International Journal of Forecasting 32 (2016) 838–848

Contents lists available at ScienceDirect

International Journal of Forecasting

journal homepage: www.elsevier.com/locate/ijforecast

Forecasting food prices: The case of corn, soybeans and wheat

H. Ahumada[a], M. Cornejo[a,b,*]

[a] Di Tella University, Argentina

[b] Buenos Aires University, Argentina

ARTICLE INFO

Keywords:
Forecast
Food prices
Equilibrium correction
Joint models
Breaks
Robust devices

ABSTRACT

Given the high correlations observed among food prices, we analyse whether the forecasting accuracies of individual food price models can be improved by considering their cross-dependence. We focus on three strongly correlated food prices: corn, soybeans and wheat. We analyse an unstable forecasting period (2008–2014) and apply robust approaches and recursive schemes. Our results indicate forecast improvements from using models that include price interactions.

1. Introduction

Food prices have shown strong correlations in the past, even before their upward co-movement over recent decades. For instance, the commodities included in the World Bank's food price index for the period 1990–2013 (on a monthly basis) show price correlations of well over 0.60, and even up to 0.85 for some subsets. We therefore explore whether or not the forecast accuracies of a subset of these commodity prices could be improved by taking their cross-dependence into account. Some might think that this is an old question that has already been answered, as simultaneous modelling did not survive after the 1973 oil crisis, due to the poor forecasting performances of macro models relative to naïve forecasts. However, we now have a better understanding of the effects of breaks on forecasting. Various different devices and methods, such as robust transformations and updating, may be useful for forecasting in the presence of breaks, and can also be applied for joint and other models that consider cross-dependence.

We focus on three food prices which are strongly correlated: corn, soybeans and wheat. These are agricultural commodities that, whether directly or indirectly, feed a large part of the world's population. There has been a special interest in understanding the common behaviours of their prices since the 2000s, when their downward trend reversed, as the demand for them started to increase significantly, driven by the unprecedented growth of emerging economies such as China and India. The demand for oilseeds has also increased greatly, due to their competing use as biofuels. Because of these effects, and also due to various macro and financial developments, their prices have experienced a long-term boom, along with many other food, mineral and energy commodities.

We are interested mainly in developing conditional forecasts of food prices in which the out-of-sample values of the weak exogenous variables will come from outside the model; that is, will be provided by the forecaster (e.g., from organizations such as the World Bank, FAO, IMF or USDA). These values should respond to conjectural scenarios about the future behaviours of the regressors, in order to quantify what would happen to corn, soybean and wheat prices if, for example, the economy of China decelerated at a given rate, or the US monetary policy changed. Thus, using the conditional forecasting models should also make it possible to project what might happen to food

* Correspondence to: Av. Figueroa Alcorta 7350, Buenos Aires, Argentina.

E-mail addresses: hahumada@utdt.edu (H. Ahumada), mcornejo@utdt.edu (M. Cornejo).

http://dx.doi.org/10.1016/j.ijforecast.2016.01.002
0169-2070/

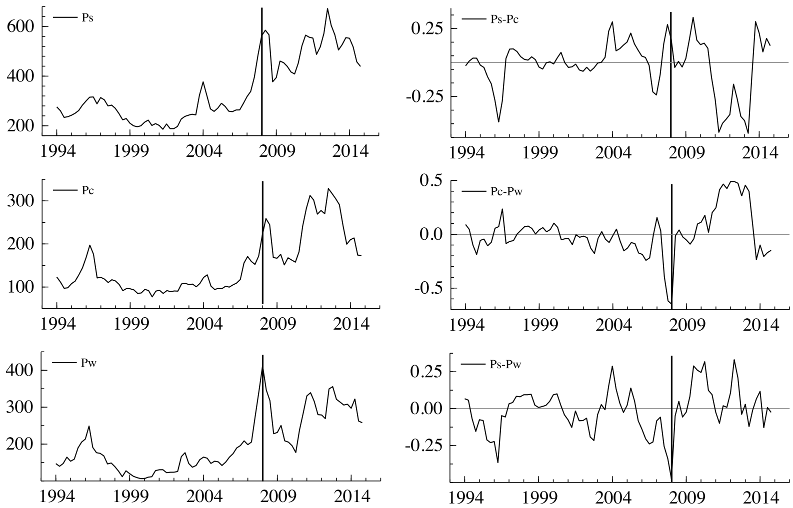


Fig. 1. Nominal absolute and relative prices from 1994 to 2014 (in US dollars). Note: The relative prices are calculated mean-adjusted. The vertical lines indicate the beginning of the out-of-sample period.

prices given a range of assumptions regarding the paths of the explanatory variables employed in the model. To allow for the effects of these variables, a necessary condition is to evaluate the forecasting accuracy of the econometric models over a given pseudo out-of-sample period and forecast horizon.

In this paper, our out-of-sample period includes the 2008–2009 world crisis and its aftermath, together with the recent reversion of the upward trend in commodity prices. As part of model selection, we pay special attention to the cross correlation of prices that motivated the discussion about single vs. joint modelling of these food prices. Forecasts are developed for one and four quarters ahead. For the multi-step forecasts, both iterated and direct approaches are tried. Robust model transformations (which keep the effects of the conditioning variables) and other robust forecast methods are also applied for comparison purposes, due to the unstable behaviour of the food commodity prices studied.

The next section describes the data and briefly reviews the empirical literature. Section 3 presents various different forecasting strategies. Sections 4 and 5 provide the estimation and forecast results, respectively. Section 6 evaluates forecasting biases and pseudo out-of-sample breaks. Finally, Section 7 concludes.

2. Data

This section describes the data used to estimate our forecasting models. Our data set of the nominal prices of corn, soybeans and wheat is quarterly over the period 1994Q3–2014Q4 (82 observations). Since our forecasting period starts in 2008Q1, we can observe what happens to the forecast accuracy of the models as the world crisis evolved and the last super boom seemed to end. Fig. 1 shows the joint behaviours of these three prices over the sample period.

Corn, soybeans and wheat behave similarly over the whole sample period, even before the upward trend that was observed from the early 2000s. In fact, the cross-correlations among them, on a quarterly basis, are higher than 0.9 for the whole sample. However, the relative prices also show instability over the last five years of the sample.

Although our estimations are on a quarterly basis, the production of corn and soybeans enter our model on a yearly basis. For them, we repeated the values of the estimated production published by the U.S. Department of Agriculture (USDA) in the second quarter of each year. The production of soybeans and corn are concentrated geographically, with the leading exporters being the United States, Brazil and Argentina. Because of this, we consider that, by the second quarter, when the southern hemisphere (Brazil and Argentina) is harvesting the bulk of these commodities and the northern hemisphere's (United States) planting season takes place, the USDA will have reliable estimates of the whole year's production. For wheat, the production of which is dispersed more widely geographically, we used the quarterly averages of the monthly estimates of annual production reported by USDA. Doing this not only improved our econometric models, but also allowed us to use the USDA production projections in a multivariate framework in order to obtain conditional forecasts of these prices.

Our information set uses several potential macro-wide and market-driven explanatory variables of commodity prices that are used commonly in the literature. It includes macroeconomic determinants (like the US monetary aggregates or the US exchange rate) as well as commodity-specific variables (such as production and inventories) and demand factors (such as the growth of emerging countries like China).[1] For further data definitions and the sources of the variables included in our models, see Appendix A.

Fig. 2 shows the behaviours of production and inventories of the commodities studied.

[1] Futures prices for explaining (spot) food prices are not included, as they would convey the same information that arises from the economic fundamentals.

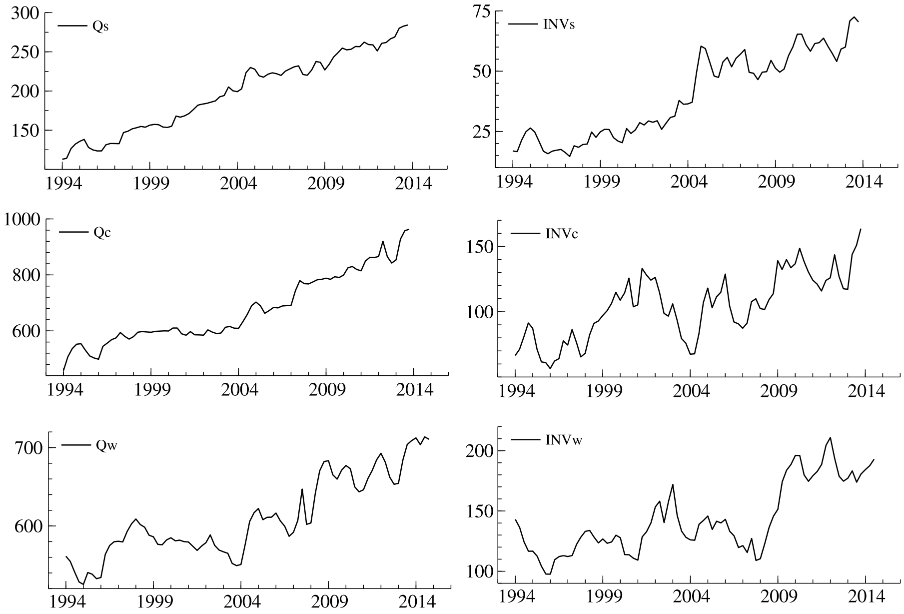


Fig. 2. World production and inventories (in millions of metric tons).

3. Model designs

This section explains the econometric approach that we followed for estimating models that can be used to obtain conditional forecasts of food prices.

(i) We started by estimating an equilibrium correction model (EqCM) for each food price (corn, soybeans and wheat) over the period 1994Q3–2007Q4 (the in-sample period), using a single-equation approach that includes levels as well as differences. These EqCMs, which nest long-run and short-run behaviours, allow us to take into account supply and demand determinants along with the effects of macro and financial variables. This representation is also useful when both (co)integrated and stationary variables are included in the models.

(ii) After estimating these models, we studied the residual cross-correlations in order to evaluate their interdependence, and carried out joint modelling if necessary. For example, agents involved in agribusiness often use the soybean to corn price ratio as a rule of thumb to guide their decisions. A general rule of thumb is that when the price of soybeans is more than 2.4–2.5 times the price of corn, farmers will probably plant more soybeans than corn. Thus, we analyse econometrically how these prices, along with the wheat prices, are related econometrically, by following a system approach for evaluating cointegration relationships and weak exogeneity. The EqCMs of the cointegrated prices may be seen as partial systems, and therefore the cointegration results will remain valid if more variables are added to a larger system (see e.g. Juselius, 2006, ch. 19).

(iii) Given the results of (i) and (ii), the models that we estimate for forecasting are special cases of the following models for the price of i ($i = 1, 2, 3$, corresponding to corn, soybeans and wheat).

$$\Delta p_{it} = \delta_i + \sum_{i=1}^{3} \theta_i EqC_{it-1} + \lambda_i' \Delta z_{it:t-l} + \varepsilon_{it},$$
$$\varepsilon_{it} \sim \mathrm{IN}(0, \sigma^2), \qquad (1)$$

for $t = 1, \ldots, T$, where EqC denotes the equilibrium correction terms that correspond to the fundamentals' long-run deviations from (i) or price differentials from (ii); $\lambda_i' \Delta z_{it:t-l}$ denotes the short-run effects of the relevant variables (where $\Delta z_{it:t-l}$ is a $k \times 1$ vector with lags 0 to l) for each i-price equation. These may include both contemporaneous and lagged effects of the own and related food prices. The results are reported in Section 5.

4. Construction of the forecasts

The estimation of the models described in the last section is the starting point of our forecasting exercise; however, other devices and models may also be useful for forecasting food prices. Examples of such are presented briefly in this section.

Unit roots and breaks are the two main sources of non-stationarity that are faced by econometric models. When dealing with integrated processes, Engle and Granger (1987) and Johansen (1988) suggested the use of equilibrium correction models (for vectors denoted VEqCM)[2] for estimating cointegrated relationships. Initially, they were recommended for forecasting as well, but both evidence (e.g. Makridakis & Hibon, 2000; Stock & Watson, 1999)

[2] Our estimated EqCMs are special cases of VEqCMs for weak exogeneous explanatory variables.

and theory (Clements & Hendry, 1998, 1999) showed that they perform poorly when economic processes are subject to structural breaks, as often occurs. As Castle, Doornik, Hendry, and Petris (2015, p. 3) explain, EqCMs "always correct back to the old equilibrium (...) irrespective of whether or not the equilibrium has shifted", thereby inducing forecast failures.

In the presence of breaks, the use of non-causal devices as double differenced (DDD) variables can outperform VEqCM; however, on the basis of this finding, a double differenced VEqCM (DVEqCM) should in turn outperform DDD (at least in mean, but not necessarily in variance or mean square forecast error, MSFE).[3] Castle, Fawcett, and Hendry (2010), Clements and Hendry (2006) and Hendry (2006) provide detailed explanations of the robustification of forecasts from equilibrium correction systems.

Alternatively, the relevant equations of a VEqCM can be adjusted after the break occurs, using residual adjustments or intercept corrections (IC).[4] This can be done by putting the forecast back on track when the forecast errors are correlated.

Furthermore, since, by construction, a DVAR does not include the equilibrium correction terms, it is another alternative that could be explored for forecasting, and will be also considered. In particular, we analyse the forecasting accuracy of a DVAR of the three prices under study.

As Hendry (2006) noted, equilibrium mean shifts are more common in lower frequency data, meaning that differencing is probably less helpful for higher data frequencies. Since we use quarterly data, the benefits of differencing are not clear.

For forecasting, we also performed direct multi-step estimation for $h = 4$ (see Chevillon, 2007; Clements & Hendry, 1996), using lagged information for lags $\geq h$ that may be useful for non-stationary cases.

To sum up, as Table 1 shows, the conditional forecasts obtained from the (individually and jointly) estimated EqCMs are compared with those obtained by following other approaches that hedge against potential breaks, such as the differenced EqCM (DEqCMs), without losing the effects of the conditioning variables. We also considered other forecasting approaches that are robust to breaks: taking into account the interdependence of prices (DVAR), and devices where no estimation is required, such as double differencing (DDD),[5] as well as the random walk itself, as usual.

Table 1
Alternative forecasting models for each food price.

Model type	Horizon	Scheme (in quarters)
EqCM	1	recursive
(single and joint)	1	fixed
	4 (iterative)	recursive
	4 (iterative)	fixed
	4 (direct)	recursive
	4 (direct)	fixed
DEqCM	1	recursive
(single and joint)	1	fixed
	4 (iterative)	recursive
	4 (iterative)	fixed
	4 (direct)	recursive
	4 (direct)	fixed
DVAR	1	recursive
	1	fixed
	4 (iterative)	recursive
	4 (iterative)	fixed
	4 (direct)	recursive
	4 (direct)	fixed
Random walk	1	–
	4	–

5. Econometric modelling results

This section begins by presenting the results from the estimation of the forecasting EqC models of corn, soybean and wheat prices. Then, it analyses the cointegration among food prices, in an attempt to improve the forecasting models by considering their interdependence.

To develop conditional models of nominal food prices based on their fundamentals, we first followed a general-to-specific approach with a single equation for each price that includes levels and up to two lagged differences.[6] We then used *Autometrics* (see Doornik, 2009; Hendry & Doornik, 2014) to help us to select the dominant congruent model, not just a best fit. Autometrics is an algorithm that uses a tree search to discard paths that are rejected as reductions of the initial model, and includes diagnostic testing.

Starting from a large information set (as in Ahumada & Cornejo, 2015), we jointly evaluate many of the explanatory variables that have been suggested in the vast body of literature that has tried to explain commodity prices. In the general unrestricted model (GUM),[7] we considered commodity-specific variables such as commodity production (q_t) and inventories (INV_t), as commodity prices tend to rise when quantities and inventories decrease. To take into account the demand effects, world and emerging and developing economic growth are captured through several variables suggested by the literature: real GDP for the OECD ($\mathrm{gdp}_t^{\mathrm{OECD}}$), India ($\mathrm{gdp}_t^{\mathrm{INDIA}}$) and China ($\mathrm{gdp}_t^{\mathrm{CHINA}}$, seasonally adjusted). To assess the possible impact of the monetary and financial environment on commodity prices, we included monetary aggregates such as the US monetary base (mb_t), M2 ($m2_t$), and the Fed's flow of funds (fof) as measures of liquidity, as well as the 3-month Treasury constant maturity rate (r_t). The US consumer price index (cpi_t^{US}) was

[3] The DVEqCM may be preferred to the DDD for conditional forecasting, as it maintains the explanatory variables of the VEqCM.

[4] As Hendry (2006) indicates, DVEqCM has the same effect as IC.

[5] Although the results are not reported here, we also calculated the RMSE and MAPE of the DDD, but no gains were found in any case.

[6] Rather than modelling the real food prices, we modelled the nominal prices as they were not statistically significant when estimating the real food prices conditional on the deflator and its lags.

[7] We keep fixed the variables that were significant in the long-run (which are needed for cointegration, according to the critical values of Ericsson & MacKinnon, 2002), and reselect the variables which turned out to have signs that differed from those expected by economic theory.

Table 2
In-sample quarterly single EqCM estimates: 1994Q3–2007Q4 (in logs).

Dependent variable:	Δp^c	Δp^s	Δp^w
Constant	12.19***	10.87***	1.8***
EqC_{it-1}	−0.37***	−0.31***	−0.36***
$\Delta_4 r_t$	−0.12***	–	–
$\Delta q_t^{\text{ethanol}}$	0.30**	–	–
$\Delta \text{gdp}_{t-1}^{\text{INDIA}}$	0.20**	–	–
Δp_{t-1}^s	–	0.28***	
Δp_{t-1}^w	–		0.14
Δe_t	–	−0.94**	–
$\Delta m2_t$	–		3.03**
$\widehat{\sigma}$	0.057	0.060	0.066
Long-run solution			
Dependent variable:	p^c	p^s	p^s
q_{t-1}^i	−2.64***	−1.50***	−2.35***
e_{t-1}	−3.68***	−1.54***	−2.89***
$\text{gdp}_{t-1}^{\text{CHINA}}$	0.89***	0.97***	–
$\text{gdp}_{t-1}^{\text{OECD}}$	–	–	2.21***
F-test[a]	0.64 [0.84] $F(31, 12)$	1.13 [0.42] $F(33, 13)$	0.83 [0.68] $F(26, 20)$

Notes: Insignificant variables are left because of autocorrelation. The selected (not reported) impulse dummies are: 1996Q2, 1996Q4 and 2006Q4 for corn prices, 2003Q4 for soybean prices and 1996Q2, 2002Q3, 2007Q3 and 2007Q4 for wheat prices.
[a] The F-test results correspond to the F statistic for the final model versus the initial GUM, with the associated tail probability (in square brackets) and the degrees of freedom (in parentheses).
*** $p < 0.01$.
** $p < 0.05$.
* $p < 0.10$.

also considered as a separate explanatory variable, along with the effect of the US real exchange rate (e_t). Finally, we also included the U.S. ethanol production (q^{ethanol}). Small letters indicate that the variables are expressed in logs (see Appendix A for further details and sources).

Table 2 shows the in-sample (1994Q3–2007Q4, $T = 54$) estimates of the selected EqCMs for each individual commodity price. All estimated models are consistent with theory, and pass all diagnostic tests at traditional levels.[8] We use a 5% target size for *Autometrics*.

For the individual models of corn, soybean and wheat prices, we found significant effects of individual productions in the long run, the US real exchange rate, and, as demand variables, the GDP for China (corn and soybeans) or the OECD (wheat).

Regarding commodity production, the negative sign of the estimated coefficient suggests that we could assume a valid conditional model of prices on production. However, we also tested weak exogeneity, following the approach of Johansen (1996), and found that prices (not production) adjust to reach the long-run relationship.[9]

Our estimates also showed that, in the short run, changes in corn prices can be explained by ethanol production growth, India's GDP growth of the previous year, and changes in the real interest rate. Corn models have richer short run effects than those for either soybeans or wheat. Soybeans and wheat each have an autoregressive term apart from the effect of the US exchange rate depreciation or the US changes in broad money stock ($\Delta m2$), respectively.

A key finding was that we still found high residual cross-correlations after the individual EqCMs had been estimated for corn and soybeans (0.47), but not for wheat (<0.13).

Given the observed cross-dependence between corn and soybean prices, their models may be enriched by estimating joint EqCMs that consider the interactions between them. We therefore studied the cointegration among these nominal prices, along with wheat (all in logs), during the in-sample period (1994Q3–2007Q4), as is shown in Table 3.

A VAR(2) estimation showed that only one cointegration relationship exists at the 5% significance level (considering the trace statistic), between corn and soybeans, as wheat prices were not significant in this vector. This indicates that the price differential ($p^s - p^c$) is stationary with a linear trend in the cointegration space. Only soybeans adjust to deviations from the long-run relationship. This finding implies that corn is weakly exogenous, and thus, a conditional model of soybean prices on corn prices is a valid representation of this system. We also analysed the case of $r = 2$, as suggested by the max statistic at the 5% level. In this case, the first vector indicated a stationary price differential between p^s and p^w, and the second vector a stationary price differential between p^c and p^w plus a trend. The only weak exogenous variable was p^w.[10] Evaluating irreductable cointegration (see Davidson, 1998, p. 95), these vectors could represent a reduced form where p^w is the instrument in the conditional relationship of p^s on p^c (as was indicated for $r = 1$). Furthermore, the relationship between soybean and corn prices is supported by the behaviour of agribusiness agents, whose economic decisions (such as planting) are based mainly on this price differential. Both commodities are also experiencing increasing demand for biofuels, and our single equation estimations showed that they have similar long-run determinants (e.g. China's GDP).

Given the system results for $r = 1$, we included the same explanatory variables as in the individual EqCM in the case of corn, while for soybeans, the long-run relationship between corn and soybean prices showed better forecasting results than including its long-run determinants from the individual EqCM. The joint EqCM estimation of corn and soybean prices is reported in Table 4.

A contemporaneous effect of the corn price on the price of soybeans turned out to be highly significant, and the residual cross-correlation decreased to just 0.09. Even if corn prices were found to be weakly exogenous, it should be noted that we need to estimate the corn and soybean prices jointly for forecasting purposes.

8 The diagnostic statistics are not reported, but can be obtained from the authors upon request.

9 However, the unrestricted systems did still show some misspecification problems, due mainly to non-normality that was not resolved by the inclusion of dummies.

10 The conditional EqCMs of corn on wheat and soybeans on wheat showed a residual cross-correlation of 0.7, suggesting that the structural relationship between soybeans and corn is not accounted for.

Table 3
Cointegration analysis, 1994Q3–2007Q4.

r	Eigenvalue	Trace	p-value	Max	p-value
0	0.49	60.72	0.00	36.62	0.00
1	0.33	24.09	0.08	21.90	0.02
2	0.04	2.19	0.94	2.19	0.94

For $r = 1$
Adjustment coefficients (α)

Variable	Unrestricted		Restricted	
	Coeff.	S.E.	Coeff.	S.E.
p^s	-0.17^{**}	0.07	-0.31^{**}	0.05
p^c	0.18	0.10	–	–
p^w	-0.06	0.10	–	–

Eigenvectors (β)

Variable	Coeff.	S.E.	Coeff.	S.E.
p^s	1.00	–	1.00	–
p^c	-1.47^{**}	0.26	-1.03^{**}	0.10
p^w	0.36	0.25	–	–
trend	-0.005^{**}	0.001	-0.004^{**}	0.001

Note:
** Indicates significance at the 1% level.

6. Forecast results

This section compares the forecast accuracies of the different econometric models. The results are divided into three subsections: forecast comparisons in terms of the root mean squared errors (RMSE) and mean absolute percentage errors (MAPE), a time-varying forecast bias evaluation, and the time-varying forecast ability.

For the forecast evaluation, we define T as the total in-sample observations, H as the total out-of-sample observations, $T^* = T + H$ as the total number of observations in the sample, and h as the forecast horizons.

Given our interest in forecasting the levels of prices, we convert the growth rates (expressed as differences of logarithms) obtained from the estimation of the EqCMs presented in the last section to (median) levels in order to evaluate the forecasting accuracy. Similar transformations were performed for other devices.

Our initial pseudo out-of-sample period is from 2008Q1 to 2014Q4, $H = 28$. We consider horizons of $h = 1$ and $h = 4$ quarters (one year) using a fixed estimation sample (1994Q3–2007Q4, $T = 54$) and a recursive scheme (the sample expands every quarter), as summarised in Table 1.

6.1. RMSE and MAPE

The forecast results are first evaluated using the RMSE and MAPE, as is shown in Table 5.

For corn prices, we can see that the joint EqCM is the best for $h = 4$.[11] In the case of $h = 1$, the lowest RMSE and MAPE values are obtained using DVAR, which also considers price interdependence.

For the price of soybeans, the joint EqCM produces forecasts that are better than those from the single

Table 4
In-sample quarterly joint EqCM estimation: 1994Q3–2007Q4 (in logs).

Dependent variable:	Δp^c	Δp^s
Constant	11.91^{***}	0.41^{*}
EqC_{it-1}	-0.36^{***}	-0.22^{***}
$\Delta_4 r_t$	-0.10^{***}	–
$\Delta q_t^{\text{ethanol}}$	0.25^{**}	–
$\Delta \text{gdp}_{t-1}^{\text{INDIA}}$	0.22^{***}	–
$\Delta \text{gdp}_t^{\text{CHINA}}$	–	0.77^{**}
Δp_t^c	–	0.39^{***}
Δp_{t-1}^s	–	0.21^{**}
Δp_{t-2}^s	–	-0.21^{**}
Δp_{t-5}^s	–	0.12
$\widehat{\sigma}$	0.054	0.042
Long-run solution Dependent variable:	p^c	p^s
q_{t-1}^i	-2.64^{***}	–
e_{t-1}	-3.68^{***}	–
$\text{gdp}_{t-1}^{\text{CHINA}}$	0.89^{***}	–
p_{t-1}^c	–	0.76^{***}
LR test[a]	20.78 [0.47] $\chi^2(21)$	

Notes: Insignificant variables remain due to autocorrelation. The selected (not reported) impulse dummies are: 1996Q2, 1996Q4, 2002Q3 and 2006Q4 for corn, 2001Q4, 2003Q4, 2007Q3 and 2007Q4 for soybeans. We allowed for a linear trend in the long-run, but it was not statistically significant.
*** $p < 0.01$.
** $p < 0.05$.
* $p < 0.10$.
[a] The LR test results correspond to the LR statistic for testing over-identifying restrictions, the associated tail probability (in square brackets) and the degrees of freedom (in parentheses).

equation model in all cases. The joint EqCM outperforms the other forecasts for $h = 1$ and $h = 4$.

In the case of wheat prices, the EqCM is the best for $h = 4$, but not for $h = 1$, for which the RMSE and MAPE are similar to those from the random walk.

11 Similar results are also found for long horizons in other empirical cases, e.g. Bårdsen, Eitrheim, Jansen, and Nymoen (2005, Ch. 11).

Table 5
1- and 4-step-ahead forecast evaluation over the period 2008Q1–2014Q4 (median level).

	Model type	$h = 1$				Model type	$h = 4$			
		Fixed scheme		Recursive scheme			Fixed scheme		Recursive scheme	
		RMSE	MAPE	RMSE	MAPE		RMSE	MAPE	RMSE	MAPE
	EqCM	32.10	12.30	30.80	11.43	EqCM (iterative)	56.01	24.43	55.99	24.18
	Joint EqCM	31.07	11.92	30.19	11.11	Joint EqCM (iterative)	**53.19**	**23.35**	**54.21**	**23.40**
	DEqCM	30.42	11.74	30.30	11.89	DEqCM (iterative)	70.83	27.20	–	–
	Joint DEqCM	30.43	11.79	30.38	11.92	Joint DEqCM (iterative)	67.03	25.08	–	–
Corn	DVAR	**28.64**	**10.48**	**29.17**	**10.87**	DVAR (iterative)	70.48	26.74	76.37	28.84
						EqCM (direct)	69.17	28.52	72.47	29.64
						DEqCM (direct)	75.28	28.18	77.87	22.41
						Joint EqCM (direct)	68.33	25.97	80.92	30.88
						Joint DEqCM (direct)	94.07	28.49	94.26	28.57
						DVAR (direct)	90.80	34.87	84.97	32.19
	Random walk	33.59	11.15	33.59	11.15	Random walk	72.79	26.16	72.79	26.16
	EqCM	71.02	11.40	57.05	9.14	EqCM (iterative)	179.23	31.61	130.43	23.74
	Joint EqCM	**49.33**	**6.69**	**52.86**	**7.53**	Joint EqCM (iterative)	**75.00**	**12.76**	**77.48**	**13.46**
	DEqCM	60.83	8.75	61.14	8.83	DEqCM (iterative)	255.34	43.89	–	–
	Joint DEqCM	68.68	10.01	67.35	9.72	Joint DEqCM (iterative)	85.58	13.38	–	–
Soybeans	DVAR	57.13	8.86	57.43	9.08	DVAR (iterative)	96.94	17.07	101.72	17.76
						EqCM (direct)	128.92	20.76	124.74	21.12
						DEqCM (direct)	130.50	21.57	130.04	21.23
						Joint EqCM (direct)	246.17	39.93	201.78	34.27
						Joint DEqCM (direct)	135.83	21.11	141.70	21.44
						DVAR (direct)	175.67	30.51	159.18	26.44
	Random walk	59.89	8.64	59.89	8.64	Random walk	120.38	18.54	120.38	18.54
	EqCM	40.19	10.53	**39.75**	10.82	EqCM (iterative)	**55.73**	**15.86**	**57.45**	**15.94**
	DEqCM	57.97	16.09	55.45	15.65	DEqCM (iterative)	64.02	19.63	–	–
Wheat	DVAR	41.06	11.04	41.36	11.01	DVAR (iterative)	81.86	25.12	84.69	25.80
						EqCM (direct)	74.40	18.91	72.87	18.56
						DEqCM (direct)	89.81	26.08	89.58	26.06
						DVAR (direct)	96.35	27.76	94.22	28.03
	Random walk	**39.98**	**10.35**	39.98	**10.35**	Random walk	95.93	26.04	95.93	26.04

Note: the numbers in bold correspond to the lowest RMSE and MAPE values.

In all three cases, EqCMs (single for wheat and joint for corn and soybeans) perform best for $h = 4$. Robust devices, such as double-differentiations, are usually a preferred solution to forecast failure, and dominate for the mean forecasting errors of EqCMs, but not necessarily for the variance or MSFE.

However, it is important to note that these results are valid on average for the whole out-of-sample period, and may differ over time. Focusing solely on the average performances of the models may result in a loss of information, and possibly incorrect forecast selection decisions. Therefore, the following subsections first analyse the forecast biases (as per Ericsson, forthcoming), then conduct a fluctuation test (as per Giacomini & Rossi, 2010), in order to evaluate the relative forecasting performances of the models with respect to the benchmark (the random walk).

6.2. Time-varying biases

To evaluate the forecasting performances of the models with the lowest RMSE and MAPE values further, we tried to determine whether or not they show forecast biases, and whether these potential forecast biases are systematic or time-varying.

Following Mincer and Zarnowitz (1969), we tested the forecast bias by regressing each forecast error term on an intercept and testing whether the intercept was statistically significant or not (as in Eq. (2) with $c_t = c$).

However, as the forecast bias may vary over time, Ericsson (forthcoming) proposed testing the time dependence of the forecast bias by regressing the forecast error term on the impulse indicator dummies; that is, a 1–0 dummy for each out-of-sample observation:

$$p_{t+h/t} - \widehat{p}_t^{(h)} = c_t + \varepsilon_t \tag{2}$$

$$= \sum_{i=1}^{T} c_i I_{it} + \varepsilon_t, \tag{3}$$

where the impulse indicator dummy I_{it} is one for $t = i$ and zero otherwise. In this context, a test that all coefficients c_i are jointly equal to zero is a test of forecast unbiasedness. As the number of coefficients to be estimated is equal to the number of observations, and therefore cannot be estimated directly, this can be achieved by using impulse indicator saturation (IIS at a 1% significance level), see Hendry, Johansen, and Santos (2008). We also perform "super saturation" (IIS+SIS at a 0.5% significance level), which searches across all possible one-step functions to capture permanent or long-lasting changes (for details of step indicator saturation, SIS, see Castle et al., 2015),[12] along with IIS. As Doornik, Hendry, and Petris (2013) showed in their Monte Carlo simulation, SIS improves the

[12] Commodity price shifts were studied using SIS by Mariscal and Powell (2014).

Table 6
Testing for bias in the forecast errors of corn prices.

Test (target size)	Model type	$h = 1$		Model type	$h = 4$	
		Fixed	Recursive		Fixed	Recursive
Mincer–Zarnowitz	DVAR	0.02 [0.89] $F(1, 27)$	0.13 [0.72] $F(1, 27)$	Joint EqC	5.15* [0.03] $F(1, 23)$	2.43 [0.13] $F(1, 23)$
IIS (1%)	DVAR	0.02 [0.89] (1, 27)	0.13 [0.72] (1, 27)	Joint EqC	16.02** [0.00right] $F(12, 12)$ I2010.3:I2013.2	22.56** [0.00] $F(14, 10)$ I2008.4:I2009.1, I2010.3:I2013.2
IIS+SIS (0.5%)	DVAR	0.02 [0.89] $F(1, 27)$	Joint EqC	11.44** [0.00] $F(1, 26)$ S2012.3	20.15** [0.00] $F(13, 11)$ I2010.3:I2013.2, S2009.1	45.38** [0.00] $F(11, 13)$ I2011.2,I2013.1, I2013.2,S2009.3, S2010.2,S2010.3, S2011.3,S2012.2, S2012.3,S2013.2, S2013.3

Notes: The entries within a given block of numbers are the F-statistic for testing the null hypothesis against the designated maintained hypothesis, the tail probability associated with the F-statistic (in square brackets), the degrees of freedom for the F-statistic (in parentheses), and, for IIS and IIS+SIS, the retained impulse (I) and step (S) dummies. An unrestricted constant is included in all cases, and we used HAC standard errors for $h = 4$.
* Indicate significance at the 5% level.
** Indicate significance at the 1% level.

Table 7
Testing for bias in the forecast errors of soybean prices.

Test (target size)	Model type	$h = 1$		$h = 4$	
		Fixed	Recursive	Fixed	Recursive
Mincer–Zarnowitz	Joint EqC	0.03 [0.86] $F(1, 27)$	0.55 [0.46] $F(1, 27)$	1.27 [0.27] $F(1, 23)$	0.92 [0.35] $F(1, 23)$
IIS (1%)	Joint EqC	18.92** [0.00] $F(4, 23)$ I2008.4, I2011.4, I2012.3, I2013.2	15.78** [0.00] $F(2, 25)$ I2008.4, I2012.3	5.58** [0.00] $F(8, 16)$ I2008.4,I2009.1,I2010.4, I2011.1,I2012.3:I2013.1: I2013.3	14.57** [0.00] $F(15, 9)$ I2008.4:I2009.4,I2010.4 I2011.1,I2011.4:I2012.4 I2013.2,I2013.3,I2014.3 I2014.4
IIS+SIS (0.5%)	Joint EqC	19.34** [0.00] $F(2, 25)$ I2008.4 I2012.3	17.83** [0.00] $F(1, 26)$ I2008.4	15.79** [0.00] $F(13, 11)$ I2009.1,I2010.4,I2011.1, I2011.4,I2012.1,I2012.3, I2012.4,I2013.2,I2013.3, I2014.4,S2009.1:S2009.3	6.78** [0.00] $F(10, 14)$ I2010.4,I2011.1,I2011.4 I2012.1,I2012.3,I2012.4 S2009.1,S2009.3,S2013.2 S2013.3

Notes: The entries within a given block of numbers are the F-statistic for testing the null hypothesis against the designated maintained hypothesis, the tail probability associated with the F-statistic (in square brackets), the degrees of freedom for the F-statistic (in parentheses), and, for IIS and IIS+SIS, the retained impulse (I) and step (S) dummies. An unrestricted constant is included in all cases, and we used HAC standard errors for $h = 4$.
** Indicate significance at the 1% level.
* Indicate significance at the 5% level.

non-null rejection frequency relative to the corresponding IIS-based test. However, it is important to note that IIS can also handle other types of breaks (blips, outliers and trending shifts), and therefore, the use of super saturation allows us to obtain different results which may also be useful for forecasting.

Tables 6–8 report the forecast bias tests just described for each commodity price. We can see that Mincer–Zarnowitz tests do not detect biases over the period 2008Q1–2014Q4 for soybean prices, or when using a recursive scheme, for corn. However, a systematic bias remains for wheat even when using a recursive scheme, which suggests that other updating schemes may be helpful.

Regarding biases over time, the start of the financial crisis of 2008–09 is detected as an outlier for the EqCM of soybeans and wheat for the shorter horizon ($h = 1$). However, if we extend the forecasting horizon to a year ($h = 4$), there are various impulse and step dummies which appear

Table 8
Testing for bias in the forecast errors of wheat prices.

Test (target size)	Model type	$h = 1$		$h = 4$	
		Fixed	Recursive	Fixed	Recursive
Mincer–Zarnowitz	EqC	4.37* [0.05] $F(1, 27)$	2.62 [0.12] $F(1, 27)$	10.43** [0.00] $F(1, 23)$	7.13** [0.01] $F(1, 23)$
IIS (1%)	EqC	12.61** [0.00] $F(1, 26)$ I2008.4	11.83** [0.00] $F(1, 26)$ I2008.4	10.43** [0.00] $F(1, 23)$	7.13** [0.01] $F(1, 23)$
IIS+SIS (0.5%)	EqC	12.61** [0.00] $F(1, 26)$ I2008.4	11.83** [0.00] $F(1, 26)$ I2008.4	58.88** [0.00] $F(14, 10)$ I2008.4,I2010.1,I2010.2, I2011.3,I2012.3,I2012.4 I2014.2,S2009.1,S2010.3 S2010.4,S2011.3,S2011.4 S2012.3,S2014.2	19.29** [0.00] $F(10, 14)$ I2010.2,I2010.4,I2011.3 I2012.3:I2013.1,I2014.2 S2009.1,S2010.4,S2011.3

Notes: The entries within a given block of numbers are the F-statistic for testing the null hypothesis against the designated maintained hypothesis, the tail probability associated with the F-statistic (in square brackets), the degrees of freedom for the F-statistic (in parentheses), and, for IIS and IIS+SIS, the retained impulse (I) and step (S) dummies. An unrestricted constant is included in all cases, and we used HAC standard errors for $h = 4$.
* Indicate significance at the 5% level.
** Indicate significance at the 1% level.

to be highly statistically significant that may be associated with the crisis, its aftermath, and the recent reversion of the upward trend in commodities prices.

6.3. *Time-varying forecasting ability*

The results in the previous subsection showed that there are time-varying biases in the forecast models selected based on the lowest RMSE and MAPE values. Given the instability of the out-of-sample period, it is also worth evaluating the evolution of the models' relative performances. We therefore apply the fluctuation test developed by Giacomini and Rossi (2010) to evaluate the local relative forecasting performances of these models against the random walk as a benchmark, for $h = 4$.

We define the local relative loss for the two models (the forecast model and the random walk) as the sequence of out-of-sample loss differences over centered rolling windows of size m (in our case, $m = 5$):

$$m^{-1} \sum_{j=t-m/2}^{t+m/2-1} \Delta L_j(\hat{\theta}_{j-h,T}, \hat{\gamma}_{j-h,T}),$$
$$t = T + h + m/2, \ldots, T^{*} - m/2 + 1. \tag{4}$$

Under the null hypothesis, $E[\Delta L_t(\hat{\theta}_{t-h,T}, \hat{\gamma}_{t-h,T})] = 0$ for all $t = T + h, \ldots, T^{*}$, and the test statistic $F^{OOS}_{t,m}$ is

$$F^{OOS}_{t,m} = \hat{\sigma}^{-1} m^{-1/2} \sum_{j=t-m/2}^{t+m/2-1} \Delta L_j(\hat{\theta}_{j-h,T}, \hat{\gamma}_{j-h}, T), \tag{5}$$

for $t = T + h + m/2, \ldots, T - m/2 + 1$, where $\hat{\sigma}^2$ is a HAC estimator of σ^2.[13]

Fig. 3 shows the results of the fluctuation test for the EqCM of corn, soybean and wheat prices against the random walk. In each case, the graph reports both the fluctuation test statistic (constructed using a centered moving window) and the one-sided critical value at the 5% level (the constant line). Positive values of the test statistic indicate that the model with fundamentals is better than the random walk.

Overall, the results of the fluctuation test indicate that models with fundamentals perform better than the random walk during a large part of the out-of-sample period. However, the forecast gains are time-varying, with the EqCM for wheat showing a weaker result by the end of the sample.

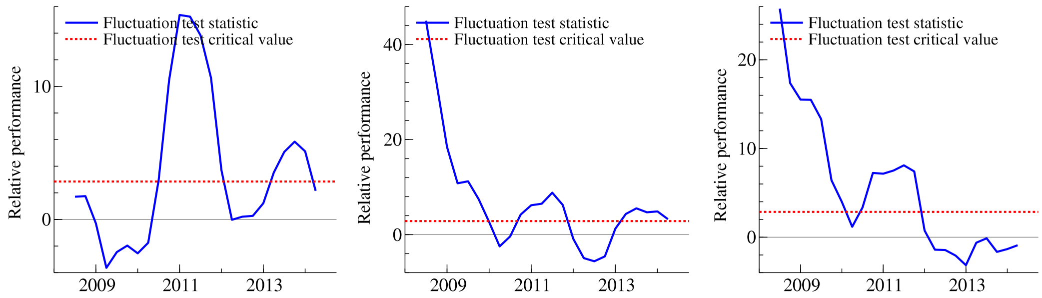
7. Final remarks

We have analysed whether the forecasting accuracies of individual food price models can be improved by taking into account the cross-dependence of the commodities studied. Our out-of-sample period was quite unstable, and therefore the effects of potential breaks were dealt with by robust approaches and recursive estimations.

Overall, we found forecast improvements from using models that take into account the interactions of the prices, that is, joint EqCMs and DVARs.

The evaluation of the non-constancy of the biases shows that our forecasting models considered the latest world crisis as an outlier rather than a break when forecasting one quarter ahead. However, when we consider a longer forecasting horizon, there is evidence of time-varying biases which would require the use of other updating schemes, particularly at the end of the sample. Ongoing studies of these models suggest that it may be helpful to re-select the explanatory variables that enter as short-run determinants. Multiplicative indicator saturation may also help to capture changes in the slope coefficients.

[13] The HAC standard errors are as per Andrews (1991).



(a) Corn (h=4). (b) Soybeans (h=4). (c) Wheat (h=4).

Fig. 3. Fluctuation test results (EqCM vs. random walk). Note: Positive values of the fluctuation statistic indicate that the EqCM is better than the random walk.

Table A.1
Data description.

Symbol	Description	Units	Source
P^s	Nominal soybean price	US dollars per metric ton	Pink Sheet, World Bank
P^c	Nominal corn price	US dollars per metric ton	Pink Sheet, World Bank
P^w	Nominal wheat price	US dollars per metric ton	Pink Sheet, World Bank
Q^s	World soybean production	Millions of metric tons	USDA
Q^c	World corn production	Millions of metric tons	USDA
Q^w	World wheat production	Millions of metric tons	USDA
INV^s	World soybean inventories	Millions of metric tons	USDA
INV^c	World corn inventories	Millions of metric tons	USDA
INV^w	World wheat inventories	Millions of metric tons	USDA
$\text{GDP}^{\text{CHINA}}$	China's real GDP	Billions of yuans	National Bureau Statistics of China
$\text{GDP}^{\text{INDIA}}$	India's real GDP	index 2005 = 100	IMF
GDP^{OECD}	OECD's real GDP	Millions of US dollars	OECD Statistics
Q^{ethanol}	U.S. ethanol production	Millions of gallons	U.S. EIA
R	3-month treasury constant maturity rate	Percentage	Federal Reserve Board
E	US real exchange rate	index	Federal Reserve Board
$M2$	M2, real monetary aggregate	Billions of US dollars	Federal Reserve Board
MB	real monetary base	Billions of US dollars	Federal Reserve Board
FOF	Fed's flow of funds[a]	Billions of US dollars	Federal Reserve Board
CPI^{US}	US consumer price index	index	IMF

[a] The FRED series Total Credit Market Debt Owed is now known as All Sectors; Credit Market Instruments; Liability.

In regard to relative forecasting abilities in such an unstable forecasting period, the models with fundamentals performed better than the random walk for a large part of the out-of-sample period.

The use of forecasting models that are based on fundamentals allows us to analyse the ways in which different scenarios for the main determinants would influence food prices, and also to determine whether examining the cross-dependence among corn, soybean and wheat prices is beneficial for forecasting purposes. A similar approach may be applied to a different set of commodities in future research.

Acknowledgments

Earlier versions of this paper were presented at the Econometrics Conference EMoD in Oxford, the XLIX Annual Meeting of the Argentine Economic Association, and the XX Latin American and Caribbean Economic Association Meeting. We would like to thank those who participated in these meetings, Andrew Powell and Fernando Navajas, as well as two anonymous referees, for many valuable comments and suggestions. The usual disclaimer applies.

Appendix A. Data definitions and sources

See Table A.1.

Appendix B. Direct multi-step estimations

We considered the same long-run relationship, but lagged four periods. For soybeans, we considered the cointegration relationship between soybean and corn prices. In this case, soybeans adjust 100% to deviations from the steady-state in the first period. See Table B.1.

Appendix C. Supplementary data

Supplementary material related to this article can be found online at http://dx.doi.org/10.1016/j.ijforecast.2016.01.002.

Table B.1
Direct in-sample quarterly estimations: 1994Q3–2007Q4.

Dependent variable:	$\Delta 4p^c$	$\Delta 4p^s$	$\Delta 4p^w$
Constant	17.51**	2.68**	3.76**
EqC_{it-4}	−0.53**	−1.00**	−0.72**
$\Delta 4p^c_{t-8}$	−0.51**	–	–
$\Delta q^{\text{ethanol}}_{t-4}$	0.71**	–	–
$\Delta q^{\text{ethanol}}_{t-5}$	0.58**	–	–
$\Delta q^{\text{ethanol}}_{t-8}$	−0.65**	–	–
$\Delta \text{gdp}^{\text{CHINA}}_{t-4}$	0.46**	–	–
$\Delta \text{gdp}^{\text{CHINA}}_{t-5}$	1.10**	–	–
$\Delta \text{gdp}^{\text{CHINA}}_{t-8}$	6.81**	–	–
$\Delta_4 r_{t-4}$	−0.13**	–	–
Δe_{t-4}	–	−3.03**	–
Δe_{t-5}	–	−2.38**	–
Δe_{t-6}	–	−2.05**	–
$\widehat{\sigma}$	0.13	0.13	0.14

** Indicate significance at the 1% level.
* Indicate significance at the 5% level.

References

Ahumada, H., & Cornejo, M. (2015). Explaining commodity prices by a cointegrated time series-cross section model. *Empirical Economics*, *48*, 1667–1690.

Andrews, D. W. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica: Journal of the Econometric Society*, *59*, 817–858.

Bårdsen, G., Eitrheim, O., Jansen, E. S., & Nymoen, R. (2005). *The econometrics of macroeconomic modelling*. Oxford: Oxford University Press.

Castle, J. L., Doornik, J. A., Hendry, D. F., & Petris, F. (2015). Detecting location shifts during model selection by step-indicator saturation. *Econometrics*, *3*, 240–264.

Castle, J. L., Fawcett, N. W. P., & Hendry, D. F. (2010). Forecasting with equilibrium-correction models during structural breaks. *Journal of Econometrics*, *158*, 25–36.

Chevillon, G. (2007). Direct multi-step estimation and forecasting. *J. Economic Surveys*, *21*, 746–785.

Clements, M. P., & Hendry, D. F. (1996). Multi-step estimation for forecasting. *Oxford Bulletin of Economics and Statistics*, *58*, 657–684.

Clements, M. P., & Hendry, D. F. (1998). *Forecasting economic time series*. Cambridge University Press.

Clements, M. P., & Hendry, D. F. (1999). *Forecasting non-stationary economic time series*. Cambridge, MA: MIT Press.

Clements, M. P., & Hendry, D. F. (2006). Forecasting with breaks. In G. Elliott, C. W. J. Granger, & A. Timmermann (Eds.), *Handbook of economic forecasting*. Amsterdam: Elsevier.

Davidson, J. (1998). Structural relations, cointegration and identification: some simple results and their application. *Journal of Econometrics*, *87*, 87–113.

Doornik, J. A. (2009). Autometrics. In J. L. Castle, & N. Shephard (Eds.), *The methodology and practice of econometrics: a festschrift in honour of David F. Hendry*. Oxford: Oxford University Press.

Doornik, J. A., Hendry, D. F., & Petris, F. (2013). Step-indicator saturation. Draft, University of Oxford.

Engle, R. F., & Granger, C. W. (1987). Co-integration and error correction: representation, estimation, and testing. *Econometrica: Journal of the Econometric Society*, *55*, 251–276.

Ericsson, N. R. (2016). How biased are US government forecasts of the federal debt? *International Journal of Forecasting*, (forthcoming).

Ericsson, N. R., & MacKinnon, J. G. (2002). Distributions of error correction tests for cointegration. *The Econometrics Journal*, *5*, 285–318.

Giacomini, R., & Rossi, B. (2010). Forecast comparisons in unstable environments. *Journal of Applied Econometrics*, *25*, 595–620.

Hendry, D. F. (2006). Robustifying forecasts from equilibrium-correction systems. *Journal of Econometrics*, *135*, 399–426.

Hendry, D. F., & Doornik, J. A. (2014). *Empirical model discovery and theory evaluation: automatic selection methods in econometrics*. Cambridge: MIT Press.

Hendry, D. F., Johansen, S., & Santos, C. (2008). Automatic selection of indicators in a fully saturated regression. *Computational Statistics*, *23*, 317–335.

Johansen, S. (1988). Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control*, *12*, 231–254.

Johansen, S. (1996). *Likelihood-based inference in cointegrated vector autoregressive models*. Oxford: Oxford University Press.

Juselius, K. (2006). *The cointegrated VAR model: methodology and applications*. Oxford: Oxford University Press.

Makridakis, S., & Hibon, M. (2000). The M3-Competition: results, conclusions and implications. *International Journal of Forecasting*, *16*, 451–476.

Mariscal, R., & Powell, A. (2014), Commodity price booms and breaks: detection, magnitude and implications for developing countries. IDB Working Paper No. IDB-WP-444.

Mincer, J. A., & Zarnowitz, V. (1969). The evaluation of economic forecasts. In *Economic forecasts and expectations: analysis of forecasting behavior and performance* (pp. 1–46). NBER.

Stock, J. H., & Watson, M. W. (1999). Forecasting inflation. *Journal of Monetary Economics*, *44*, 293–335.

Hildegart Ahumada holds a M.Phil. in Economics from the University of Oxford. She is currently Professor of Econometrics at the School of Government, Di Tella University. Member of the Argentinian Academy of Economic Sciences, fellow of the John S. Guggenheim Memorial Foundation, former Head of the Economic Research Department at the Central Bank of Argentina and Past President of the Argentine Economic Association.

Magdalena Cornejo holds a Master Degree in Econometrics from Di Tella University. She is currently a Ph.D. (candidate) in Economics at Buenos Aires University.