Productivity and unemployment scale-by-scale relationship

Marco Gallegati[*], Mauro Gallegati[*], James B. Ramsey[†] and Willi Semmler[‡]

October 13, 2011

1 Introduction

From a macroeconomic perspective the question of productivity growth is related to the effects that productivity gains may display on the overall behavior of the economy (e.g. output, employment, wages, costs, and prices) and, furthermore, as a source of a rising standard of livings. But if economists generally agree on the long run positive effects of labor productivity on real wages and output growth, much more controversial is the issue whether productivity growth is good or bad for employment: Does productivity growth increase unemployment or does it reduce it? The empirical and theoretical results have been mixed.

Empirically there is the literature on the employment and productivity differentials between Europe and the US since 1970. The early literature states a possible *trade-off* between employment and productivity growth (Gordon, 1997). But such empirical findings have been complicated by the recent behavior of (un)employment and productivity growth in the Europe and the US in the 90's: The increase of productivity growth in the US in the second half of the 90's can be shown to be associated with low and falling unemployment (Staiger *et al.* 2001) and with rising employment. Yet, in Europe there was the opposite tendency visible: Productivity growth seems to have reduced employment and increased unemployment. Over the business cycle the empirical evidence seems to show strong pro-cyclical comovements of employment and productivity with output (counter-cyclical for unemployment).[1]

The nexus of productivity and employment is also important for the empirical study of Okun's Law. If employment is correlated with output,

[*]DESS and SIEC, Università Politecnica delle Marche, Ancona, Italy

[†]Department of Economics, New York University, New York, USA

[‡]Department of Economics, University of Bielefeld, Germany and New School University, New York, USA

[1]See Backus and Kehoe (1992) among the others.

but does not reveal a one-to-one relationship as Okun (1962) states it, the relationship may change over time, due to changing growth rates of productivity. Thus, the study of the impact of productivity on employment becomes a relevant issue here too. After Okun's study was published in 1962, many authors have been involved in this discussion of the relationship of productivity and employment either from the short run or long run perspective. Particularly relevant authors are Tobin (1993), Kaldor (1985) and Solow (1997).

The relationship between productivity and employment has also become important in recent evaluations of Real Business Cycle (RBC) models. RBC theorists have postulated technology shocks as driving force of business cycles. In RBC models technology shocks, output and employment (measured as hours worked) are then predicted to be positively correlated. This claim has been made the focus of numerous econometric studies. [2] Employing the Blanchard and Quah (1989) research agenda by using VAR estimates, studies by Gali (1999), Gali and Rabanal (2005), Francis and Ramey (2004) and Basu et al. (2006) find a negative correlation of employment and productivity growth, once the technology shocks have been purified, taking out demand shocks affecting output. Yet most of the econometric work has studied the effects of productivity growth on employment (hours worked) in a "one time scale model"—for aggregate time series data.

Although our paper is relevant for the above mentioned empirically as theoretically controversies, we here focuses not on employment but unemployment. Although the nexus of unemployment and productivity growth rates may be impacted by population growth, demographic shifts, changing labor market participation rates of certain segments of the population and so on, one might presume that the demand side of labor, the offered employment by firms, is the most essential factor for driving the unemployment rate. By following up this line of thinking, we want to study the relationship of unemployment and productivity growth at multiple time scales.

The idea that time scales can be relevant in this context has already been expressed in Landmann (2004, p.3) who states that "it is useful to distinguish between an analysis of the forces shaping long-term equilibrium paths of output, employment and productivity on the one hand and the forces causing temporary deviations from these equilibrium paths on the other hand. However, ..., the need for this distinction is not universally accepted by macroeconomists." and that (Landmann, 2004, p.35) "the nature of the mechanism that link them [cfr.(un)employment and productivity growth] changes with the time frame adopted". In this perspective it may be useful to distinguish between the short, medium and long-run effects of changes in productivity growth, as their effects of productivity growth may appear contradictory, as it is, for example, in the case of the process of job

[2] For details, see Gong and Semmler (2006)

creation (see Landmann, 2004 and Walsh, 2004). [3]

The labour market provides an example of a market in which the agents involved, firms and workers (through unions), interact at different time horizons, and thus both the time horizon of economic decisions and the strength and direction of economic relationships among labor market variables, *i.e.* wages, prices and (un)employment, are likely to vary across time scales.[4] Thus, the long run effects of technological decisions maybe different from short run effects. In the short run new technology is likely to be labor reducing, and thus adding to unemployment, [5] as was visible in Europe since the 1990s. In the long run, however, new technology replacing labor increases productivity and makes firms and the economy more competitive and may reduce unemployment, and thus increase employment. [6]

Such relationships, and, in particular, the medium and long-run relationships between productivity growth and unemployment are generally analyzed in the empirical literature looking at average aggregate data, generally decades, because from a time series perspectives the rate of growth of labor productivity is a very volatile series whose implications in terms of the movements of the other supply-side variables are difficult to interpret, particularly in the short-run. [7] There a number of econometric techniques, both in time series and frequency domain, that have been used to disentangle short-term from long-term changes in productivity growth.

Yet, while standard econometric techniques (both time and frequency domain) may face some difficulties in separating long-run trends from short-run phenomena, other analytical tools, such as wavelets, may reveal useful, as wavelet analysis with respect to other filtering methods is able to decompose macroeconomic time series and data in general, into their time scale components. After the first applications of wavelet analysis in economics and finance provided by Ramsey and his co-authors (Ramsey and Zhang, 1995, 1996, Ramsey and Lampart, 1998a, 1998b), the number of wavelet applications in economics is rapidly growing in the last few years as a result of the increasing interest in this new tool to study economic relationships at different time scales.

[3]Most of the attention of economic researchers who work on productivity has been devoted to measurement issues and to resolve the problem of data consistency, as there are many different approaches to the measurement of productivity linked to the choice of data, notably the combination of employment, hours worked and GDP (see for example the OECD Productivity Manual, 2001).

[4]E.g. in Gallegati *et al.* (2009), (2011) where wavelet analysis is applied to the wage Phillips curve for the US.

[5]A statement like this gos back to David Ricardo who has pointed out that if machinery is substituted for labor unemployment is likely to increase.

[6]This point is is made clear in a simple text book illustration by Blanchard (2005)

[7]Indeed, the relationship between productivity and the unemployment rate may appear weaker when we reduce the time period used for aggregating data (see Steindel and Stiroh, 2001).

The objective of this paper is to provide evidence on the nature of the relationship between labor productivity growth and the unemployment rate on a scale-by-scale basis for the US, as it may help to isolate some key relationships over different times scales and thereby may provide some information about the challenging theoretical frameworks and the conduct of monetary policy. Thus, after decomposing both variables into their time-scale components using to the *maximum overlap discrete wavelet transform (MODWT)*, we analyze the relationship between labor productivity and unemployment at the different time scales using parametric and nonparametric approaches, as this latter framework may enable us to characterize the dynamic relationships among these variables without making any a priori explicit or implicit assumption about the shape of the relationship.

The paper proceeds as follows. In Section 2, we apply the Continuous Wavelet Transform tools, which include the Wavelet Power Spectrum, Wavelet Coherency and Phase-Difference, to productivity growth and the unemployment rate. In Section 3, we apply the MODWT to analyze the scale-by-scale relationships between the unemployment rate and productivity by using parametric and nonparametric time scale regression analyses. Section 4 concludes the paper.

2 Continuous wavelet transform

The role of wavelets has by now become familiar in empirical economic analysis. Wavelets, their generation,and their potential use are discussed in intuitive terms in Ramsey {Palgrave}. Gencay *et al.* (2002) generate an excellent development of wavelet analysis and provide many interesting economic examples. Percival and Walden (2000) provides a more technical exposition with many examples of the use of wavelets in a variety of fields, but not in economics. A variety of economic examples of the use of wavelets can be obtained from reviewing the list of joint articles shown below. In particular, see Ramsey and Zhang (1995, 1996), Ramsey and Lampart (1998a, 1998b), Ramsey *et al.* (2010) and Gallegati *et al.* (2006, 2009, 2011).

The essential characteristics of wavelets are best illustrated through the development of the continuous wavelet transform (CWT). We seek functions $\psi(t)$ such that:

$$\begin{aligned} \int \psi(u)du &= 0 \\ \int \psi(u)^2 du &= 1 \end{aligned}$$

The cosine function is a "large wave" because its square does not converge to 1, even though its integral is zero; a wavelet, a "small wave" obeys both constraints. An example would be the Haar wavelet function:

$$\psi^H(u) = \begin{cases} -\frac{1}{\sqrt{2}} & -1 < u < 0 \\ \frac{1}{\sqrt{2}} & 0 < u < 1 \\ 0 & otherwise \end{cases}$$

Such a function provides information about the variation of a function, $f(t)$, by examining the differences over time of partial sums. As will be illustrated below general classes of wavelet functions compare the differences of **weighted averages** of the function $f(t)$. Consider a signal, $x(u)$ and the corresponding "average":

$$\frac{1}{b-a}\int_a^b x(u)du = \alpha(a, b)$$

Let us choose the convention that we assess the value of the "average" at the center of the interval and let $\lambda \equiv b - a$ represent the scale of the partial sums. We have the expression:

$$\begin{aligned} A(\lambda, t) &\equiv \alpha(t - \lambda/2,\ t + \lambda/2) \\ &= \frac{1}{\lambda}\int_{t-\lambda/2}^{t+\lambda/2} x(u)du \end{aligned}$$

A(λ, t) is the average value of the signal centered at "t" with scale "λ". But what is of more use is to examine the differences at different values for λ and at different values for "t." We define:

$$\begin{aligned} D(\lambda, t) &= A(\lambda, t + \lambda/2) - A(\lambda, t - \lambda/2) \\ &= \frac{1}{\lambda}\int_t^{t+\lambda} x(u)du - \frac{1}{\lambda}\int_{t-\lambda}^{t} x(u)du \end{aligned}$$

This is the basis for the continuous wavelet transform, CWT, as defined by the Haar wavelet function. For an arbitrary wavelet function, ψ the wavelet transform is:

$$\begin{aligned} W(\lambda,\ t) &= \int_{-\infty}^{\infty} \psi_{\lambda,\ t}(u)x(u)du \\ \psi_{\lambda,\ t}(u) &\equiv \frac{1}{\sqrt{\lambda}}\left(\frac{u-t}{\lambda}\right) \end{aligned}$$

see Percival and Walden (2000).

2.1 Wavelet power spectrum

Time-scale analysis using continuous wavelet transform (CWT) can provide us with the (local) wavelet power spectrum (sometimes called scalogram or wavelet periodogram) which can be interpreted as the energy density in the time-frequency plane. Let $W_x(s, \tau)$ be the continuous wavelet transform of a signal $x(t)$, with respect to the wavelet ψ, where s is a scaling or dilation factor that controls the length of the wavelet and τ a location parameter that indicates where the wavelet is centered, $|W_x|^2$ represents the wavelet power which depicts the local variance of $x(t)$. Among the several types of wavelet families available such as, Morlet, Mexican hat, Haar, Daubechies, etc. we choose to employ a widely used wavelet such as the Morlet wavelet,[8] defined as

$$\psi_\eta(t) = \pi^{-\frac{1}{4}} e^{i\eta t} - e^{-\frac{t^2}{2}}.$$

The Morlet wavelet is a complex wavelet that produces complex transforms and thus can provide us with information on both amplitude and phase.[9] We use the Morlet wavelet with $\omega_0 = 6$ (where ω_0 is dimensionless frequency) since this particular choice provides a good balance between time and frequency localization (see Grinsted *et al.* 2004) and also simplifies the interpretation of the wavelet analysis because the wavelet scale, s, is inversely related to the frequency, $f \approx 1/s$.

In Figures 1 and 2, we see the continuous wavelet power spectra of the labor productivity growth and the unemployment rate, respectively.[10] Time is recorded on the horizontal axis and the vertical axis gives us the periods (and the corresponding scales of the wavelet transform). Reading across the graph at a given value for the wavelet scaling, one sees how the power of the projection varies across the time domain at a given scale, while reading down the graph at a given point in time, one sees how the power varies with the scaling of the wavelet (see Ramsey *et al.*, 1994). The wavelet power spectrum provides time varying analysis of the characteristics of a process in the scale-space plane and can be quite revealing about the structure of a particular process. Because wavelets indicate the presence of multiscale features, one can identify their temporal locations. The power of the projection of the signal onto the wavelet transform at the indicated level of scaling is indicated by color coding, so that we may evaluate the scaling characteristics of the

[8] The Morlet wavelet has optimal joint timefrequency concentration as it attains the minimum possible uncertainty of the corresponding Heisenberg box.

[9] As will be shown in the next subsection, detecting the phase difference is important because the phase difference characterizes phase relationships between two time series.

[10] We use quarterly data for the US between 1948:1 and 2010:4 from the Bureau of Labor Statistics. Labor productivity is defined as output per hour of all persons of the Nonfarm Business Sector (Index 2005 = 100) and measured as percentage change quarter ago at annual rate. Unemployment rate is defined as percent Civilian Unemployment Rate.

data by examining the color plots of the continuous wavelet transform. This color coding can provide an objective method for determining the principal timescales present in a signal and also for providing information about the scales at which important features provide a significant contribution. The color code for power ranges from blue (low power) to red (high power). Regions with warmer colors (red, orange and bright green) correspond to areas of high power, that is regions with wavelet transform coefficients of large modulus.

The statistical significance of the results obtained through wavelet power analysis was first assessed by Torrence and Compo (1998) by deriving the empirical (chi-squared) distribution for the local wavelet power spectrum of a white or red noise signal using Monte Carlo simulation analysis. A black contour line testing the wavelet power 5% significance level against the null hypothesis that the data generating process is generated by a stationary process is displayed, as is the cone of influence, represented by a shaded area corresponding to the region affected by edge effects.[11]

The first thing we can note is that the two series have very different wavelet power spectra. In the case of labor productivity growth there is evidence of highly localized patterns at certain scales, with high power regions concentrated in the first part of the sample (until late eighties) and at scales corresponding to periods up to 4 years. Otherwise, for the unemployment rate high power regions reveal the presence of dominant scales of variation at scales corresponding to the medium and long-run periods, since the coefficients of maximal energy are concentrated at the highest scales (lowest frequencies) and, starting from the seventies, also at intermediate scales, *i.e.* scales 3 and 4.

Although useful for revealing potentially interesting features in the data like characteristic scales, the wavelet power spectrum is not the best tool to deal with the time-frequency dependencies between two time-series. Indeed, even if two countries share a similar high power region, one cannot infer that their business cycles look alike. To detect and quantify relationships between variables, cross-wavelet tools like wavelet coherency and wavelet phase-difference have to be used.

[11] As with other types of transforms, the CWT applied to a finite length time series inevitably suffers from border distortions; this is due to the fact that the values of the transform at the beginning and the end of the time series are always incorrectly computed, in the sense that they involve missing values of the series which are then artificially prescribed; the most common choices are zero padding extension of the time series by zeros or periodization. Since the effective support of the wavelet at scale s is proportional to s, these edge effects also increase with s. The region in which the transform suffers from these edge effects is called the cone of influence. In this area of the time-frequency plane the results are unreliable and have to be interpreted carefully (see Percival and Walden, 2000).

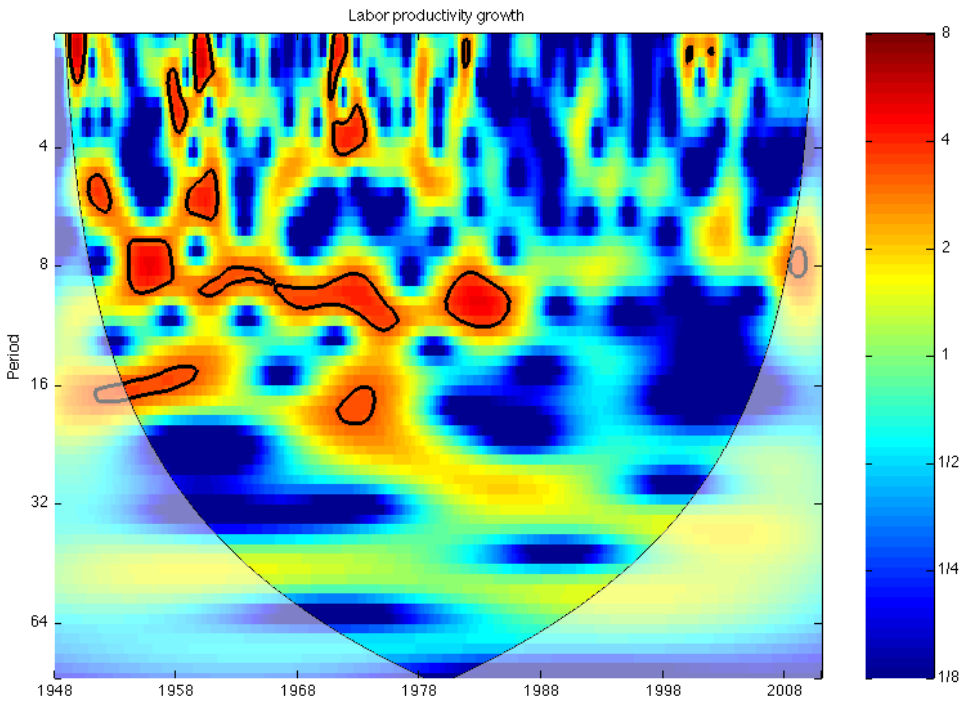

Figure 1: Wavelet power spectrum for labor productivity growth

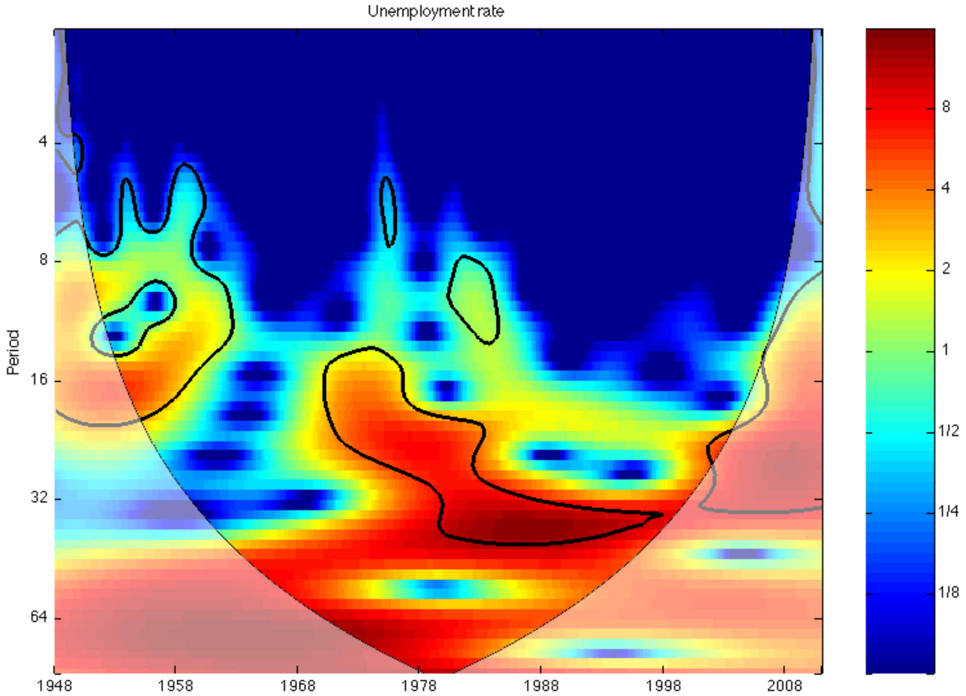

Figure 2: Wavelet power spectrum for the unemployment rate

2.2 Wavelet coherency

Wavelet tools suitable for the analysis of time-frequency dependencies between two time series are the cross-wavelet power, wavelet coherency and wavelet phase difference. Let W_x and W_y be the continuous wavelet transform of the signal $x(t)$ and $y(t)$, the cross-wavelet power of the two series is given by $|W_{xy}|=|W_x W_y|$ and depicts the local covariance of the two time series at each scale and frequency (see Hudgins *et al.*,1993). The wavelet coherence is defined as the modulus of the wavelet cross spectrum normalized to the single wavelet spectra and is especially useful in highlighting the time and frequency intervals where two phenomena have strong interactions. It can be considered as the local correlation between the time series in time frequency space. Finally, the phase difference can be useful to characterize the phase relationships between two time series as a function of frequency, *i.e.* phase synchronization of two time series.

As for the wavelet power spectrum, the wavelet coherency power is indicated by color coding. The color code for power ranges from blue (low coherency) to red (high coherency), with regions of high coherency between two time series corresponding to areas of strong local correlation. The statistical significance level of the wavelet coherence is estimated using Monte Carlo methods. The 5% significance level against the null hypothesis of red noise is shown as a thick black contour. The cone of influence is marked by a black thin line: again, values outside the cone of influence should be interpreted very carefully, as they result from a significant contribution of zero padding at the beginning and the end of the time series.

The phase difference between the two series is indicated by arrows. Indeed, the "phase arrows" show the relative phasing of the two time series and can also be interpreted as a lead/lag relationship. Right arrow means that the two variables are in-phase. If the right arrow points up (down) means that unemployment rate is lagging (leading). At the opposite, left arrow means that the two variables are out-of-phase (in-antiphase). If the left arrow points down (up) means that unemployment rate is lagging (leading).

Regions of strong coherency are evident at business cycles scales, *i.e.* at scales corresponding to periods between 2 and 8-years, except for the mid80s-mid90s period. The phase difference reveals that around the 4-year frequency the two series are generally in phase, while at frequencies approximating 2- and 8-years we can see that the unemployment rate was leading in the mid60s-mid80s period. Otherwise, in the last decade of the sample productivity growth is, respectively, leading and slightly leading the unemployment rate at scales corresponding to periods between 4- and 8 years . That both series are highly correlated at the longest scale (corresponding to periods greater than 16 years) with a stable antiphase relationship until mid 80s is also revealed by the high coherency (about 0.6) emerging throughout the sample around these frequencies.

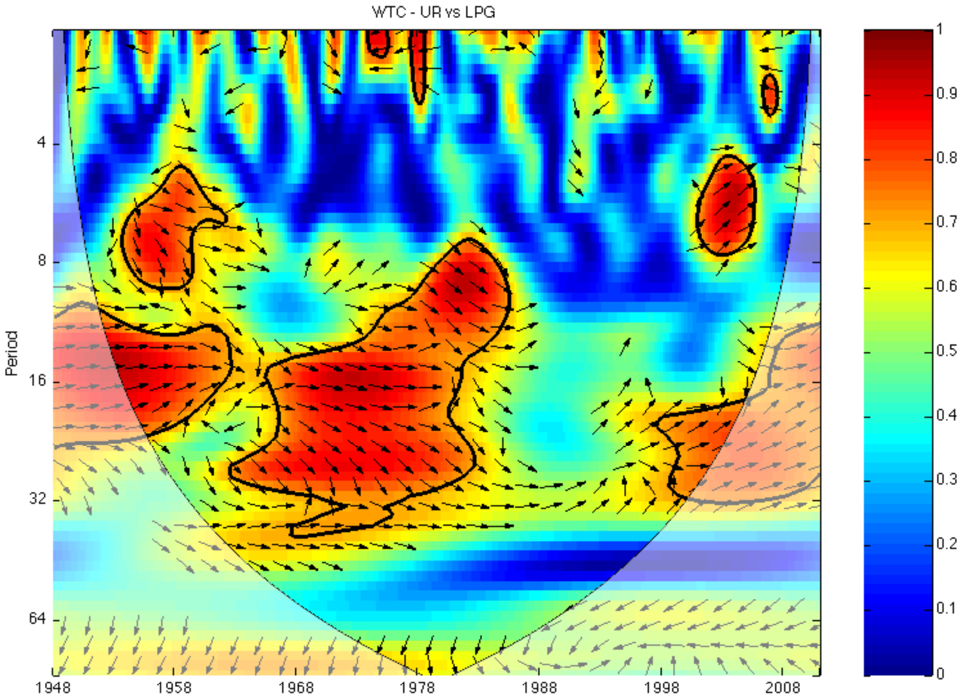

Figure 3: Wavelet coherency between the unemployment rate and productivity growth

In sum, we can conclude that not only the wavelets are adequate to capture characteristic scales and transient relations, but that they can also distinguish between different relations that occur at the same time but at distinct frequencies.

3 Discrete wavelet transform

So far we have considered only continuously labeled decompositions. But there are several difficulties with the CWT, since the CWT is a highly information redundant transform representing each datum by a pair of data, designing time, or space, and scale. First, it is computationally impossible to analyze a signal using all wavelet coefficients. Second, as noted by Gencay *et al.* (2002), $W(\tau; s)$ is a function of two parameters and as such it contains a high amount of redundant information. As a consequence, although the CWT provides a useful tool for analyzing how the different periodic components of a time series evolve over time, both individually (single wavelet power spectrum) and jointly (cross-wavelet power, wavelet coherency and phase-difference), in practice a discrete analogs of these techniques is developed. We therefore move to the discussion of the discrete wavelet transform (DWT), since the DWT, and in particular the MODWT, a variant of the DWT, is largely predominant in economic applications.[12]

The DWT is based on similar concepts as the CWT, but is more parsimonious in its use of data (Gencay *et al.*, 2003). In order to implement the discrete wavelet transform on sampled signals we need to discretize the transform over scale and over time through the dilation and location parameters. Indeed, the key difference between the CWT and the DWT lies in the fact that the DWT uses only a limited number of translated and dilated versions of the mother wavelet to decompose the original signal. The idea is to select τ and s so that the information contained in the signal can be summarized in a minimum number of wavelet coefficients. The number of observations at each scale is given by $N/2^j$ $j = 1, 2, ...J$. The discretized transform is known as the discrete wavelet transform, DWT.

The discretization of the continuous time-frequency decomposition creates a discrete version of the wavelet power spectrum in which the entire time-frequency plane is partitioned with rectangular cells of varying dimensions but constant area, called Heisenberg cells.[13] Higher frequencies can be

[12]The number of the papers applying the DWT is far greater than those using the CWT. As a matter of fact, the preference for DWT in economic applications can be explained by the ability of the DWT to facilitate a more direct comparison with standard econometric tools than is permitted by the CWT (e.g. time scales regression analysis, homogeneity test for variance, nonparametric analysis,).

[13]Their dimensions change according to their scale: the windows stretch for large values of s to measure the low frequency movements and compress for small values of s to measure the high frequency movements.

well localized in time, but the uncertainty in frequency localization increases as the frequency increases, which is reflected as taller, thinner cells with increase in frequency. Consequently, the frequency axis is partitioned finely only near low frequencies. The implication of this is that the larger-scale features of the signal get well resolved in the frequency domain, but there is a large uncertainty associated with their location. On the other hand, the small-scale features, such as sharp discontinuities, get well resolved in the time domain, even if there is a large uncertainty associated with their frequency content. This trade-off is an inherent limitation due to the Heisenbergs uncertainty principle that states that the resolution in time and frequency cannot be arbitrarily small because their product is lower bounded. Therefore, owing to the uncertainty principle, an increased resolution in the time domain for the time localization of high-frequency components comes at a cost of an increased uncertainty in the frequency localization, that is one can only trade time resolution for frequency resolution, or vice versa.

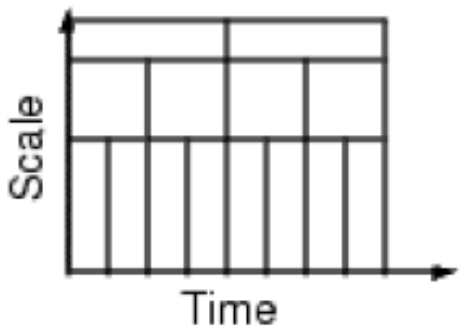

Figure 4: DWT time-scale partition

The general formulation for a continuous wavelet transform can be restricted to the definition of the "discrete wavelet transform", the properties of which can be summarized by the equation:

$$\psi_{j,k}(t) = 2^{-j/2}\psi\left(\frac{t-2^j k}{2^j}\right) \tag{1}$$

which is known as the "mother wavelet." This function represents a sequence of rescaleable functions at a scale of s $= 2^j, j = 1, 2, ...J$, and with time index k, k=1,2, 3,...N/2^j. The wavelet transform coefficient of the projection of the observed function, f(t), i = 1,2,3, ...N, N= 2^J on the wavelet $\psi_{j,k}(t)$ is given by:

$$\begin{aligned} d_{j,k} &\approx \int \psi_{j,k}(t) f(t) dt, \\ j &= 1, 2, ..J \end{aligned} \tag{2}$$

For a complete reconstruction of a signal f(t), one requires a scaling function, $\phi(.)$, that represents the smoothest components of the signal. While the wavelet coefficients represent weighted "differences" at each scale, the scaling coefficients represent averaging at each scale. One defines the scaling function, also know as the "father wavelet," by:

$$\phi_{J,k}(t) = 2^{-J/2} \phi \left(\frac{t - 2^J k}{2^J} \right) \tag{3}$$

And the scaling function coefficients vector is given by:

$$s_{J,k} \approx \int \phi_{J,k}(t) f(t) dt, \tag{4}$$

By construction, we have an orthonormal set of basis functions, whose detailed properties depend on the choices made for the functions, $\phi(.)$ and $\psi(.)$, see for example the references cited above as well as Daubechies (1992) and Silverman (1998). At each scale, the entire real line is approximated by a sequence of "non-overlapping" wavelets. The deconstruction of the function f(t) is therefore:

$$\begin{aligned} f(t) \approx & \sum_k s_{J,k} \phi_{J,k}(t) + \sum_k d_{J,k} \psi_{J,k}(t) + \\ & \sum_k d_{J-1,k} \psi_{J-1,k}(t) + ... + \sum_k d_{1,k} \psi_{1,k}(t) \end{aligned} \tag{5}$$

The above equation is an example of the Discrete Wavelet Transform, DWT based on an arbitrary wavelet function, $\phi(.)$. Using economic variables, the degree of relative error is approximately on the order of 10^{-13} in many cases, so that one can reasonably claim that the wavelet decomposition is very good. For the DWT, where the number of observations is N, N $= 2^J$, the number of coefficients at each scale is:

$$N = N/2^J + N/2^J + N/2^{J-1} + ...N/4 + N/2 \tag{6}$$

That is, there are $N/2^J$ coefficients $s_{J,k}$, $N/2^J$ coefficients $d_{J,k}$, $N/2^{J-1}$ coefficients $d_{J-1,k}$... and $N/2$ coefficients $d_{1,k}$.

While it would appear that wavelets involve large numbers of coefficients, it is also true that the number of coefficients greater than zero is very small;

the arrays are said to be "sparse". In the literature quite complicated functions are approximated to a high level of accuracy with a surprisingly small number of coefficients. As a corollary to this general statement, other scholars have noted the extent to which the distribution of coefficients under the null hypothesis of zero effect, rapidly approaches the Gaussian distribution.

Further, the approximation can be re-written in terms of collections of coefficients at given scales. Define;

$$\begin{aligned} S_J &= \sum_k s_{J,k}\phi_{J,k}(t) \\ D_J &= \sum_k d_{J,k}\psi_{J,k}(t) \\ D_{J-1} &= \sum_k d_{J-1,k}\psi_{J-1,k}(t) \\ &\cdots\cdots \\ D_1 &= \sum_k d_{1,k}\psi_{1,k}(t) \end{aligned} \quad (7)$$

Thus, the approximating equation can be restated in terms of coefficient crystals as:

$$f(t) \approx S_J + D_J + D_{J-1} + ...D_2 + D_1 \quad (8)$$

S_J contains the "smooth component" of the signal, and the D_j, j=1,2,..J, the detail signal components at ever increasing levels of detail. S_J provides the large scale road map, D_1 shows the pot holes. The previous equation indicates what is termed the multiresolution decomposition, MRD.

3.1 Time scale decomposition analysis

The orthonormal discrete wavelet transform (DWT), even if widely applied to time series analysis in many disciplines, has two main drawbacks: 1) the dyadic length requirement (i.e. a sample size divisible by 2^J), and 2) the wavelet and scaling coefficients are not shift invariant. Because of the practical limitations of DWT we perform wavelet analysis by applying the *maximal overlap discrete wavelet transform* (MODWT) using the Daubechies least asymmetric (LA) wavelet filter of length $L = 8$ based on eight non-zero coefficients (Daubechies, 1992), with reflecting boundary conditions. The MODWT is a non-orthogonal variant of the classical discrete wavelet transform (DWT) that, unlike the DWT, is translation invariant, as shifts in the signal do not change the pattern of coefficients, can be applied to data sets of length not divisible by 2^J and returns at each scale a number of coefficients equal to the length of the original series.

The application of the maximal overlap discrete wavelet transform (MODWT) with a number of levels (scales) $J = 5$ produces six individual crystals[14]: one vector of smooth coefficients s_5, representing the underlying smooth behavior of the data at the coarse scale, and five vectors of details coefficients d_5, d_4, d_3, d_2, d_1, representing progressively finer scale deviations from the smooth behavior. The synthesis or reconstruction operation reassembles the original signal from the wavelet coefficients, by using the inverse stationary wavelet transform. We reconstructed detail and smooth components of the original signals.[15] Specifically, with $J = 5$ we come up with five wavelet details vectors D_5, D_4, D_3, D_2, D_1 and one wavelet smooth vector, S_5, each associated with a particular time period 2^{j-1} In particular, since we use quarterly data the first detail level D_1 captures oscillations between 2 and 4 quarters, while details D_2, D_3, D_4 and D_5 capture oscillations with a period of 1-2, 2-4, 4-8 and 8-16 years, respectively.[16]

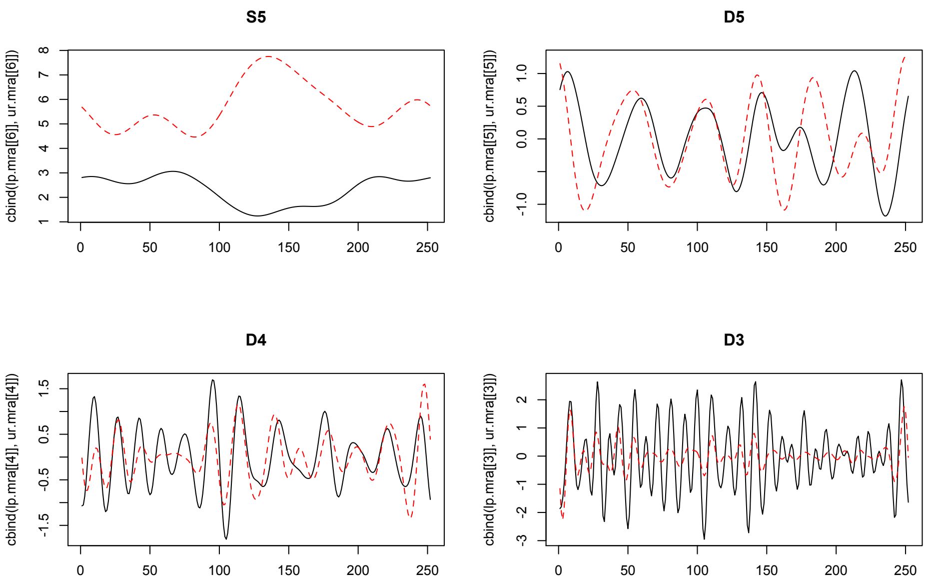

Figure 5: Phase shift relationships for unemployment (red dotted lines) and productivity (black solid lines)

[14]Each set of wavelet transform coefficients is called a crystal in wavelet terminology.

[15]Since the J components obtained by the application of MODWT are not orthogonal, they do not sum up to the original variable.

[16]Detail levels D_1 and D_2, represent the very short-run dynamics of a signal (and contains most of the noise of the signal), levels D_3 and D_4 roughly correspond to the standard business cycle time period (Stock and Watson, 1999), while the medium-run component is associated to level D_5. Finally, the smooth component S_5 captures oscillations with a period longer than 16 years corresponding to the low-frequency components of a signal.

In Figure 5 we plot the smooth component S_5 and the highest level deviations from the smooth component at an increasing level of detail, *i.e.* D_5, D_4 and D_3, as a sequence of pairs of time series between the unemployment rate (red dotted lines) plotted against labor productivity growth (black solid lines). The visual inspection of the long-run components indicate an antiphase relationship between the variables, with productivity growth slightly leading the unemployment rate. The pattern displayed by the top right panel in Figure 5 reveals that the two components are mostly in phase at the D_5 scale level, with unemployment slightly leading productivity growth. Nonetheless, the plot also shows that the two series at this level have been moving into antiphase at the beginning of the nineties, as a consequence of a shift in the phase relationship (structural break), and then have been moving in-phase again in the last part of the sample. At the D_4 scale level unemployment and productivity are in-phase throughout the sample with the exception of the sixties. Finally, at the D_3 level the most notable feature is represented by the different amplitude of the two components, with productivity growth displaying the larger amplitude.

3.2 Parametric analysis

Wavelets provide a unique tool for the analysis of economic relationships on a scale-by-scale basis. Indeed, through the time scale decomposition property of wavelet analysis it is possible to test directly for parameters instability across frequencies in a regression model by using time scale regression analysis. Time scale regression analysis allows the researcher to examine the relationship between the variables at each j scale where the variation in both variables has been restricted to the indicated specific scale. In order to perform a time scale regression analysis we need to partition each variable into a set of different components by using the discrete wavelet transform (DWT), such that each component corresponds to a particular range of frequencies, and then run regression analysis on a scale-by-scale basis (e.g. Ramsey and Lampart, 1998 and Gallegati *et al.* 2009).[17] Therefore, after decomposing the regression variables into their different time scale components using the MOWDT we estimate a sequence of least squares regressions using

$$ur[S_J]_t = \alpha_J + \beta_J lp[S_J]_t + \epsilon_t \tag{9}$$

and

$$ur[D_j]_t = \alpha_j + \beta_j lp[D_j]_t + \epsilon_t \tag{10}$$

[17]Thus, we test for frequency dependence of the regression parameter by using timescale regression analysis since the approaches used to detect and model frequency dependence such as spectral regression approaches (Hannan, 1963, Engle, 1974, 1978) present several shortcomings because of their use of the Fourier transformation. For examples of the use of this procedure in economics, see Ramsey and Lampart (1998a, 1998b) and Gallegati *et al.* (2006), (2009).

where $ur[S_J]_t$, and $lp[S_J]_t$ represent the components of the variables at the longest scale, and $ur[D_j]_t$, and $lp[D_j]_t$ represent the components of the variables at each j scale, with j=1,2,....,J.

Table 1: Time scale regression analysis for unemployment on productivity growth

Dependent variable: ur			
Aggregate	Δlp	Adj.R^2	$s.e.$
	0.0481 (1.678)	0.0072	1.613
Dependent variable: $ur[D_j]$			
Details	$\Delta lp[D_j]$	Adj.R^2	$s.e.$
D_1	-0.0126 (-2.49)	0.0203	0.1855
D_2	0.0419 (4.50)	0.0713	0.1961
D_3	0.2462 (12.03)	0.3641	0.3854
D_4	0.5052 (12.53)	0.3833	0.4203
D_5	0.3677 (5.62)	0.1085	0.5717
Dependent variable: $ur[S_J]$			
Scaling	$\Delta lp[S_J]$	Adj.R^2	$s.e.$
S_5	-1.4208 (-26.37)	0.7345	0.5023

Note: t-statistics in paranthesis.

In Table 1 we present the results from least squares estimates at the aggregate and each scale level. First of all, we notice that while the relationship between the unemployment rate and productivity is insignificant at the aggregate level, it seems to be significantly different from zero at all scales. Moreover, the scale-by-scale regressions of the unemployment rate on labor productivity growth indicates that the effects of productivity on unemployment rate differ across scales in terms of sign, significance, and estimated size. Indeed, the effect of productivity on unemployment rate is positive at the detail scale levels, with the relationship being mostly significant at the D3 and D4 scale levels, but negative at the smooth scale level, thus suggesting that labor productivity growth is associated with an increase in the unemployment rate in the short and medium term, and to a decrease in the long run.[18] In sum, the results from timescale regression analysis indicate the usefulness of disentangling short, medium and long-run effects of changes in productivity growth.

Finally, as anticipated by some scholars, we find that unemployment is decreased by an increase in productivity for the very short run and in the very long run. These are the forces that stimulate innovation and growth in the economy. These positive effects are offset in part by the increase in unemployment through productivity gains in the intermediate business cycle periods.

3.3 Nonparametric analysis

In what follows we apply a methodology that allows us to explore the robustness of the issues related to the relationship between labor productivity growth and the unemployment rate without making any *a priori* explicit or implicit assumption about the form of the relationship: nonparametric regression analysis. Indeed, nonparametric regressions can capture the shape of a relationship between variables without us prejudging the issue, as they estimate the regression function $f(.)$ linking the dependent to the independent variables directly, and without providing any parameters estimate. [19]

There are several approaches available to estimate nonparametric regression models,[20] and most of these methods assume that the nonlinear functions of the independent variables to be estimated by the procedures are *smooth* continuos functions. One such model is the locally weighted polynomial regression, i.e. *loess*, pioneered by Cleveland (1979). This procedure fits the model $y = f(x_1, ..., x_k) + \epsilon$ nonparametrically, i.e., without assuming a parametric form for $f(x_1, ..., x_k)$ The low-degree polynomial, generally first or second degree (that is, either locally linear or locally quadratic), is fit using weighted least squares, where the data points are weighted by a smooth function whose weights decrease as the distance from the center of the window increases. The value of the regression function is obtained by evaluating the local polynomial at each particular value of the independent variable, x_i where a fixed proportion of the data is included in each given local neighborhood, called the *span* of the local regression smoother (or the smoothing parameter),[21] and the fitted values are then connected in a nonparametric

[18]A similar result is obtained in Chen *et al.* (2007) which, using three different econometric approaches, find that productivity growth affects unemployment positively in the short and negatively in the long run.

[19]The traditional nonlinear regression model introduce nonlinear functions of dependent variables using a limited range of transformed variables to the model (quadratic terms, cubic terms or piecewise constant function). An example of a methodology testing for nonlinearity without imposing any a priori assumption about the shape of the relationship is the smooth transition regression used in Eliasson (2001).

[20]See Fox (2000a, 2000b) for a discussion on nonparametric regression methods.

[21]The smoothing parameter controls the flexibility of the loess regression function: large

regression curve. The main advantage the local regression (*loess*) method is that it does not require the specification of a function to fit a model to all of the data in the sample. In addition, it provides robust fitting when there are outliers in the data, support multiple dependent variables and computes confidence limits for predictions when the error distribution is symmetric, but not necessarily normal. On the other hand, loess, being a method that fits models to localized subsets of the data, requires reasonably large, densely sampled datasets in order to produce good models.

In Figure 5 we report the six scatter plots of the unemployment-productivity growth relationship at the different scale levels, from S_5 (top left panel) to D_1 (top right panel). In each panel of Figure 5 a solid line drawn by connecting the points of the fitted values for each function against its regressor is superimposed on each scatter plot. The smooth plots represented by the solid lines depict the loess fit using a smoothing parameter value of 2/3.[22] These lines can be used to reveal the nature of the estimated relationship between the dependent (unemployment rate) and the independent variable (labor productivity).

The loess fits shown on the plots in Figure 5 confirms the conclusions obtained from the time scale regression results reported in Table 1. Indeed, the analysis of the nonparametric fitted functions in the top left panel of figure 5 suggests that the long-run relationships between labor productivity and the other variables denotes some kind of non-linearity. In particular, the shape of the nonparametric fitted regression function suggests a negative long-run relationship of labor productivity with the unemployment rate. Such long-run relationship between labor productivity and unemployment rate, in contrast to the medium-term result, do not provide support to the hypothesis of a trade-off between unemployment and productivity growth. In this case the sign of the relationship is the one expected as higher labor productivity growth is expected to be associated with lower unemployment rates in the long-run.

4 Interpretation

As evidenced by the literature discussed in section 1 the empirical and theoretical results on the relationship of labor productivity and unemployment are quite mixed. Most of the studies are rather inconclusive, or contradicted by other studies. Yet, most of the studies undertaken have, however, only employed a "one-time-scale" model and thus employed only aggregate time series data. As shown in section 3, if we indeed follow this methodology for

values of produce the smoothest functions that wiggle the least in response to fluctuations in the data, the smaller q is, the closer the regression function will conform to the data

[22] We use different smoothing parameters, but our main findings do not show excess sensitivity to the choice of the span in the loess function within what appear to be reasonable ranges of smoothness (*i.e.* between 0.4 and 0.8).

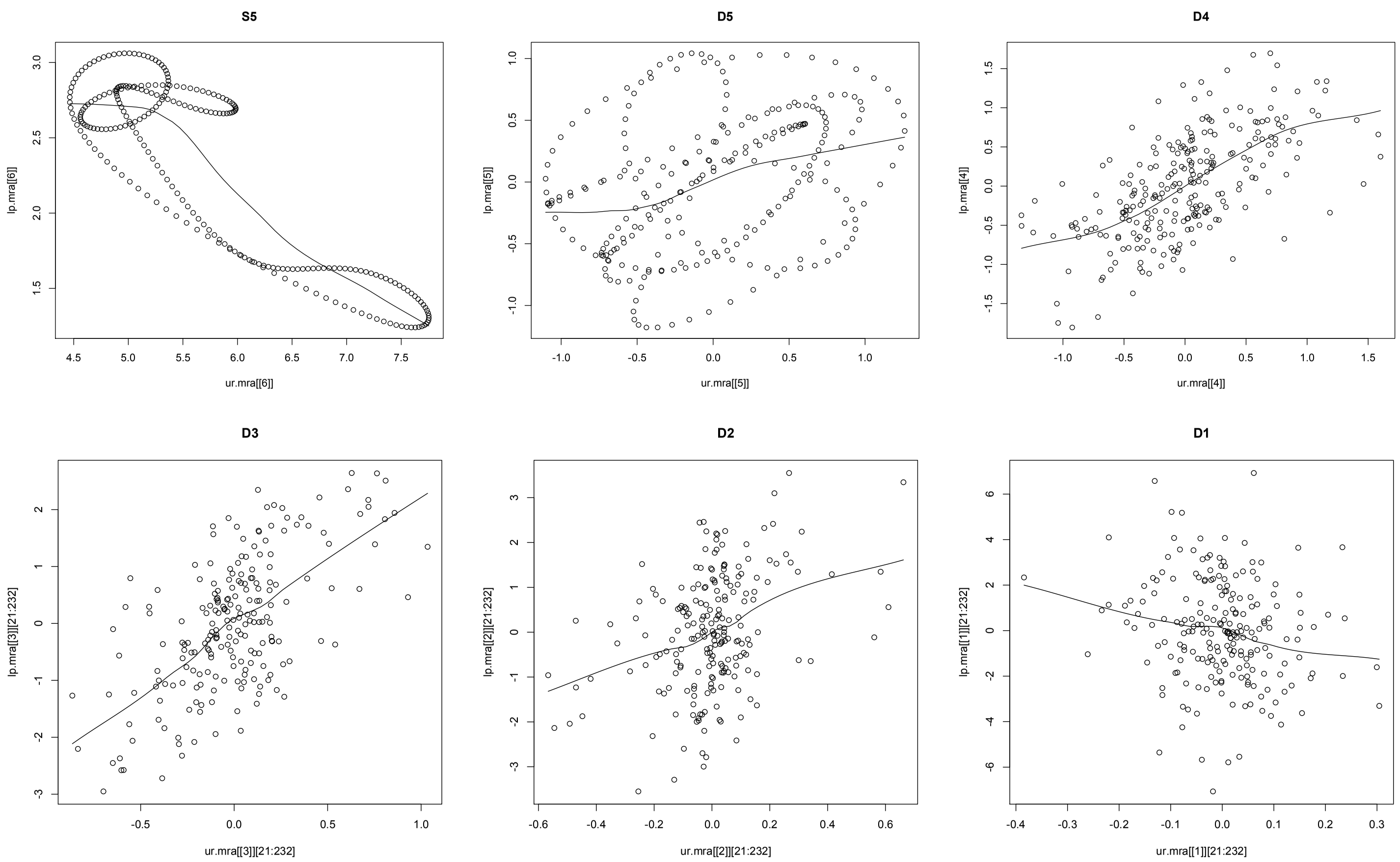

Figure 6: Scatter plot and loess fit at different scale levels

a US aggregate time series data set, with quarterly data, we find that the relationship between productivity and unemployment is insignificant.

Yet if we use wavelets to decompose the time series in multiple frequencies we find in the linear as well as in nonlinear regression analysis that there is, at certain time scales, a positive relationship between labor productivity and unemployment. The positive relationship appears to hold for short and medium run fluctuations, ranging from 2 to 4, 4 to 8 and 8 to 16 years. On the other hand, at the longest scale we find a negative relationship between productivity and unemployment. (The fact that there is also a positive relationship for the very short run, may just reflect some noise in the data). All in all, unemployment is positively associated with productivity in the short and medium term, but negatively in the long term. This demonstrates the usefulness of time scale regression analysis in disentangling short, medium and long run effects of changes in productivity growth on unemployment.

These results have relevant economic implications. First of all, ss regards the Okun's law, the US employment seems to be decoupled from economic growth. In the US there is a slowly recovering unemployment rate, though the annual growth rates of productivity are higher than in Europe. Due to high productivity growth rates, in the US one can observe some kind of jobless growth. Yet, this might be a short run phenomenon. In the long term this could be turned into a negative relationship of productivity and unemployment thus unemployment falling and employment rising with productivity growth in the long run.

Moreover, as to the controversial hypothesis of the RBC models that employment is rising with positive productivity shocks, the critics (such as Basu *et al.*, 2006) are presumably correct to state a nonsignificant relationship between (purified) technology shocks and employment or even a negative relationship of those variables. Yet, in the long run, since productivity makes the firms and the country more competitive, the increase in productivity may make unemployment falling and employment rising. So in some sense the RBC postulate of a positive relationship of productivity and employment seems to be incorrect in the short and medium run, but, given our results, is likely to hold on a long time scale.

5 Conclusion

In this paper, we used the wavelets methodology to analyze the productivity-unemployment relationship on a scale-by-scale basis. Specifically, by applying both continuous and discrete wavelet transform tools we are able to uncover relationships which would be otherwise very difficult to detect using classical econometric techniques. In a nutshell, we find that there is, for the short (long) time scales, a positive (negative) relationship between labor productivity and unemployment. The positive relationship holds for short and medium run fluctuations, from 2 to 16 years, while, for a longer time scale the relation becomes negative. These results suggest some relevant economic implications as to, for example, the Okun's law and the controversial hypotheses of the RBC models,

When Thomas More (Utopia, 1561) was asserting: sheep are eating men, he was, in the short run, right. Due to agriculture innovations, profits in the primary sector were rising, less labor force was employed in agriculture and more lands were devoted to pastureland. People had to "invent" new jobs, to increase their purchasing power, but it needed time to adjust.

6 References

Backus, D.K., and P.J. Kehoe (1992), "International Evidence on the Historical Evidence of Business Cycles", *American Economic Review*, 82 (4), 864-888.

Basu, S., J. G. Fernald, and M. S. Kimball (2006), "Are technology improvement contractionary?", *American Economic Review* 96, 1418–1448.

Blanchard, O.J. (2005), *Macroeconomics*, Fourth Edition, Prentice Hall International.

Blanchard, O. and D. Quah (1989), "The dynamic effects of aggregate demand and supply disturbances", *American Economic Review* 79, 655–673.

Bosworth B. and G.L. Perry (1994), "Productivity and Real Wages: Is There a Puzzle?", *Brooking Papers on Economic Activity*, 1994/1, 317-335.

Bowman A. W. and A. Azzalini (1997), *Applied Smoothing Techniques for Data Analysis*, Clarendon Press, Oxford.

Chen P., Rezai A. and W. Semmler (2007), "Productivity and Unemployment in the Short and Long Run", *SCEPA Working Paper*, 2007-8.

Cleveland W. S. (1979), "Robust Locally-Weighted Regression and Scatterplot Smoothing", *Journal of the American Statistical Association*, 74, 829-836.

Daubechies I. (1992), *Ten Lectures on Wavelets*, CBSM-NSF Regional Conference Series in Applied Mathematics, vol. 61, SIAM, Philadelphia.

Eliasson A.C. (2001),"Detecting Equilibrium Correction with Smoothly Time-Varying Strength", *Studies in Nonlinear Dynamics & Econometrics*, Vol. 5: No. 2, Article 2.

Engle, R. F. (1974), "Band Spectrum Regression, *International Economic Review*, 15, 1-11.

Engle, R. F. (1978), "Testing Price Equations for Stability Across Spectral Frequency Bands", *Econometrica*, 46, 869-881.

Estrella A. (2004), "Decoding Productivity: Business Cycle Properties of Labor Productivtiy Growth", *Working paper*, Federal Reserve Bank of New York.

Fox J. (2000a), *Nonparametric Simple Regression: Smoothing Scatterplots*, Sage, Thousands Oaks CA.

Fox J. (2000b), *Multiple and Generalized Nonparametric Regression*, Sage, Thousands Oaks CA.

Francis, N. and V. A. Ramey (2004), "The source of historical economic fluctuations: An analysis using long-run restrictions", *NBER Working Papers 10631*, National Bureau of Economic Research.

Francis, N. and V. A. Ramey (2005), "Is the technology-driven real business cycle hypothesis dead? Shocks and aggregate fluctuations revisited", *Journal of Monetary Economics* 52, 1379–1399.

Gali, J. (1999), "Technology, employment, and the business cycle: Do technology shocks explainaggregate fluctuations?", *American Economic Review* 89, 249–271.

Gali, J. and P. Rabanal (2005), "Technology shocks and aggregate fluctuations: How well does the rbc model fit postwar u.s. data?", *IMF Working Papers 04/234*, International Monetary Fund.

Gallegati, Marco, Gallegati, Mauro, Ramsey, J.B. and W. Semmler (2006), "The decomposition of the inflation-unemployment relationship by time-scale using wavelets", in Chiarella C. Franke R. Flaschel R. and Semmler W. (eds.), *Quantitative and Empirical Analysis of Nonlinear Dynamic Macromodels*, Amsterdam: Elsevier, 93-111.

Gallegati, Marco, Gallegati, Mauro, Ramsey, J.B. and W. Semmler (2009), "The US wage Phillips curve over different time horizons", *Giornale degli Economisti*, 68 (2), 113-148.

Gallegati, Marco, Gallegati, Mauro, Ramsey, J.B. and W. Semmler (2011), "The US wage Phillips curve across frequencies and over time", *Oxford Bullettin of Economics and Statistics*, 73(4), 489-508,.

Gallegati, Marco, Ramsey, J.B. and W. Semmler (2011), "Forecasting output using interest rate spreads: a wavelet decomposition analysis", *mimeo.*

Gençay, R., Selçuk, F., and Whitcher, B. (2001), "Scaling properties of exchange rate volatilities", *Physica A*, 289, 89-106.

Gençay, R., Selçuk, F., and Whitcher, B. (2002), *An Introduction to Wavelets and Other Filtering Methods in Finance and Economics*, San Diego Academic Press, San Diego.

Gordon R.J. (1997), "Is there a trade-off between unemployment and productivity growth?", in Snower D. and G. de la Dehesa *eds.*, *Unemployment policy: Government options for the labor market*, Cambridge, Cambridge University Press, 433-463.

Gong, G. and W. Semmler (2006), *Stoachstic Dynamic Macroeconomics: Theoy and Empirical Evidence*, Oxford University Press.

Grinsted, A., J. C. Moore, and S. Jevrejeva (2004), "Application of the cross wavelet transform and wavelet coherence to geophysical time series", *Nonlinear Processes in Geophysics*, 11, 561-566.

Hannan (1963), "Regression for Time Series with Errors of Measurement", Biometrika, 50 (3-4), 293-302.

Hastie T. and R. Tibshirani (1986), "Generalized Additive Models", *Statistical Science*, 1, 297-318.

Hastie T. and R. Tibshirani (1990), *Generalized Additive Models*, Chapman & Hall, London.

Hudgins, L., Friehe, C. A. and M. E. Mayer (1993), "Wavelet transforms and atmospheric turbulence", *Phys.Rev. Lett.*, 71(20), 3279-3282.

Kaldor N. (1985), "Economics Without Equilibrium (The Arthur M. Okun memorial lectures). M.E. Sharpe.

Landmann O. (2004), "Employment, Productivity and Output Growth", background paper for *World Employment Report 2004*, International Labour Organization (ILO, Geneva).

OECD (2001), *Measuring Productivity OECD Manual*, OECD, Paris.

Okun, A. (1962), "Potential GNP: Its measurement and significance", *Proceedings of the Business and Economic Statistics Section, American Statistical Association.*

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

Ramsey J.B. (2010), "Wavelets ",in Durlauf S.N. and L.E. Blume (eds), The New Palgrave Dictionary of Economics, Palgrave Macmillan.

Ramsey J.B. and Z. Zhang (1995), "The Analysis Of Foreign Exchange Data Using Waveform Dictionaries", *Journal of Empirical Finance*, 4, 341–372.

Ramsey, J. B. and Z. Zhang (1996), "The application of waveform dictionaries to stock market index data", in Kravtsov Y. A. and J. Kadtke (eds.), *Predictability of Complex Dynamical Systems*, Springer-Verlag, 189-208.

Ramsey, J.B. and C. Lampart (1998a), "The decomposition of economic relationship by time scale using wavelets: money and income", *Macroeconomic Dynamics and Econometrics*, 2, 49-71.

Ramsey, J.B. and C. Lampart (1998b), "The decomposition of economic relationship by time scale using wavelets: expenditure and income", *Studies in Nonlinear Dynamics and Econometrics*, Vol. 3, No. 4, 23-42.

Ramsey J.B., Uskinov D., and G.M. Zaslavsky (1995), "An analysis of U.S. stock price behavior using wavelets", *Fractals*, 3:377389

Ramsey, J.B., Gallegati, Marco, Gallegati, Mauro, and W. Semmler (2010), "Instrumental variables and wavelet decomposition", *Economic Modeling*, 27 (6), 1498-1513.

Schleicher C. (2002), "An Introduction to Wavelets for Economists", *Bank of Canada Working Paper,* No. 02-3.

Silverman (1999) "Wavelets in statistics: beyond the standard assumptions", *Phil. Trans. R. Soc. Lond. A*, 357 (1760), 2459-2473

Solow, R. (1997), *Learning from Learning by Doing: Lessons for Economic Growth*, Stanford Univerisity Press.

Steindel and Stiroh (2001), "Productivity: What Is It, and Why Do We Care About It?", *Working paper*, Federal Reserve Bank of New York.

Staiger D., Stock J.H. and M.W. Watson (2001), "Prices, Wages and the U.S. NAIRU in the 1990s", NBER Working Papers no. 8320.

Stock J.H. and M.W. Watson (1999), "Business Cycle Fluctuations in US Macroeconomic Time Series", in Taylor J.B. and M. Woodford *eds.*, *Handbook of Macroeconomics*, Amsterdam: North-Holland.

Tobin, J. (1993), "Price flexibility and output stability: An old Keynesian view", *Journal of Economic Perspectives*, 7 (1), 45–65.

Torrence and Compo (1998), "A Practical Guide to Wavelet Analysis", *Bulletin of the American Meteorological Society*, 79 (1), 6178.

Zavodny M. (1999), "Unions and the Wage-Productivity Gap", *Federal Reserve Bank of Atlanta Economic Review*, 2nd Quarter 1999, 44-53.

Walsh C.E. (2004), "The Productivity and Jobs Connection: The Long and the Short Run of It", *FRBSF Economic Letter*, Number 2004-18.