WHAT EXPLAINS HIGH COMMODITY PRICE VOLATILITY? ESTIMATING A UNIFIED MODEL OF COMMON AND COMMODITY-SPECIFIC, HIGH- AND LOW-FREQUENCY FACTORS

Berna Karali
Assistant Professor
Department of Agricultural and Applied Economics
The University of Georgia
Athens, GA 30602-7509
Telephone: (706) 542-0750
Fax: (706) 542-0739
E-mail: bkarali@uga.edu

Gabriel J. Power
Assistant Professor
Department of Agricultural Economics and Faculty of Agribusiness
Texas A&M University
College Station, TX, 77843-2124
Telephone: (979) 845-5911
Fax: (979) 845-6378
E-mail: gpower@tamu.edu

Selected Paper prepared for presentation at the Agricultural & Applied Economics Association 2009. AAEA & ACCI Joint Annual Meeting, Milwaukee, Wisconsin, July 26-29, 2009.

What Explains High Commodity Price Volatility? Estimating a Unified Model of Common and Commodity-Specific, High- and Low-Frequency Factors

Abstract

We estimate a model of common and commodity-specific, high- and low-frequency factors, built on the spline-GARCH model of Engle and Rangel (2008) to explain the period of exceptionally high price volatility in commodity markets during 2006-2008. We find that decomposing realized volatility into high- and low-frequency components reveals the impact of slowly-evolving macroeconomic variables on the price volatility. Further, we find that while macroeconomic variables have similar effects within the same commodity category (e.g., storable agricultural), they have different effects across commodity groups (e.g., live stock versus energy).

Keywords: volatility, spline-GARCH, futures markets

1. Introduction

The price volatility of agricultural commodities has been exceptionally high during the "commodity boom" of 2006-2008. Consider, for example, that since 1980 the volatility of futures prices for corn has averaged 19.7%, but it reached record high levels of 28.8% in 2006, 31.4% in 2007, and 41% in the first quarter of 2008 (Schnepf 2008). For wheat, volatility increased from a historical average of 22.2% to record values of 30.4% in 2006, 32.7% in 2007, and 73% in early 2008. This exceptionally high level of price volatility has complicated established agribusiness practices, particularly the management of price risk (Mark et al. 2008). For instance, the availability of forward contracting has sharply decreased in some regions, futures margin calls have become large enough to cause bankruptcy, and options (whose premium increases with volatility) have become prohibitively expensive. Although prices and volatility have decreased in late 2008, understanding the underlying factors behind this period of high volatility should be helpful to producers, commodity traders, and policy makers to better prepare for the next period of high price volatility.

Even though there is a vast number of studies on volatility, most use extremely reduced-form time series models that are not based on economic theory and are driven only by the statistical properties of the historical data. Indeed, models that are more structural have been found to perform poorly, i.e. they can explain only a small fraction of the observed volatility.

The main contribution of this paper is to estimate a model of common and commodity-specific, high- and low-frequency factors to explain the period of exceptionally high volatility during 2006-2008. To this end, we build upon the recently proposed volatility model and estimation framework of Engle and Rangel (2008), which may be described as semi-structural. This model combines time series dynamics (high-frequency volatility) with slowly-evolving macroeconomic effects (low-frequency volatility). Thus, it can capture the effect and magnitude of both the very short-term factors, which are typically best modeled using purely statistical models, e.g., GARCH, as well as the longer-run, macroeconomic variables that are found in more structural models. The low-frequency component is specified as a function of log real GDP, risk-free interest rate, and inflation rate (general variables) as well as a function of inventories and seasonal factors (commodity-specific variables). It has great potential to track and forecast commodity price volatility, since commodity prices have been found to fluctuate with demand and supply shocks, seasonality in the production cycle, weather, and inventories (see e.g., Anderson 1985; Pindyck 1994; Black and Tonks 2000).

We estimate, using all available futures price data, a model of volatility determinants in which commodities can be categorized by sub-groups: agricultural, live stock, energy, precious metals, and industrial metals. Using futures instead of cash price data is motivated by several key advantages. First, futures price settlement prices are available for every business day and therefore provide a high frequency of sampling which is ideal to study both high- and low-frequency components of volatility. Also, the contract is standardized, e.g., it is defined for a specific commodity grade. Second, the literature has concluded that price discovery generally occurs in futures markets, assuming futures trading volume is reasonably high. Third, futures markets serve as hedging tools and as instruments for capitalizing commodity price fluctuations.

We study the variation across commodities in the same group, e.g. storable agricultural, as well as across commodity groups (precious metals versus industrial metals). For example, gold is considered by many traders to be hedges against inflation and a weak U.S. dollar, but industrial metals are not. Moreover, inventories are essential in the case of storable commodities but do not play a role for non-storable commodities. Estimating this model will allow us to break down volatility into high-frequency (both "noise" and "news") and low-frequency (macroeconomic) components and further to determine what volatility determinants are common across commodities or commodity sub-groups as opposed to commodity-specific. Further, we ask whether a model that captures both high- and low-frequency components of volatility could better explain past volatility and predict future volatility than traditional volatility models, e.g. standard GARCH(1,1). We produce one-step ahead (one month ahead) forecasts of realized volatility and compare the results with those obtained from a simple GARCH(1,1).

In a future version of this paper, we will estimate a single, unified model of high- and low-frequency volatility using data for all commodities described. The main challenge involves the large number of parameters associated with multivariate GARCH models.

In sum, our paper aims to answer the following questions:

1. How much of the high price volatility can be explained by the common economic factors and how much by the commodity-specific factors?

2. How similar are commodities, in terms of volatility determinants, within sub-groups?

3. Does the addition of low-frequency components into the high-frequency time series model improve forecasting accuracy?

2. Model

Modeling and forecasting volatility has attracted a large amount of attention in the financial time series literature, particularly since the development of the GARCH framework for modeling volatility conditional on the available information at a given time period (see e.g., Bollerslev 1986; Poon and Granger 2003). The GARCH model has largely replaced more traditional volatility estimators such as a moving window of standard deviations of price log-returns (or innovations from an ARMA filter).

Volatility models are high-frequency; they use daily or intra-day price observations to estimate a model and produce forecasts. In this literature it is often assumed that low-frequency variables such as macroeconomic indicators are unlikely to improve the model fit because the information is quickly incorporated by the market.

Engle and Rangel (2008) propose instead a model where high-frequency, e.g. daily, volatility is the product of high-frequency news or noise as well as market reactions to lower-frequency events, such as changes in measured inflation, real GDP, etc. In their model, low-frequency volatility, captured using low-order quadratic splines, plays an important role in determining a smooth, nonlinear trend in the volatility of asset prices over long periods of time. Their results suggest that there are nontrivial gains from the inclusion of such low-frequency components into a volatility model. We present in this section a description of the spline-GARCH model developed by them.

Consider a time series of zero-mean, white noise residuals or innovations r_t from a regression of asset prices. (We describe in the following section how we obtain these residuals from a dataset of overlapping commodity futures prices.) We would like to forecast the variance of these innovations to improve our price forecasts and to narrow our confidence intervals. The innovations are found to display volatility clustering, i.e. autoregressive conditional heteroskedasticity (ARCH).

Assume that innovations to returns depend on the state of the macroeconomy, therefore we can write:

$$r_t = \sqrt{\tau_1(z_t)} u_t, \tag{1}$$

where it is assumed $E_{t-1} r_t = 0$, u_t captures "news," and $\tau_1(z_t)$ is a function of a vector of macroeconomic state variables z_t, affecting the impact of news on the asset price. Furthermore, the news itself depends on a different function τ_2 of macroeconomic variables as follows:

$$u_t = \sqrt{\tau_2(z_t) g_t} \epsilon_t, \tag{2}$$

where g_t is a unit mean GARCH process capturing volatility clustering and ϵ is an i.i.d. (0,1) process. Substituting equation (2) into equation (1), we obtain:

$$r_t = \sqrt{\tau_1(z_t)\tau_2(z_t) g_t} \epsilon_t. \tag{3}$$

Recall the traditional GARCH model specification. Residuals from a regression of asset prices are fitted to a conditional mean equation, resulting in the following:

$$r_t = \sqrt{h_t} \epsilon_t, \tag{4}$$

$$h_t = \omega + \alpha \epsilon_{t-1}^2 + \beta h_{t-1}, \tag{5}$$

where h_t is the conditional variance, ω is the volatility constant, α is the ARCH term, and β is the GARCH term. The unconditional variance is $\sigma^2 = \frac{\omega}{(1-\alpha-\beta)}$.

Engle and Rangel propose a method to estimate the low-frequency variance component, $\tau = \tau_1\tau_2$, using a number of macroeconomic variables as well as a quadratic spline approach to make the component high-frequency. Therefore we rewrite the GARCH model to include the low-frequency component:

$$r_t = \sqrt{\tau_t g_t}\epsilon_t \tag{6}$$

$$g_t = (1-\alpha-\beta) + \alpha\left(\frac{r_{t-1}^2}{\tau_{t-1}}\right) + \beta g_{t-1} \tag{7}$$

$$\tau_t = c\exp\left(w_0 t + \sum_{i=1}^{k} w_i\left((t-t_{i-1})_+\right)^2 + z_t\gamma\right) \tag{8}$$

where ϵ, conditional on the filtration F_{t-1}, is distributed as i.i.d. Normal (0,1), c is a constant, $w_0 t$ is a time trend in the low-frequency volatility, $\sum_{i=1}^{k} w_i\left((t-t_{i-1})_+\right)^2$ is a low-order quadratic spline, $(t-t_{i-1})_+ = \max\{0, t-t_{i-1}\}$, and $z_t\gamma$ represents the impact of macroeconomic variables. The number of knots in the spline, k, is determined by comparing the AIC of each specification. A larger k implies more cycles, while the sharpness of the cycles is determined by the coefficients w_i. As Engle and Rangel note, in this model the low-frequency volatility equals the unconditional volatility:

$$E\left[r_t^2\right] = \tau_t E(g_t) = \tau_t. \tag{9}$$

One of the present paper's objectives is to compare how macroeconomic variables explain low-frequency volatility and realized volatility. In particular, we claim that breaking down realized volatility into its high- and low-frequency components allows us to better understand the relationship between macroeconomic fundamentals and high-frequency volatility.

Both low-frequency volatility (LV_t) and realized volatility (RV_t) must be computed using the same sampling frequency as the macroeconomic variables, i.e. monthly. To this end, we first define realized variance over a time period $t = 1, ..., T$ as the sum of squared daily returns (again, assuming a zero long-run mean return):

$$\hat{\sigma}_T^2 = \sum_{t=1}^{T} r_t^2.$$

Realized volatility, RV_t, is defined as the square root of the realized variance and can therefore be interpreted as a measure of realized standard deviation, over a given period, e.g. one month or one year. Here for a specific month t, we use observations $d = 1, ..., N$, where N is the number of daily observations in one month. Thus, realized volatility is computed as:

$$RV_t = \sqrt{\sum_{d=1}^{N} r_{t,d}^2}.$$

In contrast, low-frequency volatility, LV_t, is defined as the sample average over a given period, e.g. one month or one year, of the low-frequency component τ_t estimated within the spline-GARCH model. The variable τ_t is sampled daily because this is the frequency used for the GARCH model. Then, low-frequency volatility is defined as:

$$LV_t = \sqrt{\frac{1}{N} \sum_{d=1}^{N} \tau_{t,d}}.$$

The new time series RV_t and LV_t are then used separately in regressions over a set of macroeconomic variables, as described in section 5.

3. Data

We analyze eleven different futures markets that can be categorized into five commodity groups: agricultural (corn, soybeans, wheat), live stock (live cattle, lean hogs), energy (crude oil, natural gas, heating oil), precious metals (gold, silver), and industrial metals (copper). For each commodity, we construct time series of daily settlement prices of the first three nearby contracts from April 1990 through December 2007 by rolling over contracts 15 days prior to their maturity. We combine three price series and run the following regression for each commodity separately:

$$F_{it} = a_i + b_i t + \sum_{j=1}^{3} c_{ij} F_{i,t-j} + e_{it}, \tag{10}$$

where F_{it} is the commodity i's futures price on day t. Because first three nearby contracts' price series are combined for each commodity, there are three price observations on day t. To account for the contemporaneous correlation among the same-day observations we apply the Generalized Least Squares (GLS) method of Karali and Thurman (2009). Briefly, the steps are: (1) run the above regression via Ordinary Least Squares (OLS); (2) compute the variance-covariance matrix of OLS residuals; (3) transform the data through the Cholesky factor of the variance-covariance matrix; (4) run the regression again with the transformed variables; (5) eliminate insignificant regressors and repeat the same procedure with the new set of regressors. This way, we obtain a contemporaneously-uncorrelated GLS residuals for each commodity to use in the spline-GARCH estimation.

To explain the economic determinants of low-frequency volatility, we regress monthly low-frequency volatility obtained from the spline-GARCH estimation on both the fundamental macroeconomic variables that are common across commodities and the commodity-specific variables. Common macroeconomic variables include Consumer Price Index (CPI), real GDP (RGDP), and 3-month Treasury Bill rate (T-Bill). CPI and T-Bill data are available monthly from the Bureau of Labor Statistics and the Board of Governors of the Federal Reserve System, respectively, while RGDP data are available only quarterly from the U.S. Department of Commerce. To match the frequency of observations we interpolate quarterly RGDP series with a cubic spline method and obtain monthly series.

For commodity-specific variables, we consider the level of inventories for the storable commodities. Even though gold and silver are storable, we could not obtain inventory data on these commodities. For corn, soybeans, and wheat we interpolated quarterly inventory series published in Grain Stocks reports by the National Agricultural Statistics Service. For inventories in energy markets, we use monthly series of "U.S. Crude Oil Ending Stocks Excluding SPR," "U.S. Natural Gas Underground Storage Volume," and "Stocks of Distillate and Residual Fuel in the United States" series from the Energy Information Administration. Finally, for copper we use monthly series of "Stocks of Refined Copper in the United States" from the American Bureau of Metal Statistics.

4. Spline-GARCH Estimation

This section presents details on the spline-GARCH approach and the results of the model estimation. For each commodity, we use a pseudo-continuous time series of daily observations, residuals from the GLS decorrelation approach outlined in the previous section. With the contemporaneous correlation removed, the residuals can be assembled into a single time series and used for estimation without concern of a "splicing" bias resulting from combining data from different futures contracts.

The first step is to fit the GLS-detrended residuals to a conditional mean equation. For each commodity, the residuals are fitted to a univariate ARMA filter to remove serial correlation. The number of AR and MA terms is selected using Likelihood Ratio rests, beginning with a (10,10) specification and removing terms until the restriction (fewer terms) cannot be rejected by the LR test. Goodness-of-fit and autocorrelation tests (Ljung-Box) are used to determine that the filtered residuals are close to zero-mean white noise. We then apply Engle's LM test to the filtered residual series separately for all commodities and we conclude that we must reject the null hypothesis of no ARCH at the 99th percentile for all commodities, for three or more lags, and at the 95th percentile for two lags. At one lag, the test results are mixed with no evidence of ARCH for some series. These results are omitted in the interest of conserving space, but are available upon request.

For each commodity, the filtered residuals denoted r_t are fitted to the spline-GARCH model. For most commodities, the optimal number of knots is between nine and twelve, with the exception is high grade copper, for which the optimal number of knots is only four. The interpretation is that copper is less cyclical, and that its cycles are longer. The best specification for the error distribution is Student-t except for gold and heating oil, for which it is Gaussian Normal (see table 1). A (1,1) specification appears to fit the data very well. Likelihood ratio tests are used to obtain the most suitable specification. A model using Gaussian errors is estimated as a restricted case against a model with Student-t errors, as the Gaussian Normal distribution obtains from the Student-t as a limiting case with degrees of freedom ν going to infinity. In most cases, the restriction is rejected by the LR test. Likewise, higher-order GARCH(p,q) specifications are considered but for all commodities we fail to reject the (1,1) model as a restricted case. In many cases, the sum $(\alpha + \beta)$ is almost one, indicating that shocks are very persistent, but not permanent. This is a frequently-observed finding in the asset price volatility literature.

The results presented in table 1 describe the GARCH model parameter estimates, robust standard errors and tests of significance, the preferred distribution of errors (with degrees of freedom ν for the Student-t), as well as the optimal number of spline knots k, and the half-life of shocks (in weeks) as predicted by the model. Standard errors are computed using the Bollerslev-Wooldridge method to provide robustness against misspecification of the errors.

Volatility is highly persistent for many, but not all, commodities. The half-life of a shock is greater than one year for corn, soybeans, natural gas, and copper, and almost one year for crude oil and for gold. For these commodities, the GARCH process is nearly integrated. On the other hand, a shock's half-life is less than a month for wheat, lean hogs, and heating oil. For most commodities, the unconditional variance $\sigma^2 = \frac{\omega}{1-\alpha-\beta}$ is very large, as suggested by the finding of near-integration.

5. Macroeconomic Effects on Low-Frequency Volatility

To study the effects of common- and commodity-specific variables on low-frequency volatility, we run the following regression in Seemingly Unrelated Regressions (SUR) framework:

$$LV_{im} = a_i + b_i CPI_m + c_i lnRGDP_m + d_i TBill_m + s_i DS_m + f_i DF_m + w_i DW_m + h_i S_{im} + e_{im},$$
$$i = 1, 2, \cdots, I, \quad m = 1, 2, \cdots, M, \quad (11)$$

where LV_{im} is the low-frequency volatility of commodity i in month m, CPI_m is the consumer price index in month m, $lnRGDP_m$ is the natural logarithm of the real GDP in month m, $TBill_m$ is the 3-month Treasury Bill rate in month m, and S_{im} is the inventory level of commodity i in month m. DS_m is a dummy variable which takes the value of one if m is in summer quarter (July, August, September), zero otherwise; DF_m is a dummy variable which takes the value of one if m is in fall quarter (October, November, December), zero otherwise; DW_m is a dummy variable which takes the value of one if m is in winter quarter (January, February, March), zero otherwise. We use this simple specification to account for seasonality because the commonly used alternatives, sinusoidal or polynomial functions, require higher-frequency observations. Naturally, for nonstorables and the commodities for which we do not have inventory data, we exclude the inventory term from the regression. The total number of commodities I is 11, and total number of months M is 213, amounting to a total of 2,343 observations in the SUR system.

The results from the SUR estimation are presented in table 2. Except for corn and live cattle, CPI coefficient is statistically significant. However, its estimated sign varies across commodities. For soybeans, wheat, crude oil, natural gas, heating oil, and copper, inflation rate is found to have a positive impact on low-frequency volatility. At times when inflation rate is higher, these commodities experience an increase in their volatility. This finding is consistent with Engle and Rangel (2008), who showed that countries experiencing higher inflation have larger expected volatilities. On the other hand, when the U.S. experiences higher inflation, the low-frequency volatility of lean hogs, gold, and silver are lower. Precious metals are seen as a protection to inflation, therefore it makes sense to observe a negative relationship between CPI and volatility in these markets.

Real GDP is negatively related to volatility of soybeans, wheat, crude oil, natural gas, and heating oil futures. As the economy grows, the low-frequency volatility in these markets fall. However, the volatility in lean hogs and silver futures markets increase as the real GDP increases.

The effect of 3-month Treasury Bill rate on the low-frequency volatility is found statistically significant for all commodities except soybeans and gold. The relationship is positive for wheat, natural gas, and copper, implying higher volatilities in high-interest-rate periods. There is, however, a negative relationship between the risk-free interest rate and the low-frequency volatility for corn, live cattle, lean hogs, crude oil, heating oil in the period. In these markets, low-frequency volatility decreases as interest rates rise.

Seasonal dummy variables are found not to affect low-frequency volatilities. Only summer and fall dummy variables have negative significant coefficient estimates for the wheat futures. This finding is interesting because it is well established in the literature that volatility in grain markets exhibits seasonal pattern due to seasonality in the production cycle. Our results show that seasonality might be inducing high-frequency volatility rather than low-frequency volatility. Other treatments for seasonality, for instance, using periodic functions, might help to clearly identify the low-frequency volatility component contained in seasonal effects.

Interestingly, inventories is an important factor in explaining low-frequency volatility only for wheat, crude oil, and copper. Further, wheat and copper have an opposite sign what the theory of storage predicts. The reason for not finding significant inventory effect in agricultural markets might be again because of the seasonal pattern in inventories. Crop inventories vary much during a year due to the production cycle. However, the inter-year change in crop inventories might not be much pronounced as the intra-year change, and therefore inventories might not be associated with the low-frequency volatility. Inventories might also have a larger effect on high-frequency volatility, which is modeled separately. [1]

We also estimate equation (11) with the realized volatility as the dependent variable. As seen in table 3, fewer coefficient estimates are statistically significant in this case. Decomposing volatility into high- and low-frequency shows that even though the CPI does not affect overall volatility of soybeans, lean hogs, and copper it does affect low-frequency volatility in these markets. Similarly, while log of real GDP does not explain overall volatility in soybeans, wheat, lean hogs, and natural gas futures, it does explain low-frequency volatility. While 3-month Treasury Bill rate is a determinant of overall volatility in only live cattle, crude oil, heating oil, and silver markets, it is an important factor of low-frequency volatility in all but soybeans and gold markets. Inventories inversely affect the realized volatility in natural gas market, confirming the theory of storage. However, they do not affect the

[1] Note that we have considered alternate specifications such as using inverse of inventories, to represent "scarcity", and also log-inventories, to have a more symmetrical distribution in the inventory variable. However, the results do not improve.

low-frequency volatility of natural gas futures. Thus, changes in the natural gas inventories cause changes in the volatility but they are not associated not with the changes in the low-frequency volatility.

In order to determine whether the effects of macroeconomic variables vary within and across commodity groups we perform hypothesis tests. First, we restrict the coefficient estimate of a macroeconomic variable within a commodity group to be the same. For example, we restrict the CPI coefficient to be the same for corn, soybeans, and wheat to represent agricultural commodities group. All within-group restrictions hold for all but seasonal dummy variables. Then, imposing within-group restrictions for all commodity groups at the same time, we test if the parameter estimates across commodity groups are the same. The results from the low-frequency volatility estimation are presented in table 4. The hypothesis that the macroeconomic variables have the same effect on the volatilities of different commodity groups is strongly rejected. The evidence for inventories is a little weaker but still it has a p-value of 0.03.

6. Comparison of Out-of-Sample Forecasting Ability

In the present analysis, observations for 2008 were omitted because it was expected that the commodity bull cycle would bias the underlying fundamental economic relationships, and also because we wanted to reserve some observations for an out-of-sample forecasting comparison between a typical GARCH volatility model and the spline-GARCH model used in the paper. For both models, the objective is to minimize a loss function associated with a k-step ahead out-of-sample forecast of realized volatility.

At the present time, the results are inconclusive. A thorough discussion of out-of-sample forecasting performance will be provided in a future version of this paper.

7. Conclusions

Models of price volatility, being high-frequency, traditionally neglect the impact of lower-frequency influences such as macroeconomic indicators. In this paper, we build on the spline-GARCH model of Engle and Rangel (2008) to obtain a model of high- and low-frequency volatility that includes both common and commodity-specific determinants. We find that using low-frequency volatility is useful firstly because several macroeconomic state variables are found to have a significant effect, which is not the case if one uses a measure of realized volatility. This suggests realized volatility, being a combination of high- and low-frequency components, obscures some important relationships that would otherwise not be detected. Therefore, as found in the case of equities by Engle and Rangel (2008), the spline-GARCH framework appears promising to better understand commodity price volatility.

The tentative answers to the questions we asked in the paper are as follows:

1. How much of the high price volatility can be explained by the common economic factors and how much by the commodity-specific factors? For most commodities, volatility is significantly affected by inflation, economic growth (log real GDP), and the risk-free interest rate (Treasury Bill rate). Volatility for most, but not all, commodities increases with inflation, decreases with economic growth, and decreases with the risk-free rate. Surprisingly, volatility is affected by inventories only in the case of wheat (increasing) and crude oil (decreasing), which suggests that a different specification might be considered.

2. How similar are commodities, in terms of volatility determinants, within sub-groups? To answer this question, we first consider F-tests of the null hypothesis that the effect of a macroeconomic state variable, e.g. inflation, is the same for all commodities in the same commodity subgroup. If we find that we cannot reject the hypothesis, then we consider the hypothesis that the effect is the same across different subgroups. We find that for all commodity subgroups and for all macro variables except for seasonality dummies, we cannot reject the restriction within each subgroup. This implies commodities within the same category, e.g. storable agricultural, are largely affected similarly by macroeconomic influences. Next, we find that we cannot reject the null (at the 95% percentile) that the effect of inventories is the same across commodity groups, but we must reject ($p < 0.0001$) the null of a common effect across commodity groups in the case of inflation (CPI), economic growth (log real GDP), and the risk-free rate (3-month Treasury Bill rate) have different effects across commodity groups.

3. Does the addition of low-frequency components into the high-frequency time series model improve forecasting accuracy? Our preliminary results are inconclusive, but future work will address this question to determine whether macroeconomic variables can be used to improve high-frequency volatility forecasts.

References

Anderson, R.W. 1985. "Some Determinants of the Volatility of Futures Prices." *The Journal of Futures Markets* 5:331–348.

Black, J., and I. Tonks. 2000. "Time Series Volatility of Commodity Futures Prices." *The Journal of Futures Markets* 20:127–144.

Bollerslev, T. 1986. "Generalized Autoregressive Conditional Heteroskedasticity." *Journal of Econometrics* 31:307–327.

Engle, R.F., and J.G. Rangel. 2008. "The Spline-GARCH Model for Low-Frequency Volatility and Its Global Macroeconomic Causes." *The Review of Financial Studies* 21:1187–1222.

Karali, B., and W.N. Thurman. 2009. "Announcement Effects and the Theory of Storage: An Empirical Study of Lumber Futures." *Agricultural Economics,* forthcoming.

Mark, D.R., B.W. Brorsen, K.B. Anderson, and R.M. Small. 2008. "Price Risk Management Alternatives for Farmers in the Absence of Forward Contracts with Grain Merchants." *Choices* 23:22–25.

Pindyck, R.S. 1994. "Inventories and the Short-Run Dynamics of Commodity Prices." *The RAND Journal of Economics* 25:141–159.

Poon, S.H., and C.W.J. Granger. 2003. "Forecasting Volatility in Financial Markets: A Review." *Journal of Economic Literature* 41:478–539.

Schnepf, R. 2008. No. "High Agricultural Commodity Prices: What Are the Issues?", Congressional Research Service (CRS) Report for Congress, updated May 29, 2008. Order Code RL34474.

Table 1: Results of the Spline-GARCH Model Estimation

	Corn	Soybeans	Wheat	Live Cattle	Lean Hogs	Crude Oil	Natural Gas	Heating Oil	Gold	Silver	Copper
ω	8.799 [8.069]	2.0123 [7.716]	30.573 [2.304]	66.550 [1004.85]	41.017 [20.833]	2.4590 [1.535]	32.774 [21.863]	1.8921 [.1798]	0.0001 [.0001]	1.9139 [0.2651]	16.445 [57.842]
α	.1805 [.00020]	.0458 [.00708]	.0027 [.00270]	.4187 [.08543]	.0527 [.00943]	.0382 [.00043]	.4018 [.00083]	.0344 [.00180]	.0570 [.00027]	.0325 [.00015]	.1468 [.00073]
β	.8193 [.00130]	.9532 [.00160]	.8588 [.00001]	.5564 [.00194]	.7124 [.00120]	.9585 [.00017]	.5962 [.00113]	.9321 [.00219]	.9396 [.00028]	.9572 [.00013]	.8507 [.00283]
k	12	13	9	12	12	10	11	10	9	13	4
ν	2.249 [.0005]	2.110 [.0844]	6.460 [.2487]	2.225 [.0172]	2.834 [.0187]	2.100 [.0045]	2.210 [.0003]	∞ N/A	∞ N/A	2.112 [.0006]	2.100 [.0004]
F	t	t	t	t	t	t	t	N	N	t	t
$(\alpha+\beta)$	.999	.999	.861	.975	.765	.996	.998	.966	.997	.990	.998
λ	>1 year	>1 year	1	5.5	0.5	42	> 1 year	4	41	13.4	> 1 year

Notes: The model is $r_t = \sqrt{\tau_t g_t}\epsilon_t$, $g_t = (1-\alpha-\beta)+\alpha\left(\frac{r_{t-1}^2}{\tau_{t-1}}\right)+\beta g_{t-1}$, and $\tau_t = c\exp\left(w_0 t + \sum_{i=1}^{k} w_i\left((t-t_{i-1})_+\right)^2 + z_t\gamma\right)$, estimated individually for all commodity series. Note that g_t is the conditional variance of the innovations given a filtration F_{t-1}, ω is a constant term, β is the GARCH term, α is the ARCH term, F is the distribution of errors, t stands for Student-t, N stands for Gaussian Normal, λ is the half-life of shocks, in weeks. Standard errors are given in the brackets and are computed using Bollerslev and Wooldridge's approach for robustness. All estimates of α, β and ν are significant at the 99th percentile. Estimates of ω are significant at the 99th percentile for wheat, lean hogs, heating oil, copper and silver, but are not significant ($p < 0.10$) for all other commodities.

Table 2: Macro Determinants of Low-Frequency Volatility

	Corn	Soybeans	Wheat	Live Cattle	Lean Hogs	Crude Oil	Natural Gas	Heating Oil	Gold	Silver	Copper
Constant	1.130	2.078	1.935	0.405	-0.475	9.572	5.179	9.882	-0.572	-3.081	0.916
	[0.61]	[0.84]	[0.29]	[0.30]	[0.27]	[1.51]	[0.96]	[1.19]	[0.56]	[0.87]	[0.55]
	(1.84)	(2.47)	(6.55)	(1.34)	(-1.74)	(6.34)	(5.38)	(8.30)	(-1.03)	(-3.52)	(1.66)
CPI	0.060	0.144	0.185	0.040	-0.053	0.595	0.538	0.676	-0.131	-0.264	0.082
	[0.06]	[0.00]	[0.03]	[0.03]	[0.02]	[0.134]	[0.09]	[0.11]	[0.05]	[0.08]	[0.05]
	(1.07)	(1.88)	(6.87)	(1.48)	(-2.15)	(4.34)	(6.15)	(6.25)	(-2.58)	(-3.33)	(1.65)
lnRGDP	-0.122	-0.244	-0.233	-0.041	0.078	-1.129	-0.647	-1.188	0.094	0.404	-0.105
	[0.08]	[0.11]	[0.04]	[0.04]	[0.03]	[0.19]	[0.12]	[0.15]	[0.07]	[0.11]	[0.07]
	(-1.58)	(-2.29)	(-6.26)	(-1.09)	(2.28)	(-5.94)	(-5.34)	(-7.92)	(1.33)	(3.67)	(-1.51)
T-Bill	-0.310	0.148	0.164	-0.481	-0.161	-0.863	0.507	-0.687	0.085	-0.458	0.446
	[0.13]	[0.00]	[0.06]	[0.06]	[0.06]	[0.33]	[0.21]	[0.259]	[0.12]	[0.19]	[0.12]
	(-2.32)	(0.80)	(2.62)	(-7.35)	(-2.71)	(-2.63)	(2.42)	(-2.65)	(0.71)	(-2.41)	(3.79)
Summer	0.002	0.002	-0.006	0.000	0.000	0.007	0.007	0.009	0.002	0.001	-0.000
	[0.01]	[0.01]	[0.00]	[0.00]	[0.00]	[0.01]	[0.01]	[0.01]	[0.00]	[0.01]	[0.00]
	(0.26)	(0.23)	(-1.98)	(0.11)	(0.05)	(0.54)	(0.77)	(0.80)	(0.46)	(0.14)	(-0.09)
Fall	0.000	-0.001	-0.008	0.000	0.000	0.008	0.010	0.010	0.004	0.000	-0.002
	[0.06]	[0.01]	[0.00]	[0.00]	[0.00]	[0.01]	[0.01]	[0.01]	[0.00]	[0.01]	[0.00]
	(0.04)	(-0.18)	(-2.35)	(0.19)	(0.07)	(0.59)	(0.97)	(0.97)	(0.73)	(0.06)	(-0.35)
Winter	0.000	-0.002	-0.003	0.000	-0.001	0.005	0.010	0.009	0.002		-0.003
	[0.01]	[0.01]	[0.00]	[0.00]	[0.00]	[0.01]	[0.01]	[0.01]	[0.00]		[0.00]
	(0.08)	(-0.20)	(-0.95)	(0.07)	(-0.25)	(0.38)	(1.11)	(0.85)	(0.43)		(-0.58)
Inventories	0.000	0.000	0.006			-0.028	0.000	0.001			0.003
	[0.00]	[0.00]	[0.00]			[0.00]	[0.00]	[0.00]			[0.00]
	(0.39)	(0.87)	(3.17)			(-5.71)	(0.17)	(0.13)			(9.12)

Notes: The model is $LV_{im} = a_i + b_i CPI_m + c_i lnRGDP_m + d_i TBill_m + s_i DS_m + f_i DF_m + w_i DW_m + h_i S_{im} + e_{im}$, for $i = 1, 2, \cdots, I$, $m = 1, 2, \cdots, M$. LV_{im} is the low-frequency volatility of commodity i in month m, CPI_m is the consumer price index in month m, $lnRGDP_m$ is the natural logarithm of the real GDP in month m, $TBill_m$ is the 3-month Treasury Bill rate in month m, and S_{im} is the inventory level of commodity i in month m. DS_m is a dummy variable which takes the value of one if m is July, August, or September, and zero otherwise; DF_m is a dummy variable which takes the value of one if m is October, November, or December, and zero otherwise; DW_m is a dummy variable which takes the value of one if m is January, February, or March, and zero otherwise. The model is estimated via Seemingly Unrelated Regressions method. Standard errors and t-values of estimates are given in the brackets and parentheses, respectively.

Table 3: Macro Determinants of Realized Volatility

	Corn	Soybeans	Wheat	Live Cattle	Lean Hogs	Crude Oil	Natural Gas	Heating Oil	Gold	Silver	Copper
Constant	1.086	1.676	1.632	0.820	-0.052	7.604	3.593	8.247	-0.607	-3.279	0.280
	[1.04]	[1.11]	[1.03]	[0.76]	[1.40]	[1.86]	[2.07]	[1.69]	[0.79]	[1.45]	[1.41]
	(1.04)	(1.51)	(1.58)	(1.08)	(-0.04)	(4.10)	(1.74)	(4.89)	(-0.76)	(-2.27)	(0.20)
CPI	0.065	0.112	0.162	0.086	-0.014	0.440	0.397	0.543	-0.130	-0.282	0.013
	[0.09]	[0.10]	[0.09]	[0.07]	[0.13]	[0.16]	[0.19]	[0.15]	[0.07]	[0.13]	[0.13]
	(0.69)	(1.11)	(1.71)	(1.24)	(-0.11)	(2.68)	(2.12)	(3.54)	(-1.80)	(-2.14)	(0.10)
lnRGDP	-0.121	-0.194	-0.196	-0.096	0.023	-0.898	-0.423	-0.981	0.097	0.429	-0.021
	[0.13]	[0.14]	[0.13]	[0.10]	[0.18]	[0.23]	[0.26]	[0.213]	[0.10]	[0.18]	[0.18]
	(-0.92)	(-1.38)	(-1.50)	(-1.00)	(0.13)	(-3.89)	(-1.63)	(-4.60)	(0.97)	(2.35)	(-0.12)
T-Bill	-0.023	0.124	0.161	-0.285	-0.219	-1.072	0.466	-0.671	0.049	-0.572	0.067
	[0.23]	[0.25]	[0.22]	[0.166]	[0.30]	[0.39]	[0.45]	[0.37]	[0.17]	[0.31]	[0.29]
	(-0.10)	(0.50)	(0.74)	(-1.72)	(-0.72)	(-2.71)	(1.04)	(-1.83)	(0.28)	(-1.82)	(0.23)
Summer	-0.001	0.006	-0.011	-0.009	-0.009	-0.009	0.026	-0.016	0.004	-0.004	0.010
	[0.01]	[0.01]	[0.01]	[0.01]	[0.01]	[0.02]	[0.02]	[0.02]	[0.01]	[0.01]	[0.01]
	(-0.13)	(0.51)	(-0.93)	(-1.35)	(-0.71)	(-0.58)	(1.05)	(-0.96)	(0.52)	(-0.33)	(1.05)
Fall	0.016	0.007	0.005	-0.002	0.000	-0.005	0.038	-0.016	-0.005	-0.014	-0.010
	[0.01]	[0.01]	[0.01]	[0.01]	[0.01]	[0.02]	[0.03]	[0.02]	[0.01]	[0.01]	[0.01]
	(1.60)	(0.60)	(0.35)	(-0.27)	(0.03)	(-0.34)	(1.13)	(-0.92)	(-0.74)	(-1.08)	(-1.01)
Winter	0.005	0.016	0.010	-0.010	0.015	-0.001	0.020	-0.001	-0.008	-0.013	-0.013
	[0.01]	[0.01]	[0.01]	[0.01]	[0.01]	[0.02]	[0.02]	[0.02]	[0.01]	[0.01]	[0.01]
	(0.50)	(1.28)	(0.99)	(-1.39)	(1.14)	(-0.04)	(0.96)	(-0.07)	(-1.14)	(-0.98)	(-1.31)
Inventories	-0.000	-0.000	-0.005			0.011	-0.038	-0.023			0.002
	[0.00]	[0.00]	[0.01]			[0.02]	[0.02]	[0.03]			[0.00]
	(-0.07)	(-0.23)	(-0.55)			(0.51)	(-2.15)	(-0.67)			(1.15)

Notes: The model is $RV_{im} = a_i + b_i CPI_m + c_i lnRGDP_m + d_i TBill_m + s_i DS_m + f_i DF_m + w_i DW_m + h_i S_{im} + e_{im}$, for $i = 1, 2, \cdots, I$, $m = 1, 2, \cdots, M$. RV_{im} is the realized volatility of commodity i in month m, CPI_m is the consumer price index in month m, $lnRGDP_m$ is the natural logarithm of the real GDP in month m, $TBill_m$ is the 3-month Treasury Bill rate in month m, and S_{im} is the inventory level of commodity i in month m. DS_m is a dummy variable which takes the value of one if m is July, August, or September, and zero otherwise; DF_m is a dummy variable which takes the value of one if m is October, November, or December, and zero otherwise; DW_m is a dummy variable which takes the value of one if m is January, February, or March, and zero otherwise. The model is estimated via Seemingly Unrelated Regressions method. Standard errors and t-values of estimates are given in the brackets and parentheses, respectively.

Table 4: F-tests on Restrictions Across Commodity Groups

	F-statistic	p-value
CPI	93.00	0.0001
lnRGDP	100.49	0.0001
T-Bill	66.23	0.0001
Inventories	3.52	0.0297

Notes: The F-tests are performed on the parameter estimates from the low-frequency volatility estimation.