Department of

PERSISTENCE AND CYCLES
IN THE US FEDERAL FUNDS RATE

Guglielmo Maria Caporale
Brunel University, London, UK

Luis A. Gil-Alana

1. Introduction

The Federal Funds rate is the interest rate at which depository institutions in the US lend each other overnight (normally without a collateral) balances held at the Federal Reserve System (the Fed), which are known as Federal Funds. Such deposits are held in order to satisfy the reserve requirements of the Fed. The rate is negotiated between banks, and its weighted average across all transactions is known as the Federal Funds effective rate. It tends to be more volatile at the end of the reserve maintenance period, the so-called settlement Wednesday, when the requirements have to be met.[1] The Federal Funds target rate is instead set by the Chairman of the Fed according to the directives of the Federal Open Market Committee (FOMC), which holds regular meetings (as well as additional ones when appropriate) to decide on this target. It is therefore a policy rate, used to influence the money supply, and to make the effective rate (which by contrast is determined by the interaction of demand and supply) follow it. Specifically, the Trading Desk of the Federal Reserve Bank of New York conducts open market operations on the basis of the agreed target. This is considered one of the most important indicators for financial markets, whose expectations can be inferred from the prices of option contracts on Federal Funds futures traded on the Chicago Board of Trade.

Given the fact that the Fed implements monetary policy by setting a target for the effective Federal Funds rate which also affects other linked interest rates and the real economy through various transmission channels, it is not surprising that both the theoretical and the empirical literature on this topic are extensive. Theoretical contributions include a well-known paper by Bernanke and Blinder (1988), who propose a model of monetary policy transmission which they then test in a follow-up study (Bernanke and Blinder, 1992) showing that the Federal Funds rate is very useful to forecast real

[1] In empirical studies, therefore, the series is often adjusted to eliminate this effect (see, e.g., Sarno and Thornton (2003).

macroeconomic variables,

with permanent effects of shocks. This is a rather strong assumption that is not justified on theoretical grounds. The classic alternative is to assume that the Federal Funds rate and interest rates in general are stationary I(0) variables

$$y_t = \beta^T z_t + x_t, \quad t = 1, 2, \ldots, \qquad (1)$$

$$(1 - L)^{d_1}(1 - 2\cos w_r L + L^2)^{d_2} x_t = u_t, \qquad (2)$$

where y_t is the observed time series; β is a (kx1) vector of unknown parameters, and z_t is a (kx1) vector of deterministic terms, that might include, for example, an intercept (i.e. z_t = 1) or an intercept with a linear trend ($z_t = (1,t)^T$); L is the lag operator (i.e., $L^s x_t = x_{t-s}$); d is a

Further, note that the polynomial $(1-L)^{d_1}$ can be expressed in terms of its Binomial expansion, such that, for all real d

$$C_{j,d_2}(\mu) = 2\mu\left(\frac{d_2-1}{j}+1\right)C_{j-1,d_2}(\mu) - \left(2\frac{d_2-1}{j}+1\right)C_{j-2,d_2}(\mu), \quad j = 2, 3, \ldots.$$

(see, inter alia, Magnus et al., 1966, or Rainville, 1960, for further details). Gray et al. (1989, 1994) showed that this process is stationary if $d_2 < 0.5$ for $|\mu = \cos w_r| < 1$ and if $d_2 < 0.25$ for $|\mu| = 1$. If $d_2 = 1$, the process is said to contain a unit root cycle (Ahtola and Tiao, 1987; Bierens, 2001); other applications using fractional values of d_2 can be found in Gil-Alana (2001), Anh, Knopova and Leonenko (2004) and Soares and Souza (2006).

In the empirical analysis we use a very general testing procedure to test the model given by equations (1) and (2). It was initially developed by Robinson (1994) on the basis of the Lagrange Multiplier (LM) principle that uses the Whittle function in the frequency domain. It can be applied to test the null hypothesis:

$$H_o: \; d \equiv (d_1, d_2)^T = (d_{1o}, d_{2o})^T \equiv d_o, \qquad (5)$$

in (1) and (2) where d_{10} and d_{20} can

(weekly data). The case of AR(2) disturbances is interesting because it allows to capture the cyclical pattern of the series through a short-memory I(0) process for u_t.[6]

Likelihood Ratio (LR) tests and other likelihood criteria (not reported) suggest that the model with AR(2) disturbances outperforms the others. These results, however, might be biased owing to the long memory in the cyclical structure of the series having been overlooked. Thus, we next consider a model such as (1) and (2) with z_t $(1,t)^T$, i.e., the null model now becomes

$$y_t \;=\; \beta_0 \;+\; \beta_1 t \;+\; x_t, \qquad (9)$$

$$(1-L)^{d_1}(1-2\cos w_r L \;+\; L^2)^{d_2}\, x_t \;=\; u_t, \qquad (10)$$

again with I(0) (potentially ARMA) u_t. The results, for the case of an intercept, which is the most realistic one on the basis of the t-values (not reported), are displayed in Table 2.

[Insert Table 2 about here]

The estimated values of r and thus j = T/r (the number of periods per cycle) for the four series is now close to 8 years. Specifically, j is found to be 8 in the case of the annual data; 97 (and thus 97/12 = 8.089 years) for the monthly data; and 7.57 years (212/28 and 424/56) for bi-weekly and weekly data. This is consistent with the plots of the periodograms displayed in Figure 3. Focusing now on the fractional differencing parameters, it can be seen that d_1 is close to (although below) 1 and d_2 is slightly above 0 for the four series. For d_1 the unit root null is rejected in favour of mean reversion in the case of annual, bi-weekly and weekly data; however, for monthly data, even though d_1 is still below 1, the unit root null cannot be rejected at conventional significance levels. As for the cyclical fractional differencing parameter, d_2, is estimated to be 0.094 in the annual case and the I(0) null hypothesis cannot be rejected. In the remaining three cases, d_2 is significantly above 0 (thus displaying long memory), ranging from 0.145 (weekly data) to

[6] The estimates of the AR(2) coefficients (not reported) were in all cases in the complex plane, which is consistent with the cyclical pattern observed in the data.

0.234 (monthly data). Very similar values for d_1 and d_2 are obtained in the case of autocorrelated disturbances; LR and no-autocorrelation tests strongly support the white noise specification for u_t for each of the four series.[7]

Finally, we investigate which of the two specifications (the I(d) one with AR(2) disturbances or the one with the two fractional differencing structures) has a better in-sample performance, and also better forecasting properties. For the first of these two purposes we employ several goodness-of-fit measures based on the likelihood function. For the forecasting experiment, we use instead various statistics including the modified Diebold and Mariano (1995) (M-DM) statistic. Remember that the two models considered are:

$$y_t = \beta_0 + x_t; \qquad (1-L)^{d_1} x_t = u_t; \qquad u_t = \phi_1 u_{t-1} + \phi_2 u_{t-2} + \varepsilon_t, \qquad \text{(M1)}$$

and

$$y_t = \beta_0 + x_t; \quad (1-L)^{d_1}(1-2\cos w_r L + L^2)^{d_2} x_t = \varepsilon_t, \qquad \text{(M2)}$$

and therefore they differ in the way the cyclical component is modelled, model (M1) and (M2) adopting respectively an AR(2) process and a Gegenbauer (fractional) specification for the d_1-differenced (

precisely that for the annual series. Other likelihood criteria (AIC and SIC) lead essentially to the same conclusions.[8]

Next we focus on the forecasting performance of the two models. For this purpose we calculate one- to twenty-step ahead forecasts over 20 periods for each of the four series at different data frequencies. The forecasts were constructed according to a recursive procedure conditionally upon information available up to the forecast date which changes recursively.

We computed the Root Mean Squared Errors (RMSE) and the Mean Absolute Deviation (MAD) for the two specifications of each series. The results (not reported here for reasons of space, but available from the authors upon request) indicate that the fractional structure outperforms the AR(2) model in practically all cases.

However, the above two criteria and other methods such as the Mean Absolute Prediction Error (MAPE), Mean Squared Error (MSE), etc., are purely descriptive devices.[9] Several statistical tests for comparing different forecasting models are now available. One of them, widely employed in the time series literature, is the asymptotic test for a zero expected loss differential due to Diebold and Mariano (1995).[10] Harvey, Leybourne and Newbold (1997) note that the Diebold-Mariano test statistic could be seriously over-sized as the prediction horizon increases, and therefore provide a modified Diebold-Mariano test statistic given by:

$$M - DM = DM \sqrt{\frac{n + 1 - 2h + h(h-1)/n}{n}},$$

[8] Note, however, that these criteria might not necessarily be the best criteria in applications involving

where DM is the original Diebold-Mariano statistic, h is the prediction horizon and n is the time span for the predictions. Harvey et al. (1997) and Clark and McCracken (2001) show that this modified test statistic performs better than the DM test statistic, and also that the power of the test is improved when p-values are computed with a Student t-distribution.

We further evaluate the relative forecast performance of the different models by making pairwise comparisons based on the M-DM test statistic. We consider 5, 10, 15, 20 and 25-period ahead forecasts. The results are displayed in Table 3.

[Insert Table 3 about here]

They show that for the 5-step and 10-step ahead predictions it cannot be inferred that one model is statistically superior to the other. By contrast, over longer horizons there are several cases where the fractional model (M2) outperforms (M1). However, these forecasting methods may have very low power under some circumstances, especially in the case of non-linear models (see, e.g., Costantini and Künst, 2011). Thus, these results should be taken with caution.

4. Conclusions

This paper uses long-range dependence techniques to analyse two important features of the US Federal Funds effective rate, namely its persistence and cyclical behaviour. In particular, it examines annual, monthly, bi-weekly and weekly data, from 1954 until 2010. The main results are the following. Whstithe folz1 85.104 484.23da17,e40 1 45.104 235.85 head pr

in4(s)-1u7

second model considered uses a Gegenbauer-type of process for the cyclical component, and therefore has two fractional differencing parameters, one corresponding to the long-run or zero frequency (d_1), and the other to the cyclical structure (d_2). When using this specification the results indicate that the order of integration at the zero frequency ranges between 0.802 (bi-weekly frequency) and 0.966 (monthly), whilst that of the cyclical component ranges between 0.094 (annual) and 0.234 (bi-weekly). Both the in-sample and out-of-sample evidence suggest that the long memory model with two fractional structures (one at zero and the other at the cyclical frequency) outperforms the other models.

Our results are not directly comparable to those of Sarno et al. (2005), who model the difference between the effective and the target rate, whilst we focus only on the former. Nevertheless, our analysis, based on letting the data speak by themselves to find the most suitable specification, produces valuable evidence for interest rate modelling, since it shows that an I(d) specification including a cyclical component outperforms both classical I(0) and simple I(d) models. This confirms the importance of adopting an econometric framework such as the one chosen here, which explicitly takes into account both persistence and cyclical patterns, to model the behaviour of the US Federal Funds effective rate and interest rates in general.

Appendix

The test statistic proposed by Robinson (1994) for testing H_o (5) in the model given by equations (1) and (2) is given by:

$$\hat{R} = \frac{T}{\hat{\sigma}^4} \hat{a}' \hat{A}^{-1} \hat{a},$$

where T is the sample size, and

$$\hat{a} = \frac{-2\pi}{T} \sum_j^{*} \psi(\lambda_j) g_u(\lambda_j; \hat{\tau})^{-1} I(\lambda_j); \qquad \hat{\sigma}^2 = \sigma^2(\hat{\tau}) = \frac{2\pi}{T} \sum_{j=1}^{T-1} g_u(\lambda_j; \hat{\tau})^{-1} I(\lambda_j),$$

$$\hat{A} = \frac{2}{T} \left(\sum_j^{*} \psi(\lambda_j)\psi(\lambda_j)' - \sum_j^{*} \psi(\lambda_j)\hat{\varepsilon}(\lambda_j)' \left(\sum_j^{*} \hat{\varepsilon}(\lambda_j)\hat{\varepsilon}(\lambda_j)' \right)^{-1} \sum_j^{*} \hat{\varepsilon}(\lambda_j)\psi(\lambda_j)' \right)$$

$$\psi(\lambda_j)' = [\psi_1(\lambda_j), \psi_2(\lambda_j)]; \qquad \hat{\varepsilon}(\lambda_j) = \frac{\partial}{\partial \tau} \log g_u(\lambda_j; \hat{\tau}); \qquad \psi_1(\lambda_j) = \log \left| 2 \sin \frac{\lambda_j}{2} \right|;$$

with $\lambda_j = 2\pi j/T$, and the summation in * is over all frequencies which are bounded in the spectrum. $I(\lambda_j)$ is the periodogram of $\hat{u}_t = (1-L)^{d_{1o}} (1 - 2\cos w_r L + L^2)^{d_{2o}} y_t - \hat{\beta}' \bar{z}_t$, with

$$\hat{\beta} = \left(\sum_{t=1}^{T} \bar{z}_t \bar{z}_t' \right)^{-1} \sum_{t=1}^{T} \bar{z}_t (1-L)^{d_{1o}} (1 - 2\cos w_r L + L^2)^{d_{2o}} y_t;$$

$\bar{z}_t = (1-L)^{d_{1o}} (1 - 2\cos w_r L + L^2)^{d_{2o}} z_t$, evaluated at $\lambda_j = 2\pi j/T$ and $\hat{\tau} = \arg\min_{\tau \in T^*} \sigma^2(\tau)$, with T^* as a suitable subset of the R^q Euclidean space. Finally, the function g_u above is a known function coming from the spectral density of u_t:

$$f(\lambda) = \frac{\sigma^2}{2\pi} g_u(\lambda; \tau), \qquad -\pi < \lambda \leq \pi.$$

Note that these tests are purely parametric and, therefore, they require specific modelling assumptions about the short-memory specification of u_t. Thus, if u_t is white noise, $g_u \equiv 1$, and if u_t is an AR process of the form $\phi(L)u_t = \varepsilon_t$, $g_u = |\phi(e^{i\lambda})|^{-2}$, with $\sigma^2 = V(\varepsilon_t)$, so that the AR coefficients are a function of τ.

Clarida, R.H., Gali, J. and M. Gertler (2000), “Monetary policy rules and macroeconomic stability: evidence and some theory”, Quarterly Journal of Economics, 115, 147-180.

Clarida, R.H., Sarno, L., Taylor, M. and L. Valente (2006), “The role of asymmetries and regime shifts in the term structure of interest rates”, Journal of Business, 79, 3, 1193-1224.

Clark, T.E. and M.W. McCracken, 2001, Tests of forecast accuracy and encompassing for nested models, Journal of Econometrics 105, 85-110.

Costantini, M. and R.M. Künst, 2011, On the usefulness of the Diebold and Mariano test in the selection of prediction models, Econometric Series 276, Institute for Advanced Studies.

Couchman, J., R. Gounder and J.J. Su, 2006, Long memory properties of real interest rates for 16 countries, Applied Financial Economics Letters 2, 25-30.

Dahlhaus, R. (1989), “Efficient Parameter Estimation for Self-similar Process”, *Annals of Statistics*, 17, 1749-1766.

Diebold, F.X. and R.S. Mariano, 1995, Comparing predictive accuracy, Journal of Business, Economics and Statistics 13, 253-263.

Diebold, F.X. and G.D. Rudebusch, 1989, Long memory and persistence in the aggregate output. Journal of Monetary Economics 24, 189-209.

Gil-Alana, L.A., 2001, Testing stochastic cycles in macroeconomic time series. Journal of Time Series Analysis 22, 411-430.

Gil-Alana, L.A. and P.M. Robinson, 1997, Testing of unit roots and other nonstationary hypotheses in macroeconomic time series. Journal of Econometrics 80, 241-268.

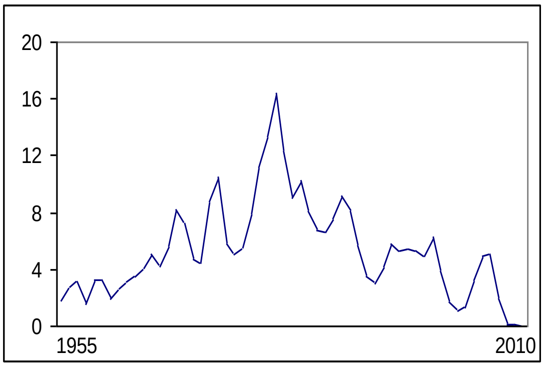
Gray, H.L., Yhang, N. and Woodward, W.A., 1989, On generalized fractional processes, Journal of Time Series Analysis 10, 233-257.

Gray, H.L., Yhang, N. and Woodward, W.A., 1994, On generalized fractional processes. A correction, Journal of Time Series Analysis 15, 561-562.

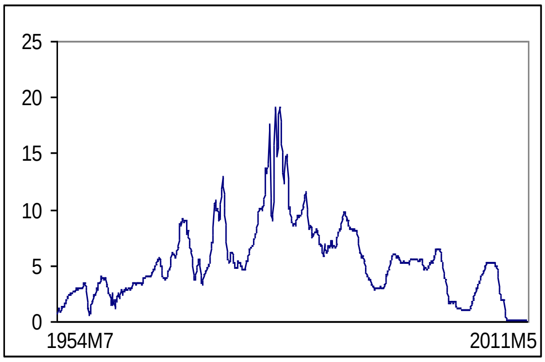
Sarno, L. and D.L. Thornton (2003), “The dynamic relationship between the Federal Funds rate and the Treasury bill rate: an empirical investigation”, Journal of Banking and Finance, 27, 1079-1110.

Figure 1: Original time series data

Annual data



Monthly data



Bi-Weekly data

Weekly data

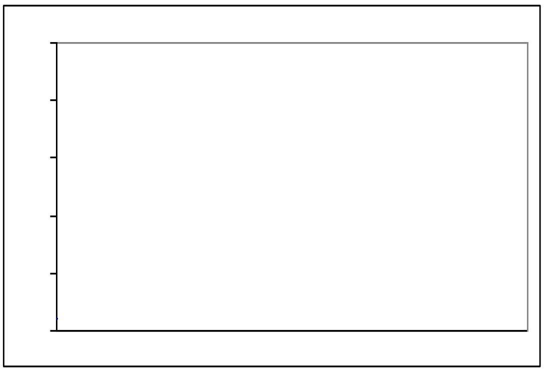
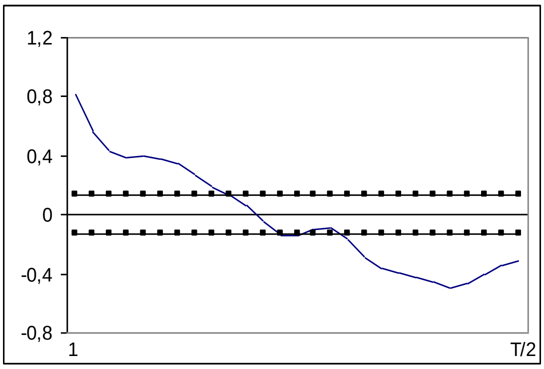
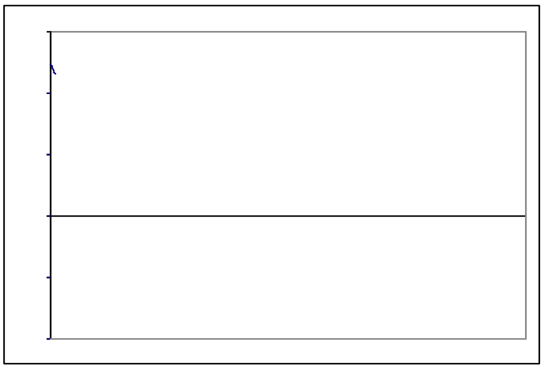
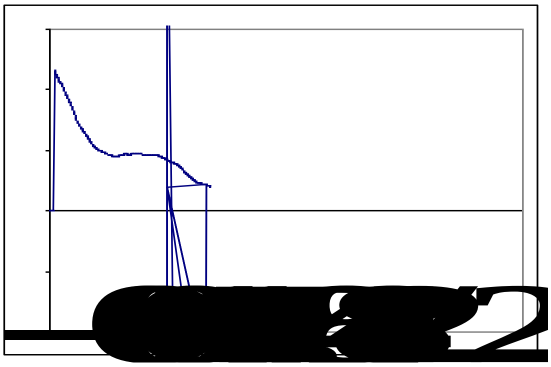
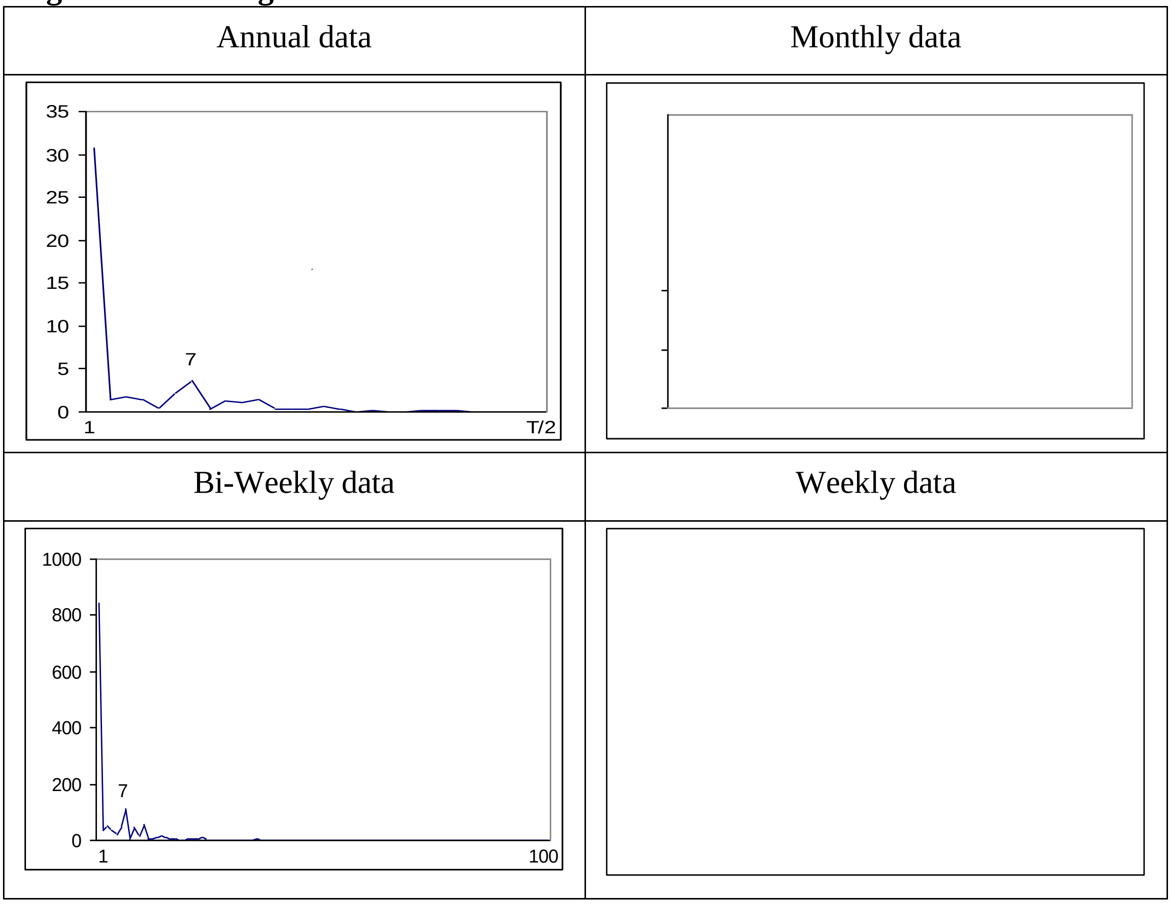


Figure 2: Correlogram of the time series

Annual data	Monthly data
	
Bi-Weekly data	Weekly data
	

Note: The thick lines refer to the 95% confidence band for the null hypothesis of no autocorrelation.

Figure 3: Periodogram of the time series



Note: The horizontal axis refers to the discrete Fourier frequencies $\lambda_j = 2\pi j/T$, j = 1, …, T/2.

Table 1: Estimates of d and 95% confidence interval in an I(d) model with an intercept

	White noise	AR(1) disturbances	AR(2) disturbances
Annual	0.937 (0.704, 1.450)	0.544 (0.429, 0.700)	0.722 (0.334, 1.495)
Monthly	1.277 (1.189, 1.383)	0.821 (0.742, 0.913)	0.852 (0.679, 1.016)
Bi-Weekly	1.168 (1.122, 1.213)	1.025 (0.891, 1.146)	0.824 (0.633, 1.008)
Weekly	0.973 (0.954, 0.994)	1.086 (1.044, 1.127)	1.045 (0.984, 1.101)

The values are Whittle estimates of d in the frequency domain (Dahlhaus, 1989). Those in parentheses are the 95% confidence interval of the non-rejection values of d using Robinson (1994).

Table 2: Estimates of d_1 and d_2 in the model with two fractional structures

Frequency	r (j)	d_1	d_2
Annual	j = 7 (r = 8)	0.932 (0.561, 0.983)	0.094 (-0.008, 0.233)
Monthly	j = 683 (r = 97)	0.966 (0.895, 1.128)	0.145 (0.109, 0.217)
Bi-Weekly	j = 1486 (r = 212)		