

Estimating the Competitive Storage Model with Trending Commodity Prices

Christophe Gouel*

Nicolas Legrand†

Online Appendix

A1 Computational details

The results were obtained using MATLAB R2014a on a PC with two quad-core processors Intel Xeon E5345 (2.33 GHz) with 32 GB of RAM running Ubuntu 12.04.5 64 bits. The maximization of the log-likelihood was done using Vaz and Vicente's (2007) free particle swarm pattern search software PSwarm version 2.1¹ and the MATLAB function `fmincon` available in MATLAB Optimization Toolbox. The Monte-Carlo simulations have been performed using MATLAB's default random number generator with the seed set to 1. The Gaussian quadrature was calculated using a MATLAB's function from John Burkardt's website.² For a unit normal distribution truncated at five standard deviations, the Gaussian quadrature with 10 nodes has nodes $\epsilon^n = \{\pm 4.4576, \pm 3.3999, \pm 2.3838, \pm 1.4132, \pm 0.4684\}$ and weights $\pi^n = \{1.9834 \times 10^{-5}, 1.2876 \times 10^{-3}, 2.3048 \times 10^{-2}, 0.14029, 0.33536\}$.

A2 Small Sample Properties

In this section, we assess using Monte Carlo experiments the small sample properties of the simulated Unconditional Maximum Likelihood estimator (UML) we developed, and compare them to those of the Conditional Maximum Likelihood estimator (CML) proposed in Cafiero et al. (2015). From equation (20), the conditional log-likelihood without trend is obtained by removing the terms corresponding to the marginal likelihood:

$$\begin{aligned} \log L^C(\theta; P_{1:T}^{\text{obs}}) = & -\frac{T-1}{2} \log 2\pi - (T-1) \log [\Phi(5) - \Phi(-5)] \\ & + \sum_{t=2}^T \log \left| \mathcal{D}^{-1'}(P_t^{\text{obs}}) \right| - \sum_{t=2}^T (1_{|\epsilon_t| \leq 5} \cdot \epsilon_t^2 + 1_{|\epsilon_t| > 5} \cdot \infty) / 2. \quad (\text{A1}) \end{aligned}$$

*INRA, Économie Publique, and CEPII (christophe.gouel@grignon.inra.fr)

†INRA, Économie Publique, and Climate Economics Chair (nicolas.legrand@grignon.inra.fr)

¹<http://www.mat.uc.pt/~lnv>

²http://people.sc.fsu.edu/~jburkardt/m_src/truncated_normal_rule/truncated_normal_rule.html

Following [Michaelides and Ng \(2000\)](#) and [Cafiero et al. \(2015\)](#), we conduct four Monte Carlo experiments varying the parameterization and the length of the samples. The first set of parameters are $a = 1$, $b = -1$, and $k = 0.02$, which implies a storage cost of 2% of the mean price and, for supply shocks with a coefficient of variation of 5%, a demand elasticity of -0.05 in the range of the best estimates obtained by [Roberts and Schlenker \(2013\)](#) on a caloric aggregate of major crops, but slightly higher in absolute value than our estimated elasticities. The second parameterization only differs by the value of b , now equal to -2 . This rotation of the slope of the demand function around its mean halves the demand elasticity making this parameterization more favorable to storage, and closer to the values found in the article. For each set of parameters, we solve for the equilibrium price function on a grid of 1,000 points, and obtain 3,000 prices series of length $T = 50$ and $T = 100$ from the asymptotic distribution. The price series are obtained by the simulation of 3,000 trajectories starting from the steady-state availability and discarding the first 50 periods as burn-in periods.

The numerical methods follow what was described previously, but differ on two aspects. Firstly, to prevent the availability corresponding to the cutoff price to be below the lower bound of the grid of interpolation points, the lower bound is changed from being -2 to being -5 , the minimum availability. Secondly, since the log-likelihood optimization behaves better on simulated samples, we use a faster optimization algorithm: the generalized pattern search algorithm implemented by the MATLAB function `patternsearch` available in MATLAB Global Optimization Toolbox. The optimization starts from initial values randomly drawn in the range between 80% and 120% of the true values. If the optimization solver fails to converge for one of the two estimators, we discard the corresponding samples for both estimators. The results of valid estimates obtained on common samples are given in table [A1](#) and [A2](#).

The results of the Monte Carlo experiments are similar to those obtained for the CML in [Cafiero et al. \(2015\)](#). They show that the two maximum likelihood estimators yield precise estimates of the parameters of the model, especially for a and b . The storage cost, k , is less precisely estimated with Root Mean Square Errors (RMSE) always above 26%. For all parameters, the bias is small, most of the RMSE coming from the standard deviation of the estimations.

For all parameterizations, the estimators perform better when the sample length increases. For the UML, a doubling of the sample length from 50 to 100 observations reduces the RMSE by 17% for both parameterizations. The CML benefits slightly more than the UML from an increase in the sample size. Indeed, they have similar RMSE for the long samples, but the UML performs better on the short samples. Regarding the influence of the parameterization, we observe that the parameterization more favorable to storage yields less precise estimates as all the RMSE of table [A2](#) are higher than in table [A1](#).

Table A1: Comparison of Monte Carlo Experiment Results with Parameterization $a = 1$, $b = -1$, and $k = 0.02$

	UML			CML		
	a	b	k	a	b	k
$T = 50$						
Mean	0.9930	-0.9720	0.0193	0.9932	-0.9745	0.0194
Standard deviation	0.0650	0.1297	0.0059	0.0668	0.1344	0.0058
Bias	-0.0070 (0.70%)	0.0280 (2.80%)	-0.0007 (3.40%)	-0.0068 (0.68%)	0.0255 (2.55%)	-0.0006 (3.14%)
RMSE	0.0654 (6.54%)	0.1326 (13.26%)	0.0059 (29.67%)	0.0671 (6.71%)	0.1368 (13.68%)	0.0058 (29.24%)
$T = 100$						
Mean	0.9951	-0.9854	0.0196	0.9944	-0.9875	0.0194
Standard deviation	0.0510	0.1066	0.0053	0.0516	0.1070	0.0052
Bias	-0.0049 (0.49%)	0.0146 (1.46%)	-0.0004 (2.17%)	-0.0056 (0.56%)	0.0125 (1.25%)	-0.0006 (2.75%)
RMSE	0.0512 (5.12%)	0.1076 (10.76%)	0.0053 (26.70%)	0.0519 (5.19%)	0.1077 (10.77%)	0.0053 (26.26%)

Notes: The price samples for which one estimator does not converge are discarded. For $T = 50$, the total number of valid replications is 2,737 for UML and 2,733 for CML. For $T = 100$, it is 2,841 for UML and 2,846 for CML. The table reports the 2,593 and 2,764 valid estimates obtained on common samples for the short and long samples.

Table A2: Comparison of Monte Carlo Experiment Results with Parameterization $a = 1$, $b = -2$, and $k = 0.02$

	UML			CML		
	a	b	k	a	b	k
$T = 50$						
Mean	0.9793	-1.9639	0.0194	0.9772	-1.9695	0.0193
Standard deviation	0.1280	0.3202	0.0085	0.1351	0.3250	0.0085
Bias	-0.0207 (2.07%)	0.0361 (1.80%)	-0.0006 (3.14%)	-0.0228 (2.28%)	0.0305 (1.52%)	-0.0007 (3.69%)
RMSE	0.1296 (12.96%)	0.3222 (16.11%)	0.0085 (42.51%)	0.1370 (13.70%)	0.3265 (16.32%)	0.0085 (42.45%)
$T = 100$						
Mean	0.9848	-1.9838	0.0195	0.9825	-1.9860	0.0194
Standard deviation	0.1055	0.2592	0.0074	0.1090	0.2611	0.0076
Bias	-0.0152 (1.52%)	0.0162 (0.81%)	-0.0005 (2.56%)	-0.0175 (1.75%)	0.0140 (0.70%)	-0.0006 (3.14%)
RMSE	0.1065 (10.65%)	0.2597 (12.99%)	0.0074 (36.94%)	0.1104 (11.04%)	0.2615 (13.08%)	0.0076 (38.02%)

Notes: The price samples for which one estimator does not converge are discarded. For $T = 50$, the total number of valid replications is 2,820 for UML and 2,767 for CML. For $T = 100$, it is 2,899 for UML and 2,890 for CML. The table reports the 2,674 and 2,839 valid estimates obtained on common samples for the short and long samples.

A3 Production variation

Table A3: Production variation, 1961–2011

Commodity	Production CV (%)	Commodity	Production CV (%)
Banana	3.67	Palm oil	4.65
Cocoa	7.13	Rice	2.75
Coffee	7.68	Sugar	3.65
Copper	4.46	Tea	2.07
Cotton	7.20	Tin	5.52
Jute	11.35	Wheat	4.34
Maize	5.84		

Notes: The coefficients of variation (CV) are obtained by calculating the standard deviation of the detrended logarithm of observed production, modeling the trend using a restricted cubic splines with five knots.

A4 Parameters estimates

Table A4: Parameter estimates with 3-knot spline trend

Commodity	g_1	g_2	a	b	k	$\log L$	# Stockouts
Banana	-0.3323 (0.2804)	-1.2359 (0.2011)	1.3237 (0.2013)	-1.6729 (0.4397)	0 -	170.6770	0
Cocoa	-1.5599 (0.8803)	0.7076 (0.3317)	0.3679 (0.1595)	-1.7978 (0.9285)	0.0002 (0.0008)	215.6308	0
Coffee	0.0925 (0.3314)	-0.6813 (0.1557)	0.2522 (0.0455)	-0.5265 (0.1072)	0.0036 (0.0025)	194.9321	1
Copper	-0.4176 (0.1943)	-0.2464 (0.1424)	0.7348 (0.0789)	-1.1835 (0.1589)	0.0073 (0.0051)	122.1660	0
Cotton	-0.6422 (0.8240)	-2.0878 (0.3217)	1.5367 (0.4016)	-3.7632 (1.6051)	0.0053 (0.0065)	123.0208	0
Jute	-0.3974 (0.0756)	-0.9164 (0.0554)	0.6917 (0.0481)	-0.7573 (0.0401)	0.0163 (0.0062)	102.1980	4
Maize	-0.6588 (0.1101)	-0.8850 (0.0815)	0.9504 (0.0816)	-1.5377 (0.1432)	0.0128 (0.0075)	84.9322	3
Palm oil	-1.7887 (0.1181)	-1.2443 (0.0539)	1.0411 (0.0981)	-1.7009 (0.1865)	0.0146 (0.0091)	119.2428	1
Rice	-1.5405 (0.3035)	-1.4251 (0.1417)	1.2128 (0.1914)	-2.0786 (0.3684)	0.0037 (0.0066)	127.1550	1
Sugar	-1.2627 (0.1028)	-0.2394 (0.1652)	0.9701 (0.0809)	-1.8482 (0.1738)	0.0235 (0.0098)	60.7431	8
Tea	-0.9290 (0.3585)	-1.1775 (0.1357)	1.1089 (0.2085)	-1.6789 (0.4149)	0.0105 (0.0053)	144.1925	0
Tin	1.7957 (0.4707)	-0.7677 (0.2843)	0.1940 (0.0244)	-0.7567 (0.1185)	0 -	209.3322	0
Wheat	-1.2473 (0.1482)	-0.8441 (0.1517)	1.2069 (0.1069)	-1.5449 (0.2021)	0.0143 (0.0096)	104.0772	1

Note: Asymptotic standard errors in parenthesis.

Table A5: Parameter estimates with 4-knot spline trend

Commodity	g_1	g_2	g_3	a	b	k	$\log L$	# Stockouts
Banana	-0.5940 (0.2995)	-0.2663 (0.3033)	-1.5903 (0.2712)	1.2033 (0.1735)	-1.5061 (0.3602)	0 -	171.7633	0
Cocoa	0.2862 (0.2857)	-2.6193 (0.5337)	0.3417 (0.4282)	0.4498 (0.0981)	-1.1874 (0.2765)	0.0024 (0.0028)	219.7236	0
Coffee	0.1199 (0.1792)	0.0281 (0.3230)	-0.7565 (0.1351)	0.2339 (0.0495)	-0.4885 (0.0898)	0.0033 (0.0024)	194.9185	2
Copper	0.6214 (0.3452)	-0.8595 (0.3155)	-0.2444 (0.1968)	0.7416 (0.1155)	-1.7526 (0.3542)	0.0005 (0.0044)	123.6516	4
Cotton	-0.2545 (0.7621)	-1.1035 (0.8597)	-2.2689 (0.4857)	1.4300 (0.5264)	-3.5214 (2.0232)	0.0049 (0.0069)	122.8690	0
Jute	-0.0613 (0.1039)	-0.6035 (0.1805)	-0.9908 (0.0710)	0.6967 (0.0797)	-0.7614 (0.0715)	0.0163 (0.0066)	102.1845	8
Maize	0.0815 (0.1073)	-1.1899 (0.0887)	-1.1229 (0.0576)	0.9358 (0.0481)	-0.8819 (0.0508)	0.0304 (0.0068)	88.8597	12
Palm oil	-0.9208 (0.1402)	-1.6398 (0.3924)	-1.5053 (0.1190)	0.9971 (0.1744)	-1.6352 (0.3267)	0.0142 (0.0095)	119.6359	1
Rice	-0.8285 (0.2131)	-1.7697 (0.4977)	-1.6808 (0.1837)	1.2815 (0.3139)	-2.2062 (0.5310)	0.0042 (0.0070)	127.1070	1
Sugar	-0.6992 (0.1369)	-1.3979 (0.2635)	-0.3698 (0.2479)	1.1136 (0.1329)	-2.0999 (0.2176)	0.0262 (0.0106)	60.8484	16
Tea	-1.4922 (0.3475)	-0.4250 (0.4818)	-1.8818 (0.2257)	1.1959 (0.2308)	-1.7916 (0.4785)	0.0107 (0.0051)	145.6865	0
Tin	1.7411 (0.5157)	1.1081 (0.2943)	-0.2441 (0.1589)	0.1517 (0.0113)	-0.5043 (0.0414)	0 -	210.3823	6
Wheat	-0.1494 (0.1744)	-1.9057 (0.3087)	-0.9026 (0.1902)	1.3447 (0.1606)	-2.3618 (0.3405)	0 -	105.8230	0

Note: Asymptotic standard errors in parenthesis.

A5 Figures of price trends

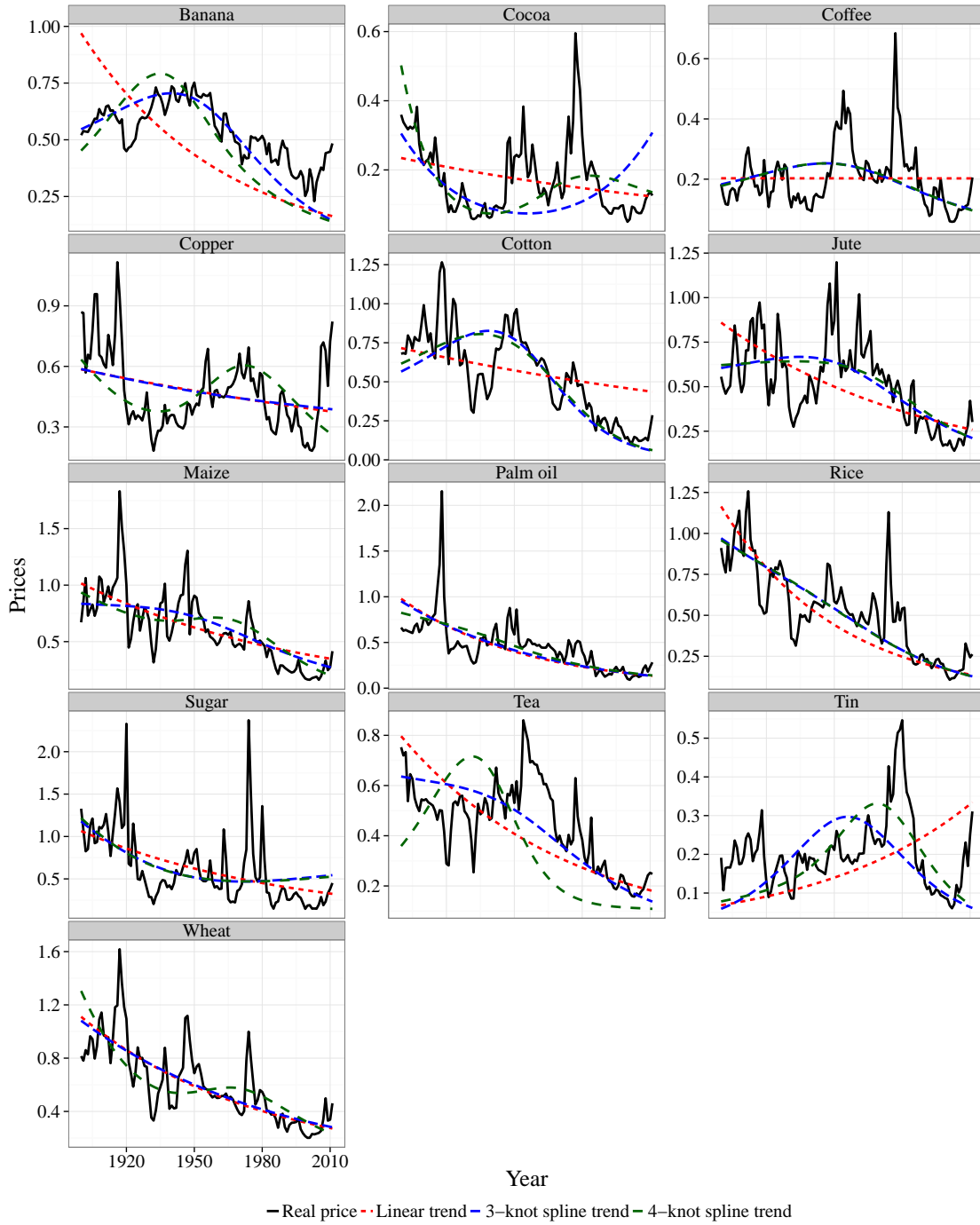


Figure A1: Price trends

References

- Cafiero, C., Bobenrieth, E. S. A., Bobenrieth, J. R. A. and Wright, B. D. (2015). Maximum Likelihood estimation of the standard commodity storage model: Evidence from sugar prices. *American Journal of Agricultural Economics*, 97(1), 122–136.
- Michaelides, A. and Ng, S. (2000). Estimating the rational expectations model of speculative storage: A Monte Carlo comparison of three simulation estimators. *Journal of Econometrics*, 96(2), 231–266.
- Roberts, M. J. and Schlenker, W. (2013). Identifying supply and demand elasticities of agricultural commodities: Implications for the US ethanol mandate. *The American Economic Review*, 103(6), 2265–2295.
- Vaz, A. I. F. and Vicente, L. N. (2007). A particle swarm pattern search method for bound constrained global optimization. *Journal of Global Optimization*, 39(2), 197–219.