Mr Phillips and the medium-run: temporal instability vs. frequency stability

Michele Fratianni Marco Gallegati
Federico Giri

Working paper no. 155

July 2019

Mr Phillips and the medium-run: temporal instability vs. frequency stability

Michele Fratianni*

Marco Gallegati†

Federico Giri‡

July 18, 2019

Abstract

The paper goes back to the original insight by Phillips and investigates the negative relationship between money wage inflation and the unemployment rate occurring at frequency bands that stretch beyond those of the business cycle. We use UK annual data for the period 1861-2015, and post-WWII quarterly data from 1960 to 2016. The two critical findings are that the wage Phillips Curve is predominantly a medium-run phenomenon, comprised in the 8-to16-year frequency band, and that the curve disappears beyond this range. Similar conclusions are reached using the post-WWII quarterly sample: at the aggregate level and at high frequencies the PC relationship is unstable over time, whereas in the frequency range between 32 to 64 quarters (our medium run timescale), this time dependency disappears.

Keywords: Phillips Curve, frequency bands, wavelet, medium run

JEL classification codes: E00, E30, E31, E32.

*Michele Fratianni,Kelley School of Business, Bloomington, Indiana, USA and Universitá Politecnica delle Marche, Ancona, Italy. E-mail: fratiann@indiana.edu.

†Marco Gallegati, Universitá Politecnica delle Marche, Piazzale Martelli n 8 Ancona, Italy. E-mail: marco.gallegati@staff.univpm.it.

‡Corresponding author: Federico Giri, Universitá Politecnica delle Marche, Piazzale Martelli n 8 Ancona, Italy. E-mail: f.giri@staff.univpm.it.

1 Introduction

Sixty-one years after A.W. Phillips 1958 published the "The relation between unemployment and the rate of change money wage rates in the United Kingdom, 1861-1957", the empirical validity of its Curve (PC for short) remains a hotly debated issue both in the world of academic research and policy applications. The vastness of the contributions defies any attempt to provide a comprehensive review of the literature; a short list of the highlights of this research includes Samuelson and Solow (1960), Phelps (1967), Friedman (1968), Lucas (1972, 1973), Sargent and Wallace (1975), Gordon (1982, 2011), King and Watson (1994), Gali and Gertler (1999), Staiger et al. (1997a,b), Haldane and Quah (1999), Mankiw (2001), Mavroeidis et al. (2014), Galí (2011) and Hall and Sargent (2018).

Much of this research, possibly reflecting the influence of Samuelson and Solow (1960) on the economics profession, focused on price inflation rather than wage inflation. That held true also in the policy world. A very early policy application of the PC was the Kennedy-Johnson tax cuts of 1964 and 1965 in the US, the objective being to reduce the unemployment rate to 4 per cent (Economic Report of the President (1962, p.46).[1] According to Walter Heller 1964, p.237, the then chairman of the President's Council of Economic Advisers, "... we cannot say we have a conclusive test of our 4 per cent full-employment bench mark and its attendant Phillips curve assumptions. What we can say is: so far, so good, adding that the 5-percenters have already been proved wrong, while we (interim) 4-per- centers have not. And the progressive stepup in manpower programs...may well be reducing the critical unemployment level at which we have to choose between further shrinkage of unemployment and accelerating increases in the price level."[2] Also central banks did accept the relevance of the relationship between rates of inflation and level of real economic activity (Mavroeidis et al., 2014, p.124), despite evidence that the relationship had shifted and become less reliable over time (Blanchard et al., 2015; BIS (2017, p. 63-67); Carney, 2017; Cunliffe, 2017; FOMC, 2018).[3]

This paper goes back to the origin and focuses on the negative relationship between money wage inflation and the unemployment rate occurring at frequency bands stretching beyond the business cycle. Our point of departure is Phillips' unorthodox data transformation, consisting in reducing the 53 observations from 1861 to 1913 –rearranged in ascending order in terms of the unemployment rate– in six average values of arbitrarily selected intervals, which were used to estimate the hyperbola, $\Delta w = -0.90 + 9.64U^{-1.39}$,

[1]It includes also the report written by the President's Council of Economic Advisers which included high-caliber Keynesian economists James Tobin, Arthur Okun and Otto Eckstein

[2]But, as (Gordon, 2011, p.14-15) notes, "the zero-inflation unemployment rate was higher than 4 percent and the tax cuts would have implied an acceleration of inflation even without the further loosening of the fiscal floodgates due to the Vietnam War."

[3]The durability of the PC in the literature can be assessed with the use of scholarly search engines. According to EconLit, 674 scholarly journal articles have PC in the document title (search date: March 4, 2018). The data bank of RePeC (Research Papers in Economics) listed, on the same date, 1,019 articles and 1,669 working papers mentioning "Phillips Curve" in their abstract. In a Google Scholar search, done again on the same date, the response to the query "Phillips Curve is vertical" was 431,000 documents. "Phillips Curve is dead" 94,000, "Phillips Curve and central banking" 55,500, "Phillips Curve is steeper" 55,100, "Phillips Curve is alive" 32,700, "Phillips Curve is flatter" 22,100.

where Δw denotes the rate of change of money wages and U = the rate of unemployment.[4] Phillips dismissed price inflation as a contributing factor for his empirical model arguing that an excess demand for labor would have induced firms to bid up money wages independently of inflation. This assumption and the restriction that changes in unemployment were essentially set to zero by averaging observations over the selected time intervals suggests that Phillips was interested in a long-run relationship (Desai, 1975).[5] Phillips' procedure was unfamiliar to economists of the time, but today can be interpreted it as a simple moving average with non-overlapping variable-width windows. Since a moving average removes high-frequency fluctuations of the time series, Phillips' averaging procedure is somewhat equivalent to applying a low-pass filter to the data. In brief, Phillips performed a crude form of wavelet analysis (Gallegati et al., 2011, footnote 2; Gallegati and Ramsey, 2019). The more popular replication of the PC by Lipsey (1960) relied, instead, on familiar aggregated time-scale estimation techniques. It "became the standard method of analysis in economics by 1960...[and] inspired a booming industry in Phillips curve estimation" (Wulwick, 1996, p.410).

Wavelet methodology was applied by Gallegati et al. (2011) to estimate the wage PC for the US along time and frequency domain. Under this methodology, one first decomposes the variables of interest into their time-scale components and then estimates the PC on a scale-to-scale level (disaggregation by frequency bands) so as to study simultaneously statistical properties in both time and frequency domain. Using the same methodology, we will revisit Phillips' original PC in the UK. The data will be Phillips' own and an expansive annual and quarterly samples. Recently, Aguiar-Conraria et al. (2019) and Mutascu (2019) use wavelet methodology to investigate the short-run instability of the US PC and find that the PC is vertical in the long run. Our paper differs from theirs in that: we focus on the original Phillips' 1958 paper; we use UK data covering a time span from 1861 until 2016; and concentrate on the *medium-run* frequencies rather on the difference between the short run and the long run.

The main motivation of the paper is not to assess the time dependency (or temporal instability) of the PC, which is a dominant aspect in literature, but its frequency dependency (or frequency instability). To this end we perform two exercises. The first is to replicate the original work by Phillips using the modern wavelet method with the intent to corroborate whether the PC is a medium-run relationship between money wage inflation and the rate of unemployment instead of a short-run tradeoff, as postulated by Samuelson and Solow (1960). A medium-run PC relationship would refute the notion of an exploitable short-run tradeoff between inflation and unemployment. The second is to extend the analysis beyond the

[4] For subsequent periods, the intercepts of the curves were adjusted by hand using the estimate of the 1861-1913 years.

[5] Phillips' argument goes as follows: "For suppose that productivity is increasing steadily at a rate of, say, 2 per cent per annum and that aggregate demand is increasingly similarly so that unemployment is remaining constant at say, 2 per cent. Assume that with this level of unemployment and without any cost of living adjustment wage rates rise by, say, 3 per cent per annum as a result of employers' competitive for labour and that import prices and the prices of other factor services are also rising by 3 per cent per annum. Then retail prices will be rising on average at the rate of 1 per cent per annum...Under these conditions the introduction of cost of living adjustments in wage rates will have no effect, for employers will merely be giving under the name of cost of living adjustment part of the wage increases which they would in any case have given as a result of their competitive bidding for labour." (p.284).

original sample to check whether the medium-run PC survives till present days. Our expanded data set goes from 1861 to 2015 for yearly observations and from 1960:Q1 to 2016:Q4 for quarterly observations.

Our main findings are that, for the period 1861-1913, Phillips detected a negative relationship between wage inflation and unemployment that is mainly present beyond the business-cycle frequency band. This medium-run relationship, not only emerges from the original Phillips dataset, but survives until present days using a long stretch of historical annual data. The other significant finding is that the curve disappears in the long run after WWII; see Figure 3, panel c. Similar conclusions are reached for the post-WWII quarterly data: at the aggregate level and at high frequencies the PC relationship is unstable over time. But, in the frequency range between 32 to 64 quarters (our medium run timescale), the relationship is strong and stable throughout the sample.

The structure of the paper is as follows. We start in Section 2 by estimating a canonical relationship between money wage inflation and the unemployment rate using aggregate-level data. In Section 3 we analyze the PC in the time-frequency domain over the long stretch of annual data, and estimate wage PCs at different time-scales and over different time intervals. We repeat the exercise for post-WWII quarterly data in Section 4. Section 5 discusses the significance of our findings in relation to the literature. Conclusions are drawn in Section 6. An Appendix contains a description of the data, with accompanying graphs, a concise summary of wavelet analysis, and a robustness exercise using bandpass filter methodology.

2 The canonical PC

We start our empirical work estimating a canonical PC relationship at the aggregate level, that is with crude data and ignoring a decomposition in frequency bands. Our specification follows Staiger et al. (1997a,b):

$$\Delta w_t = \alpha_0 + \alpha_1 U_t + \alpha_2 \Delta LP_t + \alpha_3 \pi_t + \alpha_4 CTR_t + \epsilon_t, \tag{1}$$

where Δw_t = quarter-to-quarter difference of the money wage rate, U_t = the rate of unemployment, ΔLP_t = delta log of labor productivity, π_t= price inflation, CTR_t = control variables, and ϵ_t = error term.

The nominal wage rate is measured as the composite average of weekly earnings; labor productivity is measured as output per hour worked, and price inflation is measured in terms of the GDP deflator at market prices. All data come from *The Bank of England's collection of historical macroeconomic and financial statistics*; see Ryland et al. (2010) and the Appendix A for details. Controls are defined as dummy variables: WWI stands for the first world war years (1914-1918), GD for the Great Depression years (1929-1933), WWII for the second world war years (1939-1945), Oil for the two oil shocks of the Seventies (1973-1980) and GFC for the Great Financial Crisis years (2007-2010).

Table 1 presents the annual estimates of Equation 1, except for the intercept, with their associated

Table 1: Dependent Variable: Δw_t, Bank of England, annual data, 1861-2015

	1861-2015	1861-1913	1923-1948	1960-2015
U_t	-0.095	**-0.622**	-0.042	-0.089
	(0.082)	(0.116)	(0.072)	(0.068)
ΔLP_t	**0.371**	-0.091	**0.523**	0.226
	(0.076)	(0.083)	(0.109)	(0.122)
π_t	**0.834**	0.225	**0.735**	**0.923**
	(0.082)	(0.116)	(0.088)	(0.101)
WWI	**0.064**			
	(0.010)			
GD	0.001		0.003	
	(0.007)		(0.004)	
WWII	0.009		**0.023**	
	(0.005)		(0.004)	
OIL	0.016			-0.002
	(0.012)			(0.015)
GFC	-0.002			**-0.013**
	(0.004)			(0.005)
R^2	0.863	0.553	0.907	0.882
$\mathcal{L}$	386.2	173.9	78.06	154.3

Coefficients in bold are significant at a 5 % level. Standard errors (in parenthesis) are robust to autocorrelation and heteroskedasticity using an HAC estimator for the variance and covariance matrix.

standard errors adjusted for possible serial correlation and heteroskedacity in the error term. Four samples are considered: the longest sample, 1861-2015, the Phillips' sample, 1861-1913, the interwar period, 1923-1948, and the post-WWII era, 1960-2015. The salient findings are that the PC effect, the impact of the rate of unemployment on wage inflation is statistically significant (at least at the 5 percent confidence level) only in the pre-WWI period: Phillips was judicious, or lucky, in selecting the sample. For the rest, the PC, with aggregate data, is time dependent. The rate of price inflation is statistically significant in all periods, except in the Phillips' period, thus confirming the original outcome that prices were not relevant in the PC. The coefficient of is statistically not different from unity in the post-WWII period and very close to unity in the expansive sample, a result that is consistent with a vertical PC (Benati, 2015). There is some evidence that the war years raised wage inflation, while the Great Financial Crisis lowered them.

Table 2 presents the quarterly estimates of Equation 1 for the post-WWII period. Four time windows of decreasing size are considered: 1960-2016, 1970-2016, 1980-2016, and 1990-2016 . In all four cases, the quarterly estimates are in line with the post-WWII annual estimates of Table 1: no variable is statistically significant, again at the 5 percent confidence level, with the exception of price inflation, a finding that is consistent with those reported by Gallegati et al. (2011) for the US.

Table 2: Dependent Variable: Δw_t, Bank of England, quarterly data, 1960:1- 2016:4

	1990:1- 2016:4	1980:1- 2016:4	1970:1- 2016:4	1960:1- 2016:4
U_t	-0.104	0.187	-0.103	-0.129
	(0.134)	(0.109)	(0.135)	(0.092)
ΔLP_t	0.179	0.029	-0.015	0.046
	(0.170)	(0.144)	(0.155)	(0.136)
π_t	**0.268**	**0.622**	**0.681**	**0.717**
	(0.134)	(0.119)	(0.073)	(0.083)
OIL			2.168	**2.927**
			(1.371)	(1.420)
GFC	-1.465	**-2.216**	**-2.790**	**-2.831**
	(0.887)	(0.962)	(0.902)	(0.824)
R^2	0.061	0.381	0.570	0.521
$\mathcal{L}$	-278.6	-388.1	-528.6	-638.4

Coefficients in bold are significant at a 5 % level. Standard errors are robust to autocorrelation and heteroskedasticity using an HAC estimator for the variance and covariance matrix.

Figure 1: Wage inflation and unemployment, annual (upper panel) and quarterly data (lower panel)

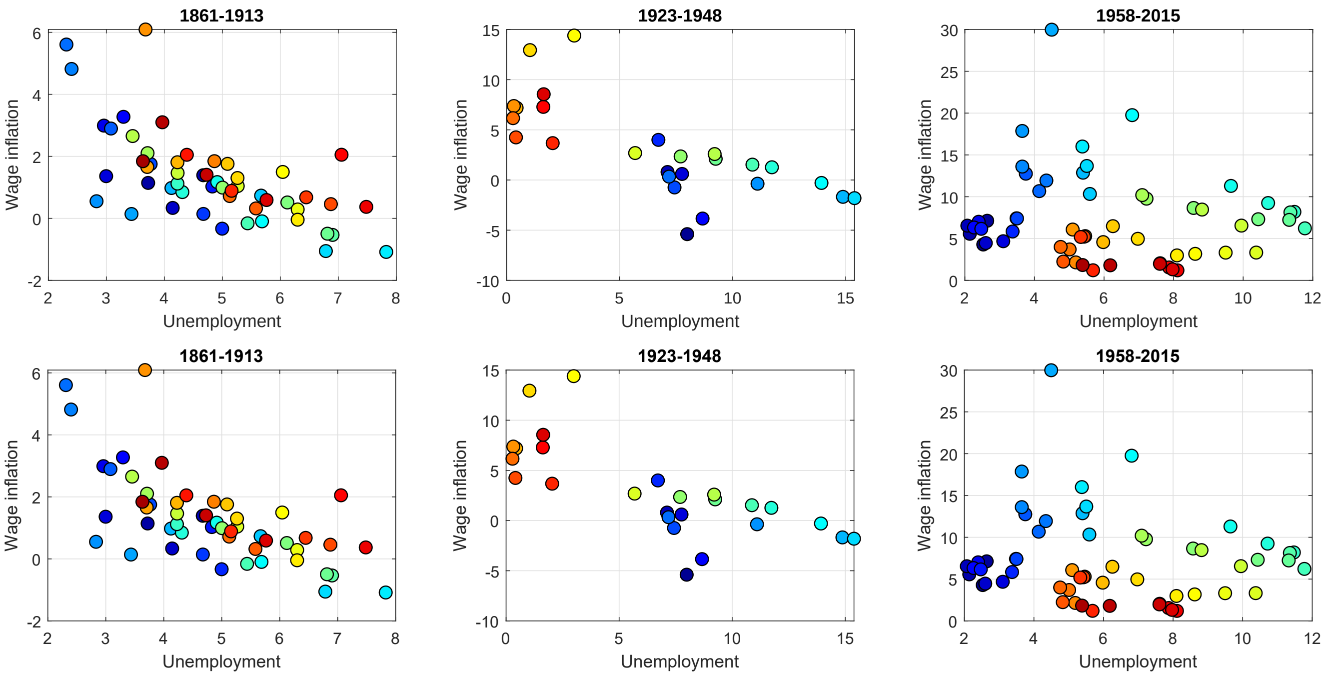
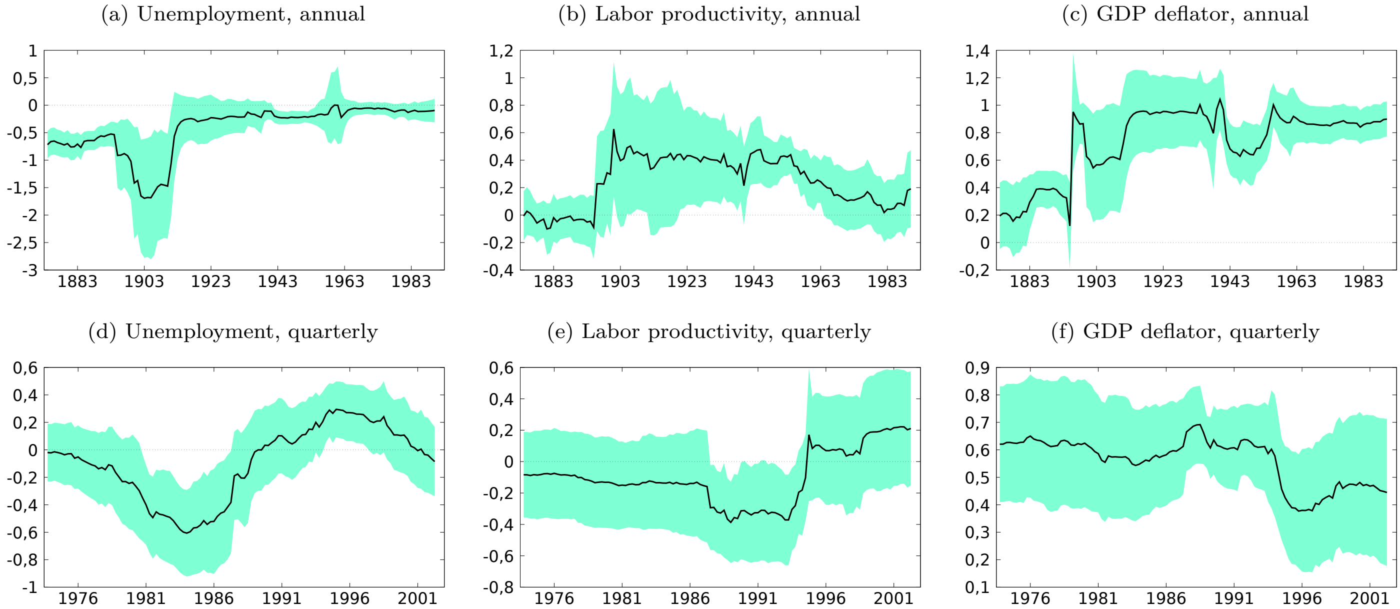


Figure 2: Rolling window coefficients, UK, annual (upper panel) and quarterly data (lower panel)



Note: annual data are from 1856 to 2015. We use a rolling window using the % 25 of the observations. Years on the x-axis are the middle value of the interval. Quarterly data are from 1960:Q1 to 2016:Q4. We use a rolling window using the % 50 of the observations. Years on the x-axis are the middle value of the interval. In both cases standard errors are robust to autocorrelation and heteroskedasticity using an HAC estimator for the variance and covariance matrix.

Figure 1 displays scatter plots of the PC relationship at aggregate frequencies for annual and quarterly data, respectively. The canonical negative relationship between money wage inflation and unemployment emerges in the period 1861-1913 and during the interwar period (1923-1948). It then disappears after WWII, a result that is corroborated by quarterly data for various sub-sample (1960-1979, 1980-1999 and 2000-2016).

Figure 2 displays the estimates of rolling-window regressions of Equation 1. The size of the windows draws from 25 percent of the total annual data and 50 percent of the total quarterly data. The windows are moved forward by an observation so that we can plot, at mid-point of the window size, a continuous time line of estimated coefficients for the unemployment rate, productivity growth and the rate of inflation. The coefficient lines are surrounded by a confidence band set at 1.96 times the standard error. For annual data, the PC relationship holds in the pre-WWI period and then disappears; for quarterly data it holds from the end of the 1970s to the mid 1980s and then disappears.

In sum, the relationship between money wage inflation and the unemployment rate, with aggregate data, suffers from a high degree of time dependency. Time dependency holds also for the coefficient of labor productivity. The impact of price inflation on wage inflation, instead, is consistently positive and tends towards unity under annual data.

3 The wage PC in time-frequency domain over the long stretch

After the pioneering contributions of Ramsey and his co-authors (1995; 1996), the wavelet methodology has gained currency in analyzing economic and financial time series (e.g. Gallegati et al., 2014). The main advantages of the wavelet transform are threefold: first, being a local rather than global transform, it allows handling both stationary and non-stationary time series.[6] Second, it provides a systematic way of performing band-pass filtering, in the sense of not being committed to any particular class of model.[7] Third, the multiresolution nature of the wavelet decomposition analysis attains an optimal trade-off between time and frequency resolution levels. In particular, the wavelet transform provides a flexible time-scale window that narrows when focusing on small-scale features and widens on large-scale features, thus displaying good time resolution (and poor frequency resolution) for short-lived high-frequency phenomena and good frequency resolution (and poor time resolution) for long-lasting low-frequency phenomena.

Multiresolution decomposition analysis may be performed using two types of wavelet transform: the continuous (CWT) and the discrete wavelet transform (DWT). The first computes wavelet transform coefficients at all times and scales, the latter provides its discretized version by using only a limited number of translated and dilated versions of the wavelet basis function in order to decompose the original signal. A brief review of wavelet analysis is provided in the Appendix B.

The wavelet coherence and phase difference represent the two main tools of the CWT. The wavelet coherence measures how the strength of the correlation between two time series changes in the time-frequency plane and represents an expansion of the correlation analysis into the time and the frequency domains, so that it is possible to detect those time scales at which the relationship is statistically significant from those scales at which it is not. The phase difference provides the relative phase between signals' variations, that is their lead/lag relationship.

Figure 3 shows the squared wavelet coherence between wage inflation and unemployment rate for the UK over different dataset and samples obtained using the MatLab package developed by Grinsted et al. (2004). Time is recorded on the horizontal axis and periods, with the corresponding scales of the wavelet transform, on the vertical axis. Hence, reading across the graph at a given value for the wavelet scaling, one sees how the power of the projection varies over time at a given scale, while reading down the graph at a given point in time one sees how the power varies with the wavelet scale (see Ramsey et al., 1995).[8] For comparison with Phillips' findings, in panel (a) we use his original dataset for the period

[6]This property may be particularly useful when dealing with complex, non-stationary signals like historical time series, since their secular movements are likely to exhibit structural changes due to shocks such as wars, economic and/or financial crises.

[7] This contrasts with Baxter and King (1999) and Christiano and Fitzgerald, 2003 approximate band-bass filters whose optimizing criteria implicitly define the specific class of model for which the approximating filter is optimal.

[8]The color code ranges for power ranges from blue (low power) to red (high power), with regions with warmer colors corresponding to areas with wavelet transform coefficients of large modulus. 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. Right (left) arrow means that the two variables are in-phase (anti-phase). If the right arrow points up (down)

1861-1947 (reproduced by Wulwick, 1996) and transform the data applying a centered first difference, $(W_{t+1} - W_{t-1})/2W_t$, In Panel (b) wage inflation is measured as the first difference of the logs. Panel (c) covers the entire sample 1856-2015. Compare first panels (a) and (b) of the figure. In both there is significant negative correlation in the 4 to 16 year range, but in (a) another significant correlation emerges in the lower 16 to 32 year range. This is due to Phillips' definition of wage inflation, a centered first difference that smooths the series further than a traditional first difference. The coherence of the longest sample period in panel (c) confirms that the medium-run results hold both before and after WWII, whereas the long-run relation does not survive the post-WWII era.

The CWT contains a high amount of redundant information. A more parsimonious representation is provided by the discrete wavelet transform (DWT) where only a limited number of translated and dilated versions of the wavelet basis functions are used to decompose the original signal. Empirical applications apply MODWT is a compromise between the continuous and discrete wavelet transforms with a relative ease of implementation and computational costs (Aguiar-Conraria and Soares, 2014).[9]

The MODWT decomposes a time series into different frequency bands. Table 3 shows the frequency domain interpretation for a five- and three-level decomposition that will be employed in our analysis with quarterly and annual data, respectively. D1, for example, includes a frequency range of between 2 to and 4 years for annual data and between 2 to and 4 quarters for quarterly data. On the other extreme, D5 captures fluctuations in the range between 32 and 64 quarters and corresponds to the D3 component with annual data (8 to 16 year). The business cycle frequency band, defined by the 2 to 8 years range, and denoted with A1 is given by D1 + D2 with annual data and by D3+D4 with quarterly data. Otherwise, the medium run, defined by the range of 8 through 16 years and denoted with A2, is equal to D3 with annual data and D5 with quarterly data. Finally, the long run, defined by the range beyond 16 years and denoted by A3, corresponds to S3 with annual data and S5 with quarterly data.

Next, we perform a sequence of least square regressions at different frequency bands:

$$\Delta w_t^{A_j} = \beta_{0,j} + \beta_{1,j} U_t^{A_j} + \beta_{2,j} \Delta LP_t^{A_j} + \beta_{3,j} \pi_t^{A_j} + \beta_{4,j} CTR_t^{A_j} + \tau_t, \tag{2}$$

where variables have the same meaning as in Equation (1). $\beta_{i,j}$ are coefficients whose first subscript refers to the corresponding right-hand side variable and the second to the timescale details $A_j (j = 1, ..., 3)$ defined above; for example $\beta_{1,2}$ is the coefficient of the unemployment rate in the timescale medium-run component defined by the range of 8 through 16 years. τ_t is an error term. Table 4 reports the estimates of Equation (2) over the entire sample, the pre-war WWI era (1861-1913), the interwar years

it means that global temperature is leading (lagging). If the left arrow points up (down), it means that global temperature is lagging (leading).

[9]The MODWT holds several interesting properties. It can be applied to any sample size, not only to datasets of length multiple of 2J as with the DWT, and returns at each scale a number of wavelet and scaling coefficients equal to the length of the original series. Although the MODWT is highly redundant, so that transformations at each scale are not orthogonal, the offsetting gain is that applying the transform leaves the phase invariant, a very useful property in analyzing transformations (Percival and Walden, 2000).

Table 3: Frequency domain interpretation in years of detail and approximation levels for a 5- (quarterly data) and 3-level (annual data) decompositions

Detail level Dj	Quarterly data	Detail level Dj	Annual data
D1	2-4 Q		
D2	4-8 Q		
D3	8-16 Q	D1	2-4 Y
D4	16-32 Q	D2	4-8 Y
D5	32-64 Q	D3	8-16 Y
S5	> 64 Q	S3	> 16 Y
A1	D3+D4	D1+D2	
A2	D5	D3	
A3	S5	S3	

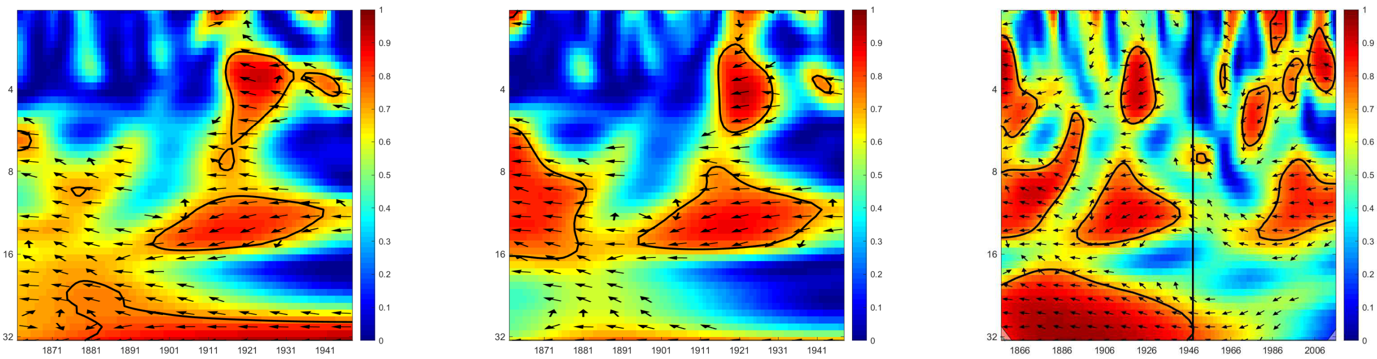
Note: A1= business cycle frequencies, A2 = medium run, A3 = long-run

Figure 3: Coherence money-wage inflation and unemployment rate

(a) Original transformation 1861-1947

(b) Log deviation, original data 1861-1947

(c) Log deviation, Bank of England data 1856-2016



Time is recorded on the horizontal axis and the scale of the wavelet transform is recorded on the vertical axis with frequency converted to time units (quarters) to ease interpretation. The color code for power ranges from blue (low coherence) to red (high coherence). A pointwise significance test is performed against an almost process independent background spectrum. 90% confidence intervals for the null hypothesis that coherency is zero are plotted as contours in black in the figure. The cone of influence, represented by a shaded area corresponding to the region affected by edge effects, is marked by black lines. The end of Phillips' sample is marked by a black line in panel c)

(1923-1948), and the post-WWII period (1960-2015). Two strong results emerge. The first is that the wage PC is predominantly a medium-run phenomenon. In fact, the unemployment coefficient is negative, statistically significant and stable in all four time periods: ranging from -0.707 in the post-WWII era to -0.883 in the pre-WWI period. In contrast, at the business-cycle frequency, the unemployment coefficient is significant in the sub-sample estimated by Phillips but not in the other two. Even after WWI, a period that includes the sharp deflationary policy of the Bank of England to restore sterling to pre-war parity gold convertibility (Moggridge, 1969), the PC effect emerges primarily at the medium-run frequency.

The second is that the wage PC is vertical in the long run, an outcome occurring in the longest available sample (1861-2015), when we would expect a real variable such as unemployment to be invariant to all the varieties of monetary policies that have been carried for more than 150 years (Lucas, 1973; Hall and Sargent, 2018, p.126). To corroborate this point, the point estimate of the price inflation coefficient is 1.002 with a standard error of 0.06; cf. third panel of Table 4. The case of a vertical wage PC is clearest when the impact of price inflation on wage inflation is unity. For the three sub-samples in the long-run time scale, price inflation is positive and statistically significant, but its estimates are either much less accurate than the one indicated above or differ from unity, or a unit price inflation coefficient occurs concomitantly with a non-zero unemployment coefficient.

Table 4: Wage PC time-scale regressions, annual data, 1861-2015

	Business cycle frequencies (D1+D2)				Medium run frequencies (D3)				Long run frequencies (S3)			
	1861-2015	1861-1913	1923-1948	1960-2015	1861-2015	1861-1913	1923-1947	1960-2015	1861-2015	1861-1913	1923-1948	1960-2015
U_t	**-0.719**	**-0.555**	-0.603	-0.567	**-0.883**	**-0.756**	**-0.798**	**-0.707**	-0.043	0.748	**-0.136**	**0.084**
	(0.205)	(0.204)	(0.358)	(0.416)	(0.210)	(0.180)	(0.103)	(0.176)	(0.053)	(0.555)	(0.057)	(0.037)
LP_t	**0.237**	-0.108	**0.463**	0.172	0.269	-0.389	**0.353**	-0.276	**0.454**	**0.961**	**-2.517**	**0.827**
	(0.079)	(0.081)	(0.118)	(0.092)	(0.190)	(0.247)	(0.122)	(0.174)	(0.162)	(0.461)	(0.863)	(0.112)
π_t^{gdpdef}	**0.664**	0.219	**0.704**	**0.883**	**0.633**	**0.651**	**0.682**	0.251	**1.002**	**1.382**	**1.194**	**0.840**
	(0.094)	(0.106)	(0.065)	(0.172)	(0.079)	(0.106)	(0.097)	(0.173)	(0.060)	(0.197)	(0.032)	(0.041)
WWI	**0.020**				**0.047**				0.003			
	(0.005)				(0.011)				(0.004)			
GD	0.0014		-0.000		0.005		0.003		0.004		0.0005	
	(0.003)		(0.003)		(0.003)		(0.003)		(0.004)		(0.003)	
$WWII$	0.002		0.002		-0.000		-0.002		-0.004		**-0.008**	
	(0.002)		(0.004)		(0.001)		(0.006)		(0.003)		(0.002)	
OIL	0.003		-0.001	0.001	-0.001			0.000	-0.008			0.004
	(0.004)			(0.005)	(0.010)			(0.004)	(0.005)			(0.004)
GFC	0.000			-0.001	**0.005**			0.000	-0.004			-0.003
	(0.002)			(0.005)	(0.002)			(0.002)	(0.002)			(0.001)
R^2	0.591	0.246	0.640	0.617	0.875	0.793	0.940	0.702	0.966	0.899	0.991	0.993
$\mathcal{L}$	443.6	170.4	75.80	169.5	477.7	193.7	98.93	219.8	542.5	201.1	112.4	243.7

Coefficients in bold are significant at a 5 % level. Standard errors are robust to autocorrelation and heteroskedasticity using an HAC estimator for the variance and covariance matrix.

Figure 4: Wage inflation and unemployment, annual data, aggregate frequencies

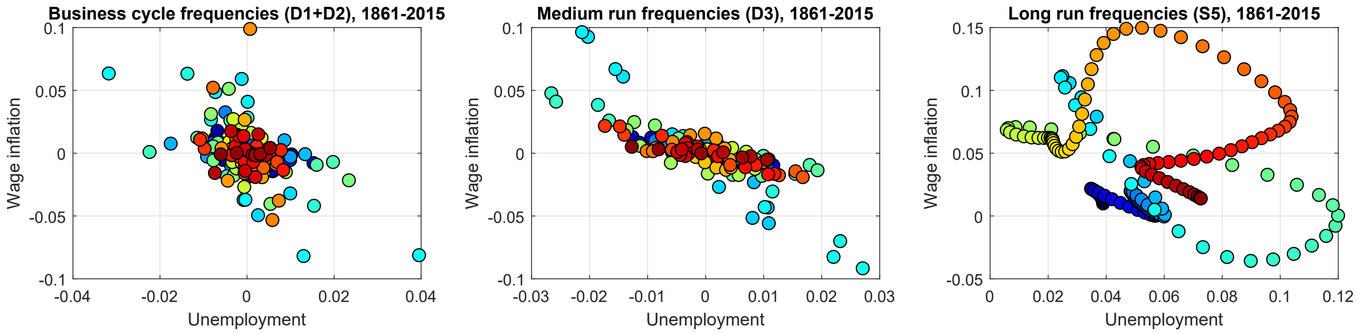
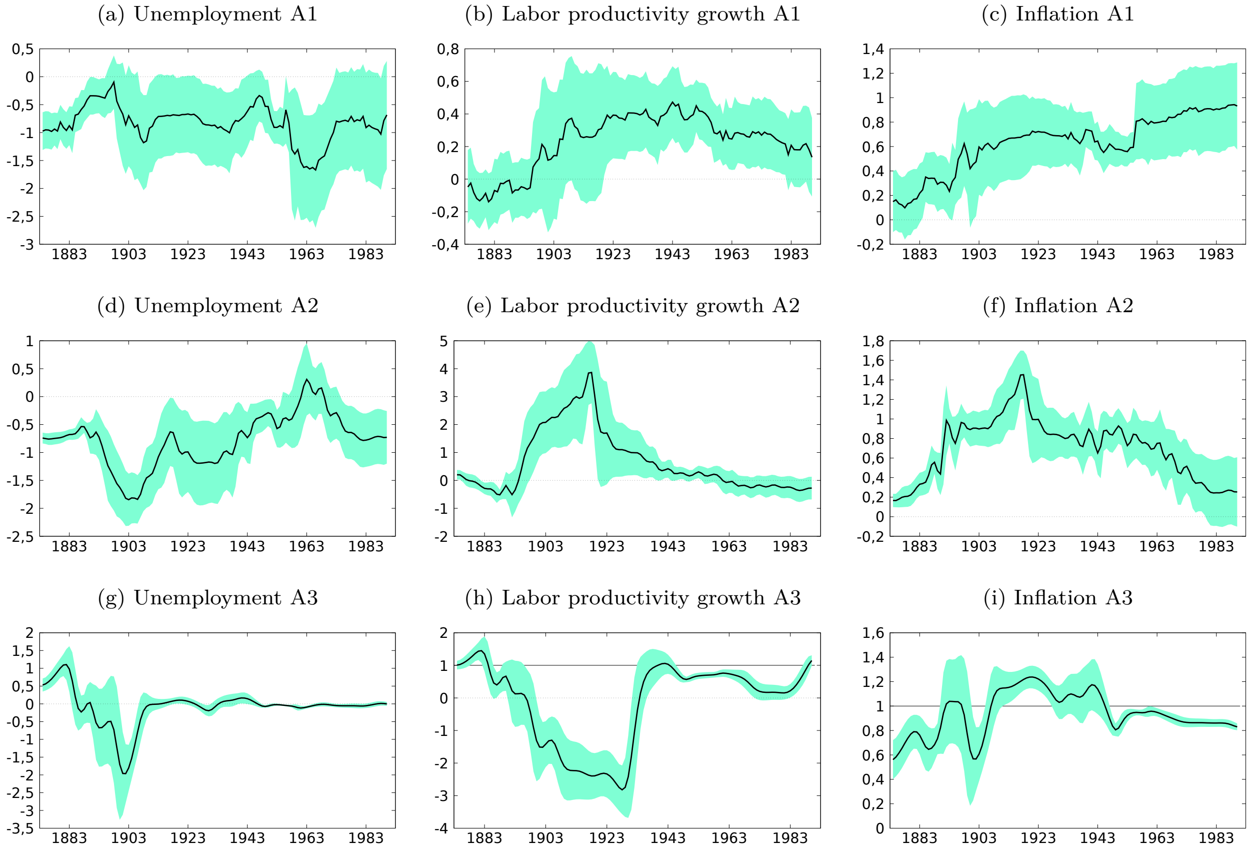


Figure 4 shows the PC relationship decomposed at the business cycle (left column), medium run (middle column) and long run (left column) frequencies. That relationship is weak at the BC frequency, but strong in the medium run, suggesting that what shows up in the aggregate data is mainly due to the medium run frequency.

In Figure 5 on rolling-window coefficients, the estimates indicate that, at a frequency band beyond 16 years (i.e, our long run), the vertical wage PC matches well with the unit tendency of the price inflation coefficient. In the medium run, instead, for almost half of the period a rising U coefficient is associated with a price inflation coefficient trending from unity towards zero.

Labor productivity growth impacts wage inflation in an unstable fashion: two out four coefficients are positive in the business-cycle range, one is significant in the medium-run range, and three are significantly positive and one is significantly negative in the long-run range. Figure 5 highlights the wandering time profile of the labor productivity growth coefficient. Particularly puzzling are the opposite time profiles of labor productivity coefficients in the medium run and the long run; see the last two rows of Figure 5. We attribute these odd outcomes in part to measurement errors due, not only to poor data coverage, but also to objective difficulties in measuring overall labor productivity as the economy shifts, over the long stretch of 154 years, massive resources and inputs from agriculture to industry, from industry to services and finally from services to the so-called fourth industrial revolution of cyber-physical system interacting with the Internet of things and systems.

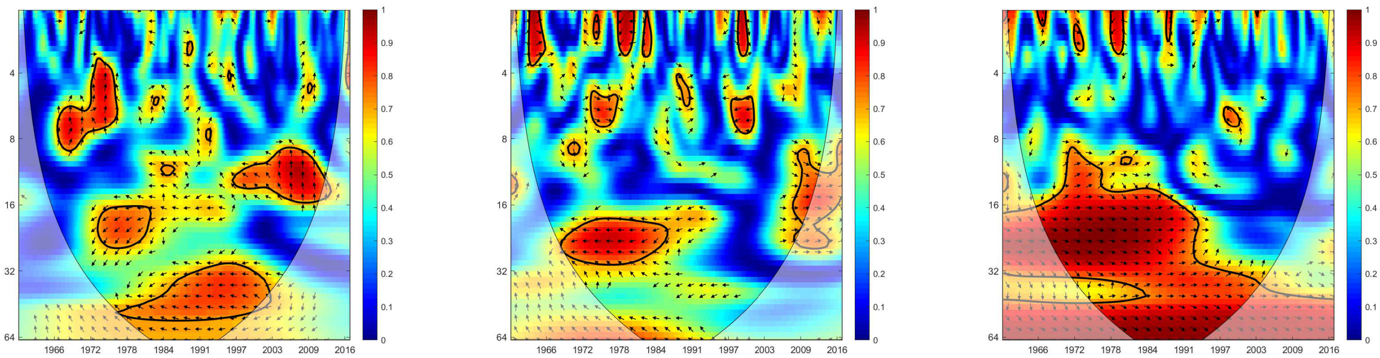
Figure 5: Rolling window coefficients, UK, Bank of England, annual data, frequency decomposition, 1856-2016



We use a rolling window using the % 25 of the observations. Years on the x-axis are the middle value of the interval. Standard errors are robust to autocorrelation and heteroskedasticity using an HAC estimator for the variance and covariance matrix.

Figure 6: Wavelet coherence, quarterly data, 1960-2016

(a) Wage growth and unemployment (b) Wage and labor productivity growth (c) Wage growth and inflation



Time is recorded on the horizontal axis and the scale of the wavelet transform is recorded on the vertical axis with frequency converted to time units (quarters) to ease interpretation. The color code for power ranges from blue (low coherence) to red (high coherence). A pointwise significance test is performed against an almost process independent background spectrum. 90% confidence intervals for the null hypothesis that coherency is zero are plotted as contours in black in the figure.The cone of influence, represented by a shaded area corresponding to the region affected by edge effects, is marked by black lines.

4 The Post WWII wage PC

The post-WWII period gives us an opportunity to verify whether our findings drawn from a long stretch of annual data hold as well for a shorter sample of quarterly data; that is, it works as a sort of robustness exercise. Figure 6 displays the output of the continuous wavelet transform over the post-war period 1960 to 2016. The negative coherence between money wage inflation and unemployment is concentrated at the medium-run frequency (32 through 64 quarters), with few scattered red areas in the business-cycle range –1965-1972, 1972-1980 and 1997-2009–; see panel (a). Wage inflation is highly correlated with price inflation at the business-cycle timescale up to the mid-1990s and throughout the sample at the medium-run timescale. The growth of labor productivity appears to be less important for money inflation than either unemployment than price inflation.

Next, we repeat the same exercise we did for annual data:[10]

$$\Delta w_t^{D_j} = \delta_{0,j} + \delta_{1,j} U_t^{D_j} + \delta_{2,j} \Delta LP_t^{D_j} + \delta_{3,j} \pi_t^{D_j} + \delta_{4,j} CTR_t^{D_j} + \kappa_t, \tag{3}$$

where variables have the same meaning as in in Equation (2). $\delta_{i,j}$ are coefficients, where the first subscript refers to the corresponding right-hand side variable and the second subscript to the timescale details $D_j (j = 2, 5)$ defined in Table 3 (e.g. $\delta_{1,2}$ is the coefficient of the unemployment rate in the timescale component 4 to 8 quarters or 1 to 2 years); and κ_t is an error term. Table 5 reports the estimates of equation (3) over the entire sample 1960:Q1-2016:Q4 and of three subsamples, each ten years shorter than the previous one. Separate regressions are performed for each timescale D_j (j = 2,...,5). In addition,

[10]As robustenss check, we estimate the same model using quarterly data and the Christiano and Fitzgerald (2003) bandpass filter. See Appendix C.

Table 5: Dependent Variable: Δw_t^{Dj}, 1960:1- 2016:4, UK

	1990:1- 2016:4	1980:1- 2016:4	1970:1- 2016:4	1960:1- 2016:4
U_t^{D2}	0.460	-0.408	-0.825	-2.183
	(1.752)	(1.837)	(1.903)	(1.950)
ΔLP_t^{D2}	0.058	0.149	**0.433**	**0.380**
	(0.162)	(0.125)	(0.166)	(0.148)
π_t^{D2}	-0.135	-0.145	-0.140	-0.142
	(0.115)	(0.098)	(0.119)	(0.103)
U_t^{D3}	-0.539	**-1.212**	0.160	0.574
	(0.369)	(0.483)	(0.609)	(0.580)
ΔLP_t^{D3}	**0.275**	**0.206**	**0.238**	**0.361**
	(0.113)	(0.074)	(0.074)	(0.068)
π_t^{D3}	**-0.309**	0.256	**0.823**	**0.851**
	(0.066)	(0.135)	(0.189)	(0.176)
U_t^{D4}	**-1.171**	-0.150	**-0.714**	**-0.702**
	(0.162)	(0.191)	(0.312)	(0.277)
ΔLP_t^{D4}	**0.775**	**0.569**	**0.631**	**0.671**
	(0.058)	(0.082)	(0.109)	(0.071)
π_t^{D4}	**0.490**	**1.203**	**1.257**	**1.264**
	(0.060)	(0.107)	(0.070)	(0.058)
U_t^{D5}	**-0.644**	**-0.679**	**-0.588**	**-0.415**
	(0.068)	(0.070)	(0.109)	(0.133)
ΔLP_t^{D5}	**-0.422**	**-0.418**	-0.341	0.018
	(0.112)	(0.105)	(0.177)	(0.234)
π_t^{D5}	**0.395**	**0.283**	**0.266**	**0.531**
	(0.089)	(0.068)	(0.115)	(0.157)
U_t^{S5}	**-0.295**	-0.046	-0.026	**0.028**
	(0.009)	(0.027)	(0.016)	(0.011)
ΔLP_t^{S5}	**1.071**	**1.164**	**1.110**	**0.925**
	(0.012)	(0.049)	(0.050)	(0.045)
π_t^{S5}	**1.347**	**0.947**	**0.947**	**0.971**
	(0.020)	(0.014)	(0.006)	(0.008)

Coefficients in bold are significant at a 5 % level. Standard errors are robust to autocorrelation and heteroskedasticity using an HAC estimator for the variance and covariance matrix.

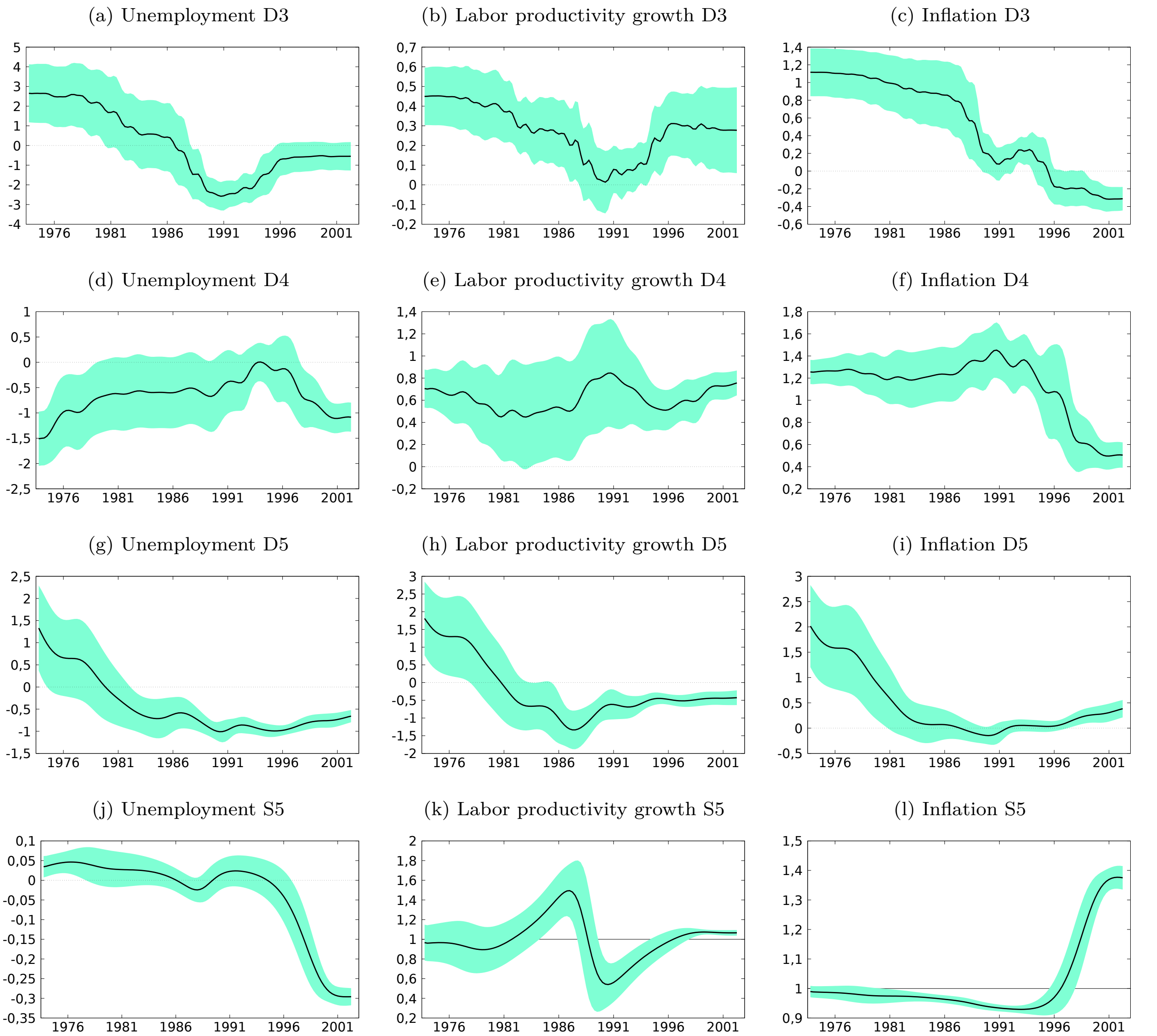
we report the estimates for the long-run timescale S_5 that goes beyond 64 quarters. The key findings are similar to those we found in the previous section. A negative and stable relationship between nominal wage growth and unemployment occurs in the medium–run frequency D_5 (32 to 64 quarters). While the PC effect is also present in the business-cycle frequency D_4 (16 to 32 quarters), the pattern is time dependent. In fact, the rolling-window coefficients of Figure 7 show that the estimates of $\delta_{1,5}$ are on the whole negative, whereas the majority of the $\delta_{1,4}$ estimates is not statistically different from zero.[11]

The other important point is that the Phillips Curve is vertical in the long run. It is clearly the case with the estimates over the 1970Q1-2016Q4 and 1980Q1-2016Q4 periods (last row of Table 5).

Then, in the last decade of the 20th century the estimates of δ_{S5} turn negative (last row of Figure 7), yielding a conventional PC effect over the shortest sub-sample, 1990Q1-2016Q4. When Equation (3)

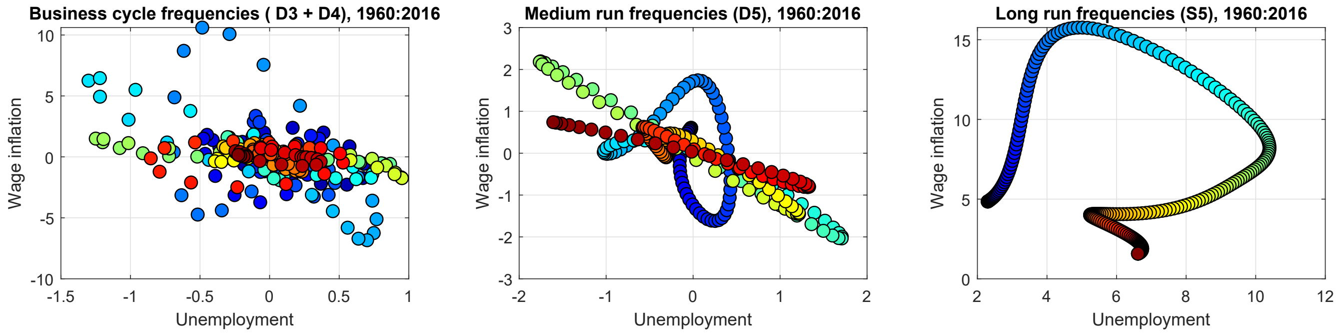
[11] PC effects at details D_2 and D_3 are even more erratic, swinging from positive to negative areas.

Figure 7: Rolling window coefficients, UK, quarterly data, 1960:1-2016:4



We use a rolling window using the % 50 of the observations. Years on the x-axis are the middle value of the interval. Standard errors are robust to autocorrelation and heteroskedasticity using an HAC estimator for the variance and covariance matrix.

Figure 8: Wage inflation and unemployment, quarterly data, aggregate frequencies



is estimated over the entire quarterly sample 1960Q1-2016Q4, $\delta_{1,S5}$ shows up with the wrong sign and statistically significant, but its impact is trivial in an economic sense. This is the reason for calling the PC vertical. Just as importantly, the growth of labor productivity and price inflation impact money wage inflation with statistically unit coefficients, as one would expect to hold in the long run.

Figure 8 shows the post WWII scatterplot between wage growth and unemployment in the business, medium and the long run. While there is no presence of the PC at the business cycle frequencies, it emerges at the medium run.

5 Discussion

With the Samuelson and Solow (1960) article, the focus of the Phillips Curve shifted from money wage inflation to price inflation, with the consequent corollary that policy could exploit a tradeoff between inflation and unemployment. The tradeoff issue was later dismissed by Phelps (1967), Friedman (1968), Sargent (1971) and Lucas (1972, 1973) and put to rest by the inflation of the 1970s. By the early Nineties, the state of the arts was that observable US data were inconsistent with a PC. Yet, "… a strikingly stable negative correlation exists over the business cycle, and recent theory indicates the Lucas-Sargent critique may not be empirically relevant. When we estimate the long-run trade-off as Gordon and Solow did, we find it is roughly one-for-one." (King and Watson, 1994, p.168-170).[12] The opposite holds for the UK. Haldane and Quah (1999, Fig 3) produce a conventionally sloped price inflation-unemployment scatter plot from 50 years of post-WWII monthly data.[13] On the other hand, Broadbent (2014, Chart 2) shows a scatter plot of money wage inflation-unemployment for the years 1968-1992 that looks like a meandering curve drifting eastwardly.[14] When we extend Broadbendt's scatter plot to the entire sample available, 1960Q1-2016Q4, the meandering pattern generally holds across the whole sample; see Figure 8. The obvious inference is that wage and price PC are not the mirror image of each other. Our findings show that a wage PC relationship, over the long historical stretch, is observable in UK data only for certain periods, whereas it is negative, stable and strong in the medium-run frequency band.

The wage PC reemerged with Galí (2011) who embedded a staggered wage contract setup in a New Keynesian (NK) model.[15] The implied NK wage PC is that money wage inflation responds positively to the expected one-period ahead wage inflation and negatively to the deviation of the average markup from its desired value.[16] However, in the empirical work based on post-war US quarterly data, money wage inflation is explained by one-period lagged price inflation and an autoregressive unemployment rate. Forward-looking NK price PC models have difficulty in explaining inflation dynamics in the data without an assumption of inflation persistence (Mavroeidis et al., 2014, p 130).[17] In the NK wage PC the lagged price inflation term is justified as a wage indexation factor, but in fact captures persistence,

[12] The methodology employed for detecting effects at various frequencies is to filter the data with moving averages of different lengths and by placing constraints on specific points (Baxter and King, 1999; Christiano and Fitzgerald, 2003). King and Watson (1994) define the business-cycle component of the series by applying a band-pass filter isolating periodic components between one and half and eight years; for the long run, they apply a low-pass filter isolating periodicities beyond eight years; see their Figure 2 on page 168 and discussion on pages 169-170.

[13] The authors also report that the relationship is not stable over the 50 years: the Curve is virtually vertical from 19848 to 1980 and virtually flat afterwards. Furthermore, the Curve becomes vertical at business-cycle frequencies (Haldane and Quah, 1999, Fig 5), further highlighting the difference between wage and price PCs.

[14] Money wage inflation is adjusted for a measure of expected price inflation, against the unemployment rate.

[15] Staggered wage contracts–such as those in Fischer (1977), Taylor (1980) or Calvo (1983)—gave impetus to the development of the New Keynesian Phillips Curve, where current price inflation responds positively to the expected one-period price inflation and a measure of real economic activity.

[16] The markup refers to the markup of the money wage with respect to the marginal rate of substitution between consumption and working. If the markup falls below its desired level, those workers who reset their money wages in a given year will experience a negative wage inflation. The NK wage PC is vertical when all workers renew their contracts each period. The Curve has the conventional negative slope: the smaller the proportion of workers renewing their contracts each period, the flatter will be the Curve.

[17] More generally, a lagged rate of inflation may be interpreted as a proxy of the expected rate of inflation or price stickiness (inertia) stemming staggered wage contracts.

and so does the autoregressive structure of the unemployment rate.

The extent of persistence in the US wage PC emerges in Gallegati et al. (2011) who test Equation (3) with the same methodology of our paper over quarterly data from 1948Q1 to 2009Q2:

> "...the highest and most significant effect of [the] unemployment rate on nominal wage growth occurs at intermediate scales...between 2 and 8 years, thus suggesting that the time of 'pass through' of the unemployment rate to wage inflation is in the business cycle frequency range...at the longest scale, corresponding to periods longer than 8 years, the estimated coefficient indicates a flattening of the wage Phillips curve, being otherwise more significant than any other frequency." (Gallegati et al., 2011, p 497).

Our findings for the UK are in line with those of the US, except that the time of pass through of the unemployment rate to money wage inflation occurs in the medium-run range (8 to 16 years) rather than in the business-cycle range.

Much of the discussion about the Phillips Curve centers on its temporal instability. The high inflation rates of the 1970s and the disinflation transition in the first half of the 1980s are typically blamed for the breakdown of the PC relationship. In fact, in both sub-periods the observed scatter plots describe a positively sloped curve (Galí, 2011, Fig 6). Blanchard et al. (2015) estimates price PCs for 20 countries and finds that, for given expected inflation, the curve flattens until the early 1990s and then stabilizes. However, for the UK the slope is not statistically significant for either the years 1990-2014 or the years 2007-2014 (see their Table 6). Cunliffe (2017, Chart 6) fits trend lines through scatter plots of the UK money wage inflation-unemployment rate for the periods 1971-1997, 1998-2012 and 2013-2017 and concludes that the PC relationship has progressively shifted inward and flattened. Our estimates of $\delta_{1,2}$ (in the frequency range of 4 to 8 quarters) of Table 5, while not statistically significant, are suggestive of a flattening wage PC as one moves from longest sample of 1960Q1-2016Q4 to the shortest sample of 1990Q1-2016Q4. The rolling window estimates of $\delta_{1,2}$ in Figure 7 describe a sinusoid curve around a zero line. The point is that both at the aggregate level and at high frequencies the PC relationship is unstable over time. But, in the frequency range between 32 to 64 quarters (our medium run timescale), this time dependency disappears: the estimates of $\delta_{1,5}$ in the four different time periods of Table 5 range from -0.415 (with standard error of 0.133) to -0.679 (0.068).

The last and thorniest issue is why the PC relationship is stable primarily in the medium run. We do not have a satisfactory or complete answer. We limit ourselves to some observations on the matter. To start, labor is more a "quasi-fixed factor" than the variable input described in conventional microeconomics texts (Oi, 1962). Fixed hiring, firing and training costs–reflecting in part the power of the unions– are responsible for this semi-fixity. Wage contracts and monopolistic price settings impart stickiness. When staggered wage contracts were introduced in the macroeconomics literature, the objective was to mimic persistence in the data and create at the same time sufficient flexibility so that a credible monetary disinflation could be carried out with little or no economic contraction (Ball, 1994); a sort of having a cake and eat it too. But Fuhrer and Moore (1995) discovered that contracts a la Taylor (1980), where the price level is an average of current and past contract wages, generate persistence in the level of prices but not in the rate of inflation.[18] Consequently, we are left with imperfectly credible monetary policy or its interaction with wage contracts as an explanation for inflation persistence in the New Keynesian PC models, where money wage inflation responds to the expected rate of inflation (Ball, 1995). The specification of our paper is not consistent with those models and relies instead on the current rate of inflation, which is very strongly correlated with past rates of inflation. It may be an "ugly duckling" model but it does justice to the data.[19]

6 Conclusion

Our paper goes back to this insight and uses an expansive UK sample of annual data, from 1861 to 2015, and quarterly post-WWII quarterly data, from 1960Q1 to 2016Q4, to explore the negative relationship between money wage inflation and the unemployment rate occurring at different frequency and over time. The two critical findings from the annual sample are that the wage PC is predominantly a medium-run phenomenon, comprised in the 8-to16-year frequency band, and that the curve disappears in the long run. Similar conclusions are reached for the post-WWII robustness sample of quarterly data. A negative and stable relationship between money wage inflation and unemployment occurs, again, in the medium–run frequency range of 32 to 64 quarters. While the PC effect is also present in the business-cycle frequency (16 to 32 quarters), it is unstable over different time intervals.

A wage PC is not a mirror image of a price PC. Scatter plots of price inflation-unemployment in the post-WWII era produce a conventionally sloped curve, but not for money wage inflation-unemployment. Possible causes of this difference may reside in variation in the labor share of total income, changes in productivity growth and a breakdown in the relationship between wage inflation and price inflation. The essential point is that over a long stretch of years a wage PC relationship is observable in UK data intermittently, whereas it is negative, stable and strong in the medium-run frequency band.

The paper does not provide a theoretically satisfactory answer of why the PC relationship is stable in the medium-run frequency range, although it is in line with the literature on the medium-run macroeconomics where wages (and prices) and real variables adjust to each other.[20] We refer to ongoing work in the literature on the source of wage and price stickiness, staggered wage contracts. The earlier promise that such contracts could explain nominal stickiness and retain sufficient flexibility to allow

[18] Driscoll and Holden (2004) start from the behavioral assumption that workers care more about being paid less than other workers than being paid more (a sort of asymmetric fairness) to generate a model of inflation persistence that occurs within certain bounds of the unemployment rate. Also near-rational expectations models have tried to arrive at inflation persistence.

[19] On the point that rational expectations cannot explain persistence in NKPC models, see Mankiw (2001, C59) and Mavroeidis et al. (2014, p 172)

[20] Representative works of this literature include Blanchard (1997), Solow (2000) and Beaudry (2005).

a credible monetary disinflation to be carried out with little or no economic loss of unemployment or output has not panned out. There is the alternative that monetary policy, the big driver in the growth of nominal magnitudes, suffers from very imperfect credibility. Clearly, additional research is required on the subject.

References

Aguiar-Conraria, L., Martins, M. M., Soares, M. J., May 2019. The Phillips Curve at 60: time for time and frequency . Nipe working paper, Universidade do Minho. Núcleo de Investigação em Políticas Económicas (NIPE).

Aguiar-Conraria, L., Soares, M. J., April 2014. The Continuous Wavelet Transform: Moving Beyond Uni- And Bivariate Analysis. Journal of Economic Surveys 28 (2), 344–375.

Ball, L., 1994. Credible Disinflation with Staggered Price-Setting. American Economic Review 84 (1), 282–289.

Ball, L., 1995. Disinflation with imperfect credibility. Journal of Monetary Economics 35 (1), 5–23.

Baxter, M., King, R. G., November 1999. Measuring Business Cycles: Approximate Band-Pass Filters For Economic Time Series. The Review of Economics and Statistics 81 (4), 575–593.

Beaudry, P., 2005. Innis lecture: Explorations in medium-run macroeconomics. The Canadian Journal of Economics / Revue canadienne d'Economique 38 (4), 1136–1159.

Benati, L., 2015. The long-run Phillips curve: A structural VAR investigation. Journal of Monetary Economics 76 (C), 15–28.

BIS, 2017. 87th Annual report.

Blanchard, O. J., 1997. The Medium Run. Brookings Papers on Economic Activity 28 (2), 89–158.

Blanchard, O. J., Cerutti, E. M., Summers, L., Nov. 2015. Inflation and Activity – Two Explorations and their Monetary Policy Implications. IMF Working Papers 15/230, International Monetary Fund.

Broadbent, B., 2014. Unemployment and the conduct of monetary policy in the UK. Speech, Federal Reserve Bank of Kansa City 38^{th} Economic Symposium, Jackson Hole, Wyoming.

Calvo, G. A., 1983. Staggered prices in a utility-maximizing framework. Journal of Monetary Economics 12 (3), 383 – 398.

Carney, M., 2017. [De]Globalization and inflation. Speech, Speech at the IMF Michel Camdessus Central Banking Lecture.

Chatfield, C., 1989. The Analysis of Time Series: An Introduction. Chapman & Hall.

Christiano, L. J., Fitzgerald, T. J., 05 2003. The Band Pass Filter. International Economic Review 44 (2), 435–465.

Cunliffe, J., 2017. The Phillips curve: lower, flatter or in hiding? Speech, Speech at the Oxford Economics Society.

Daubechies, I., et al., 1992. Ten lectures on wavelets. Vol. 61. SIAM.

Desai, M. J., February 1975. The Phillips Curve: A Revisionist Interpretation. Economica 42 (165), 1–19.

Driscoll, J. C., Holden, S., 04/05 2004. Fairness and Inflation Persistence. Journal of the European Economic Association 2 (2-3), 240–251.

Economic Report of the President, 1962. Economic Report of the President, 1962. Together with the Annual Report of the Council of Economic Advisrers.

Fischer, S., 1977. Long-term contracts, rational expectations, and the optimal money supply rule. Journal of Political Economy 85 (1), 191–205.

FOMC, 2018. Minutes of the meeting of January 30-31, 2018.

Friedman, M., 1968. The role of monetary policy. American Economic Reviewl (58), 1–17.

Fuhrer, J., Moore, G., 1995. Inflation persistence. The Quarterly Journal of Economics 110 (1), 127–159.

Galí, J., 2011. The Return Of The Wage Phillips Curve. Journal of the European Economic Association 9 (3), 436–461.

Gali, J., Gertler, M., 1999. Inflation dynamics: A structural econometric analysis. Journal of Monetary Economics 44 (2), 195–222.

Gallegati, M., Gallegati, M., Ramsey, J. B., Semmler, W., August 2011. The US Wage Phillips Curve across Frequencies and over Time. Oxford Bulletin of Economics and Statistics 73 (4), 489–508.

Gallegati, M., Ramsey, J. B., May 2019. Phillips' averaging procedure as a 'crude' version of the Haar wavelet filter. Working Papers 437, Universita' Politecnica delle Marche (I), Dipartimento di Scienze Economiche e Sociali.

Gallegati, M., Ramsey, J. B., Semmler, W., 2014. Interest rate spreads and output: A time scale decomposition analysis using wavelets. Computational Statistics & Data Analysis 76 (C), 283–290.

Gençay, R., Selçuk, F., Whitcher, B. J., 2001. An introduction to wavelets and other filtering methods in finance and economics. Academic press.

Gordon, R. J., 1982. Price inertia and policy ineffectiveness in the united states, 1890-1980. Journal of Political Economy 90 (6), 1087–1117.

Gordon, R. J., 2011. The history of the phillips curve: Consensus and bifurcation. Economica 78 (309), 10–50.

Grinsted, A., Moore, J. C., Jevrejeva, S., 2004. Application of the cross wavelet transform and wavelet coherence to geophysical time series. Nonlinear Processes in Geophysics 11 (5/6), 561–566.

Haldane, A., Quah, D., 1999. UK Phillips curves and monetary policy. Journal of Monetary Economics 44 (2), 259–278.

Hall, R. E., Sargent, T. J., February 2018. Short-run and long-run effects of milton friedman's presidential address. Journal of Economic Perspectives 32 (1), 121–34.

Heller, W. W., 1964. The Future of Our Fiscal System. The Journal of Business 38, 235–235.

Hudgins, L., Friehe, C., Mayer, M., 1993. Wavelet transforms and atmospheric turbulence. Physics Review Letter 71, 3279–3282.

King, R., Watson, M., 1994. The post-war u.s. phillips curve: a revisionist econometric history. Carnegie-Rochester Conference Series on Public Policy 41 (1), 157–219.

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 27 (105), 1–31.

Lucas, R., 1972. Expectations and the neutrality of money. Journal of Economic Theory 4 (2), 103–124.

Lucas, R., 1973. Some international evidence on output-inflation tradeoffs. American Economic Review 63 (3), 326–34.

Mankiw, N. G., 2001. The inexorable and mysterious tradeoff between inflation and unemployment. The Economic Journal 111 (471), C45–C61.

Mavroeidis, S., Plagborg-Møller, M., Stock, J. H., 2014. Empirical evidence on inflation expectations in the new keynesian phillips curve. Journal of Economic Literature 52 (1), 124–188.

Moggridge, D. E., 1969. The Return to Gold 1925: The Formulation of Economic Policy and Its Critics. Cambridge University Press.

Mutascu, M., 2019. Phillips curve in US: New insights in time and frequency. Research in Economics 73 (1), 85–96.

Oi, W., 1962. Labor as a quasi-fixed factor. Journal of Political Economy 70.

Percival, D. B., Walden, A. T., 2000. Wavelet Methods for Time Series Analysis. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press.

Phelps, E. S., 1967. Phillips curves, expectations of inflation and optimal unemployment over time. Economica 34 (135), 254–281.

Phillips, A. W., 1958. The relation between unemployment and the rate of change of money wage rates in the united kingdom, 1861-1957. Economica 25 (100), 283–299.

Ramsey, J. B., Usikov, D., Zaslavsky, G. M., 1995. An analysis of u.s stock price behavior using wavelets. Fractals 03 (02), 377–389.

Ramsey, J. B., Zhang, Z., 1996. The Application of Wave Form Dictionaries to Stock Market Index Data. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 189–205.

Ryland, T., Sally, H., Nicholas, D., 2010. The UK recession in context — what do three centuries of data tell us? Bank of England Quarterly Bulletin 50 (4), 277–291.

Samuelson, P. A., Solow, R. M., 1960. Analytical aspects of anti-inflation policy. The American Economic Review 50 (2), 177–194.

Sargent, T. J., August 1971. A Note on the 'Accelerationist' Controversy. Journal of Money, Credit and Banking 3 (3), 721–725.

Sargent, T. J., Wallace, N., 1975. "rational" expectations, the optimal monetary instrument, and the optimal money supply rule. Journal of Political Economy 83 (2), 241–254.

Silverman, B. W., 1998. Wavelets in statistics: some recent developments. In: COMPSTAT. Springer, pp. 15–26.

Solow, R. M., 2000. Toward a Macroeconomics of the Medium Run. Journal of Economic Perspectives 14 (1), 151–158.

Staiger, D., Stock, J. H., Watson, M. W., Winter 1997a. The NAIRU, Unemployment and Monetary Policy. Journal of Economic Perspectives 11 (1), 33–49.

Staiger, D. O., Stock, J. H., Watson, M. W., 1997b. How Precise Are Estimates of the Natural Rate of Unemployment? In: Reducing Inflation: Motivation and Strategy. NBER Chapters. National Bureau of Economic Research, Inc, pp. 195–246.

Strang, G., Nguyen, T., 1996. Wavelets and Filter Banks. Wellesley-Cambridge Press.

Taylor, J. B., 1980. Aggregate Dynamics and Staggered Contracts. Journal of Political Economy 88 (1), 1–23.

Torrence, C., Compo, G. P., 1998. A practical guide to wavelet analysis. Bulletin of the American Meteorological Society 79 (1), 61–78.

Torrence, C., Webster, P. J., 1999. Interdecadal changes in the enso–monsoon system. Journal of Climate 12 (8), 2679–2690.

Wulwick, N. J., 1996. Two econometric replications: The historic phillips and lipsey-phillips curves. History of Political Economy 28 (3), 391–439.

Appendix A: Data

- The original database used by Phillips for the period 1860-1947 can be found in Wulwick (1996) Table 1 at page 427. For the left panel of Figure 1, we calculate:
  - Wage inflation is calculated as the log difference of the *Money-Wage index* (column 1).
  - Unemployment rate is the *unemployment* (column 3).
- We employ the *A Millennium of macroeconomic data for the UK. The Bank of England's collection of historical macroeconomic and financial statistics* based on the work by Ryland et al. (2010).
  - Annual data for the period 1861-2015 are calculated as:
    - Wage inflation calculated as the log difference of *Composite Average Weekly Earnings series* taken from worksheet A.47
    - Unemployment rate is the *Unemployment rate* taken from the worksheet A.50
    - Labor productivity is the log difference of the *Labour productivity per hour* taken from worksheet A.56
    - Price inflation is calculated as the log difference of the *GDP deflator at market prices* taken from worksheet A.47
  - Quarterly data for the period 1960:1-2016:4 are calculated as:
    - Wage inflation calculated as the log difference of the *Spliced Average Weekly Earnings series, 1919-2015* taken from worksheet Q.1
    - Unemployment rate is the *Monthly unemployment rate* taken from worksheet M.1 and transformed into quarterly frequency.
    - Labor productivity is calculated as the log difference of the *Output per workers, seasonal adjusted.*
    - Price inflation is calculated as the log difference of the *GDP deflator at market prices* taken from worksheet Q.1

Appendix B: A brief introduction to the wavelet transform

We begin this section by recalling the basic structure of the wavelet approach to modeling time series data. Wavelets are particular types of function $\psi(.)$ that are localized both in time and frequency domain and used to decompose a function $f(x)$, *i.e.* a surface, a series, etc.., in more elementary functions which include informations about the same $f(x)$. The essential characteristics of wavelets are best illustrated through the development of the continuous wavelet transform (CWT). At the simplest level, a "wavelet" is a function $\psi(.)$ such that:

$$\int_{-\infty}^{\infty} \psi(t)dt = 0 \quad (4)$$
$$\int_{-\infty}^{\infty} \psi(t)^2 dt = 1 \quad (5)$$

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. We see from these relationships that wavelets involve weighted differences, integral of $\psi(.)$ is zero; and has a normalized scale. The import of the latter is that for any $\varepsilon > 0$, there exists a finite T such that:

$$\int_{-T}^{T} \psi(t)^2 dt < 1 - \varepsilon \quad (6)$$

In short, the squared contribution to the variance is vanishingly small outside the range (-T, T). In addition, there is a requirement on the choice of wavelet function that is called the "admissibility condition" which ensures that the original function can be reconstructed from the wavelet decomposition (see Percival and Walden, 2000).[21]

For any arbitrary mother wavelet function ψ, the continuous wavelet transform of the function $x(t)$ is given by:

$$W_x(\lambda, \tau) = \int_{-\infty}^{\infty} \psi_{\lambda,\tau}(t)x(t)dt$$

where

$$\psi_{\lambda,\tau}(t) \equiv \frac{1}{\sqrt{\lambda}}\psi\left(\frac{t-\tau}{\lambda}\right).$$

$W_x(\lambda, \tau)$ represents a set of wavelet daughters obtained by scaling and translating the mother wavelet ψ with parameters λ and τ denoting the dilation (scale factor) and translation (time shift), respectively (see Percival and Walden, 2000).

Continuous wavelet transform (CWT) can provide us with the (local) wavelet power spectrum which displays, in time (space) and frequency (scale), the frequency content of signals as a function of time and discover the way the dynamics of the sequence evolve. Let $W_x(\lambda, t)$ be the continuous wavelet transform of a signal $x(.)$, with respect to the wavelet ψ, where λ is a scaling or dilation factor that controls the length

[21] By means of the inverse transform, the original signal can be recovered by integrating over all scales λ and locations τ, respectively.

of the wavelet and t a location parameter that indicates where the wavelet is centered. $|W_x|^2$ represents the wavelet power which depicts the local variance of $x(.)$. The wavelet power spectrum provides an unfolding of the characteristics of a process in the scale-space plane and can be quite revealing about the structure of a particular process, as it can clearly show the presence of multiscale features and can easily identify their temporal locations.

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, Torrence and Compo, 1998, Grinsted et al., 2004). The wavelet coherence is defined as the modulus of the wavelet cross spectrum normalized by the wavelet spectra of each signal,

$$R^2_{xy} = \frac{|S(\lambda^{-1}W_{xy})|^2}{S(\lambda^{-1}|W_x)|^2)S(\lambda^{-1}|W_y)|^2)},$$

where S is a smoothing operator (see Torrence and Webster, 1999). The squared wavelet coherence coefficient R^2_{xy}, ranging between 0 and 1, can be considered a direct measure of the local correlation between two time series in time-frequency domain (Chatfield, 1989) and is especially useful in highlighting the time and frequency intervals where two phenomena have strong interactions. The wavelet coherence 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. The wavelet coherence phase difference θ is the ratio between the imaginary and the real part of the wavelet coherence:

$$\theta_{xy} = arctan(\frac{\Im[S(\lambda^{-1}W_{xy})]}{\Re[S(\lambda^{-1}W_{xy})]}).$$

The phase of a given time-series $x(t)$ can be viewed as the position in the pseudo-cycle of the series and it is parameterized in radian ranging from $-\pi$ to π.

The CWT is a highly information redundant transform representing each datum by a pair of data, designating time, or space, and scale. As a consequence it it is computationally impossible to analyze a signal using all wavelet coefficients. In addition, as noted by Gençay et al. (2001), $W(\lambda;\tau)$ is a function of two parameters and as such it contains a high amount of redundant information. A one to one transformation is obtained by discretizing the transform over scale and over time. 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.[22]

The DWT is based on similar concepts as the CWT, but is more parsimonious in its use of data (Gençay et al., 2001). 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 λ 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 $W(\lambda, \tau)$ in which the entire time-frequency plane is partitioned with rectangular cells of varying dimensions but constant area, called Heisenberg cells.[23] Higher frequencies can be 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 Heisenberg's 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.

The general formulation for a continuous wavelet transform can be restricted to the definition of the "discrete wavelet transform", by discretizing the time-scale parameters λ and τ. In order to obtain an orthonormal basis a transform of the scaling parameter, $\lambda = \lambda_0^j$, and the Nyquist sampling rule, $\tau = k\lambda_0^j T$, are used. The generating wavelet functions are then

$$\psi_{j,k}(t) = \lambda_0^{-j/2} \psi \left(\frac{t - k\lambda_0^j T}{2^j} \right) \tag{7}$$

where j and k are integers. If T is small enough and the computation is done octave by octave, i.e.

[22] The number of 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,).

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

$\lambda_0 = 2$, the “mother wavelet”results in the following equation

$$\psi_{j,k}(t) = 2^{-j/2}\psi\left(\frac{t-2^j k}{2^j}\right) \tag{8}$$

This function represents a sequence of rescalable 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, $f(t)$, $k = 1, 2, 3, ..., N, N = 2^J$ on the wavelet $\psi_{j,k}(t)$ is given by:

$$\begin{aligned} d_{j,k} &\approx \int \psi_{j,k}(t) f(t) dt, \\ j &= 1, 2, ..J \end{aligned} \tag{9}$$

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. The scaling function, also know as the ”father wavelet”, is defined by:

$$\phi_{J,k}(t) = 2^{-J/2}\phi\left(\frac{t-2^J k}{2^J}\right) \tag{10}$$

and the scaling function coefficient is given by:

$$s_{J,k} \approx \int \phi_{J,k}(t) f(t) dt, \tag{11}$$

The discrete wavelet transform DWT re-expresses a time series in terms of coefficients that are associated with a particular time and a particular dyadic scale. For the DWT, where the number of observations is $N = 2^J$, the number of coefficients at each dyadic scale is:

$$N = N/2^J + N/2^J + N/2^{J-1} + ... + N/4 + N/2 \tag{12}$$

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 $d_{1,k}$.

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 et al. (1992), Silverman (1998) and Strang and Nguyen (1996).[24]

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:

[24] The orthonormal properties can be demonstrated by the equations:

$$
\begin{aligned}
x(t) \approx & \sum_k s_{J,k}\phi_{J,k}(t) + \sum_k d_{J,k}\psi_{J,k}(t) + ... + \\
& \sum_k d_{j,k}\psi_{j,k}(t) + ... + \sum_k d_{1,k}\psi_{1,k}(t)
\end{aligned} \tag{14}
$$

The degree of relative error in many cases of economic variables is approximately on the order of 10^{-13}, so that one can reasonably claim that the wavelet decomposition in (10) is very good. The above equation is an example of the Discrete Wavelet Transform (DWT) based on an arbitrary wavelet function, $\phi(.)$.

Further, the approximation can be re-written in terms of collections of coefficients at given scales as:

$$
\begin{aligned}
S_J &= \sum_k s_{J,k}\phi_{J,k}(t) \\
D_J &= \sum_k d_{J,k}\psi_{J,k}(t) \\
D_{J-1} &= \sum_k d_{J-1,k}\psi_{J-1,k}(t) \\
& ...... \\
D_1 &= \sum_k d_{1,k}\psi_{1,k}(t)
\end{aligned} \tag{15}
$$

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

$$
x(t) \approx S_J + D_J + D_{J-1} + ...D_2 + D_1 \tag{16}
$$

where 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. Hence, S_J provides the large scale road map, while D_1 shows the pot holes. Since any term in (12) represents a component of the signal $x(t)$ with a different resolution, the previous equation indicates what is termed the multiresolution decomposition (MRD).

Finally, in this short introduction to wavelet analysis, we might mention the maximum overlap, or non-decimated wavelet transform, MODWT. This transform is a compromise between the CWT, with continuous variations in scale, and DWT where the power of the transform is highly localized. The MODWT is highly redundant, so that the transformations at each scale are not orthogonal; but

---

$$
\begin{aligned}
\int \phi_{J,k}(t)\phi_{J,k^*}(t)dt &= \delta_{k,k^*} \\
\int \psi_{j,k}(t)\psi_{j^*,k^*}(t)dt &= \delta_{j,j^*}\delta_{k,k^*} \\
\int \psi_{j,k}(t)\phi_{J,k^*}(t)dt &= 0
\end{aligned} \tag{13}
$$

where $\delta_{i,j}$ is the Kronecker delta.

the offsetting gain is that applying the transform leaves the phase invariant, a very useful property in analyzing transformations, and the transform is not restricted to limitations imposed by the dyadic expansion used by the DWT.

Figure 9: Time series decomposition, UK, annual data, 1861-2016

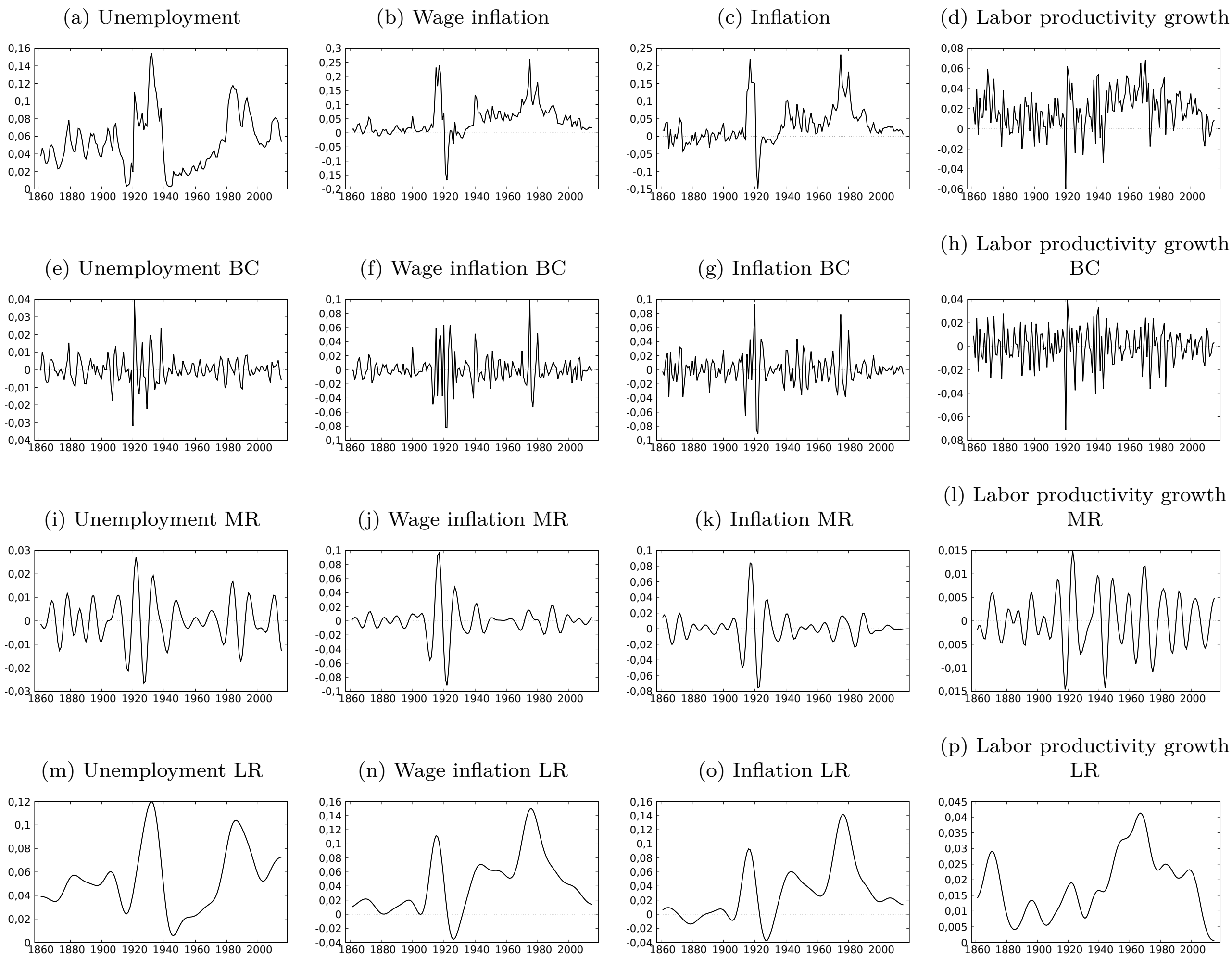
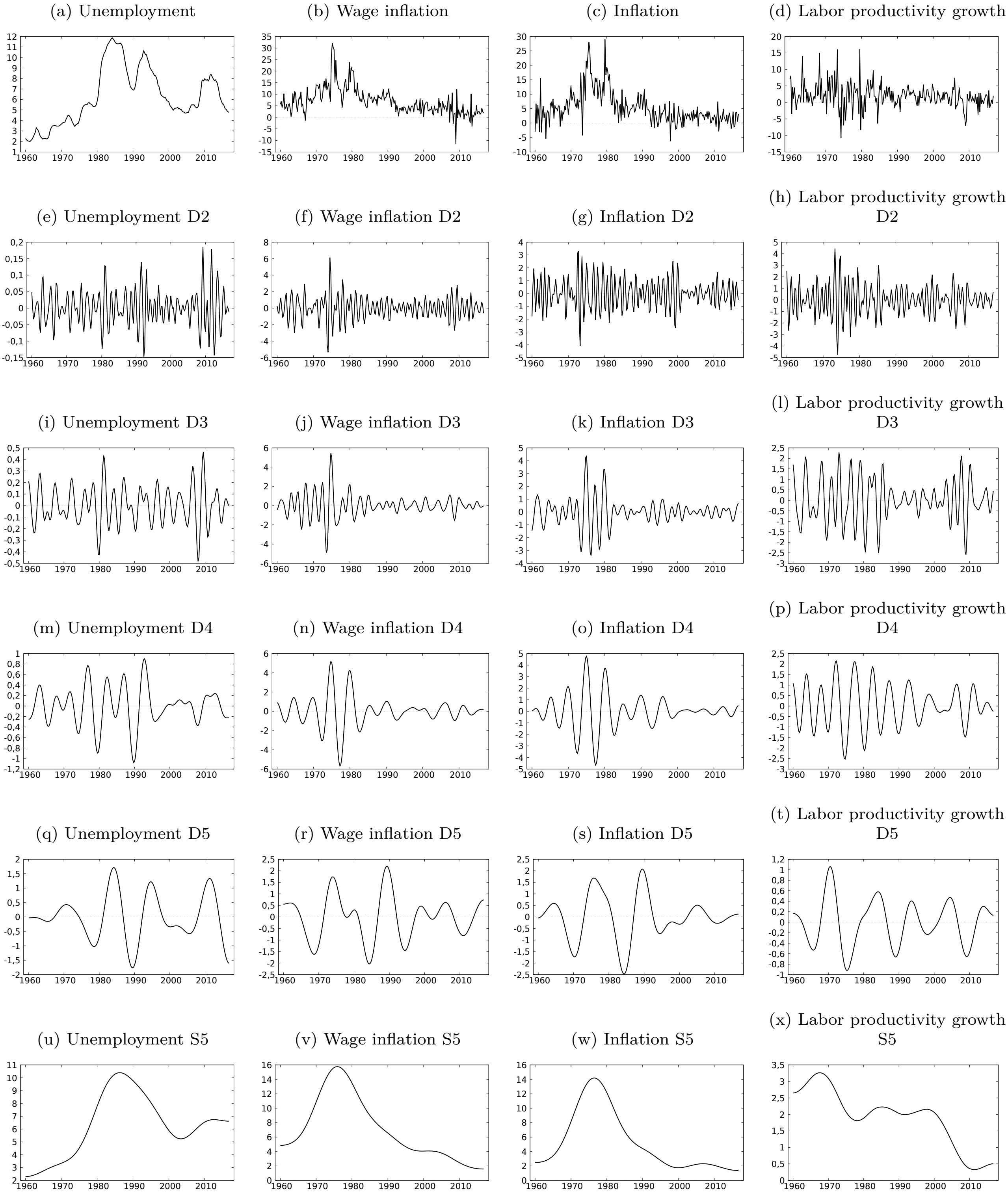


Figure 10: Time series decomposition, UK, quarterly data, 1960-2016



Appendix C: Bandpass filter estimation

Table 6: Dependent Variable: Δw_t^j, 1960:1- 2016:4, UK, bandpass filter, Christiano and Fitzgerald (2003)

	1990:1- 2016:4	1980:1- 2016:4	1970:1- 2016:4	1960:1- 2016:4
U_t^{D2}	0.683	-0.541	-1.940	-3.328
	(2.416)	(2.265)	(2.132)	(2.152)
U_t^{D3}	-0.540	**-1.140**	0.428	0.842
	(0.427)	(0.503)	(0.644)	(0.642)
U_t^{D4}	**-1.113**	0.134	**-0.868**	**-1.005**
	(0.275)	(0.351)	(0.405)	(0.339)
U_t^{D5}	**-1.349**	**-1.392**	**-1.297**	**-0.977**
	(0.075)	(0.083)	(0.127)	(0.124)
ΔLP_t^{D2}	0.024	0.206	**0.529**	**0.453**
	(0.175)	(0.146)	(0.144)	(0.133)
ΔLP_t^{D3}	**0.227**	**0.238**	**0.238**	**0.372**
	(0.086)	(0.058)	(0.060)	(0.077)
ΔLP_t^{D4}	**0.770**	**0.420**	**0.639**	**0.839**
	(0.072)	(0.137)	(0.141)	(0.114)
ΔLP_t^{D5}	**-0.560**	**-0.524**	**-0.693**	**-0.741**
	(0.078)	(0.078)	(0.177)	(0.185)
π_t^{D2}	-0.101	-0.165	-0.236	-0.222
	(0.112)	(0.104)	(0.129)	(0.120)
π_t^{D3}	**-0.279**	0.202	**0.757**	**0.810**
	(0.107)	(0.135)	(0.170)	(0.147)
π_t^{D4}	**0.517**	**1.137**	**1.210**	**1.284**
	(0.093)	(0.147)	(0.078)	(0.071)
π_t^{D5}	**-0.319**	**-0.417**	**-0.455**	-0.143
	(0.066)	(0.074)	(0.118)	(0.130)

Coefficients in bold are significant at a 5 % level. Standard errors are robust to autocorrelation and heteroskedasticity using an HAC estimator for the variance and covariance matrix.