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

Richard Ashley
Department of Economics
Virginia Tech
Blacksburg, VA 24061-0316
Email: ashleyr@vt.edu

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

This version: March 27, 2007

JEL Classification Codes: E00, E31, C22, C32, E5

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

Acknowledgement: We thank Marianne Baxter, Larry Christiano, Walter Enders, Jonas Fisher, Josh Gallin, Andy Levin, Athanasios Orphanides, Steve Reed, Jeremy Rudd, John Williams, and seminar participants at the 2003 North American Summer Meetings of the Econometric Society, at the 2004 Midwest Macroeconomics Conference, at the Board of Governors, and at Virginia Tech. 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.

Abstract

The Phillips curve has long been a focus 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 standard Phillips Curve regression specification 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 used in previous frequency domain studies otherwise confounds the analysis, leading to inconsistent parameter estimates. Using one-sided filtering to allow for observed feedback in the relationship, we find statistically significant frequency dependence. In particular, using monthly data from 1984:1-2003:12, we find an economically and statistically significant inverse relationship between inflation and unemployment for high-frequency unemployment rate fluctuations – with periods less than about one year – but no evidence for an effect of lower frequency unemployment rate fluctuations. In contrast, a model ignoring frequency dependence finds no statistically significant relationship at all between inflation and unemployment rates 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.[1] This relationship is central to contemporary monetary policy: one cannot hope to appropriately conduct such policy without an adequate understanding of short-run inflation dynamics. This relationship is also central to "New Keynesian" macroeconomic modeling: a Phillips curve relation forms one of the three key equations in most New Keynesian macro models. Yet despite its importance to macroeconomics, the nature of the Phillips curve relationship remains strongly contested.

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.[2] Even as late as the mid-1990s, some observers (Fuhrer 1995, Gordon 1997) 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 (see, e.g., Beaudry and Doyle, 2000); in post-1980 data, it is difficult to detect a link between the unemployment rate and inflation. This prompted many (e.g., Eisner, 1998) to argue that the Philips curve relationship was "dead," while others (e.g., Brayton, Roberts and Williams, 1999; Staiger, Stock and Watson, 2001) attempted 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

[1] 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).

[2] 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 provide an explanation for this phenomenon.

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; thus, 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 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[3] – indeed, changes in the natural rate are often blamed when a particular Phillips Curve specification appears to be breaking down. Although there is no reason to expect that this natural rate is a fixed constant, much previous research has made this assumption. In contrast, a number of recent studies have implicitly assumed that the natural rate is an I(1) process, estimating the relationship in differences. 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 relationship 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

[3] Not all "Phillips curve" specifications relate inflation to the unemployment rate; others relate inflation to "output gaps" (where, however, the identical issue arises) or – particularly in the New Keynesian tradition – to marginal costs. In this paper, we don't address these other specifications directly; however, in a companion paper (Ashley and Verbrugge 2005b) we explore New Keynesian Phillips curve formulations. We find significant frequency dependence in those formulations as well.

counterfactual, assumptions about the manner in which the natural rate evolves over time. Although pre-filtering approaches don't suffer from this particular criticism, they are *ad hoc*; and furthermore there is still no guarantee that splines or low-pass filtering accurately recover the time variation in the natural rate either. Finally, we demonstrate below that any two-sided filtering of a Phillips curve relationship will induce parameter estimation inconsistency in this context if there is feedback from inflation to the unemployment rate.

Nevertheless, 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, and researchers may well draw erroneous conclusions.

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, our approach does not require any *a priori* imposition of assumptions regarding the relevant frequency ranges. And because 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. Note that we need not take a stand concerning the various theories underlying the Phillips curve's dynamics – worker misperceptions, cost-push inflation, Lucas supply curve stories, and so on. Rather, the question we pose is: does frequency dependence exist in the typical

[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.

empirical Phillips curve specification?

We show below that all presently-available methods for detecting and modeling frequency dependence fail when 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, as detailed in Section 3.6 below, 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. This suggests caution in interpreting the coefficients from *any* dynamic regression which involves two-sided filtered data.

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 standard 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, the data indicate that there is a statistically significant inverse relationship betwen inflation and unemployment – but this significant relationship is restricted to rather high-frequency unemployment rate fluctuations, i.e. fluctuations with periods less than a year. The implied natural rate is far from smooth, calling into question the "smoothness" criterion that has often been imposed in empirical work (e.g., Eller and Gordon, 2002).

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

$$\pi_t = \pi_t^e + \beta \left(un_t - un_t^N \right) + \varepsilon_t \tag{1}$$

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 is 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 two unobserved components in each the above relationships: π_t^e and un_t^N.

First consider the treatment of expected inflation. 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 2002). 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 generally led to regression models of the following form:

$$\pi_t = \beta \left(un_t - un_t^N \right) + \sum_{j=1}^{m} \delta_j \pi_{t-j} + \theta Z_t + \varepsilon_t \tag{2}$$

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 and the relative price of imports. Shocks to such variables arguably create positive

[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 "meaningful" 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.

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; but each of these has severe shortcomings. 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:

$$\pi_t = \widetilde{\alpha} + \beta un_t + \sum_{j=1}^{m} \delta_j \pi_{t-j} + \theta Z_t + \varepsilon_t \tag{3}$$

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. Of course, it is unlikely that the natural rate is constant for extended periods of time; and below, we present evidence suggesting that this is far from true. (Two other simple approaches which allow un_t^N to vary slowly over time are to estimate un_t^N using weighted sample means, or to estimate (3) repeatedly over rolling sample ranges; see Williams 2004.)

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 (3), from which a time-varying natural rate estimate can be recovered. A variant of this method (e.g., Beaudry and Doyle 2000, Ball and Mankiw, 2002, Rudd and Whelan, 2005) involves identifying a filtered version of the unemployment rate with the natural rate for use in equation (2). We argue in Section 3.6 below that, in addition to being subject to considerable measurement error, such two-sided pre-filtering approaches must in this context lead to inconsistent parameter estimates.

An alternative method uses the Kalman filter to estimate the natural rate as an unobserved component; see, e.g., Staiger, Stock and Watson (1997, 2002), Gordon (1998), Gruen, Pagan and Thompson (1999), and Williams (2004). Typically, the natural rate is modeled as a unit root process in this framework, yielding the two-equation system:

$$\begin{aligned} \pi_t &= \alpha + \beta\left(un_t - un_t^N\right) + \sum_{j=1}^{m} \delta_j \pi_{t-j} + \theta Z_t + \varepsilon_t \qquad (4) \\ un_t^N &= un_{t-1}^N + \upsilon_t \end{aligned}$$

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 estimated ... very imprecisely (see Laubach and Williams, 2003).

In practice, these methodologies have tended to estimate a natural rate which closely tracks the univariate trend in the unemployment rate – e.g., see Brayton, Roberts and Williams (1999), Staiger, Stock and Watson (2002), or Williams (2004). But this does not necessarily imply accurate tracking of the natural rate dynamics. Furthermore – as noted above – we demonstrate below that if the relationship between inflation and $\left(un_t - un_t^N\right)$ is itself frequency dependent, any estimate of un_t^N deriving from a two-sided filter will lead to inconsistent OLS estimates of β, even if the identifying assumptions on the dynamics are correct.[6]

This paper presents a new approach to the specification of the Phillips curve relationship. We begin with a standard Phillips curve relationship specification as embodied in equation (3), which includes exogenous variables Z_t. We account for variation in the natural rate by allowing the coefficient β to vary across frequencies. This approach frees us from making strong assumptions regarding the natural rate data generating process. 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.

[6] Orphanides and Williams (2002) argue that it is highly unrealistic to assume knowledge of the true natural rate data generating process. Making strong assumptions on this DGP is concomitantly undesirable.

3 Methodology

3.1 Characterizing frequency dependence

In Sections 3.1–3.3 we explain what frequency dependence is, what it is not, and why it makes a difference. Sections 3.4 and 3.5 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 3.6 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. Section 3.7 addresses the issue of frequency band specification: how to select the number of frequency bands to consider, and the particular set of frequencies to be included in each band.

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:

$$c_t = \gamma_0 + \gamma_1 y_{t-1} + \gamma_2 y_{t-2} + \gamma_3 c_{t-1} + \varepsilon_t \tag{5}$$

where c_t and y_t are the deviations from trend of the log of aggregate consumption spending and 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 <u>*not*</u> what we mean by frequency-dependence.

What we <u>*do*</u> 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 (5), 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. For example, with adaptive expectations, the implication is that the coefficient γ_1 will be larger if the deviation y_{t-1} has the same sign as the deviation y_{t-2}, so that the deviation y_{t-1} is part of a smooth pattern. Note that this dependence of γ_1 on the recent history of y_{t-1} (and the resulting frequency dependence in γ_1) can thus 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 (1999a) and in Ashley and Verbrugge (2005a).) In principle one could thus eliminate the frequency dependence in γ_1 by modeling this nonlinearity in the relationship; and, indeed, one constructive aspect of the detection of frequency dependence in a relationship using the method proposed here would be to motivate and inform such an effort. The detection and modeling of frequency dependence is of interest in its own right, however, because such frequency dependence often "matches up" so nicely with economic theory, such as the permanent income hypothesis in the present example. In any case, in terms of the topic of this section, the essential point is that 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.

Similarly, viewing equation (5) as part of a bivariate VAR model, the impulse response function for c_t will be a function of past innovations in both equations, and c_t will depend differently on different lags in the y_t innovations. Frequency dependence alters the nature of the impulse response functions. In particular, if there is no frequency dependence in the $c_t - y_t$ relationship, then the moving average representation of the c_t process will be a linear function of serially independent innovations; this leads to a set of conventional *linear* impulse response functions in which the change in the expected value of c_{t+n} induced by an innovation in the y_t process of size δ is unrelated to the values of previous innovations. Conversely, frequency dependence in the $c_t - y_t$ relationship implies that the full moving average representation of the $c_t - y_t$ relationship (and hence, the impulse response functions also) consist of *nonlinear* functions of serially independent innovations. Thus, in that case, the change in the expected value of c_{t+n} induced by an innovation in the y_t process of size δ *does* depend on the values of previous innovations. (Of course, the Wold Theorem still guarantees the existence of a linear MA(∞) representation for c_t and y_t – and hence of a set of linear impulse response functions for these variables – but the innovations in *this* linear MA(∞) representation are not serially independent.)

The following explicit example clarifies this point.[7] Consider the particular case in which the linear moving average (Wold) representation for a series c_t can be approximated by the MA(1) process:

$$c_t = v_t + \gamma_1 v_{t-1}$$

in which the v_t innovation series is generated by the bilinear process:

$$v_t = 0.7 v_{t-2} u_{t-1} + u_t$$

where u_t is serially independent. It is easy to verify that the v_t generated by this bilinear process are serially uncorrelated, so this MA(1) process could in principle be the Wold representation for c_t. Now rewrite the moving average representation of c_t as a function of the current and past values of the serially independent innovations – $u_t, u_{t-1}, \ldots$ – by repeatedly substituting the bilinear model in to eliminate v_t, v_{t-1}, etc. from the model for c_t. In this way one obtains:

$$c_t = u_t + (\gamma_1 + 0.7u_{t-2} + \text{higher order terms})\, u_{t-1} + (0.7\gamma_1 u_{t-3} + \text{higher order terms})\, u_{t-2} + \ldots$$

where the higher order terms involve $(0.7)^2\, v_{t-4}u_{t-3}$, $(0.7)^2\, v_{t-5}u_{t-4}$, and so forth. Continued substitution would further elaborate these terms, but the point is clear: the coefficient on the serially independent innovation u_{t-1} is no longer a constant. Instead, it is $(\gamma_1 + 0.7u_{t-2})$ plus higher order terms. Consequently, the impulse response function at lag one is frequency dependent in the sense discussed here: the coefficient on u_{t-1} will be different when the previous innovation (u_{t-2}) is of the same sign as u_{t-1}. Thus, estimating a linear moving average model for c_t yields an impulse response coefficient estimate at lag one which cannot be stable over time or across frequencies, since c_t responds differently to a lag-one shock which is part of a smooth pattern than to a lag-one shock which has just changed sign from the previous period.

[7] See Potter (2000) for a formal treatment of nonlinear impulse response functions.

Finally, we conclude this section with a warning from McCallum (1984): "...the association of low-frequency time series statistics with 'long-run' economic propositions is not generally warranted. Instead, many so-called long-run relationships involve expectational relationships which have little or nothing to do with frequencies per se." Thus, any conclusion about the low frequency behavior of a model parameter, such as γ_1 in equation (5) is best viewed as an assertion as to how c_t responds to smooth fluctuations in y_{t-1}, not as a statement with regard to the long run relationship between c_t and y_t.

3.2 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 $dE[y_t|x_t]/dx_t$. However, 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, for example – 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 single well-defined value to test. Similarly, the least-squares estimate of β is in that case clearly neither a consistent estimator for β_0, nor for 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 simple consumption function example 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 of doubtful value.

Frequency dependence in the unemployment rate coefficient of equation (3) 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.

3.3 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 – since the underlying model coefficients are assumed constant – 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 3.4 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 (β), 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.[8]

3.4 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

$$Y = X\beta + \varepsilon \qquad \varepsilon \sim N\left[0, \sigma^2 I\right] \tag{6}$$

[8] 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.

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

$$WY = WX\beta + W\varepsilon \tag{7}$$

$$\widetilde{Y} = \widetilde{X}\beta + \widetilde{\varepsilon} \qquad \widetilde{\varepsilon} \sim N\left[0, \sigma^2 I\right] \tag{8}$$

where $\widetilde{Y} = WY$, etc., and where the $(j,k)^{th}$ element of W is given by $w_{jk} = \frac{1}{\sqrt{T}} \exp\left(\frac{2\pi ijk}{T}\right)$, 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 equation (6) and equation (8). What *has* changed, however, is that the T sample observations in Y and in each of the K columns of X are replaced by T observations on each $\widetilde{Y}$ and each column of $\widetilde{X}$, each of which now corresponds to a frequency in the interval $[0, 2\pi(T-1)/T]$. 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 (8) requires that $corr(\widetilde{x}_{k,j}, \widetilde{\varepsilon}_j)$ is zero for all values of j and k, where $\widetilde{x}_{k,j}$ is used to denote the j^{th} observation on $\widetilde{x}_k$. Since W embodies a two-sided transformation – i.e., $\widetilde{x}_{k,j}$ depends upon all of $x_{k,1}, ..., x_{k,T}$ and $\widetilde{\varepsilon}_j$ depends upon all of $\varepsilon_1, ..., \varepsilon_T$ – this condition requires that $x_{t,k}$ is strictly exogenous – i.e., 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 temporarily restricting attention to relationships in which $x_{k,1}, ..., x_{k,T}$ are strictly exogenous.

Spectral regression 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 – central to 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 (8).

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.[9]

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, cointegration, or 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 still applicable only to non-feedback relationships.

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 relationships. 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:

$$a_{j,t} = \begin{cases} \frac{1}{\sqrt{T}} & j = 1 \\ \sqrt{\frac{2}{T}} \cos\left[\frac{\pi j(t-1)}{T}\right] & j = 2, 4, ..., (T-2) \text{ or } (T-1) \\ \sqrt{\frac{2}{T}} \sin\left[\frac{\pi (j-1)(t-1)}{T}\right] & j = 2, 4, ..., (T-1) \text{ or } T \\ \frac{1}{\sqrt{T}} (-1)^{t+1} & j = T \text{ and } T \text{ is even, } t = 1, ..., T \end{cases} \tag{9}$$

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

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

[9] 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.

or

$$Y^* = X^*\beta + \varepsilon^* \qquad \varepsilon^* \sim N\left[0, \sigma^2 I\right] \tag{10}$$

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, $Var(\varepsilon^*) = Var(\varepsilon)$ because A is an orthogonal matrix.

Since the elements of the A matrix are all real-valued, equation (10) 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), Farley, Hinich and McGuire (1975), Ashley (1984), Bai (1997), or 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 (10) into m equal frequency bands and estimate how β_k varies by replacing X_k^*, the k^{th} column of X^*, with m appropriately constructed dummy variables:[10]

$$Y^* = X_{\{k\}}^*\beta_{\{k\}} + D^*\gamma^* + \upsilon^* \tag{11}$$

where $X_{\{k\}}^*$ is X^* omitting the k^{th} column, and $\beta_{\{k\}}$ is β omitting the k^{th} component. The columns $\left[D^{*1}...D^{*m}\right]$ comprising 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,

[10] 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.

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.[11]

3.5 The Tan-Ashley approach in the time domain

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 (11) through by A^T yields

$$A^T Y^* = A^T X^*_{\{k\}} \beta_{\{k\}} + A^T D^* \gamma + A^T \upsilon^* \tag{12}$$

and hence

$$Y = X_{\{k\}} \beta_{\{k\}} + D\gamma + \upsilon \tag{13}$$

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.[12] Consequently, the error vector υ is identical to the original error term in (6) if the m components of γ are all equal to β_k.

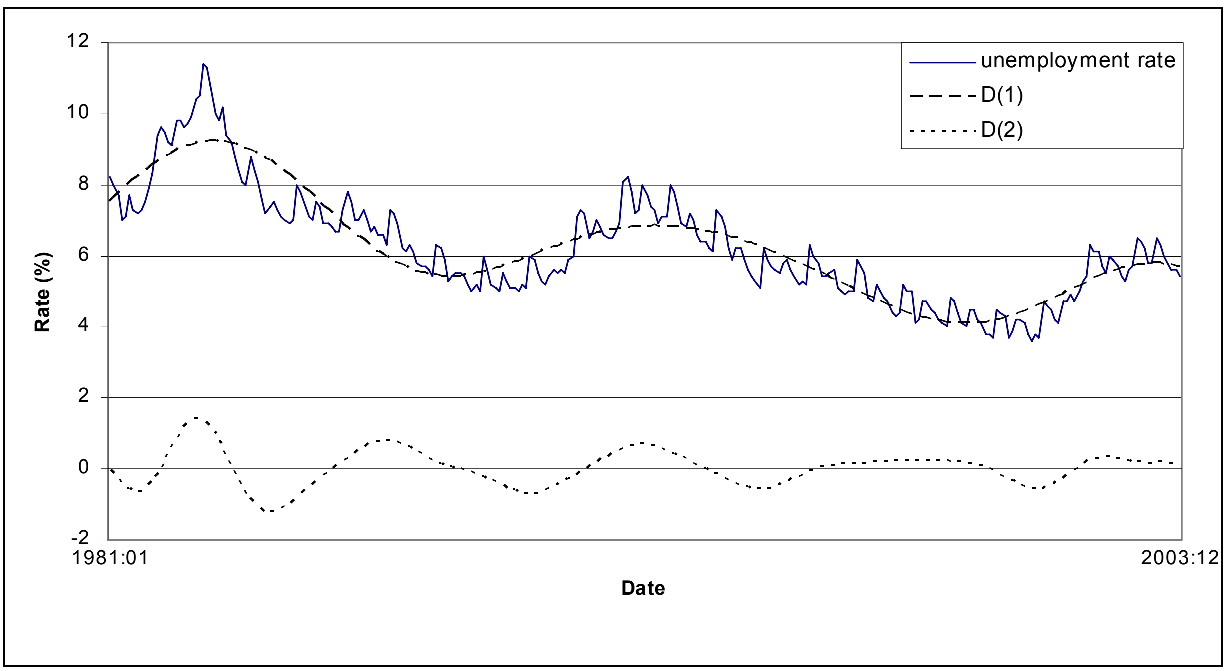
The column vectors $\left[D^1...D^m\right]$ are in essence bandpass filtered versions of X_k which partition this variable into m orthogonal components, one for each frequency band. For example, suppose that one were to partition the monthly US unemployment rate into three frequency components: D_t^1, comprising the fluctuations corresponding to low frequencies (periods greather than 60 months);

[11] We note that the foregoing analysis could all be applied substituting *any* orthogonal matrix for A, so long as its rows pick out components of increasing smoothness. The finite Fourier transforms used here are both familiar and compellingly unique, so long as one assigns the same coefficient value to both the sine and cosine rows corresponding to a particular frequency. But this is not to say that a useful transformation matrix could not be formulated in other ways – e.g. using wavelets, as defined defined by Ramsey and Lampert (1998a,b).

[12] 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.

D_t^2, a medium-frequency ("business cycle") component, corresponding to periods between 18 and 60 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 1981 through 2003.

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



No one of these m implied bandpass filters is an optimal bandpass filter. One might choose a Baxter-King (1999) or Christiano-Fitzgerald (2003b) bandpass filter for that purpose. But $\left[D^1 ... D^m\right]$ have the desirable property of partitioning X_k in an intuitively appealing way into m orthogonal frequency components that add up exactly to X_k. Consequently, 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.

In contrast, note that failing to replace $X_k \beta_k$ 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 in that case have a unique value to estimate.

Note also that, since X_k equals $D^1+...+D^m$, replacing $X_k\beta_k$ by $D\gamma$ 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. In contrast, a typical bandpass filtering analysis – e.g., Christiano and Fitzgerald (2003a), Comin and Gertler (2003), or Den Haan and Sumner (2004) – applies a bandpass filter to both Y and to some of the columns of X. Thus, one ends up in such analyses quantifying the relationship between these filtered time series, rather than the relationship between the actually observed variables. Such analyses also require an *a priori* selection of the frequency band to consider, which (as is described below) our approach does not.

Finally, note 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 use this approach to investigate the frequency dependence of the coefficient on X_k in more complex settings by replacing $X_k\beta_k$ 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 econometric techniques and software one was already employing.

3.6 The Problem with Feedback – and a Solution using One-Sided Filtering

Note that $\widehat{\gamma}^{OLS}$ will be a consistent estimate of γ in equation (13) 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 to the column vector X_k, this will be the case only if X_k is strictly 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 the components of Y and X_k 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:

$$\begin{aligned} y_t &= \lambda_1 y_{t-1} + \lambda_2 x_{t-1} + \varepsilon_t \\ x_t &= \alpha_1 x_{t-1} + \alpha_2 y_{t-1} + \eta_t \end{aligned} \tag{14}$$

Clearly, this is a feedback relationship only if α_2 is nonzero. But note that the x_t equation implies that

$$\begin{aligned} x_t &= \alpha_1 x_{t-1} + \alpha_2 y_{t-1} + \eta_t \\ &= \alpha_1 x_{t-1} + \alpha_2 \left(\lambda_1 y_{t-2} + \lambda_2 x_{t-2} + \varepsilon_{t-1}\right) + \eta_t \\ &= \alpha_1 x_{t-1} + \alpha_2 \lambda_1 y_{t-2} + \alpha_2 \lambda_2 x_{t-2} + \alpha_2 \varepsilon_{t-1} + \eta_t \end{aligned}$$

so that x_t is correlated with ε_{t-1} if there is feedback from past y_t to x_t. But, two-sided filtering implies that x^*_{t-1} depends upon x_t, x_{t+1}, x_{t+2}, etc., so that x^*_{t-1} is thus correlated with $\varepsilon_{t-1}, \varepsilon_t, \varepsilon_{t+1}, \varepsilon_{t+1}, ...$, which (under two-sided filtering) are correlated with ε^*_t. Thus, in the presence of feedback, a two-sided transformation of x_{t-1} will in general produce a transformed explanatory variable, x^*_{t-1}, which is correlated with the tranformed error term, ε^*_t, yielding inconsistent least-squares parameter estimates. (Examples of such two-sided filters include a filter based on the A matrix as discussed above, or the Hodrick-Prescott (1987) filter, or bandpass filters such as those given by Baxter and King (1999) and Christiano and Fitzgerald (2003).)

To eliminate this problem, we exploit the fact that the Tan-Ashley formulation is easily adapted to use only one-sided filtering.[13] The modified calculation steps through the sample using a moving window of length τ. In the first step, observations one through τ on X_k (i.e, $X_{k,1}, ..., X_{k,\tau}$) 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 (13). Next one uses the τ sample observations $X_{k,2}, ..., X_{k,\tau+1}$ to re-compute the τ-dimensional column vectors $D^1...D^m$. Again the last (τ^{th}) element in each of these column vectors becomes the period τ+1 observation on $D^1...D^m$ for use in estimating equation (13). And so forth.[14] Thus, one could characterize $D^1...D^m$ as being the result of a set of m

[13] Christiano and Fitzgerald (2003b) also provide a one-sided version of their filter, but they do not propose stepping this filter through the sample data, as is discussed below. Also, as noted above, their filter does not have the desirable property of being able to partition X_k into m components that add up exactly to X_k.

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

one-sided bandpass filters obtained using a moving block of τ observations.

The resulting $D^1 ... D^m$ columns still add up precisely to the original explanatory variable (X_k) over its last $T - \tau$ elements. These m columns are no longer orthogonal, but in practice they are not highly correlated with one another. In any case, the orthogonality is of modest importance: what is essential is that $D^1 ... D^m$ still precisely partition (sum up to) X_k and that they are now the product of a one-sided filter.

One must lose the use of $\tau - 1$ start-up observations in estimating equation (13) in this way, but this is necessary in order to avoid spurious results when feedback is present. This loss in degrees of freedom is manageable in the Phillips curve application given in Section 5 below: 60 monthly observations out of 288 at the beginning of the sample are sacrificed so as to be able to consider frequencies corresponding to periods as large as sixty months. (Twelve of these were "lost" due to lagged dependent variables in any case.)

Lastly, it must be mentioned that bandpass filters like the ones used here generically have problems near the endpoints of the sample. The standard method for addressing this shortcoming – e.g., Dagum (1978) and Stock and Watson (1999) – is to augment the sample using projected values obtained from univariate autoregressive models. Here, we adopt a suggestion from Christiano (2005) and form projections using a multivariate model (which includes four autoregressive lags, seasonal dummies, and other significant predictors such as lags of the index of help wanted ads). This model 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 an additional twelve months. The resulting $\tau + 12$ 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 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.[15]

[15] It would seem advisable to detrend the X_k 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 used to form a separate series D^0, or added back into observation τ of the lowest frequency band, D^1. In either case, $\sum D^i$ still sums to X_k. We note, however, that our empirical results (in Sections 5.2 and 5.3) were not sensitive to whether this detrending was done or not, or to whether

3.7 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 approach is to specify m 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, say, 18 months, a second band containing frequencies corresponding to periods between $1\frac{1}{2}$ and 5 years, and a third containing all frequencies corresponding to longer periods. This choice seems reasonable, but it is quite *ad hoc*: one might equally well choose one of many other calendar-based frequency band structures. Furthermore, one risks faulty inference. 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. But if not, then the calendar-based test could have relatively low power: 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 distorted to some degree.

An alternative approach is to choose the number and composition of the frequency bands so as to minimize an adjusted goodness-of-fit criterion, such as the Bayes-Schwarz Information Criterion (BSIC). However, in that case the sampling distribution of the F statistic for testing the null hypothesis of equal coefficients on all bands must be obtained by simulation so as to properly account for the extensive specification search undertaken; unfortunately, this leads to a test of very low power.

The approach adopted here is to simply allow the regression equation to estimate a distinct coefficient for every possible frequency allowed by the limited length of the window used to implement the one-sided filtering. For example, with the 72-month windows used in the Phillips Curve model estimated in Section 5 below, only 36 distinct frequencies – listed in Appendix 1 – are possible.[16]

D^0 was used as a separate regressor.

[16] There are half as many frequencies as months in the window because there is both a sine and a cosine

With, in this case, 288 sample observations, the degrees-of-freedom cost of estimating 36 frequency band coefficients is not prohibitive.[17]

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?

row in the A matrix of Section 3.4 for each distinct frequency.

[17] It does seem a bit wasteful, however, in view of the fact that one expects the frequency variation across frequencies to be fairly smooth. Consequently, we also investigated a more parsimonious approach in which the variation of the 36 coefficients is modeled by means of a low-order polynomial, as in the distributed lag literature. For the particular data used here, however, this approach – using both ordinary and Chebyshev polynomials – yielded less sharp results. In this case the benefit of estimating and testing fewer coefficients was outweighed by the cost (in terms of goodness of fit) imposed by the polynomial constraints.

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:

$$\begin{aligned} y_t &= \lambda_1 x_{t-1} + \lambda_2 y_{t-1} + \varepsilon_{y,t} \\ x_t &= \alpha_1 x_{t-1} + \alpha_2 y_{t-1} + \varepsilon_{x,t} \end{aligned} \tag{15}$$

where $\lambda_1 = 0.25$, $\lambda_2 = 0.55$, $\alpha_1 = 0.65$, and $\alpha_2 = 0.3$, and where $\varepsilon_{y,t}$ and $\varepsilon_{x,t}$ are niid(0,1); 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. For each of 1000 simulations, 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.[18]

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 $t = 1, ..., 300$. In both cases, $D_t^1 + D_t^2 + D_t^3$ exactly equals x_t.

Following this, y_t was regressed on y_{t-1} and $D_{t-1}^{1,k}, D_{t-1}^{2,k}$, and $D_{t-1}^{3,k}$ 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-1} 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}^1, D_{t-1}^2, and D_{t-1}^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; this rejection rate was even higher in a number of

[18] 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.

alternative specifications investigated. Evidently, two-sided filtering can readily lead to a spurious 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 the approximating 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:

$$\begin{aligned} y_t &= Bx_{t-1} + \lambda y_{t-1} + \sigma\varepsilon_{y,t} & \varepsilon_{y,t} &\sim niid\,(0,1) \\ x_t &= Ax_{t-1} + \gamma y_{t-1} + S\varepsilon_{x,t} & \varepsilon_{x,t} &\sim niid\,(0,1) \end{aligned} \tag{16}$$

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. 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 $\{D_t^1, D_t^2, D_t^3\}$. As in Section 4.2, 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. The dependent variable y_t was then regressed on a constant, y_{t-1}, and D_{t-1}^1, D_{t-1}^2, and D_{t-1}^3.

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

$$y_t = \underset{\underset{(3.41)}{-0.17}}{\widehat{\alpha}} + \underset{\underset{(0.96)}{0.04}}{\widehat{b}}\, x_{t-1} + \underset{\underset{(9.51)}{0.51}}{\widehat{\lambda}}\, y_{t-1} + u_t$$

$$y_t = \underset{\underset{(-1.97)}{-0.09}}{\widehat{\alpha}} + \underset{\underset{(3.82)}{0.16}}{\widehat{b}_1}\, D_{t-1}^1 + \underset{\underset{(-0.81)}{-0.16}}{\widehat{b}_2}\, D_{t-1}^2 + \underset{\underset{(-6.15)}{-0.55}}{\widehat{b}_3}\, D_{t-1}^3 + \underset{\underset{(10.94)}{0.52}}{\widehat{\lambda}}\, y_{t-1} + u_t$$

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:

$$\begin{aligned} y_t &= \lambda_1 z_{1,t-1} + \lambda_2 z_{2,t-1} + \lambda_3 y_{t-1} + \sigma \varepsilon_{y,t} \qquad (17) \\ z_{1,t} &= \rho_1 z_{1,t-1} + \gamma y_{t-1} + s\varepsilon_{x_1,t} \\ z_{2,t} &= \rho_2 z_{2,t-1} + \gamma y_{t-1} + \varepsilon_{x_2,t} \end{aligned}$$

where $\varepsilon_{y,t}$, $\varepsilon_{x_1,t}$, and $\varepsilon_{x_2,t}$ are niid(0,1). If $\gamma > 0$, this system exhibits positive feedback. Suppose that the analyst 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 non-persistent 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 $z_{2,t-1}$).

$T = 300$ 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 $\{D_t^1, D_t^2, D_t^3\}$. As in the previous sections, 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. Then y_t was regressed on a constant, y_{t-1}, and D^1_{t-1}, D^2_{t-1}, and D^3_{t-1}.

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

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

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 this process and the Markov-switching process considered in Section 4.2: 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. 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 and 3.7 is both necessary and effective in the presence of feedback.

5 Phillips Curve Estimation Results

5.1 Regression model specification

From (3), a standard Phillips curve specification is of the form

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

where un_t is the non-seasonally-adjusted total civilian unemployment rate, and where Z_t includes seasonal dummies, a measure of the change in the relative price of energy (Oil_t), and a measure of the change in relative import prices.[19] We consider monthly data over the period 1984-2003.[20] The measure of inflation used in constructing π_t is the growth rate of non-seasonally-adjusted CPI-U-RS.[21] As robustness checks, we also used the personal-consumption-expenditure (PCE) deflator, and (in a parallel quarterly model) the quarterly gdp deflator, as the dependent variable in the regression model. Results were qualitatively unchanged. Our extensive set of robustness checks are summarized in Appendix B.

The series un_t was decomposed into frequency bands $un^1 ... un^k$ using the one-sided filtering methodology described in Section 3.6 above, with k chosen as discussed below. Setting the window length τ to a number larger than 60 months (or 20 quarters) – corresponding to using more than

[19] As in, for example, Staiger, Stock and Watson (2001), and Eller and Gordon (2002). The energy series used was "energy commodities," which is then divided by the CPI-U-RS (lagged one month). The import price series used was the BLS Import Price Index (all commodities except petroleum), divided by the CPI-U-RS (lagged one month). Inflation in the trade-weighted real exchange rate is another proxy which is often used (e.g., Ball and Moffitt (200x) and Staiger, Stock and Watson (2001)), but we do not observe evidence in favor of its inclusion.

[20] We choose to start our sample in 1984:1, following convention in the literature (e.g., Stock and Watson (2005), who assert that "a large body of evidence ... informed our choice of a 1984 break," and who obtain statistical evidence for this break date.). This sample period focuses attention on the "Greenspan-era" inflation dynamics. (However, beginning the sample in 1980 does not alter our results, as noted in Appendix B.) Simultaneity bias is less likely to be problematic in monthly data. Here and following, heteroskedasticity-consistent standard error estimates are quoted.

[21] 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-equivalence homeownership costs. Note that other researchers, notably Crone, Nakamura and Voith (2001) suggest that additional adjustments may be worthwhile.

five years of data at the start of the sample in constructing the first window – seemed unreasonable, given the length of our sample period. Consequently, τ was set to five years, implying that each un_t^j observation is based on five-years' worth of prior data. Furthermore, since (as discussed in Section 3.7) the sixty months of actual data ($un_{t-59}...un_t$) are augmented by twelve months of projected data, the filtering window is 72 months long.

Equation (18) was then re-estimated using OLS in the form[22]

$$\pi_t = \alpha + \sum_{j=1}^{k} \beta_j un_t^j + \sum_{i=1}^{12} \delta_i \pi_{t-i} + \theta Z_t + \varepsilon_t \tag{19}$$

Here, un_t was fully partitioned into 37 components – one for each distinct frequency allowed using a 72 month rolling window, as discussed in Section 3.7.[23] It is necessary to use the one-sided filtering methodology discussed in Section 3.6, since we find evidence for significant feedback in the $\pi_t - un_t$ 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. Frequency-independence is rejected if the null hypothesis $H_0 : \beta_i = \beta_j, \forall i \neq j$ is rejected. The existence of a Phillips curve is contraindicated if $H_0 : \beta = 0$ cannot be rejected.

[22] Additional lags of the unemployment rate were not significant.
[23] Appendix A lists the frequencies and periods associated with a 72-observation rolling window.

5.2 Empirical results

Estimating the standard Phillips curve specification of equation (18) over the sample period 1984:1-2003:12 yields the OLS estimates:

$$\begin{aligned}\pi_t \;=\; & \underset{(-1.53)}{\underset{-0.83}{\alpha}} + \underset{(0.15)}{\underset{0.01}{\beta}}\, un_t + \underset{F-test:\,p=0.000}{\sum_{j=1}^{12}} \delta_j \pi_{t-j} + \underset{(18.24)}{\underset{0.04}{\theta_1}}\, Oil_t + \underset{(-3.33)}{\underset{-0.01}{\theta_2}}\, Oil_{t-1} \\ & + \underset{(2.56)}{\underset{0.04}{\theta_3}}\, IPE_t + \underset{F-test:\,p=0.000}{\sum_{i=1}^{11}} \theta_{i+4} month_t^i + \varepsilon_t \end{aligned} \tag{20}$$

Coefficient estimates, with their estimated t-statistics (based on robust standard error estimates) in parentheses, are given above for some coefficients; for others, we simply present the p-value for the F-test of the null hypothesis that all the coefficients in the distributed lag structure are zero. The variables Oil_t, IPE_t, and $month_t^1 ... month_t^{11}$ are the relative price of energy, the relative price of imports less petroleum, and seasonal dummy variables, respectively. Unlike many researchers (e.g., Gordon 1997; Brayton, Roberts and Williams 1999), we find that lags in π_t in excess of 12 months are not necessary to account for serial correlation in the model errors. This is likely due to our estimation period, in that we avoid the problematic 1970s. Furthermore, we find that the hypothesis $\sum_{j=1}^{12} \delta_j = 1$ is rejected by the data with a p-value of 0.002.

Note that the coefficient $\widehat{\beta}^{OLS}$ is not statistically significant. Thus, assuming no frequency dependence implies that the hypothesis $H_0 : \beta = 0$ cannot be rejected over this sample period. In other words, the estimation of a standard linear formulation of the Phillips curve over this time period suggests that, in fact, there <u>*is*</u> 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.

Re-estimating the Phillips curve in the form (19) over the same sample period – i.e., replacing

un_t with the 37 frequency components $un_t^1 ... un_t^{37}$ – yields the OLS estimates:

$$\pi_t = \underset{\underset{(-2.26)}{-3.53}}{\alpha} + \underset{F-test:\, p=0.009}{\sum_{j=1}^{37} \beta_j un_t^j} + \underset{F-test:\, p=0.000}{\sum_{j=1}^{12} \delta_j \pi_{t-j}} + \underset{\underset{(15.86)}{0.04}}{\theta_1} Oil_t + \underset{\underset{(-2.45)}{-0.01}}{\theta_2} Oil_{t-1}$$
$$+ \underset{\underset{(2.41)}{0.05}}{\theta_3} IPE_t + \underset{F-test:\, p=0.001}{\sum_{i=1}^{11} \theta_{i+4} month_t^i} + \varepsilon_t \tag{21}$$

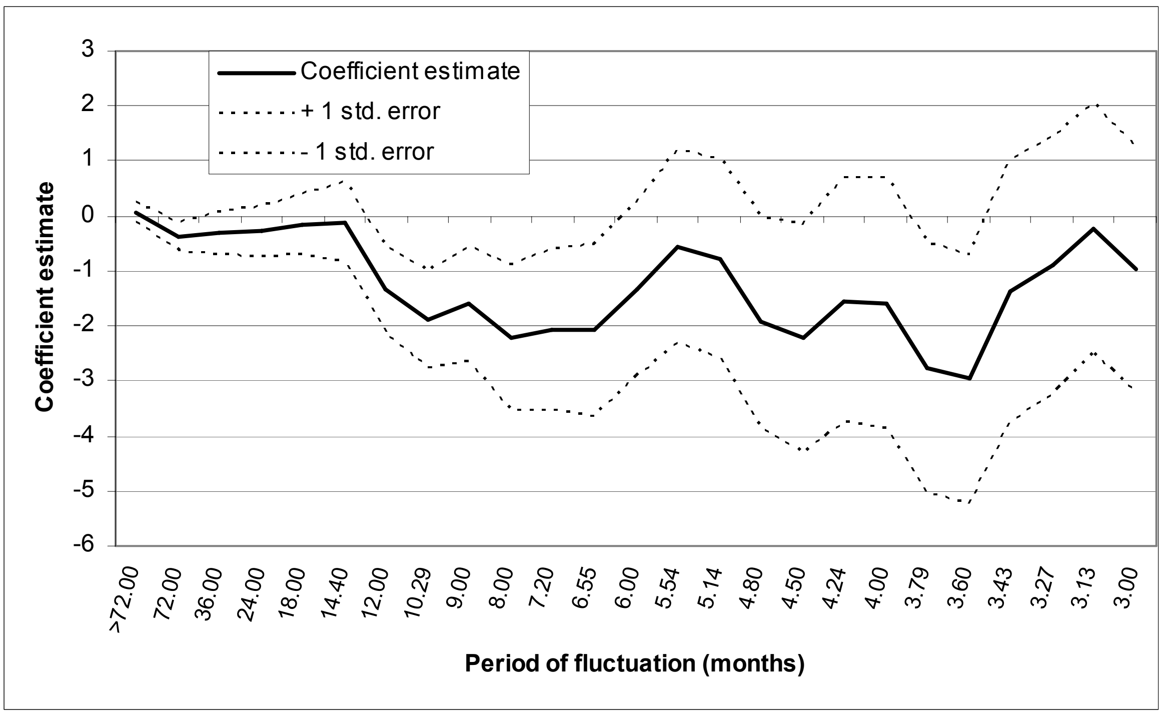
Relaxing the assumption of no frequency dependence reverses the conclusion regarding the existence of a Phillips curve; that is, $H_0 : \beta_1 = ... = \beta_{37} = 0$ is rejected at $p = 0.009$. Thus, fluctuations in the unemployment rate do have a statistically-significant relationship to fluctuations in the inflation rate once we allow for frequency dependence in the relationship. Furthermore, we can clearly reject the null hypothesis of no frequency dependence. In particular, the *p*-value for testing the null hypothesis that $\beta_1 = \beta_2 = ... = \beta_{37}$ is only 0.007. (This qualitative result – significant frequency dependence at the 4% level or better – was robust across the entire set of model specification variations listed in Appendix B.)

Since the coefficient of variation for each $\widehat{\beta}_j$ is rather large – as with raw sample periodogram estimates – some kind of smoothing must be imposed upon $\widehat{\beta}_1 ... \widehat{\beta}_{37}$ in order to learn anything about the *form* of the frequency dependence; here, it is of particular interest to determine whether this form is consistent with Friedman-Phelps theory. We discuss three smoothing approaches below, each of which corresponds to an alternative way of smoothing of the $\widehat{\beta}_j$ estimates across frequencies (i.e., across values of j): averaging, imposing an *a priori* band structure, and parameterizing the β_j as a low-order polynomial in j.

- Averaging

  One approach is to smooth the $\widehat{\beta}_j$ across the 37 values of j by averaging over adjacent coefficient estimates, using what amounts to a moving average. This averaging has no effect on the p-value at which either null hypothesis (all the $\beta_j = 0$, or all the β_j equal to each other) can be rejected. But it does impact the estimated standard deviation for each smoothed estimate. Using an equally-weighted moving average of width 7 yields the following Figure:

Figure 2: MA(7) Smoothed Coefficient Estimates



The coefficient estimates corresponding to unemployment rate fluctuations with periods less than 3 months are not plotted because they were still too noisy to interpret even at this level of smoothing. Note that the smoothed value of β_j is negligible for unemployment fluctuations with periods larger than about a year, and essentially negative for unemployment fluctuations with shorter periods.

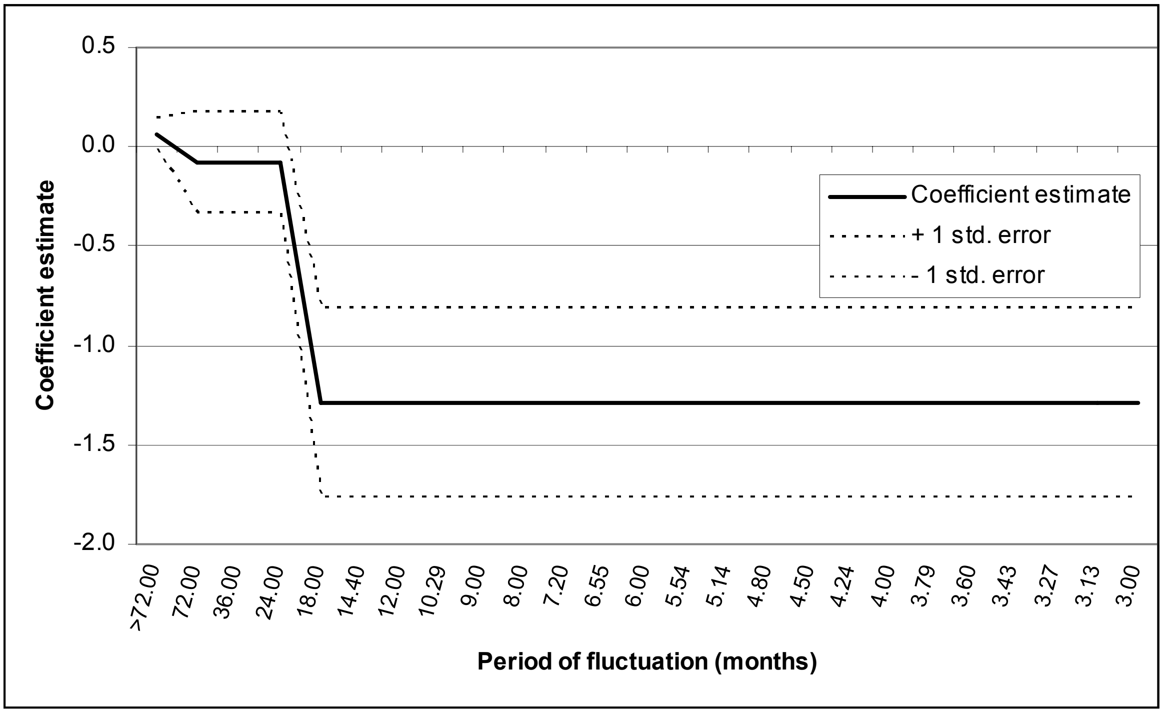
- Imposing an *a priori* band structure

  This approach, common in the real business cycle literature, imposes on the regression equation an *a priori* band structure that is suggested by economic theory. In particular, here we identify periods ranging from 18-72 months as a "business-cycle" band, and thus impose the restrictions $\beta_2 = \beta_3 = \beta_4$ (the business-cycle band) and $\beta_5 = ... = \beta_{37}$ (a "high-frequency" band, corresponding to fluctuations whose periods are less than 18 months). Our "business-cycle" band does not quite match the definition typically used in the literature – i.e., those fluctuations with periods between 18 and 60 months – since a filtering window of length 72 does not allow fluctuations with periods between 36 and 72 months in length to be distinguished (see Appendix A).[24] The lowest frequency band – comprising the single explanatory variable, un_t^1 – corresponds to unemployment rate fluctuations with periods greater than 72 months.

  Imposing these two restrictions, we find that one can still reject the hypothesis that β_j are all zero at $p = 0.04$, and that one can still reject the hypothesis that β_j are all equal at $p = 0.02$. A plot of the resulting $\widehat{\beta}_j$ across periods looks like:

[24] As noted in Appendix B, we also used a seven-year window, which allowed a "business-cycle" band of 19-48 months, closer to the standard partition used in the literature. These results still supported the existence of a Phillips curve and the rejection of frequency-independence, but were less supportive of the form of frequency-dependence imposed by the *a priori* band structure.

Figure 3: A-priori Band Structure Coefficient Estimates

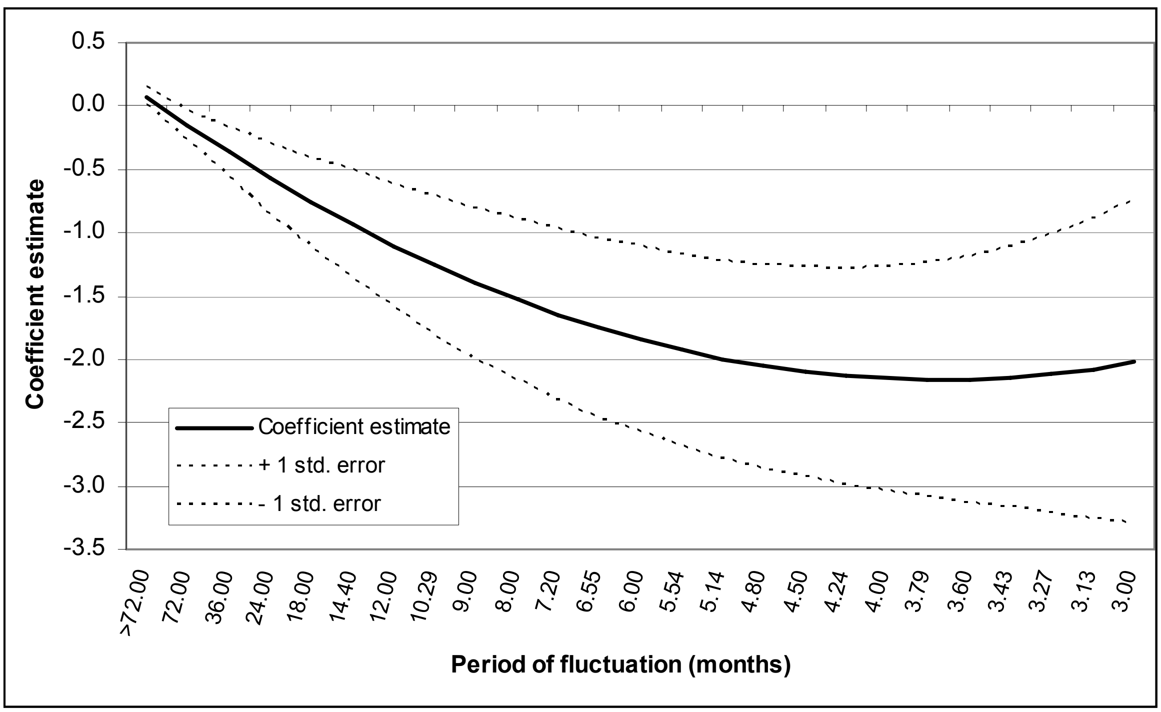


Note that this approach leads to a similar conclusion regarding the general form of the frequency dependence in the relationship; no doubt the p-values are larger because this arbitrary band structure fits the data less well.

- Parameterizing the $\widehat{\beta}_j$ as a low-order polynomial in j

  A final alternative imposes smoothness on the $\widehat{\beta}_j$ by parameterizing $\widehat{\beta}_j$ as a low-order polynomial in j. This is the same device used in the distributed lag literature. Here the polynomial order is chosen by optimizing a goodness-of-fit measure, such as FPE or BSIC. Doing so with these data yields a second-order polynomial. This smoothing restriction is evidently too restrictive, as one can now reject the hypothesis that β_j are all zero only at the 7% level. One can still reject the hypothesis that β_j are all equal at $p = 0.03$, however. A plot of the resulting $\widehat{\beta}_j$ across periods looks like:

Figure 4: 2nd-Order Polynomial Coefficient Estimates



We reiterate some salient points. Improperly imposing frequency-independence yields the conclusion that there is no Phillips curve. But once this assumption is relaxed, we can reject this null hypothesis at $p = 0.009$, and can reject the assumption of frequency-independence, i.e. $H_0 : \beta_1 = \beta_2 = ... = \beta_{37}$, at $p = 0.007$. Upon inspection of Figures 2-4, it appears that the form of frequency-dependence is consistent with the Friedman-Phelps hypothesis if one associates fluctuations in the unemployment rate with periods greater than about 12 months with movements in the natural rate.

We find that only rather high-frequency fluctuations in the unemployment rate – that is, fluctuations with periods less than a year or so – have an impact on inflation. Thus, the Phillips curve relationship is indeed inverse, but it is restricted to relatively high frequencies; we do not observe evidence for a relationship between inflation and the unemployment rate at low frequencies.

Finally, 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 the magnitude of this impact, we constructed the time series $impact_t$:

$$impact_t := \left| \left(\sum_{j=1}^{37} \widehat{\beta}_j un_t^j \right) - \widehat{\beta}^{OLS} un_t \right|$$

This series quantifies the magnitude of the estimated impact of fluctuations in un_t on the inflation rate from allowing for frequency dependence in the relationship. Because the frequency dependence in the $\pi_t - un_t$ relationship is almost entirely at high frequencies, $impact_t$ is quite noisy; consequently, it is smoothed by fitting to a fourth order polynomial. Figure 5 plots this smoothed series:

Figure 5: Smoothed magnitude of the inflation rate impact of modeling the frequency dependence in the relationship

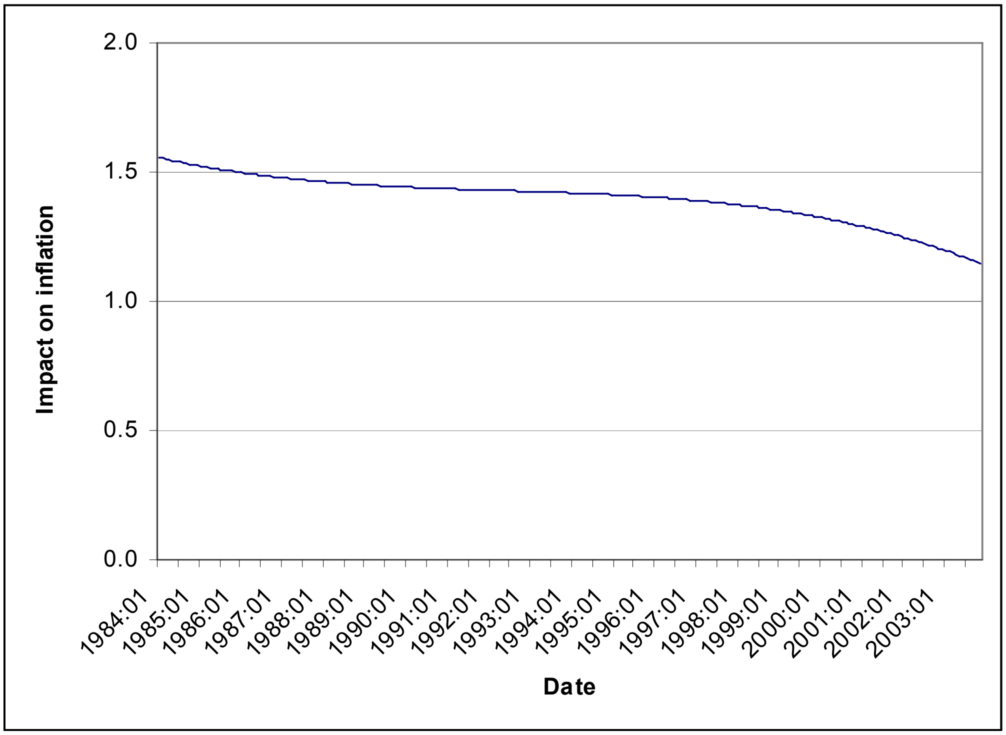


Figure 5 indicates that high-frequency fluctuations in the unemployment rate altered the inflation rate to an economically significant extent over this time period: a model ignoring the frequency dependence in the relationship would have understated the impact of unemployment rate fluctuations on the inflation rate by over one percentage point.

The existence of frequency dependence in the $\pi_t - un_t$ relationship indicates that much of the Phillips curve literature suffers from a serious mis-specification problem: 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 results obtained here indicate that fluctuations in unemployment that persist less than about a year are significantly associated with a contemporaneous fluctuation (of opposite sign) in inflation. In contrast, fluctuations in unemployment which persist longer than about a year are not significantly associated with contemporaneous fluctuations in inflation. These findings are consistent with the Friedman-Phelps formulation: one might interpret transitory un_t fluctuations, i.e. those with periods less than about one year, as deviations from the natural rate (and thus negatively associated with contemporaneous inflation); whereas more persistent un_t fluctuations (with periods larger than about one year) might be interpreted as movements in the natural rate, with the implication that such persistent unemployment fluctuations are not associated with significant inflation co-movements. Note that, as in King and Morley (2006) for example, the implicit natural rate is significantly more volatile than conventional estimates would suggest (see Williams 2004); concomitantly, under this interpretation of our results, a unit root process would be a rather poor approximation to the natural rate dynamics.

In summary, then, there *is* a Phillips curve relation – but it applies only to unemployment fluctuations with periods less than about one year. Consequently, econometric formulations of this relationship which fail to distinguish unemployment fluctuations within this range from those outside it are mis-specified. This result helps explain the apparent instability of estimated Phillips curve models across disparate time periods: for example, the Phillips curve will appear to be absent 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.[25] 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 1984-2003. Assuming that the relationship is not

[25] 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 corresponding to each distinct frequency allowed for a given window length. These components are only moderately correlated, and sum precisely to the input time series.

frequency dependent yields an estimate of the coefficient characterizing this Phillips curve relationship which is essentially zero. In contrast, allowing for the possibility of frequency dependence in this relationship, we find a significant Phillips curve relationship. In particular, our results show that there is a significant inverse relationship for high-frequency fluctuations in the unemployment rate – roughly speaking, for fluctuations whose period is less than about one year – and an insignificant relationship for more persistent unemployment fluctuations. A standard hypothesis test confirms that this pattern is statistically significant, at the 0.7% level. Our results in Figure 3, displaying the economic impact of these high-frequency fluctuations in the unemployment rate upon inflation, show that this impact is far from trivial.

What do these results mean? We draw two 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 several different frequency-specific coefficients, some negative and others negligible. In particular, 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 periods is less than about one year have an inverse relationship with inflation. In contrast, fluctuations in the unemployment rate which persist for more than about one year appear to have no relationship with inflation; the Friedman-Phelps theory would identify these more persistent fluctuations with variation in the natural rate.

7 Appendix A: 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 2 and greater of the A matrix discussed in Section 3 with a rolling filtering window 72 months in length. Row 1 corresponds to a within-window mean, through which the moving window will capture fluctuations with periods greater than 72 months. 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 Appendix B: Robustness checks

We obtained similar results:

- Imposing the restriction $\sum_{j=1}^{12} \delta_j = 1$ (Note, however, that the hypothesis $H_0 : \sum_{j=1}^{12} \delta_j = 1$ was rejected with a p-value of 0.002).
- Specifying equation (20) using $\Delta\pi$ in the place of π, as in Stock and Watson (2005).
- Including the change in the trade-weighted nominal, or real, exchange rate in equation (20).
- Using the Personal Consumption Expenditures (PCE) price index as the measure of inflation (broadly similar results were also obtained using the quarterly GDP deflator).
- Aggregating all fluctuations with period of one quarter or less into one band.
- Splitting the series Oil_t into two series, 1984:1-1986:01 and 1986:02-2001:12. (The motivation for this change is that the variance of this series appears to change markedly at the beginning of 1986.)
- Allowing the US CPI inflation process to have one structural break (in mean) in 1990:4. (The motivation for this change is that Benati and Kapetanios (2003) find compelling evidence for its existence.)
- Constructing the forecasts of the unemployment rate using only lags of the unemployment rate (and seasonal dummies), rather than on the basis of a richer multivariate specification.
- Applying the Tan/Ashley filter to a *non*-detrended unemployment rate, i.e., not detrending within each window.
- Using six, rather than twelve, lags of π in in equation (20).
- Removing IPE_t from equation (20).
- Running the regression over the period 1980:1-2003:12, rather than over 1984:1-2003:12, with and without the Benati/Kapetanios regime dummy.
- Using a seven-year window rather than a five-year window.
- Using a four-year window rather than a five-year window.
- Including the estimated trend into the lowest frequency band and using 36 bands $\left(D^1...D^{36}\right)$ as regressors in (20), rather than 37 $\left(D^0...D^{36}\right)$(See Section 3.6).

None of the above specifications overturned the basic results in this paper, as can be seen in Figures 4a and 4b, which plot the distribution of p-values for the two key hypothesis tests: $H_0 : \beta_1 = \beta_2 = ... = \beta_k$ and $H_0 : \beta_1 = ... = \beta_k = 0$. The suggested pattern of frequency dependence was also very similar across specifications.

Figure 4a: Distribution of p-values for $H_0 : \beta_1 = \beta_2 = ... = \beta_k$

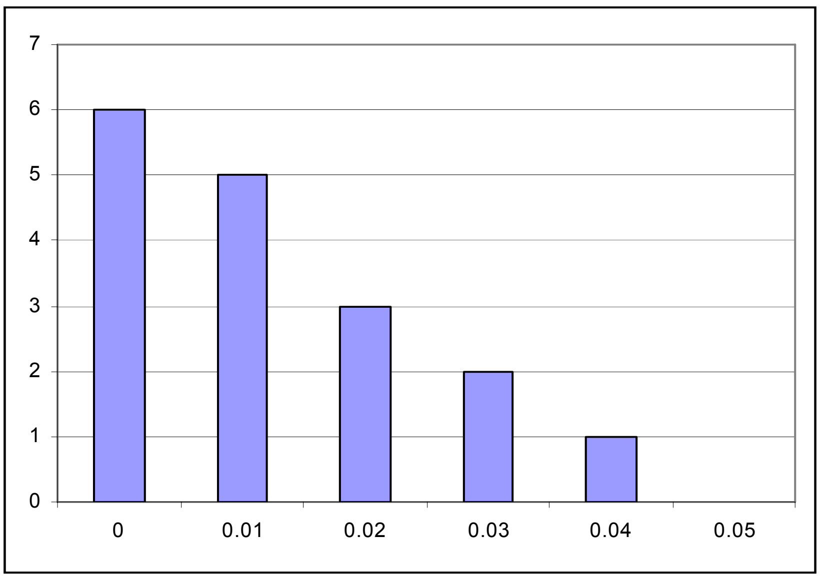
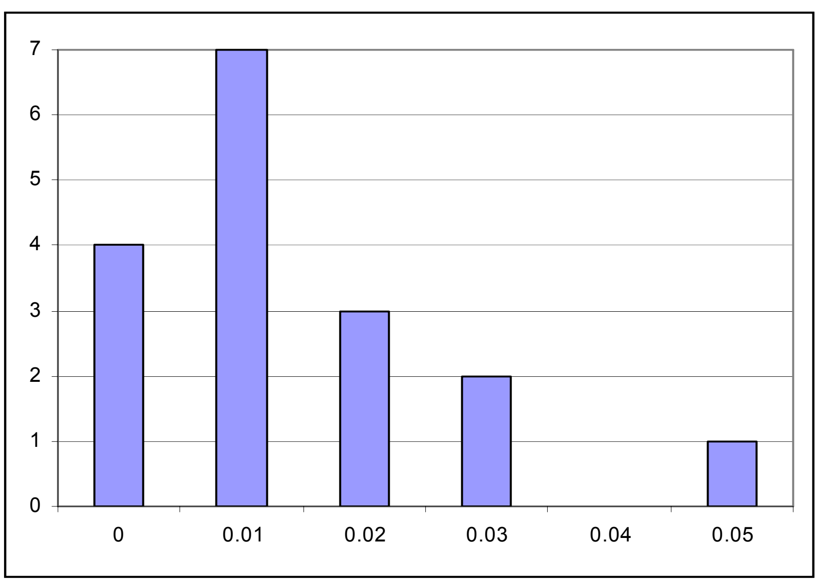


Figure 4b: Distribution of p-values for $H_0 : \beta_1 = ... = \beta_k = 0$



9 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.

Ashley, Richard, and Randal Verbrugge (2005a) "Frequency Dependence in Regression Model Coefficients: An Alternative Approach for Modeling Nonlinear Dynamic Relationships in Time Series." Mimeo, Virginia Tech.

Ashley, Richard, and Randal Verbrugge (2005b) "Frequency dependence in a New Keynesian Phillips curve." 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.

Beaudry, Paul, and Matt Doyle (2000) "What happened to the Phillips Curve in the 1990s in Canada?" *Price Stability and the Long Term Target of Monetary Policy.* Bank of Canada, 51-82.

Benati, Luca, and George Kapetanios (2003) "Structural breaks in inflation dynamics." Mimeo, Queen Mary University of London.

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

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. (2005) Private communication.

Christiano, Lawrence J., and Terry Fitzgerald (2003a) "Inflation and monetary policy in the twentieth century." Federal Reserve Bank of Chicago Economic Perspectives 1Q/2003, 22-45.

Christiano, Lawrence J., and Terry Fitzgerald (2003b) "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.

Comin, Diego, and Mark Gertler (2003) "Medium Term Business Cycles." NBER Working Paper 10003.

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

Dagum, Estela (1978) "Modelling, Forecasting and Seasonally Adjusting Economic Time Series with the X11 ARIMA Method." *The Statistician* 27.3/4, 203-16.

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.

Den Haan, Wouter J., and Steven W. Sumner (2004) "The Comovements between Real Activity and Prices in the G7." European Economic Review 48, 1333-1347.

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.

Eisner, Robert (1998) "The Economy is Booming. So Why are Economists Glum?" *The Wall Street Journal*, July 29, Editorial.

Eller, Jon, and Robert Gordon (2002) "Inflation and Unemployment in the New Economy: Is the Trade-off Dead or Alive?" Manuscript, Northwestern University.

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.

Farley, J. U., Hinich, M. and T. W. McGuire (1975) "Some Comparisons of Tests for a Shift in the Slopes of a Multivariate Linear Time Series Model." Journal of Econometrics 3, 297-318.

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.

Gali, Jordi, and Mark Gertler (1999) "Inflation dynamics: a structural econometric analysis." Journal of Monetary Economics 44, 195-222.

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

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

Gordon, Robert (1998) “Foundations 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.

Harvey, A.C. (1978) “Linear regression in the frequency domain.” *International Economic Review* 19, 507-12.

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

King, Thomas B., and James Morley (2006) “In Search of the Natural Rate of Unemployment.” Forthcoming, *Journal of Monetary Economics.*

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.

Laubach, Thomas and John C. Williams (2003) “Measuring the Natural Rate of Interest.” Review of Economics and Statistics 85.4, 1063-1070.

MaCullum, Bennett (1984) “On Low-Frequency Estimates of Long Run Relationships in Macroeconomics,” Journal of Monetary Economics 14, 3-14.

Orphanides, Athanasios, and John C. Williams (2002) “Robust Monetary Policy Rules with Unknown Natural Rate.” Brookings Papers on Economic Activity 2:2002, 63-118.

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.

Potter, Simon (2000) “Nonlinear Impulse Response Functions.” Journal of Economic Dynamics and Control 24.10, 1425-1446.

Ramsey, James B., and Camille Lampart (1998a) “The Decomposition of Economic Relationships by Time Scale Using Wavelets: Money and Income,” Macroeconomic Dynamics 2, 49-71.

Ramsey, James B., and Camille Lampart (1998b) “The Decomposition of Economic Relationships by Time Scale Using Wavelets: Expenditure and Income.” Studies in Nonlinear Dynamics and Econometrics 3.1, 23-42.

Rudd, Jeremy, and Carl Whelan (2005) “Modelling Inflation Dynamics: A Critical Review of Recent Research.” Finance and Economics Discussion Series 2005-66, Federal Reserve Board.

Rudd, Jeremy, and Carl Whelan (2006) “Can Rational Expectations Sticky-Price Models Explain Inflation Dynamics?” American Economic Review 96 (March), xx.

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 (2002) “Prices, wages and the US NAIRU in the 1990s.” in Alan B. Krueger and Robert Solow, Eds., The Roaring Nineties: Can Full Employment Be Sustained? New York: The Russell Sage Foundation and The Century Foundation Press, 3–60.

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.

Stock, James, and Mark Watson (2005) “Has Inflation Become Harder to Forecast?,” Mimeo, Princeton University.

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

Williams, John C. (2004) "Robust Estimation and Monetary Policy with Unobserved Structural Change." Mimeo, Federal Reserve Bank of San Francisco.