

## **Commodity Price Comovement: The Case of Cotton**

by

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

Suggested citation format:

Janzen, J. P., A. D. Smith, and C. A. Carter. 2012. "Commodity Price Comovement: The Case of Cotton." Proceedings of the NCCC-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management. St. Louis, MO. [<http://www.farmdoc.illinois.edu/nccc134>].

# Commodity Price Comovement: The Case of Cotton

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

Paper presented at the NCCC-134 Conference on Applied Commodity Price Analysis,  
Forecasting, and Market Risk Management  
St. Louis, Missouri, April 16-17, 2012

Copyright 2012 by Joseph P. Janzen, Aaron D. Smith, and Colin A. Carter. All rights reserved.  
Readers may make verbatim copies of this document for non- commercial purposes by any  
means, provided that this copyright notice appears on all such copies.

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. The funding support of the Economic Research Service of the US Department of Agriculture under a Cooperative Agreement is gratefully acknowledged.

\* 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](mailto:janzen@primal.ucdavis.edu).

## **Commodity Price Comovement: The Case of Cotton**

**Abstract:** During the commodity price boom and bust of 2007-2008, cotton futures prices rose and fell dramatically in spite of high levels of inventory. At the same time, correlation between cotton and other commodity prices reached historically high levels. These two observations underlie concerns that cotton prices during this period were poor signals of cotton market fundamentals and that the cotton market was 'taken along for a ride' with other commodities. The apparent coincidence of extreme price movement across a broad range of commodities requires an explanation. Were cotton prices driven by the same set of macroeconomic factors as the other commodities? Did cotton markets suffer from supply disruptions at the same time that the other commodities faced disruptions? What was the role of futures market speculators and the rise of commodity index trading?

Economists have been writing about excessive or unexplained comovement among commodity prices since at least Pindyck and Rotemberg (1990). Using this literature as a starting point, we identify potential explanations for commodity price comovement. Past studies have accounted for macroeconomic activity, and cotton-specific supply and demand changes. Tang and Xiong (2010) suggest that speculative pressure due to broad-based commodity index trading may also cause comovement among commodity prices. We develop and estimate a structural vector autoregression model to test the relative contribution of these effects to observed cotton prices. We find that supply and demand shocks specific to the cotton market are the major source of cotton price variation. There is scant evidence of comovement-type effects. While most cotton price spikes are driven by shocks to current net supply, the 2007-2008 spike was caused by higher demand for inventories.

**Key words:** commodity prices, speculation, cotton, comovement.

### **Introduction**

Since 2006, the world has experienced two commodity price booms, where the prices of many, if not all, commodities rose dramatically. Between 2006 and 2008, the price of cotton doubled, the price of wheat tripled, and the price of corn more than tripled. Price increases were not limited to agricultural commodities. Energy prices rose sharply as the price of crude oil nearly tripled and the price of natural gas doubled. Metals prices also rose over this period as the price of copper doubled and the price of silver increased by 50%. As they rose together, commodity prices fell in 2008, before commencing a second boom through 2009 and 2010 (Commodity Research Bureau, 2011).

Commodity price changes such as those observed since 2006 may be more strongly correlated as documented by Silvennoinen and Thorp (2010) because of fundamental supply and demand relationships. Commodities may be substitutes or complements in production or consumption. Commodities may be used as inputs in the production process for other commodities. In addition

to these microeconomic relationships, demand and supply for some commodities may respond to common macroeconomic shifters related to aggregate demand, exchange rates, and interest rates.

Traders active in these commodity markets understand and expect correlation or comovement based on the factors mentioned above. During the recent commodity price boom and bust market participants became concerned that correlation of prices was increasing and was not justified by market fundamentals. It was suggested that misguided sentiment and herd mentality about the general price level of commodities created excess correlation and that some commodity prices were carried along for the ride. Such irrational price behavior poses problems for those who rely upon these markets for reliable price discovery and for those who use these markets to hedge price risk.

If any agricultural commodity might have avoided the boom and bust that prevailed in 2007 and 2008, it might have been cotton. The fundamental supply and demand situation was not bullish; unlike other crops, cotton levels in storage were high. Figure 1 shows annual cotton stocks-to-use ratios since 2000. In 2007-08, the stocks-to-use ratio was above 50%, levels higher than had prevailed since the early 1980's when US government policy encouraged higher cotton stocks. In the absence of a supply-and-demand story, comovement with other commodities seems like a plausible explanation for a cotton price spike.

The idea that price movements in other, unrelated markets might spillover into the market for a commodity such as cotton is not new. We review this literature on commodity price comovement starting with the work of Pindyck and Rotemberg (1990). We ground our own analysis of comovement in the standard supply-of-storage model of commodity price determination. We then propose an empirical methodology that allows us to measure the impact of traditional supply-and-demand fundamentals including the demand for inventories, as well as comovement type effects on observed prices for cotton.

## **Literature review**

To consider the question of comovement in cotton prices, we bring together three strands of the literature on speculation and comovement in markets for storable commodities. The first defines the concept of comovement. The idea that commodity prices move together to a degree that cannot be explained by fundamental factors was initially considered by Pindyck and Rotemberg (1990), who analyzed what they termed the “excess co-movement” of prices for seven commodities. They posited that correlation among the prices of commodities whose fundamentals are unrelated that cannot be explained by macroeconomic effects is excess co-movement. They suggested that this comovement arises because “traders are alternatively bullish or bearish on all commodities for no plausible reason.”

Pindyck and Rotemberg (1990) tested for the presence of comovement after controlling for microeconomic and macroeconomic shocks. To eliminate microeconomic factors, they compared

comovement among the prices of “unrelated” commodities in a seemingly unrelated regressions framework that included macroeconomic variables such as aggregate output, interest rates, and exchange rates as controls. Cross-equation correlation of the residuals from these regressions constitutes evidence of excess comovement. Their test found evidence of persistent excess comovement amongst monthly average prices for wheat, cotton, copper, gold, crude oil, lumber, and cocoa over the period 1960-1985.

Deb, Trivedi, and Varangis (1996) questioned the validity of the test proposed by Pindyck and Rotemberg (1990). They suggested that findings of excess comovement were driven by the assumption of normal and homoskedastic errors in the seemingly unrelated regressions model. Deb, Trivedi, and Varangis (1996) proposed the use of a GARCH framework to account for the non-normality and heteroskedastic nature of commodity price changes. These models found minimal evidence of comovement for the commodities and time period used in Pindyck and Rotemberg (1990).

Ai, Chatrath, and Song (2006) revisited comovement among agricultural commodities. They pointed to the poor explanatory power of the macroeconomic variables used by Pindyck and Rotemberg (1990) and Deb, Trivedi, and Varangis (1996) and suggested that the process of considering “unrelated” commodities was inadequate. They develop a time-series model that incorporates quarterly data on net demand, that is available supply less inventories, and show that this information explains the substantive portion of comovement among prices of the commodities in their dataset. Their data spans five commodities, wheat, barley, oats, corn, and soybeans, and 5 years, from 1997 to 2002. Ai, Chatrath, and Song (2006) note that their structural model explains much of the correlation among related commodities, but it does not explain why commodity prices for fundamentally unrelated commodities may have moved together.

Two more recent developments in the economic literature on commodity price speculation may shed additional light on this problem. In their comprehensive review of the impact of oil market speculation, Fattouh, Kilian, and Mahadeva (2012) discuss the role of speculative storage in determining the price of oil and the impact of new commodity market speculators such as index traders and hedge funds. These two strands of literature point us toward a methodology to study commodity price comovement and an additional explanation for comovement among unrelated commodities. We discuss them in further detail below.

### **Sources of cotton price variability**

We want to consider the source of observed cotton futures price changes. Previous models associated price changes with two “legitimate” causes: shocks to market specific fundamentals and common shocks to macroeconomic variables. Remaining correlation with other markets was due to excessive comovement. The problem in attributing observed price changes to these or other explanations is that price determination occurs inside a set of simultaneous equations; simple attempts to estimate these equations suffer from simultaneous equations bias<sup>1</sup>.

Rather than consider only two explanations for observed prices, our model captures four alternative sources of the determination of cotton futures prices. The model admits comovement as a possible source, but we ground our empirical model in widely accepted theory of price determination for storable commodities. Our starting point is that the real price of cotton in a given month is determined by current demand and supply for cotton, related to production and use flows. We call unexpected changes to cotton-specific production and use flows “net supply” shocks. These shocks will be associated with weather-related supply disruptions, changes in demand for cotton-derived products, and anything other cotton-specific factors that affect the current flow of cotton from production to use.

The current flow demand for cotton is also driven by unexpected variation in real economic activity that affects all commodities. When the global economy is booming, consumers demand more cotton-derived products. Crucially, these shocks are common across commodities. Greater economic activity will increase demand for clothing derived from cotton, bread derived from wheat, fuel derived from crude oil, and most other commodities. Common macroeconomic influences are a legitimate and expected source of commodity price comovement. We refer to these influences as shocks to real economic activity.

When commodities are storable, the above model of price determination is incomplete. Current period supply and demand do not have to be equal because available supply consists of carryover from previous periods and current production may be stored for future use. One consequence of current period supply shocks is that cotton in storage is brought to market. Higher nearby prices induce storers to draw down inventory because returns to selling cotton now are higher relative to holding the cotton into the future. Storability also means that news about future demand and supply can influence nearby prices. News of smaller planted acreage induce traders to revise downward their expectations of future cotton production of crude oil. All else equal, this will increase the demand for cotton inventories in the current period; the current demand curve will shift along the supply curve and the price of cotton rises. This shock cannot be attributed to observed real economic activity, nor to current net supply.

The effect of inventory demand is seen in the graphical representation of the supply and demand for storable commodities due to Eastham (1939) shown in figure 2. Storability implies that when supplies are plentiful, current demand is the sum of consumption demand and the demand for inventories. When supply is tight, prices in the current period rise and buyers of the commodity cease to put new supply into storage, so that the effective demand curve is kinked. In the upper region of this demand curve, supply shocks can no longer be buffered by inventories. Prices are more volatile, in the sense that for a given shock, price movement is greater. This stylized fact is important for the identification of our empirical model.

Demand from inventory holders will affect current prices, but how do inventory holders decide how much to store? They compare the returns to storage with the cost of holding inventory. Returns to storage are determined by the difference between expected future prices and the current price. The cost of holding inventory is comprised of physical storage costs, the opportunity cost of capital, and a convenience yield, the benefit accrued from having stocks on hand when inventories are tight. Kaldor (1939), Working (1949), and others formalized this idea in the “supply-of-storage” model.

The model asserts that a competitive market will add or subtract from inventories to eliminate the opportunities for arbitrage, maintaining the following relationship between current and expected future prices:

$$(1) \quad F_{t,T} - P_t(1 + r_{t,T}) = w_{t,T} - cy_{t,T},$$

where  $P$  is the current price at time  $t$  and  $F$  is the price at time  $t$  for delivery at time  $T$ .  $r$  is the interest rate,  $w$  is the physical cost of storage, and  $cy$  is the convenience yield on inventories held between  $t$  and  $T$ . This model uses convenience yield to explain why firms may keep commodities in storage even when the price for nearby delivery exceeds the expected future price. The presence of a convenience yield is difficult to measure empirically and remains controversial (Brennan, Williams, and Wright, 1997). However, relationship between the price differential for delivery at different dates and the level of inventories is widely acknowledged as a stylized fact in storable commodity markets (Carter and Revoredo-Giha, 2007). For this reason, Geman and Ohana (2009) state that the observed spread is a suitable proxy for inventory demand.

The empirical validity of the supply-of-storage model was questioned in a series of papers by Deaton and Laroque (1992, 1996) who attempted to match predictions from the model to real world data on commodity prices. Recently, Cafiero et al. (2011) empirically confirmed the kinked demand curve for storable commodities as shown in 2 and the presence of high and low volatility periods in prices associated with low and high levels of inventories.

Inventory demand is speculation related to expectations about future prices. Many critics of futures markets and speculation in particular suggest that alternative motives for speculative trading impede the ability of markets to discover the price levels justified by the fundamental factors identified above. For example, speculative traders have often been accused of maintain a herd mentality that causes prices to wildly overshoot justified levels because traders buy or sell to follow trends. A manifestation of similar arguments relates to the recent financialization of commodity futures markets and the presence of commodity index traders. Financialization has been defined as the increased presence of “financial motives, financial markets, and financial actors in the operations of commodity markets.” (UNCTAD, 2011) and the “increased acceptance of (commodity) derivatives as a financial asset” (Fattouh, Kilian, and Mahadeva, 2012). Critics of financialization allege that the large presence of financial firms in commodity markets, trading in large volume but uninformed about commodity supply-and-demand fundamentals, causes prices to conform to the actions of these firms; when they buy, prices rise, whether such price changes are justified or not.

Many financial investors who participate in commodity markets do not have directional views on the prices of a specific commodity such as cotton. Rather they wish to gain exposure to the broad movement of commodity prices because of perceived portfolio diversification benefits. Influential papers such as Gorton and Rouwenhorst (2004) noted a negative correlation between returns from an index of commodity prices and equity or bond returns. A number of indices have been developed to measure broad commodity price movement and many managed funds and exchange-trade funds track these popular indices. 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 indices. Financial firms have developed exchange-traded funds, swap

contracts, and other vehicles to allow individual and institutional investors to track these and similar indices to 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.

Some studies purport to find a statistical linkage between the futures market positions held by CITs and commodity futures prices. Irwin and Sanders (2011) and Fattouh, Kilian, and Mahadeva (2012) provide extensive reviews of these studies, most of which assess whether increased CIT trading predicts the returns on commodity futures contracts. These reviews find that initial studies that found that CIT trading activity had causal effect on price movements (such as Singleton (2011)) were methodologically flawed. A considerable body of evidence (e.g. Stoll and Whaley, 2010; Buyuksahin and Harris, 2011; Irwin and Sanders, 2011) suggests that CIT futures market positions are not associated with price levels or price changes of the underlying futures.

Research by Tang and Xiong (2010) points to an alternative mechanism by which financial speculation and the presence of CITs may impact agricultural commodity prices. Correlation between many commodity prices and the price of crude oil, the most widely traded commodity futures market, has risen over the period in which CIT trading has become prevalent (Silvennoinen and Thorp, 2010). Tang and Xiong (2010) tested the linkage between returns for many commodities and crude oil and concluded that this comovement among prices is driven by the inclusion of commodities into major indices such as the S&P-GSCI and the DJ-UBSCI. The “index inclusion” impact of CITs alleged by Tang and Xiong (2010) follows a similar effect found in equity markets by Barberis, Shleifer, and Wurgler (2005). That study showed that inclusion in a major index leads to comovement among the prices of the index components. This is an impact of speculation that is independent of directional position taking by traditional arbitraging speculators that we have identified as an inventory demand shock above.

Recall that Pindyck and Rotemberg (1990) cited speculative herding as the reason for excessive comovement in their original comovement study. This explanation may conflate the portion of price variation due to unexpected changes to inventory demand with an irrational sort of speculation based on random or incorrect information. Some may attribute a illegitimacy to price changes based on speculation where none exists. Therefore, it is important to differentiate between speculation based upon rational expectations of future prices relative to current ones as embodied in “supply of storage”-type models and speculation based upon irrational beliefs that general commodity price levels will continue to rise or fall. We might think of this second type of speculation as being “excessive” in the sense that comovement in excess of that generated by irrational speculation may be called excessive.

Ai, Chatrath, and Song (2006) note that their model “cannot explain common movements in inventories (if any) across commodities”. However, if commodity prices move together because of common inventory demand shocks, then their model will misidentify these common movements as being part of excessive comovement. Therefore, we identify inventory demand shocks and shocks due to external markets separately. To the extent to which we can differentiate between these types



of speculative shocks, we will have improved upon Ai et al's work.

In summary, we identify four types of shocks relevant to the determination of cotton prices in a given period. These are shocks to real economic activity (flow demand), net current production (flow supply), rational speculation (inventory demand), and external market price changes (comovement). While the external market or comovement piece is relevant to compare our work with previous comovement studies, analysis of the other pieces can help us understand the genesis of price spike events. With respect to cotton, we can assess the causes of the 2007-08 price spike as compared to other price booms and busts.

### *A structural vector autoregression model for cotton prices*

We have identified four key determinants of cotton prices: real economic activity, comovement, inventory demand, and net supply. We know that there exists an exceedingly rich and complex set of factors that underly each of these determinants. A complete representation of all interacting forces would include many structural equations representing the interaction of these economic forces across time and space. The potential to misspecify these relationships is large.

Sims (1980) introduced vector autoregression (VAR) as an alternative modeling technique to strips away much of the contrivance surrounding complex large-scale models. We use a particular form the VAR model, structural VAR (SVAR) that allows for the identification of a set of orthogonal shocks, each associated with one of the variables included in the data vector,  $y_t$ , and an accompanying structural interpretation. We wish to identify four structural shocks related to real economic activity, external markets, cotton inventory demand, and cotton net supply.

Before we describe our specification of a SVAR model, we must outline some general notation. Denote the  $K \times 1$  vector of included variables as  $y_t$  and assume that  $y_t$  can be approximated by an autoregressive process of some finite order,  $p$ , so that the SVAR model can be written as:

$$(2) \quad A(L)y_t \equiv (A_0 - A_1L - A_2L^2 - \dots - A_pL^p)y_t = u_t$$

where  $B(L)$  is the the autoregressive polynomial of order  $p$ . The objective is to identify the structural shocks to this system,  $u_t$ . Since the structural model is not directly observable, the structural shocks must be related to a reduced-form VAR representation that can be estimated. This reduced form VAR is:

$$(3) \quad B(L)y_t \equiv (I - B_1L - B_2L^2 - \dots - B_pL^p)y_t = \varepsilon_t$$

Pre-multiplying both sides of the SVAR model by the inverse of the first term in the autoregressive polynomial implies that the reduced form innovations,  $\varepsilon_t$ , are equal to  $A_0^{-1}u_t$ . Essentially, the reduced-form shocks are a weighted sum of the structural shocks. We can estimate a reduced-form VAR for any  $y_t$  to estimate reduced form residuals,  $\hat{\varepsilon}_t$ . Unless we make some assumptions about the relationship between the reduced-form and structural residuals. we cannot relate  $\varepsilon_t$  to  $u_t$  because the condition relating them contains  $K$  equations and  $K \times K$  unknowns.

Two standard assumptions to identify the structural shocks are normalization and recursion. Normalization refers to setting the diagonal elements of  $A_0$  to equal one. This implies that the magnitude of the identified structural shocks are interpreted relative to each other. The recursion assumption uses a Cholesky decomposition to identify a unique matrix  $A_0$  where the above-diagonal elements are zero. Relying entirely on recursion to identify the  $A_0$  matrix implies a strict ordering to the structural shocks. The shock associated with the variable placed in the  $k$ th row of the data vector  $y_t$  cannot have a contemporaneous effect on any of the variables placed in the rows  $(1, \dots, k - 1)$ .

Using the recursion assumption in empirical SVAR work requires the modeler know the underlying process and have ready evidence that feedback effects cannot occur. Since such evidence is rarely available, Kilian (2011) outlines a series of alternatives, the most popular being the use of sign restrictions on the elements of the  $A_0$  matrix. We use an alternative method well-suited to the storable commodity price context: an estimator that relies on differences in the variance of the structural shocks. Rigobon (2003) developed an Identification through Heteroskedasticity (ItH) method to identify the parameters of the  $A_0$  matrix. The ItH estimator is a method-of-moments estimator that draws upon the second moments of the condition that relates the reduced-form and structural shocks. In general, the moment conditions are given by the following identity:

$$(4) \quad A_0 \Omega^r A_0' = \Sigma_u^r,$$

where  $A_0$  is the translation matrix,  $\Sigma_u^r$  is the variance-covariance matrix derived from the structural residuals, and  $\Omega^r$  is the variance-covariance of the reduced-form shocks. Identification of the  $A_0$  matrix by this method is impossible if all of the parameters in the matrix are unrestricted. However, if the residuals can be thought to be derived from two or more volatility regimes, such that a unique reduced-form variance-covariance matrix can be calculated for each regime, then the  $A$  matrix is identified. In the equation above, the  $r$  subscript denotes parameters that vary between the volatile and tranquil regimes. The presence of multiple volatility regimes defines additional moment conditions that can be used to solve for the unknown parameters in  $A_0$  and  $\Sigma_u^r$ . The simplest case consists of two regimes, one volatile and one tranquil. In this case the model is just-identified.

### *Choosing variables and interpreting structural shocks*

We include four variables in  $y_t$ : real economic activity, *rea*, the real price in the external market, for example the price of crude oil, *rpo*, the between distant and nearby futures prices for cotton, *sprc*, and the real price of nearby cotton futures, *rpc*. The variables we include in the model determine the structural interpretation we give to each shock. Our strategy is to include variables that measure three of the major price determinants directly along with the price of cotton, our main variable of interest. The first three variables identify the real economic activity, comovement, and inventory demand shocks respectively, which leaves residual variation in cotton prices to represent net supply shocks.

For real economic activity, we use an index of economic activity developed in Kilian (2009). This index employs ocean freight rates as a proxy for global demand for goods. Kilian notes an empirically documented correlation between freight rates and economic activity. We do not want a measure biased toward any one country or region of the world; the freight rate index will rise if economic activity rises in any part 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 can approximate commodity price movements that may be associated with speculation-induced comovement using the prices of major commodities themselves. If the implications of the financialization hypothesis are correct, then we should find that non-agricultural commodity prices have driven cotton price changes. According to Tang and Xiong (2010), cotton prices should have become increasingly related to crude oil since 2004, so we include the price of nearby West Texas Intermediate crude oil futures in our model. As a test of robustness, we repeat our analysis using the value of the silver and copper as representative measures of external market price movement for other commodity classes.

We do not have monthly measures of cotton inventories that might measure the activities of speculative storage, but economic theory provides an alternative measure. As noted above, the spread between distant and nearby prices provides a measure of the returns to storage. When stocks are tight, spreads will be low or negative to induce storers to release inventories to accommodate current demand. Conversely when stocks are plentiful, the spread between distant and nearby prices will approach the full cost of holding physical stocks between the two periods.

We include the nearby futures prices for cotton as the final variable in the model. Our model will decompose the changes observed in cotton price to three structural shocks associated with the other variables. The fourth structural shock, which we interpret as being due current cotton-specific supply and demand factors, encompasses residual variation in price.

We use a combination of normalization, recursion, and identification through heteroskedasticity to identify our SVAR model. As in previous studies of commodity price dynamics (Kilian, 2009), we make real economic activity the first variable in  $y_t$  and assume that there is no contemporaneous feedback from the other variables to the level of real economic activity. This is equivalent to setting the terms in  $A_{1i}$  equal to zero, for all  $i$ . We make a similar assumption for the external market price variable, the second variable in  $y_t$ . This assumption implies that any contemporaneous comovement effects run from other markets into the cotton market. In some sense this is restrictive, but also strengthens confidence in our conclusions if we find no evidence for comovement effects because this setup gives comovement effects precedence over cotton-specific structural shocks.

We cannot assume that inventory demand shocks take precedence over net supply shocks, or vice-versa. Contemporaneous observed changes in prices and spreads could be the result of inventory demand shocks or net supply shocks. We would expect a positive inventory demand shock to have a positive impact on price and a positive impact on the term spread. A net supply shock should also have a positive price impact but a negative impact on spreads. Importantly, net supply and inventory demand shocks may affect prices and spreads simultaneously, so assigning priority

to one or the other is not realistic.

Equations (5) and (6) show the moment conditions for our estimation procedure. If we were to leave all of the  $A$  matrix parameters unrestricted and leave the  $\sigma$  terms unrestricted across regimes, we would have  $K(K + 1) = 20$  parameters to be identified, the 12 off-diagonal terms in the  $A$  matrix plus the two sets of four structural shock variances. The presence of two regimes creates  $K(K + 1)/2 = 10$  moment conditions in each regime. (Recall that only half of the off-diagonal terms in  $\Sigma_\varepsilon$  define unique moment conditions due to the symmetry of the variance-covariance matrix.) When we set up the model in this way, we find that the *rea* and *rpo* shocks are very weakly identified across regimes, so we restrict  $\sigma^{rea}$  and  $\sigma^{rpo}$  to be constant across regimes, essentially re-adopting the recursion assumption for the first two structural shocks. This also implies that only the lower right hand corner of the  $\Omega$  matrix varies across regimes. (A similar identification strategy was used in Rigobon and Rodrik (2005)). This is shown in equations (5) and (6).

$$(5) \quad \begin{bmatrix} 1 & 0 & 0 & 0 \\ A_{21} & 1 & 0 & 0 \\ A_{31} & A_{32} & 1 & A_{34} \\ A_{41} & A_{42} & A_{43} & 1 \end{bmatrix} \begin{bmatrix} \Omega_{11} & & & \\ \Omega_{21} & \Omega_{22} & & \\ \Omega_{31} & \Omega_{32} & \Omega_{33}^r & \\ \Omega_{41} & \Omega_{42} & \Omega_{43}^r & \Omega_{44}^r \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ A_{21} & 1 & 0 & 0 \\ A_{31} & A_{32} & 1 & A_{34} \\ A_{41} & A_{42} & A_{43} & 1 \end{bmatrix}' =$$

$$(6) \quad \begin{bmatrix} \sigma_{rea} & 0 & 0 & 0 \\ 0 & \sigma_{rpo} & 0 & 0 \\ 0 & 0 & \sigma_{spr}^r & 0 \\ 0 & 0 & 0 & \sigma_{rpc}^r \end{bmatrix}$$

Restricting  $\sigma^{rea}$  and  $\sigma^{rpo}$  to be constant across regimes gives us thirteen parameters to be identified, seven in the  $A$  matrix and six structural shock variances. In this setup, the model remains just-identified; there are thirteen moment conditions. Multiplying the terms in equation (5) produces a symmetric matrix, where again only the bottom right hand corner contains terms that vary across regimes. Since the matrix is symmetric only three terms in product of (5) vary across regimes. Therefore, the thirteen unique moment conditions implied by this setup are accounted for as follows: there are seven unique conditions in the first two columns and six unique conditions represented by the three bottom right hand corner terms from each regime. This setup implies that if the variance of the reduced-form and structural shocks does not vary across regimes, then the parameters remain unidentified.

## Data

We use monthly data spanning the period 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 all 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 and removes our ability to observe price response in these periods.

As stated above, the four variables in our model are real economic activity, external commodity market prices, the intertemporal price spread in the cotton market, and the price of cotton. We use the real economic activity index developed by Kilian (2009). The index is derived from dry cargo shipping rates measured in deviations from long term trend. Data on cotton and other commodity market prices are collected 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, again deflated and adjusted for seasonality and trend. The cash price series is used because crude oil futures only began trading in 1983. We use nearby copper and silver futures prices to represent external markets for industrial and precious metals, just as crude oil is representative of the price of energy commodities.

The cotton calendar price spread, a proxy measure for the demand for inventories, is the log-difference of the six-most distant and the nearby futures contract price. The difference between the time to expiration is constant for these two contracts, so the resulting spread represents the term structure of cotton prices over a constant period of time. Because there are five delivery month contracts for each calendar year, this measure is always an old-crop/new-crop spread, representing expectations about the scarcity of cotton now compared to a future period when production can respond. The cotton price and spread variables are also adjusted for each of the months in the year 1986 to control for the effects of implementation of the 1986 Farm Bill and the Inventory Protection Certificate program, where the term structure of cotton prices shifted dramatically because of policy changes that made it less attractive for US firms to store cotton (Hudson and Coble, 1999).

## Estimation and results

We estimate a reduced-form VAR for the four variables constructed above using ordinary least squares with two lags. The lag length was selected using AIC and SBIC information criteria. 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 set corresponding to the tranquil and volatile regimes. We select dates for each regime using an *ad hoc* rule based on predictions from price-of-storage-type theoretical models, 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. This creates eight 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<sup>2</sup>.

From the reduced-form residuals for each regime, we calculate the variance-covariance matri-

ces,  $\Sigma_r$ . Because we do not use identification through heteroskedasticity to identify parameters in the first two rows or columns of  $A$ , we replace all terms the first two rows of  $\Sigma_r$  with the variance and covariance terms from the variance matrix calculated using all reduced-form residuals. We define the set of constraints on the  $A$  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  $A$  and  $\Omega_r$  by minimizing the distance function defined by the difference  $A\Omega_r A' - \Sigma_{e,r}$ .

Having solved for the model parameters, we calculate a set of orthogonal structural shocks for each period. We calculate impulse response functions for all model variables with respect to each structural shocks and generate confidence bounds for the impulse responses using a wild bootstrap procedure due to 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 impulse responses and historical decompositions below, first for the case where crude oil represents the external market following Tang and Xiong (2010) and then for alternative external markets.

### *Impulse response functions*

Figure 3 plots the time path for the response of each variable in our model to the economic activity, external market, speculative demand, and net supply shocks for the case where crude oil is the external market. The dashed lines represent the 90% confidence interval about the average response generated using 1000 bootstrap replications. These graphs demonstrate, based on the the average response observed in the data, how each variable in the model would respond to a hypothetical one unit shock from the assumed zero mean. Note that the normalization used to identify the model implies that each of the shocks causes an increase in the price of cotton. In particular, positive 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. We want to ensure that the direction of the observed structural shocks is consistent with our expectations. 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 3, we can check the validity of the identification scheme used to identify the endogenous net supply and inventory demand shocks in the cotton market. Price responds positively to inventory demand shocks and net supply shocks. If an inventory demand occurs, the spread should also 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 inventory demand shock displays some evidence of overshooting: prices increase quickly in the months following the shock and then drop towards zero. In contrast, net supply shocks have a long-lasting impact on prices.

Figure 3 shows that external forces have a relatively small impact on cotton prices, relative to the inventory 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.

Figures 4 and 5 present similar sets of impulse responses for the cases where copper and silver are used in place of crude oil as a measure of external commodity price movement. These results suggest that external shocks, due to aggregate demand for commodities or speculative comovement effects, have relatively small impacts on cotton prices.

### *Historical decomposition of cotton price shocks*

The impulse response analysis is incomplete in the sense that it 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 in any period.

Most of the variation in cotton prices is due to the the two cotton-market-specific shocks. Longer, smaller swings in price are attributed to the real economic activity and external market shocks. This 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 minute relative to the inventory 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 (2010) 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, Rausser, and Smith (2012) find that between its 2003 low and 2008 high, the contribution of real economic activity to corn prices increased by approximately 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%<sup>3</sup>. This is not trivial, 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.

### *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 four series shown are the observed cotton price and the cotton price in the absence of external market, inventory demand, and net supply shocks.

During the 2006-2011 period, comovement shocks are nearly imperceptible. We find that comovement shocks related to crude oil prices raised cotton prices only 1% at the peak of their impact. In absolute terms, the maximum cotton price impact of comovement shocks at any point in our analysis is approximately \$0.01 per pound. Relative to the price volatility observed over this period such an effect is miniscule; cotton futures prices rose nearly \$0.50 per pound in 2007-08 and by more than \$1.20 per pound in 2010-11.

Elevated cotton prices in 2010-2011 were not a repeat of the events of 2007-2008; these two price spikes have very different origins. The 2008 price spike would have been non-existent without inventory demand shocks. On the other hand, inventory demand shocks had very little to do with rising prices in 2010-11. Figure 7 shows that absent the shock to net supply in 2010-11, prices would have remained close to historic levels, rather than setting records.

In March 2008 at the peak of initial spike, we estimate that prices would have been 26% lower without inventory demand shocks and 11% higher without net supply shocks. In contrast, cotton prices would have been % lower without net supply shocks and just 12% lower without inventory demand shocks in March 2011 when prices peaked that year. The role of inventory demand in driving the 2008 price spike seems related to ongoing fear at that time of further price increases causing firms to add to inventories. Our findings for the 2011 price spike are consistent with information about supply disruptions that occurred in 2010 and 2011.



### *Checking robustness using silver and copper as external markets*

Figures 8-11 show similar historical decompositions and counterfactual analyses as shown earlier for crude oil, but with silver and copper used as the external market. These results are less conclusive about the unimportance of external market shocks. The pattern of cotton price movement from 2006 to 2011 in the absence of external market shocks is largely the same when silver and copper represent external market movements. Price spikes still occur in March 2008 and March 2011, but at lower levels.

Counterfactual analysis using silver and copper suggests that comovement raised prices by approximately 10-25% over the 2006-2011 period. One potential explanation for this result is the common influence of Chinese demand on cotton and copper prices. Variation that we have labelled an external market shock may be picking up the influence of economic growth in China, rather than speculative comovement. 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 8-11.

## **Conclusions**

We use a structural vector autoregression model to attribute observed price changes in cotton to four unique structural explanations: real economic activity, comovement induced by speculative trading, the demand for inventories, and shocks to current net supply. The comovement-driven portion of cotton prices represents the impact of speculation that is unrelated to expected future prices. Such speculation is attributed to commodity index traders who may be responsible for the recent periods of elevated prices in agricultural markets.

We find the last two of our explanations, factors specific to cotton supply and demand, are the major determinants of cotton prices. Consequently, complaints about the accuracy of cotton futures price discovery based on the belief that index traders drive prices are unfounded. We specify our model to capture comovement with non-agricultural markets as suggested by papers such as Tang and Xiong (2010). We even allow comovement shocks to take precedence over cotton-specific shocks and we find minimal comovement effects at the frequency of our data. In this sense, cotton futures markets have performed efficiently.

Unlike studies of other commodity markets using a similar approach, we find that broad trends in global commodity demand related to real economic activity matter less than cotton-specific supply-and-demand. 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 inventory demand, likely induced by expectations of higher future prices, drove prices higher in 2008.

## Notes

<sup>1</sup>Indeed, Pindyck and Rotemberg (1990) only consider variation in prices due to macroeconomic factors because this was the only explanation that could plausibly be exogenous to individual commodity markets.

<sup>2</sup>We also tested similar *ad hoc* rules for setting the regime windows and found that our results were robust to other specifications

<sup>3</sup>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.

## References

- Ai, C., A. Chatrath, and F. Song. 2006. "On the Comovement of Commodity Prices." *American Journal of Agricultural Economics* 88:574–588.
- Barberis, N., A. Shleifer, and J. Wurgler. 2005. "Comovement." *Journal of Financial Economics* 75:283 – 317.
- Brennan, D., J. Williams, and B.D. Wright. 1997. "Convenience Yield without the Convenience: A Spatial-Temporal Interpretation of Storage Under Backwardation." *The Economic Journal* 107:1009–1022.
- Buyuksahin, B., and J.H. Harris. 2011. "Do Speculators Drive Crude Oil Futures Prices?" *The Energy Journal* 0.
- Cafiero, C., E.S. Bobenrieth, J.R. Bobenrieth, and B.D. Wright. 2011. "The empirical relevance of the competitive storage model." *Journal of Econometrics* 162:44 – 54.
- Carter, C.A., G.C. Rausser, and A.D. Smith. 2012. "The Effect of Ethanol Production on Corn Prices." Working paper.
- Carter, C.A., and C.L. Revoredo-Giha. 2007. "The Working Curve and Commodity Storage under Backwardation." *American Journal of Agricultural Economics* 89:864–872.
- Commodity Research Bureau. 2011. "DataCenter CD." Futures price database, Commodity Research Bureau, Chicago, IL.
- Deaton, A., and G. Laroque. 1996. "Competitive Storage and Commodity Price Dynamics." *Journal of Political Economy* 104:pp. 896–923.
- . 1992. "On the Behaviour of Commodity Prices." *The Review of Economic Studies* 59:1–23.
- Deb, P., P.K. Trivedi, and P. Varangis. 1996. "The excess co-movement of commodity prices reconsidered." *Journal of Applied Econometrics* 11:275–291.
- Eastham, J.K. 1939. "Commodity Stocks and Prices." *The Review of Economic Studies* 6:pp. 100–110.
- Fattouh, B., L. Kilian, and L. Mahadeva. 2012. "The Role of Speculation in Oil Markets: What Have We Learned so Far?" *SSRN eLibrary*, pp. .
- Geman, H., and S. Ohana. 2009. "Forward curves, scarcity and price volatility in oil and natural gas markets." *Energy Economics* 31:576 – 585.
- Goncalves, S., and L. Kilian. 2004. "Bootstrapping autoregressions with conditional heteroskedasticity of unknown form." *Journal of Econometrics* 123:89 – 120.
- Gorton, G., and K.G. Rouwenhorst. 2004. "Facts and Fantasies about Commodity Futures." Working Paper No. 10595, National Bureau of Economic Research, June.

- Hudson, D., and K. Coble. 1999. "Harvest contract price volatility for cotton." *Journal of Futures Markets* 19:717–733.
- Irwin, S.H., and D.R. Sanders. 2011. "Index Funds, Financialization, and Commodity Futures Markets." *Applied Economic Perspectives and Policy*, pp. .
- Kaldor, N. 1939. "Speculation and Economic Stability." *The Review of Economic Studies* 7:1–27.
- Kilian, L. 2009. "Not All Oil Price Shocks Are Alike: Disentangling Demand and Supply Shocks in the Crude Oil Market." *American Economic Review* 99:1053–69.
- . 2011. "Structural Vector Autoregressions." In *Handbook of Research Methods and Applications in Empirical Macroeconomics*. Edward Elgar.
- Kilian, L., and D. Murphy. 2010. "The Role of Inventories and Speculative Trading in the Global Market for Crude Oil." *SSRN eLibrary*, pp. .
- Pindyck, R.S., and J.J. Rotemberg. 1990. "The Excess Co-Movement of Commodity Prices." *The Economic Journal* 100:pp. 1173–1189.
- Rigobon, R. 2003. "Identification through Heteroskedasticity." *The Review of Economics and Statistics* 85:pp. 777–792.
- Rigobon, R., and D. Rodrik. 2005. "Rule of law, democracy, openness, and income." *Economics of Transition* 13:533–564.
- Silvennoinen, A., and S. Thorp. 2010. "Financialization, Crisis and Commodity Correlation Dynamics." Research Paper Series No. 267, Quantitative Finance Research Centre, University of Technology, Sydney, Jan.
- Sims, C.A. 1980. "Macroeconomics and Reality." *Econometrica* 48:pp. 1–48.
- Singleton, K.J. 2011. "Investor Flows and the 2008 Boom/Bust in Oil Prices." Unpublished.
- Stoll, H.R., and R.E. Whaley. 2010. "Commodity Index Investing and Commodity Futures Prices." *Journal of Applied Finance*, pp. .
- Tang, K., and W. Xiong. 2010. "Index Investment and Financialization of Commodities." Working Paper No. 16385, National Bureau of Economic Research, September.
- UNCTAD. 2011. "Price Formation in Financialized Commodity Markets: The Role of Information." Working paper, United Nations Conference on Trade and Development.
- Working, H. 1949. "The Theory of Price of Storage." *The American Economic Review* 39:pp. 1254–1262.

## Figures

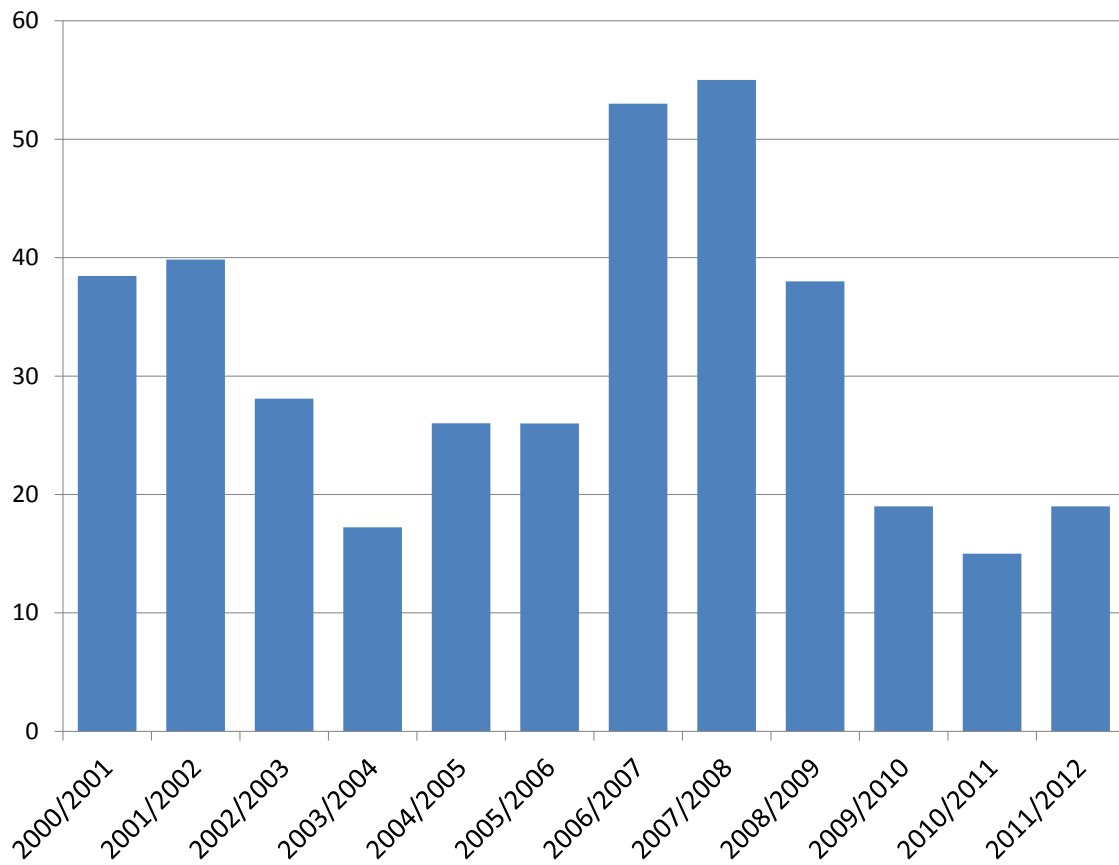


Figure 1: US crop-year ending stocks-to-use, 2000-01 to 2011-12

Source: USDA PS&D Online

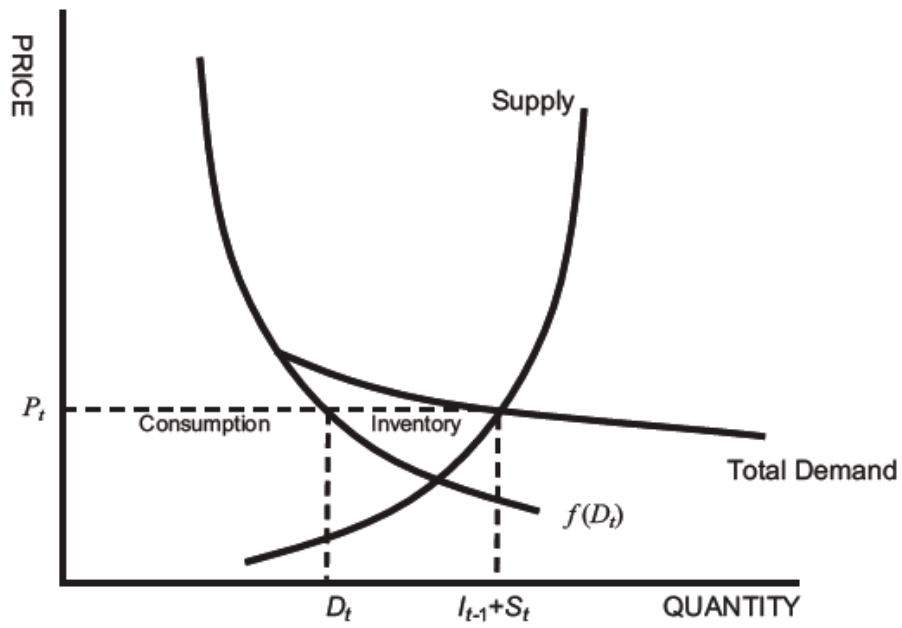


Figure 2: Equilibrium in a commodity market with storage

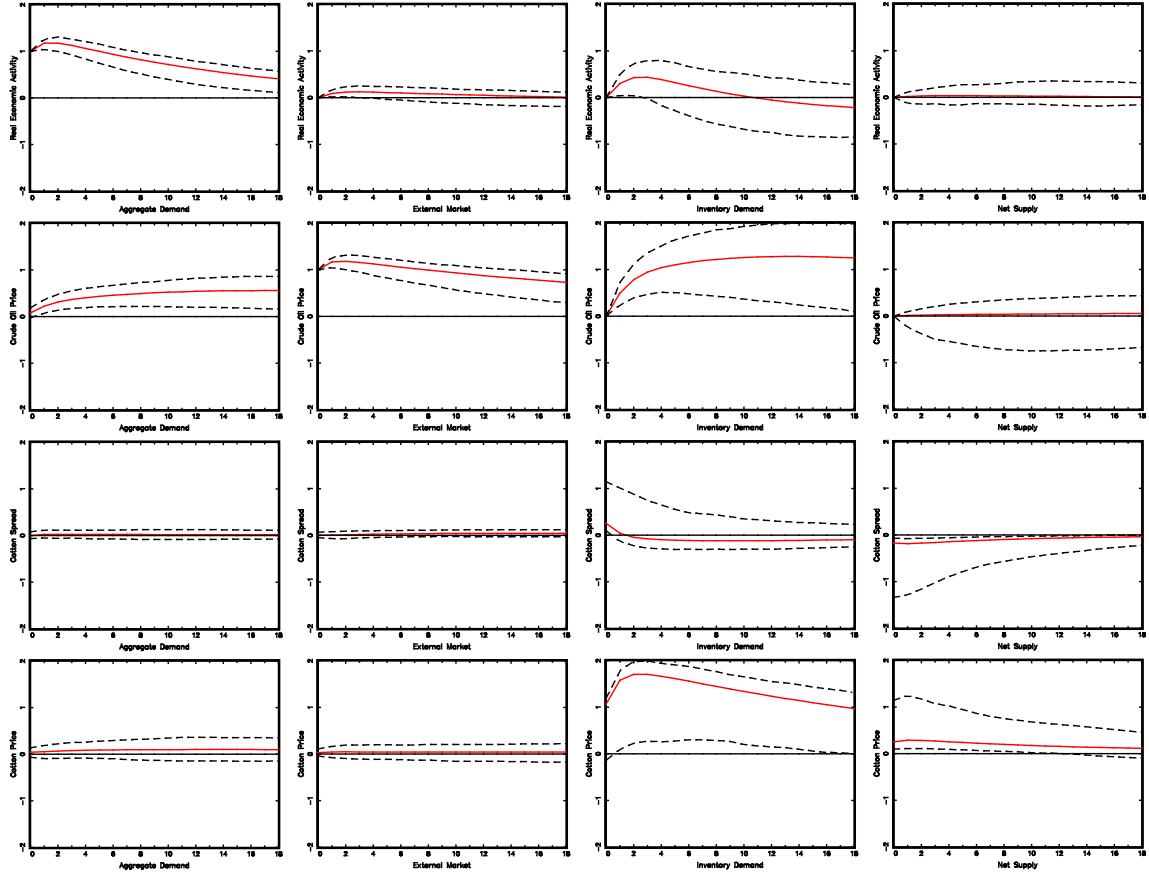


Figure 3: Impulse response functions with crude oil as external market

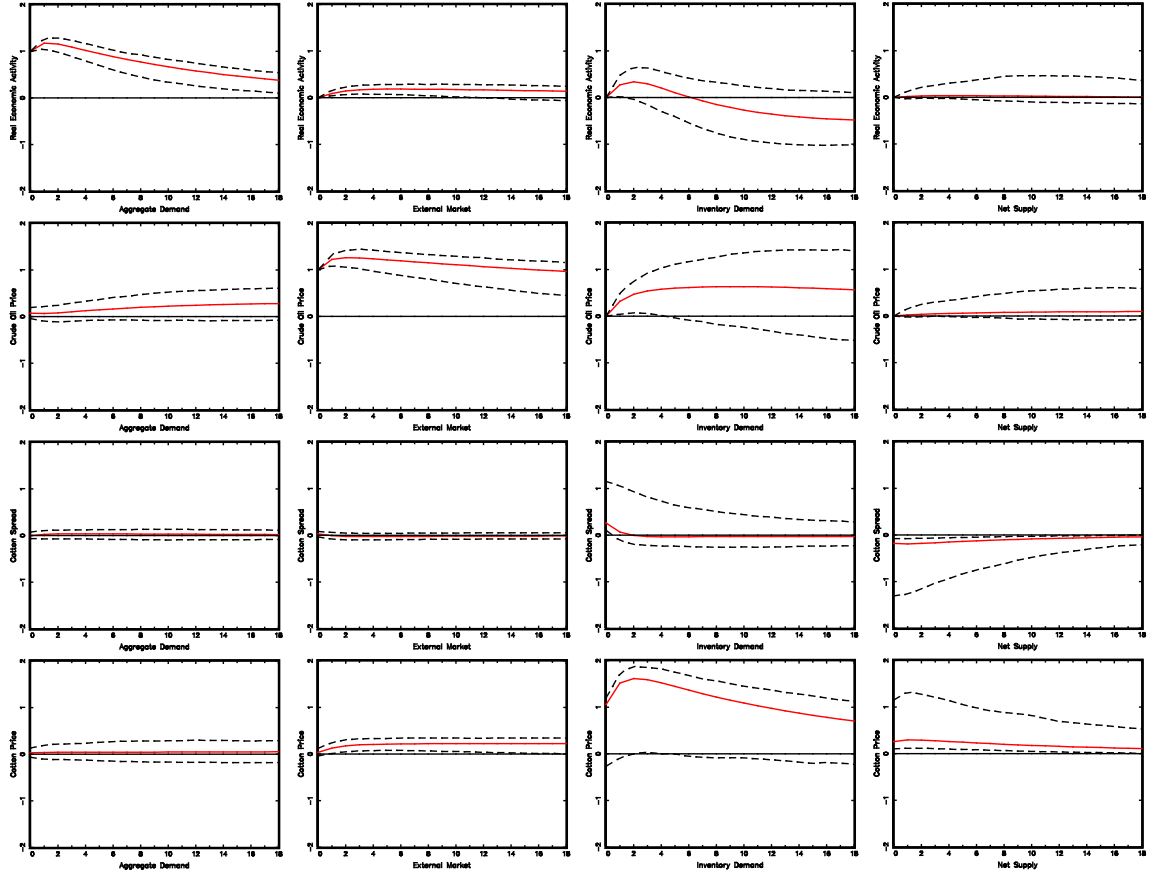


Figure 4: Impulse response functions with copper as external market



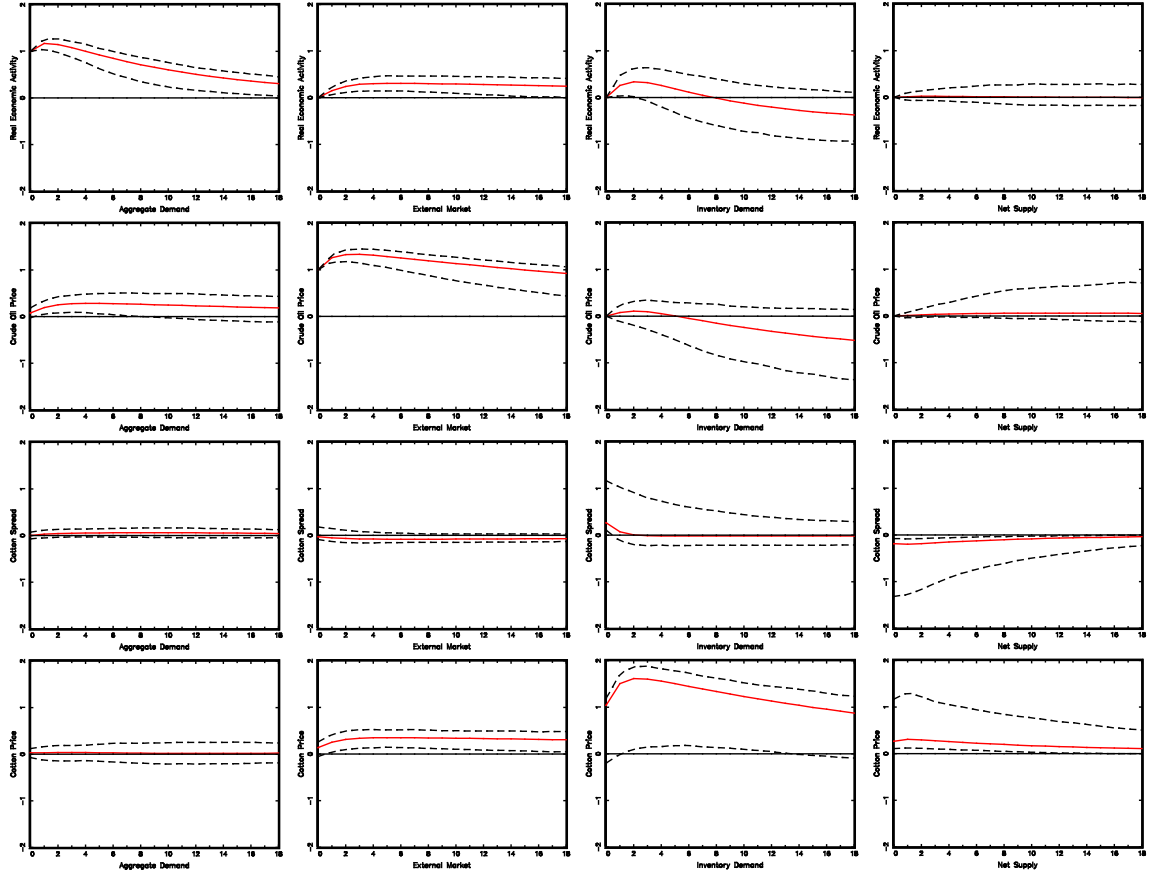


Figure 5: Impulse response functions with silver as external market

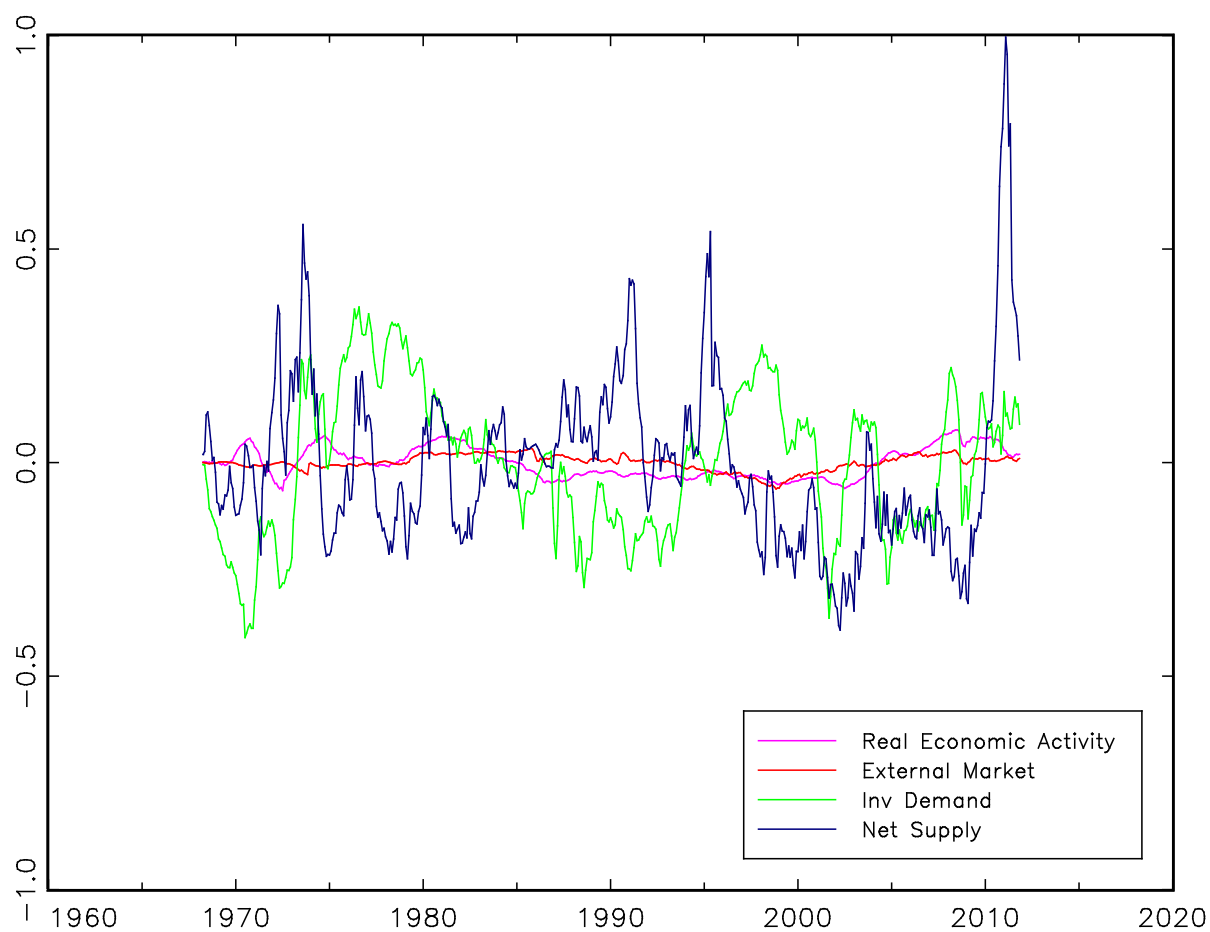


Figure 6: Historical decomposition of structural shocks with crude oil as external market

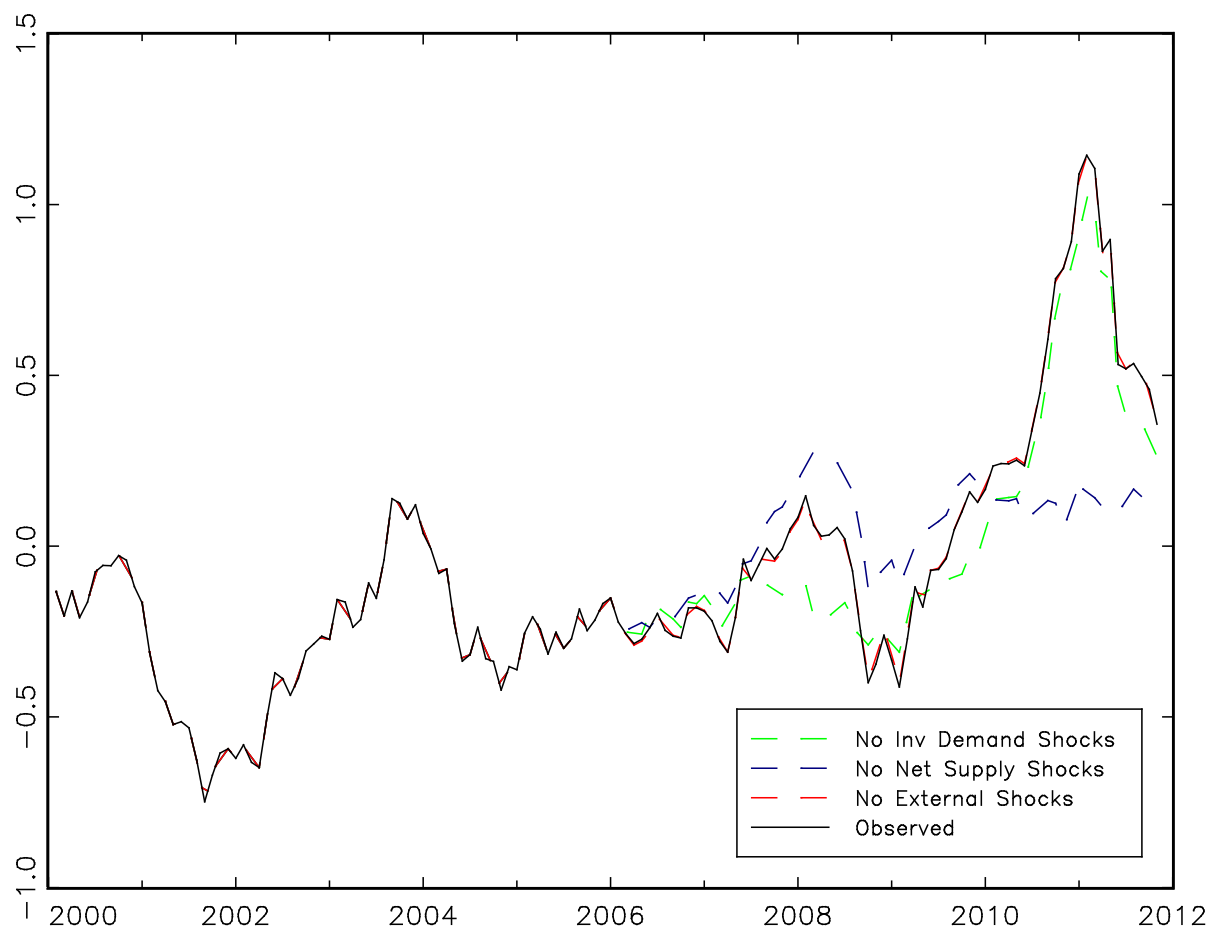


Figure 7: Counterfactual analysis for SVAR with crude oil as external market

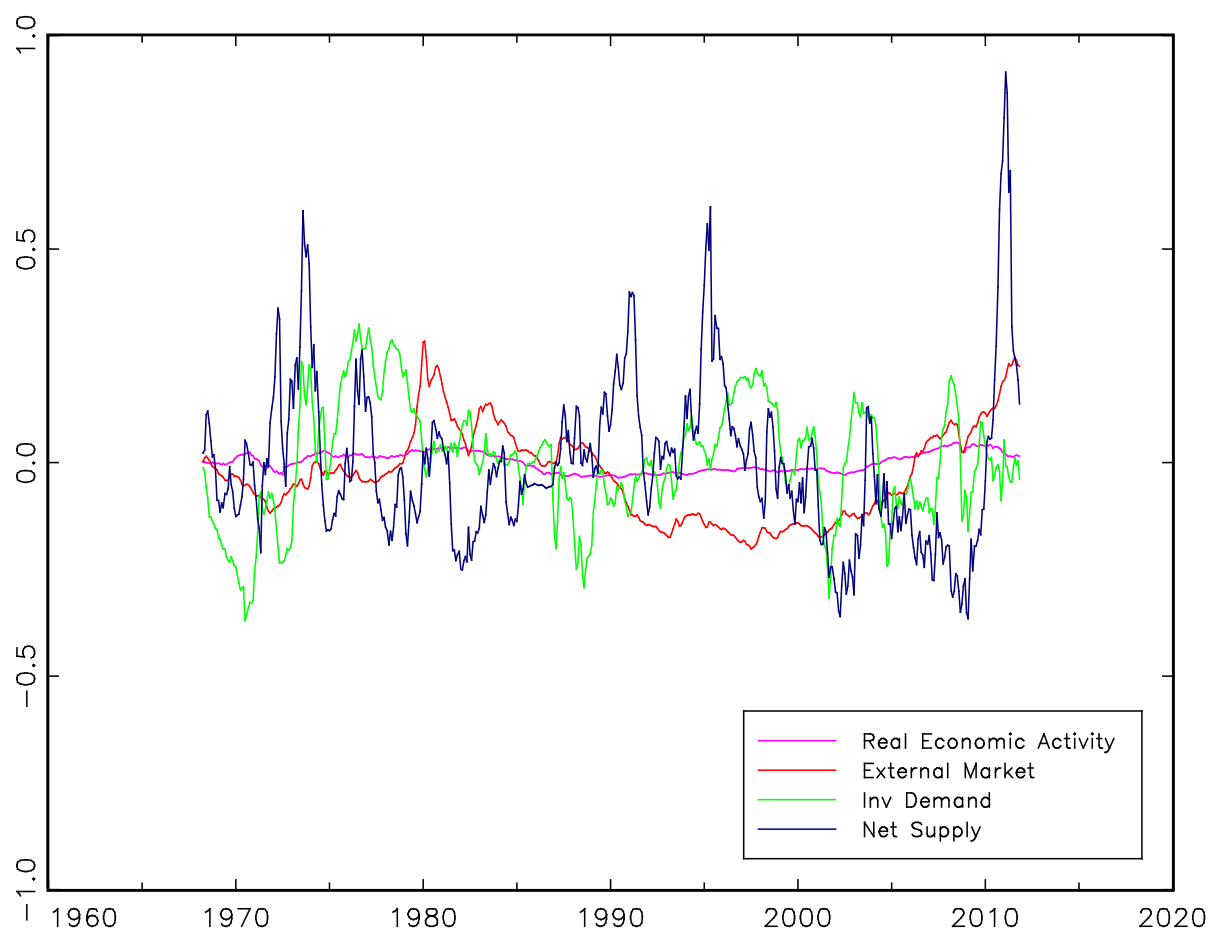


Figure 8: Historical decomposition for SVAR with silver as external market

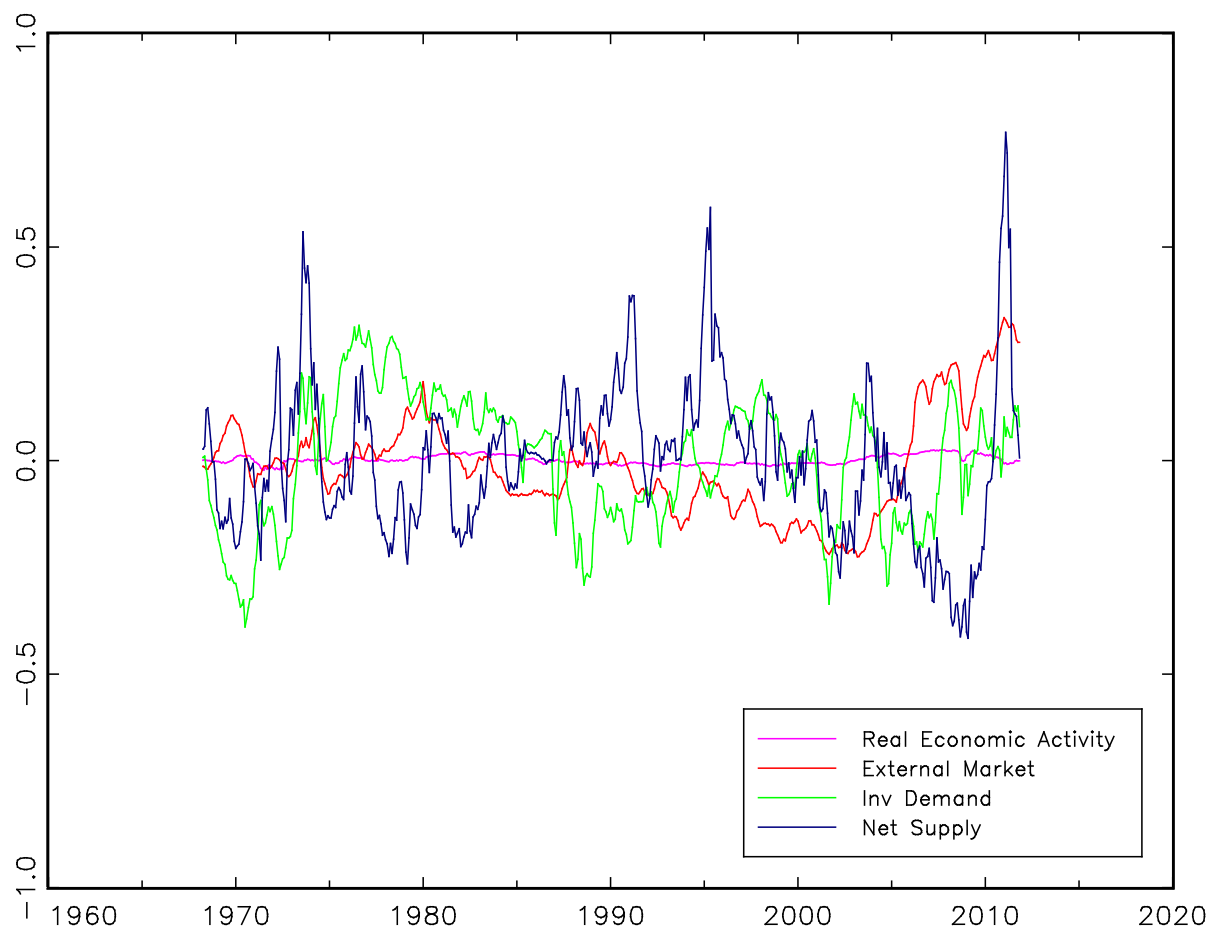


Figure 9: Historical decomposition for SVAR with copper as external market

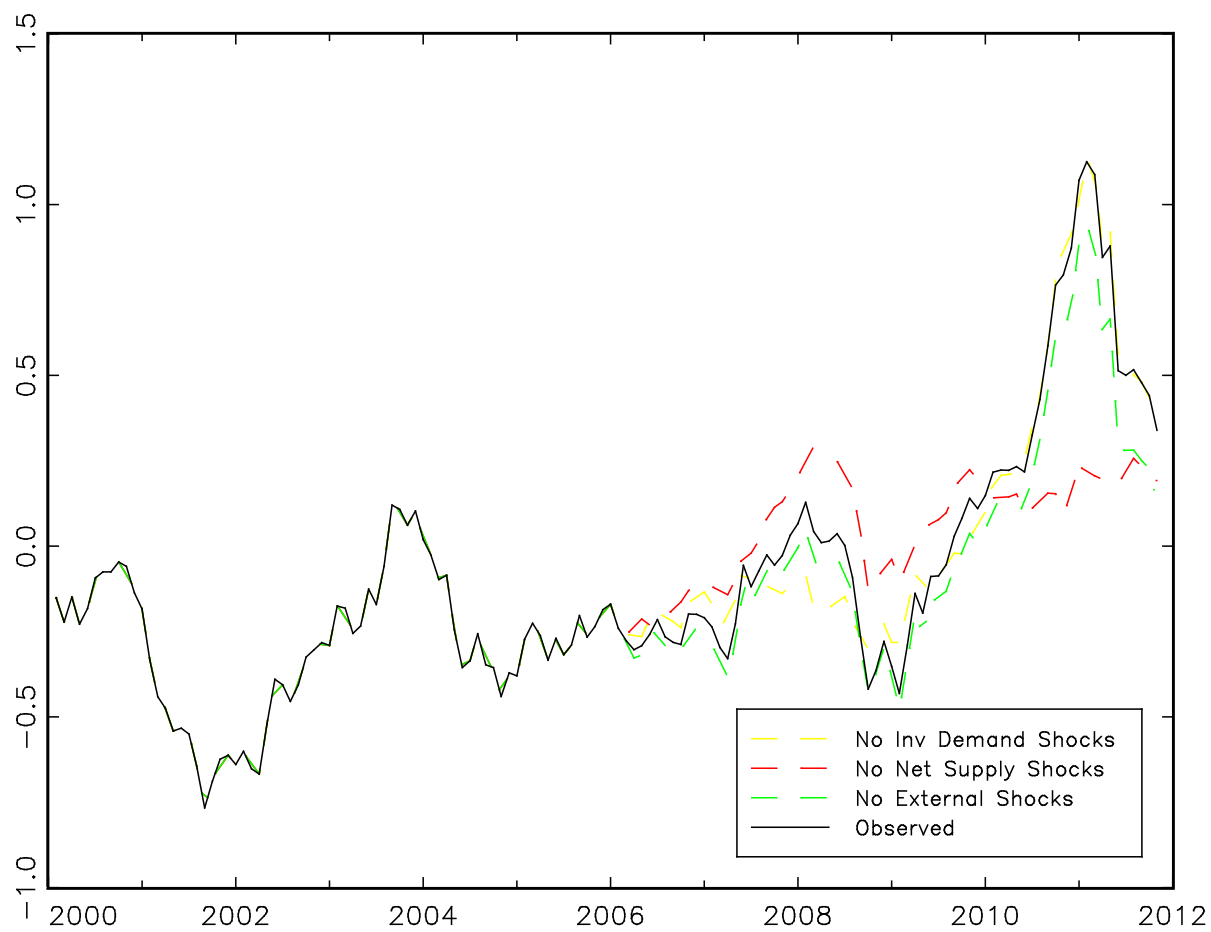


Figure 10: Counterfactual analysis for SVAR with silver as external market

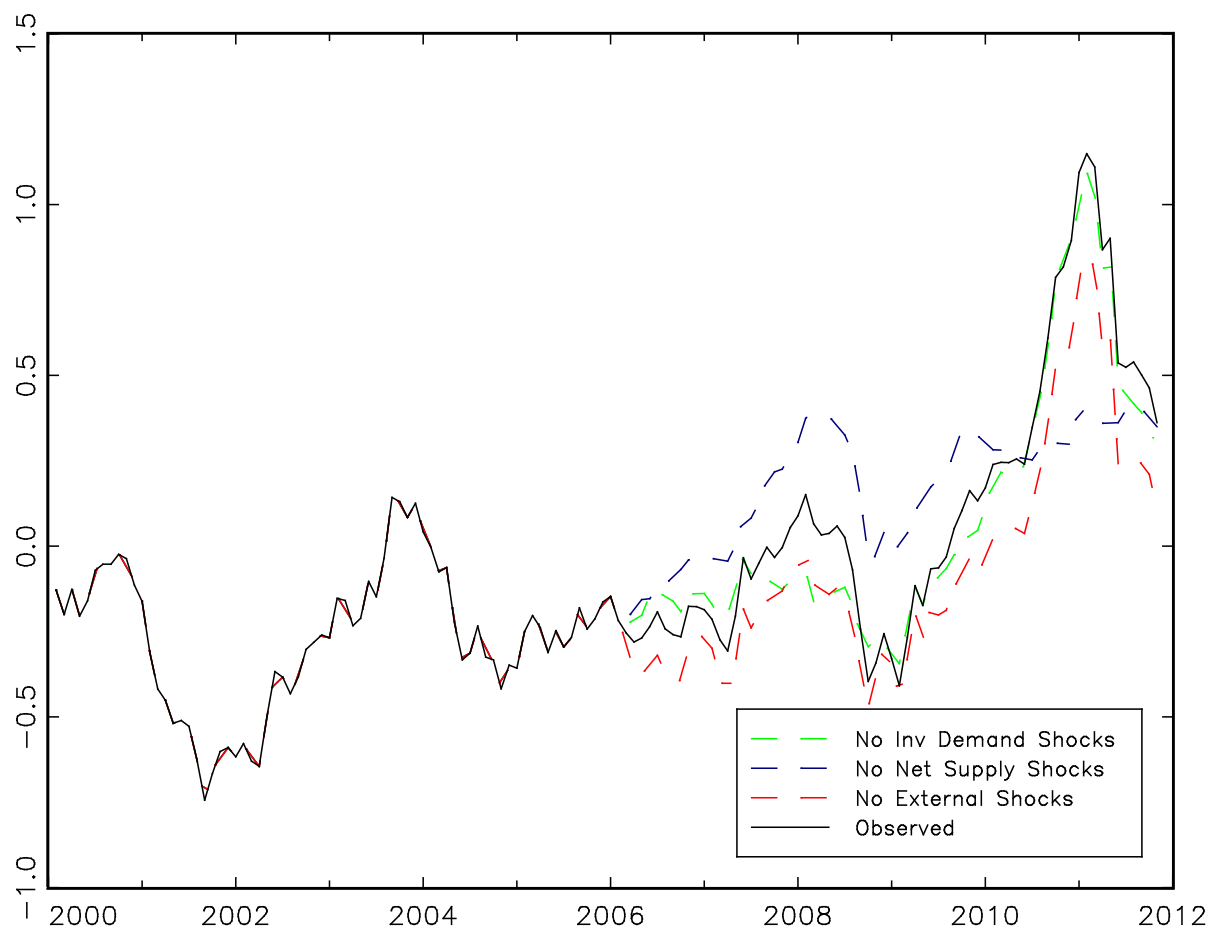


Figure 11: Counterfactual analysis for SVAR with copper as external market

## Tables

Table 1: Reduced-form VAR Regression Results

	Equation			
	REA	RPO	SPR	RPC
Intercept	-0.00054 0.00301	0.00122 0.00319	-0.00001 0.00170	0.00060 0.00270
REA(t-1)	1.16028** 0.06491	0.12170* 0.04951	0.03043 0.02549	-0.00220 0.04312
RPO(t-1)	0.08029 0.04801	1.15710** 0.07257	0.00521 0.02834	-0.00196 0.04751
SPR(t-1)	0.22797 0.13958	0.45361** 0.14245	0.84061** 0.09042	0.34338* 0.14830
RPC(t-1)	0.23048* 0.09291	0.35522** 0.10312	-0.15789** 0.05378	1.39427** 0.08993
REA(t-2)	-0.21036** 0.06460	-0.09161 0.04743	-0.02740 0.02532	0.00449 0.04314
RPO(t-2)	-0.07853 0.04767	-0.19062* 0.07115	0.00206 0.02850	-0.00094 0.04728
SPR(t-2)	-0.26925 0.14476	-0.35409* 0.14036	0.04115 0.08623	-0.27653 0.15250
RPC(t-2)	-0.25092* 0.09458	-0.29411* 0.09044	0.13683* 0.05166	-0.40805* 0.08996

Note: Heteroskedasticity-robust standard errors are in parentheses. \* and \*\* denote significance at the 5% and 1% levels.