Lund University

Master Thesis in Economics

Does PPP hold in the long run? An empirical approach using wavelets.

Author: David Berger, 870803-1532

Tutor: Fredrik NG Andersson

Place, Date: Lund, June 2012

Abstract

This paper contributes to the debate as to whether or not Purchasing Power Parity (PPP) holds in the long run. This is done by looking at fractionally integrated processes (FIP) and using wavelets in order to obtain an Ordinary Least Squares (OLS) estimate of the long run memory parameter. Firstly, drawing on the theoretical framework behind PPP, this paper outlines the statistical background and terminology relating to fractionally integrated processes, unit roots, and wavelets. Secondly, said concepts are empirically applied and tested. Based on this, the results show that real exchange rates are mean reverting and subject to long swings, which indicates that unit root tests are inappropriate for analyzing exchange rates.

Keywords: Purchasing Power Parity, Time Series Analysis, Wavelets, Fractionally Integrated Processes, Long Run Memory

Contents

"Would you like me to quote some statistics?"

"Er, well..."

"Please, I would like to. They, too, are quite sensationally dull." ~ *Life, the Universe, and Everything*

1 Introduction

The law of one price states that, in the absence of trade barriers and transaction costs, two identical goods which are bought in different markets will have an identical price, when expressed in the same currency. This results from the fact that it would otherwise be possible to buy a good cheaper in one market and sell it at a higher price in another one (this is also known as arbitrage). As a result, the real exchange rate between two countries should, according to PPP, be constant over time. Originally developed by scholars of the Salamanca school in Spain during the sixteenth century (Rogoff, 1996), the modern-day formulation of PPP has been put forth by Cassel (1918). He proposed that "[a]s long as anything like free movement of merchandise and a somewhat comprehensive trade between the two countries takes place; the actual rate of exchange cannot deviate very much from this purchasing power parity" (Cassel, 1918, pg. 413) and today PPP is widely applied when comparing price levels in different countries. More specifically, this can be expressed as

$$RE = \frac{SP}{P^*}, \tag{1.1}$$

where RE denotes the real exchange rate, P and P^* domestic and foreign price levels respectively and S the nominal exchange rate ($\frac{domestic\,currency}{foreign\,currency}$).

If PPP were to hold, the real exchange rate would be stationary over time as arbitrage would otherwise be possible. However, in the short run it is always possible to exploit real exchange rate differentials between two countries as prices are sticky and take time to adjust. This phenomenon is partly captured by Dornbusch's overshooting model and there is a broad consensus in the literature that PPP does not hold in the short run (Artus, 1978; Dornbusch, 1980; Frenkel, 1981; Taylor, 2002) and therefore follows a random walk process.

This is particularly problematic because "[i]f exchange rates are random walks, then almost everything we say about monetary policy is wrong" (Alvarez, Atkeson & Kehoe, 2007, pg. 339).

However, as the market re-prices, long and short run equilibria converge until it is no longer possible to exploit exchange rate differentials through arbitrage. The real exchange rate should therefore, in the long run, be stationary. Testing for stationarity is "traditionally" done using unit root tests. Employing these, most previous studies have detected unit roots and therefore concluded that PPP follows a random walk (Adler & Lehmann, 1983; Lopez, Murray & Papell, 2004; Belaire-Franch & Opong, 2005). However, unit root tests have low power, especially with respect to unit roots close to one (Abuaf & Jorion, 1990; Andersson, 2012), which makes it very hard to distinguish between stationary and non-stationary time series. Hence, this fact can also be seen as evidence against unit root tests rather than PPP (Abuaf & Jorion, 1990; Engel & Hamilton, 1990).

This indicates that the real exchange rate may be fractionally integrated. A FIP is considered to be stationary if its covariance matrix is finite. More specifically, stationarity implies that the order of integration, d, is given by $0 < d < 1/2$ (Percival & Walden, 2000). However, it has been shown by Granger and Joyeux (1980) that as long as $1/2 \leq d < 1$, the time series has a stationary mean even though the covariance matrix is not finite. Provided that $d < 1$, the FIP can therefore be considered mean reverting and subject to long swings.

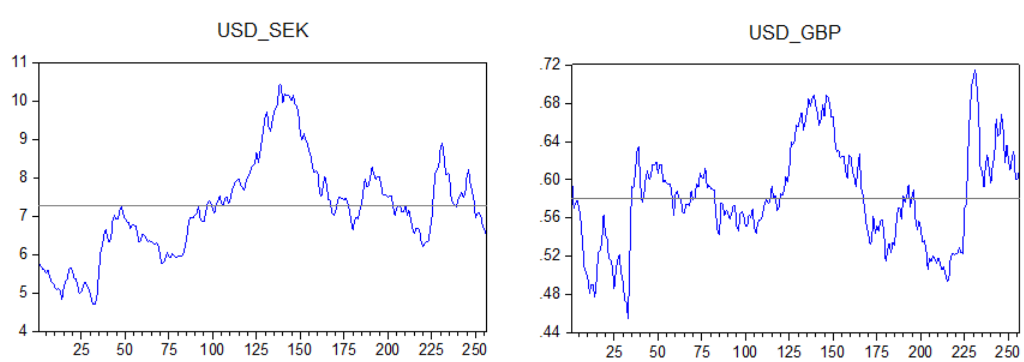

Figure 1: Real Exchange Rate, USD/SEK and USD/GBP, January 1990 - April 2011

Figure 1 shows the real USD/SEK and USD/GBP exchange rates[1] and both graphs appear as if they are mean reverting but subject to long swings ($1/2 \leq d < 1$). This process cannot be represented by an AR(p) model (Andersson, 2012) and it is therefore necessary to consider a different model, like a FIP.

It is the aim of this paper to determine the long run order of integration of a fractionally integrated process, namely PPP, using wavelet analysis and to contribute to the debate to what extent PPP holds in the long run. In accordance with Granger and Joyeux (1980), this paper utilizes the approach put forward by Jensen (1999) in order to use wavelets to estimate the long run order of integration. Based on this, it is then possible to draw conclusions with respect to the stationarity of PPP.

The subsequent sections of this paper are organized in the following way: section two outlines the theoretical foundations behind integrated processes, including unit root and wavelet theory, while section three presents the empirical results. Finally, a conclusion is drawn in the last part of the paper.

2 Integrated Processes

The data generating process (DGP) of an integrated time series is given by

$$(1-L)^d y_t = \varepsilon_t, \tag{2.1}$$

where ε_t is white noise ($var(\varepsilon) = \sigma_\varepsilon^2$), d the fractional integration order and L represents the lag operator

$$Ly_t = y_{t-1}. \tag{2.2}$$

[1]The gray line represents the average real exchange rate over the given interval.

In general, non-fractionally integrated processes are integrated of order one ($d = 1$) or zero ($d = 0$). More specifically, d can take any real value, i.e. $d \in \mathbb{R}$. It is therefore possible to model a random walk process by substituting ($d = 1$) into equation 2.1. This gives the following

$$y_t = y_{t-1} + \varepsilon_t. \tag{2.3}$$

This autoregressive process is usually referred to as random walk, has a unit root and is non-stationary. Similarly, one can model a stationary process by substituting $d = 0$ into equation 2.1, which yields

$$y_t = \varepsilon_t. \tag{2.4}$$

Given that equation 2.3 represents a non-stationary process, it will only become stationary when taking the first difference. If a time series needs to be differenced once in order to become stationary, it is said to be integrated of order one $I(1)$. Theoretically, this can be iterated but in economics one can rarely find processes that need to be differenced more than twice to become stationary ($I(2)$).

Using the DGP given by equation 2.1, one can also model a MA process but since there is no fundamental difference between the two processes[2] (Verbeek, 2008), only AR models will be considered.

2.1 Unit Roots

There are a variety of ways to test for unit roots. The most common tests are the (Augmented) Dickey Fuller ((A)DF), Phillips-Perron (PP), and Kwiatkowski, Phillips, Schmidt and Shin (KPSS). However, there are evidently alternative ways of detecting unit roots such as the tests provided by Elliot, Rothenberg, and Stock Point Optimal (ERS) and Ng and

[2] In fact, an AR(*1*) model can also be written as a MA(∞) model.

Perron (NG) but since the first three are the most widely used tests, they will be discussed in more detail below.

As the name suggests, the ADF is based on a non-augmented Dickey Fuller test. A simple AR(*1*) process

$$y_t = \rho y_{t-1} + x_t'\delta + \varepsilon_t, \tag{2.5}$$

is considered to be non-stationary if $|\rho| \geq 1$. The standard DF test modifies equation 2.5 in such a way that y_{t-1} is subtracted from both sides:

$$\Delta y_t = \alpha y_{t-1} + x_t'\delta + \varepsilon_t. \tag{2.6}$$

Since $\alpha = \rho - 1$, this yields the following hypotheses:

$$\begin{cases} H_0: & \alpha = 0,\ (non - stationary) \\ H_1: & \alpha < 1.\ (stationary) \end{cases}$$

This test only works with respect to AR(*1*) models, as ε_t will otherwise not be white noise. In order to test for higher order lags, AR(p), the ADF is employed. It works by adding p lagged difference terms of y_t to the right-hand side of equation 2.6:

$$\Delta y_t = \alpha y_{t-1} + x_t'\delta + \beta_1 \Delta y_{t-1} + ... + \beta_p \Delta y_{t-p} + \nu_t. \tag{2.7}$$

Similarly, the above hypotheses can then be employed to test for stationarity. However, one of the main points of criticism towards the ADF is that it has low power (especially in small samples) if the series is stationary but with a root close to the non-stationary boundary. This means that it will falsely suggest a non-stationary time series while it in reality is not. This can certainly be an issue with respect to PPP. In fact, Abuaf and Jorion (1990, pg. 157) claim that “the negative results obtained in previous empirical research reflect the poor power of the test rather than evidence against PPP”. Furthermore, p is assumed to be a natural number. Hence, it is not possible to look at fractionally integrated processes. Lastly, an important practical issue is selecting the correct lag length (see equation 2.7). If p is too small the test will be biased and if it is too large the power of the test will suffer. Practically, this is issue is addressed through the use of information criteria (i.e. Schwarz, Akaike or Hannan-Quinn). Based on these, the statistical software will automatically chose the most appropriate lag length.

The PP test is very similar to the ADF but one does not have to consider selecting the correct lag length. Both tests are asymptotically equivalent (University of Washington, 2012) and employ the same hypotheses. It is based on the non-augmented DF (equation 2.6) and in order to control for serial correlation when testing for unit roots, the t-ratio of the α coefficient is modified. Hence, an adjusted t-value is reported.

Contrary to the two previous tests, the KPSS is a Lagrange Multiplier (LM) test and utilizes different hypotheses

$$\begin{cases} H_0: & \sigma_\varepsilon^2 = 0, \ (stationary) \\ H_1: & \sigma_\varepsilon^2 > 0. \ (non-stationary) \end{cases}$$

It is based on the residuals from the OLS regression of y_t on x_t and in employing the opposite hypotheses it is often seen as a complementary test to the ADF and PP (Hobijn, Franses & Ooms, 2004).

Lastly, it should be noted that, for all three unit root tests, one has to make assumptions about whether or not to include a constant and/ or linear trend. This is also the case with respect to testing the first difference for unit roots. However, given that one can rarely find processes in economics that need to be differenced more than once to become stationary (Gujarati, 2004), it is sensible to include both a trend and constant in the original test and a constant when testing the first difference.

2.2 Fractionally Integrated Processes

A FIP allows d to be any real number and its spectral density function (SDF) is given by

$$S_y(f) = \frac{\sigma_\varepsilon^2}{[4sin^2(\pi f)]^d}, \tag{2.8}$$

where f represents the frequency. There are a variety of estimators available to estimate the SDF, such as the direct spectrum estimator, the (smoothed) periodogram, thresholding, and the wavelet variance. The latter "summarizes" the information in the SDF (Percival & Walden, 2000) and has been used by Jensen (1999) and is therefore also employed in this paper. More importantly, one can see that different long run parameters correspond to a change in slope.

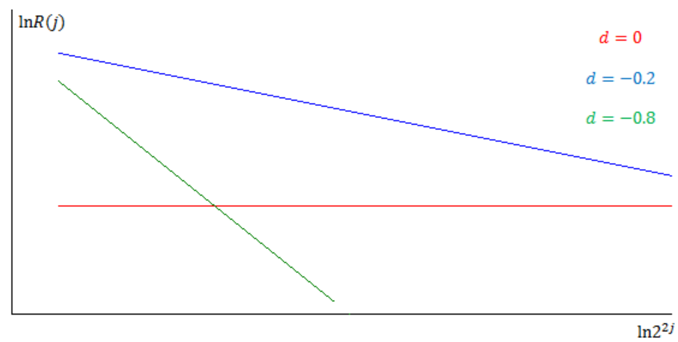

Figure 2: Wavelet OLS Estimator for different slopes of d

Figure 2 shows that a steeper slope represents a higher order of integration and vice versa[3]. As previously mentioned, the FIP is considered to be stationary if $0 < d < 1/2$ (Percival & Walden, 2000). However, it has been shown by Granger and Joyeux (1980) that as long as $1/2 \leq d < 1$, the time series is mean reverting even though the covariance matrix is not finite. Provided that $d < 1$, the FIP can therefore be considered mean reverting and subject to long swings.

The most common way of transforming a time series from the time domain to the frequency domain, is given by the Fourier transform

$$F(f) = \sum_{t=-\infty}^{\infty} x_t e^{-i2\pi ft}. \tag{2.9}$$

[3]Note the wavelet variance and its corresponding scale on the y and x-axis respectively. For a definition of scale, see section 2.3.1.

According to Euler's formula, $e^{-i2\pi ft} = cos(2\pi ft) - isin(2\pi ft)$. Hence, the Fourier transform expresses a function of time as a function of frequency, based on oscillating sine and cosine functions. It is important to note that for any $j \in \mathbb{Z}$, $F(f+j) = F(f)$. This is due to the fact that

$$F(f+j) = \sum_{t=-\infty}^{\infty} x_t e^{-i2\pi(f+j)t} = \sum_{t=-\infty}^{\infty} x_t e^{-i2\pi ft} e^{-i2\pi jt} = \sum_{t=-\infty}^{\infty} = x_t e^{-i2\pi ft} = F(f). \quad (2.10)$$

Hence, it is only necessary to consider frequencies over a *unit interval*. It is of course also possible to transform the signal back from the frequency domain to the time domain[4].

As it can be seen from the Fourier transform, the signal is now displayed in its frequency domain. However, the time domain has been entirely dropped. In order to provide a time and frequency resolution, wavelet analysis is used. This is a major advantage of wavelet analysis as "[b]y decomposing a time series into time-frequency space, one is able to determine both the dominant modes of variability and how those modes vary in time" (Torrence & Compo, 1998, pg. 61). Furthermore, wavelet analysis has advantages over Fourier analysis with respect to non-stationary and inhomogeneous systems (Abramovich, Bailey & Sapatinas, 2000). However, one should note that it is not possible to simultaneously determine the time and location of a signal as a result of Heisenberg's uncertainty principle (Battle, 1997).

2.3 Wavelets

As the name suggests, wavelets are small waves that have finite energy and grow/ decline in a limited time period. By contrast, a "big wave" is function that starts/ ends at no particular point in time and has infinite energy. An example of such a big wave is a sine function, which oscillates up and down indefinitely. While time series are usually assessed in the time domain, it is equally possible to analyze them in the frequency domain.

[4]For proof, see Andersson (2008).

2.3.1 Discrete Wavelet Transform

The most commonly used wavelet is the Haar wavelet. It is considered to be the first wavelet and is named after the Hungarian mathematician Alfréd Haar who first developed it in 1910 as an analysis tool. For different scales, it is defined as follows

$$\psi^H_{\lambda,u}(t) = \begin{cases} -\frac{1}{\sqrt{2\lambda}} & u-\lambda < t \leq u; \\ \frac{1}{\sqrt{2\lambda}} & u < t \leq u+\lambda; \\ 0 & otherwise. \end{cases} \tag{2.11}$$

Graphically this can then be displayed as shown in Figure 3.

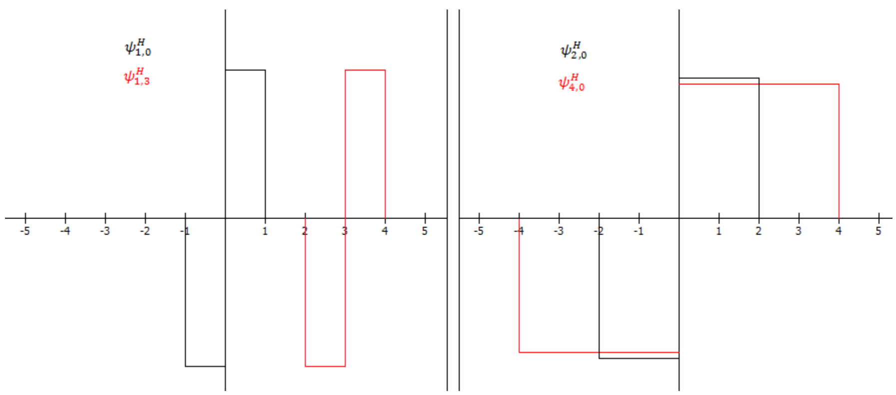

Figure 3: Haar Wavelet, different scales and translations

Figure 3 shows the Haar wavelet for different scales, λ, and translations, u. The latter shifts the wavelet along the x-axis while the former shows how averages of y_t "over many different scales are changing from one period of length λ to the next" (Percival & Walden, 2000, pg. 10). More specifically, a frequency interpretation is possible as the first scale captures frequencies $\frac{1}{4}$ to $\frac{1}{2}$, the second frequencies from $\frac{1}{8}$ to $\frac{1}{4}$ and so forth (this is sometimes referred to as a band pass filter). A wavelet filter that captures high frequencies and attenuates low ones is called high pass filter, while the scaling filter is the low pass filter.

In general, one distinguishes between Continuous Wavelet Transform (CWT) and Discrete Wavelet Transform (DWT). While the former deals with time series that are defined over the entire real axis, the latter is practically more applicable as it can be applied to a discrete set of times, e.g. $t = 0, 1, ..., T - 1$, which is also why it will be applied in this paper.

The DWT uses dyadic scales, $\lambda = 2^{j-1}$, and can therefore be computed for any amount of observations that can be expressed as a power of two. While this is not strictly true and it is also possible to "directly" use T observations through methods such as the maximum overlap discrete wavelet transform (MODWT), this goes beyond the scope of this paper and will therefore not be discussed[5].

The DWT can be obtained through the *pyramid algorithm*, which was first proposed by Mallat (1989). There are other ways of calculating the DWT but since Mallat's algorithm is the most efficient (Andersson, 2008), it will be employed in this paper.

As the name suggests, the pyramid algorithm is subdivided into several steps that resemble a pyramid[6]. The first stage "simply consists of transforming the time series $\mathbf{X}$ of length $[T = 2^J]$ into the $[T/2]$ first level wavelet coefficients $\mathbf{W_1}$ and the $[T/2]$ first level scaling coefficients $\mathbf{V_1}$" (Percival and Walden, 2000, pg. 80). This procedure is iterated until $\mathbf{W_j}$ only contains one entry. In order to obtain the wavelet and scaling coefficients, the corresponding wavelet and scaling filters are employed (which differ depending on the wavelet of choice, e.g. Haar, Daubechie, Mexican Hat, etc.). For the Haar wavelet, the wavelet filter, h, for the first scale, $\lambda = 1$, is obtained by substituting this value into equation 2.11. This gives[7]

$$h = [h_0, h_1] = [1/\sqrt{2}, -1/\sqrt{2}]. \tag{2.12}$$

Similarly, the scaling filter, g, can be obtained through its relationship with the wavelet filter

$$g_0 = -h_1 = 1/\sqrt{2}, \tag{2.13}$$

$$g_1 = h_0 = 1/\sqrt{2}. \tag{2.14}$$

Hence, the complete scaling filter for the Haar wavelet is given by

[5]For a more detailed discussion, see Percival and Walden (2000).
[6]Even though it actually more resembles an upside down pyramid.
[7]These values retain the orthonormality property of the wavelet transform.

$$g = [g_0, g_1] = [1/\sqrt{2}, 1/\sqrt{2}]. \tag{2.15}$$

The complete wavelet transform is given by

$$w = \mathbf{\Phi} x, \tag{2.16}$$

where

$$\mathbf{\Phi} = \begin{bmatrix} \mathbf{\Phi_1} \\ \mathbf{\Phi_2} \\ \vdots \\ \mathbf{\Phi_J} \\ \mathbf{\Gamma_J} \end{bmatrix} = \begin{bmatrix} \mathbf{B_1} \\ \mathbf{B_2 A_1} \\ \vdots \\ \mathbf{B_J A_{J-1}} \times \ldots \mathbf{A_2 A_1} \\ \mathbf{A_J A_{J-1}} \times \ldots \mathbf{A_2 A_1} \end{bmatrix}. \tag{2.17}$$

Here $\mathbf{B}$ and $\mathbf{A}$ contain the wavelet and scaling filters respectively.

$$\mathbf{B} = \begin{bmatrix} h_1 & h_0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & h_1 & h_0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & h_1 & h_0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 \\ & & & & & & \vdots & & & & & & \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & h_1 & h_0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & h_1 & h_0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & h_1 & h_0 \end{bmatrix}, \tag{2.18}$$

$$\mathbf{A} = \begin{bmatrix} g_1 & g_0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & g_1 & g_0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & g_1 & g_0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 \\ & & & & & & \vdots & & & & & & \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & g_1 & g_0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & g_1 & g_0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & g_1 & g_0 \end{bmatrix}. \tag{2.19}$$

Using the wavelet transform it is then possible, according to Jensen (1999), to estimate the long run memory parameter d as shown in equations 2.20 and 2.21

$$ln\, R(j) = ln\, \sigma^2 - d\, ln 2^{2j}, \tag{2.20}$$

where

$$\bar{R}(j) = \frac{1}{2^j} \sum_{k=0}^{2^j-1} w_{j,k}^2. \tag{2.21}$$

Equation 2.21 represents an estimate of the wavelet coefficient's population variance, while equation 2.20 can be estimated using OLS, where $ln\, \sigma^2$ is a constant. The long run memory parameter is then determined by the slope of equation 2.20. Depending on d, the slope will therefore differ, (see Figure 2).

Lastly, it should be noted that a variety of wavelets are available. In general, there is not one "right" wavelet to chose as they all have certain advantages/ disadvantages. The two most widely used wavelets are the Haar and Daubechie (4) wavelet (the former is used in this paper) and the following section will provide a brief comparison between the two.

When considering the D(4) wavelet filter, one can see that it is symmetric but also that it encounters problems when filtering near the boundary. The beginning and end of a series is referred to as boundary and since the DWT of the D(4) wavelet uses circular filtering, it would ideally require observations that are outside the sample (more specifically, it requires observations before/ after the sample begins/ ends). In order to address this issue, one has to make the so-called "circularity assumption". For a sample of $T = 16$, this implies that $Y_{-1} \equiv Y_{15}$ and $Y_{-2} \equiv Y_{14}$. This assumption is questionable[8], especially with respect to the quite large discontinuity between Y_{t-1} and Y_0 in the employed dataset (see Figures 5, 6 and 7). The circularity assumption does *not* have to be made when using the Haar wavelet, as it requires no boundary conditions.

Furthermore, *all* (linear) deterministic components are captured by the zero frequency component as $\mathbf{\Gamma_J}$ (see equation 2.17) is not included in estimating the long run memory parameter (see equations 2.20 and 2.21). This implies that one does not need to be concerned about

[8]For a more detailed discussion, see Percival and Walden (2000).

linear stochastic or deterministic trends. This property is shared by the Haar and D(4) wavelet[9].

On the contrary, the Haar wavelet isolates the correct frequencies less well than the D(4) wavelet. This can be seen when considering the transfer function for scale j[10],

$$H(f) = \sum_{l=0}^{L-1} h_l^j e^{-i2\pi fl}, \tag{2.22}$$

and its square, the square gain function

$$\mathcal{H}(f) = |H(f)|^2 . \tag{2.23}$$

Plotting the square gain function for the Haar and D(4) wavelet shows the relatively poor performance of the former wavelet.

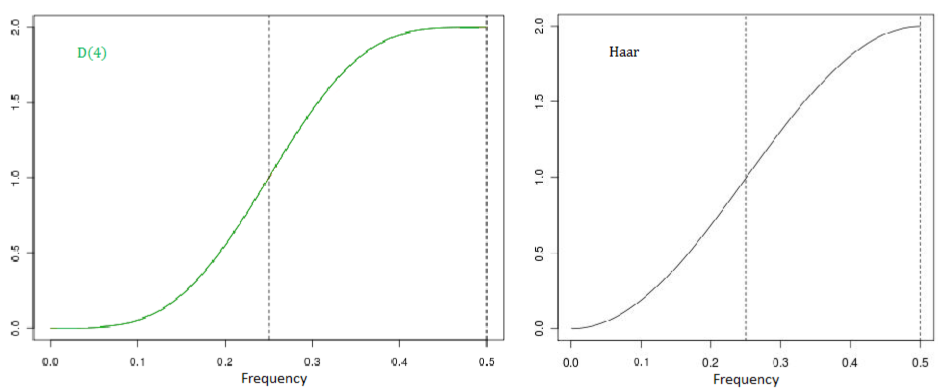

Figure 4: Squared Gain Functions, High Pass Filter H1, Haar & D(4)

In conclusion, one can see that all wavelet filters have advantages/ disadvantages. One advantage of the Haar wavelet is that it employs less boundary conditions than a D(4) wavelet (in fact, no boundary assumptions have to be made) and that one does not need to be concerned about deterministic components. On the contrary, the Haar wavelet isolates frequencies less well than the D(4) wavelet. Taking into account the advantages/ disadvantages and the used sample (see Section 3), this paper utilizes the Haar wavelet.

[9]In addition to linear trends, the D(4) wavelet also automatically removes quadratic trends.

[10]$l = 0, ..., L-1$ represents the length (width) of the wavelet filter. The Haar and D(4) wavelet filters are of width $L = 2$ and $L = 4$ respectively.

3 Empirics

In accordance with the theory, it is now possible to empirically test for unit roots and determine the long run memory parameter (section 3.2). Before doing that, section 3.1 will outline the used dataset.

3.1 Data

The used data has been split into *three* separate groups. Each group has a base currency (Group 1: USD, Group 2: GBP, Group 3: EUR), where the real exchange rate is then plotted against a basket of other currencies (SEK, JPY, NOK), as well as the remaining base currency (the EUR is a separate base currency and only plotted against the USD and GBP).

Each real exchange rate pair is plotted on a monthly basis from January 1990 to April 2011. This yields a total amount of 256 observations for each pair. Equation 1.1 shows how the real exchange rate is calculated.

The procedure is then best understood when looking at Figure 5.

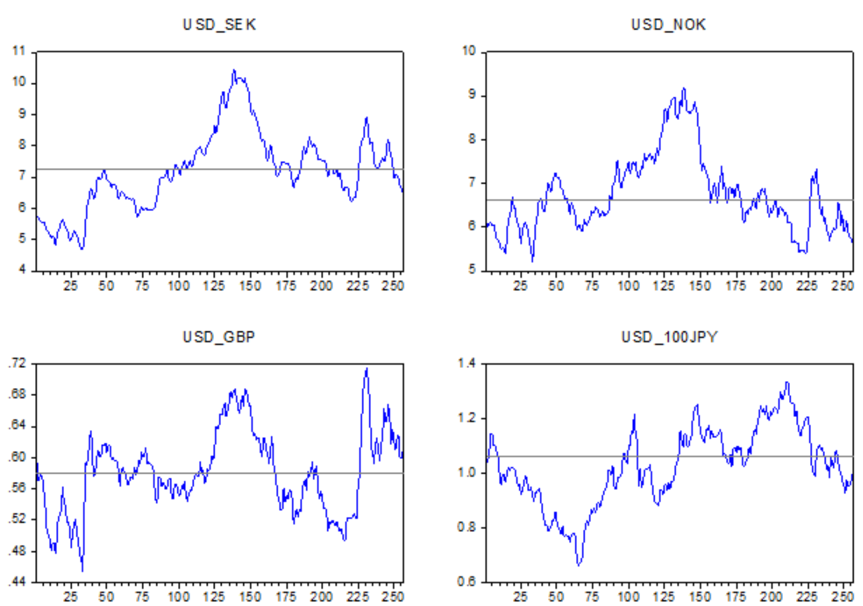

Figure 5: Group 1: USD vs Basket of currencies, January 1990 - April 2011

Considering the first pair, USD/SEK, RE is obtained by multiplying the nominal exchange rate with the ratio of domestic, P, and foreign, P^*, price levels respectively. The consumer price index (CPI) has been chosen as a proxy for the price levels. This procedure is then iterated for the remaining currency pairs.

Figure 5 shows an appreciation of the USD against the SEK/NOK/GBP/JPY at observation 125 (approximately), corresponding to May 2000. One possible explanation for this fact is that when the dot-com bubble burst in early 2000, causing the NASDAQ Composite to lose 78% of its value (Investopedia, 2012), investors were looking for safe investments. The USD is usually seen as such and has historically always been regarded as a “safe haven” (Engel & Hamilton, 1990). A similar observation can be made about the current financial crisis where the dollar appreciated against the SEK/NOK/GBP at observation 225, corresponding to September 2008.

In order to obtain the remaining groups, the base currency is changed but the procedure remains unchanged. For group two, this looks as follows.

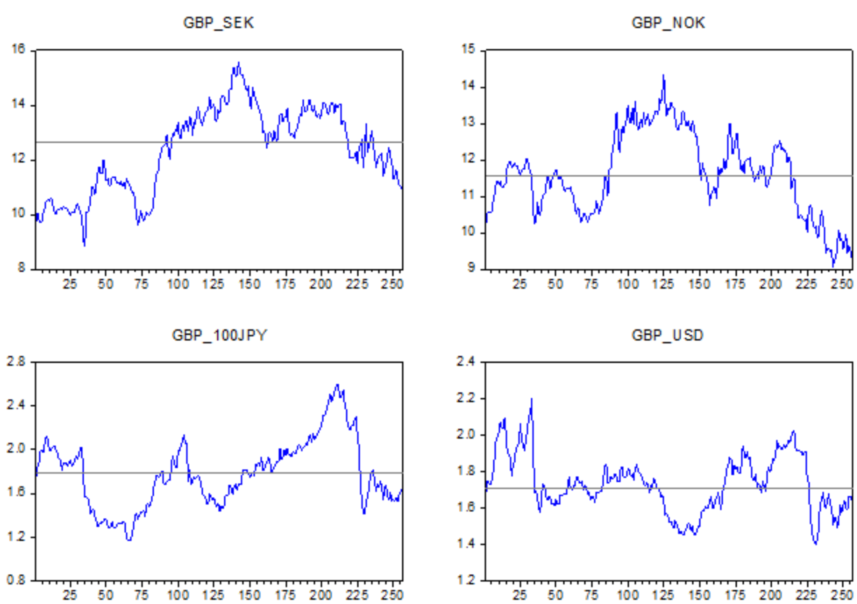

Figure 6: Group 2: GBP vs Basket of currencies, January 1990 - April 2011

Figure 6 shows that the GBP was and is significantly affected by the current financial crisis, as it depreciated against all four shown currencies. Given that Britain's financial industry accounted for nearly 10% of its GDP in 2008 (Bank of England, 2011), it is little surprising that it is severely affected by a financial crisis. This is commonly attributed to the lack of industry/ manufacturing which has been referred to as the greatest de-industrialisation of any major nation (The Guardian, 2011). Given this, a recent study by the Department for Business Innovation & Skills (2010, pg. 9) concluded that "evidence suggests that the UK lags behind its main competitors such as the United States and Germany".

Lastly, group three is displayed in Figure 7.

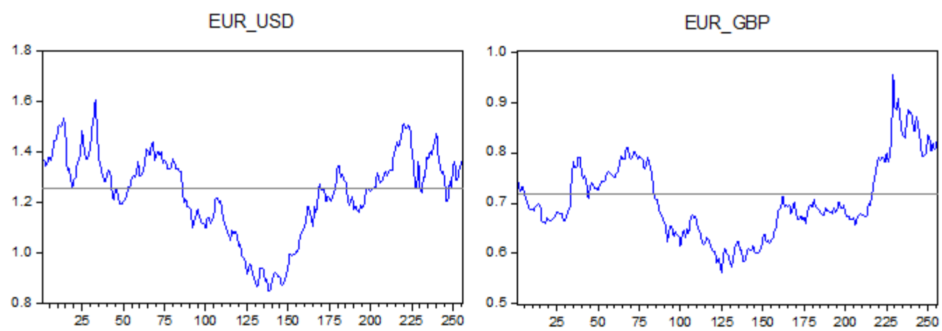

Figure 7: Group 3: EUR vs USD and GBP, January 1990 - April 2011

Figure 7 shows that after its introduction in 1999, investors were not confident with respect to the long term prospects of the EUR. As a result it depreciated against the USD and GBP. However, after this initial uncertainty the EUR appreciated. In accordance with Figure 6, the EUR appreciated sharply against the GBP as a result of the financial crisis.

All the data has been obtained using Thompson Financial Datastream 5.0, *apart* from the Swedish/ American CPI and the USD/SEK nominal exchange rate. The former can be obtained online through Statistics Sweden (2012)/ U.S. Bureau of Labor Statistics (2012) and the average nominal monthly exchange rate, S, has been accessed through `www.oanda.com` (2012).

It is important to note that all countries use different base years for their CPI indices. In order to be consistent, all CPI values have been transformed using *January 2005* (=100) as a base. This has been done in the following way:

$$New\, index\, value = \frac{old\, index\, value}{index\, value\, of\, new\, base} \times 100. \tag{3.1}$$

Since the dataset consists of monthly observations starting from January 1990, it is not possible to use the EUR as a base currency (apart from the ones in Group 3, see Figure 7) as it was officially introduced on 1 January 1999 and no other data is available. Prior to this date, all exchange rate calculations have been done using the European Currency Unit (ECU) as a base currency[11].

3.2 Results

For the first group, the results are presented in Table 1 below.

	ADF (1)	ADF (2)	PP (1)	PP (2)	KPSS (1)	KPSS (2)	I(d)	LR, d
USD/SEK	-1.81	-10.60***	-1.64	-10.65***	0.32***	0.14	1	0.95
USD/NOK	-1.76	-15.03***	-1.83	-15.06***	0.37***	0.12	1	0.91
USD/JPY	-2.22	-15.69***	-2.41	-15.69***	0.17**	0.08	1	0.93
USD/GBP	-2.64	-14.00***	-3.04	-14.02***	0.14*	0.04	1	0.68

Table 1: Results, Group 1

As mentioned previously, one has to specify whether or not to include an intercept and/ or linear trend when doing the ADF, PP and KPSS test. The (1) for each test indicates that it was done including a constant and linear trend, while (2) represents testing the first difference including a constant. The long run memory parameter, d, is presented in the last column.

[11] On 1 January 1999, the EUR replaced the ECU at a value of 1€ = 1ECU.

Table 1 shows that, based on all three unit root tests, one would conclude that the real exchange rate for all four currency pairs is integrated of order one. The ADF and PP return insignificant t-values for all currency pairs and thus do not reject the null hypothesis of a unit root. Furthermore, when the first difference is taken into account, all t-values are significant at $\alpha = 1\%$. The power of the KPSS test to detect roots close to unity seems to be slightly higher as it rejects the null hypothesis (stationarity) for USD/JPY at $\alpha = 5\%$ and for USD/GBP at $\alpha = 10\%$. Especially the last value is noteworthy as it coincides with the lowest long run memory parameter of 0.68.

When looking at the long run memory parameter, one can see that it is indeed very close to (with the exception of the real USD/GBP exchange rate) but nevertheless strictly less than unity. The results indicate that PPP is fractionally integrated and subject to long swings. Given this, one can therefore argue that it does hold in the long run as it is mean reverting.

Similar results can be observed when looking at the second group.

	ADF (1)	ADF (2)	PP (1)	PP (2)	KPSS (1)	KPSS (2)	I(d)	LR, d
GBP/SEK	-1.45	-13.76***	-1.18	-17.32***	0.39***	0.23	1	0.92
GBP/NOK	-1.96	-16.59***	-1.81	-16.69***	0.35***	0.28	1	0.92
GBP/JPY	-1.70	-14.32***	-2.04	-14.39***	0.14*	0.07	1	0.78
GBP/USD	-2.74	-14.09***	-3.11	-14.07***	0.15**	0.04	1	0.63

Table 2: Results, Group 2

Again, including an intercept and trend, the ADF and PP return insignificant t-values for all currency pairs and thus do not reject the null hypothesis of a unit root. Furthermore, all first differences are considered to be stationary, indicating an integration of order one, $I(1)$. Similar conclusions are drawn when using the KPSS test but, as in the results presented in Table 1, the KPSS test seems to be slightly more powerful with respect to detecting roots close to unity, rejecting the null hypothesis for GBP/JPY at $\alpha = 10\%$ and for GBP/USD at $\alpha = 5\%$.

Despite changing the base currency, the long run memory parameter indicates a mean reverting fractionally integrated time series. However, it should be noted that an outlier has been removed for the real GBP/NOK exchange rate (including it yields a d-value of 0.75).

Lastly, the results for group three are shown in Table 3.

	ADF (1)	ADF (2)	PP (1)	PP (2)	KPSS (1)	KPSS (2)	I(d)	LR, d
EUR/USD	-1.71	-14.31***	-1.82	-14.25***	0.39***	0.11	1	0.92
EUR/GBP	-1.65	-15.24***	-1.64	15.24***	0.34***	0.18	1	0.97

Table 3: Results, Group 3

Using the EUR as a base currency yields similar results to the ones obtained before. PPP appears to be fractionally integrated with an integration order close to unity. However, what can be seen is the exceptionally long EUR/GBP memory of 0.97. A possible explanation for this fact is that prior to the establishment of the EUR on 1 January 1999, the GBP was part of the ECU basket of goods and the long run memory parameter therefore traces parts of its own past. Furthermore, the KPSS test provides the same conclusion with respect to stationarity as the ADF and PP for all currency pairs and levels of significance.

"Now let me explain why this makes intuitive sense." ~ *Larry Wassermann*

4 Conclusion

Based on the results one can see that PPP holds in the long run, as the real exchange rate is slowly mean reverting and subject to long swings. All real exchange rate pairs are integrated of order $1/2 \leq d < 1$, which cannot be represented by an AR(p) model. Consequently, it is necessary to model RE using a FIP. Coming back to the article "If exchange rates are random walks, then almost everything we say about monetary policy is wrong" by Alvarez, Atkeson and Kehoe (2007), this also implies that exchange rates are not random walks and that we therefore can apply standard monetary models. It may take a long time for monetary policy to have an effect but the important fact is that it has an effect at all. Furthermore, the results imply that one should technically be able to, through arbitrage, exploit this mean reverting behavior to make profit. However, this is practically not applicable as the swings are very long and it takes a very long time for the time series to come back to its mean. In the case of Figure 1 (USD/SEK), one can observe a swing of more than 100 months (~ 8.5 years).

In conclusion, one can see that the indiscriminate use of standard unit root tests, like the ADF, is not always unproblematic, as they have low power with respect to accurately determining unit roots that are close to unity (though it should be noted the the KPSS test seems to have a little more power). In the case of PPP, this leads to falsely not rejecting the null hypothesis and therefore also drawing the wrong conclusions. Based on this, one can see that more advanced techniques are required. These can also handle time series that are fractionally integrated, of order $1/2 \leq d < 1$, and suggest that PPP does *not* follow a random walk process and therefore holds in the long run. This has important ramifications with respect to e.g. the effectiveness of monetary policy.

5 References

Abramovich, F., Bailey, T. C., Sapatinas, T. (2000). Wavelet Analysis and Its Statistical Applications. *Journal of the Royal Statistical Society. Series D (The Statistician), 49*(1), 1-29.

Abuaf, N., Jorion, P. (1990). Purchasing Power Parity in the Long Run. *The Journal of Finance, 45*(1), 157-174.

Adler, M., Lehmann, B. (1983). Deviations from Purchasing Power Parity in the Long Run. *The Journal of Finance, 38*(5), 1471-1487.

Alvarez, F., Atkeson, A., Kehoe, P. J. (2007). If Exchange Rates are Random Walks, Then Almost Everything We Say About Monetary Policy is Wrong. *American Economic Review, 97*(2), 339-345.

Andersson, F. N. G. (2012). *Exchange Rates Dynamics Revisited: A Panel Data Test of the Fractional Integration Order.* Under Review.

Andersson, F. N. G. (2008). *Wavelet Analysis of Economic Time Series.* Lund Economic Studies, 149.

Artus, J. R. (1978). Methods of Assessing the Long Run Equilibrium Value of an Exchange Rate. *Journal of International Economics, 8*, 277-299.

Bank of England. (2011). *Measuring financial sector output and its contribution to UK GDP.* Retrieved 2012-05-21, from: `http://www.bankofengland.co.uk/publications/Documents/quarterlybulletin/qb110304.pdf`

Battle, G. (1997). Heisenberg Inequalities for Wavelet States. *Applied and Computational Harmonic Analysis, 4*, 119-146.

Belaire-Franch, J., Opong, K. K. (2005). Some Evidence of Random Walk Behaviour of EUR Exchange Rates using Ranks and Signs. *Journal of Banking and Finance, 29*, 1631-1643.

Cassel, G. (1918). Abnormal Deviations in International Exchanges. *The Economic Journal, 28*(112), 413-415.

Department for Business Innovation & Skills. (2010). *Manufacturing in the UK: An economic analysis of the sector*. Retrieved 2012-05-21, from:
http://www.bis.gov.uk/assets/biscore/business-sectors/docs/m/10-1333-manufacturing-in-the-UK-an-economic-analysis-of-the-sector

Dornbusch, R. (1980). Exchange Rate Economics: Where Do We Stand?. *Brookings Papers, 1*, 143-206.

Engel, C., Hamilton, J. D. (1990). Long Swings in the Dollar: Are They in the Data and Do Markets Know It?. *The American Economic Review, 80*(4), 689-713.

Frenkel, J. A. (1981). Flexible Exchange Rates, Prices and the Role of News: Lessons from the 1970's. *Journal of Political Economy, 89*, 665-705.

Granger, C. W. J., Josyeux, R. (1980). An Introduction To Long-Memory Time Series Models And Fractional Differencing. *Journal of Time Series Analysis, 1*(1), 15-29.

Gujarati, D. N. (2004). *Basic Econometrics*. McGraw-Hill.

Hobij, B., Franses, P. H., Ooms, M. (2004). Generalizations of the KPSS-test for stationarity. *Statistica Neerlandica, 58*(4), 483-502.

Investopedia. (2012). *Market Crashes: The Dotcom Crash*. Retrieved 2012-05-21, from:
http://www.investopedia.com/features/crashes/crashes8.asp#axzz1vVFOuCQl

Jensen, M. J. (1999). Using Wavelets to Obtain a Consistent Ordinary Least Squares Estimator of the Long-memory Parameter. *Journal of Forecasting, 18*(1), 17-32.

Lopez, C., Murray, C. J., Papell, D. H. (2004). State of the Art Unit Root Tests and Purchasing Power Parity. *Journal of Money, Credit, and Banking, 37*(2), 361-369.

Mallat, S. G. (1989). Multiresolution Approximations and Wavelet Orthonormal Bases of $L^2(R)$. *IEEE Transactions on Pattern and Analysis and Machine Intelligence, 11*, 674-693.

Oanda. (2012). Average Exchange Rates. Retrieved 2012-05-08, from:
http://www.oanda.com/lang/de/currency/average?amount=1&start_month=1&start_year=1990&end_month=4&end_year=2012&base=USD&avg_type=Month&Submit=1&exchange=SEK&interbank=0&format=HTML

Percival, D. B., Walden, A. T. (2000). *Wavelet Methods for Time Series Analysis.* Cambridge: Cambridge University Press.

Rogoff, K. (1996). The Purchasing Power Parity Puzzle. *Journal of Economic Literature, 34*(2), 647-668.

Statistics Sweden. (2012). *Consumer Price Index (CPI).* Retrieved 2012-05-08, from: http://www.scb.se/Pages/TableAndChart____272152.aspx

Taylor, A. M. (2002). A Century Of Purchasing Power Parity. *The Review of Economics and Statistics, 84*(1), 139-150.

The Guardian. (2011). *Why doesn't Britain make things any more?.* Retrieved 2012-05-21, from: http://www.guardian.co.uk/business/2011/nov/16/why-britain-doesnt-make-things-manufacturing

Torrence, C., Compo, G. P. (1998). A Practical Guide to Wavelet Analysis. *Bulletin of the American Meteorological Society, 79*(1), 61-78.

University of Washington. (2012). *Unit Root Tests.* Retrieved 2012-05-19, from: http://faculty.washington.edu/ezivot/econ584/notes/unitroot.pdf

U.S. Bureau of Labor Statistics. (2012). *Consumer Price Index.* Retrieved 2012-05-08, from: ftp://ftp.bls.gov/pub/special.requests/cpi/cpiai.txt

Verbeek, M. (2008). *A Guide To Modern Econometrics.* Chichester, England: John Wiley & Sons Ltd.