Memory Properties of the Forward Premium: A Study on South African Exchange Rates

H Boraine and **PJ van Staden**

Department of Statistics, University of Pretoria

ABSTRACT

The forward rate unbiasedness hypothesis states that the current forward rate should be an unbiased forecaster of the future spot rate. Inference has always been done under the assumption that the forward premium is a stationary short memory series. Recent empirical results have indicated that this assumption is not valid. Standard unit root tests performed on the forward premium often indicate infinite long memory. However, in recent literature fractionally integrated models have been applied for the forward premium. Empirical analysis is usually performed on exchange rates of developed economies. In this article, the South African Rand-Dollar exchange rate is considered and the focus is therefore on a developing country. A bootstrap method for determining standard errors and confidence limits is described and implemented.

JEL C10, F31, G14

INTRODUCTION

The forward rate unbiasedness hypothesis is a logical starting point for the analysis of spot and forward exchange rates. The hypothesis states that the current forward rate should be an unbiased forecaster of the future spot rate. In terms of a regression model this is expressed as:

$$s_{t+k} = \alpha + \beta f_{t,k} + \varepsilon_{t+k} \tag{1}$$

where s_{t+k} is the logarithm of the spot rate at time $t+k$, $f_{t,k}$ is the logarithm of the forward rate at time t and k is the length of the forward contract (for example, if monthly data are used and a one-month forward contract is considered, the value of k will be one).

Following Fama (1984) a regression model in returns is often used instead of (1):

$$s_{t+k} - s_t = \alpha^* + \beta^* \left(f_{t,k} - s_t \right) + \varepsilon_{t+k}^* \tag{2}$$

where $s_{t+k} - s_t$ is the spot return and $f_{t,k} - s_t$ is the forward premium. The null hypothesis of unbiasedness in this regression is that $\alpha^* = 0$ and $\beta^* = 1$. Inference in (2) has always been done under the assumption that the forward premium is a stationary short memory series. In effect, it is assumed that the forward premium is integrated to order zero, denoted $\left(f_{t,k} - s_t \right) \sim I(0)$.

Recent empirical results have indicated that this assumption is not valid. For instance, using unit root tests, Crowder (1994) fails to reject the null hypothesis of non-stationarity in several forward premia series. Furthermore, he rejects the null hypothesis of stationarity for these series by using the KPSS test of Kwiatkowski *et al.* (1992). Crowder's empirical results therefore indicate that the forward premium has a unit root, so that $\left(f_{t,k} - s_t \right) \sim I(1)$.

As pointed out by Baillie and Bollerslev (1994), the problem with unit root tests as well as the KPSS test is that these tests force a choice between an $I(0)$ and an $I(1)$ process. They consider fractionally integrated processes that are $I(d)$ with $0 < d < 1$ and find evidence of fractional integration for the forward premia series.

Most empirical analysis is performed on exchange rates of developed economies. We consider the South African Rand-Dollar exchange rate and our focus is therefore on a developing country. In this article basic results regarding fractional integration and long memory are given. Different methods for the estimation of d are cited.

A bootstrap method for determining standard errors and confidence limits is described and implemented. The importance of standard errors and confidence limits in the interpretation of results cannot be over emphasized. If, for instance, a 95% confidence interval for d includes both the values 0 and 1, it is impossible to draw any conclusions regarding the memory properties of such a series. Even though it is almost always possible to estimate d, some time series simply do not contain enough information to estimate d accurately. It is standard practice to provide standard errors with the estimates. Confidence limits are typically calculated under the assumption of a symmetric bell-shaped distribution. This assumption is not necessarily valid. The advantage of the bootstrap is that it does not rely on any distributional form. It is used to estimate the probability distribution of an estimator, together with confidence limits.

A MODEL FOR TIME SERIES WITH LONG MEMORY

Consider the simple autoregressive model for a discrete time series x_t

$$x_t = \phi x_{t-1} + \varepsilon_t \tag{3}$$

where the ε_t's are white noise innovations with mean 0 and variance σ^2. If $\phi = 1$, x_t is integrated to order one, denoted $I(1)$, and has infinite long memory. That is, the effect of the innovations ε_t is permanent. In contrast, if $\phi < 1$, then x_t is $I(0)$ and has short memory so that the effect of the innovations disappear geometrically. The problem with the autoregressive model is that a choice must be made between these two extremes.

ARMA models are not designed to describe time series with long memory, where the effect of innovations is not permanent, but still may take relatively long to die out. A more suitable model for time series with long memory is the fractionally integrated model, which was developed independently by Granger (1980), Granger and Joyeux (1980) and Hosking (1981) and is given by

$$(1-L)^d (x_t - \mu) = u_t \tag{4}$$

where μ is the mean of x_t, u_t is a covariance stationary series, L is the backshift operator and d is the so-called fractional differencing parameter. For $d = 0$, x_t has short memory (corresponding to $\phi < 1$ in the autoregressive model), while for $d = 1$, x_t has infinite long memory (corresponding to $\phi = 1$). For $0 < d < 1$, x_t is said to be fractionally integrated and has long memory with the effect of the innovations disappearing at a slow hyperbolic rate. In particular, x_t is stationary and mean reverting for $0 < d < \frac{1}{2}$, while for $\frac{1}{2} < d < 1$, x_t is non-stationary but still mean reverting (a series is mean reverting for $d < 1$).

If u_t is a white noise series, (4) is referred to as a fractional white noise model. By substituting $u_t = \phi(L)^{-1} \theta(L) \varepsilon_t$ into (4), an autoregressive fractionally integrated moving average or *ARFIMA*(p,d,q) model is obtained:

$$\phi(L)(1-L)^d (x_t - \mu) = \theta(L)\varepsilon_t \tag{5}$$

where $\phi(L) = 1 - \phi_1 L - \ldots - \phi_p L^p$ and $\theta(L) = 1 - \theta_1 L - \ldots - \theta_q L^q$.

Using a Taylor series expansion, it can be shown that the fractional difference operator can be written as

$$(1-L)^d = \sum_{k=0}^{\infty} \frac{\Gamma(k-d)L^k}{\Gamma(-d)\Gamma(k+1)}$$

$$= \sum_{k=0}^{\infty} \pi_k L^k \tag{6}$$

with $\pi_0 = 1$, $\pi_1 = -d$ and $\pi_k = \frac{k-d-1}{k}\pi_{k-1}$.

See Baillie (1996) for a summary of long memory processes and fractional integration as well as the other important properties of the models described above.

ESTIMATION METHODS

A comprehensive discussion of estimation methods for fractionally integrated models is given by Chung and Baillie (1993). These methods include a semi-parametric estimator in the frequency domain by Geweke and Porter-Hudak (1983), a minimum conditional sum of squares (CSS) estimator, originally suggested by Hosking (1984), an approximate maximum likelihood procedure in the frequency domain by Fox and Taqqu (1986) and a full maximum likelihood estimator under the assumption of normally distributed innovations by Sowell (1992).

For our application the CSS method is used. The conditional sum of squares function is given by:

$$S(\lambda) = \frac{T}{2}\log\sigma^2 + \frac{1}{2\sigma^2}\sum_{t=1}^{T}\varepsilon_t^2 \tag{7}$$

where ε_t is obtained from (5) and $\lambda = (d,\phi_1,...,\phi_p,\theta_1,...,\theta_q,\sigma^2)$ is the vector of parameters that must be estimated. An iterative optimization algorithm is employed to minimize (7). CSS is an approximate maximum likelihood method with the advantage that it is not as computationally intensive as Sowell's full maximum likelihood. It is not based on any distributional assumptions and can be extended to estimate models with *GARCH* components.

STANDARD ERRORS AND CONFIDENCE LIMITS

The bootstrap is a method that is used to determine the standard error of an estimator and to construct confidence limits for the true parameter value – see Swanepoel (1990) for a review. In order to assess the precision of an estimate, it is of vital importance to calculate standard errors and confidence limits. If the estimate for d has a large standard error, it is not stable and the confidence interval will consequently be wide and may include a range of values that imply stationarity ($d < \frac{1}{2}$), non-stationarity and mean reversion ($\frac{1}{2} < d < 1$), and non-stationarity without mean reversion $d \geq 1$. On the other hand, if the standard error of the estimate of d is small, the confidence interval will be narrow and it will be possible to draw accurate conclusions regarding the order of differencing from the estimate of d.

The bootstrap is often applied in situations where no closed form expression for the standard error of an estimator exists, or when it is difficult to derive such an expression. In the case of *ARFIMA* models the parameters are estimated by minimizing a function that is non-linear in the parameters. Conditional sum of squares is used in this study to estimate the parameters of *ARFIMA* models. Closed form expressions for the estimators are not available, let alone expressions for their standard errors. An advantage of the bootstrap is that it does not rely on any distributional assumptions. Although it is a computationally intensive method, that is not regarded to be a serious drawback in this age of ever increasing computer power.

The bootstrap was originally proposed for cross sectional or non-time series data by Efron (1979). The idea is to draw random samples (with replacement) from the observed data. These samples are called bootstrap samples and they are regarded as independent samples from the population. The statistic is calculated for every sample. The value of the statistic will vary from sample to sample and the distribution of values is an estimate of the sampling distribution of the statistic. The standard deviation of the estimated distribution can in many cases be shown to be a consistent estimator for the standard error of the statistic under consideration. The basic assumption is that the observed data set is representative of the population.

In the case of time series data special care has to be taken to preserve the dependence structure of the time series. One of the methods applicable to dependent data is to resample from the residuals of the fitted model – see Chatterjee (1986). In a study on autoregressive processes, De Koster (1999) found that this method performs better than other methods for dependent data.

In our application, an *ARFIMA* model is fitted to the observed time series. The residuals of this fit are assumed to be representative of the error terms of the *ARFIMA* model. A large number of time series are generated using the fitted model as well as the residual terms. These time series are called bootstrap time series. Residuals are drawn randomly with replacement from the set of residual terms. For every time series an *ARFIMA* model of the same order as the original model is fitted. The estimates will vary from one series to the next. The standard deviations of the parameter estimates are then used as estimators for the standard errors of the estimates. Care should be taken to fit an appropriate model in the first place. The model may not fit well for several reasons. If the error terms have a relatively large variance, the bootstrap time series will also have large error variances and the estimates will vary relatively more across the different time series, resulting in relatively large standard errors and wide confidence intervals. In other words, if the model does not fit well in the first place, it will also be reflected by the standard errors of the parameter estimates and by the length of the confidence intervals.

In order to calculate confidence limits, it is recommended to generate at least 1000 bootstrap time series. We use the method of percentiles to calculate confidence limits. The 1000 estimates for a particular parameter are ordered and the 2.5[th] and 97.5[th] percentiles are used as a 95% confidence interval – see for example Swanepoel (1990).

APPLICATION TO SOUTH AFRICAN EXCHANGE RATES

We study the daily Rand-Dollar spot and thirty-day forward exchange rates for the period from 3 January 1995 to 5 March 2001, a time series of 1524 observations. The forward premium is calculated as the difference between the (logarithm of the) thirty-day forward rate and (the logarithm of the) spot rate (Figure 1).

Table 1 presents the results of unit root (Phillips-Perron) tests and the KPSS test for determining the order of integration of the forward premium. First consider the KPSS test, which tests $H_0 : I(0)$ against $H_1 : I(1)$. The null hypothesis is rejected at a 1% significance level so that it is clear that the forward premium is not a stationary short memory series. $H_0 : I(1)$ is tested against $H_1 : I(0)$ with the Phillips-Perron test. Since the null cannot be rejected, the indication is that the forward premium is a non-stationary infinite long memory series.

Figure 1   The forward premium: rand-dollar exchange rate
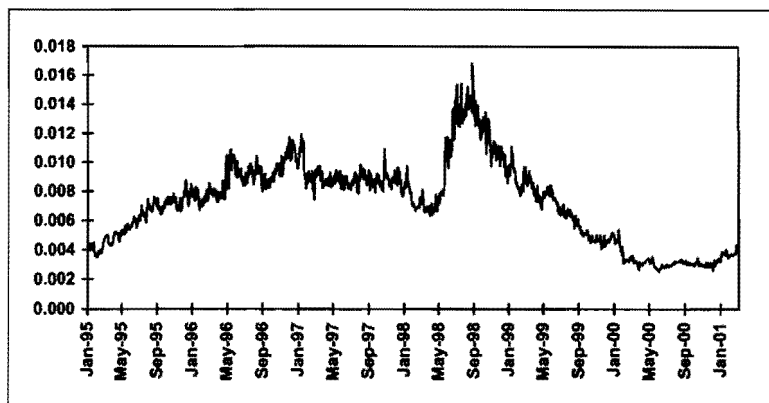
However, it is known that unit root tests have very low power against fractionally integrated alternatives – see Diebold and Rudebusch (1991) for the Dickey-Fuller test and Hassler and Wolters (1994) for the Phillips-Perron test. In contrast, Lee and Schmidt (1993) note that the KPSS test is quite powerful against fractionally integrated alternatives. Given this, the forward premium may be a long memory series that is fractionally integrated.

Table 1   Testing the order of integration of the forward premium

	$H_0 : I(0)$		Phillips-Perron test	$H_0 : I(1)$	
KPSS test	η_α	η_δ		$Z(t_\alpha)$	$Z(t_\delta)$
	5.44[a]	3.32[a]		-1.98	-2.48

Notes
1 η_α and η_δ are the KPSS test statistics based on residuals from regressions with only an intercept and an intercept and time trend respectively – see Kwiatkowski *et al.* (1992). The 0.01 and 0.05 critical values for η_α are 0.74 and 0.46 respectively. Similarly the 0.01 and 0.05 critical values for η_δ are 0.22 and 0.15.

2 $Z(t_\alpha)$ and $Z(t_\delta)$ are the Phillips-Perron adjusted test statistics of the lagged dependent variable in regressions with only an intercept and an intercept and time trend respectively – see Phillips (1987) and Phillips and Perron (1988). The 0.01 and 0.05 critical values for $Z(t_\alpha)$ are -3.44 and -2.86 respectively. Similarly the 0.01 and 0.05 critical values for $Z(t_\delta)$ are -3.97 and -3.42.

3   Results from other unit root tests (Dickey-Fuller) were similar to those
    from the Phillips-Perron test and are hence not given.
4   For both the Phillips-Perron tests and the KPSS tests the number of
    lags were truncated at lag 7 following Newey and West (1987).
5   [a] Significant at 1% level.
    [b] Significant at 5% level.

The correlogram of the forward premium gives more evidence of possible
fractional integration. The first 200 autocorrelations of the spot rate, the forward
premium and the spot return are shown in Figure 2. The spot rate is a random
walk and hence has infinite long memory. The autocorrelations of the spot rate
will therefore decrease at a very slow rate. Since the spot rate is $I(1)$, the spot
return is $I(0)$ and uncorrelated at higher lags. The spot return therefore has short
memory. The autocorrelations of the forward premium are persistent, but
considerably less so than those of the spot rate. This is consistent with long
memory behavior (the autocorrelations of a long memory series decrease at a
slow hyperbolic rate).

Figure 2   First 200 autocorrelations
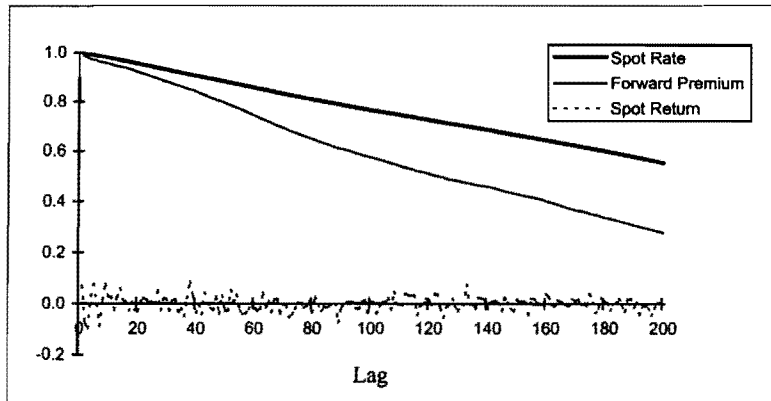
Given the above discussion, the fractional differencing parameter d of the
forward premium is estimated. CSS is used to estimate low-order
$ARFIMA(p,d,q)$ models for the forward premium with $p \leq 2$ and $q \leq 2$. Since
none of the estimated autoregressive or moving average parameters is
significant, they are eliminated from the final model, giving a fractional white
noise model. Thus, only d and σ^2 are estimated.

The empirical estimation results are given in Table 2. The estimate of d is approximately 0.8 so that the forward premium is $I(0.8)$ and has long memory. A total of 1000 bootstrap time series are used to determine the sampling distribution of d from which the standard error and confidence limits are obtained. Figure 3 shows the histogram of this distribution. The distribution is nearly symmetrical (slightly positively skewed) but not normal. Because the lower boundary of the 95% confidence interval is larger than $\frac{1}{2}$, the forward premium is non-stationary. However, the forward premium is mean reverting since the upper boundary of the interval is less than 1.

Table 2   Empirical estimation results

| Observations used	Parameter	Estimate	Standard error	95% Confidence interval
All 1524	d	0.799	0.022	(0.757, 0.845)
	σ^2	0.17 E-06	0.17 E-07	(0.14 E-06, 0.21 E-06)
First 800	d	0.792	0.030	(0.736, 0.850)
	σ^2	0.16 E-06	0.24 E-07	(0.13 E-06, 0.22 E-06)
Last 724	d	0.816	0.031	(0.757, 0.878)
	σ^2	0.18 E-06	0.29 E-07	(0.14 E-06, 0.25 E-06)

Figure 3   Sampling distribution of d
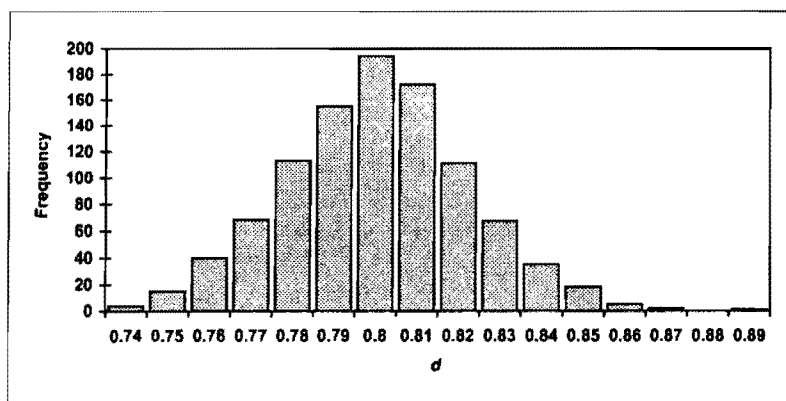
There was a large depreciation in the Rand against the Dollar in June 1998 followed by a period of high volatility. This is illustrated in Figure 4. Because of the depreciation, there was a sharp increase in the forward premium in June 1998 as can be seen in Figure 1. The question is what the effect of this shock on the forward premium series will be. Theoretically, since the estimate of d is less

than 1, the effect should not be permanent. That is, the forward premium should revert back to the mean. To analyze this problem further, and to confirm our estimation results for d, we divided the time series in two subsets. We took the first 800 observations from 3 January 1995 to 31 March 1998 and the last 724 observations from 1 April 1998 to 5 March 2001. For both these subsets we again estimated d. As can be seen in Table 2, the estimates as well as the confidence intervals correspond well with those of the complete dataset. The forward premium is therefore indeed mean reverting and we conclude that the forward premium is $I(0.8)$.

Figure 4   The daily rand-dollar spot and 30-day forward rates in 1998
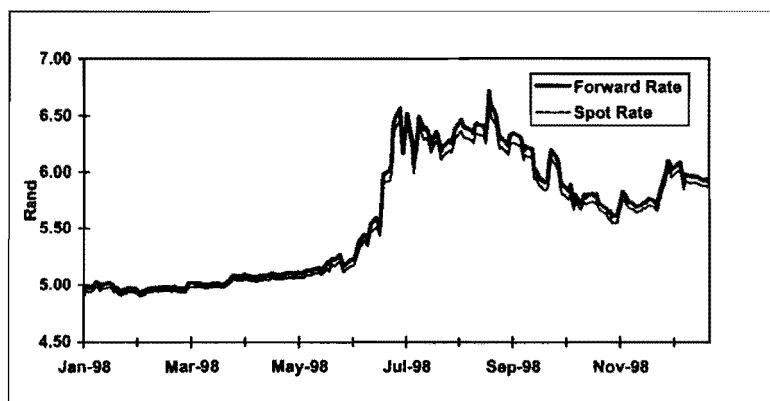
CONCLUSION

Recent empirical estimation results indicate that the widely accepted assumption of a short memory forward premium is invalid (for example Crowder, 1994). The evidence from unit root tests is that the forward premium has infinite long memory. However, Baillie and Bollerslev (1994) consider fractionally integrated models and conclude that the forward premium has long memory.

All the above studies concentrate on exchange rates from developed economies. We consider the Rand-Dollar exchange rate and hence focus on a developing economy (South Africa). Using CSS estimation we also find evidence of long memory for the South African data. Therefore, although the forward premium is non-stationary and has infinite variance, it is still mean reverting.

The long memory of the forward premium has implications for the analysis and modeling of exchange rates and is also one of the causes of the forward

premium anomaly. This anomaly refers to the fact that empirical estimates of β^* in the regression model in returns (2) are often negative. The null hypothesis of unbiasedness is then rejected (since the empirical estimate of β^* is less than unity). As explained in detail by Baillie and Bollerslev (2000), the rejection of unbiasedness is due to the persistent autocorrelations of the long memory forward premium. In fact, the long memory of the forward premium is in itself a rejection of unbiasedness.

REFERENCES

1   BAILLIE, R.T. (1996) "Long Memory Processes and Fractional Integration in Econometrics", *Journal of Econometrics*: 73, 5-59.

2   BAILLIE, R.T. & BOLLERSLEV, T. (1994) "The Long Memory of the Forward Premium", *Journal of International Money and Finance*, 13: 565-71.

3   BAILLIE, R.T. & BOLLERSLEV, T. (2000) "The Forward Premium Anomaly is not as Bad as You Think", *Journal of International Money and Finance*, 19: 471-88.

4   CHATTERJEE, S. (1986) "Bootstrapping ARMA Models: Some Simulations", *IEEE Transactions on Systems, Man and Cybernetics*, SMC-16: 294-99.

5   CHUNG, C.-F. & BAILLIE, R.T. (1993) "Small Sample Bias in Conditional Sum-of-Squares Estimators of Fractionally Integrated ARMA Models", *Empirical Economics*, 18: 791-806.

6   CROWDER, W.J. (1994) "Foreign Exchange Market Efficiency and Common Stochastic Trends", *Journal of International Money and Finance*, 13: 551-64.

7   DE KOSTER, F.H. (1999) "The Bootstrap Approach to Autoregressive Time Series Analysis", Unpublished M.Sc dissertation, University of Pretoria.

8   DIEBOLD, F.X. & RUDEBUSCH, G.D. (1991) "On the Power of Dickey-Fuller Tests against Fractional Alternatives", *Economics Letters*, 35: 155-60.

9   EFRON, B. (1979) "Bootstrap Methods: Another Look at the Jackknife", *Annals of Statistics*, 7: 1-26.

10 FAMA, E. (1984) "Forward and Spot Exchange Rates", *Journal of Monetary Economics*, 14: 319-38.

11 FOX, R. & TAQQU, M.S. (1986) "Large Sample Properties of Parameter Estimates for Strongly Dependent Stationary Gaussian Time Series", *Annals of Statistics*, 14: 517-32.

12 GEWEKE, J. & PORTER-HUDAK, S. (1983) "The Estimation and Application of Long Memory Time Series Models", *Journal of Time Series Analysis*, 4: 221-38.

13 GRANGER, C.W.J. (1980) "Long Memory Relationships and the Aggregation of Dynamic Models", *Journal of Econometrics*, 14: 227-38.

14 GRANGER, C.W.J. & JOYEUX, R. (1980) "An Introduction to Long Memory Time Series Models and Fractional Differencing", *Journal of Time Series Analysis*, 1: 15-39.

15 HASSLER, U. & WOLTERS, J. (1994) "On the Power of Unit Roots against Fractionally Integrated Alternatives", *Economics Letters*, 45: 1-5.

16 HOSKING, J.R.M. (1981) "Fractional Differencing", *Biometrika*, 68: 165-76.

17 HOSKING, J.R.M. (1984) "Modeling Persistence in Hydrological Time Series Models Using Fractional Differencing", *Water Resources Research*, 20: 1898-1908.

18 KWIATKOWSKI, D., PHILLIPS, P.C.B., SCHMIDT, P. & SHIN, Y. (1992) "Testing the Null Hypothesis of Stationarity against the Alternative of a Unit Root: How sure Are We that Economic Time Series Have a Unit Root?", *Journal of Econometrics*, 54: 159-78.

19 LEE, D. & SCHMIDT, P. (1996) "On the Power of the KPSS Test of Stationarity against Fractionally-Integrated Alternatives", *Journal of Econometrics*, 73: 285-302.

20 NEWEY, W.K. & WEST, K.D. (1987) "A Simple Positive Semi-Definite Heteroskedasticity and Autocorrelation Consistent Covariance Matrix", *Econometrica*, 55: 703-08.

21 PHILLIPS, P.C.B. (1987) "Time Series Regression with a Unit Root", *Econometrica*, 55: 277-301.

22 PHILLIPS, P.C.B & PERRON, P. (1988) "Testing for a Unit Root in Time Series Regression", *Biometrika*, 75: 335-46.

23 SOWELL, F. (1992) "Maximum Likelihood Estimation of Stationary Univariate Fractionally Integrated Time Series Models", *Journal of Econometrics*, 53: 165-88.

24 SWANEPOEL, J.W.H. (1990) "A Review of Bootstrap Methods", *South African Statistical Journal*, 24: 1-34.

25 VAN STADEN, P.J. (2001) "Fractional Integration and Co-Integration", Unpublished M.Com dissertation, University of Pretoria.