A Diagnostic for Seasonality Based Upon Autoregressive Roots

Tucker McElroy
U.S. Census Bureau

NTTS, March 13, 2019

Disclaimer

These slides are released to inform interested parties of research and to encourage discussion. The views expressed on statistical issues are those of the author and not necessarily those of the U.S. Census Bureau.

Outline

1. Background on Seasonality

2. Criteria for a Diagnostic of Seasonality

3. Persistent Oscillations

4. Seasonality Hypothesis Testing

5. Simulation Evidence

6. Data Application

Background on Seasonality

Seasonality in Official Time Series. Many official time series – such as gross domestic product (GDP) and unemployment rate data – have an enormous impact on public policy, and the seasonal patterns often obscure the long-run and mid-range dynamics.

What is Seasonality? Persistency in a time series over seasonal periods that is not explainable by intervening time periods.

- Requires persistence year to year

- Non-seasonal trending series have persistence, which comes through intervening seasons – we must screen out such cases

Background on Seasonality

The Seasonal Adjustment Task. Given a raw time series:

1. Does it have seasonality? If so, seasonally adjust.

2. Does the seasonal adjustment have seasonality? If not, publish.

Both these tasks require a seasonality diagnostic, although the properties of a time series before and after seasonal adjustment can be quite different.

Background on Seasonality

Pre-Testing. Testing for seasonality in a raw series, where the seasonality could be deterministic (stable), moving and stationary (dynamic), or moving and non-stationary (unit root). These categories are not mutually exclusive, e.g., we could have both unit root and deterministic seasonality.

Post-Testing. Testing for seasonality in a seasonally adjusted series, where the seasonality typically will only be dynamic. However, forecast-extension used in filtering introduces local non-stationarity to the beginning and end of the series.

Criteria for a Diagnostic of Seasonality

1. Rigorous statistical theory

2. Precise correspondence between seasonal dynamics and the diagnostic

3. Applicable to diverse sampling frequencies

4. Applicable to multiple frequencies of seasonal phenomena

5. Ability to assess over- and under-adjustment

Criteria for a Diagnostic of Seasonality

Correspondence. Diagnostic takes on a high value if and only if a high degree of seasonality is present (at a frequency of interest). We don't want high values occuring when seasonality is not present (spurious flagging of seasonality), or low values when seasonality is present (failure to detect).

QS. The QS diagnostic is based on autocorrelation at seasonal lags. These can take on high values for a non-seasonal AR(1) process, generating spurious indications of seasonality.

Criteria for a Diagnostic of Seasonality

Daily Time Series. A non-standard sampling frequency for official statistics. Also, there may be weekly seasonality (frequencies $2\pi j/7$ for $j = 1, 2, 3$) – corresponding to "trading day" – and annual seasonality (frequency $2\pi/365.25$ and integer multiples).

Autocorrelations. Only available at integer lags, hence not helpful for fractional periods like 365.25.

Criteria for a Diagnostic of Seasonality

Over- and Under-Adjustment. Too much seasonality removed (over-) versus too little (under-), described by Nerlove (1964) and others.

- Over-adjusment generates dips in the spectral density, corresponding to oscillatory effects in the inverse autocorrelations.

- Under-adjustment leaves peaks in the spectral density, corresponding to oscillatory effects in the autocorrelations.

Persistent Oscillations

Main Contribution. Associate persistence in a stationary time series to the presence of strong roots in its autoregressive polynomial.

1. Distribution theory for ARMA processes

2. High values of ρ (persistence) correspond to oscillatory effects in the Wold coefficients, and hence the autocorrelations

3. Adapts to any sampling or seasonal frequency (non-integer periods are fine)

4. Over-adjustment assessed through presence of strong roots in the moving average polynomial

Persistent Oscillations

Consider sequences $\{\psi_j\}_{j \geq 0}$ and their associated z-transforms $\psi(z) = \sum_{j \geq 0} \psi_j \, z^j$.

Heuristics. If $\{\psi_j\}$ satisfies a homogeneous Ordinary Difference Equation given by $\pi(B)\psi_j = 0$ for some polynomial $\pi(z)$, then $\psi_j = \sum_k a_k \, \zeta_k^{-j}$ for coefficients a_k and distinct roots ζ_k of $\pi(z)$. Then $\psi(z) = 1/\pi(z)$, and

$$\psi(r^{-1} \, e^{i\omega}) = \sum_k a_k \sum_{j \geq 0} (r^{-1} \, e^{i\omega} / \zeta_k)^j.$$

The modulus will be large if one of the roots ζ_k is close to $r^{-1} \, e^{i\omega}$ and if the corresponding a_k is not small.

Persistent Oscillations

Hence for a fixed ω, we can assess oscillatory effects as a function of persistence r.

Definition 1. *A sequence* $\{\psi_j\}_{j \geq 0}$ *has* ρ*-persistent oscillatory effects of frequency* $\omega \in [-\pi, \pi]$ *(where* $\rho \in (0,1]$*) if and only if* $|\psi(r^{-1} e^{i\omega})|$ *is maximized over* $r \in (0,1]$ *at* $r = \rho$*.*

Proposition 1. *A causal invertible ARMA(*p*,*q*) process with MA polynomial* $\theta(z)$ *and AR polynomial* $\phi(z)$ *has a causal representation with MA coefficients* $\{\psi_j\}$ *having a* ρ*-persistent oscillatory effect of frequency* ω *if and only if* $|\phi(r^{-1} e^{i\omega})/\theta(r^{-1} e^{i\omega})|$ *is minimized over* $r \in (0,1]$ *at* $r = \rho$*.*

Persistent Oscillations

The strongest types of oscillatory effects are those such that $|\psi(\rho^{-1} e^{i\omega})| = \infty$, or where $\pi(\rho^{-1} e^{i\omega}) = 0$; this occurs if and only if $\rho^{-1} e^{i\omega}$ is a root of $\pi(z)$. Such an oscillatory effect is said to be a seasonal effect.

Definition 2. *A process has ρ-persistent seasonality of frequency $\omega \in [-\pi, \pi]$ (where $\rho \in (0, 1]$) if and only if its causal representation has coefficients $\{\psi_j\}$ with a ρ-persistent oscillatory effect of frequency $\omega \in [-\pi, \pi]$, such that $\pi(\rho^{-1} e^{i\omega}) = 0$, where $\pi(z) = 1/\psi(z)$.*

Remark. The autocovariance generating function is $\gamma(z) \propto \psi(z)\, \psi(z^{-1})$, and hence the oscillatory effects of $\{\psi_j\}$ are directly inherited (via Definition 1) by the autocovariance function $\{\gamma_h\}$.

Persistent Oscillations

Example: SAR(1). The Seasonal Autoregressive process of order 1 and period s corresponds to $\phi(z) = 1 - b\, z^s$ for $b \in (0,1)$. The AR roots are $\zeta_k = b^{-1/s}\, e^{i 2\pi k/s}$ for $1 \leq k \leq s$, and $a_k = 1/s$ so that $\psi_j = s^{-1} \sum_{k=1}^{p} \zeta_k$. Hence

$$\left|\phi(r^{-1} e^{i\omega})\right|^2 = 1 + b^2\, r^{-2s} - 2b\, r^{-s}\, \cos(\omega s) = \left(1 - b\, r^{-s}\right)^2$$

when ω is a seasonal frequency (of the form $\omega = 2\pi j/s$ for some integer j), and this quantity is zero at $r = b^{1/s}$, or the reciprocal modulus of all the AR roots. Hence, the SAR(1) has $b^{1/s}$-persistent seasonality for each $\omega = 2\pi j/s$; for $b = .7$ and $s = 12$, the persistence is $.97$.

Persistent Oscillations

Extension to Anti-Persistence. Anti-seasonality, or anti-persistence of a seasonal phenomenon, is useful for diagnosing improperly constructed seasonal adjustment filters: if the filter does too much smoothing to suppress the seasonality, then there will be large troughs in the filter's squared gain function. (Note that optimal filters (Tukey, 1978) have this effect.)

Definition 3. *A process has ρ-persistent anti-seasonality of frequency $\omega \in [-\pi, \pi]$ (where $\rho \in (0, 1]$) if and only if its invertible representation has coefficients $\{\pi_j\}$ with a ρ-persistent oscillatory effect of frequency $\omega \in [-\pi, \pi]$, such that $\psi(\rho^{-1} e^{i\omega}) = 0$, where $\psi(z) = 1/\pi(z)$.*

Persistent Oscillations

Connection of Anti-Persistence to Inverse Autocorrelations. Oscillations in $\{\pi_j\}$ are governed by the roots of $\psi(z)$, since $\psi(B)\pi_j = 0$. These oscillations are also present in the inverse autocovariance function $\{\xi_h\}$, which has z-transform given by

$$\xi(z) \propto \pi(z)\,\pi(z^{-1}).$$

ARMA Summary. So for an invertible ARMA process with $\psi(z) = \theta(z)/\phi(z)$ and $\pi(z) = \phi(z)/\theta(z)$, the AR roots govern oscillations in $\{\psi_j\}$ and autocorrelations, whereas the MA roots govern oscillations in $\{\pi_j\}$ and the inverse autocorrelations.

Seasonality Hypothesis Testing

Testing Methodology. From Definition 2 we see that ρ governs a null hypothesis about the process' seasonality. Suppose an invertible ARMA model has been identified and fitted to the (differenced) data. For any given ω, the null hypothesis is that

$$H_0(\rho_0) : \pi(r^{-1}e^{i\omega}) = 0 \;\text{ has solution }\; r = \rho_0.$$

Note that $H_0(\rho_0)$ holds if and only if $\phi(r^{-1}e^{i\omega}) = 0$ has solution $r = \rho_0$. Let

$$g(r) = \left|\pi(r^{-1}e^{i\omega})\right|^2,$$

which measures departures from ρ_0-persistent seasonality.

Seasonality Hypothesis Testing

Test Statistic. Compute an estimate of $g(r)$ based upon maximum likelihood estimates (MLEs) of the ARMA parameters:

$$\widehat{g}(r) = \left|\widehat{\pi}(r^{-1}e^{i\omega})\right|^2.$$

For a sample of size T, our test statistic of $H_0(\rho_0)$ is

$$T\,\widehat{g}(\rho_0).$$

Seasonality Hypothesis Testing

Theorem 1. *Let* $\{X_t\}$ *be an invertible ARMA(*p*,*q*) process with i.i.d. inputs and autoregressive polynomial* $\phi(z)$*, and moving average polynomial* $\theta(z)$*. With* $\pi(z) = \phi(z)/\theta(z)$ *and* $\widehat{g}(r)$ *defined above, where the MLEs for the ARMA parameters are obtained from a sample of size* T*, it follows that when* $g(r) = 0$

$$T\,\widehat{g}(r) \stackrel{\mathcal{L}}{\Longrightarrow} \frac{\left|\underline{Z}'\,\underline{\zeta}\right|^2}{\left|\theta(r^{-1}e^{i\omega})\right|^2},$$

where $\underline{\zeta}_j = \left(re^{i\omega}\right)^{-j}$ *for* $1 \leq j \leq p$ *and* $\underline{Z} \sim \mathcal{N}(0, \Gamma_p^{-1})$ *such that* Γ_p *is the* $p \times p$ *Toeplitz covariance matrix corresponding to spectral density* $\left|\phi(e^{-i\lambda})\right|^{-2}$*. When* $g(r) > 0$*, instead*

$$\sqrt{T}\left(\widehat{g}(r) - g(r)\right) \stackrel{\mathcal{L}}{\Longrightarrow} \mathcal{N}(0, V),$$

where $V = \underline{\eta}' F^{-1} \underline{\eta}$, F *is the Fisher information matrix for the ARMA process, and*

$$\underline{\eta} = \left[\begin{array}{l} -(\phi(r^{-1}e^{i\omega})\,\underline{\zeta} + \phi(r^{-1}e^{-i\omega})\,\overline{\underline{\zeta}})\,|\theta(r^{-1}e^{i\omega})|^{-2} \\ (\theta(r^{-1}e^{i\omega})\,\underline{\xi} + \theta(r^{-1}e^{-i\omega})\,\overline{\underline{\xi}})\,|\theta(r^{-1}e^{i\omega})|^{-4}\,|\phi(r^{-1}e^{i\omega})|^{2} \end{array} \right],$$

where $\underline{\xi}_j = (re^{i\omega})^{-j}$ *for* $1 \leq j \leq q$.

Remark. The alternative hypothesis indicates that $g(\rho_0) > 0$, and Theorem 1 indicates that the test statistic is $O_P(T^{1/2})$ plus $T\,g(\rho_0)$ in that case, yielding a consistent test.

Seasonality Hypothesis Testing

Implementation Notes.

- May be easier to fit a high order AR model (sieve approach) using OLS
- Identify AR order using BIC (we found that AIC leads to mis-sized results)
- Use estimated parameters to obtain null limit distribution, simulating $\underline{Z}$

Over-adjustment Case. Instead define the functional $h(r) = |\psi(r^{-1}e^{i\omega})|^2$ and swap the roles of θ and ϕ. Theorem 1 can be adapted by swapping the polynomials (and $\underline{\eta}$ gets multiplied by -1).

Seasonality Hypothesis Testing

Testing Procedure. Focus on post-test (for seasonally adjusted data):

1. Remove the first and last few years of data, so as to remove local non-stationarity

2. Fit an invertible ARIMA model, and obtain the AR(∞) representation of the differenced process as $\pi(z) = \phi(z)/\theta(z)$

3. For any given ω, test $H_0(\rho_0)$ for all $\rho_0 \in (0, 1)$ at level α

4. Obtain interval $C(\alpha)$ consisting of all ρ_0 for which we failed to reject

Seasonality Hypothesis Testing

Is it Seasonal? An interval $C(\alpha)$ is obtained for each ω of interest. Seasonality exists if for at least one ω corresponding to a seasonal frequency, an interval contains $\rho = .97$ (this value is suggested by other studies, but can be modified if desired).

Simulation Evidence

Simulated Processes. We study Gaussian time series generated from

$$(1 - \phi\, B)\,(1 - 2\rho\, \cos(\pi/6)\, B + \rho\, B^2) X_t \sim \mathrm{WN}(0, \sigma^2). \qquad (1)$$

The autocovariance function and spectrum are plotted in Figure 1, where we have set $\phi = .8$ and $\rho = .9$, and $\sigma = 1$. From the plots, it is apparent that the moderate seasonality ($\rho = .9$) is somewhat attenuated by the transient effect ($\phi = .8$), so the impact of the atomic seasonality is weaker than it would be if $\phi = 0$.

Second Example. Lower the seasonal persistency to $\rho = .8$, and dampen the transient component by setting $\phi = .3$, displayed in Figure 2.

Simulation Evidence

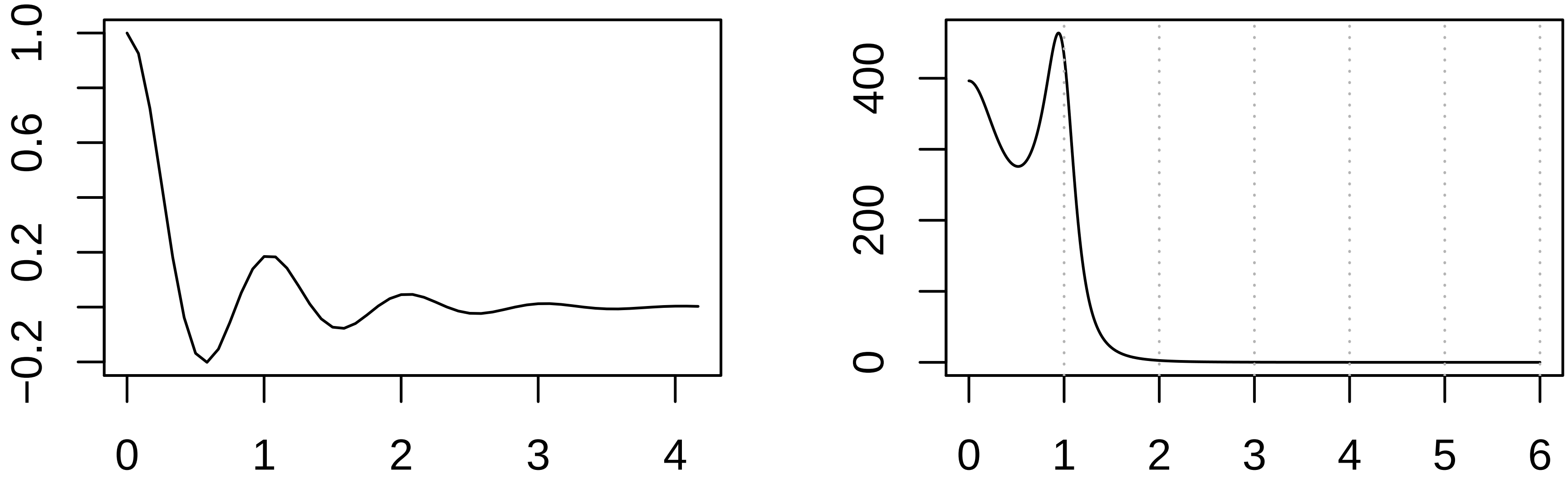

Figure 1: Autocovariance function (left panel) and spectral density (right panel) for AR(3) process ($\phi = .8$, $\rho = .9$).

Simulation Evidence

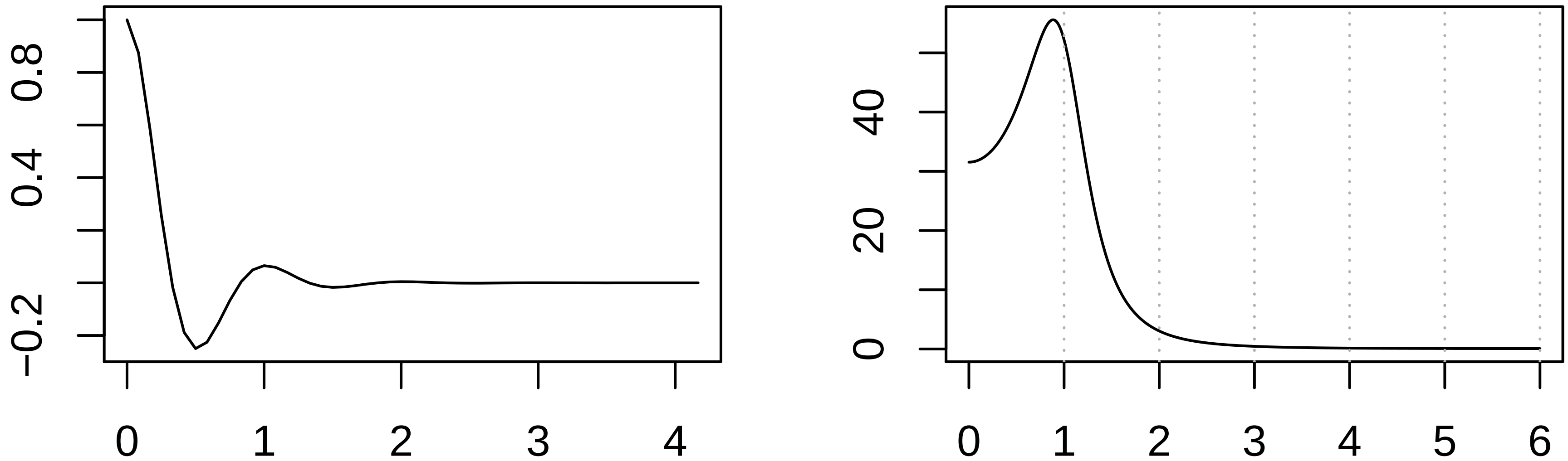

Figure 2: Autocovariance function (left panel) and spectral density (right panel) for AR(3) process ($\phi = .3$, $\rho = .8$).

Simulation Evidence

Simulations. Both processes generated with sample size $T = 12\,n$ and $n = 5, 10, 15, 20$. We take null with $\omega = \pi/6$ and either $\rho = .9, .8$, with either $p = 3$ is known or selected via BIC.

- Size: $\rho = .9$ for first process (Table 1) or $\rho = .8$ for second process (Table 2)

- Power: $\rho = .8$ for first process (Table 3) or $\rho = .9$ for second process (Table 4)

Simulation Evidence

α	5 years	10 years	15 years	20 years
.10	.149	.116	.113	.101
.05	.092	.067	.061	.054
.01	.039	.026	.021	.018
.10	.139	.116	.108	.109
.05	.080	.061	.057	.056
.01	.027	.014	.013	.012

Table 1: Size simulations from an AR(3) DGP (corresponding to Figure 1) based on a null hypothesis of $.9$-persistent seasonality at frequency $\pi/6$. Results are for known AR order (first three rows) and unknown AR order (last three rows).

Simulation Evidence

α	5 years	10 years	15 years	20 years
.10	.135	.115	.109	.104
.05	.077	.058	.057	.053
.01	.021	.012	.013	.011
.10	.132	.113	.108	.107
.05	.073	.059	.053	.054
.01	.020	.014	.009	.012

Table 2: Size simulations from an AR(3) DGP (corresponding to Figure 2) based on a null hypothesis of $.8$-persistent seasonality at frequency $\pi/6$. Results are for known AR order (first three rows) and unknown AR order (last three rows).

Simulation Evidence

α	5 years	10 years	15 years	20 years
.10	.832	.937	.983	.994
.05	.757	.907	.970	.989
.01	.564	.805	.923	.971
.10	.477	.700	.856	.930
.05	.336	.566	.755	.872
.01	.111	.277	.484	.665

Table 3: Power simulations from an AR(3) DGP (corresponding to Figure 1) with null hypothesis of $.8$-persistent seasonality at frequency $\pi/6$. Results are for known AR order (first three rows) and unknown AR order (last three rows).

Simulation Evidence

α	5 years	10 years	15 years	20 years
.10	.116	.205	.341	.473
.05	.060	.110	.203	.313
.01	.015	.028	.057	.096
.10	.281	.498	.682	.812
.05	.175	.362	.550	.702
.01	.051	.148	.294	.448

Table 4: Power simulations from an AR(3) DGP (corresponding to Figure 2) with null hypothesis of $.9$-persistent seasonality at frequency $\pi/6$. Results are for known AR order (first three rows) and unknown AR order (last three rows).

Simulation Evidence

Summary.

- Size is adequate for 10 years of data when p is known or estimated, with coverage being similar for both cases
- Power was good in the case of the first process ($g(\rho_0) = .0048$), though greatly reduced if the model order was unknown
- Power is much lower with second process ($g(\rho_0) = .0059$), although somewhat greater when the model order was unknown

Further Results. Tried to mimic dynamics of real seasonal adjustments; size and power are encouraging (not displayed).

Data Application

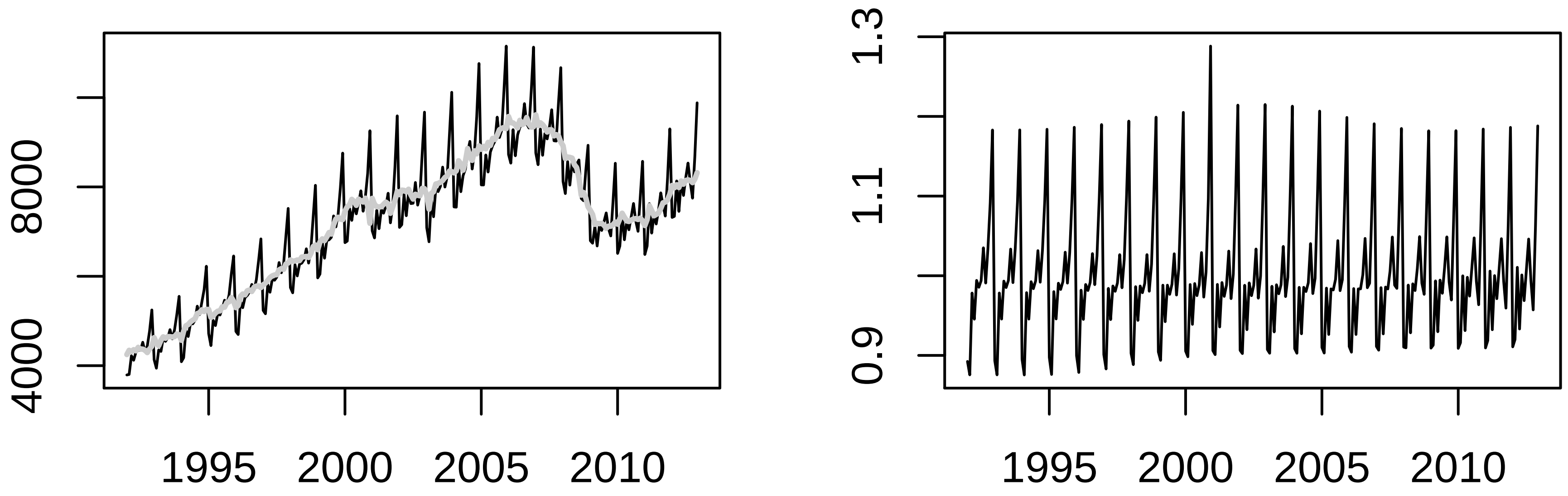

Figure 3: Retail series 442 (Furniture and Home Furnishings Stores), with seasonal adjustment (grey, left panel) and seasonal factors (right panel).

Data Application

Retail Data. We now analyze series 442 (Furniture and Home Furnishings Stores) of the Advance Monthly Retail Trade Report, covering the sample period of January 1992 through December 2012 (see Figure 3). We examine the seasonal adjustment for seasonality (from X-12-ARIMA).

Results. From BIC we obtain an ARIMA(1,1,0), having trimmed first and last three years of data. The confidence intervals for ρ, for each of the five seasonal frequencies $\pi j/6$ $(1 \leq j \leq 5)$ are

$$(0, .062) \qquad (0, .075) \qquad (0, .112) \qquad (0, .167) \qquad (0, .218)$$

(No seasonality!)

Conclusion

- A new paradigm for assessing cyclicality and seasonality is introduced, where oscillations in the Wold coefficients (and autocorrelations) are measured through root magnitudes of the polynomial Ordinary Difference Equation

- This concept addresses several of the criteria set forth, and is demonstrated in simulation to be promising as a test of residual seasonality

- Further work: need theory for case of unit-root AR and MA polynomials, to allow for testing raw data for seasonality

Email: tucker.s.mcelroy@census.gov