Testing the Rational Expectations Competitive Storage Hypothesis

Valtteri Ahti
University of Helsinki

Discussion Paper No. 267
July 2009
ISSN 1795-0562
Testing the Rational Expectations Competitive Storage Hypothesis

Abstract

A large number of low-income less-developed countries are dependent on just a few primary commodities for their export earnings. Reliance on primary commodities is particularly hazardous given the extreme behaviour and poor visibility of commodity prices. Rational expectations competitive storage theory postulates that commodity prices are determined by expected futures prices, commodity stocks and the cost of carry. This paper uses a rich data set gathered primarily from the London Metal Exchange to test for the empirical application of the theory. In-sample analysis provides support for the theory whilst the out-of-sample-results cast doubt concerning practical efficacy.

JEL Classification: O1, O13, Q3, G1

Keywords: Developing Countries, Natural Resource, Minerals, Commodity Markets

Valtteri Ahti
Department of Economics
P.O. Box 17
FI-00014 University of Helsinki
FINLAND

Email: valtteri.ahti@helsinki.fi
1 Introduction

A large number of low-income less-developed countries (LDCs) derive effectively all of their export revenue from the sale of just a few primary commodities. Table 1 presents some rather alarming statistics on export reliance on primary commodities in Sub-Saharan Africa. The statistics are particularly grim for Zambia with respect to copper and Rwanda, Uganda and Ethiopia in the case of coffee. Presumably macroeconomic policy is tantamount to understanding the dynamics of commodity prices when a single commodity generates between 99.7 to 56.9 percent of a country’s total export revenue. Decision making is hampered by poor visibility, which may produce long standing detrimental effects. For instance the current debt crisis plaguing LDCs largely traces its roots to unwarranted optimism in the late 1970’s concerning the future evolution of commodity prices.

<table>
<thead>
<tr>
<th>Percent of Total Export Earnings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zambia</td>
</tr>
<tr>
<td>copper 98</td>
</tr>
<tr>
<td>Rwanda</td>
</tr>
<tr>
<td>coffee 73</td>
</tr>
<tr>
<td>Uganda</td>
</tr>
<tr>
<td>coffee 95</td>
</tr>
<tr>
<td>Ethiopia</td>
</tr>
<tr>
<td>coffee 66</td>
</tr>
<tr>
<td>Sudan</td>
</tr>
<tr>
<td>cotton 42</td>
</tr>
<tr>
<td>Tanzania</td>
</tr>
<tr>
<td>coffee 40</td>
</tr>
<tr>
<td>Ghana</td>
</tr>
<tr>
<td>cocoa 59</td>
</tr>
<tr>
<td>Kenya</td>
</tr>
<tr>
<td>coffee 30</td>
</tr>
<tr>
<td>Zimbabwe</td>
</tr>
<tr>
<td>tobacco 20</td>
</tr>
</tbody>
</table>

Source: Oxfam (1993)

Commodity price processes are highly unstable, which augments problems stemming from export reliance. This paper concentrates on base metals, which are an important sub-group of primary commodities. Metal commodity prices exhibit many salient features of primary commodities such as persistence and severe volatility. The price processes between 1976-2009 of aluminium, copper, nickel, zinc, lead and tin are shown in Figure 1. The processes follow a sequence of sharp peaks and subsequent shallow troughs culminating in the financial crisis. The figure shows the flight to commodities from other asset classes in the initial stages of the financial crisis and the subsequent vehement bust. For instance the price of nickel rose from a historical mean of 9559 to an unprecedented high of 52179 in July 2008 only to collapse back to the historical mean. The other base metals sustained analogous extremes.

Theoretical work on commodities has a long standing history based on the theory of storage and its implications for equilibrium price evolution. The synthesis of Gustafson’s (1958) work on the optimal demand for commodity stocks and Muth’s (1961) rational expectations hypothesis may be seen as constituting

Essentially competitive storage theory postulates commodity prices as processes driven by arbitrage behaviour determined by commodity stocks, expected future prices and the cost of carry. The equilibrium pricing process may be modelled as a functional equation where competitive speculators increase or decrease stocks in anticipation of arbitrage subject to costs of carry.

A rich data set composed of information on metal commodities listed in the London Metal Exchange (LME) enables empirical examination of the validity of competitive storage theory. The data set contains information on all variables deemed pertinent by competitive storage theory for modelling commodity price processes. The data contains spot prices, futures prices and inventories. Given rational expectations, futures prices should make an excellent proxy for expected future spot prices. LME spot prices and inventories are in effect global market prices and stocks given the preeminent position of the LME as a forum for metals exchange.

A vector autoregressive model (VAR) loaded with theory prescribed variables is taken to the LME data to examine the empirical applicability of the theory. On account of nonconstant variances and since the series are found to be first order difference stationary the levels data is transformed into returns form. In-sample validity is examined via Granger and Instantaneous causality tests. The causality tests find evidence in support of the theory for all variables except for the cost of carry variable. Cost of carry is proxied by the USD Libor rate.

The storage hypothesis is evaluated using what Stock and Watson (2008) refer to as a pseudo out-of-sample evaluation. A rolling window scheme is applied with respect to each base metal where the window length is kept fixed throughout the recursive procedure. The addition of a new observation to the estimation window results in the elimination of the last observation. Additionally the model skeleton is kept fixed whilst parameters are reestimated at each iteration. The degrees of freedom is therefore held constant so that the forecast errors may be subjected to statistical testing.

One-step-ahead forecasts from the VAR are pitted against forecasts from a “no change” model, a seasonal autoregressive integrated moving average model (SARIMA) and the exponential smoothing model (ETS). Both SARIMA and ETS models have proved to be effective forecasting tools in the macroeconometric and time series literatures. Model accuracy is appraised by computing root mean squared forecast errors (RMSFE) and p values from Diebold-Mariano test of equal predictability (1995). The out-of-sample results don’t support the storage hypothesis as the VAR fails to outperform either the benchmark no change model or the univariate models.

The article is organized as follows. Section 2 presents a standard if very
basic variant of a rational expectations competitive storage model. Section 3 discusses the metal commodities data set. Section 4 develops the VAR indicated by storage theory and discusses the causality results. Section 5 presents the simulated out-of-sample framework and the results of the model horse race. Section 6 concludes.

2 A basic rational expectations competitive storage model

This section contains a simple rational expectations competitive storage model. Although the model is simple it contains all the necessary information from the point of view of empirical validation. In other words theoretical refinements do not generally perturb the set of variables. Assume that time \( t \) is discrete and the price of a commodity \( p \) is normalized with respect to some numeraire. There are two types of agents that have a demand for the commodity. Producers demand the commodity for production purposes whilst speculators carry inventories for arbitrage purposes.

Each period an amount \( z_t \) is supplied to the market. The literature often refers to \( z_t \) as a harvest. The equilibrium price is then determined as the price that clears the market \( D(p_t) = z_t \). Now denote the inverse demand function as \( P(z_t) \) and note that in the absence of inventories \( p_t = P(z_t) \). Next introduce costly inventories \( I_t \) with cost of carry \( r_t \). Then the expected profit on inventories will be \( [(1 - r_t) E_t p_{t+1} - p_t] I_t \). The appropriate control variable for a commodity speculator is then the level of inventories to hold and intertemporal profit maximization yields

\[
\begin{align*}
I_t &= 0 & \text{if } (1 - r_t) E_t p_{t+1} < p_t \\
I_t &\geq 0 & \text{if } (1 - r_t) E_t p_{t+1} = p_t
\end{align*}
\] (1)

Inventories will not be held if the expected profit from holding stocks yields a negative payoff. If speculators perceive an arbitrage opportunity they will demand inventories until parity is established between expected payoffs from holding inventories and not holding inventories. In equilibrium supply at time \( t \) and inventories carried over \( I_t \) must be equal to demand at time \( t \) and inventories to be carried over to time \( t + 1 \).

\[
z_t + I_{t-1} = D(p_t) + I_t.
\] (2)

Therefore combining (1) and (2) the equilibrium price is

\[
p_t = \max \left[ (1 - r_{t-1}) E_t p_{t+1}, P(z_t + I_{t-1}) \right]
\] (3)

and equilibrium inventories can be retrieved by inserting the equilibrium price into the equations in (1). The competitive storage hypothesis hence asserts that a commodity price \( p_t \) is determined by the variables \( E_t p_{t+1}, r_t \) and

\[
3
\]
Furthermore if rational expectations hold then the futures price of the commodity $f_t$ should reflect $E_t p_{t+1}$. Hence one is able to test the theory of storage by jointly modeling the evolution of the set of variables $y_t = \{p_t, f_t, I_t, r_t\}$.

3 Commodities data

The LME traces its lineage as a major international trade forum up to the late Victorian era. Currently it is the largest pure commodity exchange in Europe and the World’s tenth largest futures exchange. Trading features all of the important metal commodities: aluminium, aluminium alloy, copper, nickel, zinc, lead, tin and silver. Watkins and McAleer (2004) provide the following description of the exchange: The LME is used worldwide by producers and consumers as a center for spot, futures and options trading in non-ferrous metals. The LME offers three primary functions. Firstly, market participants can hedge against the risk of price volatility. Secondly, the LME settlement prices are used as reference prices around the world. For instance, approximately 95% of the world trade in copper futures takes place in the LME making it the de facto world market price. Thirdly, the LME offers the services of a global warehouse network for settlements resulting in physical delivery.

The metals data-set therefore contains all the necessary ingredients to test rational expectations competitive storage theory. Spot prices are in effect equilibrium prices, futures prices reflect expectations of future prices and LME inventories match the stocks held forward due to the exchange’s position as the premier forum for metal commodities exchange.

Figure 1 depicts the evolution of monthly metal spot and futures prices for the period 1.1976-2.2009. The series were constructed from daily data by taking the last observation of each month. Copper is the longest series stretching from 1.1976 yielding almost 400 observations whilst lead constitutes the shortest series starting from 1.1993 and therefore contributing nearly 200 observations. Metals prices were listed in sterling prior to 1989. Therefore prices listed prior to 1989 are converted into dollar denominated series using USD-GBP exchange rates. The dashed line indicating three month futures prices is practically indistinguishable from the solid line representing spot prices in Figure 1. Table 1 reports the summary statistics for the spot prices and gives an indication of price volatility in addition to giving insight to the magnitude of the price bubble prior to the financial crisis evident at the end of the sample. For instance nickel was traded at a high of over 50 000 USD per tonne compared to the median price of 7000 USD per tonne. The heights reached prior to the financial meltdown are all the more impressive given the lack of a clear trend in most of the price series.

The variables suggested by storage theory $y_t = \{p_t, f_t, I_t, r_t\}$ were inspected for unit root nonstationarity by implementing the Augmented Dickey-Fuller (ADF), Phillips-Perron (PP) and Elliot-Rothenberg-Stock (ERS) tests. The tests strongly indicate that all variables are best characterized as first order integrated processes. All variables also exhibit variance-nonconstancy. Hence
Figure 1: Metal Spot Prices
all variables were converted to differences of natural logarithms, i.e. from levels to returns form.

<table>
<thead>
<tr>
<th></th>
<th>Al</th>
<th>Cu</th>
<th>Ni</th>
<th>Zn</th>
<th>Pb</th>
<th>ld</th>
</tr>
</thead>
<tbody>
<tr>
<td>Obs</td>
<td>245</td>
<td>398</td>
<td>356</td>
<td>244</td>
<td>194</td>
<td>237</td>
</tr>
<tr>
<td>Mean</td>
<td>1688</td>
<td>2460</td>
<td>9564</td>
<td>1375</td>
<td>869</td>
<td>7144</td>
</tr>
<tr>
<td>Median</td>
<td>1550</td>
<td>1889</td>
<td>7096</td>
<td>1100</td>
<td>612</td>
<td>5828</td>
</tr>
<tr>
<td>Max</td>
<td>3086</td>
<td>8775</td>
<td>50898</td>
<td>4390</td>
<td>3690</td>
<td>23588</td>
</tr>
<tr>
<td>Min</td>
<td>1045</td>
<td>1144</td>
<td>3226</td>
<td>735</td>
<td>363</td>
<td>3705</td>
</tr>
<tr>
<td>Sd</td>
<td>476</td>
<td>1609</td>
<td>7578</td>
<td>730</td>
<td>665</td>
<td>3670</td>
</tr>
</tbody>
</table>

Table 2: Summary Statistics Metal Spot Prices

4 Tests of Causality

Since equilibrium spot prices are characterized as arising from an arbitrage relation determined by futures prices, inventories and interest rates the theory is tested using a multivariate model. One of the most successful and flexible multivariate models in time series and macroeconometrics is the VAR model. VAR models tend to outperform both univariate and more elaborate simultaneous systems of equations in terms of forecasting and description. For a textbook presentation and S-PLUS implementation see Zivot and Wang (2002).

As denoted earlier, $y_t = (p_t, f_t, I_t, r_t)'$. Then the $p$-lag VAR or VAR($p$) can be expressed as

$$y_t = \mu + \Gamma_1 y_{t-1} + \ldots + \Gamma_p y_{t-p} + u_t$$

where $\Gamma_i$ are $(4 \times 4)$ coefficient matrices for $i = 1, \ldots, p$, $\mu$ is a constant term and $u_t$ is a $K$-dimensional error process with $E(u_t) = 0$ and time invariant positive definite covariance matrix $E(u_t u'_t) = \Sigma_u$. A VAR(2) of the differenced variables would then be

$$\begin{pmatrix} \Delta p_t \\ \Delta f_t \\ \Delta I_t \\ \Delta r_t \end{pmatrix} = \begin{pmatrix} \mu_p \\ \mu_f \\ \mu_I \\ \mu_r \end{pmatrix} + \begin{pmatrix} \gamma^1_{11} & \gamma^1_{12} & \gamma^1_{13} & \gamma^1_{14} \\ \gamma^2_{11} & \gamma^2_{12} & \gamma^2_{13} & \gamma^2_{14} \\ \gamma^3_{11} & \gamma^3_{12} & \gamma^3_{13} & \gamma^3_{14} \\ \gamma^4_{11} & \gamma^4_{12} & \gamma^4_{13} & \gamma^4_{14} \end{pmatrix} \begin{pmatrix} \Delta p_{t-2} \\ \Delta f_{t-2} \\ \Delta I_{t-2} \\ \Delta r_{t-2} \end{pmatrix} + \begin{pmatrix} u_p \\ u_f \\ u_I \\ u_r \end{pmatrix}$$

The selection of lag length $p$ is generally determined by minimizing an information criterion such as the Akaike (AIC), Schwartz-Bayesian (BIC) or Hannan-Quinn (HQ) information criterion. An information criterion generally takes the
form

\[ IC(p) = \ln \left| \sum (p) \right| + c_T \cdot \varphi(n, p) \]

where \( \sum (p) = T^{-1} \sum_{t=1}^{T} \hat{u}_t \hat{u}'_t \) is the residual covariance matrix and \( \psi(n, p) \) is a function that penalizes overfitting. The idea is to estimate a large number of VAR models i.e. VAR(p), such that \( p = 1, ..., P \) and choose \( p \) such that it minimizes a prespecified information criterion. The lag length of the VAR(p) to be used in both the in-sample and out-of-sample analysis was chosen by minimizing the BIC criterion. The specification is done such that the last three years are withheld for out-of-sample analysis. All in-sample evaluation and model specification is done using only the in-sample portion of the data.

The competitive storage hypothesis indicates spot prices, futures prices, inventories and interest rates should be able to explain commodity dynamics. If that is the case then each variable in the system should Granger cause the entire system of equations. Table 3 reports the p values from the Granger and the Instantaneous causality tests with respect to each base metal. Granger causality is found with respect to all variables except the interest rate. Therefore the in-sample results for the most part lend support to the competitive storage hypothesis.

<table>
<thead>
<tr>
<th></th>
<th>Al</th>
<th>Cu</th>
<th>Ni</th>
<th>Zn</th>
<th>Pb</th>
<th>Sn</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Gra</td>
<td>Ins</td>
<td>Gra</td>
<td>Ins</td>
<td>Gra</td>
<td>Ins</td>
</tr>
<tr>
<td>( \Delta p_t )</td>
<td>0.00</td>
<td>0.00</td>
<td>0.09</td>
<td>0.00</td>
<td>0.02</td>
<td>0.07</td>
</tr>
<tr>
<td>( \Delta f_t )</td>
<td>0.00</td>
<td>0.01</td>
<td>0.09</td>
<td>0.02</td>
<td>0.01</td>
<td>0.08</td>
</tr>
<tr>
<td>( \Delta I_t )</td>
<td>0.01</td>
<td>0.00</td>
<td>0.87</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>( \Delta r_t )</td>
<td>0.78</td>
<td>0.11</td>
<td>0.30</td>
<td>0.82</td>
<td>0.58</td>
<td>0.77</td>
</tr>
</tbody>
</table>

Table 3: Granger and Instantaneous Causality p values

5 Forecasting Results

Quite possibly a more stringent test of competitive storage theory lies not in in-sample analysis, but rather in out-of-sample performance. Hence a simulated out-of-sample framework is constructed to provide an additional metric to examine the validity of the storage hypothesis.

The simulated out-of-sample forecast exercise takes the form of rolling window forecasting. The idea is to simulate a real world scenario where policymakers produce forecasts in real time. After every new observation policymakers reoptimize model parameters and then generate a new one step ahead forecast. The number of forecasts is 36 since the out-of-sample period amounts to three years. The model skeletons and the degrees of freedom for each model are kept fixed so that forecast errors may be subjected to statistical testing. This is achieved by specifying the number of parameters ex-ante by minimizing the
BIC using only the in-sample period. Additionally the length of the estimation window is held constant by dropping the earliest observation upon the addition of the most recent observation.

The VAR model is pitted against a benchmark model and a number of versatile univariate models that have in general proved successful in a number of out-of-sample forecasting applications such as the M3 competition (see Makridakis and Hibon (2000)). As is standard in the literature the benchmark model is taken to be the random walk (RW) or “no change” model. The competing univariate models are the seasonal autoregressive integrated moving average model (SARIMA) and the exponential smoothing model (ETS). The idea in using univariate models is that they represent the alternative view where the information in storage models is judged to be of no importance and hence is excluded from the model specification.

The ETS models are relatively simple and robust models where past observations are weighed exponentially less. The ETS modelling methodology is taken from Hyndman et al. (2002) and Hyndman et al. (2008). The idea is to model the ETS along three dimensions; the trend, the errors and seasonality. Each component may enter the model either additively or multiplicatively yielding a total of eight ETS combinations.

Forecast performance is evaluated by computing the root mean squared forecast error (RMSFE), the relative RMSFE and the Diebold-Mariano test of equal forecast accuracy for the one-step-ahead forecast of each model. The forecast errors are converted from returns form back into levels form prior to computing measures of forecast accuracy. Measures of accuracy are reported in Table 3. The first column relates the RMSFE of the benchmark model for all six series. Then the RMSFE, relative RMSFE and the p value of the Diebold-Mariano test are reported for each of the three models with respect to the six base metal series. The relative RMSFE is obtained by dividing the RMSFE of the model in question with the RMSFE of the benchmark model. Hence a score less (greater) than one indicates that the model in question outperformed (underperformed) the benchmark model. Similarly the columns labeled DM report the p-value of a two sided pairwise Diebold-Mariano test of equal predictability of the stated model vis-à-vis the benchmark model.

The results of the out-of-sample simulation exercise do not corroborate the in-sample causality findings as the VAR model performs poorly with respect to the competing models and the benchmark model. In fact in only one case (zinc) out of six does a VAR specification outperform the benchmark model. The Diebold-Mariano tests indicate that the forecasting ability of the VAR model is statistically indistinguishable from the benchmark model except in the case of copper. Similarly the VAR is unable to generate better forecasts than either of the univariate models. The lacklustre out-of-sample performance of the VAR counsels against operationalizing an econometric model with variables indicated by competitive storage theory.
Table 4: Out-of-sample Statistics: Spot Prices

<table>
<thead>
<tr>
<th></th>
<th>RW</th>
<th>SARIMA</th>
<th>ETS</th>
<th>VAR</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSFE</td>
<td>RMSFE</td>
<td>Relative</td>
<td>RMSFE DM</td>
<td>RMSFE</td>
</tr>
<tr>
<td>Al</td>
<td>150.23</td>
<td>149.34</td>
<td>0.99</td>
<td>0.41</td>
</tr>
<tr>
<td>Cu</td>
<td>765.64</td>
<td>765.64</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Ni</td>
<td>3496.23</td>
<td>3496.23</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Zn</td>
<td>365.58</td>
<td>347.09</td>
<td>0.95</td>
<td>0.64</td>
</tr>
<tr>
<td>Pb</td>
<td>313.60</td>
<td>313.60</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Sn</td>
<td>1229.01</td>
<td>1155.75</td>
<td>0.94</td>
<td>0.00</td>
</tr>
</tbody>
</table>

6 Conclusion

Econometric modeling of commodities prices is important given the degree of export dependency in a large number of LDCs, particularly in Sub-Saharan Africa. To address this issue theoretical insights from competitive storage models were incorporated by estimating a vector autoregression using metal commodities data on spot prices, futures prices and inventories. Granger and Instantaneous causality tests indicated that the inclusion of theory relevant variables improved the in-sample performance of the VAR. The in-sample results were somewhat negated by the poor out-of-sample performance of the VAR versus a benchmark model of no change.

Given the results of the simulation exercise one has to be cautious concerning the practical efficacy of implementing competitive storage models. However, this paper tested one potential formulation of the competitive storage model. Naturally alternate modelling specifications may prove viable and hence the results of this paper may be taken as indicative of the potential merit of incorporating storage theory prescribed variables in estimating models of primary commodities. For instance variables indicated by competitive storage theory such as inventories might be used as threshold variables in regime switching models of commodity prices. Such possibilities will be the subject of future enquiry.

References


