1997

# Wavelet Analysis of Commodity Price Behavior

Russell Davidson

Walter C. Labys

Jean-Baptiste Lesourd

*Computational Economics* **11:** 103–128, 1998.



# Wavelet Analysis of Commodity Price Behavior*

RUSSELL DAVIDSON[1], WALTER C. LABYS[2] and
JEAN-BAPTISTE LESOURD[3]
[1]*Queen's University Kingston, Ontario, Canada and GREQAM Marseille, France*
[2]*West Virginia University*
[3] *GREQAM Marseille, France*

**Abstract.** We propose a form of semi-nonparametric regression based on wavelet analysis. Traditional time series methods usually involve either the time or the frequency domain, but wavelets can combine the information from both of these. While wavelet transforms are typically restricted to equally spaced observations an integer power of 2 in number, we show how to go beyond these constraints. We use our methods to construct 'patios' for 21 important international commodity price series. These graph the magnitude of the variations in the series at different time scales for various subperiods of the full sample.

**Key words:** wavelets, semi-nonparametric regression, commodity price behavior, multi-resolution analysis

## 1. Introduction

In this study we propose a form of semi-nonparametric regression based on wavelet analysis. Traditional time series analyses rely on methods that involve either the time or the frequency domain. But wavelet transforms permit an analysis that combines both time and frequency information, the latter in terms of levels of time resolution. Usually wavelet transforms are used with equally spaced observations whose number is an integer power of two. Here, we show how to go beyond these constraints. In an application, we use our methods to construct 'patios' for some important international commodity price series. These patios show the magnitude of the variations in the series at different time scales for different subperiods of the full sample.

In the next section, we give a brief review of past econometric findings concerning commodity price behavior. Then, in Section 3, we present a discussion of multi-resolution analyses, in particular their implementation in terms of wavelets. Section 4 contains a description of how we extend traditional multi-resolution analyses to the context of nonparametric regression, and in Section 5, we present the results of applying our methods to some commodity price series. Section 6 provides an economic interpretation of the results, and, finally, Section 7 gives a

* This research was supported, in part, by grants from the Social Sciences and Humanities Research Council of Canada.

few conclusions regarding the implications of our findings on commodity prices and some suggestions for future research.

## 2. Commodity Price Behavior

The econometric analysis of commodity price behavior is based on a variety of generating processes for short-term price movements. Although these often appear completely random, arguments can be made for price regularities or cycles. Models as simple as the cobweb give the most basic explanation of these: the market clears at every point in time, but prices in the supply equation depend on expectations of future prices. The first-order differential equation describing this phenomenon can yield convergent, divergent, or continuously oscillating solutions. More complex mechanisms give rise to second-order equations, whose solutions can include not only continuous but also damped or explosive oscillations. Further explanations can be found in Ackerman (1957), Baumol (1962), or Gandolfo (1985). Development of these models specific to commodity markets appear in Brock (1988), Chavas and Holt (1991), and Mackey (1989). Studies that expand this approach to include nonlinearity and chaotic behavior include Burton (1993), Boldrin and Woodford (1990), and Jensen and Urban (1984).

The detection of regularities has proven difficult for a number of reasons. While examination of price movements may suggest cyclicality, a statistical demonstration of the existence of cycles can be difficult. Most time-domain research on this point has concentrated on short-term fluctuations. An early study of this type is Working's (1958) investigation of the random walk; he proposes that the continuous flow of many different kinds of information into commodity markets causes frequent price changes that might be nearly random. Still, this model allows for some gradualness of price changes, and thus, some degree of very short-term predictability. Samuelson (1965) further develops this theory by postulating that commodity prices follow a martingale process.

Most of the empirical tests of the random walk and martingale hypotheses look for serial correlation and trends, since both processes require price changes to be independent. Trend deviations from a random walk are first discovered for wheat and corn by Houthakker (1961) and for soybeans by Smidt (1965). Labys and Granger (1970), in performing a spectral analysis on futures and spot price series, find some evidence for a modified random walk process, mostly resembling the martingale process. Stevenson and Bear (1970) and Leuthold (1972) also find departures from random walk using filter rules that reveal positive and negative price dependence, thus casting doubt on the validity of the random walk model.

While most of these studies are based on linear models, other possibilities have been considered, including nonlinear models reflecting chaotic behavior. Drawing upon Houthakker's (1961) analysis of cotton prices, Mandelbrot (1963) replaces Gaussian probability laws with those termed 'stable Paretian'. His approach represents an attempt to discover orderly behavior within what appears to be a random

series of price fluctuations. It is in this context that Frank and Stengos (1989) investigate the martingale hypothesis, using an approach of Sims (1984) as well as chaos-based approach. Although Frank and Stengos are not able to reject the martingale hypothesis using a series of standard econometric tests involving daily and weekly silver and gold prices, they did provide correlation-dimension-based evidence of the presence of nonlinear structure. This structure is also confirmed using a similar test on soybeans by Blank (1991), and on silver, copper, sugar, and coffee futures by DeCoster, Labys, and Mitchell (1992).

Other recent studies of nonlinear dynamic processes in commodity prices involve the testing of ARCH and GARCH models. For instance, Yang and Brorsen (1992) use GARCH models for corn, pork bellies, soybeans (including meal and oil), sugar, wheat, and gold daily spot price changes. Results from these models suggest that the variance of the price changes is not constant. Departing from the linear corn/hog price cycle analysis of Jameson (1983), Chavas and Holt (1991) use a GARCH model, but suggest that the pork market may be characterised by other forms of nonlinear dynamics. Two other recent studies that investigate nonlinear dynamics in commodity prices, but employ a dynamic programming approach, are those of Deaton and Laroque (1992, 1995).

Commodity price studies in the frequency domain have usually taken a spectral analysis approach. For example, Labys and Granger (1970) analyze agriculture prices, Gelb (1979) and Parikh (1973) coffee prices, Labys, Elliott and Rees (1971) copper prices, Rausser and Cargill (1970) broiler prices, Weiss (1970) cocoa prices, and Slade (1981) several metal prices using this method. The hog cycle analysis of Talpaz (1974) is frequency related, but is limited to a Fourier representation. The advantage of the spectral approach is that it permits cycles of different frequencies to be discovered and to be tested statistically by methods based on the fast Fourier transform of a price series. This enables researchers to study not only shorter-term or higher frequencies reflecting random-walk behavior, but also medium and long-term frequencies embodying business cycles and/or growth cycles.

The disadvantage of the spectral approach is that the Fourier transform has difficulty with functions having transient components, that is, components localised in time. Another problem is that the Fourier transform of cyclical information does not provide insights into the phase relationships (leads and lags) between cycles of the same frequency in different series. In this study, we show how such problems can be overcome by combining time and resolution analysis through the use of wavelets.

## 3. A Brief Theory of Wavelets

A concept that, at least in principle, allows one to obtain local time information while still providing information in the frequency domain is *multiresolution analysis*. The notion seems to have grown out of work in which a signal is subjected to a bank of filters: high-pass filters, low-pass filters, medium-pass filters, etc. One can

imagine listening to the output of an equaliser on hifi equipment. Listening to only one band, one hears only a narrow set of frequencies. However, the output does not consist of just a continuous unvarying drone. There will normally be variations in pitch over time, limited but perceptible, and there will possibly be substantial variations in volume. A Fourier analysis would have neither of these features. At any definite frequency there would be a fixed intensity (volume), constant over the whole duration of the signal. And it would be a definite frequency, unable to vary even slightly.

What makes the difference? Principally, the fact that each *band* passed by a filter does not correspond to a definite frequency, but rather to a range of frequencies. The fundamental *uncertainty principle* excludes absolute precision of information in the time and frequency domains simultaneously. For instance, when we deal with discrete time series, the time resolution can be no better than the period between successive observations, and the frequency resolution can give no information about frequencies faster than the frequency of observation or slower than the frequency determined by the overall length of the sample.

This trade-off between information in the time and frequency domains inspires multi-resolution analysis. Much that we present below derives from the work of Daubechies (1990, 1992), Grossman and Morlet (1989), and Meyer (1992). More formal presentations of wavelet theory can be found in Beylkin et al. (1992), Chui (1992a and b), Combes et al. (1991), Meyer (1990), Meyer and Coifman (1991), and Wickerhauser (1994).

With monthly data, information on phenomena at frequencies shorter than a month will not be available. But we can certainly consider *lower* levels of resolution, and look at phenomena characterised by quarterly, or yearly frequencies, or by frequencies commonly thought to be associated with the business cycle. Notice that the concept of frequency cannot be used in this context in its usual precise sense. As noted in Priestley (1995), the term 'frequency' refers only to sines, cosines, and the exponential function of an imaginary argument. It is preferable therefore to speak of levels of resolution rather than frequencies.

A multi-resolution analysis is clearly a form of smoothing. High-frequency phenomena (those perceptible only at high levels of resolution) ride on top of phenomena visible at lower resolution, perturbing with rapid movements the slower, smoother, movements corresponding to lower resolutions. Imagine passing a signal through a low-pass filter so that only variations corresponding to annual or longer frequencies get through. This will yield a smoothly varying filtered output corresponding to an annual level of resolution. All the variation cut out by the filter, which is typically much less smooth, can be thought of as *details* that can be superimposed upon the annual level of resolution to obtain a more detailed picture at the monthly level of resolution. Now imagine that a new data source becomes available, providing daily data where before there were only monthly data. If we subtract from these daily data the monthly data, we obtain another level of detail.

Consider the space $L^2(\mathbb{R})$ of square-integrable functions on the real line. We may define a multi-resolution analysis of $L^2(\mathbb{R})$ as follows:

(1) A multi-resolution analysis of $L^2(\mathbb{R})$ is an increasing sequence $\{V_j\}_{j\in\mathbb{Z}}$ of closed subspaces of $L^2(\mathbb{R})$, with the following three properties:

$$\bigcap_{j=-\infty}^{\infty} V_j = \{0\}, \qquad \bigcup_{j=-\infty}^{\infty} V_j \text{ is dense in } L^2(\mathbb{R}); \tag{1}$$

(2) for all $f \in L^2(\mathbb{R})$, and for all integers $j \in \mathbb{Z}$, $f(x) \in V_0$ if and only if $f(2^j x) \in V_j$; and

(3) there exists a function $\phi(x) \in V_0$ such that the sequence of functions $\{\phi(x-k)\}_{k\in\mathbb{Z}}$ constitutes a Riesz basis of $V_0$.

The first of these properties defines the different levels of resolution. We see that there is a double infinity of these. The lowest level, corresponding to $j = -\infty$, admits only constant functions, while the highest level, corresponding to $j = +\infty$, encompasses the whole space $L^2(\mathbb{R})$.

The second condition specifies the relation between successive levels of resolution. Any function that belongs to level $j$ corresponds to a function at any other level by a simple process of dilation by an integer power of 2.

In the third condition, a *Riesz basis* of $V_0$ is a sequence $\{e_i\}$ of elements of $V_0$ with the property that there is an isomorphism from the Hilbert space $\ell^2(\mathbb{N})$ of square-integrable doubly infinite sequences onto $V_0$ such that $T(\varepsilon_i) = e_i$, where $\varepsilon_i$ is the sequence all of whose elements are zero except element $i$, which is unity. The condition implies, among other things of mainly technical interest, that, for an appropriate choice of normalisation,

$$\sum_{k=-\infty}^{\infty} \phi(x-k) = 1 \quad \forall x \in \mathbb{R}. \tag{2}$$

Thus the set $\{\phi(x-k)\}_{k\in\mathbb{Z}}$ gives a *partition of unity*, and it is this property that allows certain multi-resolution analyses to be *local*.

Let us take a brief look at the very simplest multi-resolution analysis, associated with the name of Haar. In this, the basic function $\phi(x)$, called the *scaling function*, or, sometimes in more recent literature, the *father wavelet*, is just the indicator function for the unit interval:

$$\phi(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1 & \text{for } 0 \le x < 1 \\ 0 & \text{for } x \ge 1. \end{cases}$$

With this choice, it is clear that (2) is satisfied. We see at once that $V_0$, being generated by the set of all integer translates of the indicator function, is just

the set of all functions that are piecewise constant on all segments of the form $[k, k+1[$, $k \in \mathbb{Z}$. Thus, at the base level of resolution, we can see nothing inside intervals of this form. All details on a scale less than that given by the distance between any two successive integers is lost at this level of resolution.

The second defining property of a multi-resolution analysis allows us to see what the other levels of resolution look like. For $j = 1$, for instance, we see that the function $\phi(2x)$ belongs to $V_1$. This function is readily seen to be the indicator function for the interval $[0, \frac{1}{2}]$. Thus $V_1$ comprises those functions that are piecewise constant on the intervals of length one half for which one end is an integer. There is just twice as much detail available in such functions. It is clear how this can be extended to arbitrary integers $j$, positive or negative. A positive $j$ gives more detail than that in $V_0$; a negative $j$ less.

The shortcomings of the Haar analysis are clear. Functions that are piecewise constant are maximally smooth *inside* the segment of constancy, but are not even continuous, let alone differentiable, at the boundaries of adjoining segments. We would clearly be prepared to sacrifice constancy inside the segments for a little continuity at the joins.

Another multi-resolution analysis can be defined for which the scaling function is the so-called *tent* function, defined as

$$\phi(x) = \begin{cases} 0 & \text{for } x < 0, \\ x & \text{for } 0 \leq x \leq 1, \\ 2 - x & \text{for } 1 \leq x \leq 2. \end{cases}$$

Property (2) is again trivially satisfied, and now we see that $V_0$ contains functions that are piecewise *linear* on the segments $[k, k+1[$, $k \in \mathbb{Z}$. Continuity at the joins is now achieved; the discontinuity is in the first derivative.

The idea of the preceding paragraphs may be extended so that $V_0$ can be made up of functions that are piecewise quadratic, or piecewise cubic, etc., and have discontinuities of the second, third, etc., derivative at the joins. Such functions are called *splines*, and their theory is well developed in the numerical analysis literature. And, although they could form the basis for the sort of procedure we develop here, something even better for our purposes is available.

One highly desirable property of spline wavelets is compact support. In the simplest case – the indicator function as scaling function – the scaling function has support only on [0, 1]. In the case of the tent function, the support is twice as long. Generally, there is a tradeoff between regularity, or smoothness, and the extent of the support of the scaling function. A short support is desirable precisely because this is what makes a multi-resolution analysis *local*. The shorter the support, the more local the information in the multi-resolution analysis for any given level of resolution. The Haar analysis achieves perfect localisation at the cost of continuity. Other setups give greater regularity at the cost of less perfect localisation.

Now we need to specify where wavelets come into a multi-resolution analysis. Recall from (1) that $V_0 \subset V_1$. It should therefore be possible to define a subset of $V_1$ complementary to $V_0, W_1$ say, so that $V_1$ is the *direct sum* of $V_0$ and $W_1$:

$$V_1 = V_0 \oplus W_1. \tag{3}$$

It is normally possible to find such a $W_1$ that is the span of the set of functions $\psi(x-k), k \in \mathbb{Z}$, for some suitable function $\psi$, called the *wavelet* of the multi-resolution analysis. (It would be the *mother wavelet* in the terminology that calls the scaling function the father wavelet.) In the Haar case, the obvious choice is the following function:

$$\psi(x) = \begin{cases} 0 & \text{for } x < 0, \\ 1 & \text{for } 0 \le x < \frac{1}{2}, \\ -1 & \text{for } \frac{1}{2} \le x < 1, \\ 0 & \text{for } x \ge 1. \end{cases}$$

It is clearly possible to construct any function that is piecewise constant on all intervals of length one half and with one endpoint an integer by taking a linear combination of the $\phi(x-k)$ – the indicator functions on the unit intervals – and of the $\psi(x-k)$. These piecewise constant functions make up $V_1$ in the Haar case.

In addition, from the last two defining properties of a multi-resolution analysis, $V_0$ is spanned by the functions $\phi(x-k), k \in \mathbb{Z}$, and $V_1$ is spanned by the $\phi(2x-k), k \in \mathbb{Z}$. Thus it must be possible to express both $\phi(x)$ and $\psi(x)$ as linear combinations of the $\phi(2x-k), k \in \mathbb{Z}$. If in addition $\phi$ and $\psi$ have compact (bounded) support, these linear combinations are necessarily finite. Thus there are coefficients $c_k$ and $g_k$ such that, for all $x \in \mathbb{R}$,

$$\phi(x) = \sum_{k=0}^{K} c_k \phi(2x-k), \quad \text{and} \tag{4}$$

$$\psi(x) = \sum_{k=0}^{K} g_k \phi(2x-k), \tag{5}$$

for some finite $K$. Equations (4) and (5) constitute what is called the *scaling filter* associated with the multi-resolution analysis. It turns out that the $g_k$ are determined once the $c_k$ are given, and that in fact the scaling function $\phi$ and the wavelet $\psi$ are fully determined by (4) and (5).

The *Daubechies wavelets* that we use in our applications make use of sets $\{c_k\}$ which turn the functions $\phi(x-k)$ and $\psi(x-k)$ into an *orthonormal basis* of

the space they span. Thus the Daubechies father and mother wavelets satisfy the orthogonality properties

$$\int_{-\infty}^{\infty} \phi(x)\psi(x-k)\,\mathrm{d}x = 0, \quad k \in \mathbb{Z},$$

$$\int_{-\infty}^{\infty} \phi(x)\phi(x-k)\,\mathrm{d}x = 0, \quad k \in \mathbb{Z},\ k \neq 0,$$

$$\int_{-\infty}^{\infty} \psi(x)\psi(x-k)\,\mathrm{d}x = 0, \quad b \in \mathbb{Z},\ k \neq 0,$$

and the normalisation properties

$$\int_{-\infty}^{\infty} \phi^2(x)\,\mathrm{d}x = \int_{-\infty}^{\infty} \psi^2(x)\,\mathrm{d}x = 1.$$

It is in fact not too difficult to find wavelets satisfying these orthonormality properties; there are many such examples. Note, however, that spline wavelets do not do so, except for the Haar wavelet.

What makes the Daubechies wavelets really useful is that they have compact support and a certain number of *vanishing moments*. For each set of Daubechies wavelets, there is an integer $m$ such that, for all integers $k$ with $0 \leq k \leq m$,

$$\int_{-\infty}^{\infty} x^k \psi(x)\,\mathrm{d}x = 0. \tag{6}$$

Since $\psi$ has compact support, the effective range of the above integral is not at all infinite. It is easy to see as well that the number of nonzero $c_k$ in the scaling filter is finite.

There is now a tradeoff between the length of the support of the wavelets and the number of vanishing moments. An increase in $m$ is bought only at the cost of extending the length of the support, that is, of lessening the degree of localisation achieved by the wavelets. However, a greater value of $m$ also provides a higher degree of continuity and differentiability of the wavelets. The Daubechies wavelets we use in this paper have $m = 2$. The resulting father wavelet is not too dissimilar from the tent function, while the mother wavelet is oscillatory, with four main crossings of the axis.

Suppose then that we choose a value of $m$ well adapted to our purposes. We may define $V_0$ as the space spanned by the functions $\phi(x-k), k \in \mathbb{Z}$. Because of (3), $V_1$ is spanned by these functions along with the $\psi(x-k), k \in \mathbb{Z}$. As we proceed to successively higher resolutions, the spaces $V_l$ are spanned by the functions already in use along with the $2^{j/2}\psi(2^j x - k), j < l, k \in \mathbb{Z}$. (The factors of $2^{j/2}$ merely serve to normalise the functions.) In this way, we can construct a multi-resolution analysis of $L^2(\mathbb{R})$ based on Daubechies wavelets.

In most applications, we neither need nor want to consider functions defined on the whole real line. If attention is restricted to functions with support in some finite interval $[a, b]$, then the compactness of the support of the Daubechies wavelets means that, at any given resolution, only a finite number of the wavelets making up the orthonormal basis at that resolution have support that intersects $[a, b]$.

Let us suppose, without loss of generality, since we can always rescale and relocate a function with compact support, that the interval we consider is [0, 1]. It turns out that the support of a Daubechies wavelet is an interval of the form $[0, 2m + 1]$, where $m$ is the highest nonvanishing moment of the wavelet $\psi$; see (6). Let $j_0$ be the smallest integer such that $2m + 1 \leq 2^{j_0}$. Then the wavelets $\phi(2^{j_0}x - k)$ have support of length $2^{-j_0}(2m+1) < 1$. Define a set of $2^{j_0}$ functions $\phi_k(\cdot), 0 \leq k < 2^{j_0}$, on [0, 1] as follows:

$$\phi_k(x) = \begin{cases} \phi(2^{j_0}x - k) & \text{if } 2^{j_0}x - k \geq 0 \\ \phi(2^{j_0}(x + 1) - k) & \text{otherwise.} \end{cases} \tag{7}$$

Thus the $\phi_k$ are just the scaling functions 'wrapped round' the interval [0, 1] when their supports extend beyond the endpoint 1. Thus the functions $\phi_k(x)$ yield a partition of unity on [0, 1]:

$$\sum_{k=0}^{2^{j_0}-1} \phi_k(x) = 1, \quad \text{for all } x \in [0, 1]; \tag{8}$$

they are mutually orthogonal, and they can be normalised to have unit $L^2$ norm at the cost of making the right-hand side of (8) a power of 2 rather than 1. We may regard them as spanning the subspace of $L^2[0, 1]$ that corresponds to the 'base' level of resolution.

Higher levels of resolution are obtained by appending to the set $\phi_k$, first, another $2^{j_0}$ functions $\psi_k$, defined just like the $\phi_k$ in (7) – but with the mother wavelet $\psi$ in place of the father wavelet $\phi$ – then, subsequently, sets of functions of the form $\psi(2^j x - k), j > j_0, k = 0, \ldots, 2^j$, with the same sort of wrapping round as in (7). It is easy to see that, at each resolution, there are exactly $2^{j+1}$ basis functions, for each $j \geq j_0$. Given the second defining property of a multi-resolution analysis, another way to span the space of the resolution with $2^{j+1}$ basis functions, based exclusively on the father wavelet, is to use the functions $\phi(2^{j+1}x - k), k = 0, \ldots, 2^{j+1} - 1$, suitably wrapped round.

The *discrete wavelet transform* can now be constructed. Let us denote by $\theta_k(\cdot), k = 0, \ldots, 2^j - 1$ the functions making up an orthonormal basis of some level of resolution of $L^2[0, 1]$ using one set of father wavelets and various sets of mother wavelets. Consider a function $f \in L^2[0, 1]$. Its representation at this resolution is just the orthonormal projection of $f$ on to the span of the $\theta_k$:

$$f^j(x) \equiv \sum_{k=0}^{2^j-1} a_k \theta_k(x),$$

where

$$a_k = \int_0^1 f(x)\theta_k(x)\,\mathrm{d}x.$$

Similarly, let the same space be spanned by the orthonormal set of functions $\eta_k(\cdot), k = 0, \ldots, 2^j - 1$, where the $\eta_k$ are translations and dilations of the father wavelet only. Then we may write

$$f^j(x) = \sum_{k=0}^{2^j-1} b_k \eta_k(x),$$

where

$$b_k = \int_0^1 f(x)\eta_k(x)\,\mathrm{d}x. \tag{9}$$

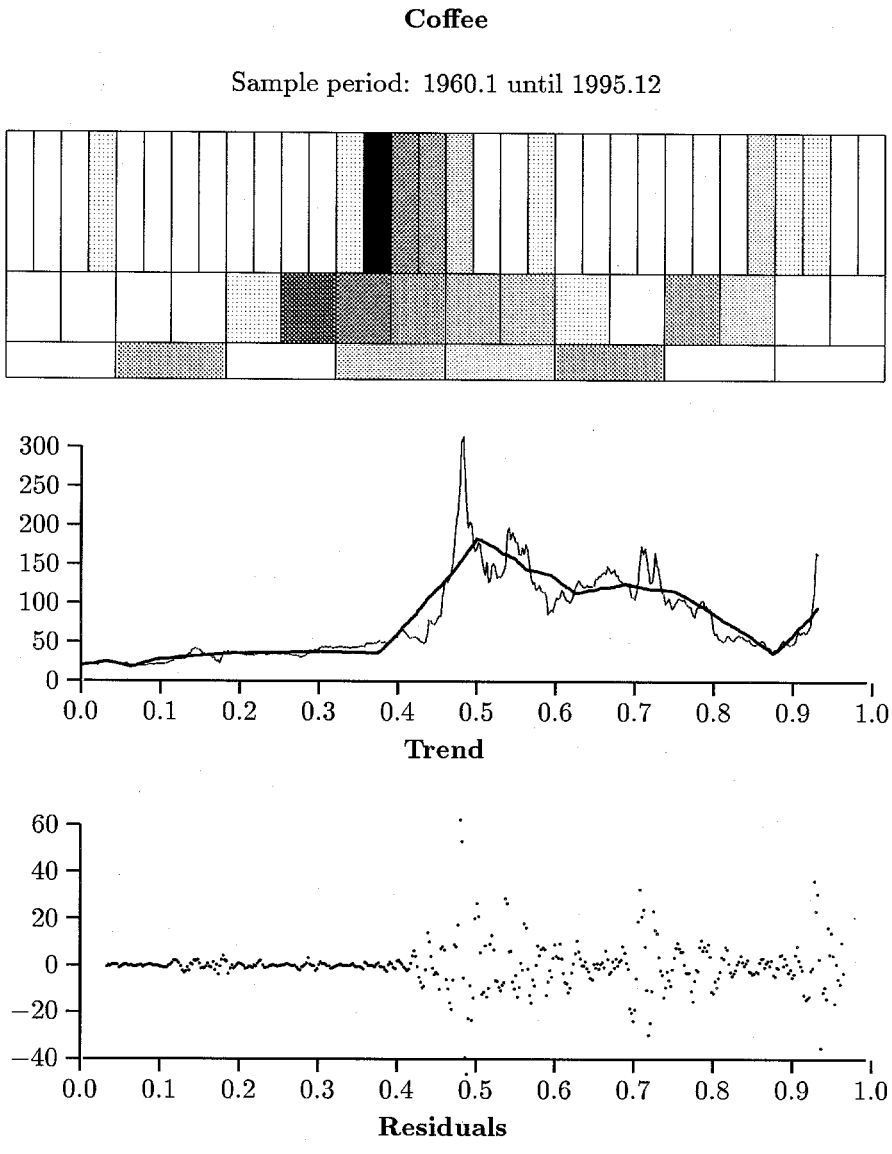
At this resolution, then, the function $f$ is represented by either one of the two $2^j$-vectors ***a*** and ***b***, with typical elements $a_k$ and $b_k$, respectively.

Since both the $\theta_k$ and the $\eta_k$ constitute orthonormal bases, the vectors ***a*** and ***b*** corresponding to any given $f$ are related by a matrix orthonormal transformation independent of $f$. This orthonormal transformation can be implemented by a fast algorithm constructed on the same lines as the fast Fourier transform. It is called the *fast wavelet transform*. Like the FET, it applies only to vectors with exactly $2^j$ components, for some positive integer $j$. The computational details for the Daubechies wavelets of low order can be found in Press et al. (1992).

The interest of the discrete wavelet transform, fast or not, is as follows. The vector ***b*** can be thought of as a *sampling* of the values of the function $f$, sufficient for a description of $f$ at resolution $j$. In fact, since the support of $\phi(2^j x - k)$ is the interval $[k2^{-j}, (k + 2m + 1)2^{-j}]$ (ignoring wrap-around), (9) can be thought of as providing (up to a scalar factor) a weighted average of the values of $f$ on this interval. On the other hand, ***a*** provides information about $f$ with respect to time and level of resolution. Unlike a Fourier frequency-domain representation, this representation is discrete. Its information content can be displayed visually by means of a 'patio' diagram, as illustrated in the top portion of Figure 1 of Section 5. Each tile of such a patio represents one of the components $b_k$ (the darkness of shading is proportional to $b_k^2$), and each of these components corresponds to a time interval and a resolution, which are, respectively, the support and the resolution level of the corresponding $\eta_k$. As Figure 1 shows clearly, the time intervals are longer at lower resolutions, on account of the uncertainty principle.

## 4. Wavelet Estimation and Tests

One of the most constraining features of the various wavelet transforms considered above is their need for data sets with observations numbering an integer power of




Proportion of variation at resolution 1 = 0.240
Proportion of variation at resolution 2 = 0.455
Proportion of variation at resolution 3 = 0.305
$R^2 = 0.835$

*Figure 1.*

2. And, in interpreting wavelet transforms, there is almost always a presumption that these observations are equally spaced in time.

There are suggestions in the literature for overcoming these constraints. For instance, one can pad data – for example, by repeating the last observation as often as needed – so that the number of observations is increased to the smallest power of 2 greater than (or equal to) the size of the observed sample. Similarly, interpretation can be straightforward if the data, rather than being equally spaced, are spaced by random drawings from a known distribution.

If for the moment we ignore these limitations, the method of data analysis known as *wavelet shrinkage* has great success in coping with widely different data sets for which the observations can be regarded as a realisation of some stochastic process. The simplest example is a nonparametric regression, in which successive observations on a random variable are generated by an unknown deterministic function of the time plus random noise; formally,

$$y_t = g(t) + u_t.$$

Here $y_t$ is the $t^{\text{th}}$ observation on the dependent variable, $g(\cdot)$ is the unknown deterministic function to be estimated, and $u_t$ is the $t^{\text{th}}$ element of a zero mean noise process, which in general may be heteroskedastic and/or serially correlated.

In order to perform wavelet shrinkage, the sequence of points $y_t, t = 1, \ldots, T$ (after padding so that $T = 2^i$ for some positive integer $i$), is subjected to a discrete wavelet transform, yielding a set of transform coefficients associated with different translations and dilations of the mother and father wavelet functions. The 'shrinkage' process then either completely eliminates or greatly reduces all the coefficients smaller in absolute value than some suitable threshold. Finally, the shrunk coefficients are used to perform an inverse transform, the result of which is the estimate of the function $g(\cdot)$.

This procedure is advocated in Donoho, Johnstone, Kerkyacharian, and Picard (1995). Many of its advantages are pointed out in the paper itself and in the accompanying discussion. In particular, the estimate of $g(\cdot)$ is usually as smooth as the unknown function $g$ itself. Examples are presented of functions with discontinuities that are well captured by the wavelet estimator. Closely related work, found in Antoniadis (1994), proves optimality results for a procedure which, while not referred to as 'wavelet shrinkage', is nevertheless clearly a variant of it.

Like all nonparametric methods, this one has an adjustable parameter – here, the threshold below which shrinkage takes place. Like most nonparametric methods, a suitable value for the adjustable parameter can often be found by cross-validation: see Nason (1996) for suggestions for implementing cross-validation in this context, and Hall and Patil (1996) for more theory about choosing the adjustable parameter.

For our present purposes, wavelet shrinkage can be applied directly to the price series we consider, since they consist of equally spaced observations. We need only use an appropriate padding to extend our data series to have $2^i$ observations. However, for more general use of nonparametric regression in econometrics, the

constraints mentioned above are too severe to allow wavelet shrinkage to be practical. What can one do if there are missing data points? or if the independent variable is not time, but rather some exogenous variable whose the distribution we do not wish to model? or if the data have trends, making untenable the usual assumption that the unknown function $g(\cdot)$ is periodic? Blind application of the usual procedure leads to strange 'edge effects' in the estimates of $g(\cdot)$ near the beginning and the end of the sample. Antoniadis (1994) alludes to ways around this involving 'tapering' of the wavelets series, but he provides no details.

The suggestion we make here is to use the translations of the scaling function and the translations and dilations of the mother wavelet function just as we would use trigonometric functions, or orthogonal polynomials, in semi-nonparametric regression. That is, we select a particular scaling function and its corresponding wavelet and treat them as functions that we can evaluate for any given argument. Then we can simply regress our dependent variable on a set of translations and dilations of the basic functions evaluated at the observed values of the explanatory variable.

One advantage of this method is that we can apply standard results on the statistical properties of semi-nonparametric regressions directly. Another is that it answers simply and naturally most of the questions posed above. There are no constraints regarding the number of observations, and missing observations simply do not appear in the sample. Unequally spaced values of the explanatory variable are handled as easily as equally spaced ones. Edge effects can still occur with trending data, as they can with any semi-nonparametric regression, although for a reason different from the one that applies to the standard wavelet transform. We will see later that a simple trick allows us to reduce these effects to a minimum.

In order to analyse the time series we consider, we proceed as follows. First, we compress the length of the observed series to the [0, 1] interval. Since we treat monthly (equally spaced) series, the values of the independent variable at which the various functions appearing in our semi-nonparametric regression are evaluated take the form $l/n$, where $n$ is the sample size and $l = 1, \ldots, n$. For present purposes, we use the Daubechies wavelet with $m = 2$ (see (6)), that is, a wavelet $\psi$ and a scaling function $\phi$ (mother and father wavelet) such that the support of both functions is the interval [0, 5]. By use of an inverse discrete wavelet transform, $\phi$ and $\psi$ are evaluated at a fine grid of $2^i$ points, where $2^i$ is much greater than the sample size $n$. The choice of $i$ simply determines the accuracy with which $\phi$ and $\psi$ are evaluated for an arbitrary argument $x$: given $x$, we take for $\phi(x)$ the value of $\phi$ at the closest point to $x$ in our grid of $2^i$ points.

At our base level of resolution, we use the functions $\phi_k$ defined in (7) for $j_0 = 3$ (since $2m + 1 = 5 < 2^3$). These functions all have support of length $5/(2^3) = 5/8$, and the different $\phi_k$ are all translates of one another by an integer multiple of 1/8. In order to take account of all translates that intersect the [0, 1] interval, therefore, we need 12 of them, namely those whose support begins at $-1/2, -3/8, -1/4, \ldots, 5/8, 3/4, 7/8$. In order to perform an analysis of a series

(of $n$ observations) at this resolution, we now form an $n \times 12$ regressor matrix – whose the elements are the 12 functions evaluated at the points $l/n, l = 1, \ldots, n$ – on which we regress the $n \times 1$ vector given by the series under analysis. Note that we use 12 different functions rather than the 'wrap-around' procedure characteristic of the discrete wavelet transform to reduce edge effects, which would otherwise be substantial with the strongly trending series we consider.

At the next stage, we need exactly the same number of translates of $\psi$ at the base resolution, since the support of $\psi$ is the same as that of $\phi$. At this resolution level, we use an $n \times 24$ regressor matrix. After that, we need $16 + 4 = 20$ translates of $\psi$ shrunk to have support 5/16, thus requiring an $n \times 44$ regressor matrix. Then we need $32 + 4 = 36$ translates of $\psi$ shrunk to support 5/32, and so on. In fact, we stop at this last level of resolution. Thus, at our highest resolution level, we need $12 + 12 + 20 + 36 = 80$ functions based on either $\phi$ or $\psi$.

How then to interpret the results of these semi-nonparametric regressions, at different resolutions? The first, using 12 translates of the scaling function $\phi$, gives a fit that we will interpret as the *trend* of the series. This interpretation is, of course, somewhat arbitrary, since we choose a specific resolution for what we call the trend. The choice is made from the appearance of the fits we actually obtain, as described in the following section. Note that the *mean* of the series being analysed is entirely captured by this first regression, since $\psi$ integrates to zero. In addition, since the Daubechies wavelets we use have vanishing first and second moments ($m = 2$), any linear or quadratic time trend in the data is also completely captured. In fact, numerical experiments, using cubic and quartic trends regressed on the 12 regressors of the first stage show that, although these trends are not exactly fit, as linear or quadratic trends are, they are nonetheless fit with a high degree of accuracy. Indeed, plots of such trends and their wavelet fits cannot be distinguished by eye.

Moving now to higher resolutions, we find another 12 regressors at the first stage, based on $\psi$. The Daubechies wavelets are shaped so that their first four translates, whose supports begin at $-1/2, -3/8, -1/4, -1/8$, are not very different from zero on [0, 1]. They appear to play an insignificant role in our analysis except to iron out the edge effects alluded to earlier. We therefore ignore them for our interpretation.

That leave us with 8 translates, which we can associate with the 8 intervals of length 1/8, beginning successively at $0, 1/8, \ldots, 7/8$. We wish to interpret each translate's contribution to the total fit as a measure of the amount of 'activity', or variation, of the series during a time interval of the form $[i/8, (i + 1)/8], i = 0, \ldots, 7$, at a time resolution equal to 1/8 of the total sample period.

This interpretation would be simple and unambiguous if we had performed a discrete wavelet transform. As we have seen, such a transform corresponds to a transformation of the data by an orthogonal matrix. But, although the wavelet functions considered as elements of $L^2(\mathbb{R})$ form an orthonormal basis, the columns of the regressor matrices we use are neither normalised nor orthogonal. In general

this occurs for two distinct reasons: the wavelet functions are evaluated only at a discrete set of points, and those whose support exceeds [0, 1] are truncated.

If the columns of the regressor matrix were orthonormal, the appropriate measure of variation in each time cell would be estimated coefficient or, better, its square, if we follow the usual practice of measuring variations as sums of squares. Without orthonormality, it is preferable to use the square of the $t$ statistic for the regressor corresponding to that cell. This quantity, of course, has its usual statistical interpretation; as such it has two virtues. First, it is invariant to the scale of the corresponding regressor, thereby removing worries about normalising the regressors, and, second, it measures the *marginal* significance of each regressor, thereby removing the necessity that the regressors be mutually orthogonal. In point of fact, the $t$ statistics for any resolution associated with the four leftmost translates are very rarely significant at conventional levels, and the other translates, except the rightmost two or three, are nearly orthogonal.

Similar interpretations are available at higher resolutions. For the highest resolution we consider, we have 32 $(36 - 4)$ time intervals in which we can measure the variation at a time scale of 1/32 of the total sample period.

We also wish to be able to compare intensities of variation across different time scales. For this, we assess the overall contribution of any given time scale as its marginal contribution to the explained sum of squares of the regression. For instance, at the second highest resolution, which corresponds to a time scale 1/16 of the sample period, the contribution is proportional to the increase in the explained sum of squares in passing from the regression at the next lower resolution (with 24 regressors) to the regression at this resolution (with 44 regressors). We can also measure the importance of the residual variation, which we naturally interpret as the variation at time scales shorter than that of the highest resolution actually considered. For this we use an $R^2$-like statistic, but one in which we ignore the variation in the trend of the series. Thus our 'total sum of squares' is the increase in the explained sum of squares as we go from the first regression, with just 12 regressors, to the final one, with 80.

In summary, then, each resolution level is assigned an overall importance proportional to its contribution to the explained sum of squares, and within that contribution, each time interval is assigned an importance proportional to the square of the $t$ statistic of the regressor corresponding to that interval at that resolution.

This allows us to construct the 'patio' plots we mentioned in the preceding section. These have three rows of subintervals, each row spanning the sample period. The first row, which has 8 subdivisions, corresponds to the first level of 'details' beyond the trend. The next row has 16 subdivisions, and corresponds to the next smallest time scale. The third row has 32 subdivisions, and corresponds to the smallest time scale we treat explicitly. Such patio plots will be presented in the next section.

## 5. Data Analysis: Patio Plots

We perform the wavelet analysis described above for twenty-one international commodity price series, covering the period January 1960 until December 1995. The series for aluminum, gold and petroleum begin in January 1970. Most of these data series can be found conveniently in a secondary source, the UNCTAD (1996) monthly *Commodity Price Bulletin*. These commodities are selected from those traded on exchanges where prices are determined competitively, or on major markets, rather than on the basis of a mark-up. We use monthly data, which provide adequate degrees of freedom to make a wavelet analysis meaningful while avoiding the high-frequency noise often found in daily or tick observations. All series are checked for consistency of definition over time and for possible recording errors. Prices are in the currency of the market in question – unnecessary conversions to U.S. dollars are avoided to minimise possible contamination through exchange-rate fluctuations. In addition, nominal prices, which tend to reflect best the emergence of information that may influence a commodity market, are used rather than real prices, which are more relevant for studying annual terms of trade than physical market adjustments. This choice also avoids possible contamination by a cyclical price deflator. Each of the series we treat in this section is defined in the Appendix.

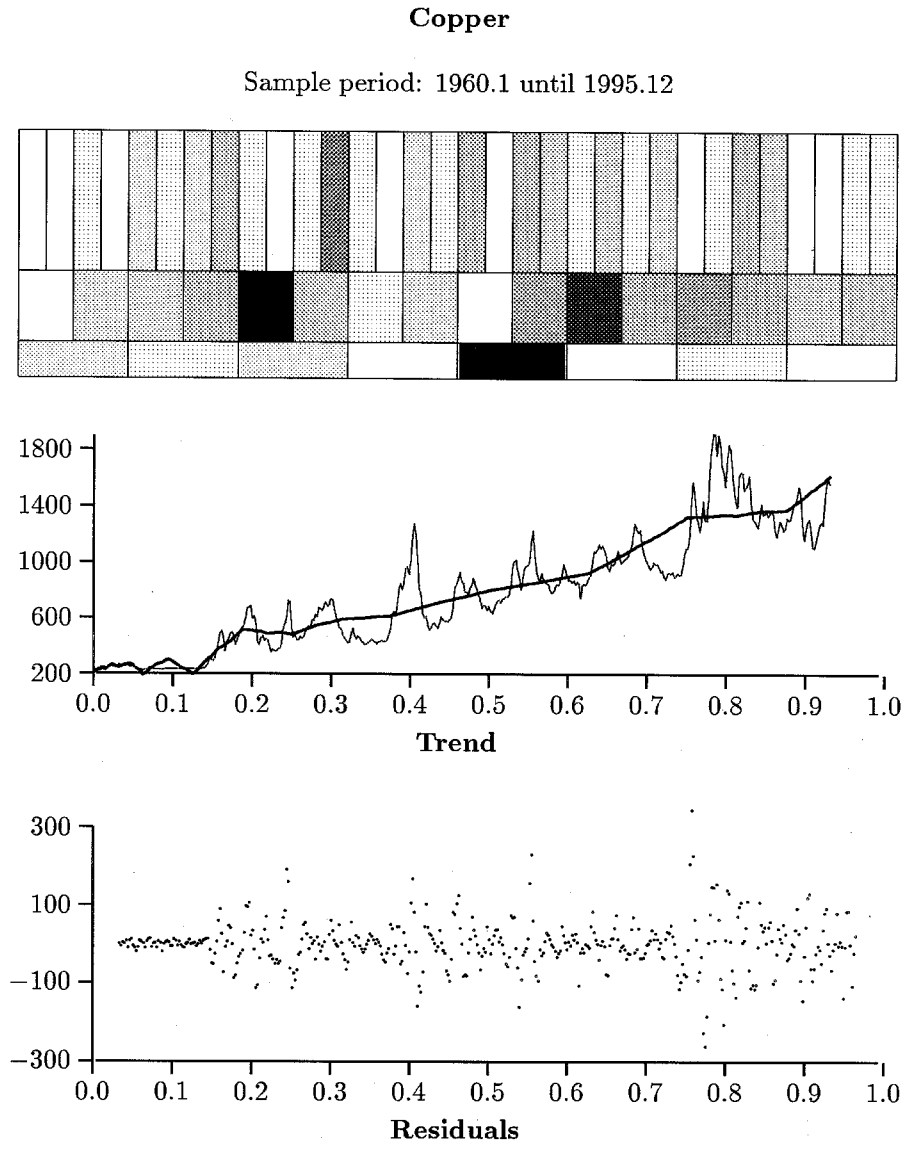
All the price series are tested for seasonality. Because of the international geographic dispersion of commodity markets and of related trading arbitrage, these prices typically do not have a marked seasonal component.

Earlier, we discussed the possibility that edge effects might be important in our wavelet analyses. Most of these appear to be accounted for satisfactorily by the inclusion of four additional regressors at each resolution, as discussed in the previous section. Another device appears to help obtain reasonable fits at the edge of the sample: a straightforward extension of the one used in Donoho, Johnstone, Keryacharian, and Picard (1995). The beginning and end of the sample are artificially extended by some number $m$ observations (we use $m = 16$) in the following way. The series, now of $n + 2m$ elements, is

$$y_m, y_{m-1}, \ldots, y_2, y_1, y_1, y_2, \ldots, y_{n-1}, y_n, y_n, y_{n-1}, \ldots, y_{n-m+1}.$$

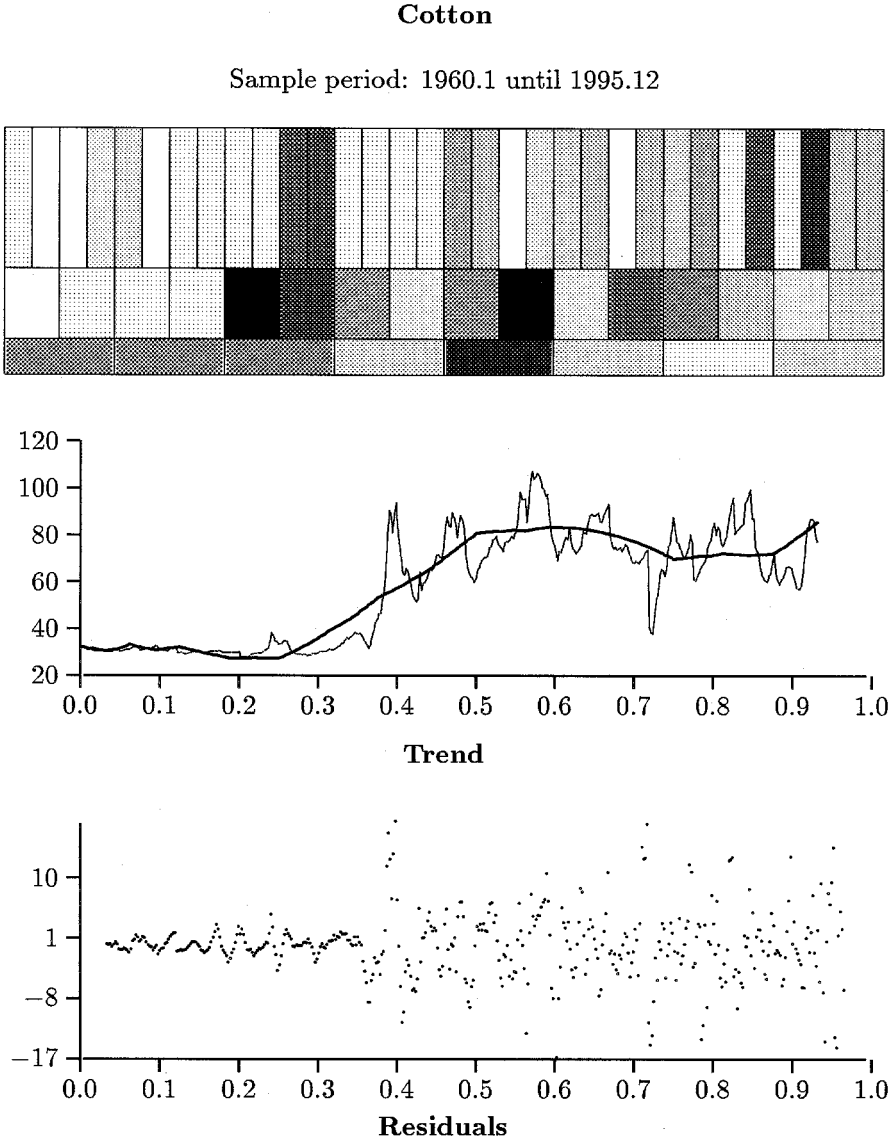
This ensures continuity at the edges, and, for all the series treated, seems to benefit the fit of the first few 'real' elements.

In Figures 1–4 we display graphically the results of the analysis. In the top portion of these displays, for four of the twenty-one commodities considerd, namely, coffee, copper, cotton, and rice, we present the patios showing the movement in the series at various resolutions in different sub-intervals. (Space considerations prevent including the patio plots for the other 17 commodities. For the interested reader, these plots, and the definitions of the series used in constructing them, are available through anonymous ftp at `russell.cnrs − mrs.fr`, in `pub/ondelettes`.) At the highest resolution, the sample period is divided into 32 subintervals, each of which



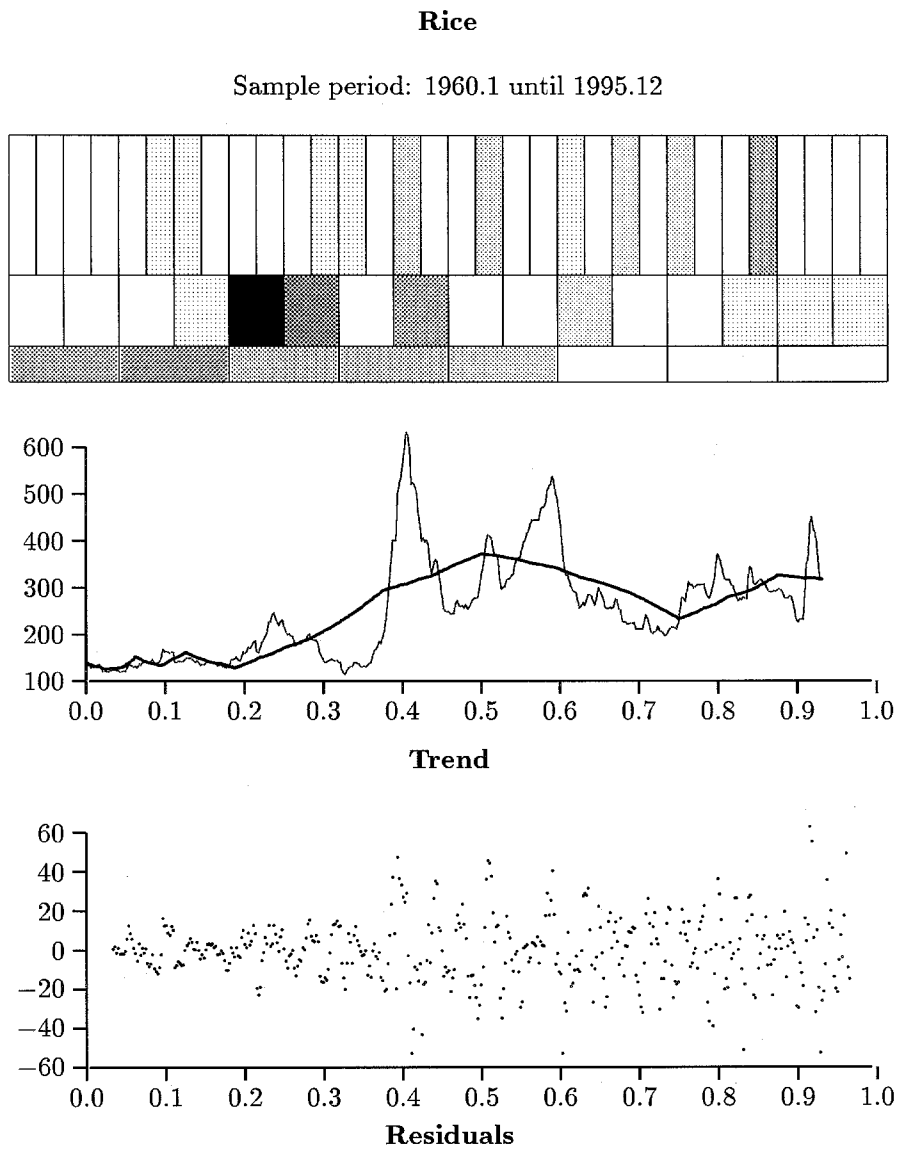

Proportion of variation at resolution 1 = 0.332
Proportion of variation at resolution 2 = 0.532
Proportion of variation at resolution 3 = 0.136
$R^2 = 0.851$

*Figure 2.*




Proportion of variation at resolution 1 = 0.353
Proportion of variation at resolution 2 = 0.439
Proportion of variation at resolution 3 = 0.207
$R^2 = 0.788$

*Figure 3.*




Proportion of variation at resolution 1 = 0.397
Proportion of variation at resolution 2 = 0.529
Proportion of variation at resolution 3 = 0.074
$R^2 = 0.946$

*Figure 4.*

is roughly a year long. The coarser resolution levels correspond to time scales of approximately 2 and 4 years.

The graph below the patio in each of these figures plots the actual series (fine line) and the trend, that is, the fit from the regression on the 12 father wavelets at the base level of resolution. The third plot in each display shows the residuals from the regression on the complete set of 80 wavelet regressors. These residuals represent movement at a level of resolution higher than one year. In most cases these residuals are evidently heteroskedastic, indicating that more information could be obtained by examining resolutions of 6-month and 3-month time scales. Further analysis with monthly data would soon lead to a perfect fit with no degrees of freedom. Finally, underneath the graphical information, we give the numerical proportions of the total variation accounted for at the various levels of resolution.

An analysis with resolution down to the 3-month level is carried out for just one commodity, coffee, and the results are shown in Figure 5. Only the two resolutions for which the sample period is subdivided into 64 and 128 subintervals are shown in this figure. These are approximately 6-month and 3-month time scales. The overall patio is now spread out over four lines, each corresponding to one quarter of the length of the sample period. The plot below shows the residuals after these two new levels have been accounted for; they may be compared with those of Figure 1.

## 6. Interpretation

The results of the previous section, in particular the patio plots, allow us to learn how commodity prices behave regarding both time location and time scale. For example, price cycles are often considered to display growing volatility over time. If so, we would expect to see darker tiles for later times at the resolution corresponding to the underlying cyclical behavior. And if the time scale of the cyclical behavior itself changes, we would expect to see the darker tiles shift from one resolution to another. Further, if cyclical behavior is present at more than one time scale, we should find dark tiles at more than one resolution. One may well observe that volatility increases over time for some time scales but decreases for others.

A shock, such as the oil shocks that clearly influence the series we study here, give rise to effects that may be more or less persistent. Such effects can be studied not only for their persistence in time but also for the time scale at which they operate. And, if detrending is an issue, the wavelet analysis allows us to distinguish between what we may call short and long run trends, once we have specified the time scales that we associate with the short or long run. If we suspect structural breaks, we would try to detect these by examining outliers. Just which observations are outliers, in the sense of being associated with large residuals, may depend on the finest level of resolution used. By comparing different resolutions, we may glean information concerning the nature of possible structural breaks.

At relatively fine levels of resolution, commodity price leads or lags may possibly be seen by comparing the relative positions of the darkest tiles for the various

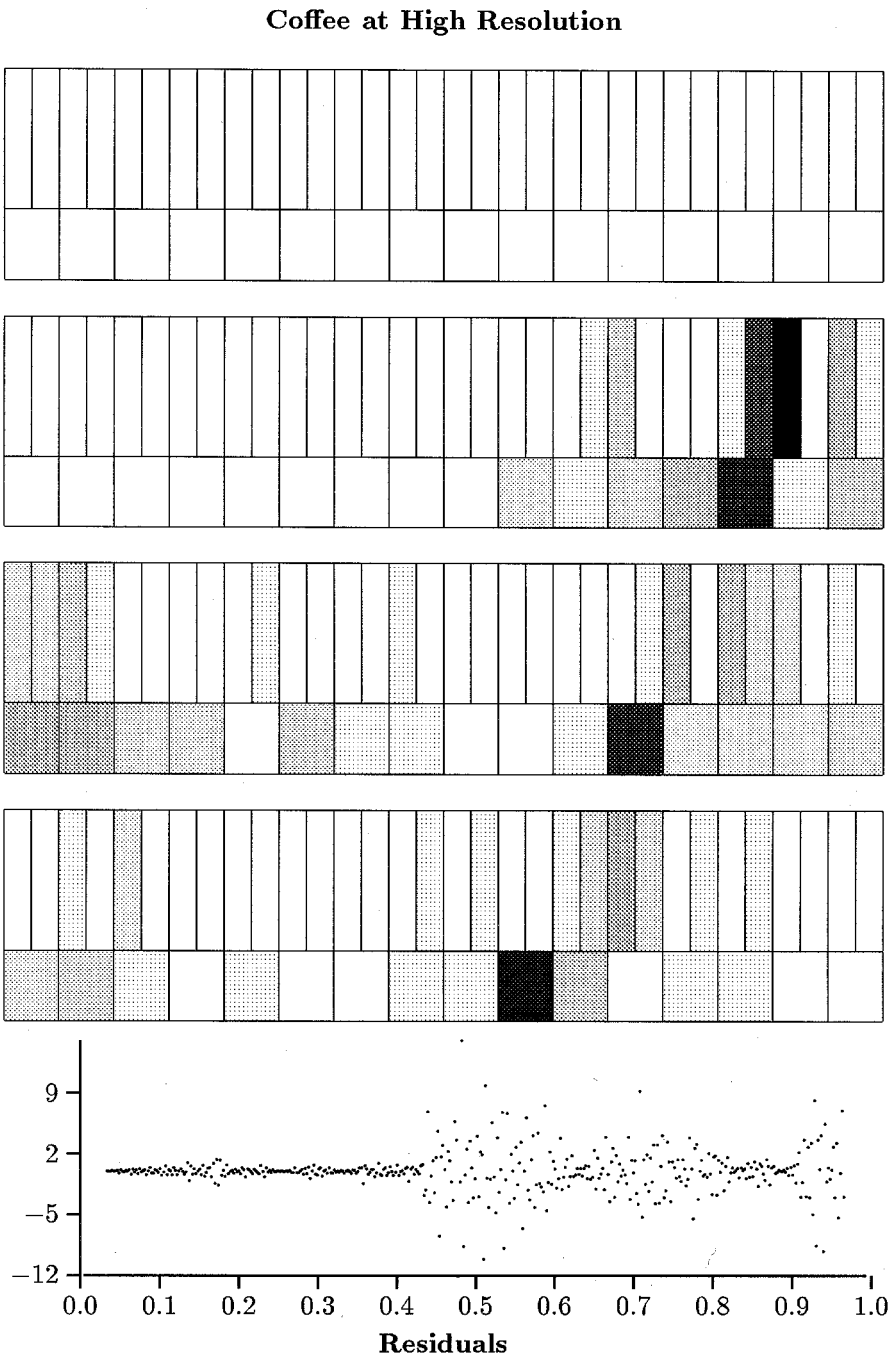

Coffee at High Resolution
9
2
−5
−12
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
Residuals

*Figure 5.*

price series. One price can be thought of as leading another if its dark tiles consistently precede those of the other. But if commodities have dark tiles at the same points in time, we may conclude that the prices of these commodities tend to move together.

Wavelet analyses *may* help to forecast commodity price movements. But one must be wary here, because the superior localisation provided by the compact support of wavelets means that no information is available outside of that support.

In the light of the above remarks, we can now look in greater detail at the properties of the twenty-one series analysed, bearing in mind that our ability to interpret patios is still rudimentary. First, some general observations. Considerable activity is observed in all the series at the annual and the next higher resolution during the market upheaval associated with the petroleum crises. Especially between 1973 and 1980, increased volatility in commodity price fluctuations is observed, though some calming appears at the end of this period. The relative shading of the tiles suggests that the intensity of the volatility is greater at different resolutions for different commodities. The fact that activity increases at the 2-year as well as at the annual resolution indicates that the effects of the upheaval over this period are more than transitory – price swings persist for more than two years before market forces produce a dampening effect. On the other hand, activity at the 4-year resolution normally appears to be more pronounced at the beginning of the period. This is probably due to the strong upward trend in prices starting late in the 1960's and peaking in 1974 or early 1975. At that point, commodity prices were higher than at any time since World War II.

These general observations change only slightly when we examine the patios of individual commodities. Beginning with the metals (exemplified here by copper), we find that 2-year activity appears strong for copper, gold, lead, tin, tungsten, and zinc. In time, the first or the fifth tile is the most important for most metals. At the 4-year resolution, there are differences among the metals. The upward trend in prices occurs earlier for tin, gold, silver, and tungsten than for copper or zinc. Copper, lead, and zinc also display similar patterns at 1-year resolution, whereas the periods of greatest activity for tin come later. This perhaps reflects the instability produced by the collapse of the tin buffer stock and the temporary suspension of tin trading on the LME.

The agricultural commodities (exemplified here by coffee, cotton, and rice) show greater dissimilarity in their patios. This is to be expected, in light of the differences in their patterns of production and consumption. For example, wheat, corn, and rice are grains grown annually. Coffee, cocoa, and tea are derived from perennial crops in order to produce beverages. Cotton, jute, rubber, and wood are raw materials, whose prices are affected by fluctuations in derived or industrial demand.

Among the grains, it appears that corn and wheat display strong movements at the four-year resolution, and this occurs early in the sample period. This may be because wheat was one of the commodities whose upward price movement actually

preceded the petroleum price jump. Soybean and wheat prices were particularly volatile during the market upheavals of the mid-seventies.

Movements at a resolution of two years are prominent for coffee, cocoa, sugar, and tea, as well as for rice, soybeans, and wheat. All but tea display volatility at the annual level, though once again it comes earlier for soybeans and wheat, perhaps because their price volatility is caused by climatic disturbances.

The raw materials, including cotton, jute, rubber, and wool show greater price activity at the longer-run or four-year resolution. This is an important result that confirms the relatively greater influence of business cycles on their demand. For petroleum, the greatest activity at the longer-run resolution comes early in the sample period, reflecting the upward shift in the price trend due to the imposed price shock. Volatility in short-term price movements comes later, reflecting the troublesome interval in which oil prices fluctuated downwards.

Our graphical results show that the commodity price volatility has not grown over time at any but the shortest of time scales and not for all commodities. Annual movements vary in intensity over the sample period, but they do not increase towards the end of the sample. At longer time scales, the most intense price movements are not associated with any particular frequency. The darkest tiles are to be found at different time scales at different periods of the sample. The residual plots, however, show that volatility at resolutions of less than a year increases markedly for almost all the commodities just after the oil shocks. Outliers occur often in the upheaval period of 1973–76, but they seem to have only short-term consequences. Lastly, differences do exist in the time locations of the most intense price activity for different commodities. Leads and lags among periods of intense activity can thus be observed as they were with tin. If this observation can be generalized, it would lend support to the traditional view that commodity prices do not move together.

## 7. Conclusions

In this paper, we have developed a semi-nonparametric approach to data analysis, based on wavelets, that to our knowledge has not been applied in other contexts. We make use of the fact that the father and mother wavelets used in multi-resolution analyses are simply new special functions, with properties that make them particularly attractive for analytical purposes. This allows us to overcome many limitations of the wavelet transforms conventionally used in signal analysis. The fast wavelet transforms that have been developed in recent years make it simple to compute the values of these new special functions for any argument. The regressors used in the semi-nonparametric regressions employed in this paper are simply vectors whose elements are such values of the wavelet functions. There is no restriction on the number of their elements, and there is no requirement for equal spacing.

The results obtained here indicate that our approach has considerable potential when applied to economic data. In particular, the patio plots provide valuable

information on both the time location and the time scale of price movements. Since this wavelet-based methodology is new, the types of interpretation given here are as yet tentative. However, because of the international importance of commodity price behavior, other econometric results are readily available that have helped us to interpret the numerical results we obtain. In particular, the wavelet analysis seems effective in analysing complex areas in price analysis that include:

(1) the tendency towards greater volatility or heteroskedasticity;
(2) the effect of random shocks in giving rise to outliers;
(3) the impact of market conditions in giving rise to structural breaks; and
(4) the possible common influence of economic events causing commodity prices to move or not to move together.

Many issues await further research. We have paid scant attention to treating edge effects – the techniques that worked well enough here may not do so in other contexts. A related issue is the precise choice of the time scales corresponding to different resolutions. Here we use the entire sample period divided into eight subintervals. We could have chosen the year as the base subinterval instead of the approximate year, and it would be interesting to see how our results would change if we did so. The Daubechies wavelet with $m = 2$ is adequate for our purposes, but other choices seem equally reasonable, and they should be investigated. The fact that our approach uses a semi-nonparametric regression means that standard inference procedures can be used for hypothesis testing. So far, we have done none of this. But, even so, we feel significant further knowledge has been added to our understanding of commodity price behavior.

## Appendix

### Data Description and Sources

COFFEE

Average of daily prices (secretariat of the International Coffee Organization, London). Robustas, weighted average of ex-dock New York (60 percent), Angola Ambriz 2 BB, Uganda standard, 1960.01–1995.12.

COPPER

London Metal Exchange, electrolytic wire bars, high grade, cash (*Metal Bulletin*, London), 1960.01–1995.12.

COTTON

Medium: U.S. Memphis Territory (medium staple), Middling 1-3/32. Prior to July 1981: S.M. 1-1/16 (USDA, Washington, DC), 1960.01–1995.12.

RICE

Thailand. White, 5% broken, and of month price, f.o.b. Bangkok, including export duty (IMF Secretariat, Washington, DC), 1960.01–1995.12.

## References


Ackerman, G. (1957). The cobweb theorem: a reconsideration, *Quarterly Journal of Economics* **71**, 151–160.

Antoniadis, A. (1994). Smoothing noisy data with tapered coiflets series, Discussion Paper RR 933-M-, LMC-IMAG, Grenoble.

Baumol, W.J. (1962). *Economic Dynamics*, New York, Macmillan Press.

Beylkin, G.R., Coifman, R., Daubechies, I., Mallat, S., Meyer, Y., Raphael, L. and Ruskai, B. (eds) (1992). *Wavelets and their Applications*, Cambridge Mass., Jones and Bartlett.

Blank, S.C. (1991). Chaos in future markets? A nonlinear dynamical analysis, *Journal of Futures Markets* **11**, 718–728.

Boldrin, M. and Woodford, M. (1990). Equilibrium models displaying endogenous fluctuations and chaos, *Journal of Monetary Economics* **25**, 189–222,

Brock, W. A. (1988). Nonlinearity and complex dynamics in economics and finance, in Anderson, P. W., Arrow, K. and Pines, D. (eds), *The Economy as an Evolving Complex System*, New York, Addison-Wesley.

Burton, M. (1993). Some illustrations of chaos in commodity models, *Journal of Agriculture Economics* **44**, 38–49.

Chavas, J. P. and Holt, M. T. (1991). On nonlinear dynamics: the case of the pork cycle, *Journal of the American Agriculture Economic Association* **73**, 820–828.

Chui, C. (1992a). *An Introduction to Wavelets*, New York, Academic Press.

Chui, C. (1992b). *Wavelets: a Tutorial in Theory and Applications*, Boston, Academic Press.

Combes, J. M., Grossman, A. and Tchamitchian, P. (eds) (1991). *Wavelets: Time-Frequency Methods and Phase Space*, New York, Springer-Verlag.

Daubechies, I. (1990). The wavelet transform, time-frequency localization and signal analysis, *IEEE Transactions on Information Theory* **36**, 961–1003.

Daubechies, I. (1992). *Ten Lectures on Wavelets*, Philadelphia, society for advanced mathematics.

Deaton, A. and Laroque, G. (1992). On the behavior of commodity prices, *Review of Economic Studies* **59**, 1–23.

Deaton, A. and Laroque, G. (1995). Estimating nonlinear rational expectations commodity price models, *Journal of Applied Econometrics* **10**, 510–540.

DeCoster, G.P., Labys, W. C. and Mitchell, D. W. (1992). Evidence of chaos in commodity futures prices, *Journal of Futures Markets* **12**, 291–305.

Donoho, D. L., Johnstone, I. M., Keryacharian, G. and Picard, D. (1995). Wavelet shrinkage: asymptopia?, *Journal of the Royal Statistical Society B* **57**, 301–369.

Frank, M. and Stengos, T. (1989). Measuring the strangeness of gold and silver rates of return, *Review of Economic Studies* **56**, 553–567.

Gandolfo, G. (1985). *Economic Dynamics*, Amsterdam, North Holland.

Gelb, A. H. (1979). A spectral analysis of coffee price oscillations, *International economic Review* **20**, 495–514.

Grossman, A. and Morlet, J. (1989). Decomposition of functions into wavelets of constant shape and related transforms, Working paper, CNRS-Luminy, Marseilles, France.

Hall, P. and Patil, P. (1996). On the choice of smoothing parameter, threshold and truncation in nonparametric regression by nonlinear wavelet methods, *Journal of the Royal Statistical Society B*, **58**, 361–377.

Houthakker, H. (1961). Systematic and random elements in short-term price movements, *American Economic Review* **51**, 164–172.

Jameson, M. H. (1983). Rational expectations and the U.S. hog cycle, in G. C. Rausser (ed) *New Directions in Econometric Practice*, Amsterdam, North Holland.

Jensen, R. V. and Urban, R. (1984). Chaotic price behavior in the nonlinear cobweb model, *Economics Letters* **15**, 235–240.
Labys, W. C., Elliot, C. and Rees, H. (1971). Copper price behavior and the London metal exchange, *Applied Economics* **3**, 99–113.
Labys, W. C. and Granger, C. W. J. (1970). *Speculation, Hedging, and Commodity Price Forecasts*, Lexington, Heath-Lexington Books.
Leuthold, R. M. (1972). Random walk and price trends: the live cattle futures market, *Journal of Finance* **27**, 879–889.
Mackey, M. C. (1989). Commodity price fluctuations: price dependent delays and nonlinearities as explanatory factors, *Journal of Economic Theory* **48**, 497–509.
Mandelbrot, B. B. (1963). The variation of certain speculative prices, *Journal of Business* **36**, 394–429.
Meyer, Y. (1990). *Ondelettes et opérateurs*, Vol. I and Vol. II, Paris Hermann.
Meyer, Y. (1992). *Wavelets and Operators*, Cambridge, Cambridge University Press.
Meyer, Y. and Coifman, R. R. (1991). *Ondelettes et opérateurs*, Vol. III, Paris, Hermann.
Nason, G. P. (1996). Wavelet shrinkage using cross-validation, *Journal of the Royal Statistical Society B*, **58**, 463–479.
Parikh, A. (1973). A study of coffee prices – a spectral approach, *Journal of Agricultural Economics*, 15–25.
Press, W. H., Teukolsky, S. A., Vetterling, W. T. and Flannery, B. P. (1992). *Numerical Recipes in C: the Art of Scientific Computing*, Cambridge, Cambridge University Press.
Priestley, M. B. (1995). Wavelets and time-dependent spectral analysis, *Journal of Time Series Analysis*, **17**, 85–103.
Rausser, G. C. and Cargill, F. (1970). The existence of Broiler cycles: an application of spectral analysis, *American Journal of Agricultural Economics* **32**, 109–121.
Samuelson, P. A. (1965). A random theory of future prices, *Industrial Management Review*, Vol 6.
Sims, C. A. (1984). Martingale-like behavior of prices and interest rates, discussion paper no. 205, Department of Economics, University of Minnesota, St. Paul.
Slade, M. (1981). Cycles in natural resource commodity prices: an analysis of the frequency domain, *Journal of Environment Economics and Management* **9**, 138–148.
Smidt, S. (1965). A test of the serial independence of price changes in soybean futures, *Food Research Institute Studies* **5**, 117–136.
Stevenson, R. A. and Bear, R. M. (1970/1977). Commodity futures: trends or random walk, as reprinted in A. Peck (ed), *Selected Writings on Futures Markets*, Chicago Board of Trade, 279–294.
Talpaz, H. (1974). Multi-frequency cobweb model: decomposition of the hog cycle, *American Journal of Agricultural Economics* **56**, 38–49.
UNCTAD (1996 and earlier issues). *Commodity Price Bulletin*, United Nations, Geneva.
Weiss, J. S. (1970). A spectral analysis of World Cocoa Prices, *American Journal of Agricultural Economics* **52**, 122–126.
Wickerhauser, M. V. (1994). *Adapted Wavelet Analysis: from Theory to Software*, Peters, Wellesley Press.
Working, H. (1958). A theory of anticipatory prices, *American Economic Review* **48**, 188–199.
Yang, S. R. and Brorsen, B. W. (1992). Nonlinear dynamics of daily cash prices, *Journal of Agricultural Economics* **74**, 706–715.