PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

The contribution of wavelets to the analysis of economic and financial data

James B. Ramsey

Phil. Trans. R. Soc. Lond. A 1999 **357**, 2593-2606
doi: 10.1098/rsta.1999.0450

THE ROYAL SOCIETY

The contribution of wavelets to the analysis of economic and financial data

By James B. Ramsey

Department of Economics, New York University, 269 Mercer St., New York, NY 10003, USA

After summarizing the properties of wavelets that are most likely to be useful in economic and financial analysis, the literature on the application of wavelet techniques in these fields is reviewed. Special attention is given to the potential for insights into the development of economic theory or the enhancement of our understanding of economic phenomena. The paper is concluded with a section containing speculations about the relevance of wavelet analysis to economic and financial time-series given the experience to date. This discussion includes some suggestions about improving our understanding and evaluation of forecasts using a wavelet approach.

Keywords: wavelets; economics; finance; time-scale; forecasting; non-stationarity

1. Introduction

Imagine a dynamical system that is characterized by the decentralized interactions of a large number of participants. Participants make decisions about their future behaviour facing enormous degrees of freedom using only local information over both time and space. Imagine further that this system is constantly evolving, open and non-isolated so that its performance over time reflects its reactions to external stimuli and even the manipulation of the fundamental relationships of the system. Imagine yet again that attempts to control the system generate anticipatory reactions by the participants, each of whom is trying to optimize their position within the system. Imagine also that each agent is operating on several time-scales at once, so that not only current events affect behaviour, but the distant past as well as the agent's anticipation of the distant future.

If you can imagine this challenge to the analyst's art, you have imagined the difficulties that face the empirical economist. As a result of these the statistical and analytical tools that work so well in other fields have limited success in economic contexts, and then perform best only as short-term approximations to more complex phenomena. The econometrician, in order to achieve some limited success, must attempt to reconcile the irreconcilable. Even though the economy is constantly evolving, one must often assume in order to forecast that the system is at least locally constant. The econometrician must assume that only a few variables are, as a practical matter, involved in any relationship. Most often, it is the violation of this technically necessary assumption that produces unanticipated and sometimes dramatic shifts in observed economic relationships.

In such a context, the search for useful analytical tools is paramount. Key issues to be considered by the putative analyst are robustness of procedure to erroneous

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

assumptions, flexibility of regression fit to deal with imprecise model formulations, the ability to handle complex relationships, efficiency of the estimators to be able to make useful distinctions on few data points, and simplicity of implementation. This, the econometrician's 'Holy Grail', may be impossible to achieve, but is nevertheless a worthwhile overall objective.

With this background, the potential promised by wavelets is readily apparent. While wavelets do not meet all the criteria required of the 'Holy Grail', they meet some of them sufficiently well to indicate that wavelets may well be able to provide new insights into the analysis of economic and financial data.

One of the first benefits of a wavelet approach is the flexibility in handling very irregular data series, as illustrated in Donoho *et al*. (1995). Given my remarks above, it is no surprise that the ability to represent highly complex structures without knowing the underlying functional form is of great benefit in economic and financial research. A corollary facility that wavelets possess is that of being able to precisely locate discontinuities, and isolated shocks to the dynamical system, in time regime shifts. Further, it is vital that the process of representation should be able to deal with the non-stationarity of the stochastic innovations that are inevitably involved with economic and financial time-series.

A critical innovation in estimation that was introduced by wavelets, although by no means necessarily restricted to wavelets, is the idea of shrinkage. Traditionally in economic analysis the assumption has universally been made that the signal, $f(t)$, is smooth and the innovations, $\varepsilon(t)$, are irregular. Consequently, it is a natural first step to consider extracting the signal $f(t)$ from the observed signal $y(t) = f(t)+\varepsilon(t)$ by locally smoothing $y(t)$. However, when the signal is as, or even more, irregular than the noise, such a procedure no longer provides a useful approximation to the signal. The process of smoothing to remove the contamination of noise distorts the appearance of the signal itself. When the noise is below a threshold and the signal variation is well above the threshold, one can isolate the signal from the noise component by selectively shrinking the wavelet coefficient estimates (Donoho & Johnstone 1995; Donoho *et al*. 1995).

The most important property of wavelets for economic analysis is decomposition by time-scale. Economic and financial systems, like many other systems, contain variables that operate on a variety of time-scales simultaneously so that the relationships between variables may well differ across time-scales.

The remainder of this paper is in two sections. The first provides a selective review of the use of wavelets in the analysis of economic and financial data. The second section speculates on the potential for wavelet analysis in the context of economic and financial data.

2. A selective review of the literature

While applications in economics have not yet extensively used the special properties of wavelets, there have been a number of interesting applications and the list is sure to grow. We might fruitfully discuss the various applications in terms of four main categories. The first and largest category in terms of numbers of papers is that emphasizing the role of non-stationarity and the ability to handle complex functions in Besov space. The second category is those papers that are most concerned with structural change and the role of local phenomena. A third involves the use of time-scale decompositions and the last is directly concerned with forecasting. I will conclude this section with a few comments about some papers that do not easily fit into any of these categories.

(*a*) *Non-stationarity and complex functions*

A first paper that warrants discussion uses applications in biology, but the statistical and methodological issues are quintessentially similar to those in economics. Von Sachs & MacGibbon (unpublished research) derived the bias and variance for wavelet coefficient estimators under the assumption of local stationarity; this paper is an extension of an earlier paper by Johnstone & Silverman (1997), who analysed the statistical properties of wavelet coefficient estimators under the assumption of stationary, but correlated, data of both short and long run. Further, von Sachs & MacGibbon (1997) incorporated into their analysis pulsatile components and simultaneously allowed for non-stationarity. An important assumption in their analysis is that while the non-stationarity is slowly changing over the entire sample period, it does so in such a manner that more observations per unit time would lead to locally asymptotic convergence of the estimators. The authors demonstrate the asymptotic properties of their estimators. In particular, they suggest under these circumstances the use of a local MAD estimator to estimate the variability of the wavelet coefficients in segments of quasi-stationarity within each level. The importance of this paper stems from the fact that the single most obvious variation in economic data, especially after first differencing, is the presence of second-order non-stationarity, which is sometimes modelled as an ARCH process; for a review in the context of financial data see Bollerslev *et al*. (1990).

A nearly contemporaneous article on a similar topic is that by Gao (unpublished research), who contemplated the model

$$y_i = f(t_i) + \sigma(t_i) z_i, \tag{2.1}$$

where $\{z_i\}$ is distributed as i.i.d. Gaussian noise. The model proposed here is a simplification of that considered by von Sachs & MacGibbon. The model is motivated by noting that heteroscedasticity arises from converting unequally spaced data to equally spaced data. When the variances are known the process of obtaining suitable estimators is straightforward, but when the variances are unknown, they must be estimated along with the wavelet coefficients. Gao's procedure extends Donoho & Johnstone (1997). Gao considers four alternative definitions of the shrinkage rule, the usual soft and hard rules, an intermediate rule that he labels 'firm', and the garrote. Under the usual assumptions supplemented by the knowledge that the variances are heteroscedastic, the wavelet coefficients w, have the distribution

$$w = Hy \sim N(\Theta, HD^2H') \tag{2.2}$$

where D is the covariance matrix for the model, H is the wavelet transform and $\Theta = Hf$. Under these conditions the near optimality in terms of expected mean-squared-error risk of the waveshrink estimators continues to hold.

Gao's solution when the variances are unknown is to use a non-decimated wavelet transform on y to obtain the variation at the finest level of detail. He next applies a running MAD estimator to obtain estimates for the standard deviation. These estimators are not distributed as Gaussian, but approximately as Student t with heavier tails. Near-optimal mean-squared-error expected risk also holds in this situation.

Pan & Wang (1998) introduce a novel approach when the data generating mechanism must be regarded as 'evolutionary', so that the wavelet coefficients are varying over time; the model is

$$y_t = W_t^{\mathrm{T}} w_t + \epsilon_t. \tag{2.3}$$

The authors interpret the model in terms of a Kalman filter, so that the pair of equations that define the time path of coefficients is given by

$$\left.\begin{aligned} y_t &= W_t^{\mathrm{T}} w_t + \epsilon_t, \\ w_t &= w_{t-1} + V_t, \end{aligned}\right\} \tag{2.4}$$

where V_t is the time-varying covariation for the wavelet coefficients. The model was applied to the monthly stock price index, y_t, that was regarded as some function $\tilde{f}(r_t)$, where r_t is the dividend yield. $\tilde{f}(r_t)$ was represented in the model by the wavelet transform $W_t^{\mathrm{T}} w_t$. The empirical results were encouraging both as to the overall degree of approximation and to the extent that turning points were correctly indicated.

An article that represents research arising from earlier concerns in the analysis of the stock market is Ramsey *et al*. (1995). In this article, the chief topics of interest were the degree of statistical self-similarity of the daily stock return data and whether there was any evidence of quasi-periodicity. The wavelet analysis indicated that while there was little evidence of scaling in the data, there was surprisingly clear evidence for quasi-periodic sequences of shocks to the system. Thus, these results confirm what is by now a commonplace statement about this type of financial data (see, for example, Ramsey & Thomson 1998); they are very complex and are more structured than mere representations of Brownian motion.

The last two papers to be discussed in this section introduce a useful generalization to wavelet analysis in those cases where one wishes to begin the analysis in a purely exploratory phase. In Ramsey & Zhang (1996, 1997), highly redundant representations in terms of 'waveform dictionaries' were used (see Mallat & Zhang 1993). The approach taken in these papers is one that generalizes both wavelets and Fourier analysis. For a given choice of scaling function $g(t)\epsilon L^2(R)$, scale parameter s, position in time u and a given frequency modulation ξ, we define the time-frequency atom by

$$g_\gamma(t) = \frac{1}{\sqrt{s}} g\left(\frac{t-u}{s}\right) \mathrm{e}^{\mathrm{i}\xi t}, \tag{2.5}$$

$$\gamma = (s, u, \xi). \tag{2.6}$$

The function $g_\gamma(t)$ is normalized to have norm 1 by the component $1/\sqrt{s}$. The function $g_\gamma(t)$ is centred at the abscissa u, where its energy is concentrated in a region that is proportional to s. The Fourier transform of $g_\gamma(t)$ is centred at the frequency ξ and has its energy concentrated in a neighbourhood of ξ with size proportional to $1/s$. Matching pursuit algorithms are used to obtain the minimum squared error collection of atoms that will most parsimoniously represent the function.

Ramsey & Zhang (1996) examined the 16 384 daily observations of the Standard and Poor's 500 stock-price index on the New York Stock Exchange from January 3rd 1928 to November 18th 1988. The time-frequency distributions indicate virtually no power for any frequencies, although there is evidence that some frequencies wax and wane in strength. However, most of the power is in time-localized bursts of activity. The bursts do not appear to be isolated Dirac delta functions, but highly localized chirps that are characterized by a rapid build-up of amplitude of signal and rapid oscillation in frequency. A plot of the squared weights of the coefficients relative to those that would be obtained from random data indicate very strong approximations with relatively few coefficients. What is even more striking is that if we restrict our attention in the time-frequency plots to those scales that will produce the highest resolution of the frequencies we observe very clear evidence of quasi-periodicity and regularity in the data.

Ramsey & Zhang (1997) applied similar techniques to tic-by-tic foreign exchange rates worldwide for a year. The three exchange rates so examined were the Deutschmark–US dollar, the Yen–US dollar and the Yen–Deutschmark.

The results obtained were qualitatively similar for all three exchange rates. Both the levels and the first differenced data were examined, because there has been some controversy in the economics literature about the appropriate data-generating mechanism; the presence of a unit root in the data being of great concern. In analysing the levels data using the waveform-dictionary approach, some evidence of structure was discovered, but only with very low power. As discovered in the stock-market data, there was evidence for frequencies that waxed and waned over the year. However, most of the power seems to be in localized frequency bursts. The economic importance of this result is that instead of viewing the transmission of information in the foreign exchange market as being represented by a swift, almost effortless, adjustment, these results indicate that there is a period of adjustment of some noticeable length and that the implied frequencies of oscillation build up and decay as in a chirp. Another key insight provided by these data is that despite the relatively low number of atoms needed to provide a very good approximation to the data, about 100 is sufficient, there is little opportunity for improved forecasting. This is because, while relatively few structures are needed to represent the data, the bulk of the power is in chirps and there does not seem to be any way of predicting the occurrence of the chirps. In short, as most of the energy of the system is in randomly occurring local behaviour, there is little opportunity to improve one's forecasts.

(*b*) *Time-scale decompositions*

For cognate disciplines such as economics and biology, one of the most useful properties of the wavelet approach is the ability to decompose any signal into its time-scale components. It is well known that physical and biological processes are phenomenalogically different across different time-scales. In economics as well, one must allow for quite different behaviour across time-scales. A simple example will illustrate the concept. In the market for securities there are traders who take a very long view (years in fact) and consequently concentrate on what are termed 'market fundamentals'; these traders ignore ephemeral phenomena. In contrast, other traders are trading on a much shorter time-scale and as such are interested in temporary deviations of the market from its long-term growth path; their decisions have a time horizon of a few months to a year. And yet other traders are in the market for whom a day is a long time.

An effort along these lines is illustrated in Davidson *et al*. (1997), who investigated US commodity prices. Even though the differences across scales were not pursued fully, the authors did consider the different properties of the wavelet coefficients across scales and calculated a measure of the relative importance of the coefficients between scales.

In two related papers (Ramsey & Lampart 1998*a*, *b*), a very different approach was taken to the use of scale in wavelet analysis. In both of those papers the concept was that the relationship between two variables may well vary across scales. For example, one may have a simple linear relationship between aggregate income and consumption, but that the coefficient relating consumption to income might well differ across scales. The objective in the two papers mentioned was to examine this issue in the context of two relationships; one was that between personal income and consumption and the other was that between money and gross domestic product, the so-called 'money velocity' relationship. Wavelets were used to provide both an orthogonal time-scale decomposition of the data and a non-parametric representation of each individual time-series. At each scale a regression was run between consumption and income, or between money and income, and the results were compared across scales. For example, using the consumption–income relationship, we have

$$\left.\begin{array}{ll} C[S_J]_t = \alpha_J + \beta_J Y[S_J]_t + \varepsilon_t, & \text{or} \\ C[D_j]_t = \alpha_j + \beta_j Y[D_j]_t + \varepsilon_t, & j = 1, 2, \ldots, J-1, \end{array}\right\} \tag{2.7}$$

where $C[S_J]_t$ and $Y[S_J]_t$ represent the components of consumption and income at the highest scale and $C[D_j]_t$ and $Y[D_j]_t$ represents consumption and income at intermediate scales.

The results were productive and informative. First, the relationship between economic variables does vary significantly across scales; that is, statistically significant differences were discovered in the estimated values of the coefficients β_j. Second, the decomposition resolved some empirical anomalies that existed in the literature. Third, the decomposition indicated that the delay in the relationship between two variables might well be a function of the state space; a result that heretofore had not been suspected. More specifically, the authors discovered that the slope coefficient relating consumption and income declines with scale, a theoretically pleasing result; and that the role of the real interest rate in the consumption–income relationship is strong and of the theoretically correct sign for the longest time-scales, but is insignificant for the shortest scales; which is another theoretically plausible result.

However, a more far-reaching result and one that has not been anticipated in the literature is that the phase of the relationship between two variables ordered in time, or between the scale components of two variables, may well vary with the state of the system. In economics a conventional assumption is that two variables, say consumption and income, are related contemporaneously; a secondary assumption that is occasionally invoked is that there is a delay between the two variables as there often is between stimulus and response. But what Ramsey & Lampart (1998*a*, *b*) discovered is that the delay between two variables might well be a function of the state of the system; that is

$$C[D_j]_t = \alpha + \beta Y[D_j]_{t-d[X_t]} + \varepsilon_t,$$

where X_t represents some component of the dynamical system within which the consumption–income relationship is embedded. An example of this in our current context is provided by recognizing that the 'timing of consumption, or certainly

Table 1. *Granger causality tests: I*

(Results of Granger causality tests on individual crystals and log differenced data for M1 and nominal personal income (NPI): P values in parentheses.)

		null hypotheses	
	results	M1 $\Rightarrow$ NPI	NPI $\Rightarrow$ M1
D6 (5 lags)	feedback	6.398 (0.000)	6.968 0.000)
D5 (20 lags)	feedback	4.491 (0.000)	5.242 (0.000)
D4 (19 lags)	M1 $\Longrightarrow$ NPI	3.334 (0.000)	0.809 (0.695)
D3 (17 lags)	M1 $\Longrightarrow$ NPI	1.838 (0.023)	1.294 (0.193)
D2 (23 lags)	M1 $\Longrightarrow$ NPI	5.620 (0.000)	1.146 (0.293)
D1 (14 lags)	M1 $\Longleftarrow$ NPI	1.558 (0.089)	5.194 (0.000)
log diff. (12 lags)	inconclusive	0.534 (0.892)	2.063 (0.186)

Table 2. *Granger causality tests: II*

(F-tests of Granger causality between M1 and nominal personal income (NPI) at [D4] across phase shifts: P values in parentheses.)

	in–out	out–in	in–in
phase shifts	4/1962–8/1967	8/1967–5/1974	4/1962–5/1974
time period	M1 $\Longrightarrow$ NPI	M1 $\Longleftarrow$ NPI4	inconclusive
NPI $\Rightarrow$ M1	6.401 (0.00)	0.997 (0.434)	0.643 (0.696)
M1 $\Rightarrow$ NPI	1.465 (0.208)	2.714 (0.020)	0.933 (0.474)

the timing of a purchase' is as much an economic decision as is the amount of the purchase.

A long-standing debate in the economics literature concerns the direction of causality between money and income; essentially does money cause income, or does income cause money? Prior empirical research had obtained conflicting results. Ramsey & Lampart (1998a, b) indicated that the decomposition of money and income into their corresponding time-scale components followed by performing a sequence of regressions between money and income at each scale level helped to resolve this debate. The main result is that the direction of the relationship depends on the level of the time-scale in a clear manner for each scale; but that the relationship between the aggregates over all time-scales is ambiguous. These results are illustrated in table 1.

Further, at the D4 scale level the 'causality tests' were ambiguous in that the results depended on the pattern of the phase relationship between money and income; that is, if money and income are moving into phase the causality goes one way, in the opposite direction if moving out of phase, and the results are anomalous if the mixture of phase relationships is random. These results are illustrated in table 2.

(c) *Forecasting*

There are only two major papers in this section, notwithstanding the importance of the topic. The first approach is discussed in Ariño (unpublished research). The main idea is relatively simple and direct. Ariño's approach is to decompose the signal into its time-scale components and then to treat each approximation at each time-scale as a separate series. He provides forecasts for each component by using a regular ARIMA formulation. The final forecast for the complete series is obtained by adding up the component forecasts. Ariño restricts his attention to three components: trend, seasonal fluctuations and noise. The technique is applied to Spanish concrete production and car sales with interesting results that indicate the procedure has potential. Ariño's paper leaves numerous distributional issues unaddressed, but the experiment promises to be useful.

The next approach is presented in a series of papers by Alex Aussem, Fionn Murtagh and co-workers (see, for example, Aussem & Murtagh 1997; Aussem *et al.* 1998). The major innovation introduced by these authors is to analyse the individual time decompositions by neural networks and to base the forecasts on the neural-network estimates. The wavelet transform used was the isotropic non-orthogonal linear à-trous wavelet.

As with Ariño, the forecast for the complete series is obtained by adding up the individual forecasts. The technique is applied to the famous sunspot series and to the daily Standard and Poor's stock-market series. For the latter the authors' provide five day ahead forecasts. As with the work reported by Ariño, the distributional properties of the combined forecasts obtained by adding up the individual time-scale-based forecasts is still an open issue, but the results to date are of interest and indicate substantial potential.

(d) *Some miscellaneous papers*

This subsection contains a discussion of some miscellaneous, but important, examples of the use of wavelets in economic analysis.

(i) *Density estimation*

The estimation of density functions has been a tradition in economics and the debate concerning the distribution of income and wealth in particular is of great topical concern. Further, because of such factors as 'minimum-wage' laws, 'poverty boundaries' or discontinuous tax relationships, income distributions frequently exhibit anomalous local behaviour that is best analysed in the context of wavelets. Härdle *et al.* (1998) illustrate the benefits of using a wavelet approach to the estimation of densities relative to the standard kernel smoothing procedures.

A recent and important line of research in economics is the analysis of fractionally integrated models; that is, models of the form

$$\Phi(L)(1-L)^{d}(x(t)-\mu)=\theta(L)\epsilon(t), \tag{2.8}$$

where $\Phi(L)$ and $\theta(L)$ are polynomials in the lag operator L, $x(t)$ and $\epsilon(t)$ are stochastic processes and d is the fractional differencing coefficient such that $|d| < 0.5$. Jensen (1997*a*, *b*) in a recent series of articles has explored the use of wavelets in the estimation of the fractional differencing coefficient, d. Jensen has several reasons for choosing wavelets as a basis for the analysis of the long-memory processes defined by the ARFIMA models. The major reasons are that the wavelet-based calculations are far less intensive to calculate than the regular exact MLE, the estimates are much more robust to modelling errors (the orders of the polynomial lags in equation (2.8) are unknown), and the estimates are robust to not knowing the mean of the process. A key element in the computational gains achieved using wavelets is the relative sparsity of the wavelet coefficients, especially for the longer time-scale components of the process.

(ii) *Structural change*

Given the concern by economists about the potential effects of sudden regime changes and isolated shocks to the system it is surprising that more research on this subject has not been carried out so far. An exception is a paper by Gilbert (unpublished research), who looked for evidence of structural change using Haar wavelets as a basis. The data used were a set of macroeconomic indicators that are seasonally adjusted, some series were quarterly and some were monthly. Using these procedures, only the inflation rate index and the index of nominal GNP growth gave any indication of an abrupt structural change.

Another example that illustrates attempts to take advantage of the ability of wavelets to resolve local features is provided by the manuscript by Jungeilges (unpublished research). The major concern in this manuscript is to provide improvements on the specification error test procedures created by Ramsey (1969) and refined in Ramsey & Schmidt (1976) and the corresponding tests generated by White (1989). The latter tests used neural networks as a device for providing semiparametric representations of any time-series. The essential idea underlying Jungeilges's manuscript is that by using wavelets to provide a robust representation of the presumed structure of the data the residual variances should be uncorrelated and distributed identically with equal variances across the sample. Unfortunately, although the test designed by Jungeilges has power against many alternative errors, it is not as powerful in general as the original tests provided by Ramsey (1969). This is one of the few examples where the use of wavelets has not, at least marginally, led to an improvement in test results or in the efficiency of estimators. The difficulty resides in the author's design of the test in that by relying mainly on the effect of the specification errors on the *heteroscedasticity* of the residuals, rather than on the *structure* of the time varying conditional mean, the author lost both power and robustness.

(iii) *Miscellaneous papers*

Spectral techniques have long been a useful, but relatively neglected, tool in economic analysis. Three important papers on this topic are Neumann (1996), Chiann & Morettin (unpublished research) and Lee & Hong (unpublished research). Neumann (1996) derived the asymptotic normal distribution of empirical wavelet coefficients used to estimate the spectral density function of a non-Gaussian process. The wavelet estimator is shown to be superior to kernel estimators in terms of mean-squared-error criteria, especially when there are spatially inhomogeneous features in the spectral density. Lee & Hong advanced this literature by proposing a consistent test for serial correlation of unknown form by wavelet methods. The simulation results confirm the theoretical prediction that when there are spatial inhomogeneities the wavelet estimator is to be preferred and when there are no spatial inhomogeneities the kernel estimator is preferred.

Finally, mention should be made of a nearly unique approach to the use of wavelets to analyse economic data. In Morehart *et al*. (1999) wavelets are used to model the *geographic distribution* of economic and financial measures used in agriculture. The idea is to provide a non-parametric framework within which multivariate spatial relationships can be described. The authors base their analysis on the non-decimated à trous wavelet. As indicated by the authors, the use of non-decimated wavelets enhances the detection of local features and anomalies. There is one other paper that uses wavelets to analyse 'spatial' differences, namely Chen & Conley (unpublished research). In this paper, the statistical context is one involving panel data where the number of observations over time is relatively short. The authors' approach is to compensate for the lack of time-series data by using information on the spatial distribution of agents. Agents that are spatially close together can be assumed to behave in a similar manner, so that a 'cluster' of agents will almost provide a set of repeated time-series experiments. The empirical results are summarized in terms of estimates of the spectral density function.

3. The research potential

In this section I speculate on the potential for wavelets in the analysis of economic and financial data. The comments can be grouped into four categories of applications.

(*a*) *Exploratory analysis: time-scale versus frequency*

Waveform dictionaries provide an excellent exploratory tool where one is particularly interested in determining the relative importance of frequency components to time-scale components. In economics and finance a preliminary examination of the data in order to assess the presence and the ebb and flow of frequency components is useful. In particular, little attention has been paid so far to the idea that frequency components may appear for a while, disappear and then reappear. Conventional spectral tools would most likely miss such frequency components altogether. Further, as will be argued below, in economics and finance, a matter of considerable concern is to separate the local from the global, the ephemeral from the permanent; wavelets provide the best tool for separating out these effects.

(*b*) *Density estimation and local inhomogeneity*

We have seen in the discussion above that in general wavelets are superior, say to kernel estimators, whenever there are local inhomogeneities. I argued in the introduction that a prominent characteristic of economic data is the presence of spatial inhomogeneities, so that one quickly concludes that wavelets are a natural choice for analysing such data. At the very least, wavelets should be used as a first effort in the exploratory stages of analysis.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

As a particular example of the argument above, the estimation of density and spectral density functions in economics and finance requires the recognition of the presence of spatial inhomogeneities. Further, these inhomogeneities are of theoretical and policy importance, so that one wants to do much more than merely note their presence. One will want to describe the inhomogeneity effects in detail. For example, the effects of minimum-wage legislation are debated frequently and at length. In order to assess the effects of minimum-wage legislation on the income distribution one requires to be able to isolate and describe analytically the spatial inhomogeneity that such legislation creates. Other inhomogeneities are created by tax legislation, by rigidities in trading rules on exchanges and commodity markets and perhaps by the process of innovation itself.

(*c*) *Time-scale decomposition*

I now come to one of the most promising opportunities for the use of wavelets in economics, decomposition by time-scale. I demonstrated above how insights were gained and the resolution of apparent statistical anomalies were removed by recognizing the potential for relationships between variables to be at the scale level, not the aggregate level. That is, one should recognize that the relationship between consumption and income varies depending on the time-scale that is involved. Much work remains to be done in this context. Gains can be made if interest is restricted to a specific scale by using non-decimated wavelets, but if one wants to observe the data at a sequence of alternative scales and to preserve orthogonality, or at least near orthogonality, of the time-scale decompositions, alternative formulations will be needed. One thought is to pursue the work of Chui *et al*. (1995), in which expansions are in terms of the scaling functions

$$\left.\begin{aligned}\phi_k^{j,n}(t) &= (2^j\alpha_n)^{1/2}\phi(2^j\alpha_n t - k),\\ \psi_k^{j,n}(t) &= (2^j\alpha_n)^{1/2}\psi(2^j\alpha_n t - k),\end{aligned}\right\} \tag{3.1}$$

and

$$\alpha_n = \frac{2^N}{n+2^N}, \quad N>0, \quad n=1,2,\dots,2^N-1$$

provides 2^N-1 additional levels between any two consecutive octave scales. For a given application, there may be a preferred value for α_n that provides the most useful time-scale decomposition.

In this context, the investigation of the potentially time-varying phase relationships between the variables of interest is an important and until now neglected aspect of empirical analysis. Wavelet decompositions are of course uniquely well suited to answering such questions. It is likely that many examples of 'complex relationships between variables' may well on further examination be found to be caused by shifting phase relationships. In short, this discovery has opened up a substantial area for future research.

(*d*) *Aspects for forecasting*

The last and equally important future use of wavelets is in the role of improving forecasts. There are in fact several characteristics of wavelets that may well enhance the forecasting of economic and even financial time-series. First, as we have already seen, gains can be made by decomposing the series to be forecasted into its time-scale components and devising appropriate forecasting strategies for each. In fact, this approach has a long history in forecasting; the wavelet approach has merely formalized old notions of forecasting trend, seasonal and business cycle components. What is of far greater importance, I believe, is that wavelets enable one to isolate the local from the global. In forecasting, one attempts to forecast the permanent components of the series, not the strictly local events. In so far as wavelets enable one to separate out the two effects, one will be able to improve forecasts by eliminating the contamination from strictly local non-recurrent events.

A particular suggestion for the enhancement of time-series forecasts in the context of wavelet analysis is to capitalize on the time-scale decomposition by devising 'test' strategies to check for variations in the structure at the time-scale analysed. At the level of the overall trend using the father wavelet, one might well speculate that if there is to be a structural change away from the predicted long-term path, that change can be characterized initially in terms of an exponential growth path either above or below the current forecast path. Consequently, one can devise a specific test strategy to test for exponential deviations away from the current path. In principle this strategy should provide a more powerful test for deviations from the trend and provide an 'earlier warning' system than is currently possible. For the higher-frequency oscillations, techniques based on detecting changes in frequency as illustrated by the Doppler example of Donoho & Johnstone (1995) is a likely beginning. In any event, one major task facing all such suggestions is to evaluate the distributional properties of the recombined forecast estimates from the individual time-scale forecasts.

The wavelet approach also facilitates an important, but frequently neglected, aspect of forecasting; the limits to forecasting. Many of us forget that not all series can be forecasted, even in principle, and that often there may well be severe limitations on the extent of the forecast horizon. Using wavelets we can resolve these issues at each scale level. It is likely, for example, at the lowest detail level that noise will predominate so that there is no opportunity for forecasting at all. At a higher level we may discover that the maximum period of forecastability is commensurate with the time-scale and not beyond. At even higher scales we may recognize that we need to separate the long-term behaviour from the episodic occurrences of local phenomena. In short, the task of forecasting is considerably more complex than has been generally recognized.

4. Conclusion

In conclusion, it is apparent that wavelets are particularly well adapted to the vagaries of the statistical analysis of economic and financial data. Only the surface has been explored so far and the future indicates that some exciting and revealing analysis will be generated by the application of wavelets. The potential benefits are far greater than the mere application of new techniques. As I have tried to indicate, wavelet-based analysis should lead to new insights into, and novel theories about, economic and financial phenomena.

References

Aussem, A. & Murtagh, F. 1997 Combining neural network forecasts on wavelet-transformed time series. *Connection Sci.* **9**, 113–121.

Aussem, A., Campbell, J. & Murtagh, F. 1998 Wavelet-based feature extraction and decomposition strategies for financial forecasting. *J. Computat. Intell. Finance* **6**(2), 5–12.

Bollerslev, T., Chou, R. Y., Jayaraman, J. & Kroner, K. F. 1990 ARCH modeling in finance: a review of the theory and empirical evidence. *J. Econometrics* **52**, 5–60.

Chui, C. K., Goswami, J. C. & Chan, A. K. 1995 Fast integral wavelet transform on a dense set of the time-scale domain. *Numerische Mathematik* **70**, 283–302.

Davidson, R., Labys, W. C. & Lesourd, J.-B. 1997 Wavelet analysis of commodity price behavior. *Computat. Economics* **11**, 103–128.

Donoho, D. L. & Johnstone, I. M. 1995 Adapting to unknown smoothness via wavelet shrinkage. *J. Am. Statist. Ass.* **90**, 1200–1224.

Donoho, D. L. & Johnstone, I. M. 1998 Minimax estimation by wavelet shrinkage. *Ann. Statist.* **26**, 879–921.

Donoho, D., Johnstone, I., Kerkyacharian, G. & Picard, D. 1995 Wavelet shrinkage: asymptopia? (with discussion). *J. R. Statist. Soc.* **57**, 301–369.

Härdle, W., Kerkyacharian, G., Picard, D. & Tsybakov, A. 1998 *Wavelets, approximation, and statistical applications.* Springer.

Jensen, M. J. 1997*a* An alternative maximum likelihood estimator of long-memory processes using compactly supported wavelets. Working paper no. 97-05. Department of Economics, University of Missouri–Columbia.

Jensen, M. J. 1997*b* Using wavelets to obtain a consistent ordinary least squares estimator of the long memory parameter. Working paper no. 97-09. Department of Economics, University of Missouri–Columbia.

Johnstone, I. M. & Silverman, B. W. 1997 Wavelet threshold estimators for data with correlated noise. *J. R. Statist. Soc.* B **59**, 319–351.

Mallat, S. & Zhang, Z. 1993 Matching pursuit with time-frequency dictionaries. *IEEE Trans. Signal Process.* **41**, 3397–3415.

Morehart, M., Murtagh, F. & Starck, J.-L. 1999 Spatial representation of economic and financial measures used in agriculture via wavelet analysis. *Int. Jl Geographic Information Systems.* (In the press.)

Neumann, M. H. 1996 Spectral density estimation via nonlinear wavelet methods for stationary non-Gaussian time series. *J. Time Series Analysis* **17**, 601–633.

Pan, Z. & Wang, X. 1998 A stochastic nonlinear regression estimator using wavelets. *Computat. Economics* **11**, 89–102.

Ramsey, J. B. 1969 Tests for specification errors in classical linear least-squares regression analysis. *J. R. Statist. Soc.* B **31**, 350–317.

Ramsey, J. B. & Lampart, C. 1998*a* Decomposition of economic relationships by time scale using wavelets: money and income. *Macroecon. Dynam.* **2**, 49–71.

Ramsey, J. B. & Lampart, C. 1998*b* The decomposition of economic relationships by time scale using wavelets: expenditure and income. *Stud. Nonlinear Dynam. Econometrics* **3**, 23–42.

Ramsey, J. B. & Schmidt, P. 1976 Some further results on the use of OLS and BLUS residuals in specification error tests. *J. Am. Statist. Ass.* **71**, 389–390.

Ramsey, J. B. & Thomson, D. J. 1998 A reanalysis of the spectral properties of some economic and financial time series. In *Nonlinear time series analysis of economic and financial data* (ed. P. Rothman), pp. 45–85. Boston, MA: Kluwer.

Ramsey, J. B. & Zhang, Z. 1996 The application of waveform dictionaries to stock market data. *Predictability of complex dynamical systems* (ed. Y. A. Kravtsov & J. B. Kadtke), pp. 189–205. Springer.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Ramsey, J. B. & Zhang, Z. 1997 The analysis of foreign exchange rate data using waveform dictionaries. *J. Empirical Finance* **4**, 341–372.

Ramsey, J. B., Usikov, D. & Zaslavsky, G. 1995 An analysis of U.S. stock price behavior using wavelets. *Fractals* **3**, 377–389.

White, H. 1989 Some asymptotic results for learning in single hidden-layer feed-forward models. *J. Am. Statist. Ass.* **84**, 1003–1013.