_{11} & 0 \\ 0 & \alpha_{22} \end{pmatrix} \begin{pmatrix} e_{t-1} \\ u_{t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_t \\ \eta_t \end{pmatrix} \tag{26}$$

$$\begin{pmatrix} \varepsilon_t \\ \eta_t \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \sigma_{\varepsilon\eta} \\ \sigma_{\varepsilon\eta} & 1 \end{pmatrix} \right) \tag{27}$$

We consider six different data generating processes, with different degrees of serial correlation in and cross-correlation between e_t and u_t, by varying α_{11}, α_{22} and $\sigma_{\varepsilon\eta}$. A description of the various parameterizations can be found in Table 2. For each parameterization we generate 5000 artificial samples of 200 observations. Necessary starting values are set equal to zero. Each replication discards the first 100 observations to eliminate start-up effects; the results are based on the remaining 100 observations. We use Hansen's approach to calculate the nuisance parameter ρ for the various parameterizations, i.e., for each parameterization we generate ten samples of 5000 observations and use these to obtain an estimate of ρ. The averages of these estimates are given in Table 2 as well.

- insert Table 2 about here -

Rejection frequencies for the modified long horizon regression tests at nominal significance levels of 1%, 5%, and 10% are reported in Table 3. Some important conclusions can be drawn from this table. First, the empirical size of all the tests varies within considerable bounds for $k = 1$, cf. Hansen (1995). On the basis of this result, no clear-cut choice between the three t-statistics considered can be made. Second, as the horizon k increases, the size of the modified Newey-West type estimators tends to increase. Therefore, these t-statistics become increasingly unreliable as k increases. In contrast, the size of the least-squares based t-statistic is fairly stable across different horizons. One possible explanation for this phenomenon is that the Newey-West type estimators have great difficulty matching the k^2 factor of the long-run variances as k increases; typically they tend to underestimate the long-run variance. The expression for the modified least-squares based t-statistic given in Corollary 2 contains several explicit references to the value of the horizon k, which apparently enables this statistic to cope more succesfully with the increase in the long-run variance.

- insert Table 3 about here -

From this limited Monte Carlo study we draw the conclusion that the modified least-squares based t-statistic can be used for inference on $\hat{\beta}_k$ for arbitrary k. In Section 5 we will use this statistic to reconsider the evidence of long horizon predictability of the nominal exchange rate given in Section 2.

4 Local Asymptotic Power of Long Horizon Regression Tests

As noted in the introduction, one of the main reasons to use long horizon regression tests is the belief that the power of the t-statistic tests is increasing in k. However, only few papers exist that investigate this question, either theoretically or in an empirical setup. Using Monte Carlo experiments, Kilian (1997) concludes that, under the null hypothesis of cointegration, no power gains are to be expected. In this section we will investigate whether this is true in case of local alternatives, under the null hypothesis of no cointegration. Our results thus are complementary to those of Kilian.

We consider the asymptotic power of the long horizon regression tests against local alternatives of the form:

$$H_a : \beta = -c/T, \tag{28}$$

where c is a positive constant and T is the sample size. The null hypothesis $\beta = 0$ holds locally if $c > 0$ and $T \to \infty$.

We use local-to-unity (in our case, local-to-zero) asymptotics to derive asymptotic power functions, cf. Phillips (1987), among others. This theory employs diffusion representations of continuous time stochastic processes. In particular, the Ornstein-Uhlenbeck process is used, which is defined as the solution $W^c(r)$ to the following stochastic differential equation:

$$dW^c(r) = cW^c(r)dr + dW(r), \tag{29}$$

where $W(r)$ is a Wiener process.

Theorem 3 compares the asymptotic local power of the long horizon regression tests for different values of the horizon. In order to be able to get a fair comparison for different values of the horizon[8], we use the "inherent link" (Berkowitz and Giorgianni (1997, p.16) between β_k and β under the alternative hypothesis of near-cointegration

$$\beta_k = 1 - (1-\beta)^k. \tag{30}$$

If we substitute the local alternative for β into this relation, we arrive at

$$\begin{aligned} \beta_k &= 1-(1-\beta)^k = 1-(1+c/T)^k \\ &= 1 - \sum_{i=0}^{k} \begin{pmatrix} k \\ i \end{pmatrix} (c/T)^i = -kc/T + O(T^{-2}). \end{aligned} \tag{31}$$

Therefore the local alternative for a general horizon k is

$$H_{a_k} : \beta_k = -kc/T. \tag{32}$$

Theorem 3 (Local asymptotic power of modified t-statistics) *The local asymptotic power of the modified $t_{LS}(\hat{\beta}_k)$ and the modified $t_{NW}(\hat{\beta}_k)$ is independent of the value of the horizon k.*

This theoretical result is in line with the empirical observations of Kilian (1997). In his empirical examples he is unable to reject the hypothesis that the power of the long horizon regression tests is independent of the value of the horizon.

5 Exchange rates and economic fundamentals revisited

The example given in Section 2 on the relation between nominal exchange rates and an economic fundamental derived from the monetary model clearly demonstrates the typical findings from long horizon regression tests. Table 1 shows that for all currencies $\hat{\beta}_k$, the least-squares t-statistic $t_{LS}(\hat{\beta}_k)$, the Newey-West based t-statistics $t_A(\hat{\beta}_k)$ and $t_{20}(\hat{\beta}_k)$, and the regression R^2 all increase in absolute value as the horizon k increase, suggesting that long-run changes in exchange rates contain a significant predictable component.

[8] Berkowitz and Giorgianni (1997, footnote 3) note that "a proper power comparison requires that the alternative is kept fixed".

In Section 3.1 it was shown that the asymptotic distribution for the t-statistic in the 'short-horizon' regression, i.e., (1) with $k = 1$, is a linear combination of the standard normal and the Dickey-Fuller distribution. In particular, after applying a Phillips-Perron type correction, the asymptotic distribution turns out to be equal to Hansen's (1995) covariate augmented Dickey-Fuller distribution. Furthermore, as the horizon k increases, the distribution of the regression coefficient is 'blown up' proportionally to k. The distribution for the naive least-squares based t-statistic is rescaled with $\sqrt{k}$, while the distribution of the Newey-West based t-statistic, which correctly accounts for the presence of serial correlation, is invariant across horizons. The regression R^2 tends to zero for all horizons. However, the rate of convergence becomes slower as k increases, which explains the fact that in finite samples it is found that R_k^2 increases with k as well.

The results of the Monte Carlo study at the end of Section 3.2 show that in small samples inference based on Newey-West type estimators proves unreliable, as they suffer from severe size distortions, cf. Newey and West (1994). However, a modified least-squares based t-statistic turns out to have a stable size over the range of horizons considered. Here we use these modified t-statistics in order to investigate the significance of the empirical results presented in Section 2.

- insert Table 4 about here -

The various modified t-statistics are shown in Table 4. Comparing the $t_{LS}(\hat{\beta}_k)$ statistics with the critical values in Table 1 in Hansen (1995), it appears that there is not much evidence for long horizon predictability of any of the currencies considered. For the Deutsche mark and the Japanese yen, none of the t-statistics is significant even at the 10% significance level. For the Canadian dollar, only $\hat{\beta}_k$ or $k = 1$ is significant at the 5% level. Only for the Swiss franc there appears to be some weak evidence for long-run predictability: whereas the t-statistics at short-horizons are not or only marginally significant, the t-statistic at the four year horizon is significant at the 1% level. These results mimic the results by Berkowitz and Giorgianni (1997) who, using critical values derived by simulations, also find some weak evidence for long-run predictability for the Swiss franc.

6 Summary and conclusions

In applications of long horizon regressions it is typically found that the point estimates of the regression parameter, the associated t-statistic to test its significance, and the regression R^2 all tend to increase in absolute value as the horizon increases. In this paper we have explored the influence of the assumption concerning the cointegration properties of the variables involved on the conclusions which are to be drawn from these outcomes. In practice, cointegration is often implicitly assumed to hold. Under this assumption, the t-statistic is asymptotically Gaussian distributed at all horizons. Consequently, the increasing significance of the regression parameter at longer horizons is interpreted as evidence for long horizon predictability. On the other hand, looking upon the long horizon regression as a conditional error-correction model and interpreting the test for long horizon predictability as a single equation test for cointegration, rather than as a test for weak exogeneity, changes the relevant asymptotic distribution theory from the conventional Gaussian to mixtures of Gaussian and Dickey-Fuller type distributions. We have derived the asymptotic distribution of the estimator of the regression parameter and its t-statistic for arbitrary horizons, under the null hypothesis of no cointegration. The dependence of these distributions on the horizon k demonstrates that these distributions provide an alternative explanation for a least part of the typical findings of long horizon regressions. If cointegration does not hold, the observed increases are not surprising at all, but rather are predicted by the asymptotic distribution theory. Hence, the question raised in the title should be answered affirmatively. Subsequently, knowing its dependence on the horizon k, it proved possible to apply a Phillips-Perron type correction to the ordinary least-squares t-statistic in order to endow it with a stable size for given, arbitrary, horizon. For Newey-West type estimators, this does not appear possible.

The results from the empirical example concerning nominal exchange rates and a fundamental value derived from the monetary model are in line with the typical results in the literature. The standard statistics suggest that the nominal exchange rate *vis-à-vis* the US dollar of the Deutsche mark, the Japanese yen, the Canadian dollar, and the Swiss franc are all predictable over long horizons. In contrast, using the modified t-statistics reveals that there is hardly any evidence of long horizon predictability of nominal exchange rates.

References

Andrews, D.W.K., 1991, Heteroskedasticity and autocorrelation consistent covariance matrix estimation, *Econometrica* **60**, 953–966.

Berkowitz, J. and L. Giorgianni, 1997, Long-horizon exchange rate predictability?, Working paper No. 97/6, International Monetary Fund.

Bilson, J.F.O., 1976, Rational expectations and the exchange rate, in J.A. Frenkel and H.G. Johnson (editors), *The economics of exchange rates: Selected studies*, Addison-Wesley, Reading.

Bollerslev, T. and R.J. Hodrick, 1995, Financial market efficiency tests, in M.H. Pesaran and M. Wickens (editors), *Handbook of applied econometrics*, Blackwell, Cambridge, MA.

Boswijk, H.P., 1992, *Cointegration, Identification and Exogeneity. Inference in Structural Error Correction Models*, PhD thesis, Thesis/Tinbergen Institute, Amsterdam.

Boswijk, H.P., 1994, Testing for an unstable root in conditional and structural error correction models, *Journal of Econometrics* **63**, 37–60.

Boudoukh, J. and M. Richardson, 1993, Stock returns and inflation: a long-horizon perspective, *American Economic Review* **83**, 1346–1355.

Boudoukh, J. and M. Richardson, 1994, The statistics of long-horizon regressions revisited, *Mathematical Finance* **4**, 103–120.

Campbell, J.Y., A.W. Lo and A.C. MacKinlay, 1997, *The econometrics of financial markets*, Princeton University Press, New Jersey.

Campbell, J.Y. and R.J. Shiller, 1988, Stock prices, earnings, and expected dividends, *Journal of Finance* **43**, 661–676.

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.

Chinn, M.D. and R.A. Meese, 1995, Banking on currency forecasts: how predicatble is the change in money?, *Journal of International Economics* **38**, 161–178.

Cutler, D.M., J.M. Poterba and L.H. Summers, 1990, Speculative dynamics and the role of feedback traders, *American Economic Review* **80**, 63–68.

De Long, J.B., A. Shleifer, L.H. Summers and R.J. Waldmann, 1990, Noise trader risk in financial markets, *Journal of Political Economy* **98**, 703–738.

Fama, E.F., 1990, Term structure forecasts of interest rates, inflation, and real returns, *Journal of Monetary Economics* **25**, 59–76.

Fama, E.F. and K.R. French, 1988, Dividend yields and expected stock returns, *Journal of Financial Economics* **22**, 3–25.

Fama, E.F. and R.R. Bliss, 1987, The information in long-maturity forward rates, *American Economic Review* **77**, 680–692.

Frankel, J. and A. Rose, 1995, A survey of empirical research on nominal exchange rates, in G. Grossman and K. Rogoff (editors), *The Handbook of International Economics, Vol. 3*, Elsevier, North-Holland.

Frenkel, J.A., 1976, A monetary approach to the exchange rate: doctrinal aspects and empirical evidence, *Scandinavian Journal of Economics* **78**, 200–224.

Goetzmann, W.N. and Ph. Jorion, 1993, Testing the predictive power of dividend yields, *Journal of Finance* **48**, 663–679.

Hansen, B.E., 1992, Convergence to stochastic integrals for dependent heterogeneous processes, *Econometric Theory* **8**, 489–500.

Hansen, B.E., 1995, Rethinking the univariate approach to unit root testing, *Econometric Theory* **11**, 1148–1171.

Hansen, L.P. and R.J. Hodrick, 1980, Forward exchange rates as optimal predictors of future spot rates: an econometric analysis, *Journal of Political Economy* **88**, 829–853.

Herrndorf, N., 1984, A functional central limit theorem for weakly dependent sequences of random variables, *Annals of Probability* **12**, 141–153.

Hodrick, R.J., 1992, Dividend yields and expected stock returns: alternative procedures for inference and measurement, *Review of Financial Studies* **5**, 357–386.

Kilian, L., 1997, Exchange rates and fundamentals: what do we learn from long-horizon regressions?, Working paper, Department of Economics, University of Michigan.

Kremers, J.J.M., N.R. Ericsson and J.J. Dolado, 1992, The power of cointegration tests, *Oxford Bulletin of Economics and Statistics* **54**, 325–348.

Mark, N.C., 1995, Exchange rates and fundamentals: evidence on long-horizon predictability, *American Economic Review* **85**, 201–218.

Mark, N.C. and D.-Y. Choi, 1997, Real exchange rate prediction over long horizons, working paper, Department of Economics, Ohio State University.

Meese, R.A., 1990, Currency fluctuations in the post Bretton Woods era, *Journal of Economic Perspectives* **4**, 117–134.

Mishkin, F.S., 1990*a*, What does the term structure tell us about future inflation, *Journal of Monetary Economics* **25**, 77–95.

Mishkin, F.S., 1990*b*, The information in the longer maturity term structure about future inflation, *Quarterly Journal of Economics* **105**, 815–828.

Mussa, M., 1976, The exchange rate, the balance of payments, and monetary and fiscal policy under a regime of controlled floating, *Scandinavian Journal of Economics* **78**, 229–248.

Nelson, C.R. and M.J. Kim, 1993, Predictable stock returns: the role of small sample bias, *Journal of Finance* **48**, 641–661.

Newey, W.K. and K.D. West, 1987, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* **55**, 703–708.

Newey, W.K. and K.D. West, 1994, Automatic lag selection in covariance matrix estimation, *Review of Economic Studies* **61**, 631–653.

Phillips, P.C.B., 1986, Understanding spurious regressions, *Journal of Econometrics* **33**, 311–340.

Phillips, P.C.B., 1987, Time series regression with a unit root, *Econometrica* **55**, 277–301.

Phillips, P.C.B., 1988, Multiple regression with integrated processes, *Contemporary Mathematics* **80**, 79–108.

Phillips, P.C.B. and P. Perron, 1988, Testing for a unit root in time series regression, *Biometrika* **75**, 335–346.

Poterba, J.M. and L.H. Summers, 1988, Mean reversion in stock returns: evidence and implications, *Journal of Financial Economics* **22**, 27–60.

Richardson, M. and J.H. Stock, 1989, Drawing inferences from statistics based on multiyear asset returns, *Journal of Financial Economics* **25**, 323–348.

White, H., 1984, *Asymptotic Theory for Econometricians*, Academic Press, London.

Zivot, E., 1996, The power of single equation tests for cointegration when the cointegrating vector is prespecified, Working paper, Department of Economics, University of Washington.

Appendix: Mathematical proofs

In the proofs, we make use of various standard results on partial sum processes, which are summarized in Lemma 1.

Lemma 1 *Let z_t be generated according to (7) and assume that (28) holds. Then*

1. $T^{-1/2} z_{\lfloor Tr \rfloor} \Rightarrow \omega_v W_v^c(r)$;

2. $T^{-2} \sum_{t=k+1}^{T} z_{t-k}^2 \Rightarrow \omega_v^2 \int_0^1 (W_v^c)^2, \quad \forall\, k > 0$;,

3. $T^{-1} \sum_{t=k+1}^{T} z_{t-k}(e_t + \ldots + e_{t-k+1}) \Rightarrow k(\omega_v \omega_e(\rho \int_0^1 W_v^c(r) dW_v(r) + \sqrt{1-\rho^2} \int_0^1 W_v^c(r) dW_{e \cdot v}(r)) + \Lambda_{ve}) - k\Lambda_{ve,k-1}, \quad \forall\, k > 0.$

Proof:
Parts 1 and 2 are standard results, see, for example, Hansen (1992, 1995) for proofs. The proof of part 3, which resembles Lemma 1(e) in Phillips (1986), is given below.

We first invoke Theorem 3.49 of White (1984), saying that a measurable function of a finite string of a mixing variable that obeys a certain summability restriction (in our case Assumption 1) is also mixing (with different mixing coefficients) and its mixing coefficients obey the same summability restriction. Second, note that as $(v_t, e_t)'$ is uniformly square integrable, a finite sum of $(v_t, e_t)'$ will also be uniformly square integrable. Taken together, these two assertions imply that (the multivariate generalisation of) Herrndorf's Corollary 1 will also apply to the finite sums considered above. This means that partial sums of these finite sums will converge to Wiener processes. Having established the existence of the limit distribution, we rewrite the expression for general k in terms of the limit for $k = 1$. Observe that $T^{-1} \sum_{t=k+1}^{T} z_{t-k}(e_t + \ldots + e_{t-k+1}) = \sum_{j=0}^{k-1} T^{-1} \sum_{t=k+1}^{T} z_{t-k} e_{t-j}$. It then follows that

$$
\begin{aligned}
\sum_{j=0}^{k-1} T^{-1} \sum_{t=k+1}^{T} z_{t-k} e_{t-j} &= \sum_{j=0}^{k-1} T^{-1} \sum_{t=k+1}^{T} (z_{t-j-1} - \sum_{i=j+1}^{k-1} v_{t-i}) e_{t-j} \\
&= \sum_{j=0}^{k-1} (T^{-1} \sum_{t=k+1}^{T} z_{t-j-1} e_{t-j} - T^{-1} \sum_{t=k+1}^{T} (\sum_{i=j+1}^{k-1} v_{t-i}) e_{t-j}) \\
&= kT^{-1} \sum_{t=2}^{T} z_{t-1} e_t - kT^{-1} \sum_{t=k+1}^{T} \sum_{i=1}^{k-1} v_{t-i} e_t \\
&\Rightarrow k(\omega_v \omega_e \int_0^1 W_v(r) d(\rho W_v(r) + \sqrt{1-\rho^2} W_{e \cdot v}(r)) + \Lambda_{ve}) - k\Lambda_{ve,k-1},
\end{aligned}
$$

where the last line follows by the strong law of large numbers of McLeish (White (1984, Theorem 3.47)).

□

Proof of Theorem 1:
Using the results of Hansen (1992,1995), it is easy to show that

$$
\begin{aligned}
T(\hat{\beta} - \beta) &= \frac{T^{-1} \sum_{t=1}^{T-1} z_t e_{t+1}}{T^{-2} \sum_{t=1}^{T-1} z_t^2} \\
&\Rightarrow R\rho \frac{\int_0^1 W_v dW_v}{\int_0^1 W_v^2} + R\sqrt{1-\rho^2} \frac{W_v dW_{e \cdot v}}{\int_0^1 W_v^2} + \frac{1}{\omega_v^2} \frac{\Lambda_{ve}}{\int_0^1 W_v^2}.
\end{aligned}
$$

The fact that Δx_t drops out can be easily seen from the following stylized example. Suppose $y = x\beta + z + u$, then $\hat{\beta} = (x'x)^{-1}x'(y-z) = (x'x)^{-1}x'(x\beta + z + u - z) = \beta + (x'x)^{-1}u$.

The limiting distribution of the least-squares based t-statistic can be derived as follows

$$\begin{aligned} t_{LS}(\hat{\beta}) &= \frac{\hat{\beta}}{\hat{\sigma}_e}\left(\sum_{t=1}^{T-1} z_t^2\right)^{1/2} \\ &= \frac{T\hat{\beta}}{\hat{\sigma}_e}\left(T^{-2}\sum_{t=1}^{T-1} z_t^2\right)^{1/2} \\ &\Rightarrow \frac{\omega_e}{\sigma_e}\left(\rho\frac{\int_0^1 W_v dW_v}{(\int_0^1 W_v^2)^{1/2}} + \sqrt{1-\rho^2}\frac{\int_0^1 W_v dW_{e\cdot v}}{(\int_0^1 W_v^2)^{1/2}}\right) + \frac{1}{\omega_v\sigma_e}\frac{\Lambda_{ve}}{(\int_0^1 W_v^2)^{1/2}}. \end{aligned}$$

If a Newey-West type estimator is employed it follows that

$$t_{NW}(\hat{\beta}) \Rightarrow \left(\rho\frac{\int_0^1 W_v dW_v}{(\int_0^1 W_v^2)^{1/2}} + \sqrt{1-\rho^2}\frac{\int_0^1 W_v dW_{e\cdot v}}{(\int_0^1 W_v^2)^{1/2}}\right) + \frac{1}{\omega_v\omega_e}\frac{\Lambda_{ve}}{(\int_0^1 W_v^2)^{1/2}}.$$

□

Proof of Theorem 2:

$$\begin{aligned} T(\hat{\beta}_k - \beta_k) &= \frac{T^{-1}\sum_{t=1}^{T-k} z_t(e_{t+k} + \ldots + e_{t+1})}{T^{-2}\sum_{t=1}^{T-k} z_t^2} \\ &\Rightarrow kR\rho\frac{\int_0^1 W_v dW_v}{\int_0^1 W_v^2} + kR\sqrt{1-\rho^2}\frac{\int_0^1 W_v dW_{e\cdot v}}{\int_0^1 W_v^2} + k\frac{1}{\omega_v^2}\left(\frac{\Lambda_{ve}}{\int_0^1 W_v^2} - \frac{\Lambda_{ve,k-1}}{\int_0^1 W_v^2}\right), \end{aligned}$$

where the second line follows from Lemma 1.

Under the null hypothesis $\beta = 0$, the least squares estimator $T^{-1}\sum(\hat{e}_{t+k} + \ldots \hat{e}_{t-1})^2$ of the residual variance converges in probability to $k\sigma_e^2 + 2\sum_{j=1}^{k-1}\Lambda_{ee,j}$, because

$$\begin{aligned} &T^{-1}\sum_{t=1}^{T-k}(\hat{e}_{t+k} + \ldots \hat{e}_{t+1})^2 = T^{-1}\sum_{t=1}^{T-k}(\hat{e}_{t+k}^2 + \ldots \hat{e}_{t+1}^2 + 2\sum_{j=0}^{k-2}\sum_{i=j+1}^{k-1}\hat{e}_{t+k-j}\hat{e}_{t+k-i}) \\ &\to k\sigma_e^2 + 2\sum_{j=1}^{k-1}\Lambda_{ee,j}, \end{aligned}$$

where the last line follows from the strong law of large numbers of McLeish. Therefore, for the least-squares based t-statistic it follows that

$$\begin{aligned} t_{LS}(\hat{\beta}_k) &= \frac{\hat{\beta}_k}{T^{-1}\sum_{t=1}^{T-k}(\hat{e}_{t+k} + \ldots \hat{e}_{t+1})^2}\left(\sum_{t=1}^{T-k} z_t^2\right)^{1/2} \\ &= \frac{T\hat{\beta}_k}{T^{-1}\sum_{t=1}^{T-k}(\hat{e}_{t+k} + \ldots \hat{e}_{t+1})^2}\left(T^{-2}\sum_{t=1}^{T-k} z_t^2\right)^{1/2} \\ &\Rightarrow \frac{k\omega_e}{\sqrt{k\sigma_e^2 + 2\sum_{j=1}^{k-1}\Lambda_{ee,j}}}\left(\rho\frac{\int_0^1 W_v dW_v}{(\int_0^1 W_v^2)^{1/2}} + \sqrt{1-\rho^2}\frac{\int_0^1 W_v dW_{e\cdot v}}{(\int_0^1 W_v^2)^{1/2}}\right) + \\ &\quad \frac{k}{\sqrt{k\sigma_e^2 + 2\sum_{j=1}^{k-1}\Lambda_{ee,j}}}\frac{1}{\omega_v}\left(\frac{\Lambda_{ve}}{(\int_0^1 W_v^2)^{1/2}} - \frac{\Lambda_{ve,k-1}}{(\int_0^1 W_v^2)^{1/2}}\right). \end{aligned}$$

Under the null hypothesis $\beta = 0$, the Newey-West estimator of the long-run variance of the error-terms $e_{t+k,k} = e_{t+k} + \ldots + e_{t+1}$ converges in probability of $k^2\omega_e^2$. Therefore

$$\begin{aligned} t_{NW}(\hat{\beta}_k) &= \frac{\hat{\beta}_k}{k\hat{\omega}_e}\left(\sum_{t=1}^{T-k} z_t^2\right)^{1/2} \\ &= \frac{T\hat{\beta}_k}{k\hat{\omega}_e}\left(T^{-2}\sum_{t=1}^{T-k} z_t^2\right)^{1/2} \\ &\Rightarrow \rho\frac{\int_0^1 W_v dW_v}{(\int_0^1 W_v^2)^{1/2}} + \sqrt{1-\rho^2}\frac{\int_0^1 W_v dW_{e\cdot v}}{(\int_0^1 W_v^2)^{1/2}} + \frac{1}{\omega_v\omega_e}\left(\frac{\Lambda_{ve}}{(\int_0^1 W_v^2)^{1/2}} - \frac{\Lambda_{ve,k-1}}{(\int_0^1 W_v^2)^{1/2}}\right). \end{aligned}$$

□

Proof of Theorem 3:
Under the local alternative $\beta = -c/T$ we can exploit the autoregressive structure of the z_t process

$$z_t = (1-\beta)^k z_{t-k} + \sum_{j=0}^{k-1}(1-\beta)^j(e_{t-j} - \Delta x_{t-j}).$$

Therefore

$$z_t - z_{t-k} = ((1-\beta)^k - 1)z_{t-k} + \sum_{j=0}^{k-1}(1-\beta)^j(e_{t-j} - \Delta x_{t-j}).$$

Using the tools of Phillips (1987) and Hansen (1992,1995), it is straightforward to show that under (32):

$$\begin{aligned} T(\hat{\beta}_k - 0) &= -kc + \frac{T^{-1}\sum_{t=1}^{T-k} z_t \sum_{j=0}^{k-1}(1-\beta)^j e_{t+k-j}}{T^{-2}\sum z_t^2} \\ &\Rightarrow -kc + kR\rho\frac{\int_0^1 W_1^c dW_v}{\int_0^1 (W_1^c)^2} + kR\sqrt{1-\rho^2}\frac{\int_0^1 W_1^c dW_{e\cdot v}}{\int_0^1 (W_1^c)^2} + k\frac{1}{\omega_v^2}\left(\frac{\Lambda_{ve}}{\int_0^1 (W_1^c)^2} - \frac{\Lambda_{ve,k-1}}{\int_0^1 (W_1^c)^2}\right), \end{aligned}$$

where the second line follows from applying Newton's binomium to $(1-\beta)^j$ and using $\beta = -c/T$. The asymptotic distributions of the least-squares based and Newey-West type covariance matrix estimators are independent of the local alternative, cf. Phillips (1988). Thus,

$$\begin{aligned} &\frac{\sqrt{k\sigma_e^2 + 2\sum_{j=1}^{k-1}\Lambda_{ee,j}}}{k\omega_e} t_{LS}(\hat{\beta}_k) - \frac{\hat{\Lambda}_{ve} - \hat{\Lambda}_{ve,k-1}}{\hat{\omega}_e(T^{-2}\sum_{t=1}^{T-k}(y_t - x_t)^2)^{1/2}} \Rightarrow \\ &\qquad \frac{c}{\omega_e\omega_v(\int_0^1 (W_v^c)^2))^{1/2}} + \rho\frac{\int_0^1 W_v dW_v}{(\int_0^1 W_v^2)^{1/2}} + \sqrt{1-\rho^2}\mathcal{N}(0,1), \\ &t_{NW}(\hat{\beta}_k) + \frac{\hat{\Lambda}_{ve} - \hat{\Lambda}_{ve,k-1}}{\hat{\omega}_e(T^{-2}\sum_{t=1}^{T-k}(y_t - x_t)^2)^{1/2}} \Rightarrow \\ &\qquad \frac{c}{\omega_e\omega_v(\int_0^1 (W_v^c)^2))^{1/2}} + \rho\frac{\int_0^1 W_v dW_v}{(\int_0^1 W_v^2)^{1/2}} + \sqrt{1-\rho^2}\mathcal{N}(0,1). \end{aligned}$$

We conclude that the power is independent of k.

□

Table 1: Long horizon regression estimates for nominal exchange rates[1]

k	$\hat{\beta}_k$	$t_{LS}(\hat{\beta}_k)$	$t_{20}(\hat{\beta}_k)$	$t_A(\hat{\beta}_k)$	R_k^2
Canadian dollar					
1	−0.019	−1.499	−1.434	−1.499	0.033
4	−0.097	−3.696	−1.884	−1.742	0.139
8	−0.204	−5.095	−2.128	−2.299	0.237
12	−0.324	−6.522	−2.415	−2.835	0.344
16	−0.434	−7.400	−2.631	−3.182	0.414
Deutsche mark					
1	−0.044	−1.863	−1.778	−1.786	0.045
4	−0.165	−3.324	−1.764	−1.706	0.117
8	−0.354	−4.655	−1.993	−2.056	0.207
12	−0.570	−6.367	−2.312	−2.350	0.334
16	−0.773	−8.335	−2.720	−2.624	0.471
Japanese yen					
1	−0.044	−1.740	−1.805	−1.628	0.041
4	−0.194	−3.478	−2.049	−1.860	0.126
8	−0.399	−4.963	−2.312	−2.355	0.228
12	−0.577	−5.953	−2.396	−2.401	0.305
16	−0.709	−6.798	−2.434	−2.369	0.374
Swiss franc					
1	−0.080	−2.329	−2.566	−2.255	0.067
4	−0.292	−4.470	−2.750	−2.434	0.196
8	−0.574	−6.715	−3.575	−3.575	0.360
12	−0.829	−9.697	−5.150	−5.039	0.549
16	−1.052	−15.640	−8.593	−8.274	0.768

[1] The table presents least squares estimates of the long horizon regression $\Delta_k s_{t+k} = \beta_k(s_t - f_t) + e_{t+k,k}$, where s_t is the log nominal exchange rate (US dollars per unit of foreign currency) and $f_t = (m_t - m_t^*) - \lambda(y_t - y_t^*)$. The data are quarterly, covering 1973Q1-1997Q3 for the Canadian dollar, Deutsche mark and Japanese yen and 1973Q1-1996Q2 for the Swiss franc (which render effective sample sizes of 94 and 89 observations for $k = 1$, respectively). The t-statistic $t_{LS}(\hat{\beta}_k)$ is computed using the least squares estimator of the residual variance. The t-statistics $t_{20}(\hat{\beta}_k)$ and $t_A(\hat{\beta}_k)$ are computed using the Newey-West estimator of the residual variance given in (6) with truncation lag $l = 20$ and l set according to the rule of Andrews (1991), respectively.

Table 2: Simulation design[1]

design	α_{11}	α_{22}	$\sigma_{\varepsilon\eta}$	ρ
1	0.0	0.0	0.0	0.71
2	0.0	0.5	0.0	0.48
3	0.5	0.0	0.0	0.89
4	0.0	0.0	0.5	0.33
5	0.5	0.5	0.0	0.71
6	0.5	0.5	0.5	0.36

[1] The table presents the design of the various data generating processes used to investigate the size of the modified t-statistics introduced in Corollary 3 in small samples. The precise meaning of the parameters α_{11}, α_{22}, $\sigma_{\varepsilon\eta}$ is given in (26)-(27), while ρ is defined in (12).

Table 3: Rejection frequencies of modified t-statistics[1]

		$t_{LS}(\hat{\beta}_k)$			$t_{20}(\hat{\beta}_k)$			$t_A(\hat{\beta}_k)$		
design	k	1%	5%	10%	1%	5%	10%	1%	5%	10%
1	1	3	7	11	5	10	14	1	5	8
	4	1	3	6	7	13	18	5	11	15
	8	1	5	8	10	17	23	10	17	22
	12	2	6	9	14	22	27	15	23	28
	16	3	7	10	18	26	32	20	28	34
2	1	5	10	14	4	9	13	1	3	7
	4	1	4	7	6	12	17	4	9	13
	8	2	5	8	8	15	19	8	14	19
	12	2	5	9	11	18	23	12	19	24
	16	2	6	9	14	21	26	16	23	28
3	1	0	1	1	1	3	5	0	1	2
	4	0	0	1	5	9	13	5	10	13
	8	0	1	2	10	17	22	15	22	29
	12	0	1	3	15	24	30	24	34	40
	16	1	2	3	22	31	37	33	42	48
4	1	1	4	6	4	8	13	1	3	6
	4	1	3	6	5	10	15	3	8	12
	8	1	3	7	7	13	18	6	12	17
	12	1	4	7	9	16	20	9	16	20
	16	1	4	7	11	18	22	12	18	23
5	1	1	3	4	2	4	6	1	2	3
	4	0	1	2	5	10	14	5	10	14
	8	0	1	2	10	16	20	13	21	26
	12	1	2	3	14	21	25	21	29	35
	16	1	2	3	18	25	30	27	36	41
6	1	2	4	6	3	6	8	1	3	6
	4	0	1	3	5	9	12	5	9	12
	8	0	1	2	8	12	16	10	16	20
	12	0	1	3	10	15	19	15	22	26
	16	0	1	3	12	18	22	18	25	30

[1] The table presents rejection frequencies of the modified t-statistics to test the null hypothesis $H_0 : \beta_k = 0$ in (7), as discussed in Corollary 3. The series (y_t, x_t) are generated as possbily correlated random walks, with the innovations generated according to (26)-(27). The specific values for the parameters α_{11}, α_{22} and $\sigma_{\varepsilon\eta}$ in the different designs are given in Table 2. The table is based on 5000 replications for sample size $T = 100$. The t-statistic $t_{LS}(\hat{\beta}_k)$ is computed using the least squares estimator of the residual variance. The t-statistics $t_{20}(\hat{\beta}_k)$ and $t_A(\hat{\beta}_k)$ are computed using the Newey-West estimator of the residual variance given in (6), with the truncation lag $l = 20$ and l set according to the rule of Andrews (1991), respectively.

Table 4: Modified long horizon t-statistics for nominal exchange rates[1]

k	$t_{LS}(\hat{\beta}_k)$	$t_{20}(\hat{\beta}_k)$	$t_A(\hat{\beta}_k)$
Canadian dollar			
1	−2.967	−1.505	−1.679
4	−1.664	−1.656	−1.744
8	−1.599	−1.892	−2.247
12	−1.682	−2.294	−2.796
16	−1.662	−2.570	−3.144
Deutsche mark			
1	−1.809	−1.879	−1.885
4	−1.392	−1.807	−1.703
8	−1.470	−1.922	−2.053
12	−1.677	−2.250	−2.346
16	−1.921	−2.686	−2.626
Japanese yen			
1	−1.648	−1.599	−1.648
4	−1.290	−1.792	−1.857
8	−1.465	−2.145	−2.352
12	−1.491	−2.182	−2.399
16	−1.500	−2.371	−2.367
Swiss franc			
1	−2.696	−2.779	−2.337
4	−2.038	−2.863	−2.430
8	−2.248	−3.481	−3.568
12	−2.688	−5.069	−5.030
16	−3.782	−8.528	−8.258

[1] The table presents results on tests for long horizon predictability using the modified t-statistics, applying the formulae in Corollary 3. The t-statistic $t_{LS}(\hat{\beta}_k)$ is computed using the least squares estimator of the residual variance. The t-statistics $t_{20}(\hat{\beta}_k)$ and $t_A(\hat{\beta}_k)$ are computed using the Newey-West estimator of the residual variance given in (6), with the truncation lag $l = 20$ and l set according to the rule of Andrews (1991), respectively. The appropriate asymptotic critical values, based on Table 1 in Hansen (1995), are (1%, 5%, 10%): Deutsche mark, -3.41, -2.82, -2.53 ($\hat{\rho} = 0.95$), Canadian dollar, -3.28, -2.70, -2.39 ($\hat{\rho} = 0.68$), Japanese yen and Swiss franc, -3.39, -2.81, -2.50 ($\hat{\rho} = 0.90$).