The Impact of Data Frequency On Stationarity Tests
Of Commodity Futures Prices

Xuedong Wu
Graduate Research Assistant
Department of Agricultural and Applied Economics
University of Georgia
bestwu@uga.edu

Jeffrey H. Dorfman
Professor
Department of Agricultural & Applied Economics
University of Georgia
jdorfman@uga.edu

Berna Karali
Associate Professor
Department of Agricultural & Applied Economics
University of Georgia
bkarali@uga.edu

Selected Poster prepared for presentation at the

Copyright 2015 by Xuedong Wu, Jeffrey H. Dorfman and Berna Karali. All rights reserved. Readers may make verbatim copies of this document for non-commercial purposes by any means, provided that this copyright notice appears on all such copies.
In this paper, we will devote efforts in the two aforementioned directions in hopes of improving unit root test results. Using data on 5 commodity futures prices (corn, soybean, cotton, live cattle and lean hog), which all display typical financial series characteristics, we first show that systematic sampling does have effects on the results of unit root testing by testing three different frequency samples: daily, weekly, and monthly.

Then, more importantly, we will test the stationarity of these series by averaging 24 models using a Bayesian Model Averaging unit root test method derived in the previous chapter to confront the model specification uncertainty issue, and compare results with traditional unit root tests to show the performance of the BMA methods, as well as its ability to handle the model specification issue.

To incorporate model uncertainty in the mean function as well as the variance structure, 24 models are averaged to come to a final comprehensive conclusion, which can be categorized into 4 groups by variance structure:
1. GARCH (1, 1) with Student’s t distribution
2. GARCH (1, 1) with Normal distribution
3. ARCH (1, 1) with Student’s t distribution
4. AR model with Student’s t distribution

And for each of the error specifications, the mean function is specified as an autoregressive model with maximum lag varying from 1 to 6.

The priors are specified as follows:
1. Dominant root: Beta (30,2), all other root: Beta (1.1,1.1),
2. GARCH/ARCH Coefficients: N(0, 3) with indicator function to make sure positive variance structure;
3. Variance of normal likelihood: Inverse gamma
4. Degrees of freedom of Student’s t distribution: truncated exponential distribution with following form:

$$p(\nu | \lambda, \delta) = \frac{\Gamma(\frac{\nu+\delta}{2})}{\Gamma(\frac{\nu}{2}) \sqrt{\nu \pi}} \left(1 + \frac{\nu \delta}{\nu + \delta} \frac{1}{\sqrt{\nu}} \right)^{\frac{\nu+\delta}{2}}$$

Table 1. Test Results of Five Commodity Futures Prices Data

<table>
<thead>
<tr>
<th></th>
<th>BMA</th>
<th>DF1</th>
<th>DF2</th>
<th>DF3</th>
<th>DF4</th>
<th>DF5</th>
<th>DF6</th>
<th>PP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Day</td>
<td>0.503</td>
<td>0.576</td>
<td>0.538</td>
<td>0.475</td>
<td>0.339</td>
<td>0.390</td>
<td>0.315</td>
<td>0.54</td>
</tr>
<tr>
<td>Week</td>
<td>0.717</td>
<td>0.530</td>
<td>0.535</td>
<td>0.497</td>
<td>0.577</td>
<td>0.648</td>
<td>0.622</td>
<td>0.55</td>
</tr>
<tr>
<td>Mon</td>
<td>0.784</td>
<td>0.576</td>
<td>0.538</td>
<td>0.475</td>
<td>0.338</td>
<td>0.390</td>
<td>0.315</td>
<td>0.52</td>
</tr>
<tr>
<td>Day</td>
<td>0.372</td>
<td>0.271</td>
<td>0.312</td>
<td>0.282</td>
<td>0.332</td>
<td>0.359</td>
<td>0.357</td>
<td>0.30</td>
</tr>
<tr>
<td>Week</td>
<td>0.406</td>
<td>0.364</td>
<td>0.381</td>
<td>0.319</td>
<td>0.296</td>
<td>0.277</td>
<td>0.256</td>
<td>0.32</td>
</tr>
<tr>
<td>Mon</td>
<td>0.481</td>
<td>0.227</td>
<td>0.098</td>
<td>0.163</td>
<td>0.113</td>
<td>0.043</td>
<td>0.026</td>
<td>0.26</td>
</tr>
<tr>
<td>Day</td>
<td>0.498</td>
<td>0.622</td>
<td>0.627</td>
<td>0.626</td>
<td>0.624</td>
<td>0.579</td>
<td>0.647</td>
<td>0.62</td>
</tr>
<tr>
<td>Week</td>
<td>0.535</td>
<td>0.593</td>
<td>0.651</td>
<td>0.662</td>
<td>0.641</td>
<td>0.550</td>
<td>0.499</td>
<td>0.62</td>
</tr>
<tr>
<td>Mon</td>
<td>0.685</td>
<td>0.489</td>
<td>0.528</td>
<td>0.216</td>
<td>0.468</td>
<td>0.463</td>
<td>0.402</td>
<td>0.52</td>
</tr>
<tr>
<td>Day</td>
<td>0.222</td>
<td>0.537</td>
<td>0.570</td>
<td>0.545</td>
<td>0.559</td>
<td>0.546</td>
<td>0.553</td>
<td>0.54</td>
</tr>
<tr>
<td>Week</td>
<td>0.254</td>
<td>0.550</td>
<td>0.353</td>
<td>0.497</td>
<td>0.577</td>
<td>0.648</td>
<td>0.622</td>
<td>0.55</td>
</tr>
<tr>
<td>Mon</td>
<td>0.408</td>
<td>0.576</td>
<td>0.538</td>
<td>0.475</td>
<td>0.338</td>
<td>0.390</td>
<td>0.315</td>
<td>0.52</td>
</tr>
<tr>
<td>Day</td>
<td>0.141</td>
<td>0.084</td>
<td>0.060</td>
<td>0.010</td>
<td>0.010</td>
<td>0.010</td>
<td>0.018</td>
<td>0.01</td>
</tr>
<tr>
<td>Week</td>
<td>0.188</td>
<td>0.010</td>
<td>0.010</td>
<td>0.010</td>
<td>0.010</td>
<td>0.010</td>
<td>0.018</td>
<td>0.01</td>
</tr>
<tr>
<td>Mon</td>
<td>0.396</td>
<td>0.010</td>
<td>0.010</td>
<td>0.010</td>
<td>0.010</td>
<td>0.010</td>
<td>0.026</td>
<td>0.01</td>
</tr>
</tbody>
</table>

The posterior density cannot be integrated directly so a numerical technique, so Gibbs sampling with a Metropolis-Hastings (MH) step is adopted to generate a sample from the posterior probability distribution of the dominant root.

For each data series the robust test is based on averaging 24 models. For each model, we used 51,000 Monte Carlo iterations of the Gibbs with MH algorithm and discarded the first 21,000 draws to achieve better convergence and better posterior sample mixture. The Geweke test (Geweke 1992) is adopted to examine the convergence of each posterior sample.

Another result to notice is that for each commodity, the probability of a dominant root greater than 1 computed by averaging 24 models using BMA method increases as the frequency of tested data decreases. This indicates that more mean-reversion information is provided by using the high frequency data. To be more specific, high frequency samples carry more information through high volatility and fat-tail behavior which can be captured by GARCH and ARCH models with Student’s t distributions in the BMA method. This information will be lost when constructing low frequency sample through systematic sampling and will be ignored by traditional methods like ADF or PP test.

The most desirable property of our BMA method is it can handle model specification uncertainty in a unit root test. Sometimes model uncertainty causes contradictory results which could lead to a misspecified model. Take soybean monthly data as an example. Using the ADF test on the monthly data and under a commonly used 10% significance level, it is confirmed nonstationary if the model specification is AR (1), AR (3) or AR (4). In contrast, for AR (2), AR (5) and AR (6) the test indicates stationarity of the data, opposite to the result using other lags as well as daily and weekly data. So the model specification uncertainty problem is important here since improper specification of the lag will lead to completely different results which will affect the following analysis. The BMA method confronts this problem by averaging all 6 possible lag specification (or more if the researcher needed) and reaching a final, more robust conclusion.

Xuedong Wu, PhD student, Department of Agricultural and Applied Economics, University of Georgia; bwu@uga.edu
Jeffrey H. Dorfman, Professor, Department of Agricultural and Applied Economics, University of Georgia; jdorfman@uga.edu
Berna Karali, Associate Professor, Department of Agricultural and Applied Economics, University of Georgia; bkarali@uga.edu

Acknowledgement to http://www.mathcaptain.com/statistics/systematic-sampling.html