4.11 Testing for Persistence

We test for the presence of long-range dependence using the modified rescaled range test and we quantify persistence by estimating the fractional integration parameter using a Bayesian ARFIMA model. The question at hand is how to distinguish between short-range and long-range dependence?

The most widely used notion of short-range dependence is the concept of 'strong mixing' due to Rosenblatt (1956). It measures the decline of statistical dependence between events separated by successively longer spans of time. As the time span increases and the maximal dependence between events becomes trivially small, then the time series is a strong-mixing one, such as the class of ARMA models wherein the autocorrelations decay exponentially. Dependence between events over a long span defines long-range dependence, such as long-memory processes (or fractionally integrated processes given the definition in equation (4.9)).

4.11.1 The Rescaled Range Statistic (R/S)

Originally due to Hurst (1951), the rescaled range statistic is set to detect long-range dependence. It is defined as R_T/s_T,

$$R_T \equiv \max_{0\leq k\leq T}\left\{\sum_{j=1}^{k}(y_j-\overline{y})\right\} - \min_{0\leq k\leq T}\left\{\sum_{j=1}^{k}(y_j-\overline{y})\right\} \tag{4.19}$$

$$s_T = \left\{\frac{1}{T}\sum_{t=1}^{T}(y_t-\overline{y})^2\right\}^{0.5} \tag{4.20}$$

where R is the range, s_T is the sample standard deviation, and $\overline{y}$ denotes the sample mean. Lo (1991, pp. 1287-1288) showed that $T^{-1/2}R_T/s_T$ is asymptotically distributed as the range of a standard Brownian Bridge on the unit interval and has expectation of $(\pi/2)^{1/2} = 1.253$ and a standard deviation of $[(\pi/2)(\pi-3)/3]^{1/2} = 0.272$.

The most important shortcoming of the rescaled range is its sensitivity to short-range dependence. For example, if the process is an AR(1), i.e., purely short-range dependent, then the mean of the rescaled range limiting distribution will be biased.

To counter and to correct for the impact of short-range dependency on the test statistic, Lo (1991, p. 1289) proposed using a modified rescaled range statistic. By correcting for short-range dependency, the limiting distribution of the modified statistic is invariant to many forms of short-range dependency but sensitive to the presence of long-range dependency. The modified statistic is robust to many forms of heterogeneity and weak dependence. Also, it is able to discriminate between short- and long-range dependency.

The modified rescaled range statistic is defined as,

$$Q_T \equiv \frac{R_T}{\sigma_T(q)} \tag{4.21}$$

where,

$$\sigma_T^2(q) = c_0 + 2\sum_{j=1}^{q} w_j(q)c_j \tag{4.22}$$

c_j denotes the jth order sample autocovariance of y_t and $w_j(q)$ are the Newey and West (1987) weights using a Bartlett window defined as,

$$w_j(q) = 1 - \left[\frac{j}{q+1}\right] \qquad q < T \tag{4.23}$$

In the presence of long memory, the normalized statistic $T^{-1/2}Q_T$ weakly converges to the range of a Brownian Bridge. The distribution is given by,

$$F(x) = \sum_{j=-\infty}^{\infty} (1 - 4x^2 j^2) \exp[-2x^2 j^2] \tag{4.24}$$

This distribution is positively skewed and its fractiles are tabulated in Lo (1991, p. 1288). The modified rescaled range statistic is robust to short-range dependence and consistent with a general class of long-range dependent stationary Gaussian alternatives (see Baillie (1996, p. 28)).

The choice of q is a subject open to debate. For our analysis, since we are using quarterly data, we computed the modified statistic at $q = 1, 2, 3, 4, 5, 6, 7, 8$ and $q = [k_T]$, where $[k_T]$ denotes the greatest integer less than or equal to k_T. As defined and proposed by Lo (1991, p. 1302), $[k_T]$ is a data-dependent approach for the choice of q,

$$k_T \equiv \left(\frac{3T}{2}\right)^{1/3} \left(\frac{2\widehat{\rho}}{1 - \widehat{\rho}^2}\right)^{2/3} \tag{4.25}$$

where $\widehat{\rho}$ is the estimated first-order autocorrelation coefficient of the data.

The following table (Table 4.E) reports the results of the modified rescaled range (Q_T) statistic for total unemployment, manufacturing and services unemployment. For each q, the first column reports the $\sum_{j=1}^{q} w_j(q) c_j$, i.e., the sum of the weighted autocovariances. Subsequent columns report the logarithm of Q_T and the normalized test statistic value $\frac{Q_T}{\sqrt{T}}$.

Given the reported critical values in Lo (1991, p. 1288), we test the null hypothesis of a simple i.i.d. process. All series are in log form and y_t denotes the log of the time series. Table 4.E computes the normalized test statistic values for Δy_t and Table 4.F computes the same statistic for the Hodrick-Prescott filtered y_t. Note that the normalizing factor $\sqrt{T}$ is different in both tables. For Tables 4.E and 4.F, the sample size is 91 and 92 observations, respectively. The reason for computing both tables is to investigate the sensitivity of the modified test statistic results to the method of detrending. Also, to check the sensitivity of the statistic to the lag length, the normalized test statistic is computed for several different values of q. Given that the normalized test statistic follows a Brownian Bridge process, the null hypothesis is examined at the 95 percent confidence level by not rejecting or rejecting according to whether the normalized test value is or is not contained in the interval [0.809 , 1.862].

Table 4.E significantly rejects the simple null hypothesis at most values of q. Long-range dependence is evident in Canadian total, manufacturing and services unemployment. Table 4.F gives similar results. However, persistence of total unemployment is less evident at the data-dependent value of q. Shorter values of q are picking up the short-range dependence. Using the Hodrick-Prescott filter increases the q lag where the first evidence of persistence is reported. For example, evidence of persistence for total unemployment is first reported at $q = 4$ when using Δy_t and at $q = 5$ when using HP filtered y_t. This one lag delay holds for total and manufacturing unemployment. For services unemployment, the lag delay is longer.

Given the strong evidence of long-range dependence in the series Δy_t, we decided to continue our analysis of long-range dependence. The next section proposes a Bayesian approach to estimate several ARFIMA models in order to quantify the fractional integration parameter.

Table 4.E

Modified Range over Standard Deviation (R/S) Test Statistic (Lo (1991))
The change of the log of Quarterly Canadian Unemployment Time Series

Total Unemployment

	Sum of Weights	Log(R/S)	Normalized Test Statistic
q=1	0.00083	0.99843	1.04450
q=2	0.00132	0.94306	0.91947
q=3	0.00170	0.90835	0.84884
q=4‡	0.00197	0.88583	0.80594*
q=5	0.00215	0.87287	0.78226*
q=6	0.00227	0.86382	0.76612*
q=7	0.00236	0.85764	0.75530*
q=8	0.00244	0.85233	0.74613*

Manufacturing Unemployment

	Sum of Weights	Log(R/S)	Normalized Test Statistic
q=1	0.00095	1.00146	1.05182
q=2	0.00257	0.86828	0.77404*
q=3	0.00468	0.76656	0.61241*
q=4	0.00717	0.68730	0.51025*
q=5	0.00991	0.62423	0.44127*
q=6	0.01282	0.57275	0.39195*
q=7‡	0.01585	0.52954	0.35483*
q=8	0.01899	0.49237	0.32572*

Services Unemployment

	Sum of Weights	Log(R/S)	Normalized Test Statistic
q=1	0.00306	0.89904	0.83085
q=2	0.00462	0.84992	0.74198*
q=3‡	0.00560	0.82397	0.69896*
q=4	0.00616	0.81050	0.67761*
q=5	0.00619	0.80973	0.67641*
q=6	0.00600	0.81438	0.68369*
q=7	0.00578	0.81970	0.69211*
q=8	0.00565	0.82271	0.69692*

‡ : Denotes the value for [k_T]

* : Indicates significance at the 5 percent level.

Table 4.F

Modified Range over Standard Deviation Test Statistic (Lo (1991)) HP Filtered Log of Quarterly Canadian Unemployment Time Series			
Total Unemployment			
	Sum of Weights	Log(R/S)	Normalized Test Statistic
q=1	0.00432	1.08805	1.27690
q=2	0.00837	1.00475	1.05403
q=3	0.01201	0.95025	0.92972
q=4	0.01513	0.91232	0.85197
q=5	0.01771	0.88538	0.80073*
q=6	0.01975	0.86611	0.76598*
q=7	0.02132	0.85241	0.74219*
q=8	0.02247	0.84294	0.72618*
q=25‡	0.01193	0.95127	0.93192
Manufacturing Unemployment			
	Sum of Weights	Log(R/S)	Normalized Test Statistic
q=1	0.00755	0.98624	1.01006
q=2	0.01415	0.90783	0.84321
q=3	0.01951	0.86003	0.75534*
q=4	0.02348	0.83040	0.70551*
q=5	0.02608	0.81297	0.67775*
q=6	0.02764	0.80315	0.66260*
q=7	0.02854	0.79771	0.65435*
q=8	0.02906	0.79460	0.64969*
q=14‡	0.02810	0.80035	0.65834*
Services Unemployment			
	Sum of Weights	Log(R/S)	Normalized Test Statistic
q=1	0.00310	1.06064	1.19881
q=2	0.00598	0.97848	0.99216
q=3	0.00853	0.92493	0.87708
q=4	0.01069	0.88799	0.80555*
q=5	0.01244	0.86213	0.75898*
q=6	0.01380	0.84391	0.72781*
q=7	0.01481	0.83134	0.70704*
q=8	0.01551	0.82308	0.69373*
q=21‡	0.00974	0.90338	0.83462*

‡ : Denotes the value for $[k_T]$

* : Indicates significance at the 5 percent level.