NBER WORKING PAPER SERIES

THE EXCESS CO-MOVEMENT OF COMMODITY PRICES

Robert S. Pindyck

Julio J. Rotemberg

Working Paper No. 2671

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 1988

This work was supported by M.I.T.'s Center for Energy Policy Research, and by the National Science Foundation under Grant No. SES-8618502 to R. Pindyck and Grant No. SES-8619004 to J. Rotemberg. We want to thank Michael Boozer and Rebecca Emerson for their superb research assistance, and John Campbell, Thomas Stoker, and Kenneth West for helpful comments and suggestions. This research is part of NBER's research program in Financial Markets and Monetary Economics. Any opinions expressed are those of the authors not those of the National Bureau of Economic Research.

NBER Working Paper #2671
July 1988

THE EXCESS CO-MOVEMENT OF COMMODITY PRICES

ABSTRACT

This paper tests and confirms the existence of a puzzling phenomenon - the prices of largely unrelated raw commodities have a persistent tendency to move together. We show that this co-movement of prices is well in excess of anything that can be explained by the common effects of past, current, or expected future values of macroeconomic variables such as inflation, industrial production, interest rates, and exchange rates. These results are a rejection of the standard competitive model of commodity price formation with storage

Robert S. Pindyck
Sloan School of Management
Massachusetts Institute of Technology
50 Memorial Drive
Cambridge, MA 02139

Julio J. Rotemberg
Sloan School of Management
Massachusetts Institute of Technology
50 Memorial Drive
Cambridge, MA 02139

1. Introduction.

This paper tests and confirms the existence of a puzzling phenomenon-the prices of raw commodities have a persistent tendency to move together. We find that this co-movement of prices applies to a broad set of commodities that are largely unrelated, i.e., for which the cross-price elasticities of demand and supply are close to zero. Furthermore, the co-movement is well in excess of anything that can be explained by the common effects of inflation, or changes in aggregate demand, interest rates, and exchange rates.

Our test for excess co-movement is also a test of the standard competitive model of commodity price formation with storage. An innovative aspect of our test, and one that distinguishes it from, say, Eichenbaum's (1983, 1984) tests of finished goods inventory behavior under rational expectations, is that we do not need data on inventory stocks. Our test relies instead on the joint behavior of prices across a range of commodities, and the fact that those prices should only move together in response to common macroeconomic shocks.

This excess co-movement casts doubt on the standard competitive commodity price model. A possible explanation for it is that commodity price movements are to some extent the result of "herd" behavior in financial markets. (By "herd" behavior we mean that traders are alternatively bullish or bearish on <u>all</u> commodities for no plausible economic reason.) Indeed, our finding would be of little surprise to brokers, traders, and others who deal regularly in the futures and cash markets, many of whom have held the common belief that commodity prices tend to move together. Analyses of futures and commodity markets issued by brokerage firms, or that appear on the financial pages of newspapers and magazines, refer to copper or oil or coffee prices rising because commodity prices in general are rising, as though increases in those prices are caused by or have the same causes as increases in wheat, cotton, and gold prices.

To conclude that prices exhibit excess co-movement, we must account for the effects of any common macroeconomic shocks. Current and expected future values of macroeconomic variables such as inflation, industrial production, etc., should have common effects on current and expected future demands (and possibly supplies) of commodities, and hence on current prices. For example, a rise in interest rates should lower commodity prices; higher interest rates depress future aggregate demand and hence commodity demands, and raise commodity carrying costs. At issue is whether the prices of unrelated commodities tend to move together after accounting for these macroeconomic effects. We find that they do.

The next section discusses our data set, and the nature of the price correlations. As we will see, price changes are correlated, and the correlations are larger the longer the intervals across which the changes are measured. In Section 3 we try to explain these correlations using a simple regression model. We find that after allowing for the common effects of current and past values of economic variables, there is still a great deal of correlation that remains. One possible explanation is that commodity demands and supplies are affected by unobserved forecasts of the economic variables. In Sections 4 and 5 we show how a latent variable model can be used to test this possibility. We find that latent variables representing unobserved forecasts of inflation and industrial production are indeed significant explanators of commodity prices. However, even after allowing for these latent variables, there is still excess co-movement left over. Section 6 concludes by discussing some limitations of our analysis and possible reasons for our findings.

2. The Correlation of Commodity Prices.

We study monthly price changes for seven commodities: wheat, cotton, copper, gold, crude oil, lumber, and cocoa. This is a broad spectrum of commodities that are as unrelated as possible. For example, the agricultural products we have chosen are grown in different climates and have different uses. None of the commodities are substitutes or complements, none are co-produced, and none is used as a major input for the production of another. Barring price movements due to common macroeconomic factors, we would expect these prices to be uncorrelated.[1]

We use United States average monthly cash prices from 1960 through 1985. Ideally, these data should correspond to a current price quotation for immediate delivery of a homogeneous good. However, all commodities are at least somewhat heterogenous, and delivery dates can vary. We have tried to obtain price data that reflect as closely as possible what sellers are charging at the time for current delivery of a well-specified commodity. Specific price series and data sources are listed in Appendix B.

Table 1 shows a correlation matrix for the monthly changes in the logarithms of these prices. Ten out of the 21 correlations exceed .1. Gold shows strong correlations with copper, crude oil, lumber, and cocoa; cotton is also correlated with copper, lumber, and wheat; and lumber is correlated with copper and cocoa.

Are these correlations as a group statistically significant? To answer this we can perform a likelihood ratio test of the hypothesis that the correlation matrix is equal to the identity matrix. It is worth discussing this test briefly because it is closely related to the tests we carry out in later sections of the paper. Consider m jointly normal random variables whose theoretical covariance matrix is given by Σ. The matrix Σ incorporates whatever restrictions are implied by the theory that is being tested, e.g., Σ would be a diagonal matrix when the variables are uncorrelated. Denote by $\hat{\Sigma}$ the maximum likelihood estimate of Σ, and let Ω be the actual covariance matrix of the variables. Then the likelihood of the data under the theoretical restrictions is given by:

$$L = |\hat{\Sigma}|^{-N/2} e^{-(N/2)\mathrm{tr}(\Sigma^{-1}\Omega)} \qquad (1)$$

where N is the number of observations. In the special case in which Σ is diagonal, the elements of $\hat{\Sigma}$ equal the corresponding elements of Ω, so that $\mathrm{tr}(\hat{\Sigma}^{-1}\Omega)$ is simply equal to m. The likelihood of the data absent any restrictions is given by (1), but with Ω substituted for $\hat{\Sigma}$.

In the case of a diagonal covariance matrix, the likelihood ratio is $|\Omega|^{N/2}$ divided by the product of the variances, also to the N/2 power. As shown in Morrison (1967), this implies that the ratio of the restricted and unrestricted likelihood functions is $\lambda = |R|^{N/2}$, where $|R|$ is the determinant of the correlation matrix. Our test statistic is therefore $-2\log\lambda$, which is distributed as χ^2 with $(1/2)p(p-1)$ degrees of freedom, where p is the number of commodities. For the seven commodities in our sample, this statistic is 114.6. With 21 degrees of freedom, this is highly significant, so we can easily reject the hypothesis that these commodity prices are uncorrelated.

The correlations of commodity price changes are much larger for longer holding periods. Tables 2 and 3 show correlations for (nonoverlapping) quarterly and annual changes, respectively, in the logarithms of prices. Observe that for annual changes, 19 out of 21 correlations exceed .2. As the χ^2 statistics below each table show, as a group the correlations remain significant at the 1 percent level. Nonetheless, the significance level for the quarterly and annual changes are lower than for the monthly ones. This occurs because there are many fewer nonoverlapping yearly than monthly observations.

A better measure of the statistical significance of the quarterly and yearly correlations is obtained by using all of the available data, i.e., using overlapping observations. χ^2 statistics computed as above using all overlapping observations give values of 194.9 for quarterly differences and 517.7 for annual differences. These statistics are not distributed as $\chi^2(21)$ because the use of overlapping data introduces serial dependence. We therefore computed, via Monte Carlo, 15,000 draws of our test statistics, $-2\log\lambda$, under the null hypothesis that the monthly price changes are i.i.d. and uncorrelated across commodities. The highest volumes that we drew for these statistics were 121.3 for quarterly price changes, and 504.1 for yearly price changes. Thus these quarterly and annual correlations that we observe are highly significant.

Of course these correlations might be due to common macroeconomic factors, such as changes in current or expected future inflation or aggregate demand. In addition, macroeconomic variables may explain more of the price movements over longer horizons, which may account for the larger correlations that we find for longer holding periods. We explore these possibilities below.

3. The Explanatory Power of Current and Past Macroeconomic Variables.

Commodity prices may have common movements because of changes in macroeconomic variables that affect demands and/or supplies for broad sets of commodities. These changes can affect prices in two ways. First, macroeconomic variables may directly affect commodity demands and supplies.

For example, an increase in the rate of industrial production will raise the demands for industrial commodities such as copper, lumber, or crude oil because these commodities are used as inputs to production, and will raise the demands for non-industrial commodities such as cocoa or wheat through the resulting increases in income.

Second, macroeconomic variables can affect commodity prices by affecting expectations about future supplies and demands. Commodities are storable, so changing expectations about future market conditions influence the demand for storage and hence current prices. This means that unexpected changes in macroeconomic variable which are useful for forecasting can have an immediate effect on commodity prices. For example, higher interest rates might reduce capital investment by suppliers of a number of commodities, thereby reducing future supplies and raising current prices. In addition, a change in interest rates might change expectations about future aggregate economic activity, which would affect expected future commodity demands, and again, current prices.

We can formalize these arguments with a simple model.[2] Write the <u>net</u> supply of commodity i at time t, $Q_{i,t}$, as:

$$Q_{i,t} = a_{i,t} + b_i p_{i,t} \tag{2}$$

where $p_{i,t} = \log P_{i,t}$, and $P_{i,t}$ is the price of commodity i at t. The index $a_{i,t}$ captures changes in both supply and demand. It depends on both commodity specific variables (e.g., a strike by copper miners or bad weather), as well as current and lagged values of x_t, a vector of macroeconomic variables (such as the index of industrial production, interest rates, inflation, etc.) that can affect many commodities. We define a set of commodities to be unrelated if there are negligible cross-price effects (so that $a_{i,t}$ does not include the prices of other commodities), and if any commodity specific variable that affects $a_{i,t}$ does not affect $a_{j,t}$, $j \neq i$.

The evolution of inventory, $I_{i,t}$, is given by the accounting identity:

$$I_{i,t} = I_{i,t-1} + Q_{i,t} \tag{3}$$

Finally, under the assumption that risk-neutral inventory holders maximize expected profits, the evolution of the price of commodity i is given by:

$$r_t = [E_t P_{i,t+1} - C_{i,t} - P_{i,t}]/P_{i,t} \tag{4}$$

where r_t is the required rate of return, E_t is the expectation conditional on all information available at time t, and $C_{i,t}$ is the one-period holding cost of the commodity, less the capitalized flow of its marginal convenience yield over the period.

The convenience yield is the flow of benefits that one obtains from holding stocks, e.g., the resulting assurance of supply as needed, ease of scheduling, etc. On the margin, this depends on the total quantity of inventory held; the larger is $I_{i,t}$, the smaller is the benefit from holding an extra unit of inventory. The convenience yield is also likely to depend on macroeconomic variables.[3] For example, an increase in the rate of industrial production implies an increase in the rate of consumption of industrial commodities, and therefore an increase in desired stocks. We model $c_{i,t}$, the logarithm of $C_{i,t}$, as a linear function of $I_{i,t}$:

$$c_{i,t} = n_{i,t} + \gamma_i I_{i,t} \tag{5}$$

where $n_{i,t}$ is a function of current and past values of x_t, the vector of macroeconomic variables.

Eqn. (4) says that prices at t depend on expected future prices. Thus current prices depend on expected future conditions in the industry, and as we show in Appendix A, they are functions of current and expected future values of x_t. We assume that forecasts of x_t are based on current and past values of x_t, and also on current and past values of a vector z_t of exogenous economic variables that do not directly affect commodity prices (e.g., the money supply and the stock market):

$$E_t x_{t+j} = \theta_j(L)x_t + \phi_j(L)z_t \tag{6}$$

As the Appendix shows, this leads to the following estimating equation:

$$\Delta p_{i,t} = \sum_{k=0}^{K} \alpha_{ik}\Delta x_{t-k} + \sum_{k=0}^{K} \beta_{ik}\Delta z_{t-k} + \epsilon_{i,t} \tag{7}$$

where $\epsilon_{i,t}$ is serially uncorrelated, and under our null hypothesis, $E(\epsilon_{i,t}\epsilon_{j,t}) = 0$ for all $i \neq j$.

It is possible for serial correlation to arise in $\epsilon_{i,t}$. We explore this by also estimating the following equation:

$$\Delta p_{i,t} = \sum_{k=0}^{K} \alpha_{ik}\Delta x_{t-k} + \sum_{k=0}^{K} \beta_{ik}\Delta z_{t-k} + \rho_i \Delta p_{i,t-1} + \epsilon_{i,t} \tag{7'}$$

The details of our model not withstanding, eqns. (7) and (7') embody a simple notion: the prices of unrelated commodities should move together exclusively in response to market participants' changing perceptions of the macroeconomic environment.

Estimation.

We estimate eqns. (7) and (7') for each of our seven commodities using OLS for the period April 1960 through November 1985. Since the results on the correlation of commodity prices are nearly the same for the two specifications, we report largely on the estimation of (7) to avoid duplication. The vector x_t includes the index of industrial production (Y), the consumer price index (π), an (equally weighted) index of the dollar value of British pounds, German marks, and Japanese yen (E), and the nominal interest rate on 3-month Treasury bills (R).[4] The vector z_t includes the money supply, M1 (M), and the S&P Common Stock Index (S). The model is first estimated with the current and one-month lagged values of these variables, and then is re-estimated with the current values and lags of one through six months.

Table 4 shows estimation results for equations that include x_t and z_t current and lagged one month. Increases in inflation and the money supply are associated with increases in the prices of all the commodities, and the interest rate with decreases. The effects of the other variables are more mixed, but as Table 5 shows, each variable has a statistically significant impact on commodity prices as a whole. That table presents likelihood ratio tests for group exclusions of explanatory variables from all seven commodity price equations. Column (1) applies to equations with one lag, and column (2) to equations with six lags. Each statistic is twice the difference of the log likelihood functions for the unrestricted and restricted models, and is distributed as χ^2 with degrees of freedom equal to the number of restrictions (14 and 49 respectively). With the exception of stock returns in column (1) and industrial production in column (2), these statistics are significant at the 1 percent level.

Denote by $\hat{\epsilon}_t$ the vector of residuals $(\hat{\epsilon}_{1,t}, \ldots, \hat{\epsilon}_{7,t})'$, and let Ω be the covariance matrix of $\hat{\epsilon}$. If our model is complete, Ω should be diagonal. We test whether Ω is indeed diagonal using the technique decribed in Section 2; the results are included in Table 5. The test statistic is significant at the 1 percent level for both versions of the model. The data reject a diagonal covariance matrix more strongly when we include six lags of the explanatory variables. This might occur because in small samples the addition of irrelevant explanatory variables automatically reduces the variance of the $\hat{\epsilon}_i$'s without necessarily reducing the covariances commensurately.

To account for serial correlation in the residuals (as reflected in the Durbin-Watson statistics in Table 4), we also estimated eqn. (7′), which includes a lagged dependent variable. To test for excess co-movement in this case, we cannot utilize the technique employed in Section 2. Instead we compare the likelihoods of models estimated both with and without the constraints imposed. Including the current values and one lag of the explanatory variables, the likelihood ratio test for a diagonal residual correlation matrix is 71.2. This is lower than for the regressions shown in Table 4, but still highly significant.

These results show that excess co-movement is statistically significant, but tell us little about its magnitude. In particular, we would like to know how much of the total variation in commodity prices is explained by this co-movement. This can be determined by comparing the R^2's for the OLS regressions in Table 4 with R^2's for regressions which explain the price change of each commodity using the current changes in the prices of all of the other commodities as additional explanatory variables. These R^2's are shown in Table 6. Except for gold, crude oil, and lumber, the R^2's for the monthly regressions on the macro variables are low; most of the variance of price changes is unexplained. When commodity prices are added as explanatory variables, the R^2's increase substantially, and for wheat, cotton, and copper, the change in R^2 (which measures the marginal explanatory power of commodity co-movements) exceeds the R^2 when only macro variables are included.

Table 6 also shows the R^2's for estimates of eqn. (7) using non-overlapping quarterly and annual data. The explanatory variables are the same, but now we use quarterly and annual changes in the logs of prices, industrial production, the money supply, etc. The marginal explanatory power of commodity price co-movements tends to increase when we use quarterly and annual data.[5] In the case of cotton, for example, the addition of other commodity price changes as explanatory variables accounts for nearly half of the total variation in annual cotton price changes. Table 6 thus shows that commodity price co-movements explains a substantial fraction of the individual price movements.

We also examined the sensitivity of our results to the choice of sample period, using monthly data and one lag of each explanatory variable. Leaving out the period October 1973 through December 1974 (during which commodity prices may have been broadly affected by OPEC, which may have also affected macroeconomic variables), the statistic for the absence of co-movements falls to 77.1. Extending the sample through October 1986 results in a statistic of 75.4, and shortening the sample so that it ends in November 1984 gives 83.0.[6] These statistics are all highly significant.[7]

After accounting for commodity price movements that are due to common macroeconomic factors, price changes remain correlated across commodities. We make a further attempt to account for this the next two sections.

4. A Latent Variable Model.

In the previous section we tested whether correlations among commodity prices can be attributed to the correlation of each price with observable macroeconomic variables that are predictors of future conditions in commodity markets. This approach is subject to a serious limitation: Individuals have more information about future x's than can be obtained from any set of current and past x's and z's which are directly observable. Thus eqn. (6) is too restrictive. Some of the news about future macroeconomic variables is of a qualitative nature which is difficult to include in the kinds of regressions reported above. This qualitative information could in principle affect all commodities and could thus be a source of correlation among their prices.

A natural way of capturing such information about the future is by incorporating a set of latent variables into our model. These latent variables represent the market's forecasts of the future values of the macroeconomic variables. Our model then becomes a MIMIC (multiple indicator multiple cause) model.[8] The "indicators," i.e., the variables which are affected by the latent variables, include both the vector of commodity prices and the actual realization of the future macroeconomic variables. The "causes" of the latent variables include any variable which is useful in forecasting macroeconomic variables. Thus the causes include our z's.

To account for market information that is unavailable to us, we first generalize eqn. (6):

$$E_t(\Delta x_{t+j}) = \theta_j(L)\Delta x_t + \phi_j(L)\Delta z_t + f_j v_t \tag{8}$$

$E_t(\Delta x_{t+j})$ is an unobserved forecast of Δx_t based on the observed current and past values of Δx_t and Δz_t, <u>and</u> on the unobserved residual vector v_t. We now consider a subset of the variables x, which we denote by y. We define the vector of latent variables J_t as follows:

$$J_t = E_t(\Delta y_{t+1}) = \theta'(L)\Delta x_t + \phi'(L)\Delta z_t + f' v_t \tag{9}$$

We now make the strong assumption that f' is of full rank. This means that

$$E_t(\Delta x_{t+j}) = \theta_j''(L)\Delta x_t + \phi_j''(L)\Delta z_t + f_j'' J_t. \tag{10}$$

In other words, knowledge of J_t is sufficient, when combined with the observable x's and z's, to generate forecasts of x_{t+j}, $j \geq 1$. We can then write the log change in the price of commodity i (which depends on all future x's) as:

$$\Delta p_{i,t} = \sum_{k=0}^{K} \alpha_{ik}\Delta x_{t-k} + g_i J_t + \epsilon_{i,t} \tag{11}$$

where g_i is a vector of coefficients.

The latent variables we include are the expectation at t of the value at t+1 of y. Therefore, the vector of residuals w_t in the equation

$$\Delta y_{t+1} = J_t + w_t \tag{12}$$

is uncorrelated with any information available at t. The system we estimate then consists of (9), (11), and (12). The vector of latent variables J has multiple causes, namely the z's, and multiple indicators, namely the current prices and future y's.

Our procedure is closely related to the more traditional instrumental variables method of estimating rational expectations models. Consistent estimates of g_i could also be obtained by using the current and lagged z's as instruments for Δy_{t+1} in a regression equation which is given by (11), where J_t is replaced by Δy_{t+1}. As in the instrumental variables approach, we assume that certain variables (the z's) affect commodity prices only through their effect on agents' expectations of certain future variables.

Like our procedure, the instrumental variables approach gives consistent estimates of g_i, even when the instrument list is not exhaustive. However, the residuals from an instrumental variables regression cannot be used directly to test for excessive co-movement of commodity prices. These residuals are constructed using the actual realized values of future macroeconomic variables. Since the market forecast must by necessity differ from these realized values, the residuals in all the equations will tend to be correlated.

We estimate (9), (11) and (12) by maximum likelihood, under the maintained assumption that the v's, w's and ϵ's are normally distributed. The contemporaneous variance-covariance matrix for the v's as well as that for the w's is left unrestricted. We assume that v's are uncorrelated with ϵ's and w's at all leads and lags, and that the same is true for the correlation between ϵ's and w's. We first estimate the model under the assumption that the covariance matrix for the ϵ's is diagonal so that our explanatory and latent variables account for all of the correlation in commodity prices. This assumption is then tested by reestimating the model with an unrestricted contemporaneous covariance matrix for the ϵ's.

We use the same variables as in the regression model of Section 3, and include two latent variables which represent the current forecasts of next period's inflation and next period's rate of growth of the Index of Industrial Production. Thus we are assuming that the money supply and the stock market affect commodity prices only via their ability to predict inflation and output.[9]

Estimation is done using LISREL.[10] Besides yielding parameter estimates, LISREL computes the value of the likelihood function given by eqn. (1), making likelihood ratio tests straightforward.

5. The Explanatory Power of Latent Variables.

Estimation results for this latent variable model are presented in Table 7. The latent variables η_π and η_y represent the market's forecasts of inflation between period t and period t+1, and growth in industrial production between t and t+1 respectively. The first seven columns of Table 7 represent the equations explaining commodity prices while the last two columns represent the equations explaining the latent variables.

As this table shows, the latent variables help explain commodity prices. In the regressions explaining prices, both latent variables have generally positive and often statistically significant coefficients. To see that the latent variables are important, note that the R^2's are much higher when the latent variables are included than in the corresponding equations of Table 4.

After estimating the model with the constraint that the covariance matrix of the ϵ's is diagonal, we reestimate it without that constraint. Even this less constrained model now incorporates some constraints since we still assume that the v's and w's are uncorrelated with the ϵ's and that the z's affect prices only through the latent variables. We test these secondary restrictions by constructing a likelihood ratio statistic which compares our less contrained model with an unconstrained alternative. This statistic is distributed as $\chi^2(25)$ when the restrictions are valid.[11] We obtain a value of 35.5, which is insignificant at the 5 percent level.

Having estimated both the restricted and less restricted models, we do a likelihood ratio test on the restrictions implied by a diagonal covariance matrix. The test statistic is 49.7 This statistic, which measures the extent to which the 21 restrictions on the off diagonal elements are violated, is smaller than the value of 88.6 that we obtained in the OLS case, but is still significant at the 1% level. Thus, even after including latent variables there is still excess co-movement of commodity prices.

We estimated several variations of this basic model, including two models with only one latent variable. The first has a latent variable for the market forecast of future inflation, and the second has a latent variable for the market forecast of growth in industrial production. The statistics of the hypothesis of no excess co-movement, which again are distribued as $\chi^2(21)$ under the null, are 48.2 and 57.0 for the first and second models respectively. Thus, forecasted inflation has more to do with joint movements of commodity prices than does forecasted production growth. Also note that the evidence against the hypothesis of no excess co-movement is slightly weaker when we include only the latent variable for inflation than when we include both. This means that simply adding latent variables may not resolve the puzzle of excess co-movement. This can occur because while the addition of latent variables raises the likelihoods of both the constrained and unconstrained models, it is the difference between these likelihoods that corresponds to our test statistic.[12]

We also estimated a latent variable version of eqn. (7') which includes lagged dependent variables. The results change very little. The test statistic for the absence of co-movement remains equal to 49.7 when there are two latent variables, and becomes 46.3 and 40.3 respectively when the only latent variables are expected inflation and the expected change in industrial production. Finally, we tried to extend the number of lags, but failed to achieve convergence of the likelihood function, presumably because of the large number of unimportant parameters being estimated.

6. Concluding Remarks.

Common movements in the prices of unrelated commodities should be traceable to changes in current or expected future values of macroeconomic variables. We have shown that these kinds of variables do not account for much of the observed co-movement of commodity prices. This is the case whether expectations are based solely on observable macroeconomic variables, or are also based on unobserved latent variables.

There are two possible explanations for this finding. One is that our model is incomplete - some important macroeconomic variables are missing from our specification. Given our extensive experimenting we doubt that this is the case, but this possibility cannot be ruled out. The other explanation is that the actors in commodity markets react in tandem to noneconomic factors. These reactions might be due to the presence of equilibrium "sunspots", "bubbles," or simply changes in "market psychology".

In any case, this would represent a rejection of the standard competitive model of commodity price formation in the presence of storage.

There are also alternative explanations for the dependence of our results on the length of the holding period. We have shown that as we increase the interval over which price changes are measured from a month to a quarter or a year, the amount of price movements which can be attributed to macroeconomic variables rises while the amount of unexplained co-movement rises as well.

One possible reason for this finding is that there is considerable high-frequency mean-reverting noise in individual commodity prices. As a result neither macroeconomic variables nor prices of other commodities explain a large fraction of individual monthly price changes.

A second possibility is consistent with the view that we have excluded relevant macroeconomic variables from our model. Suppose that changes in macro variables affect commodity price slowly. For example, an unusual monthly change in inflation might have to persist for some time before it affects perceptions about the future. Such slow effects are consistent with our finding that macro variables explain more of the movements in commodity prices over longer holding periods. Then any excluded macro variable will also explain more of the price movements for longer holding periods. This means that its exclusion increases the unexplained co-movement as the holding period is increased.

A third possibility is that common price movements are the result of liquidity effects. The fall in the price of one commodity lowers the price of others only because it impoverishes speculators who are long in several commodities at once. These liquidity effects should be larger the larger is the change in any single commodity's price. The variance of price changes is larger the longer the horizon, so we would expect liquidity effects to become more significant as the horizon increases.

A fourth possibility is that commodity prices are indeed driven partly by bubbles and fads, and that these fads are sufficiently rare that the bulk of our monthly price changes is unaffected by them. At the same time, these fads would have to be sufficiently disruptive that they explain a substantial amount of commodity price changes over longer horizons.

More research is needed to test these various hypotheses. Hopefully additional work will help to disentangle the causes of the excess co-movement of commodity prices that we have found.

APPENDIX A

Here we derive eqn. (7) from eqns. (2) through (6) and a linearization. In particular we use a linearization of (4) analogous to that employed by Campbell and Shiller (1986) to obtain a linear expression for the logarithm of price. Ignoring commodity specific subscripts, eqn. (4) becomes:

$$1 + r_t + \epsilon_t = R_t = [P_{t+1} - C_t]/P_t = (P_{t+1}/C_{t+1})(C_{t+1}/C_t)(C_t/P_t) - C_t/P_t$$

where R_t is the ex post return and ϵ_t can be thought of as the unexpected return. The logarithm of R_t is approximated at the point where C_t/P_t equals a constant h and C_{t+1}/C_t equals a constant s. Then

$$\log(R_t) \approx s - h + [(p_{t+1}-p_t-s)s - (c_t-p_t-h)h]/(s-h)$$

where c = logC and p = logP. Therefore, linearizing the log of $(1+\delta_t+\epsilon_t)$, eqn. (4) can be approximated as:

$$r_t \approx E_t\delta p_{t+1} - p_t - 2h + (1-\delta)c_t \qquad \text{(A1)}$$

where $\delta = s/(s-h)$. Using (5), we now have:

$$E_t\delta p_{i,t+1} - p_{i,t} - 2h + (1-\delta)[n_{i,t} + \gamma_i I_{i,t}] - r_t \approx 0 \qquad \text{(A2)}$$

To simplify notation, we now subsume variations in the discount rate r in $n_{i,t}$ (so that $n_{i,t}$ corresponds to $[n_{i,t}-r_t/(1-\delta)]$).

To complete the model we also require a transversality condition:

$$\lim_{T\to\infty} \delta^{(T-t)}E_t I_{i,T} = 0$$

Combining (2), (3), and (A2) gives a difference equation for $I_{i,t}$:

$$E_t I_{i,t+1} - \frac{(1+\delta+b_i\gamma_i)}{\delta} I_{i,t} + \frac{1}{\delta} I_{i,t-1} = a_{i,t+1} - \frac{1}{\delta}a_{i,t} - \frac{1}{\delta}b_i n_{i,t} \qquad \text{(A3)}$$

By factoring eqn. (A3), one can show that its non-explosive solution is:

$$I_{i,t} = k_i I_{i,t-1} + d_i E_t \sum_{j=0}^{\infty} d_i^j (a_{i,t+j} - \delta a_{i,t+j+1} + b_i n_{i,t+j}) \qquad \text{(A4)}$$

where k_i and d_i are commodity-specific constants which lie between 0 and 1 and depend on b_i, γ_i, and δ. Eqn. (A4) describes the change in inventories in terms of current and expected future values of $a_{i,t}$ and $c_{i,t}$. To see that price is also a function of current and expected future values of $a_{i,t}$ and $n_{i,t}$, combine eqns. (2), (3) and (A4):

$$p_{i,t} = \frac{1}{b_i}\left[(k_i-1)I_{i,t-1} + d_i E_t \sum_{j=0}^{\infty} d_i^j (a_{i,t+j} - \delta a_{i,t+j+1} + b_i n_{i,t+j}) - a_{i,t}\right] \quad (A5)$$

Recall that $a_{i,t}$ and $c_{i,t}$ both depend on current and lagged values of x_t. Therefore, $p_{i,t}$ depends on expected future values of x_t, so that an equation is needed to forecast x_t. Assuming that forecasts of future x's are based on (6)) we obtain:

$$p_{i,t} = \sum_{k=0}^{K} \alpha_{ik} x_{t-k} + \sum_{k=0}^{K} \beta_{ik} z_{t-k} + u_{i,t} \quad (A6)$$

The error term $u_{i,t}$ includes all commodity-specific factors, including the inventory level $I_{i,t-1}$, i.e., it includes all factors <u>not</u> explained by the macroeconomic variables x_t. For example, in the case of copper, $u_{i,t}$ might include current and past reserve levels, shocks accounting for strikes, etc. Thus under our null hypothesis, the $u_{i,t}$'s are uncorrelated across commodities. We assume that the $u_{i,t}$'s follow a random walk, so that $E_t(u_{i,t+j}) = u_{i,t}$ for $j > 0$, and changes in $u_{i,t}$ are serially uncorrelated. This leads to eqn. (7) in the text. Since the $u_{i,t}$'s could in principle have a richer temporal structure, we also allow for serial correlation by introducing a lagged dependent variable as in eqn. (7').

APPENDIX B

Monthly cash price data for January 1960 through December 1985 came from the following sources:

Cocoa: Through January 1985, Bureau of Labor Statistics, "Spot Cocoa Bean Prices in New York." February 1985 onwards, average daily cash price quoted in Chicago for Accra delivery.

Copper: Commodity Yearbook, "Producers' Prices of Electrolytic (Wirebar) Copper, Delivered U.S. Destinations," American Metal Market. Data are monthly averages of daily wholesale delivered cash prices.

Cotton: Commodity Yearbook, "Average Spot Price of U.S. Cotton, 1-1/16 inches, Strict Low Middling at Designated Markets, Agricultural Marketing Service, USDA.

Crude Oil: Platts Oil Price Handbook and Oilmanac, Annual Editions, "Average Wholesale Price of Crude Petroleum as Collected by the Independent Petroleum Association of America."

Gold: Handy and Harmon cash price. A monthly average of daily spot prices.

Lumber: Bureau of Labor Statistics, "Aggregate Price Index for Lumber and Primary Lumber Products."

Wheat: Commodity Yearbook, "Average Price of Number 1 Hard Winter Wheat, at Kansas City," Agricultural Marketing Service, USDA.

REFERENCES

Abel, Andrew B. and Olivier J. Blanchard, "The Present Value of Profits and Cyclical Movements in Investment," Econometrica, March 1986, 54, 249-74.

Aigner, Dennis J., Cheng Hsiao, Arie Kapteyn, and Tom Wansbeek, "Latent Variable Models in Econometrics," in Z. Griliches and M. Intriligator, Eds., Handbook of Econometrics, Vol. II, North-Holland, 1984.

Campbell, John Y., and Robert J. Shiller, "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors," NBER Working Paper No. 2100, Dec. 1986.

Eichenbaum, Martin, "A Rational Expectations Equilibrium Model of Inventories of Finished Goods and Employment," Journal of Monetary Economics, 1983, 12, 259-277.

Eichenbaum, Martin, "Rational Expectations and the Smoothing Properties of Inventories of Finished Goods," Journal of Monetary Economics, 1984, 14, 71-96.

Goldberger, A. S., "Maximum Likelihood Estimation of Regressions Containing Unobservable Independent Variables," International Economic Review, 1972, 13, 1-15.

Joreskog, Karl G., and Dag Sorbom, "LISREL User's Guide," 1986.

Morrison, Donald F., Multivariate Statistical Methods, McGraw-Hill, 1967.

Stein, Jerome L., The Economics of Futures Markets, Basil Blackwell, 1986.

Turnovsky, Stephen J., "The Determination of Spot and Futures Prices with Storable Commodities," Econometrica, September 1983, 51, 1363-87.

Williams, Jeffrey, "Futures Markets: A Consequence of Risk Aversion or Transactions Costs?" Journal of Political Economy, October 1987, 95, 1000-1023.

FOOTNOTES

1. Limited experimentation with other sets of commodities, including replacing gold with platinum, had little effect on our results.

2. This model is similar in structure to the finished goods inventory model of Eichenbaum (1983). It is also similar to the commodity price models of Stein (1986) and Turnovsky (1983), but more general in that they assume i.i.d. shocks, and we allow for a more general error structure.

3. For an explicit model of convenience yield that illustrates some of these general dependencies, see Williams (1987).

4. The interest rate is in level rather than first-differenced form. This is consistent with the first difference of the interest rate affecting the rate of change of commodity prices. We include the level of interest rates because it may well be a good predictor of future inflation and because equation (4) suggests that levels of interest rates may help predict individual commodity price changes.

5. The R^2's for the regressions that use only macroeconomic explanatory variables increase substantially as we lengthen the holding period, which partly explains the larger raw correlations of commodity price changes for longer holding periods shown in Tables 1 to 3.

6. We focus on the 1960:4 to 1985:11 period because of the major change in U.S. government intervention in the cotton market that occured in 1986.

7. We also considered the weather as an explanator that could affect all commodities, and included U.S. data on heating degree days, cooling degree days, temperature, and precipitation. This had virtually no effect on our results; the resulting χ^2 was 87.7.

8. See Goldberger (1972) and Aigner et. al. (1984).

9. In some sense this is more restrictive than in the earlier regression model because there the money supply and the stock market were potential predictors of all other x's as well.

10. The input is the correlation matrix Ω of all the variables of interest. Thus this matrix includes the correlations among the changes in commodity prices, the x's , the z's and the future values of inflation and production growth. See Joreskog and Sorbom (1986).

11. Ignoring the x's, the model has 7 prices, 2 future macroeconomic variables, and 4 instruments, for a total of 78 covariances. The test statistic for the less restricted model has 25 degrees of freedom because that model includes 53 free parameters: 21 covariances of the ϵ_i's, 14 g_i's in eqn. (11), 8 ϕ's in eqn. (9), 3 elements of the covariance matrix for eqn. (9), 1 covariance of the w_t's in (12), and the 6 free covariances of our instruments.

12. Eliminating the latent variable for industrial production makes the fit of the less constrained model deteriorate substantially. The test statistic for this model relative to an unconstrained alternative is 51.8. This is significant at the 1 percent level since this less constrained model imposes 27 restrictions. Thus there is less evidence against the hypothesis that money and the stock market affect commodity prices through forecasts of both inflation and output growth than there is against the hypothesis that they do so through only one of these forecasts.

TABLE 1

Correlations of Monthly Log Changes in Commodity Prices

	WHEAT	COTTON	COPPER	GOLD	CRUDE	LUMBER	COCOA
WHEAT	1.000						
COTTON	0.253	1.000					
COPPER	0.051	0.152	1.000				
GOLD	-0.020	0.045	0.322	1.000			
CRUDE	0.103	0.098	0.032	0.245	1.000		
LUMBER	-0.059	0.125	0.113	0.126	-0.085	1.000	
COCOA	-0.014	0.043	0.052	0.135	0.013	0.122	1.000

$$\chi^2(21) = 114.6$$

TABLE 2

Correlations of Nonoverlapping Quarterly Log Changes in Commodity Prices

	WHEAT	COTTON	COPPER	GOLD	CRUDE	LUMBER	COCOA
WHEAT	1.000						
COTTON	0.300	1.000					
COPPER	0.095	0.254	1.000				
GOLD	0.136	0.138	0.391	1.000			
CRUDE	0.142	0.063	0.018	0.419	1.000		
LUMBER	0.023	0.225	0.152	0.212	-0.096	1.000	
COCOA	0.050	0.085	0.228	0.214	-0.043	0.302	1.000

$$\chi^2(21) = 53.5$$

TABLE 3

Correlations of Nonoverlapping Annual Log Changes in Commodity Prices

	WHEAT	COTTON	COPPER	GOLD	CRUDE	LUMBER	COCOA
WHEAT	1.000						
COTTON	0.504	1.000					
COPPER	0.430	0.352	1.000				
GOLD	0.606	0.462	0.521	1.000			
CRUDE	0.354	0.246	0.325	0.548	1.000		
LUMBER	0.313	0.458	0.099	0.275	-0.176	1.000	
COCOA	0.272	0.289	0.241	0.233	-0.030	0.582	1.000

$\chi^2(21) = 56.3$

TABLE 4
OLS Regressions

(t-statistics in parenthesis)

	WHEAT	COTTON	COPPER	GOLD	CRUDE	LUMBER	COCOA
π	.273 (3.1)	-.081 (-0.9)	.070 (0.8)	.135 (1.7)	.333 (4.1)	-.079 (-1.0)	-.064 (-0.7)
$\pi(-1)$	-.161 (-1.8)	.204 (2.3)	-.009 (-0.1)	.203 (2.5)	.170 (2.1)	.155 (1.9)	.120 (1.4)
Y	-.001 (-0.01)	.080 (1.2)	.027 (0.4)	-.058 (-1.0)	-.088 (-1.4)	.040 (0.6)	.124 (1.9)
Y(-1)	.082 (1.3)	.045 (0.7)	.055 (0.9)	-.070 (-1.2)	-.051 (-0.9)	.066 (1.1)	.109 (1.7)
R	-.007 (-0.02)	.165 (0.4)	.421 (1.1)	-.009 (-0.03)	-.466 (-1.3)	.321 (0.9)	.264 (0.7)
R(-1)	-0.76 (-0.2)	-.254 (-0.7)	-.485 (-1.4)	-.268 (-0.8)	.298 (0.9)	-.508 (-1.5)	-.303 (-0.8)
E	-.056 (-0.9)	-.077 (-1.2)	.141 (2.2)	.325 (5.5)	-.146 (-2.4)	-.002 (-0.0)	.068 (1.1)
E(-1)	-.019 (-0.3)	.070 (1.1)	.067 (1.1)	-.064 (-1.1)	.033 (0.6)	.158 (2.7)	.051 (0.8)
M	.133 (2.0)	-.039 (-0.6)	.207 (3.2)	.120 (2.1)	.001 (0.002)	.182 (3.0)	.026 (0.4)
M(-1)	-.045 (-0.7)	.088 (1.3)	-.063 (-1.0)	.175 (2.8)	.061 (1.0)	.064 (1.0)	.018 (0.3)
S	-.003 (-0.05)	.094 (1.5)	.050 (0.8)	.077 (1.4)	.111 (1.9)	.053 (0.8)	.081 (1.3)
S(-1)	-.084 (-1.3)	-.044 (-0.7)	-.119 (-1.9)	-.097 (-1.7)	-.145 (-2.5)	.082 (1.4)	-.029 (-0.5)
R^2	.06	.05	.09	.24	.21	.18	.07
DW	1.34	1.2	1.48	1.40	1.51	1.10	1.87

TABLE 5
χ^2 Statistics for Group Exclusions of the Explanatory Variables

	(1) χ^2 with 14 degrees of freedom, 1 lag of each variable	(2) χ^2 with 49 degrees of freedom, 6 lags of each variable
(1) INF	73.22**	127.29**
(2) INDST	29.48**	71.56*
(3) TBILL	29.32**	93.24**
(4) EXCH	62.06**	166.41**
(5) MI	36.29**	81.93**
(6) STOCK	20.44	101.05**
Diagonal Correlation Matrix:	89.36**	99.44**

* Significant at 5% level
** Significant at 1% level

TABLE 6

R^2's With and Without Commodity Prices as Additional Explanatory Variables

Holding Period:	MONTHLY		QUARTERLY		ANNUAL	
Dependent Variable	Without	With	Without	With	Without	With
WHEAT	.056	.135	.136	.237	.577	.889
COTTON	.053	.154	.278	.367	.401	.887
COPPER	.090	.181	.179	.233	.701	.906
GOLD	.244	.333	.373	.398	.908	.960
CRUDE	.211	.261	.402	.482	.944	.974
LUMBER	.177	.187	.279	.298	.753	.818
COCOA	.069	.085	.159	.209	.872	.890

TABLE 7
Latent Variable Model

	WHEAT	COTTON	COPPER	GOLD	CRUDE	LUMBER	COCOA	η_π	η_y
η_π	1.334 (1.8)	1.483 (1.9)	2.037 (2.1)	1.876 (2.1)	2.247 (2.3)	-1.648 (-1.2)	0.563 (0.9)		
η_y	-0.262 (-0.6)	0.249 (0.5)	0.611 (1.1)	0.703 (1.3)	-0.324 (-0.6)	2.290 (2.4)	0.345 (0.9)		
π	-0.308 (-0.9)	-0.698 (-2.0)	-0.773 (-1.8)	-0.641 (-1.5)	-0.636 (-1.4)	0.689 (1.1)	-0.294 (-1.0)	0.425 (7.9)	-0.033 (-0.4)
$\pi(-1)$	-0.577 (-2.0)	-0.226 (-0.7)	-0.529 (-1.5)	-0.288 (-0.8)	-0.540 (-1.5)	0.892 (1.7)	0.014 (0.1)	0.298 (5.5)	-0.111 (-1.5)
Y	0.123 (0.7)	0.046 (0.3)	-0.110 (-0.5)	-0.216 (-1.0)	0.087 (0.4)	-0.699 (-2.0)	0.034 (0.2)	-0.031 (-0.8)	0.315 (5.6)
Y(-1)	0.059 (0.6)	-0.061 (-0.6)	-0.095 (-0.8)	-0.238 (-2.2)	-0.125 (-1.1)	-0.094 (-0.6)	0.043 (0.6)	0.047 (1.2)	0.108 (2.0)
R	-0.920 (-1.6)	-0.887 (-1.4)	-1.564 (-2.1)	-1.349 (-1.9)	-1.875 (-2.5)	0.476 (0.5)	-0.351 (-0.7)	0.697 (3.4)	0.428 (1.5)
R(-1)	0.651 (1.2)	0.618 (1.2)	1.268 (1.9)	0.911 (1.4)	1.378 (2.1)	-0.231 (-0.3)	0.265 (0.6)	-0.564 (-2.8)	-0.525 (-1.9)
E	-0.200 (-1.8)	-0.236 (-2.0)	-0.101 (-0.7)	0.131 (1.0)	-0.375 (-2.6)	0.151 (0.8)	0.005 (0.1)	0.106 (2.7)	0.012 (0.3)
E(-1)	0.125 (1.1)	0.199 (1.7)	0.256 (1.8)	0.103 (0.8)	0.248 (1.7)	-0.039 (-0.2)	0.097 (1.0)	-0.096 (-2.5)	0.017 (0.3)
M								0.034 (1.5)	0.104 (2.9)
M(-1)								0.021 (1.1)	0.048 (1.6)
S								0.035 (2.0)	0.034 (1.2)
S(-1)								-0.054 (-2.2)	0.009 (0.2)
R^2	0.08	0.13	0.26	0.39	0.31	0.39	0.09	0.65	0.35