UNIVERSITÀ POLITECNICA DELLE MARCHE

DIPARTIMENTO DI SCIENZE ECONOMICHE E SOCIALI

PHILLIPS' AVERAGING PROCEDURE AS A 'CRUDE' VERSION OF THE HAAR WAVELET FILTER

MARCO GALLEGATI & JAMES B. RAMSEY

QUADERNO DI RICERCA n. 437

ISSN: 2279-9575

May 2019

Collana curata da:
Massimo Tamberi

Phillips' averaging procedure as a 'crude' version of the Haar wavelet filter

Marco Gallegati* and James B. Ramsey§

Abstract

The aim of this study is to investigate the exact nature of Phillips' (1958) findings. We show that the application of the simplest type of wavelet basis function developed by Haar in 1910 allows to replicate the output of Phillips' data transformation procedure, i.e. the six mean coordinates. Specifically, the resemblance between the coarsest scale level coefficients from the Haar wavelet filter and the six crosses suggests the long-term nature of Phillips' (wage-unemployment) relationship. The application of the Haar wavelet filter allows us to examine the effects of two main features of Phillips' 'unorthodox' averaging procedure: the arbitrarily choice of variable-width intervals and the choice of sorting observations in ascending order of unemployment rate values. Our results show that the arbitrary selection of intervals affects only the smoothness (regularity) of the nonlinear pattern of the wage-unemployment relationship, but not its shape which is determined by sorting and grouping unemployment rate values in ascending order. Indeed, when observations are ordered according to a chronological sequence a simple linear relationship is evident. These findings are robust to different samples, 1861-1913 and 1861-1958.

JEL: B22, C63, E24

Key words: Phillips curve, Haar wavelet transform, Moving average filter.

*DISES, Faculty of Economics "G. Fuà", Polytechnic University of Marche, Piazzale Martelli 8, Ancona, Italy, tel:+390712207114, email: marco.gallegati@staff.univpm.it

§ Department of Economics, New York University, Mercer Street 267, New York, NY, james.ramsey@nyu.edu.

Introduction

Phillips' (1958) averaging method yields six mean coordinates from non-overlapping variable window-widths moving averages of unweighted observations.[1] As evidenced by the filtering literature, since moving average is one of the varieties of discrete low-pass filter with filtering effects depending on the window size (length), averaging corresponds to applying a low-pass filter to the signal. In particular, since the window size affects the resolution level of the analyzed signal, using different averaging lengths is equivalent to viewing data at different resolution levels, as in multiresolution decomposition analysis.

As noted by Gallegati et al.'s (2011), Phillips' averaging procedure may be considered a "crude" version of the simplest type of wavelet (filter) basis function developed by Haar in 1910, the orthogonal Haar wavelet filter.[2] The aim of the present study is to investigate the exact nature of Phillips' (1958) findings by using the Haar discrete wavelet transform (DWT).[3] The application of the Haar DWT to Phillips' original UK data for the 1861-1957 period yields a multiresolution decomposition analysis of the rate of change of money wages and unemployment rate into different components, each associated to a specific frequency resolution level.

We find several interesting results. First, we show that the approximation or scaling coefficients at the coarsest scale level, associated to fluctuations greater than 16 years, bear a striking resemblance to Phillips' six mean coordinates. We interpret this finding as evidence on the long-term nature of Phillips' wage-unemployment relationship. In addition to provide new insights into the low frequency interpretation of Phillips' results, the application of the Haar wavelet filter allows us to investigate whether and how Phillips' choices of intervals and his method of sorting observations matter. Our results show that the arbitrary selection of intervals is responsible for the regularity of the pattern formed by the averages, but not for the eye-catching hyperbolic shape of the wage-unemployment relationship. Indeed, when observations are time-ordered a "simple" linear negative relation is evident, thus implying that the non-linear pattern is strictly dependent on sorting observations of the unemployment rate in ascending order. These findings are robust to different sub-samples, 1861-1913 and 1861-1958.

The structure of the paper is as follows. In section 2 we show the similarities between

---

[1] Phillips' (1958) 'unorthodox' data tranformation procedure consists in reducing 53 observations to 6 average values by first grouping observations sorted in ascending order by the values of the uunemployment rate into several variable-width arbitrarily selected intervals, and then computing for each interval the mean values of money wage inflation and the unemployment rate.

[2] As stated in Wulwick (1996) Phillips was familiar with the ability of data averaging as a tool to filter out stochastic disturbances (Phillips and Quenouille, 1960).

[3] The same goal, that is repeating Phillips (1958) study, is performed by Wulwick and Mack (1990) using kernel regresion analysis.

Phillips' averaging procedure and the Haar wavelet filter formulation. In Section 3 we apply the Haar wavelet transform to Phillips' original UK data over the 1861-1957 period. Section 4 concludes.

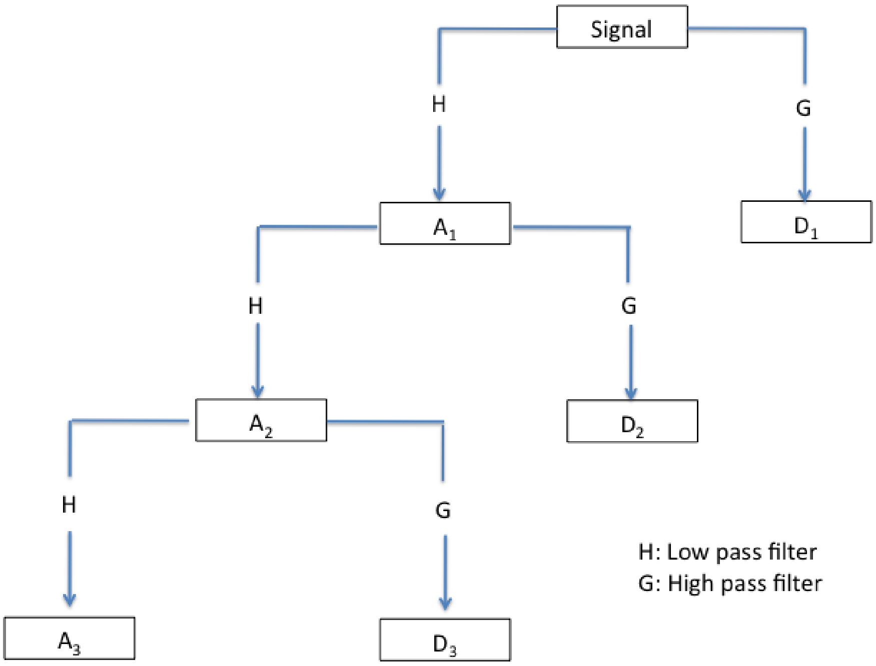
2. Phillips' averaging procedure: a "crude" version of the Haar wavelet filter

The average values computed using Phillips' 'unorthodox' data transformation procedure can be interpreted as the result of the application of a simple unweighted moving average with non-overlapping variable-width windows. Since moving average is one of the varieties of discrete low-pass filter, Phillips' averaging procedure is equivalent to applying a low-pass filter. With the filtering effects depending on the window size (length), different averaging lengths allow to viewing the data at different resolution levels, as in multiresolution decomposition analysis.

Wavelet analysis allows us to analyze a signal in multiple resolutions, each reflecting a different specific frequency range that corresponds to fine, medium and coarse time scales. In order to extract information from a signal at different scales and distinct times, wavelet analysis uses a collection of local basis functions, called wavelets, that are compactly supported, i.e. they have finite length, and are localized both in the time and the frequency domain. Since the wavelet transform can be rewritten as a convolution product, the transform can be interpreted as a linear filtering operation.

Figure 1 shows Mallat's (1989) pyramid algorithm for a 3-level wavelet decomposiiton. Wavelet-based algorithms break a signal down into different time scale components by recursively applying a sequence of filtering and downsampling steps. For the wavelet algorithm to decompose a signal into its different time scale components a dual pair of low-pass and high-pass filters is necessary at each scale level. The first is a non-overlapping moving average of the signal, the latter a non-overlapping moving difference. The first step uses a window width of two. Thus, the wavelet technique takes averages and differences on a pair of values of a signal, with the averages giving a coarse signal and the differences the fine details. Then the algorithm shifts over by two values and calculates another average and difference on the next pair. At each successive scale level the window width is dilated (doubled) and, since wavelet algorithms are recursive, the smoothed data of the previous scale level, i.e. the averages, become the input for getting new (smoothed) approximations and detail components at coarser resolution levels. At each downsampling step, the wavelet algorithm decomposes a signal into two subsignals each with a length which is half the size of the input dataset. In the end, the application of the Discrete Wavelet Transform (DWT) to a dyadic length vector of observations ($N=2^J$ for some positive integer J) yields N wavelet coefficients, that is $N = N/2 + N/4, + ... + N/2^{J-1} + N/2^J + N/2^J$, where the number of coefficients at each scale level J is (inversely) related to the width of the wavelet function.

Figure 1 - Mallat filter scheme for a 3-level wavelet decomposiiton



The Haar wavelet, proposed in 1910 by Alfred Haar, is a piecewise constant function on the real line that can take only three values: 1, -1 and 0. Therefore, Haar wavelets are the simplest orthonormal wavelet basis function with compact support (see Figure 2).[4] The Haar scaling and wavelet filters are given, respectively, by

$$H = (h_0, h_1) = 1/\sqrt{2},$$

and

$$G = (g_0, g_1) = (1/\sqrt{2}, -1/\sqrt{2}).$$

Three basic orthonormal properties characterize the Haar scaling and wavelet filters.

$$\textstyle\sum_l h_l = 0 \quad \text{and} \quad \sum_l g_l = \sqrt{2}$$

$$\textstyle\sum_l h^2_l = 1 \quad \text{and} \quad \sum_l g^2_l = 1$$

$$\textstyle\sum_l h_l\, h_{l+2n} = 0 \quad \text{and} \quad \sum_l g_l\, g_{l+2n} = 0 \qquad \text{for all integers } n \neq 0$$

The first property guarantees that g is associated with a difference operator and thus identifies changes in the data and that h may be viewed as a local averaging operator. The second property, unit energy, ensures that the coefficients from the wavelet transform preserves energy and, therefore, will have the same overall variance as the data.[5] The third property guarantees orthogonality to even shifts.

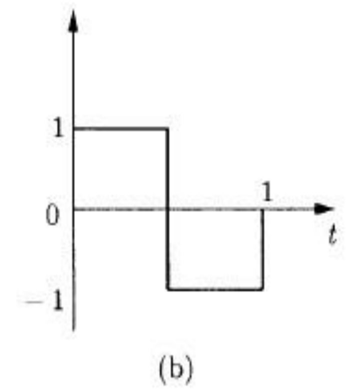
---

[4] The scaling and wavelet filters in Figure 2 assume unity values.

[5] The normalization factor $\sqrt{2}$ ensure that that the dilated and translated Haar function satisfies the second property in the wavelet definition.

The Haar functions provide the two most elementary high-pass and low-pass filters. The wavelet filter G, with filter coefficients g=[*1/√2,−1/√2]*, by computing the difference between any two adjacent samples simply accomplishes differences. The scaling filter H, with filter coefficients h=[*1/√2,1/√2]*, represents a moving average filter because it essentially computes the average of successive pairs of non-overlapping values. Thus, for the Haar scaling filter the filtered signal is a weighted average of observations with the filter coefficients (h_0, h_1) used as weights. Hence, as noted by Gallegati et al. (2011), Phillips' averaging procedure, using unity weights (as in the left panel of Figure 2), can be considered a 'crude' version of the Haar scaling filter.[6]

Figure 2 – Haar scaling (left) and wavelet (right) filters



3. Haar discrete wavelet transform of Phillips' original data

Since both the estimated regression and the smooth hyperbolic curve are based on those six average values, the 'crude' statistical method used by Phillips can be considered crucial in getting his original results (Wulwick, 1987). The analysis of the effects Phillips' averaging procedure based on the analogy with the Haar scaling filter is carried out by applying the Haar wavelet transform to Phillips' original dataset for the full sample period 1861-1957.

The application of the DWT using the Haar (1910) wavelet filter for a 3-level decomposition, i.e. J=3, produces three vectors of wavelet detail coefficients, D_1, D_2 and D_3, and three vectors of scaling coefficients, A_1, A_2 and A_3. Table 1 presents the frequency domain interpretation of each detail and approximation level component using annual data. The detail levels D_1, D_2 and D_3 represent non-overlapping changes in the rate of change of money wages and in the unemployment rate at different frequency ranges, i.e. 2-4, 4-8 and 8-16 years, respectively. The approximation levels A_1 represents non-overlapping averages of the rate of change of money wages and unemployment rate greater than 4 years. Moreover, by adding D_2 and D_3 to the lower "smooth" component A_1 and A_2 we get, respectively, the additional levels of approximation A_2 and A_3 capturing fluctuations greater than 8 and 16 years.

Table 1*: Frequency domain interpretation of multiresolution decomposition analysis*

[6] In this sense we can say that Phillips was, involuntarily, the first economist to use wavelets as a tool of analysis.

with annual data for J=3

Detail level, D_j	**Years**	**Approximation level, A_J**	**Years**	**Window width**
D_1	2-4	A_1	from 4 to ∞	2
D_2	4-8	A_2	from 8 to ∞	4
D_3	8-16	**A_3**	**from 16 to ∞**	**8**

The last column in Table 1 denotes the number of values used in calculating the scaling coefficients at that scale level. In reducing the number of observations from 52 to 6 Phillips averaged into each interval a number of raw points varying from 6 to 12 (i.e. 6, 10, 12, 5, 11, 9). Since the averaging length (window width) determines the frequency resolution of the decomposition, the approximation component at the scale level J=3, yielding 7 scaling coefficients each stemming from a window-width with length 8, can provide a useful benchmark for evaluating Phillips' averaging procedure. Indeed, the number of values of money wage rates and the unemployment rate used, on average, by Phillips for computing the mean coordinates, the *six crosses*, roughly corresponds to the number of wavelet scaling coefficients at the approximation level A_3. Thus, beyond the reasons for using moving averages,[7] the effect of Phillips' averaging procedure is to identify a "locus of long-run equilibrium points" (Desai, 1975, p.2) whose frequency resolution level corresponds to fluctuations greater than 16 years.[8]

Table 2 – Mean coordinates using Phillips' intervals and Haar A_3 scaling coefficients with observations sorted by ascending unemployment rate values (1861-1913)

Phillips' coeffs	*ur*	*1.516*	*2.351*	*3.483*	*4.49*	*5.954*	*8.372*
	dw	*5.058*	*1.547*	*0.848*	*0.346*	*-0.182*	*-0.350*
Haar A_3 coeffs	*ur*	*1.650*	*2.426*	*3.337*	*4.069*	*5.600*	*6.862*
	dw	*4.610*	*1.117*	*0.755*	*0.852*	*0.431*	*-0.991*

[7] Averaging within groups has the property that the rate of change of the unemployment rate in Phillips equation is set to zero (Wulwick, 1989).

[8] This is also consistent with Phillips' statement that "each cross (or mean coordinate) gives an approximation to the rate of change of wages which would be associated with the indicated level of unemployment if unemployment were held constant at that level (Phillips, 1958, p. 290).

Note: The values in the row "Phillips' averages" are calculated by averaging observations sorted by ascending values of the unemployment rate for Phillips' (1958) intervals (0-2, 2-3, 3-4, 4-5, 5-7, 7-11). The 7th Haar A_3 scaling coefficient, not included here, is reported in Figure 3.

Table 2 presents the values of the scaling coefficients at level A_3 and the mean coordinates of the unemployment rate and money wage rates when observations are sorted by ascending values of the unemployment rate for the period 1861-1913 and Phillips' intervals are used, i.e. 0-2, 2-3, 3-4, 4-5, 5-7, 7-11. The coefficient values shown in Table 2 are also visually displayed in the left panel of Figure 3 with filled blue circles corresponding to A_3 level approximation coefficients, while Philips' averages are marked with a cross as in his original paper. The position of the filled blue circles and the crosses is quite well aligned in the left panel of Figure 3, except for values of the unemployment rate higher than 5-6%. Similar findings[9] are provided in the right panel of Figure 3 where the same analysis is replicated for the period 1861-1957.[10] All in all, these findings are consistent with Desai's (1975) interpretation on the long-term nature of the Phillips' relationship between wage inflation and unemployment.

The previous findings allow using the approximation level component A_3 as a benchmark for the evaluation of the effects of Phillips' averaging procedure on his results. In particular, we can examine separately the effects of i) using variable window-widths, and ii) sorting observations according to increasing values of the unemployment rate. In Figure 3 the comparison of Phillips' crosses with the filled blue circles representing the Haar A3 level approximation coefficients, based on fixed regular window-widths, allows us to analyze the effect of Phillips' choice of intervals. Interestingly, the blue filled circles detect an irregular graph of averages with several ranges of unemployment characterized by a positive relationship positive. Similar findings are provided by Wulwick (1989) that, after experimenting alternative intervals similar to Phillips' intervals, her conclusion was that "only Phillips' intervals resulted in the smooth *hyperbolic* graph of averages" (Wulwick, 1989, figs. 4 and 5, p.181-2). Therefore, while the shape of the wage-unemployment relationship does not seem to be affected by using fixed or variable window widhts, the regularity of the pattern formed by the averages is strictly related to the arbitrary selection of intervals made by Phillips.[11]

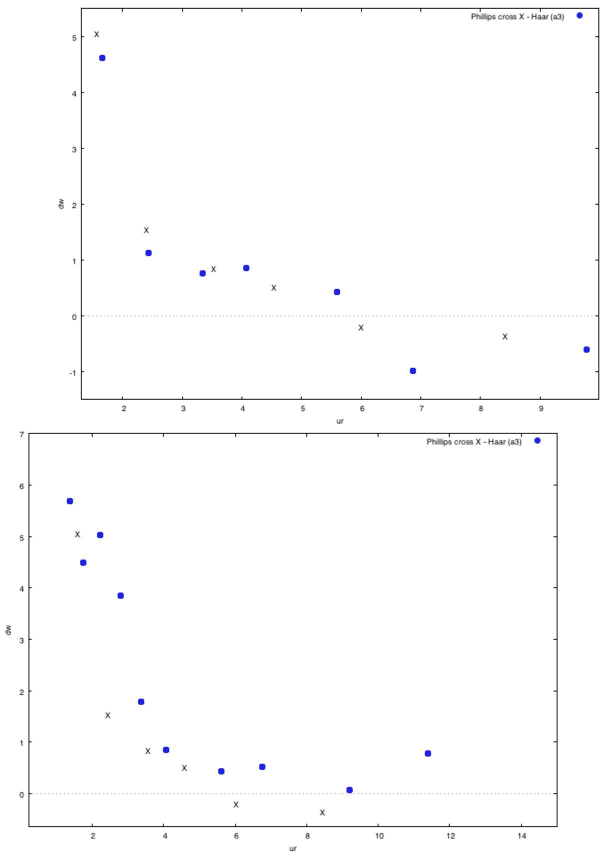
***Figure 3** – The low frequency resolution nature of Phillips' relationship*

---

[9] The main difference being that A_3 coefficients are slightly shifted upward.

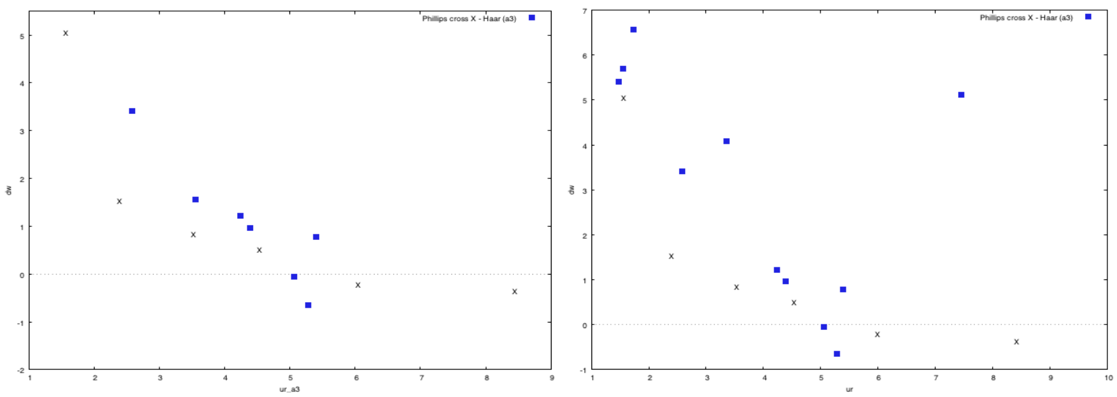
[10] In the right panel outliers have been excluded by limiting the range of x-axis and y-axis. Outliers have been defined as those values greater than 10% for the unemployment rate and greater than 15% for the money wage rate. Such values are all included in the 1918-1923 period.

[11] Although it may be appear a secondary feature it is not if we think that the failure in his 1955's book to draw or fit an eye-catching downward sloping convex curve on a scatter diagrams similar to Phillips' probably prevented from the attribution of the wage-unemployment relationship the label Brown- or Brown-Phillips curve (Corry, 2001, Button, 2018).



Note: Philips' averages (x) and A_3 scaling coefficients (•) with observations sorted by ascending unemployment rate values: 1861-1913 (left) and 1861-1957 (right). In the right panel values higher than 15% for money wage rates and the unemployment rate are excluded because they can be considered as outliers (they refer to years between 1918 and 1923).

Figure 4 *– The long-term relationship net of the effect of sorting observations in ascending order of unemployment rate values*



Note: Philips' averages (x) and A_3 Haar scaling coefficients (•) with observations sorted by time: 1861-1913 (left) and 1861-1957 (right). In the right panel values higher than 15% for money wage rates and the unemployment rate are excluded because they can be considered as outliers (they refer to years between 1918 and 1923).

The Haar wavelet transform may also be useful in order to detect the effect of sorting observations according to ascending values of the unemployment rate. To this aim we apply the Haar wavelet transform to time ordered observations. Figure 4 shows the Haar A_3 level approximation coefficients (filled blue squares) and Philips' averages (crosses) for the periods 1861-1913 (left panel) and 1861-1957 (right panel). Interestingly, the pattern diplayed by the filled blue squares in both panels of Figure 4 contrasts strikingly with that delineated by Phillips' crosses. In particular, differently from the evidence presented in Figure 3 the filled blue squares identify a simple linear pattern for the wage-unemployment relationship. This finding suggests that the nonlinear pattern (shape) of the wage-unemployment relationship may depend on grouping obervations into ascending values of the unemploment rate. What emerges from our analysis is that the ordering choice is highly influential for Phillips' results. Indeed, such ordering is responsible for obscuring the "true" linear relationship that is otherwise evident at the coarsest scale when observations are sorted througha time sequence.

In sum, our results suggest that, beyond Phillips' reasons for averaging, understanding the effects of Phillips' unorthodox data transformation procedure is crucial for assessing the precise nature of Phillips' findings in terms of frequency resolution, first of all the low frequency nature of the wage-unemployment relationship. Interestingly, in terms of the controversy between Desai (1975) and Gilbert (1976) our findings can reconcile both Gilbert's and Desai's views. Although the crude curve-fitting procedure exploited by Phillips is adopted with the only purpose to get a practical estimation result, and in this sense Gilbert's view is right, the choice of averaging observations, which is equivalent to viewing long-term information, supports Desai's view that the Phillips' procedure identifies a long-run relationship.

4. Conclusions

In this paper we provide new insights into the low frequency resolution nature of Phillips' wage-unemployment relationship. Given the similarities between Philips' averaging procedure and the simplest wavelet basis function, i.e. the Haar (1910) wavelet filter, we show that Phillips' averages (mean coordinates) resemble the long-term components of wages and unemployment corresponding to fluctuations greater than 16 years. Most importantly, we find that the choice of sorting observations by ascending values of the unemployment rate is crucial for reaching the goal of estimating the eye-catching nonlinear hyperbolic shape of the wage-unemployment relationship.

Since the statistical and economic interpretation, as well as the policy implications, of the Phillips curve are crucially related to its short- or long-run nature, we find that our results, although limited to Phillips' original dataset, may provide interesting insights into the endless debate over the form of the wage-unemployment relationship.

References

BUTTON, K. (2018). A.J. Brown, 'Phillips' Curve', and Economic Networks in the 1950s. *Journal of the History of Economic Thought*, 40, 243-64.

BROWN, A.J. (1955). *The Great Inflation, 1939-1951.* Oxford: Oxford University Press.

CORRY, B. (2001). Some myths about Phillips' Curve. in P. Arestis, M. Desai and S. Dow (eds). *Money, Macroeconomics and Keynes: Essays in Honour of Victoria Chick, Volume 1*, London: Routledge.

DESAI, M. (1975). The Phillips Curve: A Revisionist Interpretation. *Economica*, 42, 1-19.

GALLEGATI, MARCO, GALLEGATI, MAURO, RAMSEY, J.B. and SEMMLER, W. (2011). The US wage Phillips curve across frequencies and over time. *Oxford Bullettin of Economic and Statistics*, 73, 489-508.

GILBERT, C.L. (1976). The Original Phillips Curve Estimates. *Economica*, 43, 51-57.

HAAR, A. (1910). Zur Theorie der orthogonalen Funktionensysteme. *Mathematische Annalen*, 69, 331-371.

LIPSEY R.G. (1960). The Relation between Unemployment and the Rate of Change of Money Wage Rates in the United Kingdom, 1862-1957: A Further Analysis. *Economica*, New Series, 27, 1-31.

MALLAT, S. (1989). A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Transaction Pattern Analysis*, 11, 674-693.

PHILLIPS, A.W.H. (1958). The Relation between Unemployment and the Rate of Change of Money Wage Rates in the United Kingdom. 1861-1957. *Economica* 25, 283-99.

PHILLIPS, A.W.H. and QUENOUILLE, M.H. (1960). Estimation, Regulation and Prediction in Interdependent Dynamic Systems. *Bulletin de l'Institute de Statistique*, 37, 335–43.

WULWICK, N.J. (1987). The Phillips curve: Which? Whose? To do what? How? *Southern Economic Journal*, 54, 834-57.

WULWICK, N.J. (1989). Phillips' Approximate Regression. *Oxford Economic Papers*, New Series, 41, 1, *History and Methodology of Econometrics*, 170-188.

WULWICK, N.J. and MACK Y.P. (1990). A Kernel Regression of Phillips' Data. *Working Paper No. 40*, New York, Jerone Levy Economics Institute, Bard College.