The Darkside of the Moon:

Searching For The Other Half of Seasonality

Gary Cornwall Jeff Chen

Bureau of Economic Analysis

November 18, 2019

Research Question

Can we exploit variation in current testing methods to more accurately predict if a series is seasonal?

Motivation

1. Statistical agencies often evaluate tens of thousands of individual series for seasonal patterns, adjust them, and produce seasonally adjusted series (*e.g.* Gross Domestic Product).
2. Definitional malleability - every test is different because no one agrees on a strict definition of seasonality.
3. *Ad hoc* - many tests rely on rules of thumb discovered through trial and error rather than statistical rigor.
4. Residual seasonality - every wrong answer comes with a potential cost.

Motivation

"Since **several of the basic assumptions in the F test are probably violated**, the value of the F ratio to be used for rejecting the null hypothesis, i.e., no significant seasonality present, is tested at the one per thousand probability level." - (Dagum, 1980 [p. 16])

"... the statistical properties of these are **not well understood**" (Lytas, et. al., 2007) - Referring to the M7 and D8F statistic used in X-13.

"The exact null distribution of the QS-statistic is **unknown but can be approximated reasonably well** by a χ^2-distribution with two degrees of freedom, (Maravall, 2011)." (Ollech & Webel, 2017, 2018)

What are we getting at?

- We know seasonality when we see it (generally), but have a really hard time articulating it mathematically.
- This breaks down to a prediction problem.
- We propose using Random Forests to form a *pseudo*-composite test for seasonality.
- Robust simulation environment provides the basis for a horse race between our method and established statistical tests.

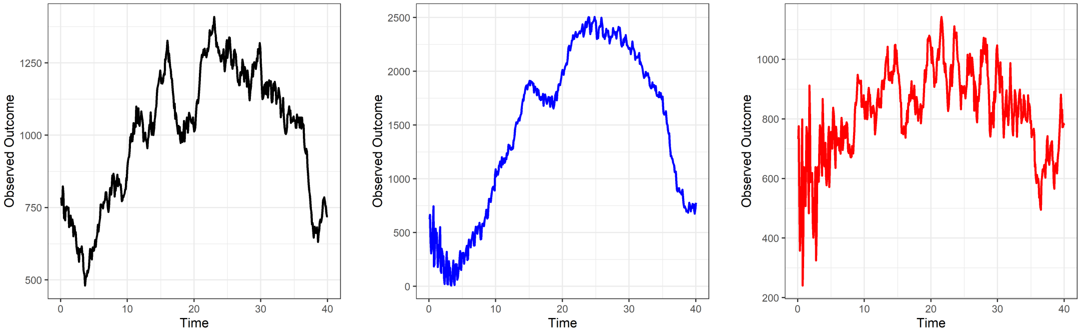
To be, or not to be?

$$(1-B)^d(1-B^s)^D\phi(B)\Phi(B^s)Y_t = \theta(B)\Theta(B^s)\epsilon_t$$

- Y_t is observed series.
- $\epsilon_t \sim \mathrm{N}(0, \sigma^2)$
- $BY_t = Y_{t-1}$
- $\phi(B)$ and $\theta(B)$ are nonseasonal polynomials.
- $\Phi(B^s)$ and $\Theta(B^s)$ are seasonal polynomials.

- $s \in \{4, 12\}$
- d, D refer to order of integration
- $(p\ d\ q)\ (P\ D\ Q)^s$ notation
- $(0\ 1\ 0)\ (P\ 0\ Q)^s$ going forward

Simulated Series: An Example



What tests are we using?

Eight main tests we will look at:

1. QS Test
2. F-stable (D8F) Test
3. F-moving (FM) Test
4. M7 Test
5. F-model (FMB) Test
6. Welch (WE) Test
7. Kruskal-Wallis (KW) Test
8. Friedman (FR) Test

For example:

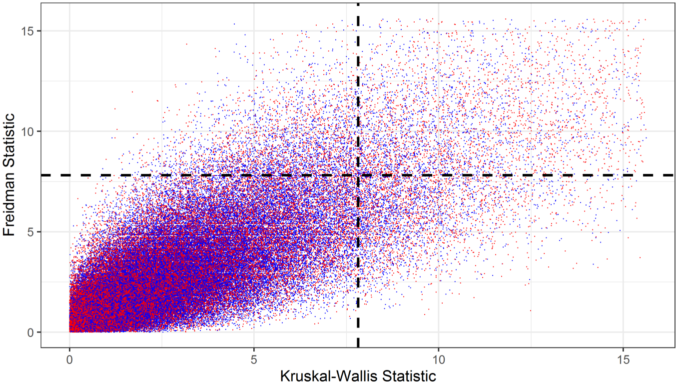
$$\mathrm{QS} = T(T+2)\left(\frac{\hat{\rho}^2(s)}{T-s} + \frac{[\max\{0, \hat{\rho}^2(2s)\}]^2}{T-2s}\right)$$

$$\gamma(g) = \mathbb{E}[y_{t+g}y_t] - \mathbb{E}^2[y_t]$$

$$\rho(g) = \gamma(g)/\gamma(0)$$

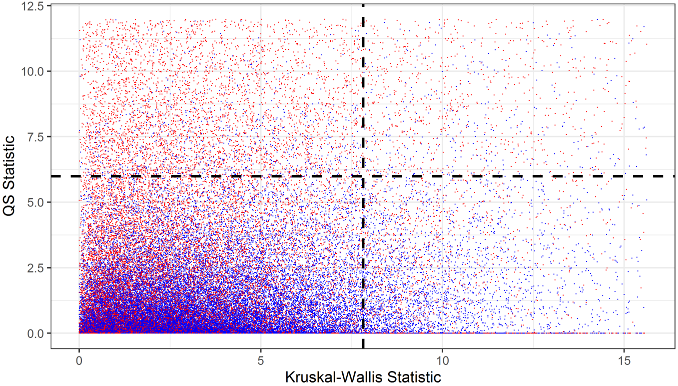
$$H_0 : \gamma(g) \leq 0 \text{ for } g \in \{s, 2s\}$$

What variation are we exploiting?



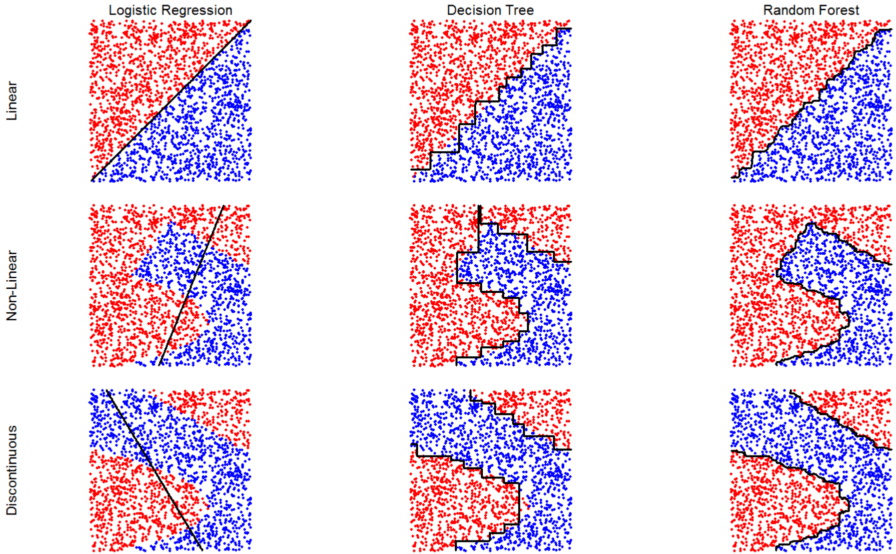
- $\approx 120,000$ simulated series.
- Each point is a single series with coordinate pair $(\mathrm{KW}, \mathrm{FR})$.
- Seasonal series are red, non-seasonal blue.

What variation are we exploiting?



- $\approx 120,000$ simulated series.
- Each point is a single series with coordinate pair $(\mathrm{KW}, \mathrm{QS})$.
- Seasonal series are red, non-seasonal blue.

Why use a Random Forest?

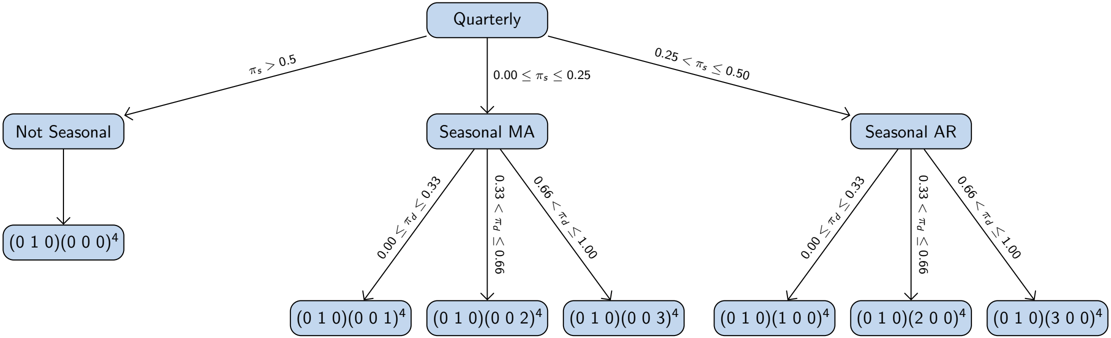


Setting up the RF

$$\underbrace{\mathcal{D}_{0,1}}_{\text{Target Variable}} = \mathscr{F}\big(\overbrace{\lambda,\ \zeta,\ \omega_m,\ \omega_q,\ \rho_m,\ \rho_q}^{\text{Feature Set}}\big)$$

- $\mathcal{D}$: an indicator variable.
- λ: The eight tests for seasonality.
- ζ: TS Characteristics (*e.g.* T, frequency, skewness, kurtosis, etc.)
- ω_m: Monthly Fourier regression coefficients.
- ω_q: Quarterly Fourier regression coefficients.
- ρ_m: Monthly autocorrelation values for two years.
- ρ_q: Quarterly autocorrelation values for two years.

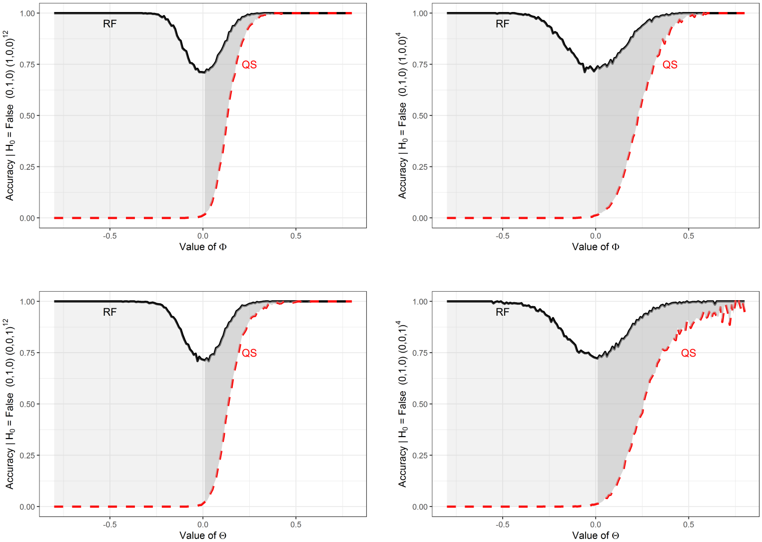
Simulation Structure



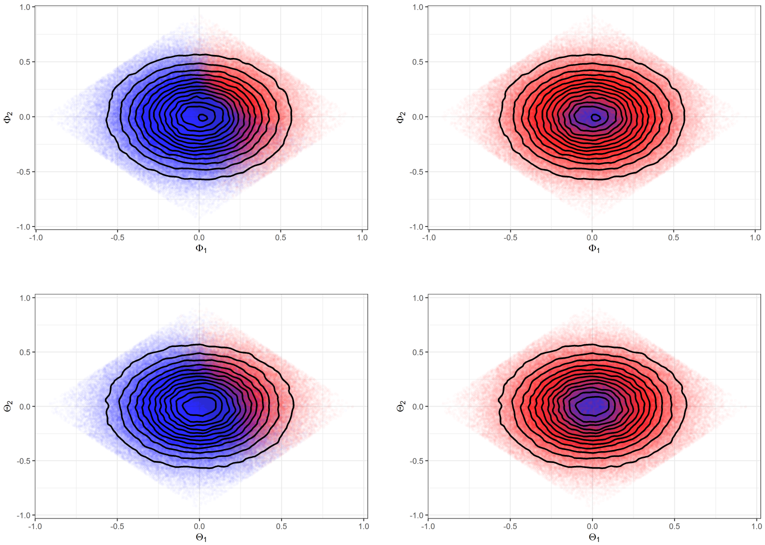
Simulation Structure

- Series length drawn with equal probability (in years), $L \in \{10, \ldots, 50\}$.
- Minimum series length is 40 (120) for quarterly (monthly) frequency.
- Maximum series length is 200 (600) for quarterly (monthly) frequency.
- $\Phi \sim \mathrm{MVN}(0, I_D\sigma^2)$, with $\sigma^2 = 0.25$, s.t. $\sum_{d=1}^{D} |\Phi_d| < 0.95$.
- $\Theta \sim \mathrm{MVN}(0, I_D\sigma^2)$, with $\sigma^2 = 0.25$, s.t. $\sum_{d=1}^{D} |\Theta_d| < 0.95$.
- Noise for all series drawn from $\epsilon_t \sim N(0, 1)$
- Cell sizes need to be big enough for RF to map hypothesis space so we draw $5,000,000$ series for train and test.

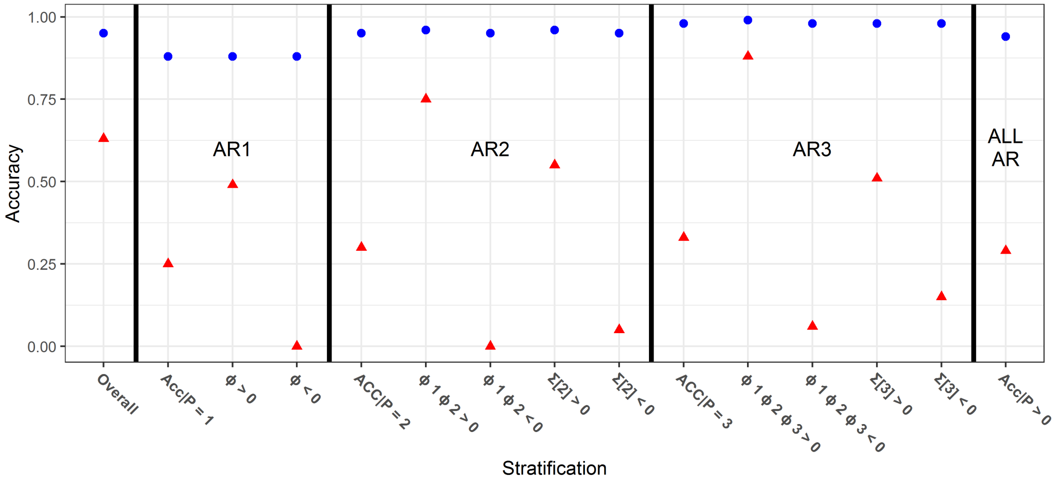
Accuracy over a single dimension.



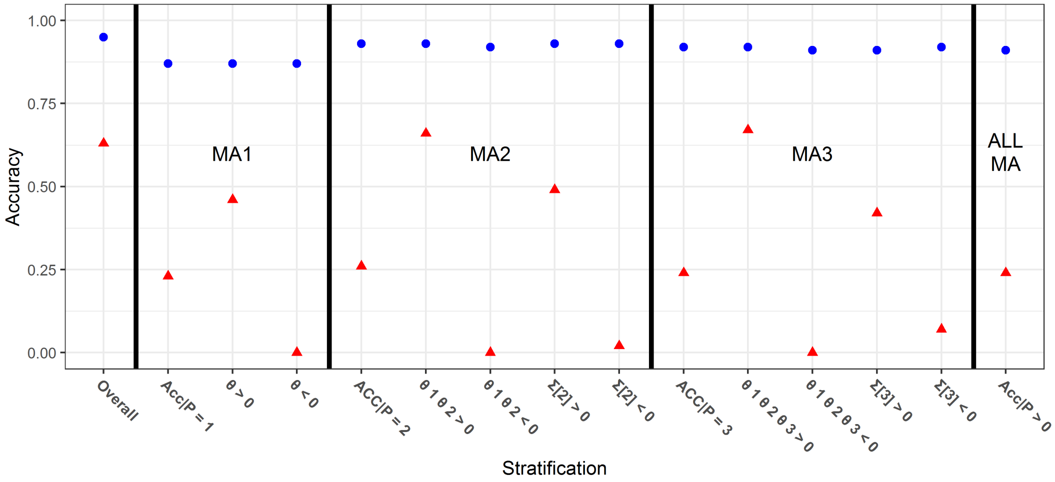
Accuracy over two dimensions.



Overall Accuracy AR Models: QS vs. RF



Overall Accuracy MA Models: QS vs. RF



Thank you!

Appendix

Table: Critical Values and Size: Frequency = 4 Years = 300

	Current CV	Current Size	Suggested CV	New Size
QS	5.991	0.023	3.602	0.059
M7	1.000	0.008	1.259	0.051
D8F	7.000	0.002	3.638	0.049
FM	2.615	0.124	3.578	0.050
FMB	2.612	0.053	2.660	0.050
WE	2.618	0.052	2.642	0.051
KW	7.815	0.052	7.877	0.047
FR	7.815	0.051	7.844	0.048

Table: Critical Values and Size: Frequency = 12 Years = 100

	Current CV	Current Size	Suggested CV	New Size
QS	5.991	0.016	3.668	0.050
M7	1.000	0.000	1.712	0.060
D8F	7.000	0.000	2.471	0.055
FM	1.792	0.824	8.956	0.049
FMB	1.797	0.059	1.791	0.060
WE	1.809	0.066	1.802	0.067
KW	19.675	0.048	19.571	0.050
FR	19.675	0.048	19.617	0.049

Table: Nominal Accuracy Table: Test Data

	RF	QS	M7	D8F	FM	FMB	WE	KW	FR	N. Series
Accuracy	0.95	0.63	0.51	0.51	0.60	0.54	0.54	0.54	0.53	5,000,000
$\mathrm{ACC} \vert P = 1$	0.88	0.25	0.02	0.01	0.31	0.18	0.18	0.13	0.12	416,529
$\mathrm{ACC} \vert P = 2$	0.95	0.30	0.04	0.05	0.62	0.22	0.21	0.18	0.16	416,795
$\mathrm{ACC} \vert P = 3$	0.98	0.33	0.06	0.09	0.85	0.23	0.23	0.21	0.20	417,260
$\mathrm{ACC} \vert P > 0$	0.94	0.29	0.04	0.05	0.59	0.21	0.21	0.17	0.16	1,250,584
$\mathrm{ACC} \vert Q = 1$	0.87	0.23	0.01	0.00	0.19	0.14	0.14	0.08	0.09	416,757
$\mathrm{ACC} \vert Q = 2$	0.93	0.26	0.02	0.01	0.19	0.16	0.16	0.10	0.10	416,714
$\mathrm{ACC} \vert Q = 3$	0.92	0.24	0.02	0.01	0.19	0.16	0.16	0.11	0.11	416,462
$\mathrm{ACC} \vert Q > 0$	0.90	0.24	0.02	0.01	0.19	0.16	0.15	0.10	0.10	1,249,933
$\mathrm{ACC} \vert P, Q = 0$	0.98	0.99	0.99	1.00	0.81	0.89	0.89	0.95	0.94	2,499,483

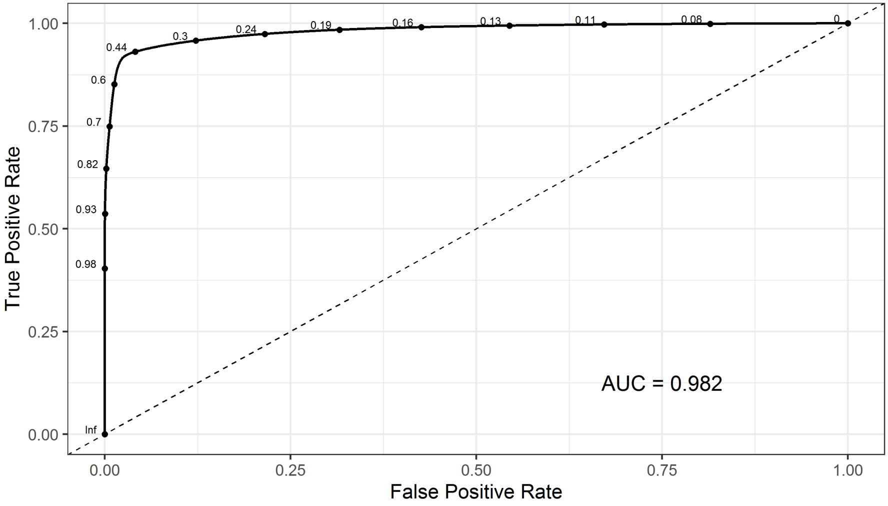
Here we calculate the accuracy based on the nominal critical values for each test and compare them to the out-of-sample prediction accuracy of the Random Forest. The breakdowns are to show accuracy when the null is false (equivalent to power) under specific seasonal dimensionality conditions. Note that there is no constraints on the parameter space of Φ or Θ beyond those which are standard.

Table: Empirical Accuracy Table: Test Data

	RF	QS	M7	D8F	FM	FMB	WE	KW	FR	N. Series
Accuracy	0.95	0.64	0.52	0.54	0.55	0.53	0.53	0.54	0.53	5,000,000
$\mathrm{ACC}\|P=1$	0.88	0.30	0.14	0.10	0.03	0.19	0.19	0.13	0.13	416,529
$\mathrm{ACC}\|P=2$	0.95	0.34	0.14	0.15	0.19	0.22	0.22	0.18	0.17	416,795
$\mathrm{ACC}\|P=3$	0.98	0.37	0.13	0.19	0.40	0.24	0.24	0.21	0.20	417,260
$\mathrm{ACC}\|P>0$	0.94	0.34	0.14	0.15	0.21	0.22	0.22	0.17	0.16	1,250,584
$\mathrm{ACC}\|Q=1$	0.87	0.28	0.13	0.06	0.00	0.15	0.15	0.08	0.09	416,757
$\mathrm{ACC}\|Q=2$	0.93	0.30	0.15	0.08	0.00	0.16	0.17	0.10	0.10	416,714
$\mathrm{ACC}\|Q=3$	0.92	0.30	0.15	0.09	0.00	0.17	0.17	0.11	0.11	416,462
$\mathrm{ACC}\|Q>0$	0.90	0.29	0.14	0.08	0.00	0.16	0.17	0.10	0.10	1,249,933
$\mathrm{ACC}\|P,Q=0$	0.98	0.95	0.91	0.96	1.00	0.88	0.87	0.95	0.93	2,499,483

Here we calculate the accuracy based on the empirical critical values for each test and compare them to the out-of-sample prediction accuracy of the Random Forest. The breakdowns are to show accuracy when the null is false (equivalent to power) under specific seasonal dimensionality conditions. Note that there is no constraints on the parameter space of Φ or Θ beyond those which are standard.

ROC Curve: Test Data



Empirical Critical Values: Quarterly Data

Time (Years)	QS	D8F	FM	M7	FMB	WE	KW	FR
10	11.205	3.893	3.083	1.158	3.219	3.156	6.420	7.533
20	18.449	3.712	2.850	1.205	2.801	2.831	7.453	7.737
30	5.472	3.551	2.880	1.239	2.681	2.704	7.585	7.634
40	3.707	3.602	2.947	1.235	2.700	2.707	7.728	7.708
50	3.458	3.603	3.045	1.246	2.667	2.706	7.722	7.751
60	3.397	3.583	3.109	1.254	2.646	2.653	7.759	7.759
70	3.428	3.598	3.132	1.258	2.665	2.670	7.727	7.661
80	3.485	3.544	3.199	1.264	2.608	2.625	7.714	7.724
90	3.496	3.617	3.215	1.258	2.634	2.637	7.755	7.800
100	3.537	3.579	3.272	1.262	2.630	2.634	7.777	7.788
200	3.569	3.598	3.513	1.271	2.608	2.614	7.854	7.830
250	3.561	3.553	3.603	1.269	2.638	2.633	7.797	7.733
260	3.616	3.569	3.614	1.271	2.626	2.628	7.829	7.740
270	3.604	3.611	3.604	1.263	2.639	2.643	7.827	7.764
280	3.634	3.618	3.599	1.255	2.643	2.642	7.896	7.804
290	3.597	3.651	3.570	1.253	2.677	2.670	7.870	7.767
300	3.602	3.638	3.578	1.259	2.660	2.642	7.877	7.844

Empirical Critical Values: Monthly Data

Time (Years)	QS	D8F	FM	M7	FMB	WE	KW	FR
5	3.356	2.804	5.697	1.422	2.126	2.895	17.801	18.269
10	3.664	2.599	5.877	1.583	1.921	2.094	19.011	19.000
15	3.563	2.512	6.405	1.641	1.865	1.976	19.288	19.330
20	3.566	2.505	6.861	1.659	1.844	1.916	19.382	19.324
25	3.543	2.486	7.222	1.683	1.824	1.878	19.425	19.379
30	3.565	2.507	7.553	1.684	1.825	1.865	19.484	19.547
35	3.608	2.494	7.757	1.701	1.826	1.855	19.500	19.516
40	3.623	2.491	7.944	1.708	1.806	1.843	19.471	19.560
45	3.571	2.478	8.083	1.704	1.812	1.837	19.521	19.626
50	3.617	2.495	8.190	1.696	1.811	1.831	19.555	19.581
55	3.640	2.488	8.305	1.705	1.807	1.827	19.576	19.601
60	3.604	2.464	8.440	1.700	1.803	1.823	19.492	19.616
65	3.630	2.467	8.454	1.714	1.803	1.820	19.575	19.616
70	3.669	2.463	8.609	1.714	1.807	1.828	19.611	19.670
75	3.640	2.469	8.673	1.718	1.802	1.815	19.555	19.561
80	3.685	2.472	8.750	1.713	1.793	1.801	19.542	19.501
85	3.635	2.454	8.835	1.723	1.796	1.813	19.549	19.500
90	3.731	2.451	8.875	1.719	1.788	1.804	19.522	19.382
95	3.694	2.466	8.955	1.718	1.792	1.812	19.564	19.442
100	3.668	2.471	8.956	1.712	1.791	1.802	19.571	19.617