Behavior of Exchange Rates and Returns: Long Memory and Cointegration

E.M. Syczewska

Warsaw School of Economics, Institute of Econometrics, Madalinskiego 6/8, 02-513 Warsaw, Poland

The aim of the paper is to present an example of analysis of exchange rate behavior with use of tools, built in GRETL econometric package, which have been developed by researchers often with background in physics or similar fields, but some (such as tests of integration and cointegration) are less known to physical audience. The series of interest is a bilateral USDPLN exchange rate; including the corresponding stock indices as additional variables can improve quality of a model even in period of crisis.

PACS: 89.65.GH

1. Introduction

In this paper we apply tools built in a GRETL econometric package (widely used in teaching and research)\textsuperscript{*} to closing daily values of bilateral USDPLN exchange rate and returns. The data set covers ten years (2000–2010), including period of the financial crisis, which makes modeling more difficult. Several methods, well known and established in time series econometrics, have been developed by researchers with background in physics or science. Some tools used here are perhaps less known to non-econometric audience, hence we describe them in greater detail than is perhaps necessary for econometricians. We note also deficiencies of some tests and methods, in hope that interested readers of different background might suggest some improvements or better choice of algorithms. All comments and suggestions of this kind are welcome\textsuperscript{†}.

This is not a detailed overview of financial econometrics, rather an application of selected methods to the financial time series of interest. The empirical example has been suggested by a research by Bauwens, Rime and Succarat [1] on bilateral exchange rates of the Norwegian crown. They managed to improve quality of their model by using stock exchange indices of respective countries as additional explanatory variables for the bilateral exchange rate. We follow their example using SP500 and WIG20 indices returns to explain the USDPLN daily returns. To check whether there is a stable dynamic economic equilibrium for exchange rate and corresponding stock indices, cointegration analysis of the series is performed. We compute the fractional integration parameter and the Hurst exponent, using algorithms built in GRETL, to check properties of the series.

Let \( \{y_t\} \), \( t = 1, 2, \ldots, N \) denote a series of closing values of exchange rate or a stock index. We use a typical definition of logarithmic returns:

\[
    z_t = 100 \ast (\ln y_t \! - \! \ln y_{t-1})
\]

where \( y_t \) – closing values of an instrument. Exchange rates and stock indices have slowly decreasing autocorrelation function ("long memory") behavior shows in their Hurst exponent and fractional integration parameter estimates. The returns series shows changing volatility (volatility clustering), excess kurtosis, asymmetry of the probability density, but is stationary in mean, hence we can apply an ARMA or pure autoregressive model with finite number of lags to the mean. However to model the volatility clustering, a second equation (according to ARCH and GARCH models, introduced respectively by Engle and Bollerslev), is needed.

The tools applied in such a research often stem originally from technical sciences or physics: the Hurst exponent, from hydrological study [5] of 1950’s; the ARMA models, from Box and Jenkins [6] fundamental monograph collecting methods of time series analysis developed by engineers. In this paper we remind definitions of stationary and integrated time series, tests for nonstationarity (ADF and KPSS) and cointegration, check for properties typical for financial data series, and apply measures such as fractional integration parameter and Hurst exponent.

2. Nonstationarity, integration and cointegration

Operational definition of stationarity, used in economic time series analysis, is the following. A series is said to be stationary (see e.g. [7], p. 12) if all three conditions hold:
1. Expected value of a series, \( E[X_t] \) is constant, independent of time;
2. Variance \( \text{Var}(X_t) \) is constant and finite, independent of time;
3. Covariance \( \text{Cov}(X_t, X_s) \) depends only on \(|t - s|\).

Stationary process is characterized in the time domain by (see e.g. [8]):

---

\textsuperscript{*} e-mail: Ewa.Syczewska@sgh.waw.pl, evams@sgh.waw.pl

\textsuperscript{†} GRETL is a free open-source cross-platform package (see gretl.sourceforge.net, and also www.kufel.torun.pl for a Polish translation of the software).

\textsuperscript{1} The author wants to express gratitude to participants of the FENS 2010 and anonymous referees for their questions and comments. All usual caveats apply.
1. Its mean: $\bar{x} = \frac{1}{N} \sum_{t=1}^{N} x_t$
2. Covariance: $C_t = \left[ 1/(N - \tau) \right] \sum_{\tau=1}^{N-\tau} (x_{t-\tau} - \bar{x})(x_{t-\tau} - \bar{x})$
3. Autocorrelation function: $R_t = \frac{C_t}{\sigma_\varepsilon^2} = \rho_\tau$
and in frequency domain by:
4. Periodogram: $I(\omega) = \frac{1}{2\pi} \sum_{\tau=0}^{\infty} \lambda_{\tau,\omega} C_\tau \cos \omega \tau$
5. Spectral density function: $f(\omega_j) = \frac{1}{2\pi} \sum_{\tau=0}^{\infty} \lambda_{\tau,\omega_j} C_\tau \omega_j \tau$ where $\omega_j = 2\pi j/N$ denote the Fourier frequencies (see [7], p. 331); $N$ – number of observations; $\lambda_{\tau,\omega}$ are appropriate weights.

Most of economic and financial time series are nonstationary. We check this with use of simplest nonstationarity tests — an Augmented Dickey-Fuller test, in-stationary. We check this with use of simplest nonsta-

In frequency domain by:

Most of economic and financial time series are nonstationary. We check this with use of simplest nonstationarity tests — an Augmented Dickey-Fuller test, in-stationary. We check this with use of simplest nonsta-

Cointegration of series $y_t, x_{1t}, x_{2t}, \ldots, x_{kt} \sim I(1)$ means that those series are nonstationary, but there is a linear combination which is stationary; in more general case, cointegration between variables exists if there is a linear combination with lower order of integration than the variables themselves (see [12]). Vector of coefficients of this linear combination is called a cointegrating vector. According to Granger, a cointegrating vector is an attractor for a trajectory of an economic system. As explained by Maddala and Kim [16], if we know from economic theory that there is a stable dynamic equi-

### 3. Typical behavior of a financial time series

Empirical distribution of the exchange rate series is asymmetric and non-normal. Periods of higher volatility correspond to increased risk of investment. Logarithmic returns show so-called volatility clustering, i.e. high autocorrelation of conditional variance. This is called an ARCH effect, after Autoregressive Conditional Heteroskedasticity model (introduced by Engle [17]; see also [18]).

Figs. 1 and 2 show respectively daily observations of USDPLN exchange rate and of its logarithmic returns, for period since 2000 until 2010. Its logarithmic returns are stationary in mean (if not in variance), as shown by the ADF and KPSS test results. Note especially more volatile behavior of the series during the last financial crisis (Fig. 2). This suggests that it is worthwhile to test for the ARCH effect.

---

1. Sir Clive William John Granger (Sept. 4, 1934 – May 27, 2009), a British economist, and Robert F. Engle (born Nov. 10, 1942), an American economist, were awarded a Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel, for their discoveries in analysis of time series data: R.F. Engle "for methods of analyzing economic time series with time-varying volatility (ARCH)"; and C.W.J. Granger for "methods of analyzing economic time series with common trends (cointegration)"; C.W.J. Granger had BSc in mathematics and PhD in statistics, R.F. Engle has B.Sc. and M.Sc. in physics and PhD in economics (see [13] and also biographical information there).
squares of residuals: regression of squared residuals on a constant and lagged

\[ \Delta \text{series} \{ f \} = \alpha_0 + \alpha_1 e_{t-1}^2 + \alpha_2 e_{t-2}^2 + \ldots + \alpha_k e_{t-k}^2 + u_t \]  

where \( e_t \) are error terms of the model in question. We check whether lagged error squares are jointly significant:

\[ \chi^2 \text{null, the test statistic} \]

\[ \frac{\text{residual squares}}{\text{degrees of freedom}} \]

where \( \chi^2 \text{null} \), the test statistic

\[ (4) = \sum \frac{(\frac{d}{k})}{\Gamma(k)\Gamma(k+1)} L^k \]

and \( L \) denotes lag operator.

Properties of a series can be classified according to \( d \), in a following way:

- If \( d = 1 \), the process is integrated and has infinite variance,
- If \( d > 1 \), the process is also nonstationary, and effects of external shocks increase in time.
- If \( 0.5 \leq d < 1 \), variance is infinite, hence the process is also nonstationary, but in long time is mean-reverting. Effects of shocks last for a long time.
- If \( 0 < d < 0.5 \), the process is stationary, mean-reverting, with finite variance.
- If \( d = 0 \), the process is mean-reverting in a short time, has finite variance, and shock effect diminish quickly.
- If \( d < 0 \), the process is antipersistent (mean-averting) and stationary.

Good overview of applications of fractional integration to financial data — exchange rates, asset returns, interest rates, inflation — is given in [19].

3.2. Estimation of fractional integration parameter

In the GRETL package, the Geweke and Porter-Hudak [20] periodogram regression method and the Whittle method are used (for the Whittle method, see [21] and [22], for other methods, see Robinson [23] or [24]).

According to Granger, for a stationary series \( X \) and white noise \( u \), if \( \Delta^d X_t = u_t \) and \( u_t \) is stationary with zero mean and continuous spectral density, \( f_u(\omega) > 0 \), then:

\[ f_u(\omega) = \frac{1}{1 - \exp(i\omega)^{-2d} f_u(\omega)} \]

Phillips [25] shows that for a nonstationary series this is a limit of periodogram ordinates. For fundamental frequencies \( \omega_s = \frac{2\pi s}{N} \), where \( N \) is number of observations, \( s = 1, 2, \ldots, m \), a regression \( \log I_x(\omega_s) = c - d \log(1 - \exp(i\omega_s)) + \text{residual} \) is estimated with OLS, hence this kind of fractional integration parameter estimates are called periodogram regressions.

As the periodogram regression is estimated with OLS method, we can use the parameter estimates and standard errors to test hypothesis concerning \( d \), namely, whether \( d = 0 \) for a stationary series, or \( d = 1 \) for a nonstationary series.

As we see, the estimates of a fractional integration parameter indicate properties of the series. Another measure, the Hurst exponent, has different origin (see [5]): it was introduced by the British hydrologist, Harold Edwin Hurst, during his research on Nile**. But the Hurst ex-

** According to [26], Hurst (Jan. 1, 1880–Dec. 7, 1978) obtained a first class honour in physics at Oxford University, for three years remained at university as a lecturer and researcher, and in
ponent can be used in a similar way to classify behavior of a series:

- $0 < H < 0.5$ indicates a series with negative autocorrelation,
- $H > 0.5$ indicates a series with positive autocorrelation,
- $H = 0.5$ indicates a random walk.

This tool was applied to financial time series by Mandelbrot in numerous papers, but its widespread use among practitioners is perhaps due to Peters books [28, 29], translated into several languages. Peters’ algorithm of computing the Hurst exponent goes along the following way. Let $r_t$ denote logarithmic returns, $m(N, t_0) = \sum_{t=t_0+1}^{t_0+N} r_t/N$ – mean of a series, then

$$ S(N, t_0) = \left\{ \frac{1}{N} \sum_{t=t_0+1}^{t_0+N} [r_t - m(N, t_0)]^2 \right\}^{1/2} $$

is a biased estimator of standard deviation.

**TABLE I**

<table>
<thead>
<tr>
<th>Size</th>
<th>$\text{RS(avg)}$</th>
<th>$\log(\text{Size})$</th>
<th>$\log(\text{RS})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2641</td>
<td>72.964</td>
<td>11.367</td>
<td>6.1891</td>
</tr>
<tr>
<td>1320</td>
<td>52.398</td>
<td>10.366</td>
<td>5.7114</td>
</tr>
<tr>
<td>660</td>
<td>32.131</td>
<td>9.3663</td>
<td>5.0059</td>
</tr>
<tr>
<td>330</td>
<td>23.909</td>
<td>8.3663</td>
<td>4.5795</td>
</tr>
<tr>
<td>165</td>
<td>16.679</td>
<td>7.3663</td>
<td>4.0060</td>
</tr>
<tr>
<td>82</td>
<td>11.353</td>
<td>6.3576</td>
<td>3.5051</td>
</tr>
<tr>
<td>41</td>
<td>7.2341</td>
<td>5.3576</td>
<td>2.8548</td>
</tr>
<tr>
<td>20</td>
<td>4.7010</td>
<td>4.3219</td>
<td>2.2330</td>
</tr>
<tr>
<td>10</td>
<td>3.0911</td>
<td>3.3219</td>
<td>1.6281</td>
</tr>
</tbody>
</table>

Coefficient | Standard error
---|---
intercept | -0.1773 | 0.0725
slope | 0.5645 | 0.0093

Source: own computations in GRETL

Partial sums and range of partial sums of deviations from a mean are defined as $X(N, t_0, \tau) \equiv \sum_{t=t_0+1}^{t_0+N} (r_t - m(N, t_0))$, for $1 \leq \tau \leq N$,

$$ R(N, t_0) \equiv \max_{\tau} X(N, t_0, \tau) - \min_{\tau} X(N, t_0, \tau). $$

Rescaled range statistics, defined as $[R/S](N) \equiv \frac{\sum_{t_0}^{t_0+N} R(N, t_0)}{\sum_{t_0}^{t_0+N} S(N, t_0)}$, is equal to $[R/S](N) \approx (aN)^H$, where $a$ — constant term, $H$ — the Hurst exponent. Hence as an estimate of the Hurst exponent the following regression results can be used: $\log[R/S](N_i) = \hat{c} + H \log N_i$, where $N_i$ correspond to several subsamples (of the original series), for which the $R/S(N_i)$ statistics are computed.

In GRETL, the Hurst exponent is computed according to this algorithm. Computations for logarithmic returns of USDPLN daily data are shown in Table I. The estimate of the Hurst exponent is equal to 0.5645, indicating the long run dependence.

4. Example for USDPLN exchange rate and returns

Table II shows results of the Dickey-Fuller test for stock indices and exchange rates under study.

They indicate nonstationarity of levels and stationarity or logarithmic returns. Similar are results of fractional integration parameter computations (Table III). Only for the USDPLN exchange rate returns the null of nonstationarity can be rejected, however value of 0.10 is still in the range of stationarity. The Hurst exponent estimates (Table IV) also show nonstationarity of series and stationarity of returns.

### 4.1. Cointegration analysis for exchange rate and indices

As exchange rates and stock indices are integrated of order 1, we next check whether there is a stable dynamic relationship between them. The Engle and Granger [12] method of cointegration testing is based on testing for stationarity of the OLS residuals from a regression of one $I(1)$ variable on the rest. Stationarity of residuals means that the series are cointegrated and the OLS estimates of parameters give the cointegrating vector (presumably describing a stable relationship between the variables).

Fig. 3. Residuals of Engle-Granger regression.

Regression of USDPLN closing values on S&;P500 and WIG20 closing values gives the following results (Table V).

If $[1, -0.00509, 0.00126]$ were a cointegration vector for USDPLN, SP500 and WIG20, then residuals of this regression should be stationary. TheADF test statistics for residuals equals $-2.813$, and has asymptotic p-value 0.056 — only slightly higher than 5%, but visual inspection of the residuals (Fig. 3) convinces us that their behavior is rather too volatile for stationarity. Hence we do not reject nonstationarity of residuals, and cannot use the OLS estimates for description of a stable relationship.
4.2. ARIMA model for USDPLN

We next estimate an ARIMA model for first differences of the exchange rate, and with corresponding stock indices as additional explanatory variables – an ARMAX model for differences of USDPLN, based on observations 2000/01/05–2010/10/18. The results are shown in Table VI. All variables are significant, roots of polynomials have moduli greater than 1, hence the model is stable.

For the above ARMAX model, the Engle test statistics $\text{LM} = 307.262$ with $p\text{-value} = \frac{\chi^2(5)}{307.26} < 0.001$, close to zero. Hence the null hypothesis of no ARCH effect is clearly rejected (as suggested by Fig. 4, which shows that the ARMAX model residuals, albeit stationary in mean, show changing volatility).

![Fig. 4: Residuals of the ARMAX model show an ARCH effect.](image)

The augmented Dickey-Fuller test for stock indices and exchange rates

<table>
<thead>
<tr>
<th>ADF test for</th>
<th>Sample up to April 30</th>
<th>Sample up to November 19, 2010</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable</td>
<td>Log returns</td>
<td>Variable</td>
</tr>
<tr>
<td>SP500close</td>
<td>–1.825 [0.369]</td>
<td>–1.900 [0.332]</td>
</tr>
<tr>
<td>WIG20close</td>
<td>–1.148 [0.699]</td>
<td>–1.152 [0.697]</td>
</tr>
<tr>
<td>USDPLNclose</td>
<td>–1.566 [0.500]</td>
<td>–1.617 [0.474]</td>
</tr>
<tr>
<td>EURUSDclose</td>
<td>–1.261 [0.650]</td>
<td>–1.280 [0.641]</td>
</tr>
<tr>
<td>EURPLNclose</td>
<td>–2.125 [0.235]</td>
<td>–2.209 [0.203]</td>
</tr>
</tbody>
</table>

Asymptotic p-values in brackets.

Source: own computations in GRETL

The Hurst exponents

<table>
<thead>
<tr>
<th>Hurst exponent for</th>
<th>Variable</th>
<th>logarithmic ret.</th>
</tr>
</thead>
<tbody>
<tr>
<td>SP500 close</td>
<td>0.9689</td>
<td>0.5510</td>
</tr>
<tr>
<td>WIG20 close</td>
<td>1.0045</td>
<td>0.5685</td>
</tr>
<tr>
<td>USDPLN close</td>
<td>1.0018</td>
<td>0.5645</td>
</tr>
</tbody>
</table>

Source: own computations in GRETL

4.3. Spectral density

Figs. 5 and 6 show respectively periodogram of logarithmic returns for USDPLN exchange rate and approximation of its spectral density function, obtained in GRETL by smoothing the periodogram with appropriate weights (Bartlett weights)\[‡‡]\. The periodogram and the spectral density approximation are also tools to detect possible periodicity of the series. Fig. 6 suggests that there is some periodicity corresponding to a cycle of half of the week, one week, and approximately two weeks (note corresponding local maxima).

\[‡‡\] Similar weights are used in computation of unbiased estimator of long-term variance of a series, in case of autocorrelation and heteroskedasticity of the disturbance.
TABLE V

<table>
<thead>
<tr>
<th>Variable</th>
<th>Estimate</th>
<th>t-ratio</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>SP500close</td>
<td>0.00509</td>
<td>133.1</td>
<td>0.0000*</td>
</tr>
<tr>
<td>WIG20close</td>
<td>-0.00126</td>
<td>-61.52</td>
<td>0.0000*</td>
</tr>
</tbody>
</table>

Source: own computations in GRETL

TABLE VI

ARMAX model results for ∆USDPLN

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. error</th>
<th>z stat.</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Const</td>
<td>-0.000448</td>
<td>0.000530</td>
<td>-0.846</td>
<td>0.3974</td>
</tr>
<tr>
<td>φ₁</td>
<td>1.30899</td>
<td>0.216640</td>
<td>6.042</td>
<td>1.52·10⁻⁹*</td>
</tr>
<tr>
<td>φ₂</td>
<td>-0.5598</td>
<td>0.159075</td>
<td>-3.519</td>
<td>0.0004*</td>
</tr>
<tr>
<td>θ₁</td>
<td>-1.2727</td>
<td>0.226735</td>
<td>-5.613</td>
<td>1.99·10⁻⁸*</td>
</tr>
<tr>
<td>θ₂</td>
<td>0.50518</td>
<td>0.168287</td>
<td>3.002</td>
<td>0.0027*</td>
</tr>
<tr>
<td>SP500cl.</td>
<td>-0.0001784</td>
<td>3.87039·10⁻⁵</td>
<td>-4.609</td>
<td>4.05·10⁻⁶*</td>
</tr>
<tr>
<td>WIG20cl.</td>
<td>-0.0001908</td>
<td>1.58094·10⁻⁵</td>
<td>-12.070</td>
<td>1.51·10⁻³³*</td>
</tr>
</tbody>
</table>

Source: own computations in GRETL

4.4. The ARCH-GARCH models for logarithmic returns, with indices returns as additional explanatory variables

The ARCH (Autoregressive Conditional Heteroskedasticity) model, introduced by Engle [17] (see also [18]), consists of one equation for expected value of a series, and second equation for conditional variance of a series. The Generalized ARCH models, introduced by Bollerslev [30–2] have an advantage of requiring smaller number of parameters to adequately represent the series (see also [33] for description and examples of other GARCH-type models and [34] for detailed analysis and examples for the Polish data).

TABLE VII

The GARCH results for a whole sample

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. error</th>
<th>z</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>rlUSD₁₋₁</td>
<td>0.0702</td>
<td>0.0205</td>
<td>3.432</td>
<td>0.0006*</td>
</tr>
<tr>
<td>rlUSD₁₋₂</td>
<td>0.0354</td>
<td>0.0204</td>
<td>1.733</td>
<td>0.0830*</td>
</tr>
<tr>
<td>α(0)</td>
<td>0.0121</td>
<td>0.0028</td>
<td>4.255</td>
<td>1.83·10⁻⁵*</td>
</tr>
<tr>
<td>α(1)</td>
<td>0.0791</td>
<td>0.0105</td>
<td>7.503</td>
<td>6.25·10⁻¹⁴*</td>
</tr>
<tr>
<td>β(1)</td>
<td>0.9053</td>
<td>0.0119</td>
<td>76.32</td>
<td>0.0*</td>
</tr>
</tbody>
</table>

Source: own computations

TABLE VIII

Additional variables can improve forecast accuracy

<table>
<thead>
<tr>
<th>GARCH model:</th>
<th>without stock indices</th>
<th>with stock indices</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Error</td>
<td>0.1779</td>
<td>0.1769</td>
</tr>
<tr>
<td>Mean Squared Error</td>
<td>1.3414</td>
<td>1.2460</td>
</tr>
<tr>
<td>Root Mean Squared Error</td>
<td>1.1582</td>
<td>1.1163</td>
</tr>
<tr>
<td>Mean Absolute Error</td>
<td>0.9747</td>
<td>0.9480</td>
</tr>
<tr>
<td>Bias proportion, UM</td>
<td>0.0236</td>
<td>0.0251</td>
</tr>
<tr>
<td>Regression proportion, UR</td>
<td>0.1431</td>
<td>0.2625</td>
</tr>
<tr>
<td>Disturbance proportion, UD</td>
<td>0.8333</td>
<td>0.7124</td>
</tr>
</tbody>
</table>

Source: own computations

We estimate a GARCH model for logarithmic returns of USDPLN exchange (with 2 lagged variables in the mean equation, and 1 lag for both variance and squared error in the conditional variance equation, see Table VII), and the GARCH model with stock indices returns as additional explanatory variables in the mean equation (not shown here), next reestimate both GARCH models for a shorter sample, and compare quality of ex-post forecasts for last month of data (Table VIII). Almost all measures (with the only exception of the regression proportion) have slightly lower values for the model with addi-
tional explanatory variable. Stock indices returns indeed slightly improve performance of the model.

5. Conclusions

There is a vast econometric literature concerning non-linear models of financial time series of changing volatility, see [34] for a detailed analysis with application to the Polish markets. The crisis period and increased volatility and risk, make the task of the exchange modeling much more difficult than usual; hence any specification which can improve the quality of the model and forecasts may be of interest. In our empirical example, as in [3] for shorter time series, use of returns of stock indices led to slight improvement of ARMA and GARCH models for the exchange rate returns. The last version of GRETL allows a choice of several GARCH-type models with variants of probability distributions for its error terms.

There are also several variants of non-stationarity tests used in applied econometrics, some of them already implemented in GRETL.

Perhaps if we look for a way of improving the analysis presented here, it would be to choose and implement better versions of the Hurst algorithms — and to this aim results of research such as [35], [36], can be of help.

References