Futures Basis for Cotton: Impact of Globalization and Structural Change

Stephen MacDonald and Leslie Meyer
Economic Research Service, USDA
stephenm@ers.usda.gov
lmeyer@ers.usda.gov

Poster prepared for presentation at the Agricultural & Applied Economics Association's 2009 AAEA and ACCI Joint Annual Meeting, Milwaukee, WI, July 26-28, 2009

Abstract

A model of commodity futures contract basis was developed based on Working's theory of the price of storage. An error-correction model was estimated for the basis for the InterContinental Exchange (ICE) #2 cotton contract maturing in December during 2000-08. The model was also extended to incorporate the impact of changes in market activity that evolved as financial markets and commodity price behavior underwent significant changes after 2005. The model captured the inversion of basis following the collapse of China's crop in 2003, but the shock realized during 2008 may have been in part driven by one-time events not included in the model. Estimates from the error-correction model suggest an extended period for the return of basis to equilibrium, spanning from about 1 ½ to 2 months.

Introduction

Basis is more stable and predictable than either cash or futures prices, and being able to predict the basis is the key to effective hedging. In turn, a futures market that meets the needs of commercial participants in physical commodity markets will be attractive to speculators. The additional liquidity speculators bring to the market sets up a virtuous cycle of effective price discovery and improved opportunities for both hedging and speculating.

At its simplest, the magnitude of the basis for a storable commodity's futures contract is a function of the time to the contract's expiration. Defining basis as the difference at time t between the current price of a futures contract (F_t) for delivery at time T ($T > t$) and the price of the same commodity on the spot market (S_t)(see Table 1 for definitions of futures market terminology used in this study):

$B_t = F_t - S_t$ Equation 1

At T, the owner of the futures contract could meet the delivery obligation with commodities purchased at time t at the price S_t and stored until the time of delivery. While only about 3 percent of futures contracts traditionally result in delivery, the possibility of delivery is regarded as key to the functioning of such markets. Assuming the spot price was not anticipated to change between t and T, then the futures price could equal the current spot price plus the cost of storage over T–t. This is consistent with "risk-management perspective" on futures markets identified by Williams (2001). In this perspective, cash and futures markets are autonomous, and risk averse hedgers purchase risk-management services from speculators. Keynes' theory of "normal backwardization" posits such a mechanism, although backwardization ($S_t > F_t$) is far from normal. In his survey of the literature on futures markets, Williams points out that the large number of papers on "optimal hedging" suggests that this risk-management perspective predominates among economists. However, Williams notes that the empirical support of this perspective is weak.

The alternative in William's dichotomy is the "arbitrage perspective." Probably one of the most important contributions to the theoretical understanding of futures markets was

Working's theory of price of storage (Working 1949). This critique of Keynes' theory culminated in,

> "..the idea that the primary function of commodity futures markets was the provision of returns for storage services, and [Working] viewed inter-temporal prices as the jointly determined price of storage."
>
> (Carter 1999)

The history of futures markets shows they emerge under the circumstances of poorly functioning spot and forward markets (Williams). This suggests that the view of spot and futures markets as autonomous misses would obscure some fundamental truths, particularly in the study of basis.

Working's theory also guides understanding of the existence of infrequent but large instances of backwardization—or steep inversions—embedded in more typical periods when $F_t > S_t$ (referred to as "contango," a term of British origin). Figure 1, with the basis on the InterContinental Exchange's (ICE) December No. 2 cotton futures contract for the years 2001-08, illustrates the characteristics of futures contract basis as described thus far. As expiration approaches each December, the difference between the futures and spot price typically diminishes. During most of the contract's lifetime, $F_t > S_t$. However, the December 2004 contract in the middle of the graph is a clear exception. Even when $T >> t$, there were occasions when $F_t < S_t$, implying a negative return to storage. However, cotton stocks, while relatively low that year, did not fall to zero either. Working's innovation was to expand the definition of basis described above to account for the services that accrue to the holder of stocks. This "convenience yield" can be considered a negative cost to stockholding, and when current supplies of a commodity are low, the value of this benefit can more than offset the positive costs of holding stocks.

Carter's 1999 survey of the literature on futures markets illustrates how the theory of the price of storage leads to the specification of a general basis model. The equilibrium relationship between the futures price and the spot price is:

$$F_{t,T} = S_t(1+r_{t,T}) + w_t + c_t \qquad \text{Equation 2}$$

Where $r_{t,T}$ is the opportunity cost of using funds to own commodity inventory from time t to T, w_t is the physical cost of storage, and c_t is the convenience yield.

Tomek (1997) clarifies Working's theory of the price of storage by noting that the cost of providing storage of a commodity includes the opportunity cost of holding the commodity, and the cost of inputs to maintain a storage facility (and account for loss and insurance), but is partly offset by a benefit to holding a commodity. This benefit varies nonlinearly with the inventory. Tomek specifies the relationship as a logarithmic function, so the first order conditions of the storage profit function specifies the convenience yield with a term that includes the inverse of inventory. Figure 2 illustrates the resulting relationship between basis (the price of supplying storage) and inventory. When inventories are large, basis is substantially a function of storage costs (opportunity cost and physical cost). When inventories are very low, the convenience yield more than offsets these costs, and the basis reverses sign, going into backwardization or inversion.

This insight allows us to use Equations 1 and 2 to specify a model of basis that is a function of interest rates, time, and the inverse of end-of-year inventories (Model 1):

$$B_t = \beta_0 + \beta_1 r_t S_t(T\text{-}t) + \beta_2 (T\text{-}t) + \beta_3 SU^{-1} + \varepsilon_t \qquad \text{Equation 3}$$

Where r_t is the interest rate at time t, the daily cost of physical storage equals β_2, and the convenience yield's effect comes into the model as a function of the inverse of the marketing year ending stocks as a ratio of annual use of the commodity. While Tomek's model included the simple inverse level inventory, the level of cotton stocks has different implications for the services that commodity ownership can have depending on the level of demand. This normalization is particularly important given the large structural changes the U.S. cotton market has endured with the end of the Multifibre Arrangement in 2005, the widespread adoption of genetically-modified cotton around the world, and the shocks introduced by the rising importance of biofuels and the gyrations of the financial sector in recent years.

While there is a longstanding tradition of blaming inexplicable price movements on speculators (Williams, p. 799), the crucial role of speculation in well functioning futures markets is widely recognized. During 2008, rising levels of commodity prices were linked by some to increased speculation and the growing role of long-only index funds.

While the impact of speculation on cash markets is clearly limited by the cost of taking physical ownership of commodities, Irwin, Garcia, and Good (1997) have pointed out how demand by long speculators could raise basis during periods close to delivery. For grains, convergence behavior near the time of contract expiration has altered in recent years, at least coinciding with the growing role of these new market participants. Figure 3 illustrates how open interest surged in cotton futures markets after 2005. While Figure 1 suggests that cotton's December contract convergence at expiration has not significantly changed, it is possible that there may have been impacts on basis earlier in the contract's life.

If open interest is added to the model as a shift variable accounting for market activity, the result is Model 2:

$$B_t = \beta_0 + \beta_1 r_t S_t(T\text{-}t) + \beta_2 (T\text{-}t) + \beta_3 SU^{-1} + \beta_3 Z_t + \varepsilon_t \qquad \text{Equation 4}$$

Where Z_t is the open interest in all cotton futures contracts at t. Open interest is measured as the sum of open interest across all expiration months. Open interest on the contract for individual months fluctuates, rising in steps as more nearby contracts approach expiration and market participants uninterested in delivery shift their interest to a later contract (Figure 4). To some extent, the shift into a new nearby contract also induces shifts across the spectrum of distant contracts, creating a complex pattern. Rather than model and/or adjust for these seasonal patterns, a total open interest variable was used to measure market activity.

The Commodity Futures Trading Commission expanded its Commitments of Traders Report in 2006 to include data on index funds. While the role of index funds is not examined explicitly in this model, the implications of the available data will be discussed in the conclusions.

Data

This study focuses on the December contract alone. This contract has by far the largest volume and open interest of any expiration month throughout most of the period before it becomes the nearby contract (Figure 4). It is a key hedging contract for U.S. producers and merchants. In the United States, cotton is largely harvested during October and November, making the December contract the reference for marketing the crop. Daily data for the December ICE # 2 cotton contract are used for the futures price (F_t) and open interest (Z_t), and daily U.S. average spot prices from USDA's Agricultural Marketing Service are the representative cash market prices (S_t). The 3-month London Interbank Offer Rate on U.S. dollar deposits is used as the representative interest rate (r_t: using the 3-month U.S. Treasury Bill rate as an alternative resulted in no significant differences). Data are from December 2000 to December 2008, covering the final year's trading in the December contracts expiring in 2001-2008. The data was structured as a panel with each equation specified for one year, and the panel's cross-section is across years.

The inverse stocks-to-use variable is calculated with annual data, and captures the fixed effects between years as well as the convenience yield. Data on daily estimates of stocks-use are unavailable, necessitating the use of a single estimate for each year. The estimate used is the current end-of-year estimate derived from USDA's PS&D database (U.S. Department of Agriculture). To account for the changes in expected stocks-use that did actually occur through a given year, and to capture the declining impact of convenience yield as the contract approaches maturity, the variable is interacted with quarterly dummies.

The magnitude of the expected cumulative physical cost of storage is captured by a trend variable that registers the number of days until contract expiration.

The behavior of futures and basis changes significantly as contract expiration draws near (MacDonald and Meyer). There are many idiosyncratic factors specific to a given year that influence basis that late in a contract's lifetime. This large variation near delivery is partly accounted for by the very large importance of product characteristics in cotton compared with grains and oilseeds. Therefore, this study's models were estimated with data that extended to T = 21 days before expiration in each year. This cutoff point was a function of the large decline observed in open interest after this point as the usefulness of the contract for the majority of market participants diminishes.

Error-Correction Model Specification

The time series properties of the data were examined first. Panel unit root tests were performed using EViews 6, and the Fisher Augmented Dickey-Fuller (Maddala and Wu, 1999) test indicated the variables were integrated of order 1. All the variables were tested under both the assumption that they both did and did not include drift in their data generating process (Table 2). Previous research has identified a trend in the basis data (MacDonald and Meyer), and the null hypothesis (Ho) of non-stationarity could not be rejected for the level of basis when a trend in the data was assumed. The opportunity cost and market activity variables were not tested for trends, but the null hypothesis of non-stationarity could not be rejected whether a trend was assumed or not for the data in levels. In each case, the Ho was rejected for the data in first differences. Therefore, all the variables were determined to be nonstationary, and I(1).

The nonstationarity of the variables indicated that it would be useful to estimate an error-correction model (Engle and Granger, 1987), if cointegration indicated the presence of a stable long run relationship. Table 3 gives the results of the panel cointegration tests using the combined Johanson test results for the panel (Maddala and Wu, 1999). In each case, the tests suggest the presence of a unique cointegrating equation. In each case, the Johanson test was performed under the assumption that the data had a linear trend. For Model 1, the null hypothesis of no cointegrating equations was clearly rejected. For Model 2, this null was rejected at the 5 percent level in the maximum eigenvalue test, but could only be rejected at the 8 percent level using the trace test. Given the mixed results, the variables were treated as cointegrated.

Granger causality testing with 10 trading day lags was undertaken to determine the degree of exogeneity. A lag of 10 trading days was chosen to allow an average of 2 weeks for the realization of impacts between variables (Table 4). Basis does not Granger-cause opportunity cost according these results, although causality flows from opportunity cost to basis. Basis and open interest can only reject the null hypothesis of no causality at the 10-percent level. However, the null hypothesis can be strongly rejected in each direction for open interest and opportunity cost. This suggests that future work could include more detailed modeling of the interactions of the variables than undertaken here.

The ECM was modeled in two steps (Engle and Granger, 1987), with a long-run equilibrium model estimated first. This was followed by a difference model using the lagged error of the long-run model as an explanatory variable. The addition of the second-step disequilbrium component results in the following full models:

Model 1

$B_t = \beta_0 + \beta_1 r_tS_t(T\text{-}t) + \beta_2 (T\text{-}t) + \beta_3 SU^{-1} + \varepsilon_t$ Equation 3

$\Delta B_t = \gamma_0 + \gamma_1 \Delta r_tS_t(T\text{-}t) + \theta (\varepsilon_{t-1})$ Equation 5

Model 2

$B_t = \beta_0 + \beta_1 r_tS_t(T\text{-}t) + \beta_2 (T\text{-}t) + \beta_3 SU^{-1} + \beta_3 Z_t + \varepsilon_t$ Equation 4

$\Delta B_t = \gamma_0 + \gamma_1 \Delta r_tS_t(T\text{-}t) + \gamma_2 \Delta Z_t + \theta (\varepsilon_{t-1})$ Equation 6

Results

For Model 1 and Model 2, the significance of the variables is high, and the signs are largely as expected (Table 5). The one exception is opportunity cost, which is negative in each model. As noted earlier, this is the case regardless of whether interest rates are measured with the LIBOR or the risk-free asset in the form of Treasury Bills. During the recent financial crisis, the behavior of these two interest rates was quite distinct, but the choice of one or the other has little impact on the estimated parameters. Hranaiova and Tomek also found negative opportunity costs in some of their estimated basis equations, but positive values are still more intuitive. Convenience yield was negative in each quarter, and the magnitude of the first quarter's parameter was about double that of the fourth quarter, consistent with expectations.

Storage costs were estimated at about 0.04 cents per day, which would translate to either $5.28 (Model 1) or $6.22 (Model 2) per bale per month. This is larger than data on monthly charges gathered by the Farm Service Agency (average of $2.62 during 2004-08) (USDA Farm Service Agency, 2009). However, fixed in and out charges averaged $9.01 during this time, and storage is unlikely to be a large number of months during most years. There actual average storage costs over this storage period would be above the marginal cost of $2.62 Alternatively, this high estimate could reflect some problems with using a simple trend to capture these costs, and may also relate to the counter-intuitive sign on the opportunity cost variable.

The parameter estimate for market activity suggests the rise in open interest between 2000-01 and 2007-08 may have accounted for a 3.8-cent increase in basis between those years. This would be 60 percent of the increase in average basis between those years of 6.2 cents (from 5.2 cents to 11.4 cents). As mentioned earlier, a fuller accounting of the dynamics between open interest and opportunity cost may be appropriate, and could alter this estimate.

The speed of adjustment to equilibrium can be drawn from the error-correction component of the model as $-\theta^{-1}$. For Model 2, the implication is that the basis adjusts back to equilibrium after 44 days, a nontrivial amount of time, but not wildly inconsistent with speed-of-adjustment estimates in other markets and circumstances (Table 6). For Model 1, adjustment is estimated to take 68 days. One possibility is that a richer lag structure needs to be examined for the equilibrium model and the ECM. Given that including the market activity variable lowered the estimated time for adjustment, additional examination of the specification may result in further changes to the model that reveals more rapid price discovery in the market than these estimates suggest.

Conclusions

Previous work has suggested that cotton basis was trending upwards during 2003-07 (MacDonald and Meyer, 2009). The model developed during this study suggests this trend could in part be a function of the structural change in world financial markets that brought new participants and increased participation by traditional participants to cotton futures trading. The model also accounts for the significant deviation from this trend that occurred in 2003. The globalization of the world economy which had brought China into world markets exposed cotton basis to the shock of a shortfall in China's cotton production that year. Figure 5 illustrates how the 2004 contract was an outlier both with respect to the level of basis and the level of stocks. China's net imports surged to more than twice their previous record that year, depleting U.S. stocks, and boosting the convenience yield of inventory.

During 2008, basis was higher than in 2007, but the degree of the increase that year was quite distinct from the behavior during previous years. At its peak, basis for the December 2008 contract was more than double its typical levels. While as early as December 2007, the December 2008 basis already exceeded the levels achieved during any day of the final year of trading for the 2001-07 December contracts, during March 2008, basis swiftly rose another 39 percent, and remained close to this peak for more than 1 month. Subsequently, basis then declined to below average levels. Large stocks in 2008 would suggest convenience yield would be small, consistent with a basis of large magnitude, but the hypothesized non-linear impact of stocks on basis is asymptotic (Figure 2), and after a certain threshold, the impact of additional stocks has little or no effect.

U.S. cotton futures trading transitioned completely to electronic trading in March 2008. Industry reports that this may have interacted with the 2004 change in exchange rules that determined that margin calls could be based on synthetic prices derived from trading in options on some occasions when futures prices were constrained by the limits on daily changes. As synthetic prices shot to nearly unprecedented levels, and buying on both options and futures markets to offset short positions drove prices sharply higher, possibly accounting for the sharply higher basis at that time (Wachovia Securities, 2009).

References

Carter, Colin, "Commodity Futures Markets: A Survey," *Australian Journal of Agricultural and Resource Economics* 43(2): 209-247.

Chicago Board of Trade, *Commodity Trading Manual*. 1985.

Enders, Walter, *Applied Econometric Time Series*, Wiley, New York. 1995.

Engle, R.F., and C.W.J. Granger, "Co-integration and Error Correction: Representation, Estimation, and Testing," *Econometrica* 53: 251-76, 1987.

Hranaiova, J., and W.G. Tomek, "Role of Delivery Options in Basis Convergence*," The Journal of Futures Markets,* Vol. 22, No. 8. Wiley Periodicals, Inc. 2002.

Irwin, S.H., P. Garcia, and D.L. Good, "The Performance of Chicago Board of Trade Corn, Soybean, and Wheat Futures Contracts After Recent Changes in Speculative Limits," University of Illinois at Urbana-Champaign, May 2007.

MacDonald, S., and L. Meyer, "Basis Trends in U.S. Cotton Futures," *Proceedings of the 2009 Cotton Beltwide Conferences*. National Cotton Council. (Forthcoming).

Maddala, G. S. and S. Wu (1999). "A Comparative Study of Unit Root Tests with Panel Data and A New Simple Test," *Oxford Bulletin of Economics and Statistics*, 61, 631–52.

Tomek, W., "Commodity Futures Prices as Forecasts," *Review of Agricultural Economics*, Vol 19(1): 23-44. 1997

U.S. Department of Agriculture, Farm Service Agency (2009). Personal communication from Timothy Murray.

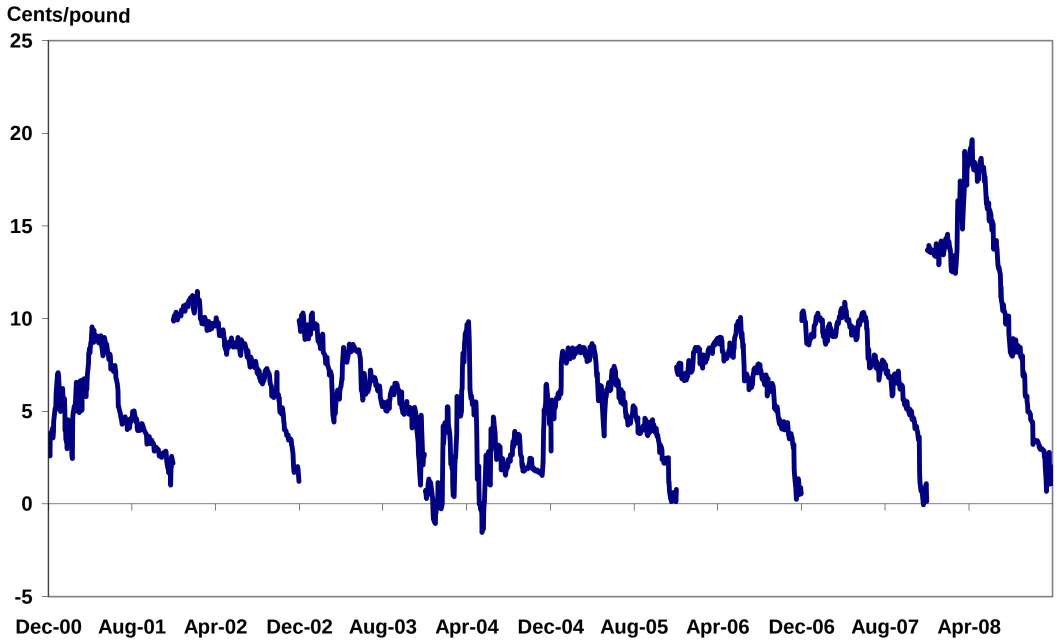
U.S. Department of Agriculture, PS&D Online, available at: http://www.fas.usda.gov/psdonline/

Williams, J. (2001), "Commodity Futures and Options." In B.L. Gardner and G.C. Rausser (Eds.) *Handbook of Agricultural Economics, Volume 1b: Marketing, Distribution, and Consumers* (pp. 745-816). Amsterdam: Elsevier Science B.V.

Wachovia Securities, Weekly Cotton Comment, January 9, 2009.

Working, H., "The Theory of Price of Storage," American Economic Review, 39(6):1254-1262. 1949.

Figure 1. Basis: December cotton futures contract, 2000-08



Source: ERS calculations based on data from Thompson/Reuters Datastream and USDA Agricultural Marketing Service

Figure 2. Cost of supply model: basis as function of inventory

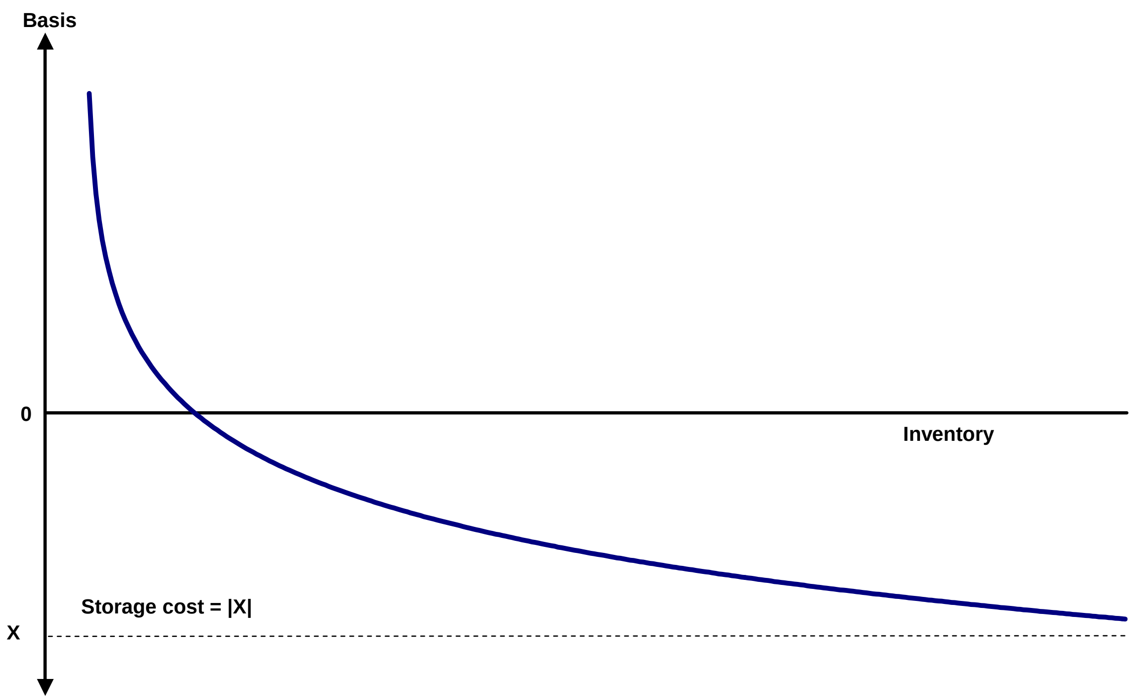
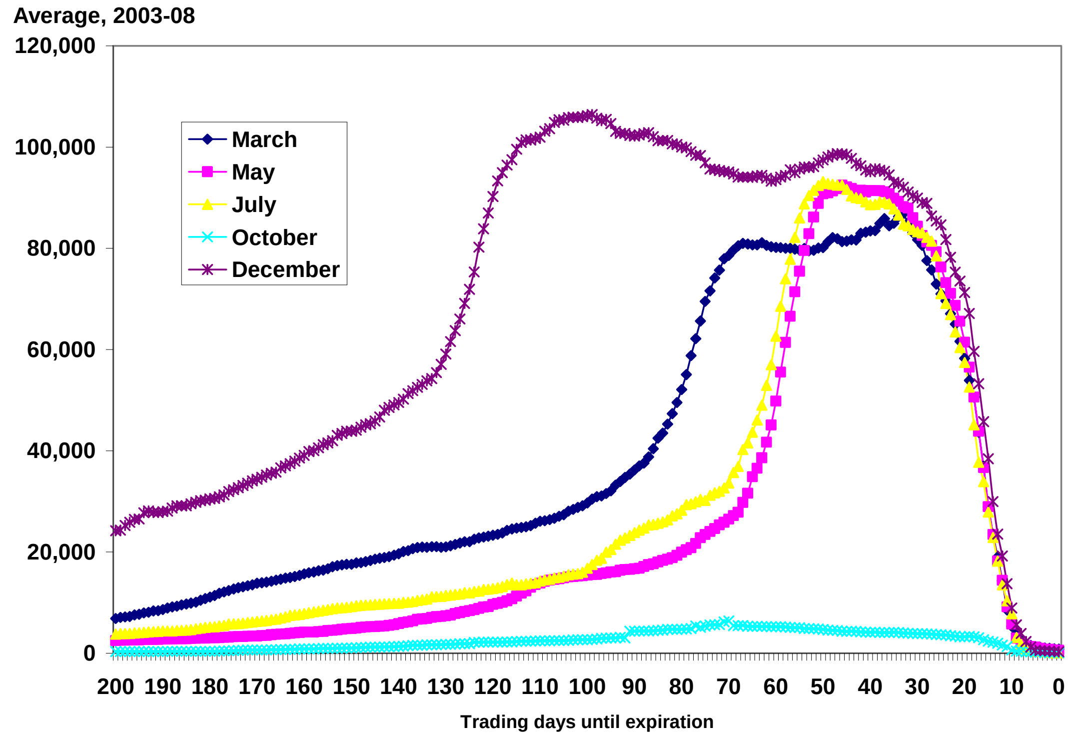
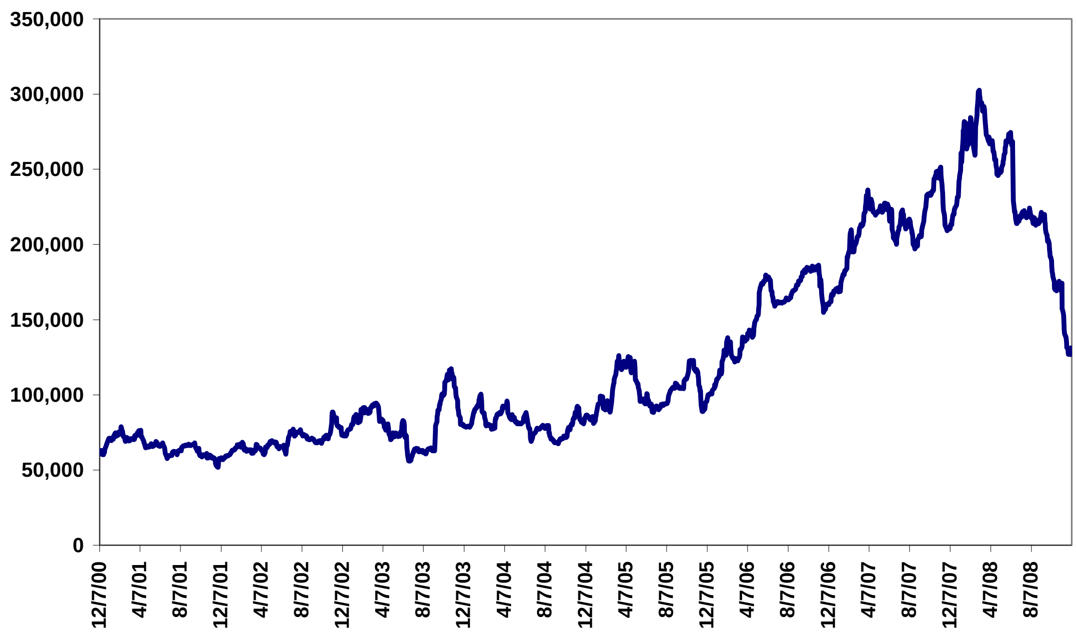


Figure 3. Open interest in ICE #2 cotton futures contracts



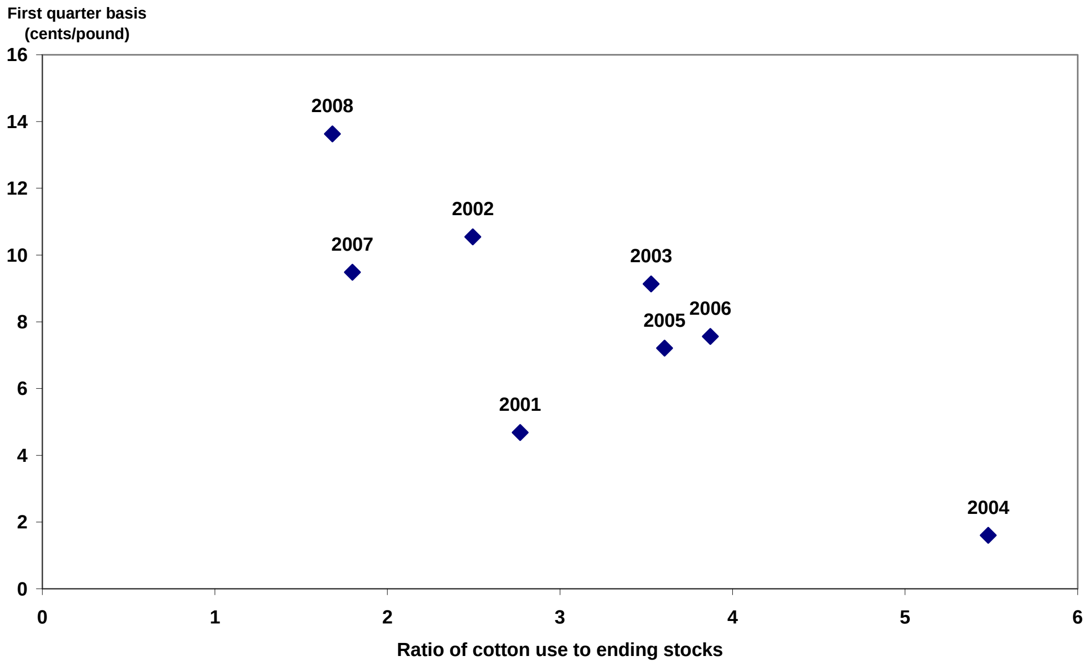
Source: Thompson/Reuters Datastream

Figure 4. Open Interest: all cotton futures contracts, 2000-08



Source: Thompson/Reuters Datastream

Figure 5. December cotton contract basis and cotton inventory



Sources: ERS calculations based on data from Thompson Reuters Datastream, U.S. Department of Agriculture Agricultural Marketing Service, and PS&D Online.

Table 1. Definition of futures markets terminology

Backwardization[1]:	When the price of a distant delivery month is below the price of a nearer delivery month, the spread between the prices is said to be in backwardization. A backwardization corresponds to an inverse carrying charge. In this study, backwardization is represented by $F_t < S_t$.
Basis[2]:	The difference between a cash price at a specific location and the price of a particular futures contract.
Contango[1]:	When the price of a distant delivery month is above the price of a nearer delivery month, the spread between the prices is said to be in contango. A contango corresponds to a positive carrying charge. In this study, contango is represented by $F_t > S_t$.
Convinience yield[3]:	The value of the convinience of having the possibility of making use of a commodity in inventory at the moment one wants.
Delivery[1]:	The changing of ownership or control of a commodity under the very specific terms and procedures established by the exchange upon which a futures contract is traded.
Long[1]:	Owning a futures contract or a volume of a commodity (opposite of short).
Nearby contract[1]:	The futures contract closest to maturity.
Open interest[1]:	The number of futures contracts that have not yet been offset by opposite futures transactions nor fulfulled by delivery of the commodity.
Opportunity cost:	The forgone returns from capital used to finance stockholding. In this study, opportunity cost is represented by $r_t S_t (T-t)$.
Short[1]:	Having sold a futures contract or a volume of the commodity (opposite of long).

[1] Williams (2001).
[2] Chicago Board of Trade (1985).
[3] Kaldor (1939).

Table 2. Stationarity testing: Fisher-Augmented Dickey-Fuller

Variable	Trend Included	Fisher Chi-square	Probabilty
Levels			
Basis	Yes	17.34	0.364
Basis	No	36.50	0.003
Opportunity cost	Yes	10.65	0.831
Opportunity cost	No	12.73	0.692
Open interest	Yes	12.25	0.727
Open interest	No	13.45	0.640
First differences			
Basis	Yes	926.13	0.000
Basis	No	933.80	0.000
Opportunity cost	Yes	996.71	0.000
Opportunity cost	No	937.92	0.000
Open interest	Yes	603.23	0.000
Open interest	No	613.37	0.000

Table 3. Cointegration testing of basis model variables (Johanson Fischer Panel Cointegration test)

Variables: Basis, opportunity cost

Number of cointegrating equations	Fischer statistic (trace test)	Probability	Fischer statistic (max-eigen test)	Probability
None	31.8	0.0106	28.51	0.0275
At most 1	21.5	0.1601	21.5	0.1601

Variables: Basis, opportunity cost, open interest

Number of cointegrating equations	Fischer statistic (trace test)	Probability	Fischer statistic (max-eigen test)	Probability
None	24.26	0.084	26.09	0.0528
At most 1	11.7	0.7646	9.818	0.876

Table 4. Exogeneity testing of basis model variables (Granger causality test)

			Null hypothesis: x does not cause y	
Variable x	Variable y	Lags	F-statistic	Probabilty
Basis	Opportunity cost	10	1.41634	0.1667
Opportunity cost	Basis	10	3.01692	0.0009
Basis	Open interest	10	1.60792	0.0984
Open interest	Basis	10	1.61174	0.0973
Open interest	Opportunity cost	10	2.95005	0.0011
Opportunity cost	Open interest	10	2.79253	0.002

Table 5. December cotton contract basis equiibrium model

Variable	Model 1	Model 2
	Coefficient estimate t-statistic	
Constant	7.272221	2.233785
	27.04	7.80
Opportunity cost	-2.78E-05	-5.25E-05
	-10.02	-21.19
Storage cost	0.036165	0.042621
	25.32	35.04
Market activity	NA	2.3E-05
	NA	28.42
Convenience yield		
First quarter	-2.889452	-2.394264
	40.42	38.44
Second quarter	-1.854559	-1.384096
	33.20	27.92
Third quarter	-1.608221	-1.006968
	29.73	-20.16
Fourth quarter	-1.448526	-0.842857
	21.15	13.79

Table 6. December cotton contract basis disequilibrium model

Variable	Coefficient estimate t-statistic	
	Model 1	Model 2
Constant	0.019208	0.019135
	3.67	3.66
Δ Opportunity cost	-5.26E-05	-5.22E-05
	5.21	5.19
Δ Market activity	NA	0.153818
	NA	0.73
Econometric error	-0.014623	-0.022946
	3.76	4.95