Commodity Price Comovement and Financial Speculation: The Case of Cotton

Joseph P. Janzen*, Aaron D. Smith, and Colin A. Carter

March 9, 2013

Abstract: Recent booms and busts in commodity prices have generated concerns that financial speculation drives prices away from the levels implied by supply and demand under rational expectations. If financial speculation has such effects, then it should be apparent in cotton where supplies were plentiful even as prices spiked. We estimate a structural vector autoregression model of the cotton futures market to explain two recent spikes in cotton prices and in doing so we make three contributions to the literature on commodity price dynamics. First, we use nonlinearities implied by the rational expectations competitive storage model to develop a new method to point identify shocks to precautionary demand for cotton separately from shocks to current supply and demand. Second, we separately identify the effects of two types of speculation: precautionary demand for the commodity and financial speculation. Third, we show empirically that most cotton price variation stems from contemporaneous unanticipated shocks to current cotton supply and demand. However, the 2008 price spike came from an increase in precautionary demand due to projections of lower future production. We find no evidence in support of claims that financial speculation causes commodity booms and busts.

Key words: commodity prices, index traders, speculation, cotton, comovement, structural VAR.

JEL codes: Q11, G13, C32

Joseph P. Janzen is a Ph.D candidate, Aaron D. Smith is an associate professor, and Colin A. Carter is a professor in the Department of Agricultural and Resource Economics, University of California, Davis. Smith and Carter are members of the Giannini Foundation of Agricultural Economics.

This paper is based on work supported by a Cooperative Agreement with the Economic Research Service, US Department of Agriculture under Project Number 58-3000-9-0076. The authors wish to thank seminar participants at ERS-USDA, the 2012 NCCC-134 Meeting, University of California, Davis, Montana State University, and Iowa State University for helpful comments on earlier versions of this work.

* Corresponding author: Department of Agricultural and Resource Economics, 1 Shields Ave, Davis, CA, 95616, USA, telephone: +1-530-848-1401, e-mail: janzen@primal.ucdavis.edu.
1 Introduction

Since 2006, the world has experienced two commodity price booms. Figure 1 shows that between 2006 and 2008, the price of cotton almost doubled and the price of wheat tripled. Price increases were not limited to agricultural commodities. Energy and metals prices also rose sharply. The price of crude oil nearly tripled, the price of copper doubled, and the price of silver increased by 50%. As they rose together, commodity prices fell together in late 2008, before commencing a second boom and bust in 2010 and 2011.

Turbulent commodity prices have significant economic and political implications, so it is important to determine the cause of these booms and busts. High and volatile prices hurt consumers, especially in countries where food constitutes a major share of household budgets or where energy is imported. When many commodity prices move simultaneously, opportunities for substitution are limited and many households are pushed into poverty. Price shocks have also been linked to subsequent political unrest (Bellemare, 2011). Policy proposals such as agricultural commodity price stabilization schemes (e.g. Von Braun and Torero, 2009) or regulatory controls on commodity futures trading (e.g. Masters, 2010) are based on particular assumptions about the cause of commodity price booms and busts. Such policies may be ineffective or even counterproductive if they incorrectly attribute the causes of observed price shocks. In this paper, we propose a modeling framework to identify the structural shocks that generate price spikes, and we apply our method to the cotton market.

Common movements among commodity prices could be caused simply by coincidental commodity-specific shocks to consumption and production or shocks to macroeconomic factors such as the global level of economic activity. In 2007-2008 and 2010-2011, adverse weather affected crop production in many regions around the world (Trostle et al., 2011). Global economic growth prior to 2008, particularly in developing countries such as China and India, led to strong demand for all commodities (Kilian, 2009). Similarly, aggregate demand for commodities collapsed during the economic downturn that followed the 2008 financial crisis (Carter et al., 2011).

Market observers such as Soros (2008) and Masters (2010) suggest that financialization of commodity markets, rather than supply-and-demand factors, is to blame for widespread commodity price spikes. This “increased acceptance of (commodity) derivatives as a financial asset” (Fattouh et al., 2012, p. 7) led to large amounts of speculative money entering commodity futures markets. According to critics, these speculative inflows caused the widespread boom and bust in commodity prices and broadened its effects across commodities such that some prices have been “taken for a ride” during this turbulent period.

If any commodity could have avoided the boom and bust of 2007 and 2008, it was cotton. The fundamental supply and demand situation was not bullish at the time; unlike other crops, large quantities of cotton were held

2
in storage. Figure 2 shows annual US and world cotton stocks-to-use ratios. In 2007-2008, the US stocks-to-use ratio was above 50%, levels higher than had prevailed since the mid-1980’s when US government policy encouraged higher cotton stocks. Nonetheless, cotton prices increased from $0.50 per pound in mid-2007 to almost $0.90 in March 2008 before dropping back to $0.50 in the ensuing six months. In contrast, cotton was scarce in 2010-2011; lower than average planted acreage combined with negative weather shocks in the US and Pakistan drastically reduced available supply. Prices increased from $0.75 per pound in mid-2010 to a peak of $2 in March 2011, but were back below $1 by the end of the year. Unlike 2007-2008, there appears to be a strong cotton-specific explanation for the boom and bust of 2010-2011.

These events make cotton a useful case study of commodity price spikes, the impact of speculation, and the potential influence of external markets on prices. Cotton is the world’s most important textile fiber, representing 40% of fiber production, and 30-40% of cotton fiber crosses borders before processing (Meyer et al., 2007). However, the cotton futures market is relatively small and thinly traded, so it may be more vulnerable to external speculative influences than other commodities. In 2011, approximately 5.3 million cotton futures contracts were traded on the Intercontinental Exchange (ICE), formerly the New York Board of Trade, which serves as the central global price discovery mechanism for cotton. This quantity represents 119 million metric tonnes of cotton with a notional value of $362 billion. By comparison, the 2011 volume of trade in the West Texas Intermediate crude oil futures market represented a notional value of $16.6 trillion (Commodity Research Bureau, 2011).

This paper makes three contributions to the applied econometrics literature on commodity price dynamics. First, we use structural vector autoregression (SVAR) to propose a novel method to identify shocks to precautionary demand for a commodity separately from current-period demand and supply. Our method uses nonlinearities implied by the canonical rational expectations competitive storage model (e.g. Gustafson, 1958; Williams and Wright, 1991) to obtain point identification. Specifically, we exploit the implication of the competitive storage model that prices are more volatile when inventories are low than when inventories are high. The lower-bound on inventories implies a kinked demand curve that is less elastic when stocks are tight, so that prices are more volatile when above the kink point and less volatile below it. This market characteristic allows us to apply the Identification through Heteroskedasticity technique developed by Rigobon (2003). In contrast, previous papers obtain only partial identification through sign restrictions and/or inequality constraints on short-run elasticities (Kilian and Murphy, 2012; Juvenal and Petrella, 2011; Lombardi and Van Robays, 2011; Carter et al., 2012).

Our second contribution is to separately identify two types of speculation: (i) precautionary demand for
cotton based on expectations about future supply and demand, and (ii) speculative effects from outside of the cotton market, such as those from the financialization of commodity markets. If the 2007-08 and 2010-11 cotton price booms were driven by financial speculation, then cotton prices would have been strongly correlated with other commodity prices during these periods. Tang and Xiong (2012) focus on this implication in their study of the effects of the recent surge in commodity index trading, but the finding of “excess comovement” among commodity prices predates the recent booms. Pindyck and Rotemberg (1990) estimate that commodity prices move together much more strongly than can be explained by fundamentals and attribute this phenomenon to financial flows driven by changing trader sentiment. Thus, to measure the possible effects of financial speculation on cotton prices, we estimate the extent to which cotton prices were driven by comovement with the price of an external commodity. We use crude oil for the external-market price in our main analysis, but we also report results using copper and silver. If we were to find that comovement is a major determinant of cotton prices, then this may be evidence of a financial-speculation effect or of a fundamental connection between the external commodity and cotton. However, if we find a small comovement effect, then it provides prima facie evidence that financial speculation was a minor factor.

Our third contribution is empirical. We find that comovement plays a minor, if any, role in cotton price determination. Rather, our model suggests that fundamental factors specific to the cotton market underlie most observed price movement. Cotton price spikes in 2008 and 2011 were caused by two different types of shocks to fundamental factors: precautionary or inventory demand led to higher prices in 2008 as cotton plantings were reduced in light of higher prices for other commodities, whereas supply shortfalls drove prices to record highs in 2011. In general, we find that cotton futures prices, while extremely volatile, have reflected supply-and-demand fundamentals rather than the machinations of financial speculators. Our results provide no support for claims that financial speculation causes commodity booms and busts.

2 Speculation and the Cotton Market

ICE cotton futures serve as a global benchmark in part because of the importance of US cotton production and exports in the global market for textiles. The US, China, India, and Pakistan grow three-quarters of the world’s cotton. The US is the third largest producer after China and India, accounting for 14% of the global total over the period from 2007-2011 (USDA-Foreign Agricultural Service, 2011). Because cotton processing has largely moved from the US to low-cost areas such as China, the majority of US production is exported and the US is the world’s leading exporter of raw cotton. After initial local processing separates cotton fiber from cottonseed, the US exports about 13 million bales of cotton fiber annually, or 36% of all global trade (USDA-Foreign
Agricultural Service, 2011).

The physical cotton underlying ICE futures contracts is US cotton deliverable at points throughout the southern and southwestern US, but the contract is traded by US and foreign cotton merchants, growers, and processors. The commercial traders who deal in both physical cotton and cotton derivatives trade with a diverse group of speculators that includes large financial firms and commodity index traders. In 2011, commodity index traders held 24% of long open interest in cotton futures contracts and index traders have held as much as 44% of long open interest (Commodity Futures Trading Commission, 2011). This large percentage suggests that cotton may also be vulnerable to spillover impacts from unrelated commodities, especially because it is a relatively small component of major commodity indexes that attract speculation. In 2011, 1.24% of the Standard & Poors - Goldman Sachs Commodity Index (S&P-GSCI) index was weighted to cotton, compared to 34.71% for WTI crude oil futures (and an additional 15.22% for Brent crude oil futures) (McGlone and Gunzberg, 2011).

The 2008 and 2011 price spikes had a serious impact on the cotton industry, which provides a further reason to understand them. In 2008, margin calls on futures positions forced several cotton merchants to exit the industry (Carter and Janzen, 2009). The Commodity Futures Trading Commission responded to the 2008 price spike with an inquiry into potential market manipulation; they found no evidence suggesting any such manipulation was present (Commodity Futures Trading Commission, 2010). Price swings in 2011 (which in absolute terms were much larger than in 2008) caused further large losses for physical cotton traders, prompting commentators to label cotton futures as the “widow maker trade” of the commodities world (Meyer and Blas, 2011).

3 Fundamental sources of price variability

To understand commodity price booms and busts, we aim to identify the source of observed price shocks. A fundamental theoretical approach models price determination as an equilibrium between supply and demand. In the case of storable commodities, such as cotton, equilibrium prices account for contemporaneous production and demand factors but also the incentive to hold current supplies in storage in anticipation of future supply and demand conditions. The canonical framework for understanding storable commodities pricing is the rational expectations competitive storage model, starting with Gustafson (1958). Williams and Wright (1991), Deaton and Laroque (1992, 1996), and Routledge et al. (2000) developed further refinements to these models.

The essential components of a rational expectations competitive storage model include a downward-sloping inverse demand function for use of the commodity at period $t$, denoted $P_t = f(D_t)$ where $P$ is the equilibrium price and $D$ the quantity demanded for current consumption. Exogenous production in each period, $S_t$,
is subject to some uncertainty over which risk-neutral stockholding firms form rational expectations. These stockholding firms maximize expected profit from storing the commodity in a competitive market. Inventories, $I_t$, are held between periods so that the quantity available at $t$, $(S_t + I_{t-1})$, equals consumption demand plus stocks carried into the next period, $(D_t + I_t)$. The final key feature of the competitive storage model is a zero lower bound on inventories, $I_t \geq 0$, as stocks cannot be borrowed from the future.

The competitive storage model implies a relationship between current and expected future prices. According to the model, stockholding firms will store the commodity according to the following no-arbitrage condition:

$$P_t = (1 + r_{t,t+1} + c_{t,t+1})^{-1} E_t(P_{t+1}) \quad if \quad I_t > 0$$

$$P_t \geq (1 + r_{t,t+1} + c_{t,t+1})^{-1} E_t(P_{t+1}) \quad if \quad I_t = 0,$$

where $r_{t,t+1}$ denotes the cost of capital and $c_{t,t+1}$ the cost of physical storage between period $t$ and a subsequent period $t+1$. This condition implies that inventories are positive when the discounted expected future price is greater than the current price and zero otherwise. The nonlinearity introduced by the zero-lower-bound on inventories requires the use of numerical methods to solve for equilibrium prices and quantities. Williams and Wright (1991) describe solution methods for such a model.

Three forces - consumption demand, inventory demand, and supply - underlie the equilibrium current price solution to the dynamic programming problem posed in the competitive storage model. Figure 3 presents these forces in a graphical representation originally developed by Eastham (1939). The solution to the stockholding firm’s dynamic programming problem is characterized by a negative relationship between inventory demand and price, so that the total demand curve in any period may be represented as the horizontal sum of the current-use demand function, $P_t = f(D_t)$, and this inventory demand relationship, $P_t = g(I_t)$. The zero lower bound on inventories implies that the total demand curve is kinked at the price/quantity point where $I_t$ equals zero. The supply curve in figure 3, $h(S_t + I_{t-1})$ represents the outcome of the current period production shock plus inventories carried in from the previous period.

From these basic elements of the competitive storage model comes information to identify the source of observed price changes. The presence of the kink point in the total demand curve, empirically confirmed by Cafiero et al. (2011) in their evaluation of the relevance of the competitive storage model, suggests that the price response to unexpected variation in current consumption, inventory demand, or production, will depend on the level of inventories. When carry-in stocks, $I_{t-1}$, are plentiful, the current period equilibrium will occur on the elastic portion of the total demand curve, as in figure 3a. The price response to shifts in consumption demand, inventory demand, or supply will be muted. In the contrasting case in figure 3b where stocks are scarce, shifts in
consumption demand and supply will result in large price changes. Shocks to inventory demand will not affect prices as the dearth of inventories breaks the link between current and future prices.

In the presence of a futures market, inventory demand may manifest itself as storage activity or speculative trading in futures contracts. Thus the no-arbitrage condition in equation (1), which describes the relationship between the expected future price, $E_t(P_{t+1})$, the current price, $P_t$, and inventories, also implies a functional relationship between the futures price, current spot prices, and inventory. In particular, competitive storage behavior implies that the calendar spread (i.e., the difference between futures and spot prices) is positively related to the inventory level. This relationship is often termed the “Working curve” after the original discovery of the relationship as an empirical regularity in Working (1933). In the context of a programming solution to the theoretical competitive storage model, Routledge et al. (2000, p. 1299) describe a similar relationship between inventories and the term structure of forward prices.

The Working curve relationship implies that information from the term structure may help infer the origin of observed price shocks when relevant quantity data regarding production, inventories, and consumption are unavailable. For example, price increases accompanied by rising inventories and an increasing calendar spread suggest an increase in inventory demand. In contrast, a temporary shock to current supply due to poor growing-season weather would raise prices, decrease the calendar spread and be associated with declining inventory. Comprehensive inventory data are typically not available at high frequencies. Thus, based upon the strictly non-decreasing nature of the Working curve, calendar spreads have been suggested as proxies for inventory (Fama and French, 1987; Ng and Pirrong, 1994; Geman and Ohana, 2009). Alquist and Kilian (2010) construct a formal model of the relationship between the spread and precautionary demand related to future supply and demand uncertainty. The distant price portion of the calendar spread in their model represents the price for delivery in some future period when supply and demand uncertainty can be resolved. Following this line of reasoning, we use a distant calendar spread in our empirical analysis to proxy for the incentive to hold inventory.

In summary, the competitive storage model implies that storable commodity prices are strongly heteroskedastic, characterized by volatile and tranquil periods conditional on inventory levels. The Working curve relationship between price spreads and inventories captures similar nonlinear response of spreads to changes in inventory, so that spreads also exhibit heteroskedasticity. Fama and French (1988) showed that when inventories are scarce, forward and nearby prices will exhibit different levels of volatility. Consequently, volatility of spreads will differ depending on relative inventory levels. We use these fundamental features in developing our empirical model in Section 5.
4 Types of Speculation

We distinguish two types of speculation. The competitive storage model describes a speculative component related to inventory demand. This component represents rational, economically-justified speculation that anticipates future supply and demand conditions. For this reason, we term this component “precautionary demand”. Importantly, speculation related to precautionary demand is a fundamental characteristic of a well-functioning storable commodity market. The second type, financial speculation, generally describes all speculative activity that is not generated by precautionary demand. It includes the excess comovement that Pindyck and Rotemberg (1990) attribute to trader sentiment, as well as the recent growth in trading by commodity index funds that provide exposure to commodities as part of a portfolio diversification strategy.

Financial Speculation and Commodity Index Trading

The recent debate about financial speculation and its effects on commodity markets has focused on commodity index trading (Irwin and Sanders, 2011). Commodity indexes are a weighted average of futures prices across a set of commodities and are designed to measure broad commodity price movement. Two industry benchmarks are the Standard and Poor’s-Goldman Sachs Commodity Index (S&P-GSCI) and the Dow Jones-UBS Commodity Index (DJ-UBSCI). Cotton is a component of both indexes. Financial firms have developed exchange-traded funds, swap contracts, and other vehicles to allow individual and institutional investors to track these and similar indexes in their investment portfolios (Stoll and Whaley, 2010). Most traders following index-trading strategies tend to take only long positions (i.e. positions that make money if prices rise); because they want exposure to commodity returns, they do not take short positions (i.e. positions that make money if prices drop.) Hereafter, we refer to firms following index-tracking trading strategies as commodity index traders, or CITs.

Critics such as Soros (2008) and Masters (2010) claim that CITs have caused boom and bust cycles in commodity markets. Unlike inventory holders in competitive storage markets, CITs do not have directional views on the prices of a specific commodity. Rather, they wish to gain exposure to the broad movement of commodity prices because of perceived portfolio diversification benefits. This lack of attention to fundamentals causes prices to be determined by investor flows. Consistent with this claim, Singleton (2011) finds that increased CIT positions in crude oil futures markets are associated with subsequent increases in prices. However, a considerable body of evidence supports the opposite conclusion, namely that CIT futures market positions are not associated with futures price levels or price changes (e.g. Stoll and Whaley, 2010; Buyuksahin and Harris, 2011; Irwin and Sanders, 2011; Fattouh et al., 2012). In addition, calculations by Irwin and Sanders (2012) indicate that Singleton (2011) significantly mismeasured CIT positions in the crude oil futures market due to his use of
a coarse imputation method.

*Identifying Financial Speculation Through Comovement*

Although the weight of evidence implies that CIT trading does not predict commodity prices, Tang and Xiong (2012) show how the effects of speculation can be revealed in the cross-section. The correlations between many commodity prices and the price of crude oil, the most widely traded commodity futures market, have risen over the period in which CIT activity has become prevalent. Tang and Xiong (2012) tested the link between returns for many commodities and crude oil and concluded that this comovement among prices is caused by the inclusion of commodities into major indexes such as the S&P-GSCI and the DJ-UBSCI. The “index inclusion” impact of CITs follows similar effects found in equity markets. According to Barberis et al. (2005), upon inclusion in the widely-followed S&P 500 index a stock’s price becomes more correlated with the index and less correlated with non-S&P 500 stocks. Barberis et al. (2005) use this result to argue that trader sentiment removed from fundamental factors specific to individual stocks is an important determinant of prices.

Finding common movement in commodity prices is not unique to the study by Tang and Xiong (2012) or the presence of CITs. Economists have documented unexplained comovement in prices at least since Pindyck and Rotemberg (1990) presented findings of “excess co-movement” among seven commodity prices. Pindyck and Rotemberg posited that correlation among fundamentally unrelated commodity prices that cannot be explained by macroeconomic factors is excess comovement. Their test for excess comovement relied upon the selection of commodities that were unrelated as substitutes or complements in production or consumption, not co-produced, and not used as major inputs in the production of others. To control for macroeconomic factors, they employed a seemingly unrelated regressions framework in which prices for cotton, wheat, copper, gold, crude oil, lumber, and cocoa were dependent variables. They used macroeconomic variables such as aggregate output, interest rates, and exchange rates as controls. Significant cross-equation correlation of the residuals from these regressions suggested evidence of excess comovement.

Pindyck and Rotemberg (1990, p. 1173) suggested that comovement arises because “traders are alternatively bullish or bearish on all commodities for no plausible reason.” Excess comovement as described by Pindyck and Rotemberg is essentially identical to descriptions of the effects of financial speculation in commodity futures markets: price changes due to trader sentiment rather than information about current and futures supply and demand conditions. For example, Masters (2010) argues that index trading creates massive, sustained buying pressure that causes prices to exceed fundamentally justified levels. Masters states: “Since the GSCI is an index of 24 commodities, it includes many commodities, such as most agriculture commodities, where there is no large
concentrated group of commodity-producers exerting selling pressure. Nonetheless, because Goldman created the GSCI, index speculators are exerting enormous buying pressure for these commodities in the absence of concentrated selling pressure. This has resulted in inflated food prices...”

Just as subsequent research countered the claims of Masters (2010) and Singleton (2011) with respect to CITs, commodity price comovement found in Pindyck and Rotemberg (1990) was contested as an artifact of methodological flaws or omitted variables. Deb et al. (1996) suggested that the excess comovement result was driven by the assumption of normal and homoskedastic errors in the seemingly unrelated regressions model of Pindyck and Rotemberg (1990). They proposed a GARCH framework to account for the heteroskedastic nature of commodity price changes and found minimal evidence of comovement for the commodities and time period used in Pindyck and Rotemberg (1990).

Ai et al. (2006) seized upon the poor explanatory power of the macroeconomic variables in the regressions of Pindyck and Rotemberg (1990) and Deb et al. (1996). They proposed that the omission of production, consumption, and inventory information led to observed cross-commodity price correlation. Their time-series model incorporating quarterly data on supply and inventories over the period 1997 to 2002 showed that this information explained comovement among prices of wheat, barley, oats, corn, and soybeans, all agricultural commodities.

This earlier comovement literature potentially explains why Tang and Xiong (2012) found a CIT-induced comovement effect: they did not consider market-specific fundamentals. However, Ai et al. (2006) cannot provide a direct refutation of Pindyck and Rotemberg (1990) and Tang and Xiong (2012) because they do not consider a comovement effect among commodities that are not direct substitutes in production and consumption. Additionally, they could not test for comovement effects resulting from the increased presence of CITs, which developed after their paper was written, and they could not consider price changes at a frequency greater than quarterly due to limited inventory data. Our econometric analysis overcomes these shortcomings.

5 Econometric Model

We propose a SVAR model of the global cotton market that jointly addresses financial speculation and commodity price formation as understood in the competitive storage model. That is, we consider an equilibrium determined by the interaction of current demand and supply, the precautionary demand for inventories based on expectations about future demand and supply conditions, and comovement with other commodities through which we capture the impacts of financial speculation. Using SVAR allows us to measure the contemporaneous contribution of these structural shocks to observed cotton prices throughout the period spanned by available
data. Most importantly, the structural shocks identified by the model represent the outcome of a useful counterfactual, namely the cumulative contribution of each factor holding all others fixed. Before presenting our model, we place it in context by discussing similar models in the literature.

*Previous literature on SVAR and commodity speculation*

SVAR methods have been used to identify similar structural shocks in the crude oil market. Kilian and Murphy (2012) use monthly data on crude oil prices, production, inventory levels, and an index of economic activity to estimate an SVAR that considers competitive-storage-model-related shocks. They call these shocks “flow demand”, “flow supply”, and “speculative demand” shocks. These are similar to the components of the competitive storage model discussed above. In the Kilian and Murphy (2012) model, the production variable accounts for variation related to current flow supply, the inventory variable addresses speculative (or precautionary) demand, and an economic activity index captures current or flow demand. By adding inventory data and the speculative demand shock, this paper extends earlier work (Kilian, 2009) that focused on current or flow shocks to oil prices.

The model in Kilian and Murphy (2012) does not extend directly to cotton for three reasons. First, their approach assumes that the flow demand shock is mainly related to commodity demand driven by the state of the global macroeconomy. In the case of cotton, demand-side forces such as tastes and preferences for cotton and substitute fibers in apparel may not coincide with changes in general economic activity. We use the same real economic activity index as Kilian and Murphy (2012), but label the associated structural shock a *real economic activity* shock. We bundle all other flow demand shocks with current supply shocks and call this hybrid a *current net supply* shock.

Second, the annual production cycle and sparse inventory data limit the ability of the Kilian and Murphy (2012) model to capture competitive-storage-model-type shocks for seasonal agricultural commodities. Pirrong (2008) notes that the competitive storage model is difficult to apply to seasonally produced agricultural commodities where production levels are observed only annually. An SVAR model using annual price and production data disregards information contained in daily, weekly, or monthly prices. Because we seek to understand rapid boom and bust price cycles such as the 2007-2008 cotton price spike, we do not directly estimate the functions contained in the competitive storage model, but use our understanding of the competitive storage model to inform our econometric identification. As in Kilian and Murphy (2012), we account for rational-expectations-based speculation. Following the discussion in Section 3, we use the calendar spread to capture the incentive to hold precautionary inventories in expectation of future supply and demand shocks. We call the
associated structural shock a *precautionary demand* shock.

Finally, the Kilian and Murphy (2012) model does not address financial speculation as a differentiated effect separate from speculation related to precautionary demand. Two subsequent papers (Lombardi and Van Robays, 2011; Juvenal and Petrella, 2011) adapt the Kilian and Murphy (2012) model to capture the effect of financial speculation on crude oil prices. Using similar variables but alternative sign restrictions to identify the structural shocks, they attribute a significant portion of crude oil price volatility to financial speculation. However, (Fattouh et al., 2012) find the sign restrictions required for identification in these papers not credible.

Following the discussion in Section 4, we identify a structural shock due to financial speculation using comovement of cotton prices with markets external to cotton. Our identification strategy relies on the fact that cotton is a small market relative to other commodities and the assumption that financial speculation effects can be approximated using the prices of commodities that are more widely held by financial speculators and more heavily weighted in major commodity indexes. If the implications of the financialization hypothesis are correct, then we should find that comovement with non-agricultural commodity prices has driven cotton price changes. According to Tang and Xiong (2012), cotton prices should have become increasingly driven by crude oil prices since 2004, so we include the price of West Texas Intermediate crude oil in our model. For robustness, we repeat our analysis using silver and copper to represent the external market. Because our identification of a financial speculation effect relies on the presence of common movements between commodity prices, we label this a *comovement* shock.

**Identification**

We include four variables in the model: (i) real economic activity, \( \text{rea} \), (ii) the real price in the external market, \( \text{ext} \), (iii) the spread between distant and nearby futures prices for cotton, \( \text{spr} \), and (iv) the real price of nearby cotton futures, \( \text{pct} \). Denoting by \( y_t \) the vector of variables, the model is

\[
A(L)y_t \equiv (A_0 - A_1 L - A_2 L^2 - \ldots - A_p L^p)y_t = Cx_t + u_t, \tag{2}
\]

where \( x_t \) denotes a vector of deterministic components that includes a constant, a linear trend, and seasonal dummy variables and the structural shocks \( u_t \) are white noise and uncorrelated with each other. We label the four structural shocks (i) real economic activity, (ii) comovement, (iii) precautionary demand, and (iv) current net supply.
An estimable reduced-form VAR is:

$$B(L)y_t = (I - B_1 L - B_2 L^2 - ... - B_p L^p)y_t = D x_t + \varepsilon_t.$$  \hspace{1cm} (3)

The reduced-form shocks, $\varepsilon_t$, are prediction errors and are a weighted sum of the structural shocks, where the matrix $A_0$ provides those weights, i.e., $\varepsilon_t = A_0^{-1} u_t$. Identifying the structural shocks requires making sufficient assumptions to enable consistent estimation of the unknown elements of $A_0$.

Given the small size of the cotton market relative to energy markets and global economy, we assume that cotton-market-specific supply and demand shocks are not transmitted to real economic activity or the external market price within a given month. Previous studies of commodity price dynamics (e.g. Kilian, 2009) have made similar assumptions about the precedence of real economic activity shocks. Giving the comovement shock precedence over the cotton-specific shocks is justified by the relative size of the cotton market and its relative weight in major commodity indexes and the cross-sectional CIT effect identified in Tang and Xiong (2012). This assumption implies that none of the unexpected variation in external market prices will be attributed to the cotton-specific structural shocks. If this assumption implies any bias in the relative size of the structural shocks measured by our model, it is an overstatement of the impact of the comovement shock on cotton prices.

With these assumptions, the reduced-form to structural shock relation is:

$$\begin{bmatrix}
1 & 0 & 0 & 0 \\
A_{21} & 1 & 0 & 0 \\
A_{31} & A_{32} & 1 & A_{34} \\
A_{41} & A_{42} & A_{43} & 1
\end{bmatrix}
\begin{bmatrix}
\varepsilon_{t}^{rea} \\
\varepsilon_{t}^{ext} \\
\varepsilon_{t}^{spr} \\
\varepsilon_{t}^{pct}
\end{bmatrix}
= \begin{bmatrix}
u_t^{REA} \\
u_t^{CM} \\
u_t^{PD} \\
u_t^{NS}
\end{bmatrix}.$$  \hspace{1cm} (4)

Further use of the recursion assumption by setting $A_{44} = 0$ would imply that the spread cannot respond contemporaneously to the net supply shock. The competitive storage model invalidates such an assumption. Switching the order of the last two shocks and imposing the recursion assumption (i.e., $A_{43} = 0$) would similarly be invalid. We expect $A_{34} < 0$ and $A_{43} > 0$, but sign restrictions alone provide only weak identification (Kilian, 2011). Rather than using sign restrictions, we exploit the nonlinearity implied by the competitive storage model to implement the Identification through Heteroskedasticity (ItH) approach developed by Rigobon (2003)).

ItH relies on differences in the variance of the structural shocks across time to identify the parameters of the $A_0$ matrix. It requires the sample to be partitioned into (at least) two volatility periods or regimes, where the variance of the structural shocks differs between regimes. In the case of two regimes, we refer to the high
variance regime as volatile and the low variance regime as tranquil. As we argue in Section 3, it is well-suited to identify structural shocks in the cotton market because the competitive storage model implies a kinked total demand curve for cotton. This kink creates two distinct volatility regimes. The same model also implies that inventory levels relative to the overall size of the market will dictate which regime applies at each point in time.

Following the presentation in Rigobon and Rodrik (2005), we partition the sample into two regimes, one volatile \( vl \) and one tranquil \( tr \), to satisfy the following properties:

\[
\begin{align*}
\var(u_t | t \in tr) &= \Sigma^{tr} \\
\var(u_t | t \in vl) &= \Sigma^{vl} \\
\Sigma^{tr} &\neq \Sigma^{vl}
\end{align*}
\]

The parameters in \( A(L) \) remain the same across regimes, which implies that the variance of the reduced form errors must also be heteroskedastic. Thus, we have two moment conditions:

\[
\begin{align*}
A_0 \Omega^{tr} A_0' &= \Sigma^{tr} \\
A_0 \Omega^{vl} A_0' &= \Sigma^{vl}
\end{align*}
\]

where \( \Omega^r \) denotes the variance-covariance matrix of the reduced-form residuals for regime \( r \in \{tr, vl\} \).

The two regimes are defined by the level of cotton inventory, which affects the volatility of shocks to prices and calendar spreads. Given the relatively small size of the cotton market, we assume that the external market and real economic activity shocks are constant across regimes. Thus, for each regime \( r \), we have:

\[
\begin{bmatrix}
1 & 0 & 0 & 0 \\
A_{21} & 1 & 0 & 0 \\
A_{31} & A_{32} & 1 & A_{34} \\
A_{41} & A_{42} & A_{43} & 1 \\
\end{bmatrix}
\begin{bmatrix}
\omega_{11} & \omega_{21} & \omega_{31} & \omega_{41} \\
\omega_{21} & \omega_{22} & \omega_{32} & \omega_{42} \\
\omega_{31} & \omega_{32} & \omega_{33} & \omega_{34} \\
\omega_{41} & \omega_{42} & \omega_{43} & \omega_{44} \\
\end{bmatrix}
= 
\begin{bmatrix}
1 & 0 & 0 & 0 \\
A_{21} & 1 & 0 & 0 \\
A_{31} & A_{32} & 1 & A_{34} \\
A_{41} & A_{42} & A_{43} & 1 \\
\end{bmatrix}
\begin{bmatrix}
\sigma_{REA} & 0 & 0 & 0 \\
0 & \sigma_{CM} & 0 & 0 \\
0 & 0 & \sigma_{PD} & 0 \\
0 & 0 & 0 & \sigma_{NS} \\
\end{bmatrix}
\]

Restricting \( \sigma_{REA} \) and \( \sigma_{CM} \) to be constant across regimes gives us thirteen parameters to be identified, seven
in the $A_0$ matrix and six structural shock variances. There are 13 free parameters in $\Omega^{tr}$ and $\Omega^{vl}$, so the model is just identified as long as the relative variance of the structural shocks ($\sigma_{PD}^r/\sigma_{NS}^r$) differs across regimes (Rigobon, 2003).

The condition that ($\sigma_{PD}^r/\sigma_{NS}^r$) differs across regimes provides intuition for how ItH works. A scatter plot of the reduced-form shocks to the calendar-spread against the reduced-form shocks to the cotton price makes a cloud of points but does not trace out either a precautionary demand or a current net supply curve. However if, say, the net supply shock is relatively more volatile in the volatile regime, then the cloud of points stretches relatively more along the precautionary demand curve, which enables its slope to be identified. Rigobon (2003) likens ItH to “probabilistic instrument” in the sense that, we cannot be certain that we are identifying shifts in a particular curve for any particular observation, but under one regime, we are more likely to observe shifts in that curve.

6 Data

We use monthly data from January 1968 to December 2011. Why do we construct such a long time series when our period of ultimate interest is only the last four years? As in most econometric work, more data allows for more precise estimation of model parameters. More importantly, we are concerned about periods of dramatic price volatility. Our time series contains arguably four periods of general boom and bust in commodity prices centered around 1973, 1996, 2008, and 2011. Using a shorter time series removes these periods, limiting our ability to observe price response during volatile periods and to identify our econometric model using ItH.

The four variables in our model are real economic activity, an external commodity market price, the calendar spread in the cotton futures market, and the price of cotton. All variables are expressed in real terms, using the US Consumer Price Index (CPI) as a deflator. To measure real economic activity, we use the real economic activity index developed by Kilian (2009). This index employs ocean freight rates as a proxy for global demand for goods and is measured in deviations from long term trend. Kilian notes an empirically documented correlation between freight rates and economic activity. Because the freight rate index will rise if economic activity rises in any part of the world, it will not be biased toward any one country or region of the world. Given the importance of demand growth in emerging economies to stimulate cotton consumption, our aggregate commodity demand measure must be a global measure.

We collect data on cotton and other commodity market prices from Commodity Research Bureau (2011). The cotton price series is the logarithm of the monthly average nearby futures price, deflated by CPI. The crude oil price series is the cash price for West Texas Intermediate crude oil. Again this series is the logarithm of
the real monthly average price. The cash price series is used because crude oil futures only began trading in 1983. To assess the robustness of our results to the use of crude oil as the external market, we use two additional series: nearby copper and silver futures prices. These prices represent external markets for industrial and precious metals, just as crude oil is representative of the price of energy commodities.

To proxy for the incentive to hold inventory, we use the calendar spread for cotton futures prices over a one-year time span. This measure is always spans a harvest, so it represents expectations about the scarcity of cotton now compared to a future period after production has responded. As such, it measures the incentive to hold inventories until future supply and demand uncertainty is resolved by a future harvest. Cotton futures contracts reach delivery five times per year in March, May, July, October, and December, so we use the log difference between the sixth most distant and the nearby futures contract. The time difference between these contracts is always a year, so the resulting spread represents the term structure of cotton prices over a constant period of time.

Data on physical inventories would be an alternative to the spread variable in our model if they existed. Some measures of stocks are available for cotton. However, these measures are either unavailable at the frequency required by our model or cover only a limited set of locations such as the warehouses licensed to receive certified stocks delivered against the ICE Futures cotton contract. Available data do not provide the global measure of cotton inventory scarcity that we require.

Figure 4 plots each variable in $y_t$, after accounting for linear trend and seasonality. We use a simple linear trend to account for productivity improvements and other factors that account for long term declines in real commodity prices. For the cotton price and spread variables, we also adjust for the impact of the 1985 Farm Bill. In early 1986, the US Department of Agriculture announced details of the cotton-specific provisions of the 1985 Farm Bill. These provisions created incentives to store cotton in the 1985-86 crop year and sell it during the 1986-87 crop year, affecting cotton term spreads and prices (Anderson and Paggi, 1987). Calendar spreads rose in early 1986 as inventories reached record levels; when the Farm Bill programs took effect in summer of 1986, spreads and prices plummeted as old-crop cotton flooded the market. We include four control variables to account for the effects of the 1985 Farm Bill: an indicator variable for the first six months of 1986, an indicator variable for the last six months of 1986, and interactions between each of these variables and a linear trend.

7 Estimation and results

We estimate a reduced-form VAR for the four variables constructed above using ordinary least squares with two lags. We select lag length using Akaike and Schwarz-Bayesian information criteria. We include a linear trend,
monthly indicator variables, and 1985 Farm Bill controls as exogenous variables in this VAR specification. The parameter estimates and their standard errors are presented in table 1. From the VAR estimates, we extract the reduced-form residuals, \( \hat{\epsilon}_t \). We divide these residuals into two sets corresponding to observations from the tranquil and volatile regimes. We select dates for each regime using a rule based on predictions the competitive storage model of commodity markets, namely that prices will be more volatile when stocks are low relative to use. We declare as volatile any crop year where projected cotton ending-stocks-to-use ratios (as defined by USDA World Agricultural Outlook Board forecasts) were below 0.25 for at least three months. Forecast stocks-to-use are a measure of inventory scarcity pertinent to contemporaneous commodity pricing at a given moment rather than in hindsight. Selecting regimes in this manner creates seven volatile windows in our sample, including the 1973-74, 1974-75, 1979-80, 1990-91, 1994-95, 1995-96, 1998-99, 2003-04, and 2010-2011 crop years\(^1\).

From the reduced-form residuals for each regime, we calculate the variance-covariance matrices, \( \Omega^{tr} \) and \( \Omega^{vl} \), as shown in table 2. Because we restrict the parameters in the first two rows or columns of \( A_0 \) to be constant across volatility regimes, we set the terms in the first two rows of \( \Omega^{tr} \) and \( \Omega^{vl} \) equal to the corresponding terms from the variance-covariance matrix calculated using the full sample. We define a set of constraints on the \( A_0 \) matrix, namely the zero terms in the first two rows. Using a constrained optimization routine subject to these constraints, we solve for the parameters in \( A_0, \Sigma^{tr}, \) and \( \Sigma^{vl} \) by minimizing a distance function equal to the sum of the squared differences \( A_0 \Omega^r A_0' - \Sigma^r \) from each regime \( r \). We report estimates of the parameters in \( \Sigma^r \) and \( A_0 \) in table 2. The variance of the reduced-form residuals for the volatile regime from the \( spr \) and \( pct \) equations, \( \omega^{vl}_{33} \) and \( \omega^{vl}_{44} \), are more than double those from the tranquil regime. The associated structural shock variances, \( \sigma^r_{PD} \) and \( \sigma^r_{NS} \) also vary across regimes. The net supply shock has variance 2.5 times greater in the volatile periods than in the tranquil periods, whereas the precautionary demand shock variance is only 25% larger in the volatile period. This result is as expected based on figure 3, and such a difference in relative variances is a key condition for identification.

Having solved for the model parameters, we calculate a set of orthogonal structural shocks for each period. We estimate impulse response functions for all model variables with respect to each structural shock and generate confidence bounds for the impulse responses using the wild bootstrap procedure of Goncalves and Kilian (2004). Since the innovation in any of the variables in the model in each period can be represented as a weighted sum of the structural shocks from that period, we can create time-series representations of the historical contribution of each structural factor to the observed innovations in each variable. We discuss our results for these

\(^1\)We also tested similar rules for setting the regime windows and found that our results were robust to other specifications.
impulse responses and historical decompositions below, first for the case where crude oil represents the external market and then for alternative external markets.

**Impulse response functions**

Figure 5 plots the time path for the response of each variable in our model to the economic activity, comovement, precautionary demand, and net supply shocks for the case where crude oil is the external market. The dashed lines represent the pointwise 90% confidence interval about the average response generated using 1000 bootstrap replications. These graphs demonstrate, based on the average response observed in the data, how each variable in the model would respond to a hypothetical one standard deviation structural shock using the observed standard deviation of the structural shocks across the entire sample. The normalization used to identify the model implies that each of the shocks causes an increase in the price of cotton. In particular, net supply shocks refer to a disruption that increases cotton prices.

The impulse response functions serve two purposes. First, they act as a check on the validity of our assumptions about the shocks we want to identify. The direction of the responses should be consistent with the theory that motivated our identification scheme. Second, the impulse response functions for the price of cotton can be compared to ascertain the magnitude and duration of the influence of each structural shock.

Focusing on the bottom-right corner of figure 5, we can check the validity of the identification scheme used to identify the endogenous net supply and precautionary demand shocks in the cotton market. If a precautionary demand shock occurs, the price and the spread should increase to encourage stockholding by providing higher returns to storage. Our results show this is the case. Similarly, a net supply shock (equivalent to a supply disruption) raises cotton prices and has a negative influence on the spread in order to draw supplies in storage to the market. The precautionary demand shock displays some evidence of overshooting: prices increase quickly in the months following the shock before declining. Both the precautionary demand and net supply shocks have an impact on prices for at least a year, which is to be expected based on the annual harvest cycle.

Figure 5 shows that external forces have a small impact on cotton prices, relative to the precautionary demand and net supply shocks specific to the cotton market. Real economic activity and external market shocks are small but long lived. However, neither of these effects are statistically significant on average. Even though our model allows these forces to take precedence over cotton-specific factors, we find no evidence in the impulse response functions to corroborate the hypothesis that external markets are driving cotton prices. The insignificance of crude oil market shocks suggests that broad-based commodity market speculation has not impacted cotton prices.
Historical decomposition of cotton price shocks

The impulse response analysis only allows us to assess the average response of cotton prices to the structural shocks. Historical decompositions in figure 6 allow us to assess the structural origins of variation in any of the variables in our model. In our case, we are most interested in the effect of structural shocks on observed cotton prices. The series in figure 6 are constructed so that the sum of the four series equals the realized price net of trend and seasonality in any month.

Most of the variation in cotton prices is due to the two cotton-market-specific shocks. Longer, smaller swings in price are attributed to the real economic activity and external market shocks. This finding suggests that these factors can contribute to periods of high and volatile prices, but they are likely to be a small component. For example, the real economic activity component increases during the period from 2000 to 2008, likely tracking commodity demand growth from emerging markets such as China, however the effect is small relative to the precautionary demand and net supply disruptions that occur over the same period. Crude oil price shocks have negligible effect across the entire time period, suggesting that external market comovement-type impacts driven by commodity market financialization are minimal and have not changed significantly since 2004.

The results of our decomposition analysis differ from analyses of crude oil prices by Kilian (2009) and Kilian and Murphy (2012) that used similar methods. These studies found that fluctuations in real economic activity related to the macroeconomic business cycle were the largest and most persistent driver of crude oil prices, particularly during period of rising prices that ended in 2008. Similarly, Carter et al. (2012) find that between its 2003 low and 2008 high, real economic activity generated an increase in corn prices of up to 50%. Our results for cotton suggest that real economic activity does not similarly impact the cotton market. Over the same period, real economic activity raised cotton prices by 5-10%\(^2\). This is not zero, but it is small relative to the net-supply generated portion of the price spike.

The net supply shock is the largest and most variable component of observed cotton futures prices over this period. It is the major driver of cotton price spikes in 1973-74, 1990-91, 1995-96 and especially the most recent spike in 2010-11. These are all periods of major supply disruptions. Each of these major positive net supply shocks is associated with lower US and world cotton production. The 2007-08 price spike is not associated with major changes in the net supply shock, a point we return to in the next section.

\(^2\)We measure this change as the log difference between prices with and without each of the shocks. These log differences approximate a percentage change, so we use percent to refer to these log differences.
Counterfactual analysis of the 2008 and 2011 price spikes

Because of the orthogonality of the shocks, we can eliminate individual shocks from our historical decomposition and use the sum of the remaining shocks to construct the price series for a counterfactual scenario: what would have happened to cotton prices in the absence of any one of the effects we identify? For example, how would the time series of observed cotton prices have differed if the external market shocks did not affect cotton prices? We use such counterfactual analysis to consider the two most recent price spikes in cotton that occurred in 2007-08 and 2010-11. The counterfactuals are plotted in figure 7. The five series shown are the observed cotton price and counterfactual cotton prices with each of the real economic activity, external market, precautionary demand, and net supply shocks set to zero for 2006 to 2011.

During the 2006-2011 period, comovement shocks are nearly imperceptible. We find that comovement shocks related to crude oil prices raised cotton prices only 2% at the peak of their impact. In absolute terms, the maximum cotton price impact of comovement shocks at any point in our analysis is less than $0.02 per pound. Relative to the price volatility observed over this period such an effect is minuscule; cotton futures prices rose nearly $0.50 per pound in 2007-08 and by more than $1.20 per pound in 2010-11.

Cotton price spikes in 2007-2008 and 2010-2011 had very different origins; elevated cotton prices in 2010-2011 were not a repeat of the events of 2007-2008. Figure 7 illustrates. The 2008 price spike would have been non-existent without shocks to precautionary demand. In March 2008 at the peak of the initial spike, we estimate that prices would have been 26% lower without precautionary demand shocks and 11% higher without net supply shocks. In contrast, in March 2011 the log-difference between observed prices and counterfactual prices with the net supply shock was 1.07, corresponding to an approximately 70% difference. In the same month, cotton prices would have been just 12% lower without precautionary demand shocks.

Market intelligence produced in early 2008 corroborates the precautionary demand explanation for the 2007-2008 spike. Early US Department of Agriculture projections for the 2008-2009 cotton crop year called for “sharply lower production and ending stocks” in the US. As projected plantings of other field crops were expected to increase, cotton planted acres were to decline by 25%. Projections also called for “strong but decelerating growth in consumption” and increased exports to China in the coming crop year (USDA-World Agricultural Outlook Board, 2008). These expectations are consistent with price increases in early 2008 due to precautionary demand. Later US Department of Agriculture forecasts were considerably less bullish, consistent with falling prices later in 2008.

In contrast, precautionary demand shocks had very little to do with rising prices in 2010-2011. Evidence from this period suggests a series of shocks to current net supply were behind the large price increase. Global
production for the 2009-2010 crop year was approximately 13% below the previous five-year average and consumption rebounded from declines following the 2008 global financial crisis. The world ending-stocks-to-use ratio fell from 56% in 2009 to 39% in 2010 (USDA-Foreign Agricultural Service, 2011). Cotton prices were extremely vulnerable to further shocks. Unexpected events, highlighted by floods in Pakistan and periodic export bans in India, limited available supplies during the 2010-2011 crop year that encompassed most of the price spike (Meyer and Blas, 2011). Reports suggest that the market suddenly became aware of “shortages in (current) mill inventories” in late 2010 (USDA-World Agricultural Outlook Board, 2010) that triggered steep price increases.

Assessing Robustness

We consider two robustness checks of our findings with respect to the influence of financial speculation on cotton prices. First, we replace the crude oil price variable in our SVAR model with prices representative of other commodity classes such as industrial and precious metals and consider the magnitude of the structural shocks due to financial speculation. Second, we test stability of the VAR parameters over time to account for the possibility that the rise in financial speculation and commodity index trading caused a break in the relationship between crude oil and cotton prices.

Figures 8-9 display similar counterfactual analyses as those shown in figure 7 for crude oil, but with silver and copper used as the external market price variable. The pattern of cotton price movement from 2006 to 2011 in the absence of comovement shocks is largely the same when silver and copper represent external market movements. In the absence of speculative comovement shocks, the cotton price would have been 7% lower in March 2008 and 18% lower in March 2011 according to the model that uses silver as the external market price; it would have been 13% lower in March 2008 and 24% lower in March 2011 according to the model that uses copper as the external market price. However, these models also predict that, in the absence of comovement shocks, prices would have initiated their run ups from similarly lower values. Thus, although comovement shocks increased somewhat absolute magnitude of the peak price, they did not change the magnitude of the price increases during the booms.

When using silver and copper as external markets, the implied comovement effects on cotton prices are somewhat larger than when we use crude oil, but they remain much smaller than the effects of cotton-specific precautionary demand and net supply shocks. One potential explanation for the larger copper effect is the common influence of Chinese demand on cotton and copper prices. Variation that we have labeled an external market shock may be picking up the influence of economic growth in China, rather than speculative comove-
ment. Given that index trading concentrates speculative capital most heavily in energy rather than precious and industrial metals, it seems unlikely that financialization-driven comovement accounts for all of the external market effect observed in figures 8-9.

Parameter instability may affect our model’s capacity to capture financial speculation impacts. Increasing activity of financial speculators may have caused a break in the relationship between cotton prices and prices for other commodities. To address the issue of model stability, we test for a structural break in the coefficients that determine the impact of external commodity prices on the cotton market using January 2004 as our break point. Irwin and Sanders (2011, p. 6) document, beginning in 2004, a steady increase in the value of assets invested in commodity index products. Using publicly unavailable data on open interest held by CITs prior to 2006, they also document a significant increase between 2004 and 2006 in the number of agricultural futures contracts and percentage on long open interest held by CITs. Tang and Xiong (2012) also select 2004 as a break point and found a change in the relationship between crude oil and other commodity prices post-2004.

Our model stability test must address the potential for structural breaks in parameters that determine the impact of lagged and current values of external market prices, \( ext \), that represent the effect of financial speculation. We cannot only test whether the coefficients on lagged \( ext \) in our reduced-form VAR model are constant across time; we must also test the stability of the parameters \( A_{32} \) and \( A_{42} \) that determine the current response of cotton spread and price to the external market shocks.

To conduct this test, we reestimate the equations in the reduced-form VAR equations for the cotton price and the calendar spread, including contemporaneous values of real economic activity and crude oil price as regressors. This specification is implied by the exclusion restrictions in the first two rows of the \( A_0 \) matrix. We include interactions between each of the crude oil price variables and an indicator variable equal to one in the period post-January 2004, \( D_{\geq 2004} \). Table 3 presents the coefficient estimates for the external market variables and the post-January 2004 interaction terms for each equation in the two-variable VAR. None of interaction terms are individually significantly different from zero according to simple t-tests. We conduct a Wald test of the joint significance of the interaction term coefficients and the post-January 2004 indicator variable from both equations. We fail to reject the joint null hypothesis that none of the external market coefficients vary across periods at the five percent significance level.

8 Conclusions

We use a structural vector autoregression model to attribute observed price changes in cotton to four factors: real economic activity, the precautionary demand for inventories, shocks to current net supply, and comovement
that reveals the effects of speculative trading. The comovement-driven portion of cotton prices represents a component that is unrelated to expected future prices and attributable to financial speculators and commodity index traders who some blame for recent periods of elevated prices in agricultural markets.

Our model is applied to the cotton futures market. Cotton is a useful case study; it is an important commodity both in the US and in major developing countries like China and India, cotton prices have recently experienced booms and busts in spite of some contrary fundamental factors, and commodity index traders, linked by other researchers to financial speculation, account for a significant volume of futures trading.

We find no evidence that financial speculators are the cause of cotton price spikes. The portion of observed price changes due to comovement has been small. Factors specific to the cotton market are the major determinants of cotton prices. Our model even allows comovement shocks to take precedence over cotton-specific shocks and we still find minimal comovement effects.

Most cotton price spikes are fundamentally driven and strongly associated with shocks to current supply. The 2008 price spike was an exception. We find that precautionary demand, likely induced by projections of lower acreage and steady demand, drove prices higher in 2008. Market intelligence from 2008 corroborates this explanation. Unlike studies of other commodity markets such as Kilian and Murphy (2012), we find that broad trends in global commodity demand related to real economic activity matter less than cotton-specific supply-and-demand.

Our results suggest that complaints about the accuracy of cotton futures price discovery based on the belief that index traders and other financial speculators are distorting prices are unfounded. Accordingly, legislative and regulatory efforts to restrict the trading activities of these traders will not prevent future price spikes. Price spikes are characteristic of storable commodity markets when inventories are unavailable to mitigate the effect of shocks.
References


Masters MW. 2010. Testimony before the u.s. commodity futures trading commission. Submitted testimony.


Figure 1: Prices for selected agricultural, energy, and metals commodities, monthly average nearby US futures prices, 2005-2011, January 2005 = 100.

Source: Commodity Research Bureau
Figure 2: US and world crop-year ending stocks-to-use, 1968-69 to 2011-12
Source: USDA PS&D Online
(a) Equilibrium with plentiful stocks

(b) Equilibrium with scarce stocks

Figure 3: Price determination in a stylized single-period representation of the competitive storage model
Figure 4: Data plot for variables in $y_t$, January 1968 to December 2011
Figure 5: Impulse response functions with crude oil as external market
Figure 6: Historical decomposition of structural shocks with crude oil as external market
Figure 7: Counterfactual analysis for SVAR with crude oil as external market
Figure 8: Counterfactual analysis for SVAR with silver as external market
Figure 9: Counterfactual analysis for SVAR with copper as external market.
### Table 1: Reduced-form VAR Regression Results

<table>
<thead>
<tr>
<th>Equation</th>
<th>$rea_t$</th>
<th>$ext_t$</th>
<th>$spr_{t-1}$</th>
<th>$pct_{t-1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.0789</td>
<td>-0.1776</td>
<td>0.0863</td>
<td>0.0615</td>
</tr>
<tr>
<td></td>
<td>(0.1032)</td>
<td>(0.1102)</td>
<td>(0.0563)</td>
<td>(0.0848)</td>
</tr>
<tr>
<td>$rea_{t-1}$</td>
<td>1.1442**</td>
<td>0.1308*</td>
<td>0.0491</td>
<td>-0.0186</td>
</tr>
<tr>
<td></td>
<td>(0.0715)</td>
<td>(0.0545)</td>
<td>(0.0284)</td>
<td>(0.0487)</td>
</tr>
<tr>
<td>$ext_{t-1}$</td>
<td>0.0736</td>
<td>1.1520**</td>
<td>-0.0061</td>
<td>0.0245</td>
</tr>
<tr>
<td></td>
<td>(0.0465)</td>
<td>(0.0677)</td>
<td>(0.0364)</td>
<td>(0.0583)</td>
</tr>
<tr>
<td>$spr_{t-1}$</td>
<td>0.3070**</td>
<td>0.5559**</td>
<td>0.4828**</td>
<td>0.7348**</td>
</tr>
<tr>
<td></td>
<td>(0.1080)</td>
<td>(0.1510)</td>
<td>(0.1683)</td>
<td>(0.2098)</td>
</tr>
<tr>
<td>$pct_{t-1}$</td>
<td>0.2764**</td>
<td>0.3513**</td>
<td>-0.2865**</td>
<td>1.5265**</td>
</tr>
<tr>
<td></td>
<td>(0.0847)</td>
<td>(0.1030)</td>
<td>(0.0728)</td>
<td>(0.1042)</td>
</tr>
<tr>
<td>$rea_{t-2}$</td>
<td>-0.1935**</td>
<td>-0.1033*</td>
<td>-0.0428</td>
<td>0.0163</td>
</tr>
<tr>
<td></td>
<td>(0.0722)</td>
<td>(0.0514)</td>
<td>(0.0284)</td>
<td>(0.0489)</td>
</tr>
<tr>
<td>$ext_{t-2}$</td>
<td>-0.0715</td>
<td>-0.1820**</td>
<td>0.0157</td>
<td>-0.0309</td>
</tr>
<tr>
<td></td>
<td>(0.0464)</td>
<td>(0.0662)</td>
<td>(0.0372)</td>
<td>(0.0591)</td>
</tr>
<tr>
<td>$spr_{t-2}$</td>
<td>-0.3444**</td>
<td>-0.4856**</td>
<td>0.3655*</td>
<td>-0.6210**</td>
</tr>
<tr>
<td></td>
<td>(0.1132)</td>
<td>(0.1507)</td>
<td>(0.1619)</td>
<td>(0.2067)</td>
</tr>
<tr>
<td>$pct_{t-2}$</td>
<td>-0.2983**</td>
<td>-0.3006**</td>
<td>0.2591**</td>
<td>-0.5296**</td>
</tr>
<tr>
<td></td>
<td>(0.0856)</td>
<td>(0.0924)</td>
<td>(0.0703)</td>
<td>(0.1026)</td>
</tr>
</tbody>
</table>

Note: Heteroskedasticity-robust standard errors are in parentheses. * and ** denote significance at the 5% and 1% levels. Coefficient estimates for seasonal indicators and time trend are not reported.
Table 2: Structural vector autoregression model parameter estimates

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Tranquil</th>
<th>Volatile</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_0$</td>
<td>1 0 0 0</td>
<td>-0.0807 1 0 0</td>
</tr>
<tr>
<td></td>
<td>-0.0347 -0.0215 1 0.6710</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.0996 -0.0796 -1.5662 1</td>
<td></td>
</tr>
<tr>
<td>$\Omega_r$</td>
<td>0.4258 0.0344 -0.0072 0.0339</td>
<td>0.4258 0.0344 -0.0072 0.0339</td>
</tr>
<tr>
<td></td>
<td>0.5191 -0.0086 0.1484 0.3371</td>
<td>0.5191 -0.0086 0.1484 0.3371</td>
</tr>
<tr>
<td>$\sigma_{RE}$</td>
<td>0.4258</td>
<td>0.4258</td>
</tr>
<tr>
<td>$\sigma_{CM}$</td>
<td>0.5163</td>
<td>0.5163</td>
</tr>
<tr>
<td>$\sigma_{PD}$</td>
<td>0.0618</td>
<td>0.0785</td>
</tr>
<tr>
<td>$\sigma_{NS}$</td>
<td>1.2475</td>
<td>3.0458</td>
</tr>
</tbody>
</table>

Note: For ease of presentation, we report the estimated variance and covariance parameter values multiplied by 100.
Table 3: Coefficient estimates for external market variables in parameter stability regressions

<table>
<thead>
<tr>
<th>Variable</th>
<th>spr</th>
<th>pct</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Intercept</strong></td>
<td>0.1069</td>
<td>0.0366</td>
</tr>
<tr>
<td><strong>ext_t</strong></td>
<td>-0.0129</td>
<td>0.0486</td>
</tr>
<tr>
<td><strong>ext_{t-1}</strong></td>
<td>0.0087</td>
<td>-0.0369</td>
</tr>
<tr>
<td><strong>ext_{t-2}</strong></td>
<td>0.0160</td>
<td>-0.0196</td>
</tr>
<tr>
<td><strong>D_{≥2004}</strong></td>
<td>0.0223</td>
<td>-0.0754</td>
</tr>
<tr>
<td><strong>D_{≥2004} × ext_t</strong></td>
<td>-0.0119</td>
<td>0.0424</td>
</tr>
<tr>
<td><strong>D_{≥2004} × ext_{t-1}</strong></td>
<td>0.0192</td>
<td>-0.0427</td>
</tr>
<tr>
<td><strong>D_{≥2004} × ext_{t-2}</strong></td>
<td>-0.0158</td>
<td>0.0239</td>
</tr>
</tbody>
</table>

Wald test statistic $\chi^2(8) = 4.21$

P-value = 0.8379

Note: Standard errors are in parentheses. * and ** denote significance at the 5% and 1% levels. Coefficient estimates for variables unrelated to parameter stability tests are not reported.