

Andersson, Fredrik N. G.

**Working Paper**

## Band Spectrum Regressions using Wavelet Analysis

Working Paper, No. 2011:22

**Provided in Cooperation with:**

Department of Economics, School of Economics and Management, Lund University

*Suggested Citation:* Andersson, Fredrik N. G. (2011) : Band Spectrum Regressions using Wavelet Analysis, Working Paper, No. 2011:22, Lund University, School of Economics and Management, Department of Economics, Lund

This Version is available at:

<https://hdl.handle.net/10419/260012>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# Band Spectrum Regressions using Wavelet Analysis

Fredrik N. G. Andersson\*  
(Lund University)

## Abstract

In economics it is common to distinguish between different time horizons (i.e. short run, medium run, and long run). Engle (1974) proposed combining the discrete Fourier transform with a band spectrum regression to estimate models that separates between different time horizons. In this paper we discuss possibilities and challenges using the maximal overlap discrete wavelet transform instead of the Fourier transform when estimating band spectrum regressions.

**Keywords:** band spectrum regression, wavelet transform, frequency domain, economic modeling

**Jel-Codes:** C14, C32, C51

---

\* **Corresponding author:** email: [ngf.andersson@nek.lu.se](mailto:ngf.andersson@nek.lu.se)

Fredrik Andersson gratefully acknowledges funding from the Jan Wallander and Tom Hedelius foundation (project number P2009-0189:1) and the Swedish Research Council (421-2009-2663).

## 1. Introduction

In economics in general and in macroeconomics in particular, it is common to distinguish between different time horizons (see e.g. Friedman, 1957; King and Watson, 1996; Ramsey and Lampart, 1998). The neoclassical synthesis, for example, combines (New-) Keynesian theories with (dynamic-) neoclassical theories to distinguish between the causes of short term volatility and the causes of long term growth (Samuelson, 1955; Goodfriend and King, 1997; Woodford, 2003). Similarly, Friedman and Kuznets (1954) and Englund et. al. (1992) suggests separating between three time horizons: short run shocks, medium run business cycle fluctuations, and the long run equilibrium.

To estimate models where the relationship between variables depend on the time horizon, Engle (1974), Phillips (1991) and Corbae, Ouliaris and Phillips (2002), among others, have proposed a band spectrum regression. The band spectrum regression is a two step estimator. First, all variables are transformed to the frequency domain where the respective time horizons are easily identified. Second, parameter estimates for each time horizon are obtained by regressing on a sub-set of frequencies rather than the entire frequency band.

Each time horizon is in the frequency domain represented by a unique set of frequencies. The long run, for example, is represented by the low frequencies close to the zero frequency and the short run is represented by the high frequencies close to frequency one half. By identifying the frequencies representing a given time horizon, one can estimate a model for that time horizon by regressing on just those frequencies that represent that time horizon.

Common for most band spectrum estimators is that they employ the Fourier transform to transform the time series from the time domain to the frequency domain. An alternative transform is a wavelet transform. Unlike the Fourier transform (which only contains frequency resolution) the wavelet transform combines time and frequency resolution, which

makes it possible to obtain an unbiased and consistent estimate of a time series spectral density for time series that contain structural breaks, outliers, and other non-recurring events (Percival and Walden, 2006). Without time resolution (e.g. Fourier transform), it is necessary to pre-whiten the data to remove outliers, and to split the sample into smaller samples to control for structural breaks. Because economic time series often include outliers and structural breaks the wavelet transform is well suited for economic analyses.

For discrete data, two wavelet transforms are available: the discrete wavelet transform, and the maximal overlap discrete wavelet transform. The discrete wavelet transform (DWT) imposes a restriction on the sample size: only time series with  $T=2^J$  observations, where  $J$  is an integer, can be transformed. The maximal overlap discrete transform (MODWT) does not impose such restrictions on the sample size. The MODWT also has better small sample properties than the DWT. Most economic studies therefore use the MODWT, but there are cases when the DWT is used as well (see e.g. Andersson, 2008). For a more detailed discussion on wavelet analysis see Percival and Walden (2006), Crowley (2007) or Andersson (2008).

In this paper we demonstrate how the MODWT can be used to estimate band spectrum regressions. We also discuss the econometric consequences for not distinguishing between different time horizons in empirical analyses. These effects are also illustrated in a simulation study.

The rest of the paper is organized as follows: Section 2 and Section 3 discuss the band spectrum regressions and the MODWT, Section 4 contains simulations illustrating the band spectrum regression, and Section 5 concludes the paper.

## **2. Modeling Time Horizons**

Following Samuelson (1955) and Woodford (2003), consider a model with two time horizons: the short run and the long run,

$$y_t = x_{LR,t}\beta_{LR} + x_{SR,t}\beta_{SR} + \varepsilon_t, \quad (1)$$

where  $t=1, \dots, T$  is time,  $x_{LR,t}$  is a vector with the long run component of the explanatory variables,  $x_{SR,t}$  is a vector with the short run component of the explanatory variables and  $\varepsilon_t$  is a random disturbance. Some variables can have both a long run effect and a short run effect, but it is also possible that a given variable either have just a long run effect or just a short run effect. Moreover, if a variable have both a long run and a short run effect, the long run and the short run effects can either be the same or different. In this example, let the long run be represented by the frequencies 0 to  $a$  and the short run by the frequencies  $a$  to 0.5, where  $0 \leq a \leq 0.5$ , in the frequency domain.

Before we consider how to estimate the parameters in (1), consider the case when we (wrongly) estimate the model,

$$y_t = x_t\beta + \eta_t. \quad (2)$$

where  $x_t = x_{LR,t} + x_{SR,t}$ , and we in effect assume that there is only one time horizon (i.e.  $\beta_{LR} = \beta_{SR}$ ). In this case the OLS estimates equals,

$$\hat{\beta} = (\mathbf{x}'\mathbf{x})^{-1}\mathbf{x}'\mathbf{y} = (\mathbf{x}'\mathbf{x})^{-1}\mathbf{x}'(\mathbf{x}_{LR}\beta_{LR} + \mathbf{x}_{SR}\beta_{SR} + \boldsymbol{\varepsilon}). \quad (3)$$

All frequencies are orthogonal to each other whereby the expected value of the parameter is equal to,

$$\begin{aligned} E[\hat{\beta}] &= E[(\mathbf{x}'\mathbf{x})^{-1}\mathbf{x}'(\mathbf{x}_{LR}\beta_{LR} + \mathbf{x}_{SR}\beta_{SR} + \boldsymbol{\varepsilon})] = \\ &= (\mathbf{x}'\mathbf{x})^{-1}\mathbf{x}'_{LR}\mathbf{x}_{LR}\beta_{LR} + (\mathbf{x}'\mathbf{x})^{-1}\mathbf{x}'_{SR}\mathbf{x}_{SR}\beta_{SR}. \end{aligned} \quad (4)$$

As can be seen from (4), the estimated parameter vector is a weighted average of the long run and the short run parameters. The size of the weights  $(\mathbf{x}'\mathbf{x})^{-1}\mathbf{x}'_{LR}\mathbf{x}_{LR}$  and  $(\mathbf{x}'\mathbf{x})^{-1}\mathbf{x}'_{SR}\mathbf{x}_{SR}$ , moreover, depend on the statistical properties of  $\mathbf{x}$ . For example, in a bivariate model with only one explanatory variable the weights in (4) are equal to the share of the variance in  $x_t$  explained by the long run and the short run respectively. If the long run information dominates  $x_t$ ,  $\hat{\beta}$  is close to  $\beta_{LR}$ . If the short run information dominates  $x_t$ ,  $\hat{\beta}$  is close to  $\beta_{SR}$ .

A second consequence for estimating (2) instead of the true model (1) is that the variance of the disturbances in (2) is greater than the variance of the disturbances in (1). This can be seen by considering the difference between  $\varepsilon_t$  from (1) and  $\eta_t$  from (2)

$$\eta_t - \varepsilon_t = x_{LR,t}(\beta_{LR} - \beta) + x_{SR,t}(\beta_{SR} - \beta). \quad (5)$$

Rewriting (5) we get,

$$\eta_t = \varepsilon_t + x_{LR,t}(\beta_{LR} - \beta) + x_{SR,t}(\beta_{SR} - \beta). \quad (6)$$

As can be seen by (6), the disturbances in (2) thus include both the true disturbances ( $\varepsilon_t$ ) and a miss-specification component:  $x_{LR,t}(\beta_{LR} - \beta) + x_{SR,t}(\beta_{SR} - \beta)$ . Consequently,

$$var(\eta_t) = \sigma^2 + var(x_{LR,t})(\beta_{LR} - \beta)^2 + var(x_{SR,t})(\beta_{SR} - \beta)^2. \quad (7)$$

To correctly estimate the model we must separate between long and short run effects.

### *2.1 Band Spectrum Regressions Using the MODWT*

To estimate (1) all variables have to be decomposed into their short run component and a long run component. In the literature several different techniques are employed to make the separation of the time series into time horizons: for example a Fourier transform, a Hodrick-Prescott filter (HP-filter) or 5-year averages. Problems with the Fourier transform have already been discussed. The HP-filter separates short and long run variations in the variables, but suffers from two problems. First, it only separates between two time horizons, and secondly, it is somewhat of a black box concerning which frequencies are defined as short-run and long run. In the growth literature 5-year averages are commonly used (Islam, 2003), but simple averages may induce cycles and artifacts whereby the time horizon decomposition can be spurious (Slutsky, 1938; Percival and Walden, 2006).

Considering that the MODWT solves the problems other techniques suffers from (at least asymptotically), let us consider the case where we estimate a band spectrum regression using the MODWT. Let  $\mathbf{W}$  be the MODWT matrix (see Percival and Walden, 2006, for a detailed

description of how to construct the transform matrix). The transform coefficients are obtained by multiplying the transform matrix with the vector containing the variable that is to be transformed,

$$\boldsymbol{\omega}_y = \mathbf{W}y, \quad (8)$$

where  $\boldsymbol{\omega}_y$  are the transform coefficients. The transform matrix and the transform coefficients can be partitioned as,

$$\boldsymbol{\omega}_y = \begin{bmatrix} \boldsymbol{\omega}_{1y} \\ \boldsymbol{\omega}_{2y} \\ \vdots \\ \boldsymbol{\omega}_{Jy} \\ \mathbf{v}_{Jy} \end{bmatrix} = \begin{bmatrix} \mathbf{W}_1 \\ \mathbf{W}_2 \\ \vdots \\ \mathbf{W}_J \\ \mathbf{V}_J \end{bmatrix} \times \mathbf{y}, \quad (9)$$

where the transform coefficients  $\boldsymbol{\omega}_{jy}$ ,  $j=1, \dots, J$ , represent the frequency band  $\frac{1}{2^{j+1}}$  to  $\frac{1}{2^j}$ , and the transform coefficients  $\mathbf{v}_{Jy}$  represents the frequency 0 to  $\frac{1}{2^{J+1}}$ .

Assume  $a = \frac{1}{2^{J+1}}$ , such that the long run is defined by the frequencies 0 to  $\frac{1}{2^{J+1}}$  and the short run by the frequencies  $\frac{1}{2^{J+1}}$  to  $\frac{1}{2}$ . In this case  $\mathbf{V}_j \mathbf{x} = \mathbf{V}_j \mathbf{x}_{LR}$  and  $\mathbf{V}_j \mathbf{x}_{SR} = \mathbf{O}$ , where  $\mathbf{O}$  is a matrix of zeroes.

Assume we are interested in the long run relationship between  $y_t$  and  $x_t$ . By applying the transform matrix  $\mathbf{V}_j$  to  $\mathbf{y}$  and  $\mathbf{x}$  and estimating the following model with an OLS,

$$\mathbf{V}_j \mathbf{y} = \mathbf{V}_j \mathbf{x} \boldsymbol{\beta} + \mathbf{V}_j \boldsymbol{\varepsilon}. \quad (10)$$

we get the OLS parameter estimate,

$$\begin{aligned} \hat{\boldsymbol{\beta}} &= (\mathbf{x}' \mathbf{V}_j' \mathbf{V}_j \mathbf{x})^{-1} \mathbf{x}' \mathbf{V}_j' \mathbf{V}_j \mathbf{y} = (\mathbf{x}' \mathbf{V}_j' \mathbf{V}_j \mathbf{x})^{-1} \mathbf{x}' \mathbf{V}_j' \mathbf{V}_j (\mathbf{x}_{LR} \boldsymbol{\beta}_{LR} + \mathbf{x}_{SR} \boldsymbol{\beta}_{SR} + \boldsymbol{\varepsilon}) = \\ &(\mathbf{x}' \mathbf{V}_j' \mathbf{V}_j \mathbf{x})^{-1} \mathbf{x}' \mathbf{V}_j' \mathbf{V}_j \mathbf{x}_{LR} \boldsymbol{\beta}_{LR} + (\mathbf{x}' \mathbf{V}_j' \mathbf{V}_j \mathbf{x})^{-1} \mathbf{x}' \mathbf{V}_j' \mathbf{V}_j \mathbf{x}_{SR} \boldsymbol{\beta}_{SR} + (\mathbf{x}' \mathbf{V}_j' \mathbf{V}_j \mathbf{x})^{-1} \mathbf{x}' \mathbf{V}_j' \mathbf{V}_j \boldsymbol{\varepsilon} = \\ &(\mathbf{x}' \mathbf{V}_j' \mathbf{V}_j \mathbf{x})^{-1} \mathbf{x}' \mathbf{V}_j' \mathbf{V}_j \mathbf{x}_{LR} \boldsymbol{\beta}_{LR} + (\mathbf{x}' \mathbf{V}_j' \mathbf{V}_j \mathbf{x})^{-1} \mathbf{x}' \mathbf{V}_j' \mathbf{V}_j \boldsymbol{\varepsilon}. \end{aligned} \quad (11)$$

The expected value of  $\hat{\boldsymbol{\beta}}$  is equal to,

$$E[\hat{\boldsymbol{\beta}}] = (\mathbf{x}' \mathbf{V}_j' \mathbf{V}_j \mathbf{x})^{-1} \mathbf{x}' \mathbf{V}_j' \mathbf{V}_j \mathbf{x}_{LR} \boldsymbol{\beta}_{LR} + (\mathbf{x}' \mathbf{V}_j' \mathbf{V}_j \mathbf{x})^{-1} \mathbf{x}' \mathbf{V}_j' \mathbf{V}_j E[\boldsymbol{\varepsilon}] = \boldsymbol{\beta}_{LR}. \quad (12)$$

The expected value of  $\hat{\beta}$  is hence equal to  $\beta_{LR}$ . The covariance matrix of the parameter estimate is given by,

$$E [(\hat{\beta} - \beta_{LR})(\hat{\beta} - \beta_{LR})'] = E \left[ (\mathbf{x}'\mathbf{V}_j'\mathbf{V}_j\mathbf{x})^{-1} \mathbf{x}'\mathbf{V}_j'\mathbf{V}_j\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'\mathbf{V}_j'\mathbf{V}_j\mathbf{x}(\mathbf{x}'\mathbf{V}_j'\mathbf{V}_j\mathbf{x})^{-1} \right], \quad (13)$$

given that  $\varepsilon_t$  is white noise with variance  $\sigma^2$  we get,

$$E [(\hat{\beta} - \beta_{LR})(\hat{\beta} - \beta_{LR})'] = \sigma^2 (\mathbf{x}'\mathbf{V}_j'\mathbf{V}_j\mathbf{x})^{-1} \mathbf{x}'\mathbf{V}_j'\mathbf{V}_j\mathbf{V}_j'\mathbf{V}_j\mathbf{x}(\mathbf{x}'\mathbf{V}_j'\mathbf{V}_j\mathbf{x})^{-1}. \quad (14)$$

Similarly we can replace  $\mathbf{V}_j$  with

$$\mathbf{w} = \begin{bmatrix} \mathbf{W}_1 \\ \mathbf{W}_2 \\ \vdots \\ \mathbf{W}_j \end{bmatrix}, \quad (15)$$

and estimate the short run parameters. Given estimate of the short run and the long run we obtain estimates of the error term,

$$\hat{\boldsymbol{\varepsilon}} = \mathbf{y} - \mathbf{V}_j'\mathbf{V}_j\mathbf{x}\hat{\beta}_{LR} - \mathbf{w}'\mathbf{w}\mathbf{x}\hat{\beta}_{SR}. \quad (16)$$

It should be noted that it is not the variance of the long term residual that enters into (14), but the variance of the error term in (1) which includes both the short and the long terms. To obtain an estimate of the residuals, i.e. (16), it is therefore necessary to estimate both a short run and a long run model.

Instead of separately estimating one model for the short run and one model for the long run, we can estimate short and long run parameters in one step by estimating,

$$\mathbf{y} = \mathbf{V}_j'\mathbf{V}_j\mathbf{x}\beta_{LR} + \mathbf{w}'\mathbf{w}\mathbf{x}\beta_{SR} + \boldsymbol{\varepsilon}. \quad (17)$$

### 3. Modeling with Only One Time Horizon

So far we have assumed that we are able to specify one model for each time horizon. At times economic theory may only suggest a model for some of the time horizons. For example, growth theory is oftentimes used to model long term economic growth, but is also commonly considered inappropriate for modeling short term growth variations.



Assume we cannot estimate a short run model; either because there is no suitable economic model or because we cannot obtain the necessary data (i.e. variables that have a short run effect on the dependent variable). In this case we cannot estimate of the disturbances in (1) from (16) and consequently we cannot obtain an estimate of  $var(\hat{\beta})$ .

However, using,

$$\hat{\boldsymbol{\varepsilon}}_{LR} = \mathbf{V}_j \mathbf{y} - \mathbf{V}_j \mathbf{x} \hat{\boldsymbol{\beta}}_{LR}, \quad (18)$$

we can, under the assumption of white noise, use  $\hat{\boldsymbol{\varepsilon}}_{t,LR}$  to obtain an estimate of  $var(\hat{\boldsymbol{\varepsilon}}_{t,SR})$ .

Naturally we would prefer to use  $\hat{\boldsymbol{\varepsilon}}_t$ , and using  $\hat{\boldsymbol{\varepsilon}}_{t,LR}$  is only a second option. However, we can use  $\hat{\boldsymbol{\varepsilon}}_{t,LR}$  to test if  $\boldsymbol{\varepsilon}_t$  is normally distributed and homoskedastic since it is still possible to estimate  $var(\hat{\boldsymbol{\varepsilon}}_t)$  under these circumstances. The only remaining problem is that we cannot use  $\hat{\boldsymbol{\varepsilon}}_{t,LR}$  to test for autocorrelation.

### 3.1 Variance Estimation

The variance of the short term disturbances can be estimated using  $\hat{\boldsymbol{\varepsilon}}_{t,LR}$ . Assuming white noise, we know that the spectral density function for white noise is given by,

$$s(f) = \sigma^2. \quad (19)$$

We can therefore write,

$$\sigma^2 = 2 \int_0^{1/2} s(f) df. \quad (20)$$

In other words, white noise is characterized by each frequency having the same variance.

Given our assumptions we know that  $\boldsymbol{\varepsilon}_{t,LR}$  is represented by the frequencies 0 to  $a$ , and  $\boldsymbol{\varepsilon}_{t,SR}$  by frequencies  $a$  to 0.5. Moreover,

$$\sigma^2 = var(\boldsymbol{\varepsilon}_{t,LR}) + var(\boldsymbol{\varepsilon}_{t,SR}), \quad (21)$$

Let,

$$Q = \frac{(0.5-a)}{(a-0)}, \quad (22)$$

be the ratio between the short run and long run frequencies. Using the assumption that  $\varepsilon$  is white noise, from (19) we obtain,

$$\begin{aligned}\sigma^2 &= \text{var}(\varepsilon_{t,LR}) + \text{var}(\varepsilon_{t,SR}) = \text{var}(\varepsilon_{t,LR}) + Q\text{var}(\varepsilon_{t,LR}) = (Q + 1)\text{var}(\varepsilon_{t,LR}) = \\ &= \frac{1}{2a}\text{var}(\varepsilon_{t,LR}).\end{aligned}\tag{23}$$

From (23) we obtain an expression we can use to estimate the residual variance without estimating a short run model.

### 3.2 Testing for Normality, Heteroskedacity and Autocorrelation

Testing for normality using the long run residuals is possible because the MODWT is performed by taking linear combinations of the variables. Thus, if  $\varepsilon_t$  is normally distributed then  $\mathbf{V}_j\boldsymbol{\varepsilon}$  is also normally distributed. Heteroskedacity can be tested using the long run residuals because the variance-covariance matrix of white noise is given by,

$$\begin{aligned}E[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'] &= \sigma^2\mathbf{I} = [\text{var}(\varepsilon_{t,LR}) + \text{var}(\varepsilon_{t,SR})]\mathbf{I} = E[\boldsymbol{\varepsilon}_{LR}\boldsymbol{\varepsilon}'_{LR}] + E[\boldsymbol{\varepsilon}_{SR}\boldsymbol{\varepsilon}'_{SR}] = \\ &= \sigma^2\mathbf{V}_j\mathbf{V}'_j + \sigma^2\mathbf{w}'\mathbf{w}.\end{aligned}\tag{24}$$

The diagonal elements of  $\mathbf{V}_j\mathbf{V}'_j$  and  $\mathbf{w}'\mathbf{w}$  are both equal to constants. We can therefore test for heteroskedacity using either just the long run or the short run residuals.

In theory we could also test for autocorrelation using the estimated long run residuals, but in practice this is not feasible. Consider first the case when  $\varepsilon_t$  is white noise. The autocovariance function for white noise is given by,

$$E[\varepsilon_t\varepsilon_{t-h}] = 2\sigma^2 \int_0^{1/2} e^{i2\pi fh} df,\tag{25}$$

whereby the autocovariance function for the long run residuals equals,

$$E[\varepsilon_{LR,t}\varepsilon_{LR,t-h}] = 2 \int_0^a s(f)e^{i2\pi fh} df,$$

As  $a \rightarrow 0$ , the correlation  $corr[\varepsilon_{LR,t}, \varepsilon_{LR,t-h}] \rightarrow 1$ . When we estimate an AR(1) model using just the long run disturbances, the estimated AR-parameter is going to be close to 1 even though  $\varepsilon_t$  is white noise.

Now, consider the case when  $\varepsilon$  follows AR(1) process,

$$\varepsilon_t = \rho\varepsilon_{t-1} + \varphi_t. \quad (26)$$

For this process the autocovariance function is given by,

$$E[\varepsilon_t \varepsilon_{t-h}] = 2 \int_0^{1/2} \frac{\sigma^2}{|1 - \rho e^{-i2\pi f}|^2} e^{i2\pi f h} df, \quad (27)$$

where  $\rho$  is the autoregressive parameter. Similar to the case with white noise, the correlation  $corr[\varepsilon_{LR,t}, \varepsilon_{LR,t-h}] \rightarrow 1$  when  $a \rightarrow 0$ . Therefore without the high frequency (short run) variations tests of autocorrelation become meaningless because we cannot distinguish between different autoregressive parameters.

Figure 1 shows the expected autoregressive parameter from an OLS for white noise and an AR(1) process with different autoregressive parameters when estimated using only a subset of low frequencies. The parameter  $a$  is shown along the  $x$ -axis (we use frequencies 0 to  $a$  to estimate the parameter), and the expected AR(1) parameter is shown along the  $y$ -axis. When  $a=0.5$  (i.e. we use all frequencies), all short and along run information is used to estimate the AR parameter. When  $a < 0.5$  some of the short run information is excluded and the closer  $a$  is to zero the more of the short run information has been excluded.

[FIGURE 1]

As can be seen in the Figure, when  $a=0.5$ ,  $\rho$  is equal to its true value. As  $a$  gets smaller,  $\rho$  increases and approaches 1 as  $a$  approaches 0. Although there is a difference in  $E[\hat{\rho}]$  for different values of  $a$ , these are too small to be distinguishable in practice.

#### 4. Simulations

To illustrate the band spectrum regression using a MODWT we perform two sets of simulations. In the first set of simulations we consider a model with two time horizons, and in the second set of simulations we consider a model with three time horizons.

#### 4.1 Two time horizons

In the first experiment the underlying model includes two time horizons: the short run and the long run. For simplicity we use a bivariate model with only one explanatory variable,

$$y_t = x_{LR,t}\beta_{LR} + x_{SR,t}\beta_{SR} + \varepsilon_t, \quad (28)$$

where  $\beta_{LR} = 1$  and  $\beta_{SR} = 0.2$ ,  $\varepsilon_t \sim N(0,1)$  and we assume that the long run is defined by the frequencies 0 to 1/32 and the short run by the frequencies 1/32 to 1/2. In the case of monthly data, this would imply that the short run lasts up to 32 months and the long run is 32 months and beyond. In the case of quarterly data, the short run would be lasting up to 8 years and the long run from 8 years and beyond.

In this experiment we simulate  $x_t$  as an autoregressive process,

$$x_t = \rho x_{t-1} + \theta_t, \quad (29)$$

where we consider three choices of  $\rho = \{0, 0.5, 1\}$ , and we let  $\theta_t \sim N(0,2)$ . The greater the autoregressive parameter the more the long run influences  $x$ . Due to these variations between which component of  $x$  (short run or long run) that dominates, depending on  $\rho$ , the expected value of  $\beta$  when estimating the model,

$$y_t = x_t\beta + \varepsilon_t, \quad (30)$$

varies as well. For example, the greater the value of  $\rho$  the closer  $\hat{\beta}$  is to  $\beta_{LR}$ .

As an illustration, we estimate both the miss-specified model (30) and a correctly specified model based on (28). Each experiment is carried out 1000 times and the average estimated parameter and the standard error are presented in Table 1.<sup>1</sup>

---

<sup>1</sup> As wavelet filter we use the Haar-wavelet.

As expected the parameter estimate from (30) is an average of the two time horizons. For smaller  $\rho$ , the high frequencies contain a larger proportion of the variance in  $x$  than the low frequencies do. The estimated parameter is therefore closer to the short run parameter than the long run parameter: the average estimate is 0.281. As  $\rho$  increase the low frequencies become more influential and the estimated parameter comes closer to the long run parameter. When  $\rho = 1$ ,  $x_t$  is a random walk and  $\hat{\beta}$  should asymptotically approach  $\beta_{LR}$ . As can be seen in Table 1, when  $T=100$  the estimated parameter is 0.916, but as  $T$  increase to 500, the average estimate is close to 1: the average estimate is 0.981 when  $T=500$ . Yet, there is still a small bias even for  $T=500$ .

[TABLE 1]

The band spectrum regression estimates are on average unbiased. They are also more efficient than the estimates from (30). The variance of the residual is also correctly estimated by the band spectrum MODWT regression. The true value is 1 and the average estimate is, irrespective of  $\rho$  and  $T$ , within the interval 0.975 and 0.999. The estimate from the misspecified model (28) is between 1.124 and 5.454, which is expected considering (7). The closer the parameter estimate is to the long run parameter the greater the influence of the misspecification component. The more  $\hat{\beta}$  deviates from  $\beta_{SR}$  the higher the variance of the disturbance in the misspecified model.

#### 4.2 Three time horizons

Next we consider a case with three time horizons: the long run, the medium run, and the short run. As before we let the frequencies 0 to 1/32 represent the long run, the medium run is represented by the frequencies 1/32 to 1/8 and the short run by the frequencies 1/8 to 1/2. Using monthly data the short run is defined as cycles with a periodicity from 2-8 months, the medium run by cycles with a periodicity from 8 to 32 months and the long run as more

persistent variations than 32 months. Using quarterly data this definition of the time horizons imply that the short run lasts up to 2 years, the medium run 2 to 8 years, and the long run is 8 years and beyond.

The data generating process is given by,

$$y_t = x_{LR,t}\beta_{LR} + x_{MR,t}\beta_{MR} + x_{SR,t}\beta_{SR} + \varepsilon_t, \quad (31)$$

where  $x_{MR,t}$  is the medium run variations in the explanatory variables and  $\beta_{MR}$  the medium run parameter. In the simulations we set,  $\beta_{LR} = 1$ ,  $\beta_{MR} = 0.5$ , and  $\beta_{SR} = 0.2$ . As before  $\varepsilon_t \sim N(0,1)$  and  $x_t$  is simulated as an autoregressive process. Simulation results are presented in Table 2.

[TABLE 2]

The results with three time horizons are similar to the results with two time horizons. The miss-specified model estimates a weighted average of the three time horizons and consequently the residuals include a miss-specification component. As indicated by the simulations, the closer the estimated parameter is to the long run parameter, the greater the size of the miss-specification component.

The band spectrum regression is on average unbiased and more efficient than the miss-specified OLS estimator. A difference compared to the two time horizon model is that the standard error of the parameter estimates has increased. The increase is small for the long run parameter, and higher for the short run parameter estimates.

As for the two time horizon case, the estimates of the variance of the shocks are biased for the miss-specified model, but close to the true value for the band spectrum regression. Moreover, the closer  $\hat{\beta}$  is to  $\beta_{LR}$  the larger the bias for the miss-specified model.

## 5. Summary

In this paper we have shown how the maximal overlap discrete wavelet transform can be used to estimate band spectrum regressions. We have also illustrated the econometric consequences of assuming that the model only contains one time horizon when there are several in the true model.

## References

- Andersson, F.N.G. (2008). Wavelet Analysis of Economic Time Series. *Lund Economic Studies*, 149.
- Corbae, D., S. Ouliaris, and P.C.B. Phillips (2002). Band Spectral Regression with Trending Data. *Econometrica*, 70, 1067-1109.
- Crowley, P. M. (2007). A Guide to Wavelets for Economists. *Journal of Economic Surveys*, 21, 207-267.
- Engle, R.F. (1974). Bandspectrum Regression. *International Economic Review* 15(1), 1-11.
- Englund, P., Persson, T. and L.E.O. Svensson (1992). Swedish Business Cycles: 1861-1988. *Journal of Monetary Economics* 30, 343-371.
- Friedman, M. (1957). *A Theory of the Consumption Function*. Princeton University Press.
- Friedman, M. and S. Kuznets (1954) *Income from Independent Professional Practice*. New York: National Bureau of Economic Research.
- Islam, N. (2003). What Have we Learnt from the Convergence debate? *Journal of Economic Surveys* 17(3): 309-356.
- King, R.G. and M.W. Watson (1996). Money, Prices, Interest Rates and the Business Cycle. *The Review of Economics and Statistics* 78, 35-53.
- Percival, D. and T. Walden (2006). *Wavelet Methods for Time Series Analysis*. Cambridge: Cambridge University Press.
- Phillips, P.C.B. (1991). Spectral Regression for Co-integrated Time Series, in *Non-parametric and Semiparametric Methods in Economics and Statistics*, editors Barnett W., J. Powell, and G. Tauchen. Cambridge: Cambridge University Press.
- Ramsey, J.B. and C. LAMPART (1998). Decomposition of Economic Relationships by Timescale Using Wavelets. *Macroeconomic Dynamics* 2, 49-71.
- Samuelson, P.A. (1955). *Economics*. New York: McGraw and Hill.
- Slutzky, E. (1938). The Summation of Random Causes as the Source of Cyclic Processes. *Econometrica*, 5, 105-146.
- Goodfriend, M. and R.C. King (1997). The New Neoclassical Synthesis and the Role of Monetary Policy. *NBER Macroeconomics Annual* 12, 231-283.
- Woodford, M (2003). *Interest and Prices: Foundation of a Theory of Monetary Policy*. New Jersey: Princeton University Press.



$T$	<i>Miss-specified model</i>		<i>Correctly specified model</i>		
	$\hat{\beta}$	$\hat{\sigma}^2$	$\hat{\beta}_{LR}$	$\hat{\beta}_{SR}$	$\hat{\sigma}^2$
$\rho = 0$					
100	0.281 (0.061)	1.124 (0.174)	1.017 (0.211)	0.199 (0.056)	0.980 (0.147)
200	0.281 (0.041)	1.132 (0.120)	0.997 (0.142)	0.200 (0.040)	0.985 (0.101)
500	0.281 (0.025)	1.146 (0.081)	0.995 (0.085)	0.200 (0.024)	0.995 (0.064)
$\rho = 0.5$					
100	0.412 (0.076)	1.362 (0.229)	1.000 (0.107)	0.200 (0.054)	0.975 (0.138)
200	0.412 (0.052)	1.400 (0.168)	1.000 (0.075)	0.200 (0.038)	0.990 (0.100)
500	0.416 (0.035)	1.425 (0.109)	1.000 (0.045)	0.200 (0.025)	0.995 (0.063)
$\rho = 1$					
100	0.916 (0.065)	4.728 (2.819)	1.000 (0.013)	0.200 (0.044)	0.980 (0.138)
200	0.957 (0.033)	5.267 (3.250)	1.000 (0.006)	0.200 (0.030)	0.992 (0.102)
500	0.981 (0.016)	5.454 (3.176)	1.000 (0.002)	0.200 (0.019)	0.999 (0.064)

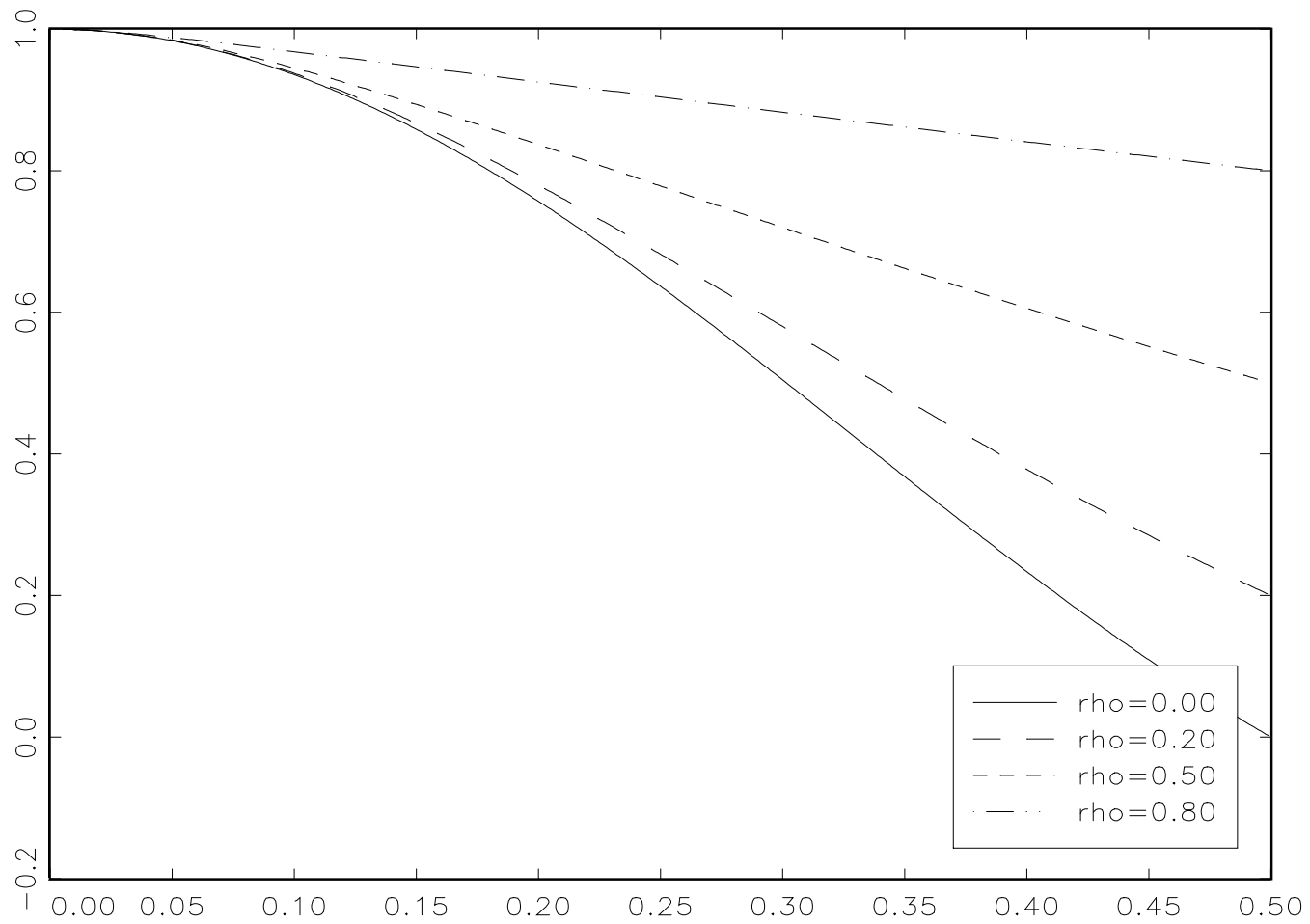
Table 1: Simulation Results, Two Time Horizons

**Note:** The standard error of the parameter estimates are presented below the average parameter estimates.

$T$	<i>Miss-specified model</i>		<i>Correctly specified model</i>			
	$\hat{\beta}$	$\hat{\sigma}^2$	$\hat{\beta}_{LR}$	$\hat{\beta}_{MR}$	$\hat{\beta}_{SR}$	$\hat{\sigma}^2$
$\rho = 0$						
100	0.326 (0.062)	1.151 (0.177)	1.018 (0.222)	0.499 (0.235)	0.199 (0.066)	0.971 (0.146)
200	0.325 (0.042)	1.157 (0.122)	0.999 (0.151)	0.495 (0.164)	0.202 (0.047)	0.981 (0.101)
500	0.325 (0.026)	1.170 (0.082)	0.995 (0.091)	0.500 (0.104)	0.200 (0.029)	0.992 (0.064)
$\rho = 0.5$						
100	0.485 (0.072)	1.361 (0.215)	1.004 (0.114)	0.496 (0.169)	0.202 (0.079)	0.965 (0.137)
200	0.485 (0.050)	1.393 (0.158)	0.999 (0.079)	0.503 (0.108)	0.198 (0.054)	0.985 (0.100)
500	0.490 (0.033)	1.411 (0.096)	1.000 (0.044)	0.501 (0.072)	0.202 (0.036)	0.994 (0.062)
$\rho = 1$						
100	0.934 (0.052)	3.690 (2.006)	1.000 (0.014)	0.496 (0.092)	0.202 (0.075)	0.971 (0.137)
200	0.966 (0.026)	4.037 (2.294)	1.000 (0.006)	0.503 (0.064)	0.197 (0.054)	0.987 (0.100)
500	0.985 (0.012)	4.143 (2.223)	1.000 (0.003)	0.500 (0.041)	0.200 (0.036)	0.997 (0.064)

Table 2: Simulation Results, Three Time Horizons

**Note:** The standard error of the parameter estimates are presented below the average parameter estimates.



*Figure 1: Autoregressive parameter using a sub-set of all frequencies*