Econometrics of Commodity Markets

Aaron Smith

UC Davis

August 21, 2014
Before we begin

- Slides and references will be posted on my website http://asmith.ucdavis.edu

- We will do two replication exercises. Data and code will be posted on my website.
Outline

1. Introduction
   ▶ start with the data

2. Models
   ▶ Structural (rational storage model estimated by ML)
   ▶ Structural Time Series (linear VAR)
   ▶ Financial Time Series (linear factor model)
   ▶ Mostly Harmless (linear regressions with carefully chosen instruments)

3. Example
   ▶ What was the effect of biofuel mandates on agricultural commodity prices?

4. Conclusion
Where to Begin?

- **You have a question**
  - What was the effect of biofuel mandates on food and fuel prices?
  - What explains the run-up in oil prices in the 2000’s?
  - How much does financial speculation affect commodity prices?
  - How much would an X% carbon tax reduce carbon emissions?
  - etc

- Economists usually start with a model
  - a model is some equations and identification assumptions

- Let’s learn from statisticians

- Start with the data!
Start with the Data

What is this series?
Crude Oil Price (U.S. Refiner Acquisition Cost)

What do you notice? Look at trend, volatility, autocorrelation.
Taking logs makes variation proportional to the level — now volatility looks relatively constant over time.
Deflating by CPI reduces trend

Autocorrelation still substantial
What is this Series?

- No trend; substantial autocorrelation; increasing volatility
- Large shocks more likely to be positive than negative
- Seasonality is significant
Natural Log of Real Natural Gas Price (Wellhead)

- No trend; substantial autocorrelation; increasing volatility
- Large shocks more likely to be positive than negative
- Seasonality is significant
Average 6-month Change in Log of Real NG Price

- Prices are higher in December than June due to heating demand
- Vertical bars are 95% confidence intervals for the average change
What is this Series?

- Large shocks more likely to be positive than negative
- Doesn’t appear to be much autocorrelation or trend
New England Average Hourly Electricity Price in 2000

- Large shocks more likely to be positive than negative
- Doesn’t appear to be much autocorrelation or trend
Numerous large positive shocks

 Doesn’t appear to be much autocorrelation or trend
Autocorrelation and trend were obscured by the noise!
Seasonality is still obscured
Seasonality: Prices higher in afternoon than morning

Clear autocorrelation/trend
What is this Series?

- Significant downward trend
- Constant volatility
- Large positive shocks more common than large negative shocks
Log of Real Corn Price in March in Central Illinois

- Significant downward trend
- Constant volatility
- Large positive shocks more common than large negative shocks
Significant seasonality

March price exceeds Sept price, at least since 1960

Why?
Which price?

- Commodities vary in **attributes** and **quality**

Log Real Price of Natural Gas
Which price?

- Several varieties of wheat are grown in the US, each with different characteristics and used for different purposes.
Which price?

- WTI crude oil prices have dropped below Brent prices in recent years due to shale oil glut
Which price?

- Commodities vary in **attributes** and **quality**
  - Partial substitutability means that prices of different varieties may be similar most of the time
  - ... but price spread **can be very large** if one variety is scarce
  - Same true for prices across different locations or delivery dates

- Commodities traded worldwide, but not typically in **auctions**
  - Retail prices are very sticky
  - Many “spot” price series compiled by surveying firms
  - Futures markets typically provide the only venue for open trade

- **So, what is “the” market price?**
  - **Answer:** There’s no such thing, so be explicit about what you are measuring
Common themes and why they matter

- **Trend** and **autocorrelation** can cause spurious inference
  - Can your variable of choice explain anything after controlling for past prices and a trend?

- **Seasonality** affects dynamics
  - Correlation between one observation and the next depends on the time of day/year

- **Volatility** can distort results
  - OLS puts most weight on large observations
  - Outliers and high-volatility periods may dominate
  - Skewness affects dynamics — prices change differently after large positive shocks than large negative shocks

- **Prices differ** by time, quality, and location
  - Use the relevant series for your research problem
We have a question and we know our data

Now, we’re ready for a model
  - But which model?

Possibilities include
  1. Structural (rational storage model estimated by ML)
  2. Structural Time Series (linear VAR)
  3. Financial Time Series (linear factor model)
  4. Mostly Harmless (linear regressions with carefully chosen instruments)
Rational Storage Model (e.g., Wright (2011))

- Take supply and demand curves as given
  - $Q_S = S(P)$ and $Q_D = D(P)$
  - Net supply: $I = g(P) \equiv S(P) - D(P)$
  - Inverse net supply: $P = g^{-1}(I)$

- A profit-maximizing competitive firm chooses inventory level based on current and expected future net supply
Representative Storage Firm

- Net supply is subject to shocks ($\varepsilon_t$). Each period, firm chooses how much inventory to hold ($I_t$) and pays storage fees ($\delta$).

- Revenue equals price times net sales
  - net sales = reduction in inventory = $I_{t-1} - I_t$

- Maximize profit:

$$\max \{ I_t \} \sum_{t=0}^{\infty} \frac{1}{(1+r)^t} E_0 \left[P_t(I_{t-1} - I_t) - \delta I_t \right]$$

subject to $I_t \geq 0$ for all $t$

- First order conditions imply:

$$P_t \geq \frac{1}{1+r} E_t [P_{t+1}] - \delta$$

- Equilibrium condition:

$$P_t = g^{-1}(I_t, \varepsilon_t)$$
Rational Storage Model Implies Kinked Demand

- Inventory mitigates price shocks — demand is elastic

\[ P_t = \frac{1}{1+r} E_t [P_{t+1}] - \delta \quad \text{if } I_t > 0 \checkmark \]

\[ P_t > \frac{1}{1+r} E_t [P_{t+1}] - \delta \quad \text{if } I_t = 0 \]
Temporary Supply Shock Depletes Inventory

- Zero inventory implies price spike and inelastic demand

\[ P_t = \frac{1}{1+r} E_t [P_{t+1}] - \delta \quad \text{if } I_t > 0 \]

\[ P_t > \frac{1}{1+r} E_t [P_{t+1}] - \delta \quad \text{if } I_t = 0 \]
Simulating Prices (Deaton and Laroque (1992))

- Large shocks more likely to be positive than negative
- Spikes occur when inventory goes to zero (i.e., market is “tight”)
What does the model imply for econometrics?

- Prices are **non-Gaussian**, even if shocks are Gaussian
- Empirical applications typically **univariate**
  - goal is to fit price dynamics
- Computationally intensive **maximum likelihood**
  - identification based on the mapping from the shocks, which have an assumed distribution, to prices through the specified demand function and storage
    - i.e., identification based on dynamics
  - \( P_{t+1} = (1 + r) (P_t + \delta) + u_{t+1} \) .... except when price is high
- Deaton and Laroque (1992) found that the model didn’t fit the data
  - Cafiero, Bobenrieth, Bobenrieth, and Wright (2011) showed that this result was based on an **incorrect** solution of the model. A correct solution provides a much better fit.
Extensions

- **supply of storage curve**
  - The “Working curve” plots the price of storage \( \delta \) as an increasing function of inventory \( I \)
  - At low inventory, the price of storage can be negative.
    - This is known as a **convenience yield** — firms are willing to hold some inventory at a loss to take advantage of potential merchandising opportunities.

- **shock dynamics**
  - model fit may be improved by adding autocorrelated shocks
    - e.g., an *iid* shock may represent weather and an autocorrelated shock may represent demand shifts

- **trend** can be added to account for population and technology growth
We have a question and we know our data

- Now, we’re ready for a model
  - But which model?

- Possibilities include
  1. Structural (rational storage model estimated by ML)
  2. **Structural Time Series** (linear VAR)
  3. Financial Time Series (linear factor model)
  4. Mostly Harmless (linear regressions with carefully chosen instruments)
Linear VAR Models

\[ X_t = \Phi_1 X_{t-1} + \Phi_2 X_{t-2} + \ldots + \Phi_p X_{t-p} + c_0 + c_1 t + \varepsilon_t \]

\[ \varepsilon_t \sim WN(0, \Omega) \]

- Multivariate system
  - \( X_t \) may include time series on production, consumption, inventory, demand and supply shifters, and futures
  - \( WN \equiv \text{“white noise”} \equiv \text{no autocorrelation} \)

- Everything is endogenous

- How to identify causal effects?
  1. “Granger causality,” which is another word for predictability — requires assumption that effects occur after causes
  2. Assumptions on meaning of \textit{contemporaneous correlations}, i.e., off-diagonal elements of \( \Omega \)
    - If \( E[\varepsilon_{1t}\varepsilon_{2t}] \neq 0 \), did \( \varepsilon_{1t} \) cause \( \varepsilon_{2t} \) or the other way around?
  3. Other. See Kilian (2013)
VAR Identification

- Strip the problem down to its most basic form
  \[ X_t = \varepsilon_t, \quad \varepsilon_t \sim \text{WN}(0, \Omega) \]

- Define structural errors \( \nu_t = A\varepsilon_t \), and write \( AX_t = \nu_t \)

- **Example:** commodity supply and demand with temperature

\[
\begin{bmatrix}
1 & \alpha_{12} & \alpha_{13} \\
\alpha_{21} & 1 & \alpha_{23} \\
\alpha_{31} & \alpha_{32} & 1
\end{bmatrix}
\begin{bmatrix}
 p_t \\
 q_t \\
w_t
\end{bmatrix} = \nu_t
\]

\[
E[\nu_t\nu'_t] =
\begin{bmatrix}
\sigma_1^2 & 0 & 0 \\
0 & \sigma_2^2 & 0 \\
0 & 0 & \sigma_3^2
\end{bmatrix}
\]

- **Identification problem:** which variable causes which?

- With 3 variables, the data give us 3 covariances, so we can identify 3 parameters
VAR Identification: Triangular system

\[ AX_t = \nu_t, \quad \nu_t \sim WN(0, \Sigma) \]

- Impose zeros in lower triangle of A
- Direction of causation: temperature → quantity → price

\[
\begin{bmatrix}
1 & \alpha_{12} & \alpha_{13} \\
0 & 1 & \alpha_{23} \\
0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
p_t \\
q_t \\
w_t
\end{bmatrix}
= \nu_t
\]

\[
E[\nu_t\nu_t'] = 
\begin{bmatrix}
\sigma_1^2 & 0 & 0 \\
0 & \sigma_2^2 & 0 \\
0 & 0 & \sigma_3^2
\end{bmatrix}
\]

- Identification assumption: Short-run supply perfectly inelastic
VAR Identification: Known Supply Elasticity

\[ AX_t = \nu_t, \quad \nu_t \sim WN(0, \Sigma) \]

- Impose a known value in A
- Could also get partial identification by specifying a range for \( \alpha_{21} \)

\[
\begin{bmatrix}
1 & \alpha_{12} & \alpha_{13} \\
0.1 & 1 & \alpha_{23} \\
0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
p_t \\
q_t \\
w_t
\end{bmatrix}
= \nu_t
\]

\[ E[\nu_t \nu_t'] =
\begin{bmatrix}
\sigma_1^2 & 0 & 0 \\
0 & \sigma_2^2 & 0 \\
0 & 0 & \sigma_3^2
\end{bmatrix}
\]

- Identification assumption: Short-run supply elasticity equals 0.1
VAR Identification: Instrumental Variables

\[ AX_t = \nu_t, \quad \nu_t \sim WN(0, \Sigma) \]

- Assume that weather does not shift demand
- Standard regression IV would allow \( \nu_{1t} \) to contain some supply shocks, i.e., \( \sigma_{12}^2 \neq 0 \)

\[
\begin{bmatrix}
1 & \alpha_{12} & 0 \\
\alpha_{21} & 1 & \alpha_{23} \\
0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
p_t \\
q_t \\
w_t
\end{bmatrix} = \nu_t
\]

\[ E[\nu_t\nu_t'] =
\begin{bmatrix}
\sigma_1^2 & \sigma_{12}^2 & 0 \\
\sigma_{12}^2 & \sigma_2^2 & 0 \\
0 & 0 & \sigma_3^2
\end{bmatrix}
\]

- System is not identified, even though \( \alpha_{12} \) is identified
VAR Identification: Instrumental Variables

\[ AX_t = \nu_t, \quad \nu_t \sim WN(0, \Sigma) \]

- Assume that weather does not shift demand
- Get system identification if assume either
  1. \( \sigma^2_{12} = 0 \) — all supply shocks come from \( w_t \), or
  2. \( \alpha_{21} = 0 \) — supply is perfectly inelastic

\[
\begin{bmatrix}
1 & \alpha_{12} & 0 \\
\alpha_{21} & 1 & \alpha_{23} \\
0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
p_t \\
q_t \\
w_t
\end{bmatrix} = \nu_t
\]

\[
E[\nu_t \nu_t'] =
\begin{bmatrix}
\sigma^2_1 & \sigma^2_{12} & 0 \\
\sigma^2_{12} & \sigma^2_2 & 0 \\
0 & 0 & \sigma^2_3
\end{bmatrix}
\]

- Or identify system by adding an independent demand shifter
A Common Difference Between IV and VAR

**IV:** Exclusion restriction, but don’t label shocks

\[
\begin{align*}
\text{demand equation} & : \begin{bmatrix} 1 & \alpha_{12} & 0 \end{bmatrix} \begin{bmatrix} p_t \\ q_t \\ w_t \end{bmatrix} = \nu_t \\
\text{supply equation} & : \begin{bmatrix} 0 & 1 & \alpha_{23} \end{bmatrix} \begin{bmatrix} p_t \\ q_t \\ w_t \end{bmatrix} = \nu_t \\
\text{temperature equation} & : \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} p_t \\ q_t \\ w_t \end{bmatrix} = \nu_t \\
\end{align*}
\]

\[
E[\nu_t \nu'_t] = \begin{bmatrix}
\sigma^2_1 & \sigma^2_{12} & 0 \\
\sigma^2_{12} & \sigma^2_2 & 0 \\
0 & 0 & \sigma^2_3
\end{bmatrix}
\]

**VAR:** No exclusion restriction, but label shocks

\[
\begin{align*}
\text{demand equation} & : \begin{bmatrix} 1 & \alpha_{12} & \alpha_{13} \end{bmatrix} \begin{bmatrix} p_t \\ q_t \\ w_t \end{bmatrix} = \nu_t \\
\text{supply equation} & : \begin{bmatrix} 0 & 1 & \alpha_{23} \end{bmatrix} \begin{bmatrix} p_t \\ q_t \\ w_t \end{bmatrix} = \nu_t \\
\text{temperature equation} & : \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} p_t \\ q_t \\ w_t \end{bmatrix} = \nu_t \\
\end{align*}
\]

\[
E[\nu_t \nu'_t] = \begin{bmatrix}
\sigma^2_1 & 0 & 0 \\
0 & \sigma^2_2 & 0 \\
0 & 0 & \sigma^2_3
\end{bmatrix}
\]
Linear VAR Models for Commodity Prices

- **Advantage:** brings information from other observables
  - estimating *univariate* rational storage models is so hard that no-one tries to bring in other variables

- **Advantage:** controls for *autocorrelation* and *trend*

- **Advantage:** identifies shocks so enables dynamic counterfactual

- **Disadvantage:** linear dynamics misspecified
  - could add nonlinearity through regime switching
  - Working curve implies linearity reasonable for $\ln(P)$ vs $\ln(I)$
We have a question and we know our data

- Now, we’re ready for a model
  - But which model?

- Possibilities include
  1. Structural (rational storage model estimated by ML)
  2. Structural Time Series (linear VAR)
  3. Financial Time Series (linear factor model)
  4. Mostly Harmless (linear regressions with carefully chosen instruments)
Financial Time Series Models (Schwartz (1997))

- **Goal:** jointly fit the distribution and dynamics of futures and spot prices

- **Model:**
  \[ dS = (\mu - \delta) Sdt + \sigma_1 Sdz_1 \]
  \[ d\delta = \kappa (\alpha - \delta) Sdt + \sigma_2 dz_2 \]

- **Discrete-time version (approximate)**
  \[ X_t = X_{t-1} + \left( \mu - \delta_t - 0.5\sigma_1^2 \right) + \sigma_1 \varepsilon_{1t} \]
  \[ \delta_t = \kappa \alpha + (1 - \kappa) \delta_{t-1} + \sigma_2 \varepsilon_{2t} \]

  where \( X_t = \ln S_t \)

- **Recall** \( P_t = \frac{1}{1+r} E_t [P_{t+1}] - \delta \)
  - price equals last period’s price plus price of storage plus shock
Financial Time Series Models (Schwartz (1997))

\[
dS = (\mu - \delta) S dt + \sigma_1 S dz_1 \\
\quad \quad d\delta = \kappa (\alpha - \delta) S dt + \sigma_2 dz_2
\]

- **Advantage:** potentially richer dynamics than VAR
  - but typically use only price data

- **Disadvantage:** no attempt at causal identification
  - typically used for derivatives pricing
We have a question and we know our data

- Now, we’re ready for a model
  - But which model?

- Possibilities include
  1. Structural (rational storage model estimated by ML)
  2. Structural Time Series (linear VAR)
  3. Financial Time Series (linear factor model)
  4. Mostly Harmless (linear regressions with carefully chosen instruments)
Mostly Harmless Method (i.e., Reduced Form)

▶ **Goal:** Regress log quantity on log price and interpret the coefficient as a supply (or demand) elasticity

\[ q_t = \alpha + \beta p_t + u_t \]

▶ **Problem:** Is \( \beta \) the supply elasticity, the demand elasticity, or some hybrid?
  ▶ Put another way, what is the source of the variation in \( p_t \)?

▶ **Solution:** instrumental variables

▶ **Advantage:** explicit treatment of causation
  ▶ but verbal arguments to justify instruments sometimes leads to sloppy thinking

▶ **Disadvantage:** no dynamics
Is $\beta$ the supply or demand elasticity?

$$q_t = \alpha + \beta p_t + u_t$$

- Answer depends on sources of price variation

- Earliest IV application was to agricultural **supply and demand**
  - Specifically, in Appendix B of a book called *The Tariff on Animal and Vegetable Oils*, written by Philip Wright

- **Weather** is a common instrument
  - In agriculture, it is plausible that growing-season weather shifts supply but not demand
    - Roberts and Schlenker (2013) note that past weather affects inventories and thereby affects the demand for current supply

- In energy, supply shocks such as refinery outages may provide plausible identification

- **Be careful:** verbal arguments to justify instruments can lead to sloppy thinking (Hendricks, Janzen, and Smith (2014))
Example: What was the effect of biofuel mandates on food commodity prices?
History of Ethanol in the U.S.

- **mid 1800s**: Internal combustion engine invented
- **1920**: USGS estimates that peak oil is imminent
  - European agriculture recovers from WWI — ag prices drop
- **1922**: Henry Ford’s autobiography
  - “There is fuel in corn; oil and fuel alcohol are obtainable from corn, and it is high time that someone was opening up this new use so that the stored-up corn crops can be moved.”
- **1920s and beyond**: Texas oil boom
  - Oil became cheap; no need for ethanol
- **1970s**: Oil prices spiked
  - 1978 Energy Policy Act - $0.40 per gallon excise-tax exemption for blending ethanol into gasoline
History of Ethanol in the U.S.

- **1983:** USDA and DOE report on feasibility of ethanol mandate
  - “though technologically attainable, is not economically feasible even under optimistic market scenarios”

- **1990:** Clean Air Act requires oxygenates added to gasoline
  - Big fight between two potential oxygenates: MTBE and Ethanol
  - MTBE (a natural gas derivative) wins

- **Early 2000s:** Ethanol FINALLY gets an opening
  - MTBE pollutes water

- **2005:** Energy Act established the Renewable Fuel Standard
  - RFS mandates a minimum volume of biofuels to be used in the national transportation fuel supply

- **2007:** Energy Independence and Security Act doubled the mandate
The 2007 RFS raised the mandate by 5.5 billion gallons per year
Research Question: What was the effect on the price of agricultural commodities of the 5.5bgal ethanol-demand shock?
What was the effect of the RFS on the price of agricultural commodities

Possible models

1. Structural (rational storage model estimated by ML)
   ▶ Bobenrieth, Wright, and Zeng (2014)

2. Structural Time Series (linear VAR)
   ▶ Carter, Rausser, and Smith (2013)
   ▶ Hausman, Auffhammer, and Berck (2012)

3. Financial Time Series (linear factor model)
   ▶ None

4. Mostly Harmless
   ▶ Roberts and Schlenker (2013)
Which price?

- **Structural.** Bobenrieth, Wright, and Zeng (2014) use a calorie-weighted average of corn (US Gulf), wheat (hard red winter, US Gulf), and rice (Thailand 5% broken, milled) prices, deflated by a composite index of prices for manufactured exports from the fifteen major developed and emerging economies to low- and middle-income economies, valued in US dollars. Phew!

- **VAR.** Carter, Rausser, and Smith (2013) use the average daily price of corn in Central IL in March, deflated by CPI

- **Mostly harmless.** Roberts and Schlenker (2013) use a calorie-weighted average of average corn, wheat, and soybean futures prices in the delivery month, deflated by CPI
Price effect depends on supply and demand elasticity

Supply equation
\[ \ln(Q_{st}) = \alpha_s + \beta_s \ln(P_t) + u_t \]

Demand equation
\[ \ln(Q_{dt}) = \alpha_d + \beta_d \ln(P_t) + v_t \]

Price effect
\[ \text{Price effect} = \frac{\Delta \ln(Q)}{\beta_s - \beta_d} \]
Roberts and Schlenker (2013)

Supply equation
\[ \ln(Q_{st}) = \alpha_s + \beta_s \ln(P_{st}) + \gamma_s \omega_t + f_s(t) + u_t \]

Demand equation
\[ \ln(Q_{dt}) = \alpha_d + \beta_d \ln(P_{dt}) + f_d(t) + v_t \]

- Use global calorie-weighted aggregate of corn, soybeans, rice, and wheat
- \( Q_{st} \) is production; \( Q_{dt} \) is use (production minus change in inventory)
- \( P_{dt} \) is November/December expiring US futures price
- \( P_{st} \equiv E_{t-1}[P_{dt}] \) is US futures price last December for delivery this year. Using futures helps with endogeneity problem.
- \( \omega_t \) is average yield shock (proxy for weather)
- \( \omega_{t-1} \) is instrument for price in supply equation
- \( f(t) \) is a flexible trend
Supply equation \[ \ln(Q_{st}) = \alpha_s + \beta_s \ln(P_{st}) + \gamma_s \omega_t + f_s(t) + u_t \]
Demand equation \[ \ln(Q_{dt}) = \alpha_d + \beta_d \ln(P_{dt}) + f_d(t) + v_t \]

- Model has no dynamics
- Because they control for flexible trend, most price variation in the model comes from short-run shocks
  - they are estimating the effects of the average shock to supply and demand, which is of short duration
- The RFS is a long-run shock, it shifts both the current demand curve and the inventory demand curve
- May under-estimate the effect
Replication of Roberts and Schlenker (2013)

Supply equation \[ \ln(Q_{st}) = \alpha_s + \beta_s \ln(P_{st}) + \gamma_s \omega_t + f_s(t) + u_t \]
Demand equation \[ \ln(Q_{dt}) = \alpha_d + \beta_d \ln(P_{dt}) + f_d(t) + v_t \]

▸ Data available from AER website

▸ We will replicate their main results in Table 1, and we will estimate the effects for corn only and for US corn only

▸ In our replication, we take the policy shock to be 5.5bgal, which translates to calories and tons of corn as below
  ▸ R&S estimate the effect of total ethanol demand (11bgal in 2009)

<table>
<thead>
<tr>
<th></th>
<th>Quantity Shock</th>
<th>% Shock</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global Calories</td>
<td>112 mil. people</td>
<td>1.6</td>
</tr>
<tr>
<td>Global Corn</td>
<td>33.4 MMT</td>
<td>4.7</td>
</tr>
<tr>
<td>US Corn</td>
<td>33.4 MMT</td>
<td>12.5</td>
</tr>
</tbody>
</table>
Results of replication of Roberts and Schlenker (2013)

<table>
<thead>
<tr>
<th></th>
<th>Quantity Shock</th>
<th>% Shock</th>
<th>Estimate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global Calories</td>
<td>112 mil. people</td>
<td>1.6</td>
<td>11.3</td>
</tr>
<tr>
<td>Global Corn</td>
<td>33.4 MMT</td>
<td>4.7</td>
<td>18.1</td>
</tr>
<tr>
<td>US Corn</td>
<td>33.4 MMT</td>
<td>12.5</td>
<td>20.5</td>
</tr>
</tbody>
</table>

- **Little difference** whether we estimate US or global corn model
  - World net demand enters US model through export demand

- **Effect on wheat, rice, and soybeans is almost half** the corn effect
  - Corn makes up about a third of calories, so zero effect on other commodities would imply 6% total effect
    - Implied effect on others is \((11.3 - 0.33 \times 18.1) / 0.67 = 8\%\)

- Futures prices **not endogenous** in supply equation
Carter, Rausser and Smith (2013)

- Model **prices** and **inventory** using a VAR

- Supply and demand
  - **Supply** of inventory comes from today’s net supply
    - e.g., bad weather reduces production and thereby reduces available inventory
  - **Demand** for inventory comes from expected future net demand

- Length of run of shocks
  - **Transitory** supply/demand shocks shift the supply of inventory
  - **Permanent** supply/demand shocks shift both the supply of and demand for inventory

- Can identify **inventory-demand** shocks from the difference between spot and futures prices (convenience yield)
Why length of run matters: base case
Transitory demand shift
Permanent demand shift

Price

Inventory

g(P)

inventory demand

Price

Quantity

D(P)

S(P)
total demand

Price

Inventory

g(P)

inventory demand
1. **Real futures price of corn**
   - Chicago futures Price in March of December futures contract
   - Harvest occurs in Sept/Oct; March is middle of crop year
   - March is before weather realizations that determine yield
   - Deflate by CPI for all items and take logs

2. **Convenience yield (negative price of storage)**
   - \( F_{t,T} = (P_t(1 + r_{t,T}) + c_{t,T})(1 - y_{t,T}) \)
   - Futures equals spot plus price of storage
   - \( r_{t,T} \) is yield on one-year Treasury notes plus 200 basis points
   - Set \( c_{t,T} \) to 5c/bu/mo in 1982-83 dollars
   - We use \( cy_t \equiv -\ln(1 - y_{t,T}) \)

3. **Crop-year ending inventory in U.S.**

4. **Index of real economic activity (REA) (Kilian, 2009)**
   - index is based on dry-cargo shipping rates and is designed to capture shifts in global demand for industrial commodities
Carter, Rausser and Smith (2013) — Detrended Data

Real Economic Activity Index

Log Inventory

Log Real Futures Price

Convenience Yield
Econometric Model

\[ AX_t = BX_{t-1} + \Gamma Z_t + U_t \]

\[ X_t = \begin{bmatrix} REA_t \\ i_t \\ f_t \\ cy_t \end{bmatrix} \quad Z'_t = \begin{bmatrix} 1 \\ t \end{bmatrix} \quad U_t = \begin{bmatrix} u_{1t} \\ u_{2t} \\ u_{3t} \\ u_{4t} \end{bmatrix} \]

\[ i_t = \alpha_{23}(f_t + cy_t) + \alpha_{21}REA_t + B'_1X_{t-1} + \Gamma'_1Z_t + u_{2t} \]

\[ f_t = -\alpha_{32}i_t + \alpha_{31}REA_t + B'_2X_{t-1} + \Gamma'_2Z_t + u_{3t} \]

\[ cy_t = -\alpha_{42}i_t + \alpha_{43}f_t + \alpha_{41}REA_t + B'_3X_{t-1} + \Gamma'_3Z_t + u_{4t} \]

inventory supply

inventory demand

supply of storage
Identification Problem: Endogenous Inventory

\[ AX_t = BX_{t-1} + \Gamma Z_t + U_t \]

\[ A = \begin{bmatrix}
  1 & 0 & 0 & 0 \\
  -\alpha_{21} & 1 & -\alpha_{23} & -\alpha_{23} \\
  -\alpha_{31} & \alpha_{32} & 1 & 0 \\
  -\alpha_{41} & \alpha_{42} & -\alpha_{43} & 1
\end{bmatrix} \quad \begin{bmatrix}
  REA_t \\
  i_t \\
  f_t \\
  cy_t
\end{bmatrix} \]

\[ i_t = \alpha_{23} (f_t + cy_t) + \alpha_{21} REA_t + B_1' X_{t-1} + \Gamma_1' Z_t + u_{2t} \]

\[ f_t = -\alpha_{32} i_t + \alpha_{31} REA_t + B_2' X_{t-1} + \Gamma_2' Z_t + u_{3t} \]

\[ cy_t = -\alpha_{42} i_t + \alpha_{43} f_t + \alpha_{41} REA_t + B_3' X_{t-1} + \Gamma_3' Z_t + u_{4t} \]

inventory supply

inventory demand

supply of storage
Problem: Inventory is Endogenous to Prices

- Short-run inventory supply difficult to identify because inventory demand varies little

- We use “set identification”
  - Rather than estimate the short-run inventory supply elasticity, we assume only that it lies in a pre-specified range

- We assume
  - Current-year demand elasticity for corn exceeds -0.1
  - Next year’s net demand is more elastic than current year net demand
  - Stocks to use ratio never exceeds 0.4

- Blanchard and Perotti (2002), Faust (1998), and Kilian and Murphy (2011) use similar approaches
Impulse Responses

REA Shock

Inv. Supply Shock

Inv. Demand Shock

Supp. Storage Shock

REA Response Inventory Response

Fut. Price Response
Historical Decomposition of Futures Prices

- Futures price decomposed into components emanating from each of the four shocks
- Real economic activity an important driver of prices since 2003
- Large inventory supply shocks in 2011 and 2012
- Large inventory demand shock in 2006
Counterfactual: No Inv. Demand Shocks from 2006-12

Log Difference Between Observed and Counterfactual

<table>
<thead>
<tr>
<th></th>
<th>2006</th>
<th>2007</th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
<th>2011</th>
<th>2012</th>
<th>Ave</th>
</tr>
</thead>
<tbody>
<tr>
<td>No ID Shocks</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No ID/IS Shocks</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prod Shocks/No ID Shocks</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Counterfactual: No Inv. Demand Shocks from 2006-12

![Graph showing log(real price) and Futures Price with two lines: one for No ID Shocks and one for Observed.]

<table>
<thead>
<tr>
<th>Year</th>
<th>2006</th>
<th>2007</th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
<th>2011</th>
<th>2012</th>
<th>Ave</th>
</tr>
</thead>
<tbody>
<tr>
<td>No ID Shocks</td>
<td>0.20</td>
<td>0.34</td>
<td>0.11</td>
<td>0.00</td>
<td>0.34</td>
<td>0.26</td>
<td>0.17</td>
<td>0.20</td>
</tr>
<tr>
<td>No ID/IS Shocks</td>
<td>0.00</td>
<td>0.34</td>
<td>0.26</td>
<td>0.17</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prod Shocks/No ID Shocks</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Carter, Rausser & Smith – Ethanol on Corn
Counterfactual: No ID or IS Shocks from 2006-12

Log Difference Between Observed and Counterfactual

<table>
<thead>
<tr>
<th></th>
<th>2006</th>
<th>2007</th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
<th>2011</th>
<th>2012</th>
<th>Ave</th>
</tr>
</thead>
<tbody>
<tr>
<td>No ID Shocks</td>
<td>.20</td>
<td>.34</td>
<td>.11</td>
<td>.00</td>
<td>.34</td>
<td>.26</td>
<td>.17</td>
<td>.20</td>
</tr>
<tr>
<td>No ID/IS Shocks</td>
<td>.28</td>
<td>.41</td>
<td>.14</td>
<td>.02</td>
<td>.58</td>
<td>.66</td>
<td>.75</td>
<td>.41</td>
</tr>
<tr>
<td>Prod Shocks/No ID Shocks</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Counterfactual: No ID or IS Shocks apart from U.S. Production Shocks

Futures Price

Log Difference Between Observed and Counterfactual

<table>
<thead>
<tr>
<th></th>
<th>2006</th>
<th>2007</th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
<th>2011</th>
<th>2012</th>
<th>Ave</th>
</tr>
</thead>
<tbody>
<tr>
<td>No ID Shocks</td>
<td>.20</td>
<td>.34</td>
<td>.11</td>
<td>.00</td>
<td>.34</td>
<td>.26</td>
<td>.17</td>
<td>.20</td>
</tr>
<tr>
<td>No ID/IS Shocks</td>
<td>.28</td>
<td>.41</td>
<td>.14</td>
<td>.02</td>
<td>.58</td>
<td>.66</td>
<td>.75</td>
<td>.41</td>
</tr>
<tr>
<td>Prod Shocks/No ID Shocks</td>
<td>.28</td>
<td>.46</td>
<td>.16</td>
<td>.13</td>
<td>.56</td>
<td>.55</td>
<td>.33</td>
<td>.35</td>
</tr>
</tbody>
</table>
What was the effect of the RFS on the price of agricultural commodities?

1. Rational storage model (Bobenrieth, Wright, and Zeng (2014))
   - Approx 50% (numbers not reported)
   - Results based on univariate estimates — no controls

2. Structural VAR (Carter, Rausser, and Smith (2013))
   - 35% effect on corn on average for 2006-12 (95% CI: 15%-57%)
   - 29% effect on corn in the long run (95% CI: 6%-80%)
   - implied effect of transitory shock is $35 - 20 = 15\%$

3. Financial Time Series (linear factor model)
   - None

4. Mostly Harmless (Roberts and Schlenker (2013))
   - 11.3% effect on calories from corn, soybeans, rice, and wheat (95% CI: 6.9%-15.7%)
   - 18.1% effect on corn (95% CI: 8.9%-27.3%)
Conclusion

- Storage matters for commodity price dynamics
  - prices spike when inventory is low
  - think about whether you want to estimate the effect of short- or long-run shocks

- Know your data

- **Trend, autocorrelation, seasonality, and volatility** matter