The Role of Storage in Commodity Markets: Indirect Inference Based on Grains Data

Christophe Gouel[1,2] Nicolas Legrand[3]

[1]INRAE, Paris-Saclay Applied Economics

[2]CEPII

[3]INRAE, SMART

Motivations

The high stakes in understanding the commodity price behavior stem from:

- The typical volatility;
- The episodic and recent price spikes;
- The income, political and social stability of countries;
- The importance for economic agents decisions;

Motivations

The high stakes in understanding the commodity price behavior stem from:

- The typical volatility;
- The episodic and recent price spikes;
- The income, political and social stability of countries;
- The importance for economic agents decisions;

$\Rightarrow$ Quantitative analyses of commodity price volatility need a consistent model explaining the commodity price formation.

The storage model

- One of the first rational expectations model (Gustafson, 1958);
- Workhorse economic model for analyzing commodity prices: simple extension of a supply/demand model accounting for speculative storage;
- Two types of demand: immediate consumption & speculative for storage;
- Able to explain the main features of commodity prices (nonlinearity, positive skewness, volatility clustering, . . .);

The storage model

- One of the first rational expectations model (Gustafson, 1958);
- Workhorse economic model for analyzing commodity prices: simple extension of a supply/demand model accounting for speculative storage;
- Two types of demand: immediate consumption & speculative for storage;
- Able to explain the main features of commodity prices (nonlinearity, positive skewness, volatility clustering, ...);
- However, not empirically validated:
  - Deaton and Laroque (1992, 1996) rejected it because of inability to match the high autocorrelation of prices;
  - Solutions to the autocorrelation puzzle in Cafiero et al. (2011, 2015) and Gouel and Legrand (2017);
  - But model estimated on prices only, so silent on its ability to match other moments.

This paper

- Reframe the empirical debate differently:
  - The point of this model is to explain jointly price **and** quantity movements (not only price dynamics);
  - Fitting prices with wrong movements in the markets fundamentals is unhelpful;
  - Integrate the storage model with methods from the modern macro literature;

This paper

- Reframe the empirical debate differently:
  - The point of this model is to explain jointly price **and** quantity movements (not only price dynamics);
  - Fitting prices with wrong movements in the markets fundamentals is unhelpful;
  - Integrate the storage model with methods from the modern macro literature;
- New approach for structurally estimating the storage model:
  - Exploit information in the joint price and quantity dynamics;
  - Indirect inference approach (Gourieroux et al., 1993; Smith, 1993);
  - Using Roberts and Schlenker's (2013) econometric model as auxiliary model.

Related Literature

- **Storage model's extensions**: Williams & Wright (1982); Chambers & Bailey (1996); Routledge et al. (2000); Osborne, (2004); Dvir & Rogoff (2014); Knittle & Pindyck (2016); Gouel (2020); Bobenrieth et al. (2021);
- **Storage model's estimations with quantities**: Roberts & Schlenker (2013); Hendricks et al. (2015); Steinwender (2018); Ghanem & Smith (2022);
- **(S)VAR models for commodity prices:** Kilian (2009); Kilian & Murphy (2014); Baumeister & Hamilton (2019); Caldara et al. (2019);
- **Indirect Inference in macro:** Rotemberg & Woodford (1997); Christiano et al. (2005); Cooper & Haltiwanger (2006); Bansal et al. (2007); Guvenen & Smith (2014); Low & Pistaferri (2015).

Main takeaways

- Richer specification of the storage model;
- New structural estimation method for the storage model;
- Full empirical test of storage theory and diagnostic of empirical failures;
- Credible solution to the long-standing puzzle of a lack of induced persistence in prices (AR1 = 0.87);

Main takeaways

- Richer specification of the storage model;
- New structural estimation method for the storage model;
- Full empirical test of storage theory and diagnostic of empirical failures;
- Credible solution to the long-standing puzzle of a lack of induced persistence in prices (AR1 = 0.87);
- New puzzle: Model unable to match the magnitude of $\text{cor}(P, D)$ and $\text{cor}(P, Q)$
- Likely misspecification on the demand side.

Outline

Introduction

Storage model

Econometric strategy

Data

Estimations

Sensitivity analysis

Applications

Conclusion

Outline

Introduction

Storage model

Econometric strategy

Data

Estimations

Sensitivity analysis

Applications

Conclusion

Storage model

Dynamic stochastic setting

- Nonlinear rational expectations model;
- Elastic supply with one-period lag and subject to 3 shocks of different timings;
- A single autocorrelated shock on demand;
- Isoelastic functional forms and multiplicative shocks to make it compatible with Roberts and Schlenker (2013);
- Trending model:
  - Trend g_q for consumption & production;
  - Trend g_p for prices;
- Model is stationarized to work with detrended variables.

Storage model

Producers

Assumptions

- Production planned in period t, realized in $t+1$;
- 3 normally distributed shocks; ▸ VCOV Intuitions

$$\begin{Bmatrix} \epsilon_t \\ \eta_t \\ \omega_t \end{Bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_\epsilon^2 & 0 & 0 \\ 0 & \sigma_\eta^2 & \rho_{\eta,\omega}\sigma_\eta\sigma_\omega \\ 0 & \rho_{\eta,\omega}\sigma_\eta\sigma_\omega & \sigma_\omega^2 \end{bmatrix} \right)$$

Harvest-time yield shock
Planting-time yield shock
Planting-time cost shock

- Marginal production cost function:

$$\gamma'(h_t) = \beta\bar{p}\left(\frac{h_t}{\bar{d}}\right)^{1/\alpha_S};$$

Producer's problem

$$\max_{h_t} \beta \, \mathrm{E}_t \left(p_{t+1} h_t \, \mathrm{e}^{\eta_t + \epsilon_{t+1}} \right) - \gamma(h_t) \, \mathrm{e}^{\omega_t}.$$

▸ FOC Details

Model's equations

Producers

$$\beta\, e^{\eta_t}\, E_t\left(p_{t+1}\, e^{\epsilon_{t+1}}\right) = \gamma^{'}\left(h_t\right) e^{\omega_t};$$

Model's equations

Producers

$$\beta\, e^{\eta_t}\, E_t\left(p_{t+1}\, e^{\epsilon_{t+1}}\right) = \gamma^{'}\left(h_t\right) e^{\omega_t};$$

Storers

$$(1-\delta)\,\beta\, e^{g_p}\, E_t\, p_{t+1} - p_t - k\bar{p} \leq 0, \quad = 0 \text{ if } x_t > 0;$$

Model's equations

Producers

$$\beta\, e^{\eta_t}\, E_t \left(p_{t+1}\, e^{\epsilon_{t+1}}\right) = \gamma' \left(h_t\right) e^{\omega_t};$$

Storers

$$(1-\delta)\, \beta\, e^{g_p}\, E_t\, p_{t+1} - p_t - k\bar{p} \leq 0, \quad = 0 \text{ if } x_t > 0;$$

Final demand

$$c_t = d\left(p_t\right) e^{\mu_t} = \bar{d}\left(p_t/\bar{p}\right)^{\alpha_D} e^{\mu_t},$$

$$\mu_t = \rho_\mu \mu_{t-1} + \upsilon_t \text{ with } \upsilon_t \sim \mathcal{N}\left(0, \sigma_\upsilon^2\right);$$

Model's equations

Producers

$$\beta\, e^{\eta_t}\, E_t\left(p_{t+1}\, e^{\epsilon_{t+1}}\right) = \gamma^{'}\left(h_t\right) e^{\omega_t};$$

Storers

$$(1-\delta)\,\beta\, e^{g_p}\, E_t\, p_{t+1} - p_t - k\bar{p} \leq 0, \quad = 0 \text{ if } x_t > 0;$$

Final demand

$$c_t = d\left(p_t\right) e^{\mu_t} = \bar{d}\left(p_t/\bar{p}\right)^{\alpha_D} e^{\mu_t},$$

$$\mu_t = \rho_\mu \mu_{t-1} + \upsilon_t \text{ with } \upsilon_t \sim \mathcal{N}\left(0, \sigma_\upsilon^2\right);$$

Market equilibrium

$$s_t = x_t + d\left(p_t\right) e^{\mu_t};$$

Model's equations

Producers

$$\beta \, e^{\eta_t} \, E_t \left(p_{t+1} \, e^{\epsilon_{t+1}} \right) = \gamma' \left(h_t \right) e^{\omega_t};$$

Storers

$$(1-\delta) \, \beta \, e^{g_p} \, E_t \, p_{t+1} - p_t - k\bar{p} \leq 0, \quad = 0 \text{ if } x_t > 0;$$

Final demand

$$c_t = d \left(p_t \right) e^{\mu_t} = \bar{d} \left(p_t / \bar{p} \right)^{\alpha_D} e^{\mu_t},$$

$$\mu_t = \rho_\mu \mu_{t-1} + \upsilon_t \text{ with } \upsilon_t \sim \mathcal{N}\left(0, \sigma_\upsilon^2\right);$$

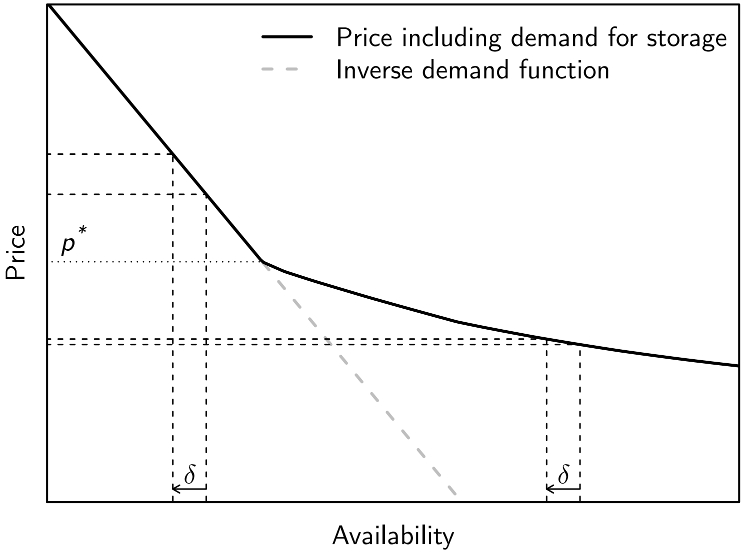
Market equilibrium

$$s_t = x_t + d \left(p_t \right) e^{\mu_t};$$

Market availability

$$s_t \equiv (1-\delta) \, x_{t-1} \, e^{-g_q} + h_{t-1} \, e^{\eta_{t-1} + \epsilon_t}.$$

Kinked market demand



Storage model

Parameters

Fixed:

- $\beta = 1/(1+r) = 1/1.02$ – Discount factor;
- $\bar{p} = \bar{d} = 1$ – Steady state values;

Estimated ex-ante:

- $g_q = 2.5\%$ – Growth rate of quantities;
- $q_p = -2\%$ – Growth rate of prices;

To estimate: 10 parameters gathered in vector θ:

- α_D, α_S – Demand and supply elasticities;
- $\rho_\mu, \rho_{\eta,\omega}$ – Shock correlations;
- $\sigma_\upsilon, \sigma_\epsilon, \sigma_\omega, \sigma_\eta$ – Shock sizes;
- k, δ – Storage costs.

Outline

Introduction

Storage model

Econometric strategy

Data

Estimations

Sensitivity analysis

Applications

Conclusion

Observed variables

- Shocks are not observable
- Observables
  - Yield shock: $\Psi_t = \exp(\eta_{t-1} + \epsilon_t) = \exp(\psi_t)$;
  - Consumption: $c_t = d(p_t)\exp(\mu_t)$;
  - Production: $q_t = h_{t-1}\exp(\eta_{t-1} + \epsilon_t)$;
  - Price: p_t;
  - Expected price: $\mathrm{E}_{t-1}\, p_t$ from futures;

Observed variables

- Shocks are not observable
- Observables
  - Yield shock: $\Psi_t = \exp(\eta_{t-1} + \epsilon_t) = \exp(\psi_t)$;
  - Consumption: $c_t = d(p_t) \exp(\mu_t)$;
  - Production: $q_t = h_{t-1} \exp(\eta_{t-1} + \epsilon_t)$;
  - Price: p_t;
  - Expected price: $\mathrm{E}_{t-1} p_t$ from futures;
- 5 observables > 4 shocks (=3 supply + 1 demand) $\Rightarrow$ misspecified model
  - Stochastic singularity $\Rightarrow$ Impossible to use a likelihood-based method.

IV approach

Based on Roberts and Schlenker (2013)

Demand

Log consumption demand:

$$\ln c_t = \ln\left(\bar{d}/\bar{p}^{\alpha_D}\right) + \alpha_D \ln p_t + \mu_t,$$

IV approach

Based on Roberts and Schlenker (2013)

Demand

Log consumption demand:

$$\ln c_t = \ln\left(\bar{d}/\bar{p}^{\alpha_D}\right) + \alpha_D \ln p_t + \mu_t,$$

With $\mu_t = \rho_\mu \mu_{t-1} + \upsilon_t$, this gives

Estimating equation

$$\begin{aligned} \ln c_t = {} & (1-\rho_\mu)\ln\left(\bar{d}/\bar{p}^{\alpha_D}\right) + \alpha_D \ln p_t \\ & - \alpha_D \rho_\mu \ln p_{t-1} + \rho_\mu \ln c_{t-1} + \upsilon_t; \end{aligned}$$

IV approach

Based on Roberts and Schlenker (2013)

Demand

Log consumption demand:

$$\ln c_t = \ln\left(\bar{d}/\bar{p}^{\alpha_D}\right) + \alpha_D \ln p_t + \mu_t,$$

With $\mu_t = \rho_\mu \mu_{t-1} + \upsilon_t$, this gives

Estimating equation

$$\ln c_t = (1 - \rho_\mu) \ln\left(\bar{d}/\bar{p}^{\alpha_D}\right) + \alpha_D \ln p_t - \alpha_D \rho_\mu \ln p_{t-1} + \rho_\mu \ln c_{t-1} + \upsilon_t;$$

OV bias because $E(\upsilon_t | p_t) \neq 0$, so estimated by 2SLS using ψ_t as instrument for $\ln p_t$

IV approach

Based on Roberts and Schlenker (2013)

Demand

Log consumption demand:

$$\ln c_t = \ln\left(\bar{d}/\bar{p}^{\alpha_D}\right) + \alpha_D \ln p_t + \mu_t,$$

With $\mu_t = \rho_\mu \mu_{t-1} + \upsilon_t$, this gives

Estimating equation

$$\ln c_t = (1 - \rho_\mu) \ln\left(\bar{d}/\bar{p}^{\alpha_D}\right) + \alpha_D \ln p_t - \alpha_D \rho_\mu \ln p_{t-1} + \rho_\mu \ln c_{t-1} + \upsilon_t;$$

OV bias because $E(\upsilon_t | p_t) \neq 0$, so estimated by 2SLS using ψ_t as instrument for $\ln p_t$

Supply

Log production:

$$\ln q_t = \ln\left(\bar{d}/\bar{p}^{\alpha_S}\right) + \alpha_S (\eta_{t-1} - \omega_{t-1}) + \alpha_S \ln\left(E_{t-1}\left(p_t\, e^{\epsilon_t}\right)\right) + \psi_t$$

IV approach

Based on Roberts and Schlenker (2013)

Demand

Log consumption demand:

$$\ln c_t = \ln\left(\bar{d}/\bar{p}^{\alpha_D}\right) + \alpha_D \ln p_t + \mu_t,$$

With $\mu_t = \rho_\mu \mu_{t-1} + \upsilon_t$, this gives

Estimating equation

$$\ln c_t = (1-\rho_\mu)\ln\left(\bar{d}/\bar{p}^{\alpha_D}\right) + \alpha_D \ln p_t - \alpha_D \rho_\mu \ln p_{t-1} + \rho_\mu \ln c_{t-1} + \upsilon_t;$$

OV bias because $E(\upsilon_t | p_t) \neq 0$, so estimated by 2SLS using ψ_t as instrument for $\ln p_t$

Supply

Log production:

$$\ln q_t = \ln\left(\bar{d}/\bar{p}^{\alpha_S}\right) + \alpha_S\left(\eta_{t-1} - \omega_{t-1}\right) + \alpha_S \ln\left(E_{t-1}\left(p_t\, e^{\epsilon_t}\right)\right) + \psi_t$$

Estimating equation

$$\ln q_t = a_q + b_q \ln\left(E_{t-1}\, p_t\right) + c_q \psi_t + u_{q,t}$$

IV approach

Based on Roberts and Schlenker (2013)

Demand

Log consumption demand:

$$\ln c_t = \ln\left(\bar{d}/\bar{p}^{\alpha_D}\right) + \alpha_D \ln p_t + \mu_t,$$

With $\mu_t = \rho_\mu \mu_{t-1} + \upsilon_t$, this gives

Estimating equation

$$\ln c_t = (1 - \rho_\mu) \ln\left(\bar{d}/\bar{p}^{\alpha_D}\right) + \alpha_D \ln p_t - \alpha_D \rho_\mu \ln p_{t-1} + \rho_\mu \ln c_{t-1} + \upsilon_t;$$

OV bias because $E(\upsilon_t | p_t) \neq 0$, so estimated by 2SLS using ψ_t as instrument for $\ln p_t$

Supply

Log production:

$$\ln q_t = \ln\left(\bar{d}/\bar{p}^{\alpha_S}\right) + \alpha_S (\eta_{t-1} - \omega_{t-1}) + \alpha_S \ln\left(E_{t-1}\left(p_t\, e^{\epsilon_t}\right)\right) + \psi_t$$

Estimating equation

$$\ln q_t = a_q + b_q \ln\left(E_{t-1}\, p_t\right) + c_q \psi_t + u_{q,t}$$

OV bias because $E(\eta_{t-1} - \omega_{t-1} | E_{t-1}\, p_t) \neq 0$, so estimated by 2SLS using ψ_{t-1} as instrument for $\ln E_{t-1}\, p_t$.

IV approach

Based on Roberts and Schlenker (2013)

Demand

Log consumption demand:

$$\ln c_t = \ln\left(\bar{d}/\bar{p}^{\alpha_D}\right) + \alpha_D \ln p_t + \mu_t,$$

With $\mu_t = \rho_\mu \mu_{t-1} + \upsilon_t$, this gives

Estimating equation

$$\ln c_t = (1-\rho_\mu)\ln\left(\bar{d}/\bar{p}^{\alpha_D}\right) + \alpha_D \ln p_t - \alpha_D \rho_\mu \ln p_{t-1} + \rho_\mu \ln c_{t-1} + \upsilon_t;$$

OV bias because $E(\upsilon_t | p_t) \neq 0$, so estimated by 2SLS using ψ_t as instrument for $\ln p_t$

$\Rightarrow \alpha_D$, ρ_μ, and σ_υ

Supply

Log production:

$$\ln q_t = \ln\left(\bar{d}/\bar{p}^{\alpha_S}\right) + \alpha_S(\eta_{t-1} - \omega_{t-1}) + \alpha_S \ln\left(E_{t-1}\left(p_t e^{\epsilon_t}\right)\right) + \psi_t$$

Estimating equation

$$\ln q_t = a_q + b_q \ln\left(E_{t-1} p_t\right) + c_q \psi_t + u_{q,t}$$

OV bias because $E(\eta_{t-1} - \omega_{t-1} | E_{t-1} p_t) \neq 0$, so estimated by 2SLS using ψ_{t-1} as instrument for $\ln E_{t-1} p_t$.

$\Rightarrow \alpha_S$

Indirect inference

Principles

- Simulation based estimation methods;
- Useful for models with intractable likelihood;
- Minimize a criterion based on the distance of an auxiliary model (AM) estimated on observations and on simulations;

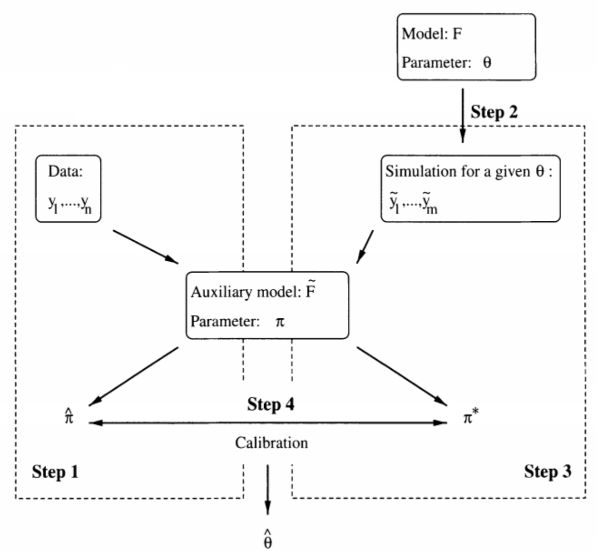
Indirect inference

Principles

- Simulation based estimation methods;
- Useful for models with intractable likelihood;
- Minimize a criterion based on the distance of an auxiliary model (AM) estimated on observations and on simulations;
- AM selects (important) aspects of the data on which to focus the analysis;
- AM can be seen as a window from which observed and simulated data can be analyzed $\Rightarrow$ Need not be an accurate description of the true DGP.

Indirect inference

Practical steps



Indirect inference

Auxiliary model

- Preferred specification based on OLS version of the supply/demand model augmented by 1st-stage IV equations:

$$\log q_t = a_q^{OLS} + b_q^{OLS} \log(\mathrm{E}_{t-1}\, p_t) + c_q^{OLS}\psi_t + u_{q,t}^{OLS},$$
$$\log(\mathrm{E}_{t-1}\, p_t) = a_{\mathrm{E}\,p} + b_{\mathrm{E}\,p}\psi_{t-1} + c_{\mathrm{E}\,p}\psi_t + u_{\mathrm{E}\,p,t},$$
$$\log c_t = a_c^{OLS} + b_c^{OLS} \log p_t + c_c^{OLS} \log p_{t-1} + d_c^{OLS} \log c_{t-1} + u_{c,t}^{OLS},$$
$$\log p_t = a_P + b_p\psi_t + c_p \log p_{t-1} + d_p \log c_{t-1} + u_{p,t},$$
$$\psi_t = a_\psi + u_{\psi,t};$$

- 15 target parameters:
$\zeta = [b_q^{OLS}, c_q^{OLS}, \sigma_{u_q^{OLS}}, b_{\mathrm{E}\,p}, c_{\mathrm{E}\,p}, \sigma_{u_{\mathrm{E}\,p}}, b_c^{OLS}, c_c^{OLS}, d_c^{OLS}, \sigma_{u_c^{OLS}}, b_p, c_p, d_p, \sigma_{u_p}, \sigma_{u_\psi}]$. 2SLS auxiliary model

Indirect inference

Objective function

$$\hat{\theta} = \arg\min_{\theta \in \Theta} \left[\hat{\zeta}_T - \frac{1}{\tau} \sum_{i=1}^{\tau} \hat{\zeta}_T^i(\theta) \right]' W \left[\hat{\zeta}_T - \frac{1}{\tau} \sum_{i=1}^{\tau} \hat{\zeta}_T^i(\theta) \right],$$

- $\hat{\zeta}_T$: estimated on T observations;
- $\hat{\zeta}_T^i(\theta)$: estimated on T simulations from the storage model with parameters θ;
- $\tau = 200$;
- W: diagonal weighting matrix $= \text{diag}(V_{\hat{\zeta}_T}^{-1})$. ▸ Details

Indirect inference

Objective function

$$\hat{\theta} = \arg\min_{\theta \in \Theta} \left[\hat{\zeta}_T - \frac{1}{\tau} \sum_{i=1}^{\tau} \hat{\zeta}_T^i(\theta) \right]' W \left[\hat{\zeta}_T - \frac{1}{\tau} \sum_{i=1}^{\tau} \hat{\zeta}_T^i(\theta) \right],$$

- $\hat{\zeta}_T$: estimated on T observations;
- $\hat{\zeta}_T^i(\theta)$: estimated on T simulations from the storage model with parameters θ;
- $\tau = 200$;
- W: diagonal weighting matrix $= \text{diag}(V_{\hat{\zeta}_T}^{-1})$. ▸ Details

Variance-covariance matrix of estimates:

$$V_{\hat{\theta}} = \left(1 + \frac{1}{\tau}\right) \left(\left\{ \frac{1}{\tau} \sum_{i=1}^{\tau} \mathrm{E} \left[\frac{\partial \hat{\zeta}_T^i(\theta)}{\partial \theta} \right] \right\}' W \left\{ \frac{1}{\tau} \sum_{i=1}^{\tau} \mathrm{E} \left[\frac{\partial \hat{\zeta}_T^i(\theta)}{\partial \theta} \right] \right\} \right)^{-1}.$$

MC analysis findings

- MC based on 500 replications of actual sample size $T = 56$ and longer sizes;
- Indirect inference more precise than either the OLS or 2SLS (RMSEs);
- α_D & α_S with II unbiased in small or large samples unlike OLS estimates;
- k, δ and $\rho_{\eta,\omega}$ not precisely estimated (but unbiased);
- Both bias and volatility vanish as T goes large $\Rightarrow$ consistency;
- α_S, σ_υ, and σ_ω estimated with greater precision in the OLS-based II.

MC tables

Outline

Introduction

Storage model

Econometric strategy

Data

Estimations

Sensitivity analysis

Applications

Conclusion

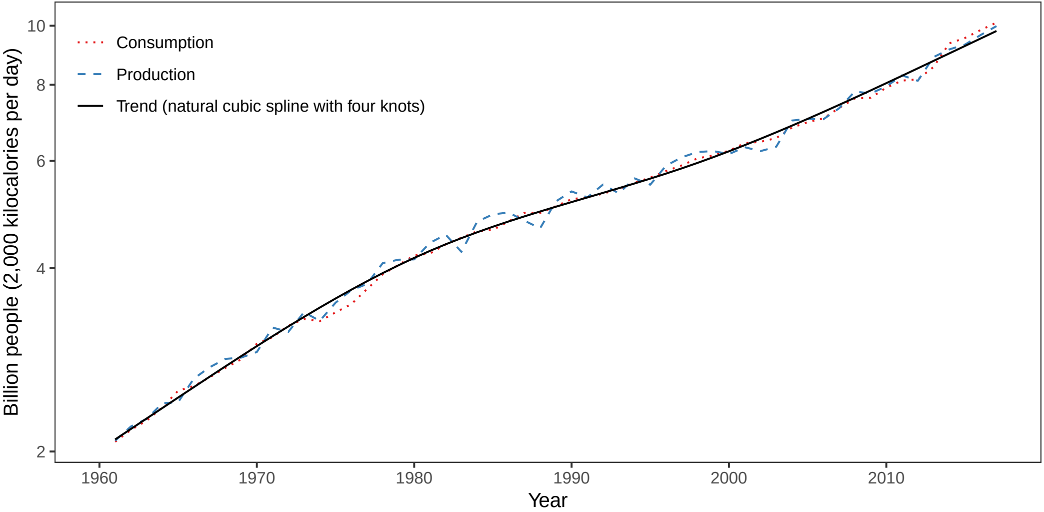
Data

Same data construction as in Roberts and Schlenker (2013)

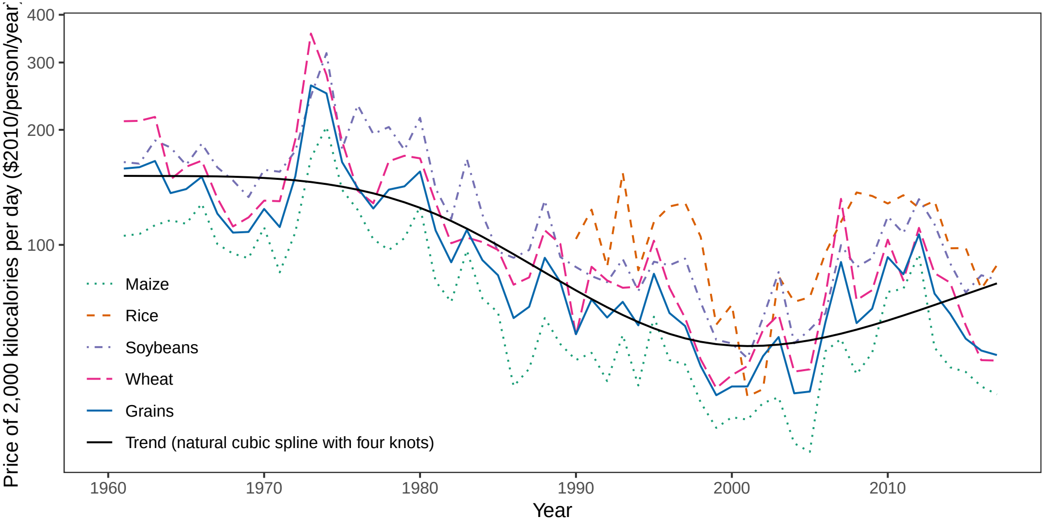
- 5 observables: Price, Expected price, Production, Consumption, and yield shock;
- 1 commodity: a caloric aggregate of maize, rice, soybeans, and wheat;
- Annual frequency from 1961 to 2017;
- Production, consumption, and yield shock from FAOSTAT;
- Prices from CBOT for futures contract at harvest (at delivery for P_t and 1-year before for $\mathrm{E}_{t-1} P_t$), deflated by US CPI;
  - Rice excluded because futures prices not available before 1986;
- Detrending: multiplicative natural cubic splines with 4 knots for quantities and prices ($\neq$ # of knots tested in the paper).

More

Caloric Production & Consumption



Caloric Prices



Descriptive statistics on detrended caloric data

Variables	1-year AC	2-year AC	CV
Demand price ($\log(p_t)$)	0.576	0.167	0.236
Supply price ($\log(\mathrm{E}_t\, p_{t+1})$)	0.652	0.236	0.192
Consumption ($\log(c_t)$)	0.642	0.302	0.019
Production ($\log(q_t)$)	0.042	-0.095	0.028
Yield shock (ψ_t)	0.148	0.050	0.023

- Production more volatile than consumption
  $\Rightarrow$ Smoothing effect of inter-annual storage;

Descriptive statistics on detrended caloric data

Variables	1-year AC	2-year AC	CV
Demand price ($\log(p_t)$)	0.576	0.167	0.236
Supply price ($\log(\mathrm{E}_t\, p_{t+1})$)	0.652	0.236	0.192
Consumption ($\log(c_t)$)	0.642	0.302	0.019
Production ($\log(q_t)$)	0.042	-0.095	0.028
Yield shock (ψ_t)	0.148	0.050	0.023

- Production more volatile than consumption
  $\Rightarrow$ Smoothing effect of inter-annual storage;
- Prices an order of magnitude more volatile than quantities
  $\Rightarrow$ Inelastic demand and supply;

Descriptive statistics on detrended caloric data

Variables	1-year AC	2-year AC	CV
Demand price ($\log(p_t)$)	0.576	0.167	0.236
Supply price ($\log(E_t\, p_{t+1})$)	0.652	0.236	0.192
Consumption ($\log(c_t)$)	0.642	0.302	0.019
Production ($\log(q_t)$)	0.042	-0.095	0.028
Yield shock (ψ_t)	0.148	0.050	0.023

- Production more volatile than consumption
  $\Rightarrow$ Smoothing effect of inter-annual storage;
- Prices an order of magnitude more volatile than quantities
  $\Rightarrow$ Inelastic demand and supply;
- Significant persistence in consumption (unlike production and yield)
  $\Rightarrow$ Persistent demand shocks.

Correlations on detrended caloric data

Variable	$\log(p_t)$	$\log(\mathrm{E}_t\, p_{t+1})$	$\log(c_t)$	$\log(q_t)$	ψ_t
Demand price ($\log(p_t)$)					
Supply price ($\log(\mathrm{E}_t\, p_{t+1})$)	0.935				
Consumption ($\log(c_t)$)	−0.488	−0.451			
Production ($\log(q_t)$)	−0.406	−0.270	0.395		
Yield shock (ψ_t)	−0.532	−0.498	0.527	0.775	

- High $\mathrm{cor}(\log p_t, \log \mathrm{E}_t\, p_{t+1})$ consistent with frequent storage arbitrage;
- $\mathrm{cor}(\log c_t, \log q_t) \neq 1 \Rightarrow$ Role of stocks;
- $\mathrm{cor}(\log c_t, \log p_t) \neq -1 \Rightarrow$ Role of demand shocks.

Price correlations of detrended caloric data

Commodity	Maize	Rice	Soybeans	Wheat	Grains
Maize					
Rice	0.662				
Soybeans	0.858	0.772			
Wheat	0.790	0.611	0.776		
Grains	0.923	0.688	0.887	0.959	

- Prices highly correlated;
  $\Rightarrow$ Hard to separate own-price from cross-price elasticities;

- High levels of correlation between the index of grains calories and individual crops;
  $\Rightarrow$ An aggregated indicator is a suitable measure of the state of the world food market.

Outline

Introduction

Storage model

Econometric strategy

Data

Estimations

Sensitivity analysis

Applications

Conclusion

Estimation results

Auxiliary parameters fit

	OLS		2SLS		Indirect inference	
Parameters	Estimate	Std. Errors	Estimate	Std. Errors	Estimate	Std. Errors
ρ_μ	0.547	(0.106)	0.535	(0.159)	0.702	(0.068)
$\rho_{\eta,\omega}$					-0.442	(0.307)
σ_ω					0.188	(0.031)
σ_η					0.014	(0.006)
σ_ϵ					0.020	(0.005)
σ_υ	0.014	(0.001)	0.016	(0.004)	0.019	(0.003)
δ					0	
k					0.037	(0.014)
α_D	-0.021	(0.010)	-0.065	(0.026)	-0.068	(0.019)
α_S	0.059	(0.011)	0.075	(0.026)	0.086	(0.016)
σ_φ					0.027	(0.005)
σ_ψ	0.023	(0.002)	0.023	(0.002)	0.025	(0.002)
σ_μ	0.016	(0.002)	0.019	(0.005)	0.026	(0.005)
σ_ϑ	0.029		0.031		0.034	(0.004)

Moments fit

Moment	Observed	Standard deviation	Simulated
$\sigma_{\ln p}$	0.236	0.023	0.262
$\sigma_{\ln c}$	0.019	0.002	0.018
$\sigma_{\ln q}$	0.028	0.002	0.031
$\sigma_{\ln \mathrm{E}\, p}$	0.193	0.018	0.180
σ_{ψ}	0.024	0.002	0.025
$\phi_{\ln p}(1)$	0.576	0.110	0.559
$\phi_{\ln c}(1)$	0.642	0.146	0.568
$\phi_{\ln q}(1)$	0.042	0.140	−0.011
$\phi_{\ln \mathrm{E}\, p}(1)$	0.652	0.116	0.607
$\phi_{\psi}(1)$	0.146	0.142	0.001

Moments fit

Moment	Observed	Standard deviation	Simulated
$\phi_{\ln p, \ln c}(0)$	−0.488	0.102	0.083***
$\phi_{\ln p, \ln q}(0)$	−0.406	0.103	−0.183**
$\phi_{\ln p, \ln \mathrm{E}\, p}(0)$	0.939	0.017	0.871***
$\phi_{\ln p, \psi}(0)$	−0.534	0.118	−0.454
$\phi_{\ln c, \ln q}(0)$	0.395	0.109	0.590*
$\phi_{\ln c, \ln \mathrm{E}\, p}(0)$	−0.452	0.106	0.283***
$\phi_{\ln c, \ln \psi}(0)$	0.529	0.116	0.463
$\phi_{\ln q, \ln \mathrm{E}\, p}(0)$	−0.271	0.115	−0.025**
$\phi_{\ln q, \psi}(0)$	0.775	0.050	0.831
$\phi_{\ln \mathrm{E}\, p, \psi}(0)$	−0.500	0.118	−0.292

Moments fit

Moment	Observed	Standard deviation	Simulated
$\phi_{\ln p,\ln c}(1)$	−0.469	0.125	0.191***
$\phi_{\ln p,\ln q}(1)$	0.104	0.156	−0.015
$\phi_{\ln p,\ln E p}(1)$	0.643	0.069	0.627
$\phi_{\ln p,\psi}(1)$	−0.274	0.142	−0.183
$\phi_{\ln c,\ln p}(1)$	−0.326	0.109	0.205***
$\phi_{\ln c,\ln q}(1)$	0.184	0.110	0.299
$\phi_{\ln c,\ln E p}(1)$	−0.300	0.118	0.181***
$\phi_{\ln c,\psi}(1)$	0.304	0.127	0.187
$\phi_{\ln q,\ln p}(1)$	−0.257	0.110	0.216***
$\phi_{\ln q,\ln c}(1)$	0.323	0.110	0.352
$\phi_{\ln q,\ln E p}(1)$	−0.212	0.116	0.092**
$\phi_{\ln q,\psi}(1)$	0.067	0.134	−0.143*
$\phi_{\ln E p,\ln p}(1)$	0.566	0.094	0.534
$\phi_{\ln E p,\ln c}(1)$	−0.508	0.116	0.293***
$\phi_{\ln E p,\ln q}(1)$	0.070	0.147	0.043
$\phi_{\ln E p,\psi}(1)$	−0.358	0.129	−0.138*
$\phi_{\ln \psi,\ln p}(1)$	−0.162	0.108	−0.120
$\phi_{\ln \psi,\ln c}(1)$	0.334	0.127	0.123
$\phi_{\ln \psi,\ln q}(1)$	−0.115	0.122	0.002
$\phi_{\ln \psi,\ln E p}(1)$	−0.203	0.115	−0.226

Additional estimation results

	FAOSTAT data		FAOSTAT data without rice				USDA-PSD data			
	II – Only shrinkage ($k = 0$)		2SLS		II		2SLS		II	
	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE
ρ_μ	0.714	(0.067)	0.530	(0.158)	0.674	(0.074)	0.533	(0.226)	0.738	(0.059)
$\rho_{\eta,\omega}$	−0.450	(0.315)			−0.396	(0.287)			−0.437	(0.341)
σ_ω	0.189	(0.032)			0.219	(0.042)			0.177	(0.026)
σ_η	0.015	(0.006)			0.021	(0.006)			0.013	(0.007)
σ_ϵ	0.020	(0.005)			0.024	(0.006)			0.020	(0.005)
σ_υ	0.018	(0.003)	0.021	(0.005)	0.024	(0.004)	0.018	(0.004)	0.022	(0.004)
δ	0.038	(0.013)			0				0	
k	0				0.032	(0.012)			0.038	(0.015)
α_D	−0.064	(0.018)	−0.083	(0.035)	−0.087	(0.026)	−0.076	(0.033)	−0.089	(0.025)
α_S	0.086	(0.016)	0.088	(0.035)	0.096	(0.020)	0.086	(0.026)	0.079	(0.012)
σ_φ	0.027	(0.005)			0.037	(0.006)			0.024	(0.005)
σ_ψ	0.025	(0.002)	0.030	(0.003)	0.032	(0.003)	0.023	(0.002)	0.024	(0.002)
σ_μ	0.026	(0.005)	0.025	(0.006)	0.033	(0.007)	0.022	(0.006)	0.033	(0.006)
σ_ϑ	0.034	(0.004)	0.040		0.044	(0.005)	0.031		0.031	(0.003)

Estimation results by commodity

	Maize				Soybeans				Wheat			
	2SLS		II		2SLS		II		2SLS		II	
	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE
ρ_μ	0.501	(0.192)	0.735	(0.062)	0.482	(0.195)	0.565	(0.102)	0.605	(0.197)	0.628	(0.085)
$\rho_{\eta,\omega}$			−0.960	(1.641)			0.201	(0.233)			−0.133	(0.245)
σ_ω			0.170	(0.028)			0.362	(0.098)			0.473	(0.160)
σ_η			0.013	(0.018)			0.025	(0.015)			0.035	(0.007)
σ_ϵ			0.039	(0.007)			0.040	(0.010)			0.024	(0.009)
σ_υ	0.028	(0.005)	0.034	(0.005)	0.048	(0.013)	0.059	(0.017)	0.029	(0.010)	0.037	(0.010)
δ			0				0				0	
k			0.048	(0.016)			0.018	(0.018)			0.060	(0.028)
α_D	−0.110	(0.031)	−0.131	(0.033)	−0.090	(0.111)	−0.168	(0.118)	−0.096	(0.074)	−0.126	(0.048)
α_S	0.162	(0.057)	0.165	(0.032)	0.226	(0.181)	0.170	(0.054)	0.060	(0.052)	0.064	(0.024)
σ_φ			0.043	(0.010)			0.063	(0.011)			0.051	(0.008)
σ_ψ	0.041	(0.004)	0.042	(0.004)	0.047	(0.004)	0.047	(0.004)	0.040	(0.004)	0.042	(0.004)
σ_μ	0.033	(0.007)	0.051	(0.010)	0.055	(0.016)	0.071	(0.025)	0.037	(0.014)	0.047	(0.015)
σ_ϑ	0.057		0.058	(0.008)	0.073		0.074	(0.008)	0.047		0.056	(0.007)

Model fit inspection

- Reminder: 40 moments tested now;
- Overall fit is good including price persistence but

Model fit inspection

- Reminder: 40 moments tested now;
- Overall fit is good including price persistence but
- Unmatched negative level of price-demand and price-production covariances;
  - $\text{cor}(P, D)$ and $\text{cor}(P, Q)$ increase with the size of demand shocks, so possibly related to too large estimated volatility of consumption;
    $\Rightarrow$ Demand side misspecifications.

Outline

Introduction

Storage model

Econometric strategy

Data

Estimations

Sensitivity analysis

Applications

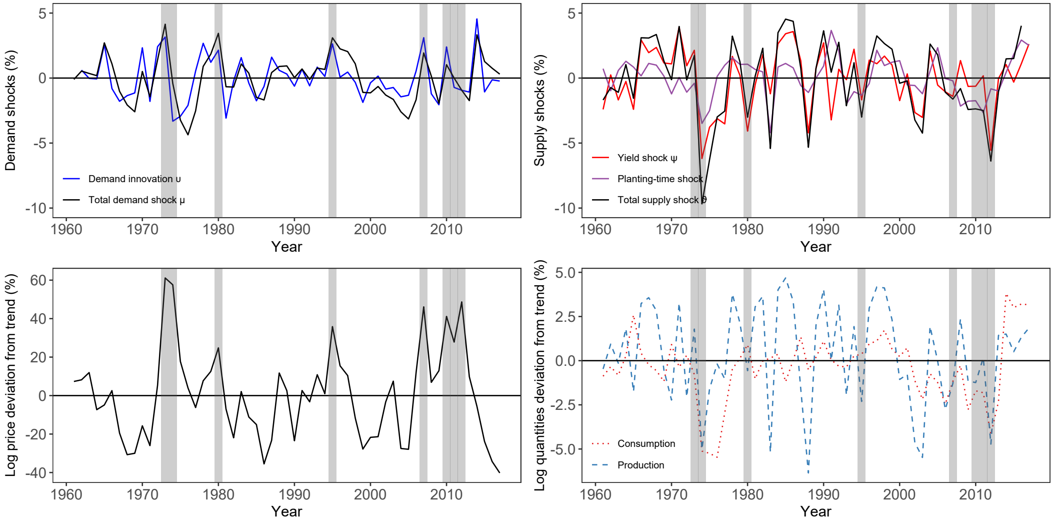
Conclusion

Model features and market dynamics

Data or model	$\phi_{\ln p}(1)$	$\sigma_{\ln p}$	$\sigma_{\ln c}$	$\sigma_{\ln q}$	$\phi_{\ln p, \ln c}(0)$	$\phi_{\ln p, \ln q}(0)$
Trending data	0.87	0.46	–	–	–	–
Detrended data	0.58	0.24	0.019	0.028	-0.49	-0.41
1. Benchmark	0.56	0.26	0.018	0.031	0.09	-0.18
2. $\rho_\mu = 0$	0.38	0.21	0.017	0.029	-0.33	-0.50
3. $\rho_\mu = 0, \sigma_\upsilon = \sigma_\mu$	0.38	0.23	0.022	0.030	-0.04	-0.38
4. $\alpha_S = 0$	0.65	0.30	0.014	0.024	0.12	-0.16
5. $g_q = 0$	0.56	0.26	0.017	0.031	0.08	-0.17
6. $g_p = 0$	0.60	0.24	0.018	0.032	0.19	-0.14
7. $k = 0.018$	0.60	0.24	0.018	0.032	0.19	-0.14
8. $\sigma_\eta = 0$	0.53	0.25	0.017	0.027	0.19	-0.12
9. $\sigma_\omega = 0$	0.54	0.25	0.016	0.026	0.20	-0.11
10. $\sigma_\eta = 0, \sigma_\epsilon = \sigma_\psi$	0.52	0.26	0.017	0.031	0.09	-0.19
11. $\sigma_\omega = \sigma_\eta = 0, \sigma_\epsilon = \sigma_\psi$	0.51	0.26	0.017	0.028	0.15	-0.16
12. $\rho_\mu = 0, \sigma_\upsilon = \sigma_\mu, \alpha_S = 0$	0.24	0.20	0.022	0.027	0.06	-0.36
13. $\rho_\mu = 0, \sigma_\upsilon = \sigma_\mu, \alpha_S = 0, \sigma_\eta = 0, \sigma_\epsilon = \sigma_\psi, g_q = 0$	0.25	0.20	0.022	0.028	0.00	-0.40
14. $\rho_\mu = 0.535, \sigma_\upsilon = 0.016$	0.47	0.23	0.015	0.030	-0.25	-0.39
15. $k = \infty$	0.16	0.45	0.025	0.025	-0.58	-0.58

Historical Decomposition

Grey areas denote price spikes: mean deviation > 23.6% (or 1 SD)



Outline

Introduction

Storage model

Econometric strategy

Data

Estimations

Sensitivity analysis

Applications

Conclusion

Conclusion

- First full structural estimation of a rational expectations storage model on prices and quantities:
  - Richer storage model;
  - New estimation method;
  - Estimation of all parameters;
- More credible solution to the price autocorrelation puzzle, because of the constraints imposed by the non-price moments;
- Highlight dimensions of the data left unexplained and directions for further model improvements.

Extensions

- New puzzle: model unable to reproduce $\text{cor}(\ln p, \ln c)$ and $\text{cor}(\ln p, \ln q)$;
  - As long as this misspecification is not solved, estimation by full-information technique may be problematic;
  - Some possible solutions
    - Measurement errors
    - Shocks on storage costs as in Knittel and Pindyck (2016);
- Approach applicable to other commodities as long as an observable shock is available as instrument (e.g., the aggregate demand shock of Kilian, 2009).

Thank you for your attention

Shocks assumptions

Intuitions & examples

- Planting time shocks
  - η: soybeans rust, crop news from southern hemisphere, seasonal weather forecasts, groundwater level, …
  - ω: variety of shocks to production costs from labor and fuel to fertilizers;
- $\psi_{t+1} = \eta_t + \epsilon_{t+1}$: observable yield shock;
- $\epsilon_{t+1} \perp \eta_t$ due to RE assumption;
  $\Rightarrow$ can be seen as a yield forecast error at planting time
- $\rho_{\eta,\omega}$ can result from production decisions in time of low yield prospect (for e.g., sowing density).

Back

Storage model

Producers

FOC:

$$\beta \, e^{\eta_t} \, E_t \left(p_{t+1} \, e^{\epsilon_{t+1}} \right) = \gamma' \left(h_t \right) e^{\omega_t} \, .$$

Substituting the marginal cost function:

$$\frac{h_t}{\bar{d}} = \left[e^{\eta_t - \omega_t} \, E_t \left(\frac{p_{t+1}}{\bar{p}} \, e^{\epsilon_{t+1}} \right) \right]^{\alpha_S} \, .$$

Storage model

Producers

FOC:

$$\beta\, e^{\eta_t}\, E_t\left(p_{t+1}\, e^{\epsilon_{t+1}}\right) = \gamma^{'}\left(h_t\right) e^{\omega_t}\,.$$

Substituting the marginal cost function:

$$\frac{h_t}{\bar{d}} = \left[e^{\eta_t-\omega_t}\, E_t\left(\frac{p_{t+1}}{\bar{p}}\, e^{\epsilon_{t+1}}\right)\right]^{\alpha_S}\,.$$

Final production = acreage $\times$ yield shock ($q_{t+1} = h_t \exp(\eta_t + \epsilon_{t+1})$):

$$q_{t+1} = \underbrace{\bar{d}\,\underbrace{e^{(1+\alpha_S)\eta_t-\alpha_S\omega_t}}_{\exp(\varphi_t)}\left[E_t\left(\frac{p_{t+1}}{\bar{p}}\, e^{\epsilon_{t+1}}\right)\right]^{\alpha_S}}_{\substack{E_t\, q_{t+1}\exp(-\sigma_\epsilon^2/2)\\ =\text{Expected production}}}\, e^{\epsilon_{t+1}}\,.$$

Back

Details on weighting matrix

- W is a diagonal matrix with elements corresponding to the inverse of the variance of the parameters of the auxiliary model;
- Calculated using
  - For regression parameters: Formulas for standard errors robust to heteroskedasticity;
  - For standard deviations:

$$\text{var}\left(\sigma^{\text{OLS}}\right) = \frac{\left(\sigma^{\text{OLS}}\right)^2}{2\left(T-l\right)} \text{ and } \text{var}\left(\sigma^{\text{2SLS}}\right) = \frac{\left(\sigma^{\text{2SLS}}\right)^2}{2\left(T-l\right)R_p^2},$$

with $T-l$ the degree of freedom and R_p^2 the partial R^2 of the 1$^{\text{st}}$ stage where endogenous variable and instrument have both been first regressed on the exogenous variables.

Back

Alternative Auxiliary model

- Based on 2SLS regressions (1st & 2nd stage of the IV model):

$$
\begin{aligned}
\log q_t &= a_q^{2SLS} + b_q^{2SLS} \log(\mathrm{E}_{t-1}\, p_t) + c_q^{2SLS} \psi_t + u_{q,t}^{2SLS}, \\
\log(\mathrm{E}_{t-1}\, p_t) &= a_{\mathrm{E}\,p} + b_{\mathrm{E}\,p} \psi_{t-1} + c_{\mathrm{E}\,p} \psi_t + u_{\mathrm{E}\,p,t}, \\
\log c_t &= a_c^{2SLS} + b_c^{2SLS} \log p_t + c_c^{2SLS} \log p_{t-1} + d_c^{2SLS} \log c_{t-1} + u_{c,t}^{2SLS}, \\
\log p_t &= a_P + b_p \psi_t + c_p \log p_{t-1} + d_p \log c_{t-1} + u_{p,t}, \\
\psi_t &= a_\psi + u_{\psi,t}.
\end{aligned}
$$

- 15 target parameters:

$\zeta = [b_q^{2SLS}, c_q^{2SLS}, \sigma_{u_q^{2SLS}}, b_{\mathrm{E}\,p}, c_{\mathrm{E}\,p}, \sigma_{u_{\mathrm{E}\,p}}, b_c^{2SLS}, c_c^{2SLS}, d_c^{2SLS}, \sigma_{u_c^{2SLS}}, b_p, c_p, d_p, \sigma_{u_p}, \sigma_{u_\psi}]$

Back

Non-stationarity tests

Summary

- Large literature on the nature of trends in commodity prices (e.g., Ghoshray, 2010; Lee et al., 2006);
- Prebish-Singer hypothesis of a secular deterioration in commodity prices relative to that of manufactured goods;
- Need to account for possible breaks in deterministic trends to avoid spurious rejection;
- LM tests allow for one or two structural breaks w/o a linear or quadratic deterministic trend under both the null and alternative hypotheses Lee and Strazicich (2003, 2013) and Lee et al. (2006);
- Stationarity further confirmed by ADF, PP and KPSS unit root tests implemented on detrended variables but 3 knots is not flexible enough.

Back

Auxiliary parameters fit

Coefficient	Observations: Estimate	Observations: Standard error	Model: Estimate
b_q	0.058	0.013	0.048
c_q	1.103	0.099	1.148
σ_{u_q}	0.015	0.001	0.015
b_c	−0.021	0.010	−0.007
c_c	−0.005	0.011	0.011
d_c	0.547	0.118	0.534
σ_{u_c}	0.014	0.001	0.014
$b_{\mathrm{E}\,p}$	−2.382	1.382	−1.687
$c_{\mathrm{E}\,p}$	−3.783	0.991	−2.303
$\sigma_{u_{\mathrm{E}\,p}}$	0.165	0.016	0.160
b_p	−4.112	0.937	−4.445
c_p	0.486	0.105	0.456
d_p	−0.130	1.690	2.881*
σ_{u_p}	0.180	0.018	0.180
σ_{u_ψ}	0.023	0.002	0.025
b_q^{2SLS}	0.075	0.021	0.086
b_c^{2SLS}	−0.065	0.026	−0.068

Back

Monte Carlo with IV approach

Back

OLS estimations of the supply and demand equations

	ρ_μ	$c_q - 1$	σ_ψ (%)	σ_ϑ (%)	σ_υ (%)	α_D	α_S
$\sigma_\omega = 5\%$							
Mean	0.36	0.049	2.49	2.64	1.28	−0.021	0.067
St. dev.	0.13	0.023	0.23	0.25	0.14	0.011	0.005
RMSE (%)	37.93	43.240	9.40	9.65	21.70	71.702	17.712
SE	0.14	0.024	0.24		0.13	0.011	0.005
$\sigma_\omega = 10\%$							
Mean	0.37	0.065	2.49	2.75	1.29	−0.022	0.053
St. dev.	0.13	0.043	0.23	0.27	0.14	0.011	0.009
RMSE (%)	36.57	55.737	9.40	10.37	20.97	69.774	35.730
SE	0.14	0.045	0.24		0.13	0.011	0.009
$\sigma_\omega = 20\%$							
Mean	0.39	0.077	2.49	3.00	1.33	−0.026	0.018
St. dev.	0.13	0.078	0.23	0.31	0.14	0.010	0.015
RMSE (%)	33.51	72.009	9.40	14.01	19.25	65.089	79.423
SE	0.13	0.083	0.24		0.13	0.010	0.014

Monte Carlo with indirect inference approach

Back

Results based on OLS auxiliary model

	ρ_μ	$\rho_{\eta,\omega}$	σ_ω	σ_η	σ_ϵ	σ_υ	δ	k	α_D	α_S
$\sigma_\omega = 5\%$	OID: 0.043									
Mean	0.50	−0.45	5.05	1.47	1.98	1.61	1.97	3.08	−0.071	0.080
St. dev.	0.11	0.31	0.66	0.33	0.28	0.26	1.34	2.27	0.016	0.008
RMSE (%)	22.19	78.65	13.31	22.14	14.13	16.26	67.26	75.82	22.380	9.539
ASE	0.09	0.39	0.67	0.36	0.30	0.26	19.33	18.19	0.020	0.008
$\sigma_\omega = 10\%$	OID: 0.049									
Mean	0.50	−0.44	10.22	1.46	1.98	1.62	2.00	3.08	−0.071	0.081
St. dev.	0.11	0.29	1.63	0.35	0.29	0.26	1.47	2.26	0.016	0.014
RMSE (%)	22.67	72.08	16.41	23.31	14.54	16.08	73.66	75.45	22.346	17.661
ASE	0.09	0.34	1.56	0.38	0.32	0.26	18.58	17.58	0.020	0.014
$\sigma_\omega = 20\%$	OID: 0.038									
Mean	0.50	−0.45	21.01	1.46	1.98	1.62	2.00	3.17	−0.072	0.083
St. dev.	0.11	0.28	5.60	0.38	0.31	0.24	1.62	2.28	0.015	0.026
RMSE (%)	22.94	70.52	28.44	25.25	15.53	15.25	81.17	76.20	21.506	32.172
ASE	0.10	0.33	5.23	0.41	0.34	0.26	18.51	17.09	0.020	0.025