Mis-Specification in Phillips Curve Regressions: Quantifying Frequency Dependence in This Relationship While Allowing for Feedback

(Preliminary; please do not quote without the authors' permission.)

Richard Ashley
Department of Economics
VPI&SU
Blacksburg, VA 24061-0316
Email: ashleyr@vt.edu

Randal J. Verbrugge[1]
Division of Price and Index Number Research
Bureau of Labor Statistics
2 Massachusetts Ave. NE
Washington, DC 20212
Email: verbrugge.randal@bls.gov

October 6, 2004

JEL Classification Codes:

Keywords: Phillips Curve, Inflation, NAIRU, frequency-dependence

Acknowledgement 1 *We thank Marianne Baxter, Walter Enders, Jonas Fisher, Steve Reed, John Williams, and seminar participants at the 2003 North American Summer Meetings of the Econometric Society, the 2004 Midwest Macroeconomics Conference, and at Virginia Tech. "All errors that remain are intended as a test of the reader's knowledge." (Paul McGreal 2002.) All errors, misinterpretations and omissions are ours. All views expressed in this paper are those of the authors and do not reflect the views or policies of the Bureau of Labor Statistics or the views of other BLS staff members.*

[1] Corresponding author.

Abstract

Phillips curve regressions have long been an integral part of empirical macroeconomic research. Here we provide compelling evidence that previous models quantifying the dynamic relationship between inflation and unemployment rates have been mis-specified in their assumption that the coefficient on unemployment is a constant. Instead, we find that this coefficient is frequency-dependent: the inflation impact of a fluctuation in the unemployment rate differs for a fluctuation which is part of a smooth pattern of changes versus a fluctuation which is an isolated event, just as Friedman's "natural rate" hypothesis suggests.

In particular, we analyze a typical Phillips Curve regression model using using a newly developed econometric technique capable of consistently estimating the frequency dependence in a feedback relationship. Explicitly allowing for feedback in such a relationship is essential because the two-sided nature of the Fourier transformations necessary for any sort of frequency domain analysis otherwise confounds the analysis, leading to inconsistent parameter estimates. Once the feedback is properly allowed for by using one-sided filtering, we find statistically significant frequency dependence in the Phillips curve relationship. In particular, using monthly US data from 1980 to 2003, we find an economically and statistically significant inverse relationship between inflation and unemployment for high-frequency unemployment rate fluctuations – with periods ≤ 9 months – but no evidence for an effect of lower frequency unemployment rate fluctuations. In contrast, a model ignoring frequency dependence finds no relationship whatever during this sample period.

1 Introduction

Few macroeconomic relationships have received as much attention as the Phillips curve, which postulates an inverse relationship between inflation and the unemployment rate.[2] This relationship is central to the conduct of contemporary monetary policy. For the first several decades since its introduction, the Phillips curve (augmented with a shifting intercept, and some additional explanatory variables such as oil prices) appeared to be a reasonable approach to understanding inflation dynamics.[3] Even as late as the mid-1990s, some observers (Fuhrer 1995, Gordon 1997) have suggested that such models had been quite successful in "explaining" or tracking inflation, both within and outside of the sample. However, the inflation experience of the 1990s proved more difficult to reconcile with standard Phillips curve models, and has resulted in attempts (e.g., Brayton, Roberts and Williams, 1999; Staiger, Stock and Watson, 2001) to "resurrect" the Phillips curve.

One of the problematic issues involved in Phillips curve estimation involves the natural rate of unemployment, often referred to as the "NAIRU", or non-accelerating inflation rate of unemployment. In 1968, Milton Friedman postulated the existence of a "natural rate" of unemployment, a notion which challenged the entire concept of the Phillips curve. Friedman suggested that the normal dynamic processes of job destruction, search, and job creation would lead to a non-zero equilibrium unemployment rate, and that, in response to macroeconomic conditions, the actual unemployment rate would fluctuate around this natural rate. For example, surprise increases in the money supply would temporarily increase output and reduce the unemployment rate. Over longer horizons, however, Friedman argued that the inflation rate could have no impact on the unemployment rate, since the public would over time adjust its inflation expectations to the new steady-state level of inflation, and the unemployment rate would return to this natural rate irrespective of the new steady-state inflation rate. In particular, summarizing Friedman (1968) and Phelps (1967, 1968), the Phillips curve must be reformulated to include the impact of the public's

[2] Although credit for the discovery of this relationship generally goes to Phillips (1958), one could argue that the original discovery was due to Fisher (1926).

[3] The distinctly positive correlation between the inflation rate and the unemployment rate in the 1970s led many researchers (e.g., Lucas and Sargent, 1978) to cast grave doubt on the existence of a Phillips curve. Indeed, in monthly data, a regression of the inflation rate on twelve lags of the inflation rate and on the unemployment rate – a reasonable-looking specification – yields a statistically insignificant coefficient estimate on the unemployment rate. Our results below yield a possible explanation for this phenomenon.

inflationary expectations, and to take into account the natural rate of unemployment. A Phillips curve thus reformulated is often referred to as an "expectations-augmented" Phillips curve. The events of the 1970s largely bore out the predictions of Friedman and Phelps. The existence of a natural rate, and the importance of inflationary expectations, are consequently no longer seriously contested.

Empirical implementations of Phillips curve must thus come to terms with a natural rate – indeed, changes in the natural rate are often blamed when a particular Phillips Curve specification appears to be breaking down. Though there is no reason to expect that this natural rate is a fixed constant, previous research has largely made this assumption. A number of recent studies have taken the opposite extreme by estimating the relationship in differences, thus tacitly assuming that the natural rate is an I(1) process. Recent studies attempt to model the time evolution of an I(1) natural rate using a Kalman filter approach (e.g., King, Stock and Watson 1995; Debelle and Vickery, 1997; Gordon, 1997, 1998; Gruen, Pagan and Thompson, 1999; Brayton, Roberts and Williams, 1999; Staiger, Stock and Watson 2001), or extract an estimate of the time evolution of the natural rate using splines or low-frequency bandpass filters, as in Staiger, Stock and Watson (1997) and Ball and Mankiw (2001).

These approaches are likely to distort the estimation of the relationships between inflation and unemployment, since they impose arbitrary assumptions as to which frequencies are important. Futhermore – as discussed more explicitly below – the Kalman filter makes specific, most likely counterfactual, assumptions about the manner in which the natural rate evolves over time. Although pre-filtering approaches don't suffer from this particular criticism, there is still no guarantee that splines or low-pass filtering accurately recover the time variation in the natural rate. In particular, we demonstrate below that such two-sided filtering will induce parameter estimation inconsistency in this context if there is any feedback from inflation to the unemployment rate.

Yet decomposing inflation and the unemployment rate by frequency is theoretically appealing. In particular, the Friedman-Phelps hypothesis strongly suggests that the relationship between the inflation rate and the unemployment rate is actually *frequency-dependent*; that is, the relationship between low-frequency movements in the inflation rate (corresponding to the prevailing steady-state inflation rate) and low frequency movements in the unemployment rate (corresponding to changes in the natural rate[4]) will likely be quite different from the relationship of higher-frequency movements in the inflation rate to higher-frequency movements in the unemployment rate. In essence, the Friedman-Phelps formulation suggests that the high frequency movements in these two time series may well have the inverse relationship suggested by Phillips, while the low frequency movements will be unrelated. Clearly, if such frequency-dependence is empirically significant, then a standard Phillips curve model which assumes that the same relationship obtains at all frequencies will yield coefficient estimates that consistently characterize neither of these two distinct relationships.

Below we present a new approach for detecting and modeling frequency dependence in an estimated regression model coefficient, and apply this approach to the Phillips curve relationship. Our approach is formulated in the time domain, so it is easy to implement using ordinary regression software. Moreover, since the new procedure does not require any specification of the dynamics of the natural rate of unemployment, its validity does not hinge on the correctness of such a specification. Indeed, our approach quantifies the frequency dependence in the relationship between inflation and unemployment arising from all sources – natural rate dynamics, policy responses, labor market frictions, etc.

We show below that all presently-available methods for detecting and modeling frequency dependence fail when substantial feedback is present in the relationship, as is the case in the inflation-unemployment relationship. This failure is due to the two-sided nature of the filtering – Hodrick-Prescott, Baxter-King, or even ordinary X-11 seasonal adjustment – used in these approaches to isolate a specific range of frequencies for analysis. Fundamentally, the two-sided filtering interacts with the feedback in the relationship to induce correlations between the filtered series and the relevant regression error terms, thus producing inconsistent parameter estimates.

In this paper we describe an extension to the Tan and Ashley (1999) frequency-dependence modeling framework which overcomes this problem. Simulations using artificially generated data demonstrate that the new technique is able to correctly detect frequency dependence in the presence of feedback, and illustrates the distortions created when feedback is not properly handled.

Applying this new technique to allow for both frequency dependence and feedback in a stan-

[4]Hall (1999) and Cogley and Sargent (2001) argue that the low frequency trend component of the unemployment rate is an estimate of the natural rate; Staiger, Stock and Watson (2001) adopt this argument.

dard Phillips curve formulation, we find statistically significant frequency dependence in the Phillips curve relationship, of a sort that is consistent with the Friedman-Phelps theory. In particular, when we allow the data to select the optimal set of frequency bands based upon a BSIC criterion, we find that a two-band model is chosen. In this model, there *is* a statistically significant inverse relationship between inflation and unemployment, but it is restricted to unemployment rate fluctuations in the high-frequency band, which includes frequencies corresponding to unemployment rate fluctuations with periods less than 9 months. This frequency dependence is significant at the 5% level, even accounting for the specification search involved in choosing the number, and extents, of the bands.

The outline of the remainder of the paper is as follows. Section 2 presents the underlying macroeconomic theory and briefly discusses prior empirical work. Section 3 describes the econometric methodology proposed here, and in particular includes a critique of two-sided filtering in the presence of feedback. Section 4 presents simulation evidence which indicates that the new methodology of Section 3 is both necessary and effective. Section 5 presents the empirical results on the Phillips curve. Section 6 concludes the paper.

2 Theory and Prior Empirical Work

As noted above, the Phillips Curve has long been the focus of empirical work. The prototypical expectations-augmented Phillips curve is the specification

$$\pi_t = \pi_t^e + \beta\left(un_t - un_t^N\right) + \varepsilon_t$$

where π_t is actual inflation (in wages, or in an appropriate price index) during period t, π_t^e is the level of inflation that was expected to occur during period t, un_t is the unemployment rate at time t, and un_t^N is the natural rate of unemployment at t. Two difficulties arise, each relating to one of the unobserved components in the above relationship: π_t^e and un_t^N.

First consider the treatment of expected inflation, π^e. The random-walk model of expectations, which specifies that $\pi_t^e = \pi_{t-1}$, has been used extensively in the literature (e.g., Gordon 1990, 1998, Fuhrer 1995, Staiger, Stock and Watson 2001). This assumption is reasonably consistent with the data but, because inflation is observed to have considerable inertia, a number of lags of inflation are required in the specification to ensure that the resulting regression model errors are serially uncorrelated. This has led to regression models of the following form:

$$\pi_t = \beta\left(un_t - un_t^N\right) + \sum_{j=1}^{m} \delta_j \pi_{t-j} + \theta Z_t + \varepsilon_t \tag{1}$$

where the condition $\sum_{j=1}^{m} \delta_j = 1$ is often imposed.[5] Since deterministic seasonal components have frequently been observed in seasonally-unadjusted inflation data, monthly dummies are often included as well. Finally, since the 1970s it has become common practice to also include in this specification price control dummy variables and measures of "supply shocks," such as the relative price of energy. Shocks to such variables arguably create positive correlation between inflation and unemployment, and would thus bias the estimate of β if omitted. All such control variables are here collected in the vector Z_t.

The second difficulty, the unobserved natural rate, has been handled in a variety of ways. Most Phillips curve regression specifications implicitly assume that the natural rate is a constant, in which case a regression of the following form is appropriate:

$$\pi_t = \widetilde{\alpha} + \beta un_t + \sum_{j=1}^{m} \delta_j \pi_{t-j} + \theta Z_t + \varepsilon_t \tag{2}$$

where the natural rate can be recovered from estimates of the coefficients $\widetilde{\alpha}$ and β. Occasional shifts in an otherwise constant natural rate have been handled by allowing for shifts in the intercept.

Recently, several authors have explored more sophisticated methods to allow for a potentially time-varying natural rate. For example, Staiger, Stock and Watson (1997) model the natural rate as a flexible polynomial, estimating a time-varying constant in (2), from which a time-varying natural rate estimate can be recovered. A variant of this method (e.g., Ball and Mankiw, 2002) involves identifying a filtered version of the unemployment rate with the natural rate for use in equation (1).

[5] This condition is related to a unit root in inflation; some authors (e.g., Gordon, 1997) assert that this restriction is necessary for a natural rate that is consistent with a constant rate of inflation. However, the existence of a unit root in inflation partly depends upon Fed policy: if the Fed stabilizes inflation around a target, there will be no unit root in inflation, and forward-looking models will not generate a unit-sum restriction.

Some authors (e.g., Stock and Watson, 1999) impose the restriction that inflation is I(1) by specifiying the Phillips curve relation using first-differences of inflation. Since, as emphasized by Baxter 1995, first-differencing removes most of the low- and medium-frequency components of the series, this will substantially distort least-squares estimates of the coefficient β if the relationship is frequency-dependent.

An alternative method uses the Kalman filter to estimate the natural rate as an unobserved component. Typically, the natural rate is modeled as a unit root process in this framework, yielding the two-equation system:

$$\begin{aligned} \pi_t &= \alpha + \beta \left(un_t - un_t^N \right) + \sum_{j=1}^{m} \delta_j \pi_{t-j} + \theta Z_t + \varepsilon_t \\ un_t^N &= un_{t-1}^N + \upsilon_t \end{aligned} \tag{3}$$

where un_t^N is a latent or unobserved variable, and $(\varepsilon_t, \upsilon_t)$ are assumed to be jointly NIID. The variance of υ_t is either imposed a priori, or – as in Gruen, Pagan and Thompson (1999)– is estimated by suitably concentrating out the log-likelihood. In practice, the estimated natural rate closely tracks the univariate trend in the unemployment rate, regardless of methodology – e.g., see Brayton, Roberts and Williams (1999) or Staiger, Stock and Watson (2001). But this does not necessarily imply accurate tracking of the natural rate dynamics. Furthermore, if the relationship between inflation and $\left(un_t - un_t^N \right)$ is itself frequency dependent, the OLS estimate of β will be inconsistent even if the natural rate dynamics *are* correct.

This paper presents a new approach to the specification of the Phillips curve relationship. We begin with the standard Phillips curve relationship specification embodied in equation (2) (including in Z_t variables to model inflation expectations) and account for variation in the natural rate (and remaining variation in inflation expectations) by allowing the coefficient β to vary across frequencies. Since feedback from inflation to unemployment rates is an important element of the Phillips curve relationship, we develop new econometric tools for quantifying frequency dependence in feedback relationships.

3 Methodology

In this section we explain what frequency dependence is, what it is not, and why it makes a difference. Subsections 3.3 and 3.4 discuss the Tan/Ashley approach to the detection and modeling of frequency dependence in the absence of feedback and its straightforward implementation in the time domain. Section 3.5 addresses the issue of how to select the number of frequency bands to consider and the particular set of frequencies to be included in each band. Finally, Section 3.6

discusses the problematic nature of two-sided filtering in the context of feedback relationships and describes how we modify the Tan/Ashley methodology appropriately to deal with this problem.

It is best to be clear at the outset as to the meaning of the term "frequency dependence" in the context of a regression coefficient. Consider the following aggregate consumption function:

$$c_t = \gamma_0 + \gamma_1 y_{t-1} + \gamma_2 y_{t-2} + \gamma_3 c_{t-1} + \varepsilon_t \tag{4}$$

where c_t is the log of aggregate consumption spending in period t, y_t is the log of disposable income in period t, and ε_t is a covariance-stationary error term. In this model γ_1 is the "short-run marginal propensity to consume," characterizing how consumption spending (on average) responds to fluctuations in y_{t-1}. In contrast, $\frac{(\gamma_1+\gamma_2)}{(1-\gamma_3)}$ is the "long-run marginal propensity to consume," the change in steady-state consumption from a one unit change in steady-state income; it answers the question, "How does average steady-state consumption spending vary across different steady-state after-tax income levels?" The distinction between γ_1 and $\frac{(\gamma_1+\gamma_2)}{(1-\gamma_3)}$ is *not* what we mean by frequency-dependence.

What we *do* mean by frequency-dependence is that, according to the permanent-income hypothesis, the value of γ_1 itself depends upon frequency. In particular, this hypothesis asserts that consumption should *not* change appreciably if the previous period's fluctuation in income is highly transitory (high-frequency), whereas consumption *should* change significantly if the previous period's fluctuation in income is part of a persistent (low-frequency) movement in income. γ_1, then, should be approximately equal to zero for high frequencies, and close to one for very low frequencies. Equation (4), in contrast, incorrectly restricts γ_1 to be the same across all frequencies.

This frequency dependence in γ_1 implied by the permanent income hypothesis concomitantly implies that γ_1 varies over time. In the special case of adaptive expectations, for example, the implication is that γ_1 will be larger if y_{t-1} has the same sign as y_{t-2}. Thus, this frequency dependence in γ_1 can be viewed as a symptom of unmodeled nonlinearity in the relationship between c_t and y_{t-1}. This aspect of frequency dependence is discussed at some length in Tan and Ashley (1999). Here, the essential point is that this frequency dependence in γ_1 further implies that the value of γ_1 is not a fixed constant; rather, it varies over time, due to its dependence on $y_{t-1}, y_{t-2}, y_{t-3}$, etc.[6]

[6] Similarly, viewing equation (4) as part of an bivariate VAR model, the impulse response function for

3.1 Consequences of frequency dependence

Now consider a simple bivariate time series model:

$$y_t = \beta x_t + \varepsilon_t \qquad \varepsilon_t \sim NIID\left[0, \sigma^2\right]$$

for $t \in \{1, ...T\}$. The parameter β can be interpreted as $dE[y_t|x_t]/dx_t$; if β actually takes on two values – β_0 in the first half of the sample and β_1 in the second half of the sample – then this regression is clearly mis-specified. In that case, the usual statistical machinery for testing hypotheses about β is invalid – indeed, the hypotheses themselves are essentially meaningless, since β does not have a well-defined value to test. Similarly, the least-squares estimate of β cannot be a consistent estimator for either β_0 or β_1. In particular, if the sign of the relationship is positive in the first part of the sample and negative later on, then the least squares estimate of β might well be close to zero, leading to the erroneous conclusion that y_t and x_t are unrelated.

One of the key implications of the spectral regression model of Engle (1974, 1978) – summarized in section 3.3 below – is that β is stable across time if and only if it is stable across *frequencies*; this was also discussed in the context of the consumption function in the previous section. Thus, if the value of β is different at low frequencies than at high frequencies, then β varies over time also, albeit in a manner which might be difficult to detect with time domain parameter stability tests. Still, this result implies that frequency variation in β yields all of the same unhappy properties as does time variation. In particular, the least squares estimator of β is an inconsistent estimator of $dE[y_t|x_t]$ with respect to x, and – since β does not have a unique value – hypothesis tests about β are problematic to interpret.

Frequency dependence in the unemployment rate coefficient of equation (2) might arise from mis-specified dynamics for the natural rate; or it could occur for other reasons. We take such frequency dependence to be an empirical issue – one which is consequential for the foregoing reasons – and below develop methods for detecting and correcting for it.

c_t will be a linear function of past innovations in both equations. While c_t may well depend differently on different lags in the y_t innovations, if there is no frequency dependence in the $c_t - y_t$ relationship, then the coefficients in this impulse response function will all be constants. In contrast, frequency dependence in the relationship implies that a coefficient on one of the y_t innovations in the c_t impulse response function itself depends on the value of previous innovations. Thus, for example, in that case the coefficient quantifying the impact of a y_t innovation on subsequent values of c_t itself depends on whether this y_t innovation was an isolated event, or part of a pattern of similar previous y_t innovations.

3.2 Pseudo-frequency dependence

It is important to distinguish 'true' frequency dependence in a relationship from a superficially similar concept in which the coefficients of the model quantifying the relationship are constant, but the *coherence* (closely related to the magnitude of the cross-spectrum of the variates) is frequency-dependent. This latter notion is used in Geweke (1982), Diebold, Ohanian and Berkowitz (1998), and a host of other studies. These decompositions are mathematically sound, but we call what they measure 'pseudo-frequency dependence' because such measures do not actually quantify frequency variation in the relationship itself.

A simple example clarifies this distinction. Consider the following consumption relation,

$$c_t = \beta y_{t-1} + u_t + \phi u_{t-1}$$

$$\begin{pmatrix} u_t \\ y_t \end{pmatrix} \sim NIID \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_u^2 & 0 \\ 0 & \sigma_y^2 \end{pmatrix} \right]$$

The marginal propensity to consume in this relationship is clearly a constant (β) and Fourier transforming both sides of this equation will do nothing to change that – it merely yields a relationship between the Fourier transform of c_t and the Fourier transform of y_{t-1}, still with a constant coefficient β. (E.g., see Section 3.3 below.) But the cross-spectrum and coherence functions relating c_t and y_t are *not* constants: by construction, they depend explicitly upon the frequency parameter ω. In particular, Geweke (1982)'s measure of the strength of the linear dependence of c_t on y_{t-1} (a generalization of the coherence function) for this model is:

$$f_{y\to c}(\omega) = \frac{1}{2} \ln \left\{ \frac{\sigma_u^2 \left(1 + \phi^2 - 2\phi \cos(\omega)\right) + \beta^2 \sigma_y^2}{\left[\sigma_u^2 \left(1 + \phi^2 - 2\phi \cos(\omega)\right)\right]^2} \right\}$$

which clearly *does* depend upon frequency so long as the moving average parameter ϕ is not zero.

Evidently, this frequency dependence in Geweke's measure (and in the other 'strength of association' measures based upon the cross-spectrum and the coherence function) is not quantifying the frequency variation in the c-y relationship itself, since there is none to quantify. So what *is* it doing? These kinds of measures are usually interpreted as quantifying the degree to which the overall R^2 for the equation is due to sample variation at low frequencies versus high frequencies.

Suppose that ϕ is positive, in which case Geweke's measure indicates that low frequencies are important to the R^2 of the relationship. This says nothing about whether consumption and income are differently related at low versus high frequencies – that depends upon the marginal propensity to consume, which is constant. Rather, it says that this dynamic relationship transforms serially uncorrelated fluctuations in y_{t-1} and u_t into positively correlated fluctuations in c_t. Alternatively, one could observe that c_t in that case has substantial spectral power at low frequencies, and interpret this result, to paraphrase Geweke (1982, p. 312), as indicating that the white noise innovations in y_{t-1} explain most of this low frequency portion of the variance in c_t.[7]

3.3 Regression in the frequency domain in the absence of feedback

The most elegant way to assess the actual frequency dependence of a regression coefficient is to estimate the regression equation in the frequency domain. Such spectral regression was originally proposed by Hannan (1963) and most clearly exposited in Engle (1974, 1978). Following Engle, spectral regression is based on the simple notion that a multiple regression model in the time domain, such as

$$Y = X\beta + \varepsilon \qquad \varepsilon \sim N\left[0, \sigma^2 I\right] \tag{5}$$

can be Fourier-transformed on both sides of the equation via multiplication by a complex-valued matrix W, yielding

$$\begin{aligned} WY &= WX\beta + W\varepsilon \qquad (6) \\ \widetilde{Y} &= \widetilde{X}\beta + \widetilde{\varepsilon} \qquad \widetilde{\varepsilon} \sim N\left[0, \sigma^2 I\right] \qquad (7) \end{aligned}$$

where $\widetilde{Y} = WY$, etc., and where the $(j,k)^{th}$ element of W is given by $w_{j,k} = \frac{1}{\sqrt{T}} \exp\left(\frac{2\pi ijk}{T}\right)$, with T equal to the sample length. The variance of $\widetilde{\varepsilon}$ is still $\sigma^2 I$ because W is an orthogonal matrix.

Note that the coefficient vector β is identical in both regression equations. What *has* changed, however, is that the T sample observations in Y and in each column of X are replaced by T 'observations' on each variable, each of which now corresponds to a frequency in the interval

[7] Note also that both the coherence and gain functions are, by construction, non-negative at all frequencies. Thus, neither of these concepts can possibly capture frequency dependence as discussed here, which can readily involve a regression coefficient having one sign at low frequencies and the opposite sign at high frequencies.

$[0, 2\pi (T-1)/T]$. In particular, one can identify the j^{th} 'observation' in this transformed regression model as corresponding to frequency $2\pi (j-1)/T$.

Note, however, that consistent least squares estimation of β in equation (7) requires that $corr(\widetilde{x}_{j,k}, \widetilde{\varepsilon}_j)$ is zero for all values of j and k. Since W embodies a two-sided transformation – i.e., $\widetilde{x}_{j,k}$ depends upon all of $x_{1,k}, ..., x_{T,k}$ and $\widetilde{\varepsilon}_j$ depends upon all of $\varepsilon_{1,k}, ..., \varepsilon_{T,k}$ – this condition requires that $x_{t,k}$ be uncorrelated with both past and future values of ε_t. This issue is taken up more explicitly in Section 3.6 below; it is side-stepped here by restricting attention to relationships in which there is no feedback between y_t and $x_{1,k}, ..., x_{T,k}$.

This framework has unique advantages over regression in the time domain. For example, missing observations and distributed lag expressions involving non-integer lags can be dealt with fairly readily in the frequency domain. And – vital for the present context – detecting and modeling frequency variation in a component of β corresponds precisely to testing for instability in this component across the sample 'observations' in equation (6).

Prior to Tan and Ashley (1999), however, this framework also had some fairly intense drawbacks, which severely limited its usefulness and acceptance. For one thing, $\widetilde{Y}$ and $\widetilde{X}$ are complex-valued, precluding the use of ordinary regression software to estimate β. An estimator for β can be expressed in terms of the cross-periodograms of Y and the columns of X – e.g., equation 10 of Engle (1974) – but the calculations still require specialized software. Consequently, Engle's approach is really only convenient for considering parameter variation over at most *two* frequency bands: in that special case it is possible to finesse the problem so that ordinary regression software suffices.[8]

Another problem with Engle's framework is really just cosmetic, but nevertheless effectively limits the credibility of the results: one cannot drop a group of, say, the five lowest frequency 'observations' without also dropping the five observations at the highest five frequencies – otherwise, the least squares estimate of β is no longer real-valued. These latter five observations, at what appear to be the five highest frequencies, in fact actually do correspond to low frequencies because of symmetries in the W matrix, but one is apt to lose one's audience in trying to explain it.

Finally, Engle's formulation does not deal with econometric complications such as simultaneity,

[8] Later work by Thoma (1992, 1994) pushes this idea a bit further by observing how the parameter estimate varies as more frequencies are added to the low frequency band.

cointegration, or, as noted above, feedback. Phillips (1991) provides a framework for estimating cointegrated systems in the frequency domain based directly on Hannan's formulation in terms of the spectra and cross-spectra of the data. But this approach again requires specialized software, and is sufficiently sophisticated as to severely limit the ability of most practitioners to modify it as needed in order to deal with the particular problems posed by individual applications.

The net result is that spectral regression methods have been applied to the frequency dependence problem for only a handful of macroeconomic relationships.

The approach developed in Tan and Ashley (1999) effectively eliminates the objections noted above, at least for non-feedback systems. This formulation is similar in spirit to Engle's except that the complex-valued transformation matrix (W) is replaced by an equivalent *real*-valued transformation matrix (A) with $(j,t)^{th}$ element:

$$a_{j,t} = \begin{cases} \frac{1}{\sqrt{T}} & j = 1 \\ \sqrt{\frac{2}{T}} \cos\left[\frac{\pi j(t-1)}{T}\right] & j = 2, 4, ..., (T-2) \text{ or } (T-1) \\ \sqrt{\frac{2}{T}} \sin\left[\frac{\pi (j-1)(t-1)}{T}\right] & j = 2, 4, ..., (T-1) \text{ or } T \\ \frac{1}{\sqrt{T}} (-1)^{t+1} & j = T \text{ and } T \text{ is even, } t = 1, ..., T \end{cases} \tag{8}$$

This transformation, which first appears in Harvey (1978), yields a real-valued frequency domain regression equation

$$AY = AX\beta + A\varepsilon \qquad A\varepsilon \sim N\left[0, \sigma^2 I\right]$$

or

$$Y^* = X^*\beta + \varepsilon^* \qquad \varepsilon^* \sim N\left[0, \sigma^2 I\right] \tag{9}$$

with $Y^* = AY$, etc. In fact, each row of A is just a linear combination of two rows in the W matrix, based on the usual exponential expressions of the sine and cosine – e.g., $\cos(x) = \frac{1}{2}e^{ix} + \frac{1}{2}e^{-ix}$. Again, $Var(\varepsilon^*) = Var(\varepsilon)$ bacause A is an orthogonal matrix.

Since the elements of the A matrix are all real-valued, equation (9) can be estimated using ordinary regression software. Moreover, the effect of the transformation on a column vector (e.g., Y) is now plain to see. The second and third rows of the A matrix ($j = 2$ and 3) correspond to the two 'observations' at the lowest non-zero frequency. The weights in these rows make one complete oscillation over the T periods in the actual sample, so any fluctuation in Y_t that is sufficiently brief as to average out to essentially zero over a period of length $T/2$ will have little impact on either Y_2^* or Y_3^*. In contrast, suppose that T is even and consider the highest frequency row of A. This row simply averages $T/2$ changes in the data; clearly, it is ignoring any slowly-varying components of the data vector and extracting the most quickly-varying component.

The "observations" in this regression model thus do correspond to frequencies. Consequently, frequency variation in, say, β_k – the k^{th} component of β – can be assessed by applying any of the variety of procedures in the literature for examining the variation in an estimated regression coefficient across the sample observations: e.g., Chow (1960), Brown, Durbin and Evans (1975), Ashley (1984), or Bai (1997) and Bai and Perron (1998, 2003). We will return to this issue in Section 3.5; for now, we observe that Tan and Ashley (1999) use the procedure given in Ashley (1984) and simply partition the T 'observations' in equation (9) into m equal frequency bands and estimate how β_k varies by replacing the k^{th} column of X^*, X_k^*, with m appropriately constructed dummy variables:[9]

$$Y^* = X^*_{\{k\}}\beta_{\{k\}} + D^*\gamma^* + \upsilon^* \tag{10}$$

where $X^*_{\{k\}}$ is X^* omitting the k^{th} column, and $\beta_{\{k\}}$ is β omitting the k^{th} component. The columns $\left[D^{*1}...D^{*m}\right]$ composing the D^* matrix consist of m new explanatory variables, one for each frequency band – D_j^{*s}, the j^{th} component of the new explanatory variable for frequency band s, is zero for each component outside the frequency band, and equal to the corresponding component of X_k^* (the k^{th} column of X^*) for each component inside the frequency band.

3.4 Time domain version of the Tan/Ashley approach

It is both helpful and instructive to re-cast the Tan-Ashley formulation in the time domain. Since A is an orthogonal matrix, A^{-1} is just its transpose, A^T. Multiplying the regression model of (10) through by A^T yields

$$A^TY^* = A^TX^*_{\{k\}}\beta_{\{k\}} + A^TD^*\gamma^* + A^T\upsilon^* \tag{11}$$

and hence

$$Y = X_{\{k\}}\beta_{\{k\}} + D\gamma + \upsilon \tag{12}$$

[9] Simulations in Ashley (1984) indicate that this modest generalization of the Chow test performs at least as well as more sophisticated alternatives with samples of moderate length.

Here Y is the original dependent variable data vector and $X_{\{k\}}$ is the original data matrix, omitting the k^{th} column.

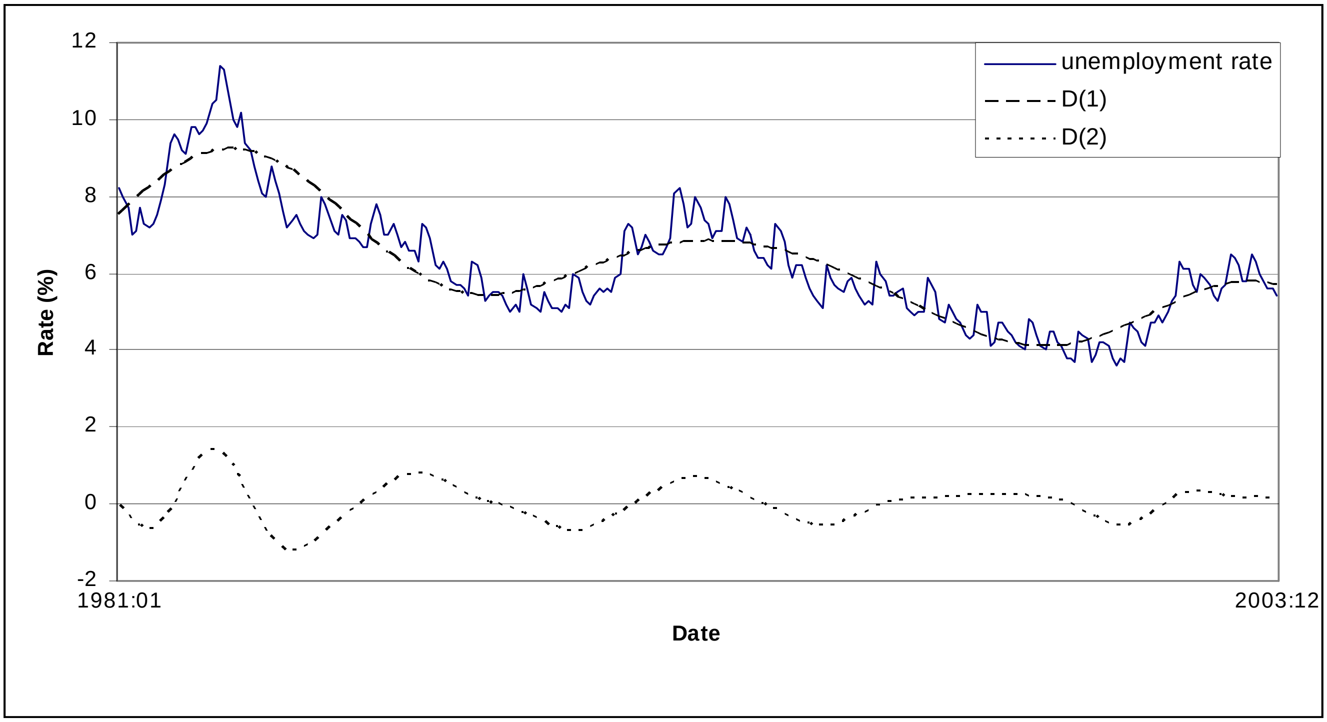
The matrix $D = \left[D^1 ... D^m\right]$ thus has as its columns the back-transforms of the frequency-domain explanatory variables $\left[D^{*1} ... D^{*m}\right]$ corresponding to each of the m frequency bands being considered. Note that, since the columns $\left[D^{*1} ... D^{*m}\right]$ are orthogonal and add up to $X_k^* = AX_k$, the column vectors comprising $\left[D^1 ... D^m\right]$ are orthogonal also and add up to X_k, the original data vector for the k^{th} explanatory variable;[10] consequently, the error vector v is identical to the original error term in (5) if the m components of γ are all equal to β_k.

Thus, the column vectors $\left[D^1 ... D^m\right]$ are in essence bandpass filtered versions of X_k which partition this variable into m orthogonal components, one for each frequency band. For example, suppose that one were to partition the monthly US unemployment rate into three frequency components: D_t^1, comprising the fluctuations corresponding to low frequencies (periods greather than 72 months); D_t^2, a medium-frequency ("business cycle") component, corresponding to periods between 18 and 72 months; and D_t^3, a high-frequency component, corresponding to periods less than 18 months. Figure 1 plots the monthly US unemployment rate, along with D_t^1 and D_t^2 – the first and second of these components – using data from 1980 through 2003.

Note than no one of these implied bandpass filters is an optimal bandpass filter – one might choose a Baxter-King (1999) or Christiano-Fitzgerald (2003) bandpass filter for that purpose – but $\left[D^1 ... D^m\right]$ have the desirable property of slicing up X_k in an intuitively appealing way into m orthogonal components that add up exactly to X_k. Therefore, replacing $\beta_k X_k$ by $D\gamma$ in the regression equation allows one to conveniently test for, and model, frequency dependence in β_k, with frequency stability corresponding to the null hypothesis that all m components of γ are equal.

Figure 1: Time Plot of the US Unemployment Rate and its Low- and Medium-Frequency Components (D_t^1 and D_t^2)

[10] Tan and Ashley (1999) give an explicit example of this with $m = 3$ frequency bands. Given their particular partitioning , they show how D^{*1} is zero except for the first third of the 'observations' (corresponding to the lowest frequencies) – yielding a smooth D^1 time domain series – whereas D^{*3} is zero except for the last third of the 'observations' (corresponding to the highest frequencies), and yields a rapidly varying D^3 time domain series. They do not, however, point out that the m filtered components $\left[D^1 ... D^m\right]$ are orthogonal.



Note that failing to replace β_k by $D\gamma$ when the m components of γ are not equal yields a mis-specified regression model for Y: $\widehat{\beta}_k^{OLS}$ cannot possibly be consistent for β_k in this model since β_k does not have a unique value to estimate.

Moreover, note also that there is nothing essential about the simple form of the original model ($Y = X\beta + \varepsilon$) in the analysis above. One could just as easily investigate the frequency dependence of the coefficient on X_k by replacing it with the weighted sum $D\gamma$ *regardless of how* X_k *enters the analysis* - linearly or nonlinearly, instrumented or not, etc. – using essentially the same techniques and software one was already employing.

Finally, note that, since $X_k = D^1 + ... + D^m$, using the $D^1...D^m$ instead of X_k in a regression model leaves the properties of the error term unaffected under the null hypothesis of no frequency dependence. No sample information is lost; the only statistical cost is a loss of $m - 1$ degrees of freedom, since more coefficients are being estimated.

The frequency decomposition proposed here has important advantages over a typical bandpass-filtering approach in which X_k is bandpass-filtered in m different ways so as to obtain m frequency components $F^1, ..., F^m$. The F^j thus obtained are not in general orthogonal to one another. More importantly, however, they do not add up to the original data vector X_k. Thus, even under the null hypothesis that these components enter the model with equal coefficients, this kind of decomposition fails to preserve the model error distribution. Moreover, since $F^1, ..., F^m$ do not sum up to X_k, an indeterminate amount of the actual sample variation in X_k is lost due to the decomposition, rendering any interpretation of the results problematic.

3.5 Frequency Band Specification

Selecting the number of frequency bands, and the particular set of frequencies to be included in each band, is an important issue in implementing the analysis described above.

One alternative is to simply specify these bands on a priori grounds; this is analogous to common practice in empirical macroeconomics, where attention is often restricted to "business cycle" frequencies. In the present context, this "calendar-based" approach might suggest a three-band formulation – one band containing all frequencies corresponding to periods of less than 18 months, a second band containing frequencies corresponding to periods between one and a half and six years, and a third containing frequencies corresponding to longer periods. This choice seems reasonable, but it is somewhat ad hoc – one might equally well choose one of many other calendar-based frequency band structures.

An alternative approach, adopted here, is to allow the data to choose the frequency band structures. Since each band structure corresponds precisely to assuming that the relevant regression coefficients are constant over the analogous sets of "observations" in equation (12) above, this amounts to a problem which has received a good deal of attention in the literature recently – e.g., Bai (1997), Bai and Perron (1998, 2003).

It is computationally feasible to search over all possible band structures with the number of bands less than some maximum value, and to choose the one which minimizes some adjusted goodness-of-fit criterion, such as the Bayes-Schwarz Information Criterion (BSIC). Of course, one must then estimate the sampling distribution of the F statistic for testing the null hypothesis of equal coefficients on all m bands – using the bootstrap, for example – so as to prevent the size distortion which this extensive specification search will surely induce. The resulting test will have unnecessarily low power, however. Instead, exploiting the fact that we expect the parameter variation in this case to be relatively smooth, the extent of the specification search is substantially reduced by using a variation of the "regression tree" approach. In particular, the search is restricted in two ways:

(1) We constrain both "observations" corresponding to a given frequency to remain in the same frequency band. (Recall from equation (8) that each non-zero frequency corresponds to both a sine and a cosine in the A matrix.)[11]

(2) The procedure begins by assuming that there is just one band, and searches for a single sample split which improves the BSIC. If none is found, the search is over. If such a sample split *is* found, then this "breakpoint" is no longer modified. Instead, each of the two bands implied by this breakpoint is examined to see if it can be split in two so as to improve the BSIC. This process continues until either the maximum number of bands to be considered is reached, or no BSIC-improving split of the existing bands can be found.

This latter search restriction is conservative in that the fully-optimal m band specification might never be examined because the best $(m-1)$ band structure is not nested within it. However, since the parameter variation test is bootstrapped to appropriately account for the amount of specification search, on balance the substantial reduction in (mostly irrelevant) search activity provided by this restriction notably increases the power of the procedure to detect parameter variation.

3.6 Dealing with feedback

Note that $\widehat{\gamma}^{OLS}$ will be a consistent estimate of γ in equation (12) if and only if the error term in this equation is uncorrelated with each of the regressors $D^1 ... D^m$. Since the t^{th} observation on each of these regressors is the result of what amounts to a two-sided nonlinear bandpass filter applied

[11] Except in the case where the number of observations is even, for the highest frequency.

to $X_{k,t}$, this will be the case only if X_k is strongly exogenous, that is, only if every observation on X_k is uncorrelated with every observation on the error term in the original regression model. (This is, of course, equally the case for *any* methodology which applies a two-sided bandpass filter to the k^{th} regressor.) Unfortunately, feedback in the relation between Y_t and $X_{k,t}$ induces exactly this kind of correlation.

For example, consider the analysis of possible frequency dependence in the parameter λ_2 of the following bivariate equation system:

$$\begin{aligned} y_t &= \lambda_1 y_{t-1} + \lambda_2 x_{t-1} + \varepsilon_t \qquad (13) \\ x_t &= \alpha_1 x_{t-1} + \alpha_2 y_{t-1} + \eta_t \end{aligned}$$

Clearly, feedback exists if and only if α_2 is nonzero. Any two-sided filter – one based on the A matrix discussed above, or bandpass filters such as those given by Baxter and King (1999) or Christiano and Fitzgerald (2003), or the Hodrick-Prescott filter – applied to x_{t-1} in equation (13) will yield a transformed value x^*_{t-1} which depends upon $x_t, x_{t+1}, x_{t+2},...x_T$. But note that equation (13) implies that

$$\begin{aligned} x_{t+1} &= \alpha_1 x_t + \alpha_2 y_t + \eta_{t+1} \\ &= \alpha_1 x_t + \alpha_2 \left(\lambda_1 y_{t-1} + \lambda_2 x_{t-1} + \varepsilon_t\right) + \eta_{t+1} \\ &= \alpha_1 x_t + \alpha_2 \lambda_1 y_{t-1} + \alpha_2 \lambda_2 x_{t-1} + \alpha_2 \varepsilon_t + \eta_{t+1} \end{aligned}$$

which is clearly correlated with the regression error term ε_t unless α_2 is zero, in which case there is no feedback. Thus, in the presence of feedback, any transformation of x_{t-1} which depends upon x_{t+1} will be correlated with the model error term, yielding inconsistent least-squares parameter estimates.

To eliminate this problem, we exploit the fact that the Tan/Ashley formulation is easily adapted to use only one-sided filtering. The only cost is a modest amount of additional computation and the loss of the use of the first τ sample observations in estimating equation (12), where τ is the period corresponding to the lowest frequency separately distinguished in the analysis. The calculation steps through the sample using blocks of length τ. In the first step, observations one through τ on X_k (i.e, $X_{1,k}, ..., X_{\tau,k}$) are used to compute the τ-dimensional column vectors $D^1...D^m$, one for each of the m frequency bands. The last (period τ) element in each of these vectors becomes the period τ observation on $D^1...D^m$ for use in estimating equation (12). Next one uses the τ sample observations $X_{2,k}, ..., X_{\tau+1,k}$ to re-compute the τ-dimensional column vectors $D^1...D^m$. Again the last (τ^{th}) element in each of these vectors becomes the period τ+1 observation on $D^1...D^m$ for use in estimating equation (12). And so forth.[12] Thus, one could characterize $D^1...D^m$ as being the result of a set of m one-sided bandpass filters obtained using a moving block of τ observations.

The resulting $D^1...D^m$ columns still add up precisely to X_k over its last $T - \tau$ elements. These m columns are no longer precisely orthogonal to one another, but in practice they are quite close to being orthogonal. In any case, the orthogonality is of modest importance: what is essential is that $D^1...D^m$ still partition (sum up to) X_k and are now the product of a one-sided filter.

It is unfortunate that one must lose the use of the $\tau - 1$ start-up observations in estimating equation (12) in this way, but this is necessary in order to avoid spurious results when feedback is present. This loss is analogous to the start-up observations "lost" in using lagged variables in an equation. Indeed, this loss is noticeable in the Phillips curve application given in Section 5 below: 60 observations out of 288 are sacrificed so as to be able to consider frequencies corresponding to periods as large as sixty months.

Lastly, it must be mentioned that bandpass filters like the ones used here generically have problems near the endpoints of the sample. This is not surprising. As Christiano and Fitzgerald (2003) put it, "it is hard to say without the benefit of hindsight whether a given change in a variable is temporary ... or more persistent." The standard method for addressing this shortcoming – as, for example, in Stock and Watson (1999) – is to augment the sample using projected values obtained from univariate autoregressive models. Here, we adopt an essentially identical procedure: an AR(4) model (plus seasonal dummy variables) is estimated using observations from the beginning of the sample through the last of the τ observations in the window, and used to forecast the series for another twelve months. The resulting τ+12 observations are then decomposed into the m frequency components, and the τ^{th}observation on each component is used to produce the values of $D^1...D^m$ from this window. The $D^1...D^m$ column vectors produced in this way still (by construction) add

[12] Windows-based, and RATS, software implementing this partitioning of a given input column vector is available from the authors.

up precisely to X_k; they are still each the product of an entirely one-sided bandpass filter; and (since their values are now no longer close to the endpoint of each window) they produce quite satisfactory decompositions.[13]

4 Detecting and modeling frequency dependence in simulated data

In Section 3 above, existing approaches for detecting and modeling frequency-dependence were reviewed, and it was shown that the usual (two-sided) pre-filtering approaches to the detection of frequency dependence will yield misleading results in the presence of feedback. Finally, in Section 3.6 we proposed a one-sided extension to the Tan/Ashley approach for analyzing frequency dependence in the presence of feedback. In this section, we summarize the results from a small simulation study which provides evidence for the efficacy of this proposed methodology. This simulation study is intended to be suggestive rather than exhaustive; consideration is limited to two rather simple data generating processes; these are intended primarily to demonstrate that the procedure can in fact correctly detect the presence and form of frequency dependence even when feedback is present, and only partially to illustrate possible sources for the frequency dependence observed below in the relationship between inflation and unemployment in U.S. data.

The simulation results reported below address three questions relating to data-generating processes which feature feedback. First, in the presence of such feedback, does two-sided filtering actually lead to a spurious finding of frequency-dependence when none actually exists? Second, does the one-sided procedure proposed in Section 3.6 avoid such spurious findings? Finally, does the one-sided procedure correctly detect, and appropriately model, frequency-dependence when such dependence is present?

[13] We note that it is also necessary to detrend the X_t data in each window, since a somewhat persistent time series can appear quite trended in each of the sequence of windows, even though it is not trended overall. Thus, a linear trend is estimated over the $\tau + 12$ observations in each window, and subtracted from the X_k values prior to decomposing it into the m frequency components. After the decomposition is performed, observation τ's estimated trend value is then added back into observation τ of the lowest frequency band, D^1. In this way, $D^1 ... D^m$ still sum to X_k.

4.1 Spurious frequency dependence detection using two-sided filtering in the presence of feedback

The data-generating process considered here is a particular bivariate VAR, given by:

$$\begin{aligned} y_t &= \lambda_1 x_{t-1} + \lambda_2 y_{t-1} + \varepsilon_{y,t} \\ x_t &= \alpha_1 x_{t-1} + \alpha_2 y_{t-1} + \varepsilon_{x,t} \end{aligned} \tag{14}$$

where $\lambda_1 = 0.25$, $\lambda_2 = 0.55$, $\alpha_1 = 0.65$, and $\alpha_2 = 0.3$; qualitatively similar results were obtained using numerous bivariate VAR specifications, however. Since $\alpha_2 \neq 0$, this bivariate system exhibits feedback; since both equations are linear, there is no actual frequency dependence in these coefficients. 1000 bootstrap simulations were conducted. For each simulation, both the one-sided and two-sided approaches were used to test for the presence of frequency dependence across three frequency bands. The frequency bands used were set such that the lowest frequency band coresponded to fluctuations with period greater than 6, the medium frequency band corresponded to fluctuations with periods between 4.5 and 6, and the highest frequency band corresponded to fluctuations with period less than 4.5.[14]

In each simulation run, the series x_t was decomposed by frequency using both the one-sided and two-sided procedures, yielding $\left\{D_t^{1,1-sided}, D_t^{2,1-sided}, D_t^{3,1-sided}\right\}$ and $\left\{D_t^{1,2-sided}, D_t^{2,2-sided}, D_t^{3,2-sided}\right\}$, with $t = 1, ..., 300$.

Following this, y_t was regressed on y_{t-1} and $D_t^{1,k}, D_t^{2,k}$, and $D_t^{3,k}$ first for $k =$ 1-sided, and then for $k =$ 2-sided. In each case, an F-test testing equality of the coefficients on the three D_t variables was performed. The resultant p-value was then recorded for each simulation. Since there is in fact no frequency dependence in this linear model, the null hypothesis of equal coefficients on the three components D_t^1, D_t^2, and D_t^3 should be rejected (at the 5% level) in only about 5% of the cases.

Although the procedures differed only in the method of decomposition, the results were starkly different. When filtered using the two-sided methodology, the null of frequency dependence was rejected, at the 5% level, in nearly 40% of the cases (and for other specifications investigated, this rejection rate was even greater). Evidently, two-sided filtering can readily lead to a spurious

[14] Since Section 4 features artificial examples, this band structure (used throughout the section) was chosen arbitrarily. See the Appendix for an example of the relationship between frequencies and periods.

detection of frequency dependence in the presence of feedback; the issues we raise in Section 3.6 are not merely a theoretical detail. Conversely, when filtered using the one-sided methodology (and regarded as a test of frequency-dependence), the "size" of the one-sided procedure was correct: the null was rejected at the 5% level of significance in ca. 5% of the cases.

4.2 Detection and modeling of frequency dependence due to unmodeled Markov-switching

We now turn to our final question: does the one-sided procedure correctly detect, and appropriately model, frequency-dependence when such dependence is actually present? Two distinct data-generating processes are considered, each of which generates frequency dependence in the coefficients of a (mis-specified) linear model one might actually estimate. The generating mechanism examined in this section is a Markov-switching process; in this case, the frequency dependence in the coefficients of the approximating linear model arises because of unmodeled nonlinearity in the relationship. A second generating mechanism is considered in Section 4.3 below. There the frequency dependence in the coefficients of a linear model arises from unmodeled heterogeneity due to aggregation.

In this section we examine a Markov-switching process which is a bivariate VAR alternating between two regimes:

$$\begin{aligned} y_t &= Bx_{t-1} + \lambda y_{t-1} + \sigma\varepsilon_{y,t} \\ x_t &= Ax_{t-1} + \gamma y_{t-1} + S\varepsilon_{x,t} \end{aligned} \tag{15}$$

where A, B, and S are random variables whose values are regime-dependent: in regime 1, $(A, B, S) = (a_1, b_1, s_1)$, while in regime 2, $(A, B, S) = (a_2, b_2, s_2)$. The process switches between regime 1 and regime 2 according to a Markov process with switching probability q. If $\gamma > 0$, this system exhibits positive feedback.

By construction, within each regime the parameter B is a fixed constant. However, if the Markov-switching is unmodeled, i.e. if one estimates a (mis-specified) regression equation which fails to account for regime switching, then the coefficient on x_{t-1} in a model for y_t is frequency (and time) dependent unless $a_1 = a_2$ and $b_1 = b_2$. For example, suppose that $a_1 = 0.8$, $b_1 = 0.5$, and $s_1 = 0.5$, whilst $a_2 = 0.0$, $b_2 = -0.5$, and $s_2 = 1.0$. In this case, when the process is in regime 1, x_t is highly persistent and y_t is positively related to past x_t. In contrast, when the economy is in regime 2, x_t is not persistent and y_t is inversely related to x_t. This cross-regime coefficient disparity can generate substantial frequency dependence in the relationship between y_t and x_{t-1}. To see why, note that y_t will be positively related to x_{t-1} when x_t is in the "low-frequency" (persistent) regime, whereas y_t will be inversely related to x_{t-1} when x_t is in the "high-frequency" (non-persistent) regime. Over the course of the sample, the low-frequency variation in x_t will be dominated by periods during which x_t was in phase 1, and the high-frequency variation in x_t will be dominated by periods during which x_t was in phase 2.[15] Note that if the values of A, B and S were the same in both regimes, then the system would be an ordinary bivariate VAR whose coefficients do not exhibit frequency dependence.

$T = 300$ observations on this process were generated using the following parameter values:

Parameter	Regime 1	Regime 2
A	0.8	0.0
B	0.5	-0.5
λ	0.2	0.2
γ	0.3	0.3
S	0.5	1.0
σ	0.6	0.6
q	0.02	0.02

The series x_t was decomposed by frequency using the one-sided procedure, yielding $\{D_t^1, D_t^2, D_t^3\}$. In this case, the three frequency bands chosen were set such that the lowest frequency band coresponded to fluctuations with period greater than 6, the medium frequency band corresponded to fluctuations with periods between 4.5 and 6, and the highest frequency band corresponded to fluctuations with period less than 4.5.[16] The dependent variable y_t was then regressed on a constant, y_{t-1}, and D_t^1, D_t^2, and D_t^3.

Regression results, with and without the allowance for frequency dependence, were as follows

[15] Crudely speaking, one might think of the high-frequency part of x_t as the part of the time series of x_t which "survives" first-differencing; conversely,the low-frequency part of x_t is essentially its stochastic trend. During regime 2, the stochastic trend is near zero, and the first-difference of x_t is large in magnitude. Conversely, during regime 1, the stochastic trend diverges from zero, while the first-difference of x_t is generally small in magnitude. Thus, the low frequency part of x_t substantially differs from zero only during regime 1, while the high frequency part of x_t substantially differs from zero only during regime 2.

[16] See footnote 14.

(coefficient estimates appear below the coefficients, with t-statistics in parentheses):

$$
\begin{aligned}
y_t &= \underset{\underset{(3.41)}{-0.17}}{\widehat{\alpha}} + \underset{\underset{(0.96)}{0.04}}{\widehat{b}} x_{t-1} + \underset{\underset{(9.51)}{0.51}}{\widehat{\lambda}} y_{t-1} + u_t \\
y_t &= \underset{\underset{(-1.97)}{-0.09}}{\widehat{\alpha}} + \underset{\underset{(3.82)}{0.16}}{\widehat{b}_1} D^1_{t-1} + \underset{\underset{(-0.81)}{-0.16}}{\widehat{b}_2} D^2_{t-1} + \underset{\underset{(-6.15)}{-0.55}}{\widehat{b}_3} D^3_{t-1} + \underset{\underset{(10.94)}{0.52}}{\widehat{\lambda}} y_{t-1} + u_t
\end{aligned}
$$

The F-test of no frequency dependence (i.e., $H_0 : b_1 = b_2 = b_3$) = 24.8, with p-value = 0.000000. The pattern of frequency-dependence in the data is clearly captured by our procedure.

4.3 Detection and modeling of frequency dependence due to aggregation

The second data-generating process considered is a trivariate VAR:

$$
\begin{aligned}
y_t &= \lambda_1 z_{1,t-1} + \lambda_2 z_{2,t-1} + \lambda_3 y_{t-1} + \sigma \varepsilon_{y,t} \qquad (16) \\
z_{1,t} &= \rho_1 z_{1,t-1} + \gamma y_{t-1} + s \varepsilon_{x_1,t} \\
z_{2,t} &= \rho_2 z_{2,t-1} + \gamma y_{t-1} + \varepsilon_{x_2,t}
\end{aligned}
$$

If $\gamma > 0$, this system exhibits positive feedback. Suppose that the econometrician is unable to observe $z_{1,t}$ and $z_{2,t}$, but can only observe their *sum* z_t, defined as $(z_{1,t} + z_{2,t})$. Unless $\lambda_1 = \lambda_2$ or $\rho_1 = \rho_2$, such aggregation will induce frequency-dependence in the resultant bivariate VAR: the coefficient on z_{t-1} in a model for the $\{y_t, z_t\}$ process will be frequency-dependent. For example, suppose that $\rho_1 = 0.8$ and $\lambda_1 = 0.5$, whilst $\rho_2 = -0.1$ and $\lambda_2 = -0.5$. In this case, y_t is positively related to the persistent variable $z_{1,t}$, and inversely related to the noisy variable $z_{2,t}$. This implies that the relationship between y_t and z_{t-1} is frequency-dependent: y_t is positively related to low-frequency variations in z_{t-1} (which are dominated by $z_{1,t-1}$), and inversely related to high-frequency variations in z_{t-1} (which are dominated by $z_{2,t-1}$). Of course, the system is still misspecified if if λ_1 equals λ_2, but the coefficient on z_t does not in that case exhibit frequency-dependence.

$T = 300$ observations on this process were generated using the following parameter values:

Parameter	Value
λ_1	0.5
λ_2	-0.5
λ_3	0.2
σ	0.6
α_1	0.8
α_2	-0.1
s	0.5
γ	0.3

The series x_t was decomposed by frequency using the one-sided procedure, yielding $\{D_t^1, D_t^2, D_t^3\}$. The three frequency bands chosen were set such that the lowest frequency band coresponded to fluctuations with period greater than 6, the medium frequency band corresponded to fluctuations with periods between 4.5 and 6, and the highest frequency band corresponded to fluctuations with period less than 4.5.;[17] then y_t was regressed on a constant, y_{t-1}, and D_t^1, D_t^2, and D_t^3.

Regression results, with and without the allowance for frequency dependence, were as follows (coefficient estimates appear below the coefficients, with t-statistics in parentheses):

$$
\begin{aligned}
y_t &= \underset{\substack{-0.21 \\ (-4.06)}}{\widehat{\alpha}} + \underset{\substack{-0.02 \\ (-0.53)}}{\widehat{b}} x_{t-1} + \underset{\substack{0.46 \\ (8.48)}}{\widehat{\lambda}} y_{t-1} + u_t \\
y_t &= \underset{\substack{-0.11 \\ (-2.18)}}{\widehat{\alpha}} + \underset{\substack{0.14 \\ (4.02)}}{\widehat{b}_1} D_{t-1}^1 + \underset{\substack{0.23 \\ (1.51)}}{\widehat{b}_2} D_{t-1}^2 + \underset{\substack{-0.41 \\ (-6.19)}}{\widehat{b}_3} D_{t-1}^3 + \underset{\substack{0.42 \\ (8.42)}}{\widehat{\lambda}} y_{t-1} + u_t
\end{aligned}
$$

The F-test of no frequency dependence (i.e., $H_0 : b_1 = b_2 = b_3$) = 24.7, with p-value = 0.000000. Again, the pattern of frequency-dependence in the data is clearly captured by our procedure.

One final remark on the simulation results for both of these processes: we find that the presence of unmodeled frequency-dependence in the relationship frequently leads to an initial linear model for y_t which includes multiple lags of x_t, even though only x_{t-1} is actually influencing y_t; furthermore the estimate of the coefficient on x_{t-1} is frequently statistically insignificant. This latter observation is not surprising, since the OLS coefficient estimate on x_{t-1} is in both cases an admixture of two different relationships, a positive one at low frequencies, and a negative one at high frequencies.

[17] See footnote 14.

This finding suggests that analysts may be missing some significant empirical relationships because of unmodeled frquency dependence.

We conclude that the procedure described in Section 3.6 is both necessary and effective in the presence of feedback.

5 Phillips Curve Estimation Results

5.1 Regression model specification

From (2), a standard Phillips curve specification is of the form

$$\pi_t = \alpha + \beta un_t + \sum_{j=1}^{12} \delta_j \pi_{t-j} + \theta Z_t + \varepsilon_t \tag{17}$$

where Z_t includes seasonal dummies and a measure of the change in the relative price of energy, Oil_t (as in, e.g., Staiger, Stock and Watson 2001). We consider the period 1980:1-2003:12. As Benati and Kapetanios (2003) find compelling evidence for the existence of structural breaks in the US CPI inflation process, we estimate (17) assuming that inflation had one structural break (in mean) in early 1990, using the date these researchers identify: 1990:4.[18] Since the behavior Oil_t appears to markedly change in character during this interval, we defined two dummy variables (Oil_t^1 and Oil_t^2), allowing the coefficients on these regressors to differ in periods 1980:1-1986:01 and 1986:02-2001:12.[19] The measure of inflation used in constructing π_t is the growth rate of non-seasonally-adjusted CPI-U-RS;[20] un_t is the non-seasonally-adjusted total civilian unemployment rate.

[18] Benati and Kapetanios (2003) also find a break in 1981:4. However, we don't find subtantial evidence for this break in our data, likely because this break occurs so early in our sample.

[19] The energy series used was "energy commodities," which is then divided by the CPI-U-RS. The coefficients on Oil^1 and Oil^2 are allowed to vary over these subperiods because the time-series properties of Oil_t (in particular, its variance) display two distinct regimes over the sample. We therefore did not want to constrain the relationship between inflation and the relative price of energy to be the same over the entire sample.

[20] The Bureau of Labor Statistics (BLS) has made numerous improvements to the CPI over the past quarter-century. For example, in 1983 the BLS adopted a rental-equivalence approach to the measurement of homeownership costs in the CPI-U; other methodological improvements have subsequently occurred. While these improvements make the present and future CPI more accurate, *historical* price index series have not been adjusted to consistently reflect all of these improvements. The CPI-U-RS (or CPI-U "Research Series," described in Stewart and Reed (1999)) comes closest to this ideal; it consistently corrects the CPI-U for all changes in methodology from 1978 onwards. Researchers seeking a (mostly) consistent series from 1967 onwards can append the CPI-U-RS to the CPI-U-X1 series, a series which at least incorporates rental-

Mis-specification in Phillips Curve regressions

Using the one-sided filtering methodology described in Section 3.6 above, the series un_t was decomposed into frequency bands $un^1...un^k$, where the number of bands, and the frequencies in each band, are selected as described below. Equation (17) was then re-estimated using OLS in the form:[21]

$$\pi_t = \alpha + \sum_{j=1}^{k} \beta_j un_t^j + \sum_{j=1}^{12} \delta_j \pi_{t-j} + \theta Z_t + \varepsilon_t \tag{18}$$

We performed two alternative tests of frequency dependence based on equation (18): one using "calendar-based" frequency bands, and one in which the frequency bands were data-selected using the regression tree specification search procedure described in Section 3.5 above. The calendar-based approach has the advantage that the sampling distribution of the F-statistic for testing the equality of $\beta_1...\beta_k$ is not distorted by a specification search, so that it is not necessary to estimate this sampling distribution using bootstrap simulations. On the other hand, the calendar-based bands are necessarily somewhat *ad hoc*.

If the chosen calendar-based bands are consistent with the actual pattern of frequency dependence present in the data, then this procedure will have high power to detect that pattern. If not, then the calendar-based test could have relatively low power, even though appropriately accounting for the specification search in the procedure in which bands are data-selected substantially lowers the apparent power of that procedure. Thus, one might unnecessarily fail to uncover an existing pattern of frequency dependence in a particular regression coefficient through a maladroit selection of a calendar-based frequency band structure. Moreover, even if one does still detect frequency dependence in spite of such a maladroit choice, the pattern of frequency dependence thus observed will surely be distored to some degree. Consequently, unless one has a specific and strongly-held prior opinion as to what frequency band structure is consistent with any actual frequency dependence, then the specification search procedure described in Section 3.5 will be more appropriate.[22] Here, for illustrative purposes, results are given using both approaches.

equivalence homeownership costs. Note that other researchers, notably Crone, Nakamura and Voith (2001) suggest that still other adjustments may be worthwhile.

[21] A residual outlier was detected in February of 1986; consequently, the regression was run with and without a dummy variable corresponding to this outlier. Upon inclusion of this dummy, the Jarque-Bera test no longer rejects normality of the residuals; however, conclusions regarding frequency-dependence are identical. Additional lags of the unemployment rate were not significant; it is possible that the additional lags found by other researchers are a by-product of ignoring frequency-dependence in this relationship, as suggested by the simulation study in Section 4.

[22] And, of course, basing one's calendar-based bands upon the results of a specification search and pretending it didn't take place is self-delusional.

Mis-specification in Phillips Curve regressions

Since "business cycle frequencies" are ordinarily taken to comprise fluctuations with periods between one and five years, our initial impulse for a calendar-based frequency band structure was to decompose the unemployment rate data into three bands corresponding to periods of 2 to 18 months, periods of 18 to 72 months, and periods in excess of 72 months. However, setting τ to a number larger than 60 – corresponding to using more than five years of data for each window – seemed unreasonable, given that we are using post-1975 data on un_t. Consequently, each un_t^j observation is based on the data $un_{t-59}...un_t$. An implication of this filtering window is that fluctuations with periods larger than 72 months cannot be distinguished from fluctuations with periods of 72 months. For that reason the low frequency band in our calendar-based model was modified to include the frequency corresponding to a period of 72 months.[23]

In the second approach we allowed the data itself to choose the frequency band structures, using the regression tree procedure described in Section 3.5 above. Since this search procedure substantially distorts the sampling distribution of the F-statistic for testing the equality of the coefficients on the k bands, the actual sampling distribution for this statistic was estimated using 2000 bootstrap simulations. In particular, we simulated T observations on π_t from our estimate of equation (18), drawing errors (with replacement) from the residuals of this equation. For each set of T observations we repeated the search procedure, re-estimated equation (18), and stored the resulting p-value corresponding to the F-statistic for testing the equality of $\beta_1...\beta_k$. The empirical p-value at which the null hypothesis of no frequency dependence can be rejected was then calculated as the fraction of these p-values which exceed the value obtained when equation (18) is estimated using the actual sample data.[24]

[23] The Appendix lists the frequencies and periods associated with a 72-observation rolling window. Recall from the discussion at the close of Section 3.6 that the sixty months of actual data ($un_{t-59}...un_t$) are augmented by twelve months of projected data, so that the filtered value for each frequency band is twelve months prior to the end of a 72-month filtering window.

[24] We thus bootstrap the distribution of the p-values rather than that of the F-statistic values. This was necessary since the procedure potentially searches over both one-band, two-band, and three-band specifications, as well as considering the composition of the bands. Thus, the null hypothesis sometimes involved one parameter restriction, and sometimes involved two parameter restrictions, rendering the F-statistics themselves non-comparable.

5.2 Empirical results

Estimating the standard Phillips curve specification of equation (17) over the sample period 1980:1-2003:12 yielded:[25]

$$\begin{aligned}\pi_t \;=\;& \underset{\underset{(-0.72)}{-0.354}}{\alpha} + \underset{\underset{(-0.92)}{-0.05}}{\beta}\, un_t + \underset{F-test:\, p=0.000}{\sum_{j=1}^{12} \delta_j \pi_{t-j}} + \underset{\underset{(8.14)}{0.10}}{\theta_1}\, Oil_t^1 + \underset{\underset{(-2.30)}{-0.03}}{\theta_2}\, Oil_{t-1}^1 \\ &+ \underset{\underset{(16.22)}{0.04}}{\theta_3}\, Oil_t^2 + \underset{\underset{(-2.13)}{-0.01}}{\theta_4}\, Oil + \underset{F-test:\, p=0.000}{\sum_{i=1}^{11} \theta_{i+5} month_t^i} + \underset{\underset{(-3.68)}{-0.76}}{\theta_{17}}\, BK_t + \varepsilon_t \qquad (19)\end{aligned}$$

For selected coefficients, we present coefficient estimates, with their estimated t-statistics in parentheses; for others, we simply present the F-test of the null hypothesis that all the coefficients in the distributed lag structure are zero. The variables Oil_t^1 and Oil_t^2, $month_t^1 ... month_t^{11}$, and BK_t are the relative price of energy, seasonal dummy variables, and inflation break dummy variable described in Section 5.1. Unlike many researchers (e.g., Gordon 1997; Brayton, Roberts and Williams 1999), we find that longer lags in π_t are not necessary to account for serial correlation. This is partly due to our estimation period – we avoid the problematic 1970s – and partly due to the inclusion of the inflation-break dummy variable. Estimating an analogous model for un_t, we find evidence for significant feedback in the $\pi - un$ relationship; in particular, the null hypothesis that the lagged inflation rate π_{t-1} is unrelated to movements in un_t is rejected at the 2% level. Consequently, it is necessary to use the one-sided filtering methodology described in Section 3.6 above.

Note that the coefficient $\widehat{\beta}^{OLS}$ is not statistically significant. The estimation of a standard linear formulation of the Phillips curve over this time period suggests that, in fact, there *is* no Phillips curve. As the simulation results in Sections 4.2 and 4.3 suggest, however, a statistically insignificant β estimate does not necessarily imply the lack of a statistically significant Phillips curve relationship since any frequency dependence in this relationship renders $\widehat{\beta}^{OLS}$ an inconsistent estimate.

Table 1 presents the coefficients of interest for the analysis of frequency-dependence in the Phillips curve equation. Three Phillips curve specifications are considered:

[25] Here and following, we quote results pertaining to the specification which included the inflation-regime dummy, BK_t, using an estimation procedure which produced White heteroskedasticity-corrected standard errors. Neither of these choices is consequential regarding inference.

- the "classical" Phillips curve (i.e., equation (17) above, which ignores frequency dependence)
- the "*a priori* calendar-based bands" model (which partitions the unemployment rate into bands with periods less than 18 months, between 18 and 72 months, and greater than or equal to 72 months),

and

- the "data-selected" model (which turns out to partition the unemployment rate into two components: fluctuations with periods greater than 9 months, and fluctuations with periods less or equal to 9 months).

Table 1 quotes both estimated t-statistics and estimated standard errors for the coefficient estimates. For the calendar-based model, we also report the p-value of the F-test whose null hypothesis is that coefficients β_1, β_2, and β_3 are all equal. And for the data-selected model we also report the bootstrapped p-value of the F-test whose null hypothesis is that coefficients β_1 and β_2 are equal; this bootstrapping procedure accounts for the specification search involved in obtaining this frequency band formulation. In either case, a rejection of the null hypothesis indicates statistically significant frequency dependence.

Table 1: Frequency Dependence in the Phillips Curve $\pi - un$ Relationship

	Classical	A priori calendar-based bands	Data-selected bands
	un_t $\widehat{\beta}^{OLS} = -0.05 \pm 0.06$ (-0.92)	un_t^1 ($\geq$ 72 months) $\widehat{\beta}_1^{OLS} = -0.04 \pm 0.06$ (-0.67) un_t^2 (18 to 72 months) $\widehat{\beta}_2^{OLS} = -0.03 \pm 0.32$ (-0.09) un_t^3 ($<$ 18 months) $\widehat{\beta}_3^{OLS} = -1.88 \pm 0.92$ (-2.05)	un_t^1 ($>$ 9 months) $\widehat{\beta}_1^{OLS} = 0.003 \pm 0.06$ (0.04) un_t^2 (2 to 9 months) $\widehat{\beta}_2^{OLS} = -4.16 \pm 1.42$ (-2.92)
F−test p-value	–	0.13[1]	0.050[2]

1. Asymptotic p-value, H_0: $\beta_1 = \beta_2 = \beta_3$
2. Bootstrapped p-value, H_0: $\beta_1 = \beta_2$

Three things are worth noting. First, the model using the data-selected frequency bands rejects the null of frequency-dependence at the 5% level: the bootstrapped p-value for the test $H_0 : \beta_1 = \beta_2$ is 0.05. Second, decomposing the unemployment rate using a priori calendar-based bands does not yield statistically-significant evidence for frequency-dependence. This negative result highlights the importance of avoiding strong priors and allowing the data to speak to the nature, and form, of the frequency dependence. These two results are robust to modifications in a variety of modeling choices.

Third, we find that the inflation impact of higher-frequency fluctuations in the unemployment rate is economically, as well as statistically, significant. To quantify and display this impact, we constructed the time series $impact_t$ as follows:

$$impact_t := \widehat{\beta}_2^{OLS} * un_t^2$$

$Impact_t$ quantifies the estimated impact of fluctuations in un_t^2 on the inflation rate. Since un_t^2 is the high-frequency component of un_t, we plot the smoothed absolute value of $impact_t$ against time in Figure 2:

Figure 2: Smoothed estimates of impact of high-frequency fluctuations in unemployment rate upon inflation

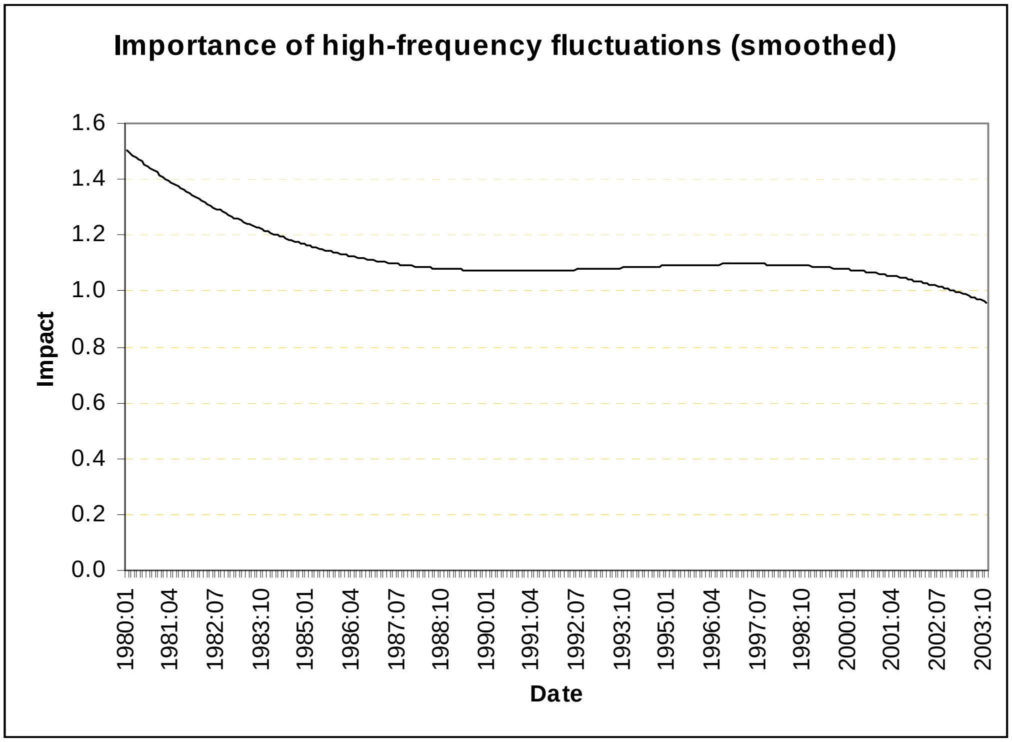


Figure 2 indicates that high-frequency fluctuations in the unemployment rate altered the inflation rate by approximately 1.5% in the early 1980's, and this impact declined to something less than 1% by the end of our sample. In contrast, lower-frequency fluctuations in the unemployment rate had no detectable impact upon inflation.

The existence of this frequency dependence indicates that much of the Phillips curve literature suffers from mis-specification: since the coefficient on un_t in a standard Phillips curve model is frequency dependent, estimates of this coefficient previously reported in the literature are actually an admixture of several different coefficients. In particular, the data indicate that fluctuations in unemployment that persist less than or equal to approximately 9 months are significantly associated with a contemporaneous fluctuation (of opposite sign) in inflation. In contrast, fluctuations in unemployment which persist longer than about 9 months are not significantly associated with contemporaneous fluctuations in inflation. These results are quite consistent with the Friedman-Phelps formulation: one might interpret transitory un_t fluctuations, i.e. those with periods less than about 9 months, as deviations from the natural rate (thus negatively associated with contemporaneous inflation); and more persistent un_t fluctuations (with periods larger than about 9 months) might be interpreted as movements in the natural rate – with the implication that such unemployment fluctuations are not associated with significant inflation co-movements.

In summary, then, there *is* a Phillips curve relation – but it applies only to unemployment fluctuations with periods less than or equal to approximately 9 months. Consequently, econometric formulations of this relationship which fail to distinguish unemployment fluctuations within this range from those outside it are mis-specified. This may help explain the apparent instability of estimated Phillips curve models across disparate time periods; for example, the Phillips curve will appear to be non-existent during periods in which un_t fluctuations are quite persistent.

6 Conclusion

This paper makes two contributions. First, we present new econometric methodology which allows one to consistently decompose a regression parameter across frequency bands, even when this regressor is in a feedback relationship with the dependent variable in the model. This technique is easy to apply and is applicable to a wide range of macroeconomic relationships.[26] We also demonstrate that two-sided filtering leads to inconsistent parameter estimates and yields unreliable inferences about the existence of frequency-dependence when feedback is present in the relationship.

The second contribution of this paper is the application of this new technique to a standard Phillips curve model using monthly US data from 1980-2003. Assuming that the relationship is not frequency dependent, the estimate of the coefficient characterizing this relationship is essentially zero. Allowing for the possibility of frequency dependence in this relationship, however, we find that such frequency dependence is a significant feature of this relationship. In particular, our

[26] Implementing RATS and FORTRAN code are available from the authors. Both of these programs use 1-sided filtering to decompose a given time series into components consisting of variation only over specified frequency bands; as noted in Section 3, these components are only moderately correlated, and sum precisely to the input time series.

results show that there is a significant negative relationship for high-frequency fluctuations in the unemployment rate –fluctuations whose periods range less than or equal to about 9 months – and an insignificant relationship for unemployment fluctuations outside this frequency range. A standard hypothesis test – bootstrapped to account for the specification search used in obtaining the BSIC-minimizing band structure – confirms that this pattern is significant at the 5% level. Our results in Figure 2 displaying the economic impact of these high-frequency fluctuations in the unemployment rate upon inflation show that this impact, which is on the order of $1 - 1\frac{1}{2}\%$, is far from trivial.

What do these results mean? We draw three conclusions. First, our finding of statistically significant frequency dependence in this relationship implies that nearly all previously estimated Phillips curve coefficients are an admixture of two different frequency-specific coefficients – one negative and the other zero. Thus, one implication of our results is that the apparent Phillips curve relationship can be expected to weaken or disappear in time periods when the unemployment rate fluctuates very smoothly.

Second, our results are supportive of the Friedman-Phelps theory. Fluctuations in the unemployment rate whose period is less than around 9 months have an inverse relationship with inflation. In contrast, fluctuations in the unemployment rate which persist for more than about 9 months have no relationship with inflation; the Friedman-Phelps theory would identify these with fluctuations in the natural rate.

Third, our work poses interesting challenges for forecasting and policy. The standard Phillips curve relationship has often been viewed as at least useful for the purposes of forecasting; yet our results indicate that only the high frequency components of unemployment rate fluctuations are related to the inflation rate. This result strongly suggests decomposing the unemployment rate into its various frequency components for use as an input to an empirical Phillips curve model for forecasting use. Furthermore, our work has implications for Taylor-type monetary policy rules; in particular, it implies that only transitory fluctuations in the unemployment rate impact inflation rates, suggesting that these linear rules need to be re-thought.

7 Appendix: Frequencies and periods associated with a 72-month rolling filtering window

The Table below indicates explicitly which frequencies (and periods, in months) will correspond to rows of the A matrix discussed in Section 3 with a rolling filtering window 72 months in length. A sinusoidal fluctuation in x_t with period equal to one of those listed here will appear entirely in the filtered series (D_t^j) containing that period; all other fluctuations will, to some degree, "leak" into the filtered series corresponding to adjacent frequency bands. Passband filters with a smaller degree of leakage can be formulated (e.g., Baxter and King (1999)), but do not yield filtered components which add up to the unfiltered series value.

allowed frequency	allowed period	allowed frequency	allowed period
0.014	72.00	0.264	3.79
0.028	36.00	0.278	3.60
0.042	24.00	0.292	3.43
0.056	18.00	0.306	3.27
0.069	14.40	0.319	3.13
0.083	12.00	0.333	3.00
0.097	10.29	0.347	2.88
0.111	9.00	0.361	2.77
0.125	8.00	0.375	2.67
0.139	7.20	0.389	2.57
0.153	6.55	0.403	2.48
0.167	6.00	0.417	2.40
0.181	5.54	0.431	2.32
0.194	5.14	0.444	2.25
0.208	4.80	0.458	2.18
0.222	4.50	0.472	2.12
0.236	4.24	0.486	2.06
0.250	4.00	0.500	2.00

8 References

Ashley, Richard. (1984) "A Simple Test for Regression Parameter Instability," Economic Inquiry 22.2, 253-267.

Ashley, Richard, and Randal Verbrugge (2001) "To difference or not to difference: a Monte Carlo investigation of spurious regressions in VARs." Mimeo, Virginia Tech.

Bai, Jushan (1997) "Estimating multiple breaks one at a time." Econometric Theory 13, 315-352.

Bai, Jushan, and Pierre Perron (1998) "Estimating and testing linear models with multiple structural changes." Econometrica 66.1, 47-78.

Bai, Jushan, and Pierre Perron (2003) "Computation and analysis of multiple structural-change models." Journal of Applied Econometrics, forthcoming.

Ball, Robert and N. Gregory Mankiw (2002) "The NAIRU in theory and practice." NBER Working Paper No. 8940.

Baxter, Marianne (1994) "Real exchange rates and real interest differentials: have we missed the business cycle relationship?" Journal of Monetary Economics 33.1, 5-37.

Baxter, Marianne, and Robert King (1999) "Measuring business cycles: approximate band-pass filters for economic time series." The Review of Economics and Statistics 81, 575-593.

Brayton, Flint, John Roberts and John C.Williams (1999) "What's happened to the Phillips curve?" Mimeo, Federal Reserve Board.

Breiman, Leo, Jerome Friedman, Richard Olshen, and Charles Stone (1984) Classification and Regression Trees. Wadsworth.

Brown, R.L., J. Durbin and J.M. Evans (1975) "Techniques for testing the constancy of regression relationships over time." Journal of the Royal Statistical Society Series B, 37,149-192.

Chow, Gregory C. (1960) "Tests of Equality between Sets of Coefficients in Two Linear Regressions." Econometrica 28, 591-605.

Christiano, Lawrence J., and Terry Fitzgerald (2003) "The band-pass filter." International Economic Review 44.2, 435-465.

Cogley, Timothy and Thomas Sargent (2001) "Evolving post-WWII US inflation dynamics." NBER Macroeconomics Annual 2001.

Corbae, Dean, Sam Ouliaris, and Peter C.B. Phillips (2002) "Band spectral regression with trending data." Econometrica 70, 1067-1109.

Crone, Theodore, Leonard Nakamura and Richard Voith (2001) "Measuring American rents: a revisionist history." Manuscript, Federal Reserve Bank of Philadelphia.

Debelle, G. and J. Vickery (1997) "Is the Phillips curve a curve? Some evidence and implications for Australia." Reserve Bank of Australia Research Discussion Paper No. 9706.

Diebold, Francis, Lee Ohanian and Jeremy Berkowitz (1998) "Dynamic Equilibrium Economies: A Framework for Comparing Models and Data," Review of Economic Studies 65.3, 433-51.

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

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

Fisher, Irving (1926) "A statistical relationship between unemployment and price changes." International Labor Review 13, 785-92.

Friedman, Milton (1968) "The role of monetary policy." American Economic Review 58.1, 1-17.

Fuhrer, Jeffrey (1995) "The Phillips curve is alive and well." New England Economic Review of the Federal Reserve Bank of Boston, March/April, 41-56.

Geweke, John (1982) "Measurement of Linear Dependence and Feedback Between Multiple Time Series." Journal of the American Statistical Association 77, 304-314.

Gordon, Robert (1982) "Inflation, flexible exchange rates, and the natural rate of unemployment." in Workers, Jobs and Inflation, Martin Baily (ed.), Washington, DC: Brookings Institution, 88-152.

Gordon, Robert (1990) “US inflation, labor’s share, and the natural rate of unemployment.” in Economics of Wage Determination, Heinz Konig (ed.), Berlin: Springer-Verlag, 1-34.

Gordon, Robert (1997) “The time-varying NAIRU and its implications for economic policy.” Journal of Economic Perspectives 11, 11-32.

Gordon, Robert (1998) “Foudnations of the Goldilocks economy: supply shocks and the time-varying NAIRU.” Brookings Papers on Economic Activity , 297-333.

Gruen, David, Adrian Pagan and Christopher Thompson (1999) “The Phillips curve in Australia.” Reserve Bank of Australia Discussion Paper No. 9901.

Hall, Robert (1999) “Comment on ‘Rethinking the role of the NAIRU in monetary policy: implications of model formulation and uncertainty.’ ” in (J.B. Taylor, ed.), Monetary Policy Rules, NBER Conference Report, Chicago: University of Chicago Press.

Hannan, E. (1963) “Regression for Time Series” in M. Rosenblatt, ed. Time Series Analysis, John Wiley: New York, 14-37.

Hodrick, Robert J., and Edward Prescott, (1997) “Postwar US business cycles: an empirical investigation.” Journal of Money, Credit and Banking 29, 1-16.

Ireland, Peter (1999) “Does the time-consistency problem explain the behavior of inflation in the United States?” Journal of Monetary Economics 44.2, 279-291.

King, Robert, James Stock and Mark Watson (1995) “Temporal instability of the unemployment-inflation relationship.” Economic Perspectives of the Federal Reserve Bank of Chicago 19.3, 2-12.

Liu, J., S. Wu, and J.V. Zidek (1997) “On segmented multivariate regressions.” Statistica Sinica 7, 497-525.

Phelps, Edmund (1967) “Phillips curves, expectations of inflation, and optimal unemployment over time.” Economica 2.3, 22-44.

Phelps, Edmund (1968) “Money-wage dynamics and labor market equilibrium.” Journal of Political Economy 76.2, 678-711.

Phillips, Peter (1991) "Spectral Regression for Cointegrated Time Series." in (Barnett, W.A., J. Powell, G. E. Tauchen, eds.) Nonparametric and Semiparametric Methods in Economics and Statistics. Cambridge University Press: Cambridge.

Staiger, Douglas, James Stock and Mark Watson (1997) "The NAIRU, unemployment, and monetary policy." Journal of Economic Perspectives 11, 33-51.

Staiger, Douglas, James Stock and Mark Watson (2001) "Prices, wages and the US NAIRU in the 1990s." NBER Working Paper No. 8320.

Stewart, Kenneth, and Stephen Reed (1999) "Consumer Price Index research series using current methods, 1978-98." Monthly Labor Review, June, 29-38.

Stock, James, and Mark Watson (1999) "Forecasting Inflation." Journal of Monetary Economics 44, 293-335.

Stock, James, and Mark Watson (1999) "Business Cycle Fluctuations in U.S. Macroeconomic Time Series," in (John Taylor and Michael Woodford, eds), Handbook of Macroeconomics. Amsterdam: Elsevier, 3-64.

Tan, Hui-Boon, and Richard Ashley (1999a) "On the Inherent Nonlinearity of Frequency Dependent Time Series Relationships." in (P. Rothman, ed.) Nonlinear Time Series Analysis of Economic and Financial Data, Kluwer Academic Publishers: Norwell, 129-142.

Tan, Hui-Boon, and Richard Ashley (1999b). "Detection and Modeling of Regression Parameter Variation Across Frequencies With an Application to Testing the Permanent Income Hypothesis." Macroeconomic Dynamics 3, 69-83.

Thoma (1992) "The Effects of Inside and Outside Money on Industrial Product Across Spectral Frequency Bands." The Review of Economics and Statistics 74, 737-741.

Thoma (1994) "The Effects of Money Growth on Inflation Rates across Spectral Frequency Bands." Journal of Money, Credit, and Banking 26.2, 218-230.

Yao, Y.Q. (1988) "Estimating the number of change-points via Schwarz' criterion." Statistics and Probability Letters 6, 181-189.