

Frequency Dependence in Regression Model Coefficients: An Alternative Approach for Modeling Nonlinear Dynamic Relationships in Time Series ^{*}

Richard Ashley ^a, Randal J. Verbrugge ^b

^a*Department of Economics Virginia Tech*

^b*U.S. Bureau of Labor Statistics*

Abstract

This paper proposes a new class of nonlinear time series models in which one of the coefficients of an existing regression model is frequency dependent – that is, the relationship between the dependent variable and this explanatory variable varies across its frequency components. We show that such frequency dependence implies that the relationship between the dependent variable and this explanatory variable is nonlinear. Past efforts to detect frequency dependence have not been satisfactory; for example, we note that the two-sided bandpass filtering used in such efforts yields inconsistent estimates of frequency dependence where there is feedback in the relationship. Consequently, we provide an explicit procedure for partitioning an explanatory variable into frequency components using one-sided bandpass filters. This procedure allows us to test for and quantify frequency dependence even where feedback may be present. A distinguishing feature of these new models is their potentially tight connection to macroeconomic theory: indeed, they are perhaps best introduced by reference to the frequency dependence in the marginal propensity to consume posited by the Permanent Income Hypothesis of consumption theory. An illustrative empirical application is given, in which the Phillips Curve relationship between inflation and unemployment is found to be negligible at low frequencies, corresponding to periods greater than a year, but inverse at higher frequencies – just as predicted by Friedman and Phelps in the 1960's.

JEL Classification Codes: C22, E00, E31

^{*} We thank Anthony Atkinson, Marianne Baxter, Larry Christiano, Walter Enders, Jonas Fisher, Steve Reed, and John Williams for helpful comments. 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.

Email addresses: ashleyr@vt.edu (Richard Ashley),

1 Introduction

Much of the literature on nonlinear time series modeling focuses on detecting and quantifying nonlinear serial dependence in a single time series. Yet surely it is obvious that the nonlinear dynamics of greatest interest from an economics perspective are usually multivariate in nature. Consequently, a new framework is proposed here for detecting and modeling nonlinear dynamic relationships *between* time series. In particular, the focus here is on modeling frequency dependence in a regression model coefficient.

A valuable and distinguishing characteristic of the frequency dependent regression model developed below is its potentially tight relationship with relevant economic theory. Indeed, the nature of this kind of frequency dependence is introduced in Section 2 with reference to a stylized consumption function embodying the well-known Permanent Income Hypothesis (PIH) of consumption theory, as in Modigliani and Brumberg (1954) and Friedman (1957). The PIH theory predicts that the partial derivative of aggregate consumption spending with respect to disposable income – which Keynes called the “marginal propensity to consume” – will differ across frequencies. In particular the PIH theory predicts that this coefficient will be large for slowly-varying (persistent or low frequency) fluctuations in household income, because these fluctuations are likely to be identified by the agents as primarily corresponding to changes in “permanent” income. In contrast, the theory predicts that the marginal propensity to consume will be small for quickly-varying (non-persistent or high frequency) fluctuations in household income, as these transitory fluctuations will be identified as primarily corresponding to changes in “temporary” income. Thus, the marginal propensity to consume – which would be a fixed parameter in an ordinary consumption function specification – is posited to be larger at low frequencies than at high frequencies: in other words, frequency dependence in this coefficient is the embodiment of the PIH theory.

Other examples of economic relationships which are likely to be frequency dependent abound: the interest rate elasticity of foreign exchange rates, price elasticities in markets for goods and services, the coefficient on unemployment in a Phillips Curve . . . the list goes on and on.

In Section 2 we also demonstrate – again using the PIH example – that this kind of frequency dependence inherently represents dynamic nonlinearity in the relationship. The frequency dependence is an intrinsic symptom of underlying nonlinearity which has not yet been explicitly modeled. In this sense, the frequency dependent regression model proposed here should be taken as a starting point rather than an ending point, in much the same way that an observation of conditional heteroscedasticity in a linear model’s errors ought actually to suggest that an investigation of a nonlinear model might be fruit-

`verbrugge.randal@bls.gov` (Randal J. Verbrugge).

ful. Still, because of its close relationship to theoretically important issues, the detection and modeling of frequency dependence is an important issue in its own right and, in particular, can yield graceful and powerful tests of the relevant theories. Moreover, while the frequency dependent regression framework does not *directly* suggest the form of an appropriate underlying nonlinear model, it is suggestive and does provide a benchmark for evaluating the essential adequacy of any proposed nonlinear model, in that one might sensibly require that data simulated by this model should exhibit the kind of frequency dependence which has been observed.

Past efforts to detect and model frequency dependence have not been satisfactory. For example, a typical approach has been to apply a bandpass filter to both the dependent and explanatory variables so as to confine the entire analysis to a specific, pre-chosen band of frequencies, often characterized as “business cycle” frequencies.¹ There are several shortcomings to this approach. For example, one might still wonder what impact fluctuations in an explanatory variable at business cycle (and other) frequencies have on the unfiltered dependent variable. And one might further wonder to what degree the results of such studies hinge on the rather *ad hoc* specification of this particular band of frequencies: perhaps important relationships in the data have been obscured by an unfortunate choice as to which frequencies to include in the analysis. Finally, we show below that the two-sided filters used in such analyses substantially distort the analysis when, as is typical in macroeconomic systems, feedback is present in the relationship.²

Our proposed modeling procedure is described below in Sections 3 and 4. Briefly, the idea is to partition the explanatory variable whose coefficient is of interest into a set of M components, such that each component corresponds to a particular set of frequencies, and such that the M components sum up precisely to the sample data on the original explanatory variable. This explanatory variable is then replaced in the model by a weighted sum of these M components. The resulting M coefficients on these frequency components are estimated in place of the original coefficient; these coefficient estimates then map out the frequency dependence in the original coefficient. In this context a test for frequency dependence is then straightforward: one merely tests the null hypothesis that these M coefficients are all equal. Moreover, this new approach is easy to implement since this linear form in the M components

¹ E.g., Christiano and Fitzgerald (2003a), Comin and Gertler (2003) and Den Haan and Sumner (2004).

² The results from spectral regression approaches to the modeling of frequency dependent coefficients – e.g., Hannan (1963), Engle (1974, 1978), and Tan and Ashley (1999b) – are similarly distorted by feedback because the Fourier transformation used in these approaches is itself a two-sided filter. An exposition of our present approach as an extension of Tan and Ashley (1999b) is given in Ashley and Verbrugge (2006).

can be substituted into whatever estimation framework was already in use, whether simple or complex.

Since two-sided filtering is problematic in the presence of feedback, we implement the partitioning of the explanatory variable into frequency components using a sequence of one-sided filters obtained from a moving window passed through the sample. The length of this window limits the size of the smallest frequencies (longest periods) which can be separately distinguished, but has the concomitant advantage of making it feasible to estimate the regression coefficient at each of the entire set of allowed frequencies. Consequently, the procedure lets the data speak freely as to the form of the frequency dependence without imposing any *ad hoc* band structure on it.

An illustration of the effectiveness of procedure is given in Section 5 using artificially generated data. An illustrative empirical example, drawn from Ashley and Verbrugge (2006), is given in Section 6. There we analyze the frequency dependence in the coefficient on the unemployment rate in a standard Phillips Curve model for the inflation rate using monthly U.S. data from 1980 to 2003. We find an economically and statistically significant inverse relationship between inflation and unemployment for high frequency unemployment rate fluctuations – with periods less than twelve months – but no evidence for an effect of lower frequency unemployment rate fluctuations. If one interprets low frequency unemployment rate fluctuations as shifts in the so-called “natural rate” of unemployment and high frequency fluctuations as temporary deviations from this natural rate, these results are supportive of the Friedman–Phelps inflation theory. In contrast, a model ignoring frequency dependence – i.e., constraining the coefficient to be the same at all frequencies – exhibits no statistically significant relationship at all between the inflation and unemployment rates over this period. Thus, a failure to recognize and model the dynamic nonlinearity in this relationship using our frequency dependent regression approach is in this case catastrophic: it leads to the erroneous conclusion that the Phillips Curve relationship does not exist at all.

2 Frequency Dependent Regression Parameters and Nonlinear Dynamics

The parameter on each explanatory variable in an ordinary regression model is a fixed, albeit unknown, constant. In contrast, the value of such a parameter in a frequency dependent regression model varies over time because – due to its variation with frequency – its value at any given time depends on the recent history of the explanatory variable it multiplies. Consider, for example, the case where this parameter’s value decreases with frequency. In that case the value of the parameter is larger when a current fluctuation in the corresponding explanatory variable is part of a smooth, persistent pattern of

similar recent changes – what might be called a low frequency fluctuation. Conversely, the value of the parameter is smaller when the current fluctuation in this explanatory variable is instead a brief, transitory (high frequency) fluctuation in the time series.

In view of the fact that such a parameter does not have a single value to estimate, least squares estimators of it cannot possibly be consistent. Thus, such estimates can easily mislead; the Phillips Curve example in Section 7 illustrates this nicely.

The Permanent Income Hypothesis (PIH) in consumption theory provides the canonical example of this kind of frequency dependence. Consider, for example, the following simple consumption function relating aggregate consumption at time t (c_t) to aggregate disposable income in the previous period (y_{t-1}), where each of these variables is expressed as a deviation from its trend value:

$$\begin{aligned} c_t &= \lambda_1 y_{t-1} + \lambda_2 c_{t-1} + \epsilon_{c,t} \\ y_t &= \alpha_1 y_{t-1} + \alpha_2 c_{t-1} + \epsilon_{y,t}. \end{aligned} \tag{1}$$

The parameter λ_1 is what would, in Keynes' terminology, be identified as the marginal propensity to consume; this parameter would be taken as constant in an ordinary regression equation. In contrast, the PIH implies that λ_1 will be larger in time periods during which y_{t-1} is primarily part of smooth fluctuation over time – identified in the theory as a change in permanent income – and smaller in periods where during which y_{t-1} is a sudden fluctuation, corresponding in the theory to a change in temporary income. In the present context, these are identified as low frequency and high frequency changes in income, respectively, so that λ_1 is indicated to be a function – in this case, an inverse function – of the degree to which y_{t-1} is a high frequency fluctuation.

One might model this kind of frequency dependence in λ_1 by observing that this dependence implies that the value of λ_1 , rather than being a constant, depends on the recent history of y_{t-1} . For example, a crude model for λ_1 exhibiting history dependence consistent with frequency dependence (and the PIH theory) would parametrize λ_1 as:

$$\begin{aligned} \lambda_1 &= l_0 / \left\{ 1 + l_1 \left(y_{t-1} - y_{t-1}^{smoothed} \right)^2 \right\} \\ &= l_0 / \left\{ 1 + l_1 \left(y_{t-1} - \left[\frac{1}{2} y_{t-1} + \frac{1}{2} y_{t-2} \right] \right)^2 \right\} = l_0 / \left\{ 1 + \frac{1}{4} l_1 (y_{t-1} - y_{t-2})^2 \right\} \end{aligned} \tag{2}$$

with l_1 a positive constant. This particular formulation specifies that λ_1 is smaller to the extent that the current value of lagged income represents a deviation from its average over the previous two periods. Clearly, this simple

model for λ_1 implies that

$$c_t = \left\{ \frac{l_0}{1 + \frac{1}{4}l_1(y_{t-1} - y_{t-2})^2} \right\} y_{t-1} + \lambda_2 c_{t-1} + \epsilon_{c,t} \quad (3)$$

so that, in the context of this particular example, the relationship between c_t and y_{t-1} is dynamically nonlinear if and only if λ_1 is frequency dependent in the sense used here.³

The particular parameterization of the history dependence of λ_1 posited above provides a useful example of what we mean by frequency dependence in a regression parameter and how this dependence is equivalent to otherwise-unmodeled nonlinear dynamics in the relationship. Indeed, this example suggests that the frequency dependent regression model described below is actually a new class of nonlinear dynamic model.

However, the particular parameterization examined above is too specific and *ad hoc* to provide an attractive general framework for detecting and modeling frequency dependence in a regression relationship. For such a framework we turn in the next Section to a consideration of how this kind of frequency dependence can be more gracefully examined using ideas based on the spectral regression literature.

3 The Frequency Dependent Regression Model

Frequency dependence in a regression coefficient is most gracefully examined, not by positing some specific, particular model for the history-dependence of the coefficient – as in the example at the end of the previous section – but by transforming the regression equation into the frequency domain, as in the spectral regression models of Hannan (1963), Engle (1974, 1978), and Tan and Ashley (1999a,b).

These early spectral frameworks typically required specialized software and (because of the two-sided nature of the fourier transforms used) were unsuitable for use in feedback relationships. Both of these limitations are eliminated in the work reported here, which is based on theory developed in Ashley and Verbrugge (2006); this framework involves only time domain regressions and effectively uses only one-sided transformations. Nevertheless, the Engle (1974) spectral regression framework is a good place to begin the exposition.

In Engle's formulation, the ordinary multiple regression model

$$Y = X\beta + \epsilon \quad \epsilon \sim N(0, \sigma^2 I) \quad \text{where } Y \text{ is } T \times 1 \text{ and } X \text{ is } T \times K \quad (4)$$

³ See Tan and Ashley (1999a) and also Ashley and Verbrugge (2006, Section 3), which includes a discussion of nonlinear impulse response functions and the Wold decomposition.

is transformed via premultiplication by the complex-valued $T \times T$ matrix W , whose $(k, t)^{th}$ element is given by $(1/\sqrt{T})\exp[(2\pi itk)/T]$. This yields

$$\begin{aligned} WY &= WX\beta + W\epsilon \\ \tilde{Y} &= \tilde{X}\beta + \tilde{\epsilon} \quad \tilde{\epsilon} \sim N(0, \sigma^2 I) \end{aligned} \quad (5)$$

which define \tilde{Y} , \tilde{X} and $\tilde{\epsilon}$; the variance of $\tilde{\epsilon}$ is still $\sigma^2 I$ because W is an orthogonal matrix.

Note that the coefficient vector β is unaffected by this transformation but that the T components of Y and $\tilde{\epsilon}$ and the T rows of \tilde{X} now correspond to the T frequencies, $\{0, 2\pi/T, 4\pi/T, \dots, 2\pi(T-1)/T\}$ rather than to the T time periods of the original regression model. Thus, finding that the j^{th} component of β ‘‘depends on frequency’’ is equivalent to finding that β_j is unstable across the T ‘‘observations’’ in the model for \tilde{Y} .

Engle’s spectral regression framework is problematic because W is complex-valued, implying that \tilde{Y} and \tilde{X} are complex-valued also. The Tan and Ashley (1999a,b) framework resolves this difficulty by replacing W with the real-valued matrix A , whose $(s, t)^{th}$ element is given by

$$a_{s,t} = \begin{cases} \left(\frac{1}{T}\right)^{\frac{1}{2}}, & \text{for } s = 1; \\ \left(\frac{2}{T}\right)^{\frac{1}{2}} \cos\left[\frac{\pi s(t-1)}{T}\right], & \text{for } s = 2, 4, 6, \dots, (T-2) \text{ or } (T-1); \\ \left(\frac{2}{T}\right)^{\frac{1}{2}} \sin\left[\frac{\pi(s-1)(t-1)}{T}\right], & \text{for } s = 3, 5, 7, \dots, (T-1) \text{ or } T; \\ \left(\frac{1}{T}\right)^{\frac{1}{2}} (-1)^{t+1}, & \text{(for } s = T \text{ when } T \text{ is even)} \end{cases} \quad (6)$$

This matrix is simply related to W by means of row manipulations based on the usual exponential expressions for the sine and cosine – e.g., $\cos(x) = 1/2e^{ix} + 1/2e^{-ix}$. Premultiplying the original regression model instead by A yields

$$\begin{aligned} AY &= AX\beta + A\epsilon \\ Y^* &= X^*\beta + \epsilon^* \quad \epsilon^* \sim N(0, \sigma^2 I) \end{aligned} \quad (7)$$

which defines Y^* , X^* and ϵ^* ; the variance of ϵ^* is still $\sigma^2 I$ because A is an orthogonal matrix. Again, the T components of Y^* and ϵ^* and the T rows of X^* now correspond to frequencies rather than to time periods. Thus, finding that the j^{th} component of β ‘‘depends on frequency’’ is again equivalent to finding that β_j is unstable across the T ‘‘observations’’ in the model for Y^* .

Note, however, that – since A , Y^* and X^* are all real-valued – the model for Y^* can be estimated using ordinary regression software.

Consequently, the constancy of β_j across the T “observations” in the model can be examined using any of the multitude of parameter instability tests available in the literature – e.g., Chow (1960), Brown, Durbin and Evans (1975), Farley, Hinich and McGuire (1975), Ashley (1984), Bai (1997), or Bai and Perron (1998, 2003).

Supposing, for the moment, that it is possible to partition these T “observations” into M “frequency bands”, the parameter instability test which is most convenient to use for this purpose is the straightforward extension of the usual Chow test for parameter instability given in Ashley (1984). This test amounts to simply assigning an appropriately defined dummy variable, $D^{*1} \dots D^{*M}$, to each band and testing the null hypothesis that the coefficients on all M of these dummy variables are equal. Thus, in testing for possible frequency dependence in the j^{th} component of β , the s^{th} “observation” on the dummy variable for the m^{th} band – denoted D_s^{*m} – would equal $X_{s,j}^*$ for values of s in the m^{th} frequency band and would equal zero for values of s outside of this band. Because these M dummy variables by construction add up to the j^{th} column of X^* , the constancy of β_j across the T “observations” can be readily tested by replacing the $\beta_j X_{s,j}^*$ term in the frequency domain regression (Equation 7) by the linear form $\sum_{m=1}^M \beta_{j,m} D^{*m}$:

$$Y^* = X_{\{j\}}^* \beta_{\{j\}} + \sum_{m=1}^M \beta_{j,m} D^{*m} + \epsilon^* \quad (8)$$

where $X_{\{j\}}^*$ denotes the X^* matrix with its j^{th} column omitted and $\beta_{\{j\}}$ denotes the β vector with its j^{th} component omitted. One then tests the null hypothesis $H_0 : \beta_{j,1} = \beta_{j,2} = \dots = \beta_{j,M}$.

For the present purpose, however, it is even more convenient to transform this regression equation back into the time domain by pre-multiplying it by A^{-1} . Since this matrix inverse is just the transpose of A in this case, we have:

$$\begin{aligned} A^t Y^* &= A^t X_{\{j\}}^* \beta_{\{j\}} + \sum_{m=1}^M A^t \beta_{j,m} D^{*m} + A^t \epsilon^* \\ Y &= X_{\{j\}} \beta_{\{j\}} + \sum_{m=1}^M A^t \beta_{j,m} D^{*m} + \epsilon \end{aligned} \quad (9)$$

where $X_{\{j\}}$ denotes the original X matrix with its j^{th} column omitted. Note that this regression model is identical to the original model (Equation 4) except that now there are M new regressors – $A^t D^{*1} \dots A^t D^{*M}$ replacing X_j , the j^{th} column of X .

Each of these new regressors can be interpreted as the result of applying a

simple passband filter to X_j – one for each of the M frequency bands.⁴ These implied filters are not optimal in terms of sharp passband cut-offs – one would choose Baxter–King (1999) or Christiano-Fitzgerald (2003b) filters instead were that the goal – but they have the desirable property of partitioning X_j into M orthogonal frequency components whose sum is precisely X_j . Thus, the stability of β_j across the M frequency bands can be readily examined by simply estimating Equation 9 and testing the null hypothesis, $H_0 : \beta_{j,1} = \beta_{j,2} = \dots = \beta_{j,M}$.

Thus, once one has decided on exactly which frequencies one wants to include in each of the M frequency bands, it is straightforward to partition the j^{th} regressor (X_j) into M frequency components which add up precisely to X_j and test β_j for frequency dependence. In principle, one can choose the number and composition of the M frequency bands to minimize some adjusted measure of the regression model’s goodness-of-fit, such as the Schwarz Criterion. However, this specification search will substantially distort the sampling distribution of the F statistics used in testing hypotheses about $\beta_{j,1} \dots \beta_{j,M}$. Unfortunately, correcting for this distortion by estimating these sampling distributions using monte carlo simulations yields a test of low power.

And there is a second problem. The bandpass filters defined above – in common with all other bandpass filters in common use – are two-sided. That is, the t^{th} observation in the component of X_j for the m^{th} frequency band – i.e., the t^{th} component of $A^t D^{*m}$ – depends on all T values of X_j , including the future values, $X_{t+1,j}, \dots, X_{T,j}$. This is of no special concern if there is unidirectional Granger causality from X_j to Y , but it induces inconsistency in least squares estimates of $\beta_{j,1} \dots \beta_{j,M}$ if, as is commonly the case in macroeconomic and financial applications, there is feedback in the $Y - X_j$ relationship.

Both of these problems are eliminated by instead partitioning X_j into frequency components using one-sided filters based on a moving window; this is addressed in the next section.

4 The Problem with Feedback – and a Solution Using One-Sided Filtering

Least squares estimators of $\beta_{j,1} \dots \beta_{j,M}$, M will be consistent if and only if the error term in the model for Y is uncorrelated with each of the explanatory variables in the model, including the M frequency components of X_j . Since, as defined in Section 3 above, each of these components is the result of what amounts to a two-sided bandpass filter applied to X_j , this will be the case only if X_j is strongly exogenous, that is, only if every observation on X_j –

⁴ See Ashley and Verbrugge (2006, sections 3.4 and 3.5) for a more detailed exposition.

i.e., $X_{1j} \dots X_{Tj}$ – is uncorrelated with every observation on the error term in the regression model for Y . (This is, of course, equally the case for *any* methodology which applies a two–sided bandpass filter to X_j .) Unfortunately, feedback in the $Y - X_j$ relationship 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:

$$\begin{aligned} y_t &= \lambda_1 y_{t-1} + \lambda_2 x_{t-1} + \epsilon_t \\ x_t &= \alpha_1 x_{t-1} + \alpha_2 y_{t-1} + \eta_t \end{aligned} \tag{10}$$

Clearly, this is a feedback relationship only if α_2 is nonzero. But note that the Equation 10 implies that

$$\begin{aligned} x_t &= \alpha_1 x_{t-1} + \alpha_2 y_{t-1} + \eta_t \\ &= \alpha_1 x_{t-1} + \alpha_2 (\lambda_1 y_{t-2} + \lambda_2 x_{t-2} + \epsilon_{t-1}) + \eta_t \\ &= \alpha_1 x_{t-1} + \alpha_2 \lambda_1 y_{t-2} + \alpha_2 \lambda_2 x_{t-2} + \alpha_2 \epsilon_{t-1} + \eta_t \end{aligned} \tag{11}$$

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 $x_{t-1}^1 \dots x_{t-1}^M$, the M frequency components of x_{t-1} , all depend on $x_t, x_{t+1}, \dots, x_{T-1}$. Consequently, $x_{t-1}^1 \dots, x_{t-1}^M$ are all correlated with $\epsilon_{t-1}, \epsilon_t, \dots, \epsilon_{T-2}$. Since $x_{t-1}^1 \dots x_{t-1}^M$ are thus all correlated with ϵ_t , replacing $\lambda_2 x_{t-1}$ in Equation 10 by a weighted sum of $x_{t-1}^1 \dots x_{t-1}^M$ will yield inconsistent parameter estimates.⁵

To eliminate this problem, we instead partition X_j into frequency components using one–sided filters based on a moving window. This calculation steps through the sample data using blocks which are τ periods in length, where a typical value for τ might be 48 or 60 with monthly data. In the first step, observations one through τ on X_j (i.e., $X_{1,j} \dots X_{\tau,j}$) are used to compute the M bandpass–filtered component series, $A^t D^{*1} \dots A^t D^{*M}$. Note that A is now a $\tau \times \tau$ matrix, so that each of these M component series is τ periods in length. The last (period τ) observation in each of these M components is retained as the filtered output for this window; the other $\tau - 1$ observations on the M components are discarded. The window is next advanced one period and the new set of τ sample observations (i.e., $X_{2,j} \dots X_{\tau+1,j}$) are bandpass filtered to again yield M component series; the last observation in each of these M series (which now corresponds to period $\tau + 1$ in each case) is retained as the filtered output for this window and, as before, the other $\tau - 1$ observations on

⁵ Note that this problem with feedback is not particular to the approach used here: the above argument also implies that applying a two–sided bandpass filter to both y_t and x_{t-1} , as in Christiano and Fitzgerald (2003), Comin and Gertler (2003) and Den Haan and Sumner (2004) will similarly lead to inconsistent least squares estimation if there is feedback in the $y - x$ relationship. Examples of two–sided bandpass filters include the Hodrick–Prescott (1987) filter, and bandpass filters such as those given by Baxter and King (1999) or Christiano and Fitzgerald (2003b), as well as the filters based on the A matrix as discussed in the previous section.

the M components are discarded. And so forth. Finally, in the last window, the sample observations $X_{T-\tau+1,j} \dots X_{T,j}$ are bandpass filtered to compute the M component series and the last observation in each (now corresponding to period T) is retained as the filtered value derived from this final window.

In this way, $X_{t,j}$ can be partitioned into M frequency components using only observations $X_{t,j}, X_{t-1,j}, \dots, X_{t-\tau+1,j}$ for each time period in the interval $[\tau, T]$. These M components are no longer precisely orthogonal, but they still, by construction, add up exactly to $X_{t,j}$ for each time period in this interval. A weighted sum of these M frequency components can now be used to replace X_j even in settings where feedback in the $Y - X_j$ relationship is a possibility because each of these components is now effectively the product of one-sided bandpass filtering.

By specifying a modest value for τ , the length of the moving-window, it becomes feasible to estimate a distinct coefficient for every possible frequency. Thus, partitioning X_j in this way not only makes it possible to filter in an effectively one-sided manner, it also eliminates the need to choose a value for M and to specify which frequencies are to go into each of the M bands. Instead, one can simply calculate a component for every possible frequency allowed by the length of the window. For example, setting τ equal to 48 – corresponding to windows four years in length with monthly data – there are only 25 possible frequencies to consider.⁶ Similarly, with windows 60 months in length, there are only 32 possible frequencies to consider. A frequency partitioning of X_j using a 60-month window cannot distinguish between variations in X_j with periods greater than 60 months in length – i.e., frequencies smaller than $1/60$ – but this is not a problem unless the important frequency variation in β_j is occurring at frequencies this small or smaller.

Thus, the total cost of this moving window partitioning procedure is a loss of $\tau - 1$ observations at the beginning of the sample (to initiate the first window), a loss of $M - 1$ degrees of freedom (in order to estimate $\beta_{j,1} \dots \beta_{j,M}$ instead of just β_j), and a loss of resolution at frequencies greater than $1/\tau$. This seems a small price to pay in exchange for robustness to feedback and a graceful solution to the problem of choosing the frequency bands.⁷

⁶ For a window 48 periods in length the A matrix of Equation 4 has 48 rows. But this does not imply that there are 48 possible frequencies. The first row corresponds to frequency zero, but rows 2 and 3 both correspond to the same frequency ($1/48$); rows 4 and 5 both correspond to the same frequency ($2/48$), and so forth. Finally, rows 46 and 47 both correspond to a frequency of $23/48$ and row 48 corresponds to a frequency of $24/48$. This yields a total of 25 distinct frequencies.

⁷ Because one expects the coefficient variation across frequencies to be fairly smooth, a more parsimonious approach is to model the variation in the M coefficients by means of a low-order polynomial, as in the distributed lag literature. This approach is discussed and implemented in Ashley and Verbrugge (2006) but omitted here due to space limitations.

Software implementing this moving–window frequency partitioning procedure is available from the authors as a stand–alone Windows program and as a RATS procedure. This software also linearly de-trends the data in each window prior to the filtering, adding the value of the trend at observation τ into the zero–frequency component so that the filtered series still add up to X_j .

It should also be noted that bandpass filters generically yield poor results near the sample endpoints. The standard method for addressing this shortcoming – as originally suggested in Dagum (1978) and described in Stock and Watson (1999) – is (with monthly data) to augment the windowed sample using twelve projected values obtained using an AR(4) model with seasonal dummy variables. The software implements this procedure, yielding filtered values twelve periods away from the end of the window.

The effectiveness and usefulness of the procedure is illustrated in the next two sections.

5 An Example Using Artificially Generated Data

So as to illustrate the results from our procedure, 450 observations were generated from a particular example of the bivariate model considered at the end of Section 2:

$$c_t = \left\{ \frac{1/2}{1 + (y_{t-1} - y_{t-2})^2} \right\} y_{t-1} + .2c_{t-1} + 1/2\epsilon_{c,t} \quad (12)$$

$$y_t = .7y_{t-1} + .2c_{t-1} + \epsilon_{y,t}$$

where the realizations of $\epsilon_{c,t}$ and $\epsilon_{y,t}$ were independent draws from the unit normal distribution. Observe that there is both nonlinear serial dependence and feedback in this specification of the $c - y$ relationship. As noted in Section 2, the coefficient λ_1 on y_{t-1} in the linear model

$$c_t = \lambda_1 y_{t-1} + \lambda_2 c_{t-1} + \epsilon_{c,t} \quad (13)$$

ought consequently to be larger for low frequency fluctuations in y_{t-1} than for high frequency fluctuations.

We used a moving window 60 periods in length to partition y_{t-1} into all 37 possible frequency components,⁸ replaced $\lambda_1 y_{t-1}$ in Equation 13 by $\sum_{j=1}^{37} \lambda_{1,j} y_{t-1}^j$ and estimated $\lambda_{1,1} \dots \lambda_{1,37}$ using OLS. Based on these estimates, the null hypothesis $H_0 : \lambda_{1,1} = \lambda_{1,2} = \dots = \lambda_{1,37}$ can be rejected with p–value .0006. Consequently, our procedure does indeed detect the frequency dependence in

⁸ The 60 months in each window were augmented with the 12 projected values as described at the end of the previous section; that is why there are 36 possible non–zero frequencies and hence 37 frequency components.

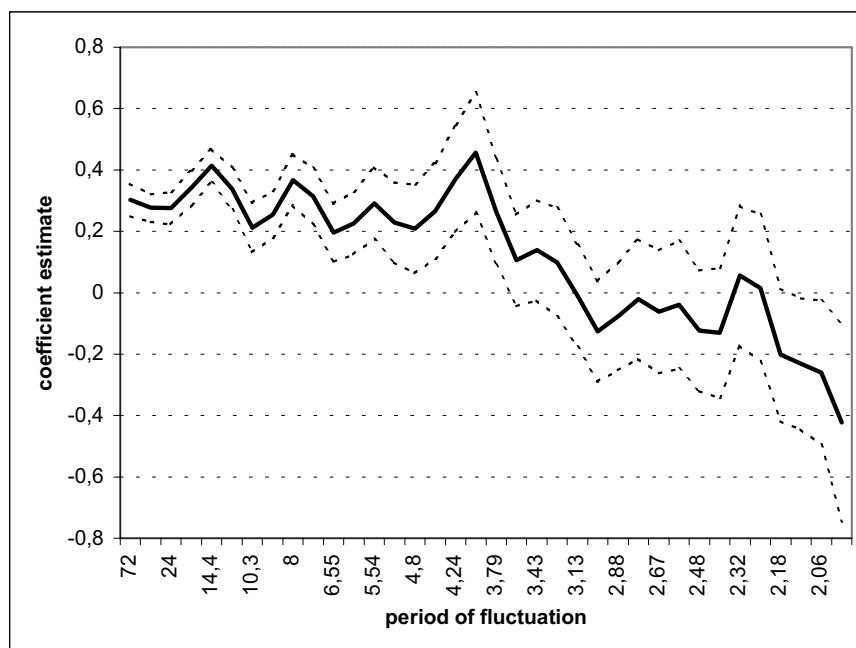


Figure 1. Smoothed parameter estimates $(\hat{\lambda}_{1,1} \dots \hat{\lambda}_{1,37})$ based on Equation 13.

λ_1 induced by the dependence of the marginal propensity to consume in this model on the squared deviation of y_{t-1} from its recent average value.

The 37 OLS parameter estimates, $\hat{\lambda}_{1,1} \dots \hat{\lambda}_{1,37}$ are plotted, plus or minus one estimated standard deviation, in Figure 1. Because these individual estimates display quite a bit of sampling variation, they are smoothed in the plot using a three-point symmetric moving average.⁹ Clearly, the coefficient on y_{t-1} in this regression model is indeed larger for low frequency (i.e., high period) fluctuations in y_{t-1} .

6 An Empirical Example: The Phillips Curve

The Phillips Curve postulates an inverse relationship between inflation and the unemployment rate; it is one of the most-studied relationships in empirical

⁹ That is, the smoothed value of $\hat{\lambda}_{1,j}$ is $\frac{1}{4}\hat{\lambda}_{1,j-1} + \frac{1}{2}\hat{\lambda}_{1,j} + \frac{1}{4}\hat{\lambda}_{1,j+1}$ for $j = 2, \dots, 36$ with double weight put on the central value at the two endpoints – i.e., for j equal to 1 or 37. The coefficient standard error estimates are adjusted accordingly. The 37 frequencies implied by the 72 month rolling window are given by $0, 1/72, 2/72, \dots, 36/72$ and the concomitant periods for the non-zero frequencies are given by $72, 72/2, 72/3, \dots, 2$; see footnote 6.

macroeconomics. A standard Phillips Curve specification is of the form:

$$\pi_t = \alpha + \gamma \text{un}_t + \sum_{i=1}^{12} \delta_i \pi_{t-i} + \Delta Z_t + \epsilon_t \quad (14)$$

where π_t is the inflation rate, un_t is the unemployment rate, and Z_t typically includes seasonal dummy variables and adjustments for structural changes, such as changes in relative energy prices.¹⁰

Estimating an equation of this form using monthly US data from 1980 : 1 to 2003 : 12, we find that the OLS estimate of γ is $-.05 \pm .06$. (Robust (White) standard error estimates are used since there is some evidence for heteroscedasticity in these data.) Thus, ignoring the possibility of frequency dependence in this coefficient, we find (as is typical) that there is no statistically significant Phillips Curve relationship over this sample period.

Decomposing un_t in the manner described in Section 5 above into 37 components – $\text{un}_{t,1} \dots \text{un}_{t,37}$ – using a sequence of 60-month moving windows yields the modified regression model:¹¹

$$\pi_t = \alpha + \sum_{m=1}^{37} \gamma_m \text{un}_{t,m} + \sum_{i=1}^{12} \delta_i \pi_{t-i} + \Delta Z_t + \epsilon_t \quad (15)$$

Estimating Equation 15 using OLS, we find that the null hypothesis that $\gamma_1 \dots \gamma_{37}$ are all equal to zero can be rejected with p-value equal to .018. Thus, once frequency dependence in γ is appropriately allowed for, the Phillips Curve relationship becomes evident. Moreover, we find that the null hypothesis that $\gamma_1 \dots \gamma_{37}$ are all equal can be rejected with p-value equal to .014. Thus, the frequency dependence in γ is statistically significant, providing strong evidence that the kind of nonlinearity discussed here is actually present in the Phillips Curve relationship.

Figure 2 below plots smoothed values of $\hat{\gamma}_1 \dots \hat{\gamma}_{25}$, corresponding to fluctuations in un_t with periods greater than or equal to 3 months, plus and minus one estimated standard deviation. This smoothing – analogous to the smoothing used in estimating a power spectrum from a sample periodogram – is necessary

¹⁰ See Ashley and Verbrugge (2006) for a thorough review of the literature on the Phillips Curve and also for a detailed description of the particular specification and data used in this section. More detailed results are discussed there. One might wonder about the exogeneity of un_t in this model, but this is the standard specification and the use of OLS estimation is common practice in this literature.

¹¹ The 60 months in each window were augmented with the 12 projected values described in footnote 8; that is why there are 36 possible non-zero frequencies and hence 37 frequency components. Similar results are obtained using 60 and 84 month windows. Also, in view of the criticism of linear de-trending in the context of spectral regression given by Corbae, Ouliaris and Phillips (2002), it is notable that the results are not sensitive to how or even whether the within-window de-trending described at the end of Section 5 above is done.

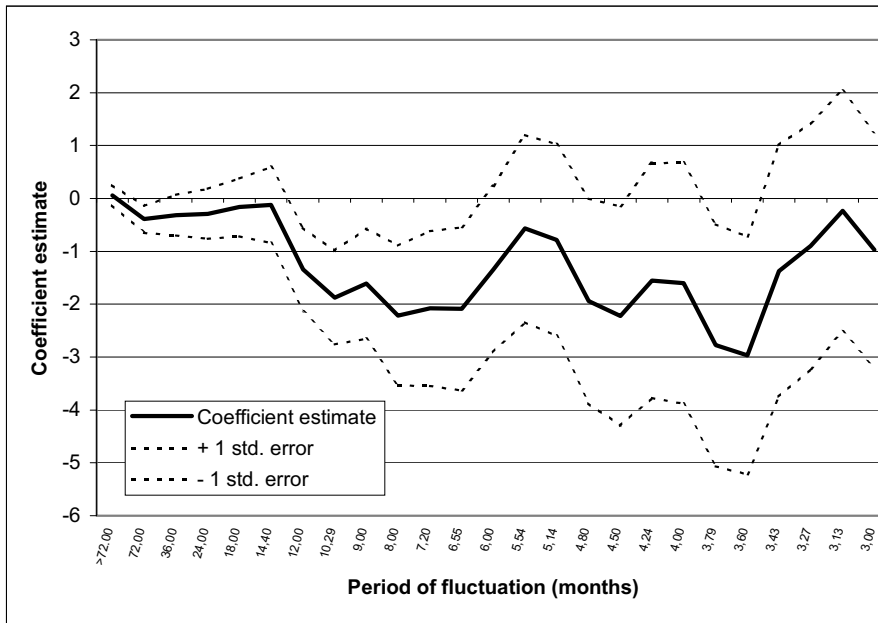


Figure 2. Smoothed parameter estimates ($\hat{\gamma}_1 \dots \hat{\gamma}_{25}$) from Equation 15.

because the $\hat{\gamma}_j$ estimates are individually quite noisy. A seven-point symmetric moving average is used for this purpose and the coefficient standard error estimates are adjusted accordingly. The estimates $\hat{\gamma}_{26} \dots \hat{\gamma}_{37}$ correspond to fluctuations in un_t with periods between 2 and 3 months; these are omitted from the plot because their standard error estimates are so large as to distort the scale of the figure.¹²

Evidently, low frequency fluctuations in un_t – i.e., fluctuations corresponding to periods greater than or equal to around a year – have no impact on inflation, whereas higher frequency fluctuations in un_t have an inverse impact. In particular, $H_0 : \gamma_1 = \dots = \gamma_4 = 0$ – corresponding to un_t fluctuations with periods greater than or equal to 12 months – can be rejected only with a p-value of .381; whereas $H_0 : \gamma_5 = \dots = \gamma_{25} = 0$ – corresponding to un_t fluctuations with periods of 3 to 12 months – can be rejected with a p-value of .019. Associating low frequency un_t fluctuations with changes in the so-called “natural rate of unemployment” and high frequency un_t fluctuations with departures from the natural rate – as in Hall (1999), Cogley and Sargent (2001), and Staiger, Stock and Watson (2001) – these results are consistent with the Friedman–Phelps theory of inflation.

¹² The null hypothesis that these twelve coefficients are all zero cannot be rejected; the p-value for this test is .912.

7 Conclusions

The frequency dependent regression modeling approach proposed here introduces a new class of nonlinear models. This new approach is also exceptionally easy to implement: once the relevant explanatory variable has been partitioned into frequency components, one merely replaces the variable by a weighted sum of these components in whatever estimation framework was already being used. Because these new models are so tightly connected to economic theory, they are both readily interpretable in terms of theory and particularly well-suited for testing economic theory. This point is illustrated with an application to the Phillips Curve, in which the observed frequency dependence in the coefficient on the unemployment variable provides strong empirical support for the Friedman–Phelps theory of inflation, whereas an ordinary regression model fails to detect any relationship between inflation and unemployment whatsoever.

References

- Ashley, R. (1984) “A Simple Test for Regression Parameter Instability.” *Economic Inquiry* 22:253–267.
- Ashley, R., Verbrugge R. (2006) “Mis-Specification in Phillips Curve Regressions: Quantifying Frequency Dependence in This Relationship While Allowing for Feedback.” Mimeo, Economics Department, Virginia Tech.
- Bai, J. (1997) “Estimating Multiple Breaks One at a Time.” *Economic Theory* 13:315–52.
- Bai, J., Perron P. (1998) “Estimating and Testing Linear Models with Multiple Structural Changes.” *Econometrica* 66(1):47-7-8.
- Bai, J., Perron P. (2003) “Computation and Analysis of Multiple Structural Change Models.” *Journal of Applied Econometrics* 18(1):1–22.
- Baxter, M., King R. (1999) “Measuring business cycles: approximate band-pass filters for economic time series.” *The Review of Economics and Statistics* 81:575–593.
- Brown, R. L., Durbin, J., Evans J. M. (1975) “Techniques for Testing the Constancy of Regression Relationships Over Time.” *Journal of the Royal Statistical Society Series B* 37:149–192.
- Chow, G. C. (1960) “Tests of Equality Between Sets of Coefficients in Two Linear Regressions.” *Econometrica* 28:591–605.
- Christiano, L. J., Fitzgerald T. (2003a) “Inflation and monetary policy in the twentieth century.” Federal Reserve Bank of Chicago *Economic Perspectives* 1Q/2003:22–45.
- Christiano, L. J., Fitzgerald T. (2003b) “The band-pass filter.” *International Economic Review* 44.2:435–465.
- Cogley, T., Sargent, T. (2001) “Evolving post–WWII US inflation dynamics.”

- NBER Macroeconomics Annual.
- Comin, D., Gertler, M. (2003) "Medium Term Business Cycles." NBER Working Paper 10003.
- Corbae, D., Ouliaris, S., Phillips, P. C. B.(2002) "Band Spectral Regression with Trending Data" *Econometrica* 70(3):1067–1109.
- Dagum, E. (1978) "Modeling, Forecasting and Seasonally Adjusting Economic Time Series with the X11ARIMA Method", *The Statistician* 27:203-216.
- Den Haan, Wouter, J., Sumner, S. W. (2004) "The Comovements between Real Activity and Prices in the G7." *European Economic Review* 48:1333–1347.
- Engle, R. (1974) "Band Spectrum Regression." *International Economic Review* 15:1–11.
- Engle, R. (1978) "Testing Price Equations for Stability Across Spectral Frequency Bands." *Econometrica* 46:869–881.
- Farley, J. U., Hinich, M.J., McGuire, T. W. (1975) "Some Comparisons of Tests for a Shift in the Slopes of a Multivariate Linear Time Series Model." *Journal of Econometrics* 3:297–318.
- Friedman, M. (1957) *A Theory of the Consumption Function*, Princeton, N.J.: Princeton
- Friedman, M. (1968) "The role of monetary policy." *American Economic Review* 58.1:1–17.
- Hall, R. (1999) "Comment on 'Rethinking the role of the NAIRU in monetary policy: implications of model formulation and uncertainty.'" in (J.B. Taylor, ed.), *Monetary Policy Rules*, NBER Conference Report, Chicago:University of Chicago Press.
- Hannan, E. (1963) "Regression for Time Series." in M. Rosenblatt, ed. *Time Series Analysis*, John Wiley: New York, 14–37.
- Hodrick, R. J., Prescott, E. (1997) "Postwar US business cycles: an empirical investigation." *Journal of Money, Credit and Banking* 29:1–16.
- Modigliani, F., Brumberg, R. H. (1954) "Utility Analysis and the Consumption Function: An Interpretation of Cross-Section Data", in *Post-Keynesian Economics* (Kenneth K. Kurihara, ed.) Rutgers University Press: Brunswick, New Jersey, 388–436.
- Phelps, E. (1967) "Phillips curves, expectations of inflation, and optimal unemployment over time." *Economica* 2.3:22–44.
- Phelps, E. (1968) "Money-wage dynamics and labor market equilibrium." *Journal of Political Economy* 76.2:678–711.
- Staiger, D., Stock, S., Watson M. (2001) "Prices, wages and the US NAIRU in the 1990s", NBER Working Paper No. 8320.
- Stock, J., Watson M. (1999) "Business Cycle Fluctuations in U.S. Macroeconomic Time Series", in (John Taylor and Michael Woodford, eds), *Handbook of Macroeconomics*. Amsterdam: Elsevier, 3–64.
- Tan, H.-B., Ashley, R. (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 Pub-

lishers: Norwell, 129–142.

Tan, H.-B., Ashley. R. (1999b). “Detection and Modeling of Regression Parameter Variation Across Frequencies With an Application to Testing the Permanent Income Hypothesis”, *Macroeconomic Dynamics* 3:69–83.