FUNDAÇÃO GETULIO VARGAS

EPGE

Escola de Pós-Graduação em Economia

Ensaios Econômicos

Escola de
Pós-Graduação
em Economia
da Fundação
Getulio Vargas

N° 744 ISSN 0104-8910

Using Common Features to Understand the Behavior of Metal-Commodity Prices and Forecast them at Di?erent Horizons

João Victor Issler, Claudia Rodrigues, Rafael Burjack

Agosto de 2013

URL: http://hdl.handle.net/10438/11176

Victor Issler, João
Using Common Features to Understand the Behavior of Metal-Commodity Prices and Forecast them at Di?erent Horizons/ João Victor Issler, Claudia Rodrigues, Rafael Burjack - Rio de Janeiro : FGV,EPGE, 2013
52p. - (Ensaios Econômicos; 744)

Inclui bibliografia.

CDD-330

Using Common Features to Understand the Behavior of Metal-Commodity Prices and Forecast them at Different Horizons*

João Victor Issler[†] Claudia Rodrigues[‡] Rafael Burjack[§]

August 22, 2013

Abstract

The objective of this article is to study (*understand* and *forecast*) spot metal price levels and changes at monthly, quarterly, and annual frequencies. Data consists of metal-commodity prices at a monthly and quarterly frequencies from 1957 to 2012, extracted from the IFS, and annual data, provided from 1900-2010 by the U.S. Geological Survey (USGS). We also employ the (relatively large) list of co-variates used in Welch and Goyal (2008) and in Hong and Yogo (2009).

*We gratefully acknowledge the comments and suggestions of Roberto Castello Branco, Robert Engle, Afonso Arinos de Mello Franco Neto, Allan Timmermann, the Editor in charge (Rabah Arezki), five anonymous referees, and seminar participants at the 2013 IMF seminar on *Understanding International Commodity Price Fluctuations* and the 2012 VALE-EPGE conference on *The Economics and Econometrics of Commodity Prices*. All remaining errors are ours. Issler thanks CNPq, FAPERJ and INCT for financial support. We thank Marcia Waleria Machado and Marcia Marcos for excellent research assistance.

[†]Graduate School of Economics - EPGE, Getulio Vargas Foundation. email: joao.issler@fgv.br

[‡]Investor Relations Department, Vale. email: claudia.rodrigues@vale.com

[§]Graduate School of Economics - EPGE, Getulio Vargas Foundation email: rburjack@fgvmail.br

We investigate short- and long-run comovement by applying the techniques and the tests proposed in the *common-feature* literature. One of the main contributions of this paper is to *understand* the short-run dynamics of metal prices. We show theoretically that there must be a positive correlation between metal-price variation and industrial-production variation if metal supply is held fixed in the short run when demand is optimally chosen taking into account optimal production for the industrial sector. This is simply a consequence of the derived-demand model for cost-minimizing firms. Our empirical evidence fully supports this theoretical result, with overwhelming evidence that cycles in metal prices are synchronized with those in industrial production. This evidence is stronger regarding the global economy but holds as well for the U.S. economy to a lesser degree.

Regarding out-of-sample forecasts, our main contribution is to show the benefits of forecast-combination techniques, which outperform individual-model forecasts – including the random-walk model. We use a variety of models (linear and non-linear, single equation and multivariate) and a variety of covariates and functional forms to forecast the returns and prices of metal commodities. Using a large number of models (N large) and a large number of time periods (T large), we apply the techniques put forth by the *common-feature* literature on forecast combinations. Empirically, we show that models incorporating (short-run) common-cycle restrictions perform better than unrestricted models, with an important role for industrial production as a predictor for metal-price variation.

1 Introduction

The purpose of this paper is twofold. The first is to improve our understanding of metal-commodity price variation either in the long run or in the short run by using standard time-series techniques. We rely on the *common-trend and common-cycle* approach put forward by Engle and Kozicki (1993), Vahid and Engle (1993, 1997), Engle and Issler (1995), Issler and Vahid (2001, 2006), Vahid and Issler (2002), Hecq et al. (2006), and Athanasopoulos et al. (2011). Here, non-stationary economic series are decomposed into an integrated trend component and a stationary and ergodic cyclical component, where their properties can be jointly investigated in a unified multivariate setting based on vector autoregressive (VAR) models. Trends and cycles can be common to a group of series being modelled, and these *common features* can be removed by independent linear combination[1]. Our second objective is to improve on current forecasts of metal-commodity prices taking into account the recent financialization of commodity markets and the role of information in commodity markets; see Hong and Yogo (2009, 2012) and Gargano and Timmermann (2012). Instead of relying on a specific model to forecast metal-commodity prices, we diversify out the risk of large forecast errors (and increase the information set used in forecasting) by combining forecasts of different models. This approach, first put forward by Bates and Granger (1969), has been shown to reduce forecast uncertainty in a variety of studies; see Hendry and Clements (2004) and Stock and Watson (2006). Recently, Issler and Lima (2009) have developed an optimal forecast-combination in a panel-data setting, where forecasts of different models (or survey results) comprise the cross-sectional dimension. In their context, the optimal forecast using a mean-squared error (MSE) risk function can be consistently estimated employing the bias-corrected average forecast (BCAF), which is a *common feature* of all forecast models.

Early modern empirical work on commodity prices focused on the behavior of trend prices – Cuddington and Urzúa (1989) and Cuddington (1992). Trends are modelled as martingale processes. As (Deaton, 1999, p. 27) puts it, referring to the drift term in commodity prices: "what commodity prices lack in trend, they make up for in variance." Cashin et al. (2002) summarize the "stylized facts about real commodity prices: they are often dominated by long periods of doldrums punctuated by sharp upward spikes (Deaton and Laroque (1992)); they have a tendency to trend down in the long run (Grilli and Yang (1988)); shocks to commodity prices tend to persist for several years at a time (Cashin et al. (2000)); and unrelated commodity prices move together (Pindyck and Rotemberg (1990))." Compared to the interest on the trends of commodity prices, little work has been done on cycles,

[1] Perhaps cointegration is the best-known example of *common features*.

the early exceptions being Labys et al. (1999), Cashin et al. (1999), and Pindyck and Rotemberg (1990).

Recently, however, there has been a renewed interest on commodity-price cycles, see Jerrett and Cuddington (2008) and IMF (2012). Our paper complements this effort. One of its main contributions is to *understand* the short-run dynamics of metal prices. We show, theoretically, that there must be a positive correlation between metal-price variation and industrial-production variation if metal supply is held fixed when demand is optimally chosen taking into account optimal production for the industrial sector. This is simply a consequence of the derived-demand model for cost-minimizing firms. The details of this models are given in Section 2. Our empirical evidence in Section 5 (monthly and quarterly data) fully supports this theoretical result, with overwhelming evidence that cycles in metal prices are synchronized with those in industrial production. This evidence is stronger regarding the global economy but holds as well for the U.S. economy to a lesser degree. As far as we know, we were the first authors to investigate and find common cycles in this way, accounting for theory and empirics, and not just describing a stylized fact[2].

Our second contribution is in *forecasting* metal prices at different horizons. In doing so, we try to incorporate the overwhelming evidence found on common cycles between metal prices and industrial production. One of the advantages of the *common-trend and common cycle* method is parsimony, with obvious benefits for building efficient forecasting models; see Issler and Vahid (2001), Vahid and Issler (2002), and Athanasopoulos et al. (2011). As argued by Vahid and Issler (2002), vector autoregressions (VARs) have been increasingly used in multivariate analysis and in forecasting economic data. One of their shortcomings is the excessive number of parameters. For example, a $VAR(p)$ for n series has $n^2 \cdot p$ parameters in the conditional mean. One can easily see the burden on degrees of freedom if the number of series being modelled (n) is large. Cointegration certainly reduces the number of parameters, but these reductions are mild. On the other hand, short-run

[2]Obviuosly, we were not firsts to investigate the cyclical behavior of metal prices (Jerrett and Cuddington, 2008). Nor were we the firsts to emphasize the importance of demand factors (Deaton and Laroque, 1996).

restrictions – or common cycles – have a much greater potential to reduce the number of parameters in the dynamic representation. For example, when dealing with post-war quarterly data, and a VAR with three variables and eight lags, there are seventy five mean parameters to be estimated from about two hundred data points on each variable. If the three-variable system has one known cointegrating vector, the number of free parameters falls from seventy five to sixty nine when estimating a vector error-correction model – VECM. *Common-cyclical features* show more potential in reducing the number of conditional-mean parameters. If the three variables in the VECM share one common cycle, then the number of mean parameters falls from sixty nine to twenty seven.

Using efficient models in forecasting metal prices is of obvious interest. However, most models are misspecified, and it has been largely documented that the average forecast using several models outperforms individual models themselves; see Hendry and Clements (2004). Hence, we apply forecast-combination methods to forecast metal prices, showing that they work in practice. We go one step beyond, resorting to a *common-feature* technique proposed by Issler and Lima (2009).

As argued by Issler and Lima (2009), Bates and Granger (1969) made the econometrics profession aware of the benefits of forecast combination when a limited number of forecasts is considered. The widespread use of different combination techniques has lead to an interesting puzzle from the econometrics point of view – the *forecast combination puzzle*: if we consider a fixed number of forecasts ($N < \infty$), combining them using equal weights ($1/N$) fare better than using "optimal weights" constructed to outperform any other forecast combination in the mean-squared error (MSE) sense. Regardless of how one combine forecasts, if the series being forecast is stationary and ergodic, and there is enough diversification among forecasts, we should expect that a weak law-of-large-numbers (WLLN) applies to well-behaved forecast combinations. This argument was considered in Palm and Zellner (1992) who asked the question "*to pool or not to pool*" forecasts? Timmermann (2006) used risk diversification – a principle so keen to finance – to defend pooling of forecasts. Of course, to obtain this WLLN result, at least the number of forecasts has to diverge ($N \to \infty$), which entails the use of asymptotic panel-data techniques. This is exactly the approach in

Issler and Lima (2009), with the added twist that now $N, T \to \infty$, with $T \to \infty$ prior than N: the sequential asymptotic approach developed by Phillips and Moon (1999), denoted by $(T, N \to \infty)_{seq}$.

Forecast combination works well in practice because of risk diversification: idiosyncratic forecast errors vanish because of the WLLN works as the number of forecasts being combined increases without bounds. However, the forecast combination puzzle also works against forecast combinations because of the *curse of dimensionality*: as N increases, if one has to estimate "optimal weights" to combine forecasts with a fixed number of observations, the estimates of these weights are inconsistent. Issler and Lima solve the curse of dimensionality by imposing equal weights that need not be estimated $(1/N)$, and perform bias correction to take MSE down to its minimum, identifying, in the limit, the conditional expectation of the series being forecast: if y_t is the series being forecast, and h is the horizon, then, what is being identified is the latent variable $\mathbb{E}_{t-h}(y_t)$, where $\mathbb{E}_{t-h}(\cdot)$ is the conditional expectation operator using all information available (observable or not) up to period $t-h$. Here, we are able to expand the information content of every individual model.

The paper is divided as follows: Section 2 presents a theoretical model that delivers common cycles among metal prices and industrial output. Sections 3 and 4 summarize the econometric techniques employed here, while the empirical results are reported in detail in Section 5. Section 6 concludes.

2 Understanding the Fluctuations of Metal-Commodity Prices

From a theoretical point-of-view, commodity-price dynamics have been studied at least since Newbery and Stiglitz (1981), Deaton and Laroque (1992, 1996) and Chambers and Bailey (1996). Early modern empirical work on commodity prices has focused on the behavior of trend prices – Cuddington and Urzúa (1989) and Cuddington (1992) – where trends were modelled as martingale processes.

Despite the proliferation of trend studies, the literature on the cyclical fluctuations of commodity prices has not been so prolific. Indeed, Jerrett and Cuddington (2008) note that "authors have analyzed the movement of metal prices over the business cycle as well as comovements among commodity prices (see Labys et al. (1999); Cashin et al. (1999); and Pindyck and Rotemberg (1990))." Regarding cycles, Deaton and Laroque (1996) is an important paper emphasizing the importance of demand shocks for short-run fluctuations. As they put it, "it is likely that demand shocks are a more plausible source of price fluctuations than has usually been supposed in the literature[3]."

We also argue here that there is an important role for demand shocks in explaining the short-run variation of metal-commodity prices. Indeed, the overwhelming empirical evidence below suggests that the short-run fluctuations of metal-commodity prices are synchronized with those of industrial production in a global scale. To a lesser degree, they are also synchronized with U.S. industrial production. In trying to understand how these stylized facts come about, we devise a simple theoretical model motivated by the fact that metal commodities are inputs in industrial production processes, which generates a *derived demand* for metal commodities.

Consider a representative industrial firm, which chooses the optimal quantity of inputs x_i, $i = 1, 2, \cdots, n$, all stacked in a vector $x = (x_1, x_2, \cdots, x_n)'$, when producing output y_0. The choice of output y_0 can be thought as an optimal decision coming from the firm's output market. The corresponding prices for inputs $i = 1, 2, \cdots, n$, stacked in a vector $w = (w_1, w_2, \cdots, w_n)'$, are considered given for the firm when choosing x. The firm's cost minimization problem in this context is:

$$\min_{x} \quad C(w, x) = w \cdot x \qquad s.t. \qquad f(x) \geq y_0. \tag{1}$$

From the first-order (interior) condition of this problem, using Shepard's Lemma, we derive the optimal derived demands for all inputs, labelled $x_i^*(w, y_0)$:

$$\frac{\partial C(w, x^*)}{\partial w_i} = x_i^*(w, y_0), \qquad i = 1, 2, \cdots, n. \tag{2}$$

[3]Recently, there has been a renewed interest on commodity-price cycles, see Jerrett and Cuddington (2008) and IMF (2012).

A critical issue in describing the equilibrium for input markets is how to model supply. Of course, this depends on the horizon at which markets are supposed to clear. In modelling short-run fluctuations, it is reasonable to assume that metal-commodity supply cannot be increased without climbing a very steep cost function. Thus, we treat supply as fixed ($\overline{x_i}$) in the short run. This assumption is consistent with the fact that mining projects are very intensive in capital and take a long time to mature. Since capital is traditionally held fixed in short-run analysis, this is similar to fixed short-run supply. Even if one considers the existence of inventories, they also cannot change in quantity in the short run. Indeed, we can think as the inventories as part of this fixed supply ($\overline{x_i}$) for metal commodities.

Thus, the short-run equilibrium condition for inputs (including metal commodities) is:

$$x_i^*(w, y_0) = \overline{x_i}. \qquad i = 1, 2, \cdots, n. \tag{3}$$

Ceteris paribus, given the equilibrium condition (3), we investigate how changes in output potentially change the price of input i, w_i, considered here to be a metal commodity used in production. Totally differentiate (3) considering only changes in w_i and in industrial production, y_0, later solving for $\frac{dw_i}{dy_0}$:

$$\begin{aligned} 0 &= \frac{\partial x_i^*(w, y_0)}{\partial w_i} dw_i + \frac{\partial x_i^*(w, y_0)}{\partial y_0} dy_0, \text{ or,} \\ \frac{dw_i}{dy_0} &= -\frac{\frac{\partial x_i^*(w,y_0)}{\partial y_0}}{\frac{\partial x_i^*(w,y_0)}{\partial w_i}}. \end{aligned} \tag{4}$$

It is straightforward to establish unequivocally that $\frac{dw_i}{dy_0} > 0$ since, from theory, we should have $\frac{\partial x_i^*(w,y_0)}{\partial y_0} > 0$ and $\frac{\partial x_i^*(w,y_0)}{\partial w_i} < 0$. This result $\left(\frac{dw_i}{dy_0} > 0\right)$ is completely intuitive: given concavity of the cost function vis-a-vis input prices $\left(\frac{\partial x_i^*(w,y_0)}{\partial w_i} = \frac{\partial^2 C(w,x^*)}{\partial w_i^2} < 0\right)$, if the representative firm wants to increase industrial production in the short run, it will put an upward pressure in the metal-commodity market, stemming from the fact that it should take more inputs to produce more $\left(\frac{\partial x_i^*(w,y_0)}{\partial y_0} > 0\right)$, otherwise it is not a cost minimizer.

In this setup, changes in industrial production have a positive correlation with changes in metal-commodity prices. Of course, this does not imply that these fluctuations will be synchronized, but that is the object of the empirical investigation in Section 5 below. It should also be noted that, as the equilibrium horizon becomes larger, supply cannot be treated as fixed, which reduces the importance of demand factors.

3 Cointegration and Common Cycles for Metal Prices

We discuss here a unified econometric framework that allows investigating the existence of short- and long-run restrictions for metal-commodity prices. An in-depth theoretical discussion of these issues can be found in Engle and Granger (1987), Vahid and Engle (1993), Vahid and Engle (1997), Hecq et al. (2006), and Athanasopoulos et al. (2011).

Assume that y_t is a n-vector of $I(1)$ metal prices[4] (or log metal prices), which can be represented by a vector autoregression (VAR) model in levels:

$$y_t = \Gamma_1 y_{t-1} + \ldots + \Gamma_p y_{t-p} + \epsilon_t. \tag{5}$$

If elements of y_t cointegrate, Engle and Granger (1987) showed that the system (5) can be written as a Vector Error-Correction model (VECM):

$$\Delta y_t = \Gamma_1^* \Delta y_{t-1} + \ \ldots \ + \Gamma_{p-1}^* \Delta y_{t-p+1} + \gamma\alpha' y_{t-1} + \epsilon_t \tag{6}$$

where γ and α are full rank matrices of order $n \times r$, r is the rank of the cointegrating space, $-\left(I - \sum_{i=1}^{p} \Gamma_i\right) = \gamma\alpha'$, and $\Gamma_j^* = -\sum_{i=j+1}^{p} \Gamma_i$, $j = 1, \ldots, p-1$.

For our purposes, testing for cointegration will be used to verify whether metal-price data share common trends (or have long-run comovement). Testing for common trends among y_t will use the maximum-likelihood approach in Johansen (1991). A

[4]Other variables of interest may also be jointly modeled with metal-commodity prices, e.g., industrial production and other co-variates that could help explaining their behavior.

key issue to assure that inference is done properly is to estimate the lag length of the VAR (5) consistently, i.e., to estimate p consistently. Athanasopoulos et al. discuss how this can be achieved by using a combination of information criteria. An alternative to way to infer p is to perform diagnostic testing to rule out the risk of underestimation of p, which leads to inconsistent estimates for the parameters in (6).

Vahid and Engle (1993) show that the dynamic representation for y_t (6) may be restricted if there exist white noise independent linear combinations of the series Δy_t, i.e., if the y_ts share *common cycles*. These white noise linear combinations of the series Δy_t can be expressed using cofeature vectors $\tilde{\alpha}_i'$, stacked in an $s \times n$ matrix $\tilde{\alpha}'$, which eliminate all serial correlation in Δy_t. Thus, $\tilde{\alpha}' \Delta y_t = \tilde{\alpha}' \epsilon_t$. This is what Hecq, Palm and Urbain (2006) have labelled *strong-form serial-correlation common features*:

$$\tilde{\alpha}'\Gamma_1^* = \tilde{\alpha}'\Gamma_2^* = \ldots = \tilde{\alpha}'\Gamma_{p-1}^* = 0, \text{ and} \quad (7)$$

$$\tilde{\alpha}'\gamma = 0. \quad (8)$$

If we only impose restrictions (7), but not (8), we obtain what they have labelled *weak-form serial-correlation common features*: $\tilde{\alpha}' [\Delta y_t - \gamma \alpha' y_{t-1}] = \tilde{\alpha}' \epsilon_t$, i.e., we only inherit an unpredictable linear combination of Δy_t once we *control* for the long-run deviations $\alpha' y_{t-1}$ stemming from cointegration.

We continue the discussion of common cycles in the case of strong-form serial-correlation common features ((7) and (8)), given that the weak-form case can be immediately inferred from it[5]. Since cofeature vectors are identified only up to an invertible transformation, without loss of generality, we can consider $\tilde{\alpha}$ to be of the form:

$$\tilde{\alpha} = \begin{bmatrix} I_s \\ \tilde{\alpha}^*_{(n-s)\times s} \end{bmatrix}$$

Completing the system by adding the unconstrained VECM equations for the re-

[5] The Appendix contains a more complete discussion.

maining $n-s$ elements of Δy_t, we obtain a *quasi-structural* model,

$$\begin{bmatrix} I_s & \tilde{\alpha}^{*\prime} \\ \underset{(n-s)\times s}{\mathbf{0}} & I_{n-s} \end{bmatrix} \Delta y_t = \begin{bmatrix} \underset{s\times(np+r)}{\mathbf{0}} \\ \Gamma_1^{**} \ \ldots \ \Gamma_{p-1}^{**} \ \gamma^* \end{bmatrix} \begin{bmatrix} \Delta y_{t-1} \\ \vdots \\ \Delta y_{t-p+1} \\ \alpha' y_{t-1} \end{bmatrix} + v_t. \tag{9}$$

Since $\begin{bmatrix} I_s & \tilde{\alpha}^{*\prime} \\ \underset{(n-s)\times s}{\mathbf{0}} & I_{n-s} \end{bmatrix}$ is always invertible, we can recover (6) from (9). However, that the latter has $s \cdot (np+r) - s \cdot (n-s)$ fewer parameters, thus, being over-identified.

The literature on common cycles proposes estimation of the system in (9) in two different ways. The first is to employ full-information maximum likelihood (FIML), constructing the likelihood function exploiting the correlation among the errors v_t. The other is to employ the generalized method of moments (GMM), exploiting the fact that the errors v_t are orthogonal to the regressors in (9). Notice that this includes the first s errors in v_t, which come from the white-noise combinations using $\tilde{\alpha}$. Analogously, testing for the existence of s cofeature vectors – vectors leading to s linearly independent white noise combinations of the elements in Δy_t – can be done by canonical-correlation analysis (likelihood based) or by over-identifying-restriction tests (GMM based).

In testing for the existence of s serial-correlation common features (SCCF), by means of canonical-correlation analysis, the null hypothesis is that the first smallest s canonical correlations are jointly zero and the test statistic is $-T\sum_{i=1}^{s}\log(1-\lambda_i)$, where λ_i, $i = 1, \cdots, n$, are the sample squared canonical correlations between $\{\Delta y_t\}$ and $\{\alpha' y_{t-1}, \Delta y_{t-1}, \Delta y_{t-2}, \cdots, \Delta y_{t-p+1}\}$. The limiting distribution of this test statistic is χ^2 with $s(np+r) - s(n-s)$ degrees of freedom.

One possible drawback of the canonical-correlation approach is that it assumes homoskedastic data, and that may not hold for metal-price (and other macroeconomic and financial data) collected at high frequency. In this case, a GMM approach is more robust, since inference can be conducted with Heteroskedastic and Auto-Correlation (HAC) robust estimates of the variance-covariance matrices of parameter estimates. The vector of instruments comprise the series in $\alpha' y_{t-1}$, $\Delta\, y_{t-1}$, $\Delta\, y_{t-2}, \cdots, \Delta\, y_{t-p+1}$, collected in a vector Z_{t-1}. GMM estimation and testing exploits the following moment restriction:

$$\begin{aligned}\mathbf{0} &= \mathbb{E}\left[v_t \otimes Z_{t-1}\right] = \qquad (10)\\ &= \mathbb{E}\left[\left(\left[\begin{array}{cc} I_s & \tilde{\alpha}^{*\prime} \\ \underset{(n-s)\times s}{\mathbf{0}} & I_{n-s}\end{array}\right]\Delta\, y_t - \left[\begin{array}{c}\underset{s\times(np+r)}{\mathbf{0}} \\ \Gamma_1^{**} \; \ldots \; \Gamma_{p-1}^{**} \; \gamma^*\end{array}\right]\left[\begin{array}{c}\Delta\, y_{t-1} \\ \vdots \\ \Delta\, y_{t-p+1} \\ \alpha' y_{t-1}\end{array}\right]\right)\otimes Z_{t-1}\right],\end{aligned}$$

i.e., the orthogonality between all the elements in v_t and all the elements in Z_{t-1}. The test for common cycles is an over-identifying restriction test – the J test proposed in Hansen (1982) – which has an asymptotic χ^2 distribution with degrees of freedom equal to the number of over-identifying restrictions. The over-identifying restrictions test checks whether the errors of the system are orthogonal to all the instruments in Z_{t-1}.

4 A Forecast-Combination Approach for Metal Prices

Here, we discuss the techniques used for optimal forecasting of metal-commodity prices. An in-depth theoretical discussion of these issues can be found in Bates and Granger (1969), Palm and Zellner (1992), Stock and Watson (2006), Timmermann (2006), and more recently in Issler and Lima (2009). The latter is our preferred approach, partly reproduced here with improved notation for completeness. It is appropriate for forecasting a weakly stationary and ergodic univariate process $\{y_t\}$ using a large number of forecasts that will be combined to yield an optimal forecast in the mean-squared error (MSE) sense. These forecasts are the result of several econometric models that need to be estimated prior to forecasting. We label forecasts of y_t, computed using conditioning sets lagged h periods, by $f_{i,t}^h$, $i = 1, 2, \ldots, N$.

Therefore, $f_{i,t}^h$ are h-step-ahead forecasts and N is the number of models estimated to forecast $f_{i,t}^h$.

Issler and Lima (2009) consider 3 consecutive distinct time sub-periods. The first sub-period E is labeled the "estimation sample", where models are usually fitted to forecast y_t subsequently. The number of observations in it is $E = T_1 = \kappa_1 \cdot T$, comprising $(t = 1, 2, \ldots, T_1)$. The sub-period R (for regression) is labeled the post-model-estimation or "training sample", where realizations of y_t are usually confronted with forecasts produced in the estimation sample, and weights and bias-correction terms are estimated. It has $R = T_2 - T_1 = \kappa_2 \cdot T$ observations in it, comprising $(t = T_1 + 1, \ldots, T_2)$. The final sub-period is P (for prediction), where genuine out-of-sample forecast is entertained. It has $P = T - T_2 = \kappa_3 \cdot T$ observations in it, comprising $(t = T_2 + 1, \ldots, T)$.

Forecasts $f_{i,t}^h$'s are approximations to the optimal forecast $(\mathbb{E}_{t-h}(y_t))$ as follows:

$$f_{i,t}^h = \mathbb{E}_{t-h}(y_t) + k_i^h + \varepsilon_{i,t}^h, \tag{11}$$

where k_i^h is the individual model time-invariant bias for h-step-ahead prediction and $\varepsilon_{i,t}^h$ is the individual model error term in approximating $\mathbb{E}_{t-h}(y_t)$, where $\mathbb{E}(\varepsilon_{i,t}^h) = 0$ for all i, t, and h. Here, the optimal forecast is a *common feature* of all individual forecasts and k_i^h and $\varepsilon_{i,t}^h$ arise because of forecast misspecification.

We can always decompose the series y_t into $\mathbb{E}_{t-h}(y_t)$ and an unforecastable component ζ_t^h, such that $\mathbb{E}_{t-h}(\zeta_t^h) = 0$ in:

$$y_t = \mathbb{E}_{t-h}(y_t) + \zeta_t^h. \tag{12}$$

Combining (11) and (12) yields the well known two-way decomposition, or error-component decomposition, of the forecast error $f_{i,t}^h - y_t$:

$$\begin{aligned} f_{i,t}^h &= y_t + \mu_{i,t}^h, \qquad i = 1, 2, ..., N, \text{ and } T > T_1, \\ \mu_{i,t}^h &= k_i^h + \eta_t^h + \varepsilon_{i,t}^h, \text{ where } -\zeta_t^h = \eta_t^h \end{aligned} \tag{13}$$

From the perspective of combining forecasts, the components k_i^h, $\varepsilon_{i,t}^h$ and η_t^h play very different roles. If we regard the problem of forecast combination as one aimed at diversifying risk, i.e., a finance approach, then, on the one hand, the risk associated with $\varepsilon_{i,t}^h$ can be diversified, while that associated with η_t^h cannot. On the other hand, in principle, diversifying the risk associated with k_i^h can only be achieved if a bias-correction term is introduced in the forecast combination, which reinforces its usefulness.

Issler and Lima propose the following non-parameteric consistent estimates for the components k_i^h, B^h, η_t^h, and $\varepsilon_{i,t}^h$: $\widehat{k}_i^h = \frac{1}{R}\sum_{t=T_1+2}^{T_2} f_{i,t}^h - \frac{1}{R}\sum_{t=T_1+2}^{T_2} y_t$, $\widehat{B^h} = \frac{1}{N}\sum_{i=1}^{N} \widehat{k}_i^h$, $\widehat{\eta^h}_t = \frac{1}{N}\sum_{i=1}^{N} f_{i,t}^h - \widehat{B^h} - y_t$, $\widehat{\varepsilon}_{i,t}^h = f_{i,t}^h - y_t - \widehat{k}_i^h - \widehat{\eta^h}_t$. They show that, under a set of conditions, the *feasible* bias-corrected average forecast (BCAF) $\frac{1}{N}\sum_{i=1}^{N} f_{i,t}^h - \widehat{B^h}$ obeys:

$$\underset{(T,N\to\infty)_{seq}}{\text{plim}} \left(\frac{1}{N}\sum_{i=1}^{N} f_{i,t}^h - \widehat{B^h} \right) = y_t + \eta_t^h = \mathbb{E}_{t-h}(y_t),$$

where $\text{plim}_{(T,N\to\infty)_{seq}}$ is the probability limit using the sequential asymptotic framework of Phillips and Moon (1999). Thus, the feasible BCAF is an optimal forecasting device.

They also show that there is an infinite number of optimal forecast combinations using deterministic weights $\{\omega_i\}_{i=1}^{N}$,such that $\omega_i \neq 0$, $\omega_i = O(N^{-1})$ uniformly, with $\sum_{i=1}^{N} \omega_i = 1$ and $\lim_{N\to\infty} \sum_{i=1}^{N} \omega_i = 1$. This allows the discussion of the well-known *forecast combination puzzle*: if we consider a fixed number of forecasts ($N < \infty$), combining them using equal weights ($1/N$) fare better than using "optimal weights" constructed to outperform all other forecast combination in the mean-squared error (MSE) sense. Optimal population weights, constructed from the variance-covariance structure of models with stationary data, are optimal. Thus, the forecast-combination puzzle must be a consequence of the lack of consistency in estimating them, and can arise when N, the number of models being combined, is high relative to the number of observations used in estimating them by OLS – R.

Finally, there is one interesting case in which we can dispense with estimation in combining forecasts: when the mean bias is zero, i.e., $B^h = 0$, there is no need to estimate B^h and the BCAF is simply equal to $\frac{1}{N}\sum_{i=1}^{N} f_{i,t}^h$, the sample average of all forecasts. This is the ultimate level of parsimony. To test the null that $B^h = 0$, Issler and Lima developed a robust t-ratio test that takes into account the cross-sectional dependence in k_i^h.

4.1 Forecast Combination for Nested Models

The potential problem of nested models is that the innovations from nested models can exhibit high cross-sectional dependence, preventing a weak law-of-large numbers (WLLN) to hold. Issler and Lima (2009) introduce nested models by considering a continuous set of models splitting the total number of models N into M classes (or blocks), each of them containing m nested models, so that $N = mM$. In the index of forecasts, $i = 1, \ldots, N$, we group nested models contiguously. Hence, models within each class (block) are nested but models across classes (blocks) are non-nested.

The number of classes and the number of models within each class to be functions of N, respectively as follows: $M = N^{1-d}$ and $m = N^d$, where $0 \leq d \leq 1$. Notice that this setup considers all the relevant cases: (i) $d = 0$ corresponds to the case in which all models are non-nested; $d = 1$ corresponds to the case in which all models are nested and; (iii) the intermediate case $0 < d < 1$ gives rise to N^{1-d} blocks of nested models, all with size N^d.

Regarding the interaction across blocks of nested models, it is natural to impose that the correlation structure of innovations across classes is such that it does not prevent a weak law-of-large numbers (WLLN) to hold, although we expect it not to hold within every block of nested models. Keeping some nested models poses no problem, since the mixture of all models will still deliver the optimal forecast. From a practical point of view, the choice of $0 \leq d < 1$ seems to be superior and is sufficient to guarantee optimality of forecasts combinations as before.

5 Empirical analysis

5.1 Data and Empirical Implementation

We employed data of different frequencies and different sources building a very comprehensive dataset of metal prices and potential co-variates that can be used either in building economics models and/or for forecasting. We have a library of data on three different frequencies: monthly, quarterly and annual.

On a monthly basis, the metal-price data consists of commodity prices for a variety of metals (or derived products) – Aluminium, Copper, Lead, Nickel, Tin and Zinc – extracted from the International Financial Statistics (IFS) of the IMF. Metal-price data at this frequency is available from 1957:1 to 2012:3. Nominal price data were deflated using the consumer price index (CPI) for the U.S., which was extracted from the FRED database of the St. Louis FED. We have also a measure of industrial production in a monthly basis, constructed by J.P. Morgan, which is used in building quasi-structural models. It includes Chinese and Indian industrial production. Because of that, this series is available only from 1992:01 through 2012:09. In forecasting, we used co-variates (predictors) which are potentially correlated to the prices of these metals or derived products. These are mostly composed by financial indices downloaded from the library kept by Welch and Goyal (2008) and by Hong and Yogo (2012), available from 1965 to 2008. This list includes: global, U.S., and Chinese industrial production, the primary metals coincident and leading indices, provided by the United States Geological Service (USGS), and a few financial-sector co-variates, such as: VIX – a volatility index, the U.S. real effective exchange rate, returns and excess returns on U.S. government bonds at different maturities, and the return on the S&P500 index.

On a quarterly basis, we collect price data for the same metals (or derived products) listed for monthly frequency. They are available from 1957:1 through 2012:1, and were extracted from IFS database. Nominal price data were deflated using the CPI for the U.S. We also employed the (relatively large) list of co-variates used in Welch and Goyal (2008) and in Hong and Yogo (2012).

On an annual basis, metal-price data were provided by the United States Geological Service (USGS), from 1900 through 2010. Annual prices were deflated by the U.S. CPI – a series put together by the St. Louis and Minneapolis Federal Reserve Economic Database. Actual annual CPI data covers the period 1913-2010, whereas the period 1900-13 uses FED estimates. We also employed the list of annual covariates used in Welch and Goyal (2008) and in Hong and Yogo (2012) and a list of financial indices and real economic variables, such as Angus Maddison's historical GDP, and Shiller's U.S. per capita real consumption.

Our analysis of common-cyclical features will focus on the GMM tests proposed in Section 3, which is an appropriate testing strategy under unknown heterogeneity and dependence of the moment restrictions in question. Cointegration analysis investigates the existence of long-run relationships among economic data. As is well known, this requires the use of long-span data. Higher frequency is not a substitute for the it. Thus, we put much more emphasis on cointegration tests using annual data, given it has the longest span – 110 years of data. Obviously, we still test for cointegration at other frequencies, but we do not emphasize the results so much. In addition to that, because monthly and quarterly data are only available since 1957 (55 years of data), cointegration results using annual, quarterly and monthly data may not necessarily match. This may be simply a consequence of the fact that the samples used in these analyses are different.

Regarding the forecasting exercise, the focus will be on monthly and annual frequencies alone – the former being appropriate to short-term forecasts and the latter to long-term forecasts.

5.2 Bivariate Analysis: Cointegration and Common Cycles for Metal Prices

5.2.1 Monthly Frequency

Data for (log) prices of metals (or derived products) – Aluminium, Copper, Lead, Nickel, Tin and Zinc – are available from 1957:01 through 2012:03, whereas data for (log) Global Industrial Production (seasonally adjusted) is available from 1992:01

through 2012:09. Data for U.S. industrial production is available from 1919:1 onwards. All these series show signs of containing a unit root, which is confirmed for all of them using Phillips and Perron (1989) test[6].

First, we analyze the pairwise behavior of metal commodity prices alone, asking whether they share common trends and/or common cycles. Results are presented in Table 1. Regarding cointegration, we find overwhelming evidence of common trends among pairwise prices (10 out of 15). Conditional on this evidence, we tested for common cycles using the GMM approach (robust to heteroskedasticity and serial correlation of unknown form), finding no signs of pairwise common cycles for the growth rates of metal-commodity prices – the only exception being the pair aluminum and lead, albeit the evidence is faint.

Next, we investigate whether prices for metal commodities cointegrate and/or share common cycles with global industrial production. The analysis is pairwise, one commodity price at a time. Results are presented in Table 2 (panel A). Regarding cointegration, with the exception of aluminium, we find no evidence of a long-run relationship between metal prices and global industrial production for the last 20 years. On the other hand, results for common cycles are very different. Using the GMM approach, at 5% significance, we found evidence of strong-form *common-cyclical features* between industrial production and the following metals: copper, nickel, tin, and zinc. In addition to that, we also found evidence of weak-form *common-cyclical features* between industrial production and aluminium.

To motivate the findings of common cycles between global industrial production and the real price of copper, nickel, tin, and zinc, we detail here the results for copper, a metal for which its price is known to be associated with economic activity – the conventional wisdom of financial and business-cycle analysts for a long time[7].

[6] A slight caveat involves (log) aluminium prices, which rejects the null of a unit root at 5% significance when a constant is included, but rejects when a constant and trend are included. It also rejects Kwiatkowski et al. (1992) stationarity test. Thus, we chose to model it as a $I(1)$ process.

[7] There is a large group of webpages advertising the relationship between the Dow Jones Index and copper prices (e.g., http://www.marketoracle.co.uk/Article27240.html) or between copper prices and business cycles (e.g., http://www.marketwatch.com/story/is-dr-copper-about-to-make-a-house-call-2013-01-07).

In Figure 1 below, we plot the growth rates of copper prices – labelled $\Delta \ln \left(P_t^{Co}\right)$ – and the growth rates of global industrial production – labelled $\Delta \ln \left(IP_t^G\right)$, both standardized (zero mean, unit variance).

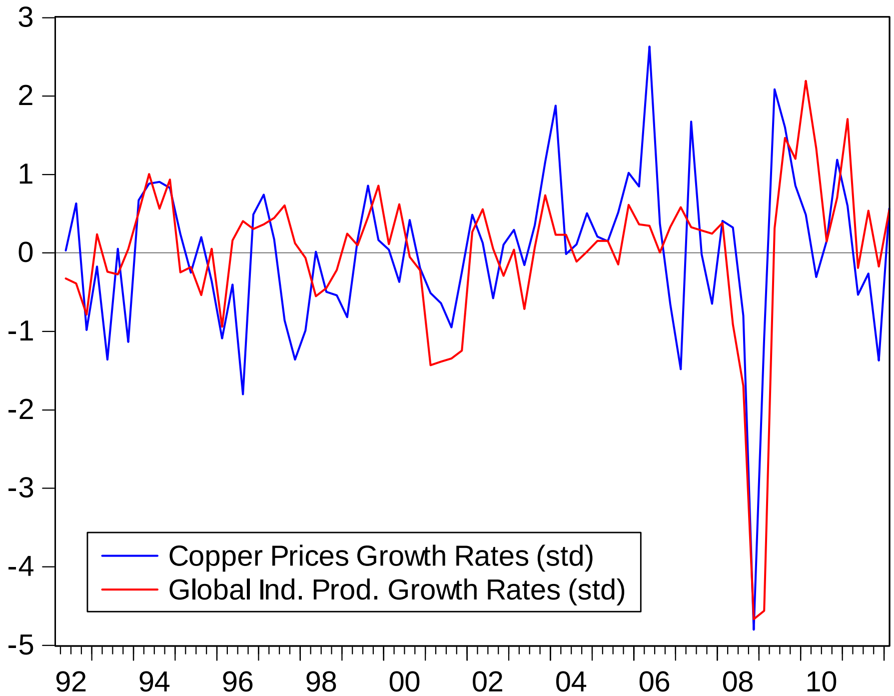

Figure 1: Standardized Growth Rates of Global Industrial Production and Copper Prices

Notice that both $\Delta \ln \left(P_t^{Co}\right)$ and $\Delta \ln \left(IP_t^G\right)$ show signs of serial correlation, as is apparent from Figure 1 above. However, using the Ljung-Box test at 10% significance, our empirical results in Table 2 found that the following linear combination is white noise (unpredictable):

$$\Delta \ln \left(P_t^{Co}\right) - \underset{(1.50)}{7.523} \times \Delta \ln \left(IP_t^G\right) + \underset{(0.006)}{0.015}, \tag{14}$$

with robust standard errors in parenthesis. This shows that $\Delta \ln \left(P_t^{Co}\right)$ and $\Delta \ln \left(IP_t^G\right)$ are synchronized and that $\ln \left(P_t^{Co}\right)$ and $\ln \left(IP_t^G\right)$ share a common cycle. As we argued in Section 2, this is consistent with a standard theory of derived demand for copper in producing industrial output for the global economy when the supply of copper is held fixed in the short run.

Next, still in Table 2 panel A, we investigate whether or not the growth rates of metal-commodity prices are synchronized with U.S. industrial production. Since we want to compare results with the previous tests using global industrial production, we employ the same sample period in the analysis (1992:1-2012:3), noting that U.S. industrial production is available from 1919:1 onwards. For the U.S., using the J-statistic, we find clear evidence of synchronized growth rates for industrial production and aluminium, tin, and zinc. Regarding lead and copper, the evidence is not so clear. In any case, overall, the only statistically significant combinations of growth rates are the ones involving aluminium and tin, respectively. Thus, the evidence regarding the U.S. economy is weaker than that of the global economy. If we apply the same tests using the whole sample – 1957:1-2012:3 – we find no evidence of common cycles at all; see panel B in Table 2.

It is interesting to contrast the evidence of common cycles between metal-commodity prices and global industrial production with that between the former and U.S. industrial production. As is well known, there is a recent migration of industrial activity from developed countries to emerging economies, especially China and India. Global industrial production is highly influenced by the industrial production of these two countries, which may explain why it is synchronized with metal-commodity prices. On the other hand, developed countries such as the U.S., have witnessed a continued decline of their industrial sector, which may explain why we did not find strong evidence of synchronicity of U.S. industrial production and metal-commodity prices.

5.2.2 Quarterly Frequency

On a quarterly frequency, data for (log) prices of metals (or derived products) – Aluminium, Copper, Lead, Nickel, Tin and Zinc – are available from 1957:01 through

2012:01. Table 3 presents results of pairwise cointegration between metal prices, reporting overwhelming evidence of cointegration between prices of different metal commodities. This is consistent with our previous finding at the monthly frequency, although our quarterly results are even stronger – 14 out of 15 pairs versus 10 out of 15.

Regarding pairwise common cycles among commodity prices, we found limited evidence that they share common-cyclical features. Using the GMM approach described in Section 3, for all possible 15 pairwise cases, we found strong-form SCCF for 6 of them – aluminium-copper, aluminium-lead, aluminium-nickel, copper-nickel, lead-zinc, and nickel-tin. So, the growth rate of prices of aluminium and nickel are synchronized with those of other metal commodities. Similar results are also obtained for weak-form SCCF.

Next, using the sample 1992:1 through 2012:1, we investigate the existence of pairwise common trends and common cycles between metal-commodity prices and global and U.S. industrial production, respectively. Results are given in Table 4. First, in panel A, we find no evidence of cointegration between metal prices and global and U.S. industrial production. Second, regarding global industrial production, we find strong evidence of common cycles for aluminium, copper, and tin. For zinc, there is a common cycle at 5% significance, but not at 10%. Third, regarding U.S. industrial production, we find strong evidence of common cycles for aluminium only. For copper, tin, and zinc, there is a common cycle at 5% significance, but not at 10%. Thus, we conclude that the evidence of common cycles is stronger regarding global industrial production. This result is consistent with our findings for the monthly frequency.

Finally, we investigate the existence of pairwise common trends and common cycles between metal-commodity prices and U.S. industrial production, using the complete sample from 1957:1 through 2012:1. Results are given in Table 4, panel B. In strong-form SCCF tests, we find hard evidence of common cycles for aluminium and nickel, and some evidence for lead and copper.

5.2.3 Annual Frequency

On an annual basis, metal-price data were provided by the United States Geological Service (USGS) from 1900 through 2010. Prices were deflated by the U.S. CPI. Table 5 presents results of pairwise cointegration between metal prices. We found overwhelming evidence of cointegration between prices of different metal commodities. Given the longer span of this annual database vis-a-vis the monthly and quarterly databases – more than twice as long – cointegrating evidence here should receive more weight vis-a-vis previous evidence. From all possible 15 cases, we found cointegration among 10 pairs of metal-commodity prices.

One interesting issue is the long-run behavior of real metal-commodity prices: while three of them displayed an obvious increase in prices (copper, nickel, and zinc) in 110 years – more than a twofold increase from 1900 to 2010 – the other three displayed an obvious decrease over time (aluminium, lead, and tin) of about 70%-90%. We conjecture here that the industrial processes of the beginning of the 20th Century used aluminium, lead, and tin in larger quantities than what was used by the end of the 20th Century. An inverse pattern being observed for copper, nickel, and zinc.

As we stressed above, cointegration analysis requires the use of long-span data. Higher frequency is not a substitute for span. Thus, our preferred results are the ones obtained in cointegration tests when using annual data, given it has the longest span – 110 years of data – since monthly and quarterly data are only available since 1957 (55 years of data).

Conditional on cointegration results, we investigate next the existence of common cycles for annual price data in pairwise analyses. Table 5 presents overwhelming evidence of common cycles for commodity prices in testing. For all possible 15 pairwise cases, we found strong-form SCCF for 14 of them, the only exception being the pair tin-zinc. Similar results are also obtained for weak-form SCCF. One point to note is that annual data for metal-prices showed much more synchronization than did quarterly and monthly data. This may be a sign that some high-frequency fluctuations that are not synchronized tend to disappear with time aggregation.

Another plausible explanation for synchronization is the fact that we are using a longer sample period for the annual analysis.

Finally, we investigated whether there are cointegration and common cycles for metal prices and (U.S. or global) industrial production. Unfortunately the instantaneous growth rates of U.S. and global industrial production showed no signs of possessing a *serial-correlation feature*, which is a necessary condition to test for common cycles. Thus, we refrained here to go any further on that regard[8].

5.3 Multivariate Analysis: Cointegration and Common Cycles for Metal Prices

We condition on previous evidence of bivariate cointegration and common cycles among metal prices and among metal prices and industrial production to build multivariate models for (log) metal prices (or derived products) – Aluminium, Copper, Lead, Nickel, Tin and Zinc – and industrial production – either for the U.S. economy or at a global level. We expect these multivariate models to display common cycles, so we construct two different sets of vector autoregressive (VAR) models to serve as reduced forms: one for metal prices and global industrial production and one for metal prices and U.S. industrial production – both with seven variables – later investigating if they possess common trends using Johansen's (1991) test and common-cyclical-feature restrictions using the GMM approach of Section 3.

We focus on monthly data, since data at the highest frequency represent best the short-run analysis which are the object of theoretical modelling of Section 2 and the empirical evidence above. We were careful in selecting the lag order of the VAR to avoid having "dynamically incomplete" models; see Vahid and Issler (2002) and Athanasopoulos et al. (2011)[9]. For monthly data and global industrial production,

[8] This has some implications for the implementation of the optimal forecast methods discussed above, the main one being that we would not be able to build restricted VECMs (with common-cycle restrictions) to be later used in forecast combinations. Notice that, for the monthly frequency, we can do just that, linking the *understanding* part of this paper with the *forecast* part.

[9] Indeed, these papers document that using standard information criteria underestimates lag order in small samples for data where common-cyclical features are present.

we selected a VAR with two lags in levels. There is no evidence of cointegration when all metal prices and global industrial production are jointly modelled. We selected five lags when U.S. industrial production was used instead, finding again no cointegration for the system.

Next, we present GMM tests for common-cyclical-feature restrictions in the two systems described above. Results are presented in Table 6. We conclude for the existence of six cofeature vectors in both cases. Thus, all metal prices share a common cycle with industrial production (U.S. or global), given the form of the contemporaneous relationships in (6). This is consistent with our previous bivariate results, although a bit stronger, since, for the former, it was not unanimous.

To get an idea of the parsimony entailed by imposing common-cyclical-feature restrictions in a multivarite setting, note that the unrestricted VAR in differences with one lag, such as (6), for six metal prices and global industrial production, has a total of 56 parameters. However, the same system where common-cyclical-feature restrictions are imposed (equation (9)) – with the existence of 6 cofeature vectors – has only 20 parameters. Testing whether it is valid to impose those restrictions leads to a p-value of 0.7376, which validates the restricted model at usual significance levels[10].

To see this explicitly, denote by Δy_t a vector stacking respectively the instantaneous growth rates of aluminium, copper, lead, nickel, tin zinc, and global industrial production, as shown in the right-hand-side of equation (15) below. The estimated

[10]The parsimony of the four-lag is even more striking, owing to the higher number of lags.

quasi-structural model took the form[11]:

$$\underset{7\times 1}{\begin{pmatrix} \Delta\ln\left(P_t^{AL}\right) \\ \Delta\ln\left(P_t^{Co}\right) \\ \Delta\ln\left(P_t^{Pl}\right) \\ \Delta\ln\left(P_t^{Ni}\right) \\ \Delta\ln\left(P_t^{TN}\right) \\ \Delta\ln\left(P_t^{Zn}\right) \\ \Delta\ln\left(IP_t^{G}\right) \end{pmatrix}} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & \underset{(1.13)}{-7.023} \\ 0 & 1 & 0 & 0 & 0 & 0 & \underset{(1.49)}{-2.13} \\ 0 & 0 & 1 & 0 & 0 & 0 & \underset{(1.50)}{-6.87} \\ 0 & 0 & 0 & 1 & 0 & 0 & \underset{(1.31)}{-5.45} \\ 0 & 0 & 0 & 0 & 1 & 0 & \underset{(2.06)}{-9.60} \\ 0 & 0 & 0 & 0 & 0 & 1 & \underset{(1.80)}{-8.13} \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}^{-1} \times \tag{15}$$

$$\begin{bmatrix} \multicolumn{7}{c}{\underset{6\times 7}{\mathbf{0}}} \\ \underset{(0.00)}{0.02} & \underset{(0.003)}{0.002} & \underset{(0.002)}{-0.005} & \underset{(0.002)}{0.004} & \underset{(0.004)}{-0.008} & \underset{(0.003)}{0.008} & \underset{(0.04)}{0.14} \end{bmatrix} \Delta y_{t-1},$$

As theory and experience has taught us, the restricted VECM forecasts much better than their unrestricted counterparts. In our previous experience, to give some idea of how much better the restricted VECM forecasts, consider the following: Issler and Vahid (2001) find a 25% reduction for the determinant of the mean-squared forecast error matrix – $|MSPE|$ – for U.S. macroeconomic aggregates, Vahid and Issler (2002) find a reduction of 20% for $|MSPE|$ when predicting U.S. coincident series using the same statistic, and Athanasopoulos et al. (2011) find a reduction of 47% for $|MSPE|$ when predicting different measures of Brazilian Inflation.

Finally, in constructing the models that will be used in the forecast-combination analysis, we employed some reduced-rank models where common-cycle restrictions are imposed. Thus, our combined forecasts will benefit from what we have learned in the empirical analysis regarding the synchronicity of cycles in metal-commodity prices and between the former and different measures of industrial production.

[11] Disregarding constant terms.

5.4 Forecasting Metal Prices using Forecast Combinations

We now implement the forecast theory discussed in Section 3 above, where forecast accuracy is measured by the root of the mean-squared forecast error. Metal-price data used here is the same one used in the cointegration and common-cycle analyses of the previous section, although we only focused on results for monthly and annual data alone. The former is appropriate to examine short-term forecast accuracy, whereas the latter is appropriate to measure long-term accuracy.

Our target variables in forecasting are commodity prices for Aluminium, Copper, Lead, Nickel, Tin and Zinc – made available from the London Mercantile Exchange (extracted from the IFS) for monthly frequency and from the USGS at annual frequency. For some of the estimated models, we used co-variates (predictors) which are highly correlated to metal prices. Some are related to economic activity, such as: the global industrial production, the U.S. industrial production, the Chinese industrial production, the primary metals coincident index (USGS), a leading index of metals price (USGS), and some other financial-sector co-variates, such as: VIX – a volatility index, the U.S. real effective exchange rate and the S&P500 index.

Our monthly data set covers the period from January 1965 through December 2008, comprising 528 observations ($T = 528$). Our annual data set covers data from 1900 to 2010, comprising 111 observations ($T = 111$). Table 7 presents the correlations between the predictors are metal price data. Since there is evidence of a unit root for the metal prices and the co-variates used here, some series were transformed to instantaneous growth rates prior to computing correlations.

In order to fit well the cross-sectional asymptotic requirement (large N) regarding the WLLN, we need to have a large set of diversified forecasts to eliminate the combination of idiosyncratic errors. For this reason, we chose a few classes of different econometric models: AR, VAR, VECM, restricted VECM (common-cycle restrictions), all using distinct co-variates (predictors), and distinct functional forms (levels, logs), and stationarity assumptions (stationarity vs. difference-stationarity) for the target variable and predictors. Considering that some of these models were fairly similar, we discarded a few of those, ending up with N between 115 and 125, i.e., between 115 and 125 distinct models and distinct forecasts for each time horizon (h). Obviously, some of them are nested within each other, and we also have classes of nested models as well. As we argued before, this will not pose as long as we have a large enough number of diverse classes.

For implementing the BCAF and other combining techniques discussed above, we split the sample in three distinct parts, each with a specific purpose: the first one, from 1 to T_1, to estimate the coefficients of each model; the second from T_1+1 to T_2, to compute the bias; and the third from T_2+1 to T, to implement truly out-of-sample forecasting, and to assess the forecast accuracy of different forecast strategies and of individual models using the root mean-squared error (RMSE) of forecasts.

To asses forecast accuracy, we constructed an algorithm which is appropriate for the bias-corrected average forecast (BCAF). For alternative forecast combinations or forecasting schemes, slight modifications are required. The algorithm runs as follows:

1. For each model (AR, VAR, VECM, restricted VECM with common-cycle restrictions, and a specific set of predictors), we estimate the coefficients of the regressors using the sub-sample from 1 to T_1.

2. Forecast h-steps ahead the models estimated in step 1 (f_{it}^h) from T_1 to T_2. Each model should be forecasted h-steps ahead $T_2 - T_1 - h + 1$ times.

3. Calculate the bias associated with each h-step ahead forecasts and each model; the bias is the average error between the h-steps ahead forecast and the observed value of the target series (from T_1 to T_2).

4. Forecast h-steps ahead the same models estimated in step 1 for only $T_2 + h$, using the same coefficients estimated in step 1.

5. Store the bias from step 3 and the forecast made in step 4, f_{i,T_2+h}^h.

6. Update $T_1 = T_1 + 1$, $T_2 = T_2 + 1$.

7. Go to step 1 until $T_2 = T$.

8. Adjust the forecasts of each model (made from T_2+1 to T) by their respective bias.

9. Combine all these adjusted forecasts using equal weights.

10. Compute the RMSE of the BCAF, considering the series of metals price index as the target series.

For the monthly dataset, we took $T_1 = 200$ and $T_2 = 378$. Since $T = 528$, this leaves 150 observations to evaluate out-of-sample performance of different models. For the annual data set, we took $T_1 = 35$ and $T_2 = 70$. Since $T = 111$, this leaves 41 observations for out-of-sample evaluation. In both cases, we kept enough data to estimate the models and two similar-size sub-samples to estimate their biases and to perform out-of-sample forecasts.

The maximum horizon was set to 6 months for monthly data and to 5 years with annual data. After computing the average bias for each forecast horizon ($\widehat{B^h}$), we tested the null $H_0 : B^h = 0$, using Issler and Lima's t-ratio test. Tables 8 and 9 present the results, respectively with monthly and annual data.

From the results in Tables 8 and 9, we conclude that, for monthly data, none of the mean biases were significant using the t-ratio test in Issler and Lima. For annual data, zinc is the only metal for which the mean bias is clearly non-zero. There is also scattered evidence of non-zero mean bias for copper and nickel at higher horizons.

Recall from Section 4 that, when the mean bias is zero, the optimal forecast collapses to the simple average across models $\frac{1}{N}\sum_{i=1}^{N} f_{i,t}^h$. In this case, it might not be worth discarding a part of the sample to compute the bias. Then, the best strategy is simply to merge the samples E and R into one (sample from $t = 1$ through $t = T_2$), where models are estimated.

5.4.1 Comparing forecast accuracy of different models

Given the theoretical results in Section 4 for optimality of combined forecasts, we consider here several forecast-combination strategies: (i) the bias-corrected average forecast (BCAF); (ii) the average forecast (AF); (iii) the weighted average forecast

(WAF), where weights are based on the inverse of the mean-squared error for each model, normalized to add up to unity; (iv) the simple average of the 5 best fitting models (by Bayesian Information Criteria – BIC, using a previous sample period); (v) the simple average of the 10 best fitting models (BIC); (vi) the median forecast[12]. We computed the RMSE for these forecast strategies. Results are presented in Tables 10 and 11, respectively for monthly and annual data. We could not make most forecast models for zinc to converge at the monthly frequency, and thus refrain from presenting its forecast-evaluation results.

For monthly data, for 1- through 6-steps ahead and across all metal prices, the best performance (by far) in terms of RMSE was achieved by the Average Forecast (AF), followed by the *best model* and then followed closely by the weighted average forecast (WAF). For 3 out of 5 metal prices, the Average Forecast (AF) was the best forecast strategy out of sample. This is exactly what one should expect from econometric theory (Section 4), given our evidence above that the average bias was statistically zero, i.e., that we could not reject the null $H_0 : B^h = 0$, using Issler and Lima's t-ratio test in Table 8. Moreover, for 4 out of 5 metal prices, forecast combinations were superior to choosing *"a best model."*

In Tables 10 and 11, we also present out-of-sample R^2 statistics (percentage) comparing forecast strategies with the random-walk model with and without drift, which are important benchmarks to be beaten in the finance literature. R^2-statistics were computed for one-step ahead forecasts only to save space. For metal price z_t, we have:

$$R^2 = 100 \times \left[1 - \frac{\sum_{t=T_2+1}^{T} \left(z_t - \widehat{z}_{t|t-1} \right)^2}{\sum_{t=T_2+1}^{T} \left(z_t - \widehat{z}_{t|t-1}^{BMK} \right)^2} \right],$$

where $\widehat{z}_{t|t-1}$ is the one-step-ahead forecast of any given strategy and $\widehat{z}_{t|t-1}^{BMK}$ is the one-step-ahead forecast of the benchmark – random-walk with and without drift.

[12]For the median forecast, we have no theoretical optimality result. For the 5-best and 10-best models, optimality can be justified as the number of best models increases with N, e.g., as a fixed quantile of models.

For monthly data – Table 10 – aluminium, copper and tin, our best strategy overall was the average forecast. It has $R^2 = 2.75\%$ for aluminium, $R^2 = 14.30\%$ for copper, and $R^2 = 14.98\%$ for tin, vis-a-vis the random walk with drift[13]. Table 10 also includes Clark and West's (2007) tests for equal predictive accuracy applicable to nested models[14]. With the exception of nickel and tin, for all other models we cannot reject the null of equal forecast accuracy between any forecast strategy and the random-walk models.

For annual data – Table 11 – the best forecast strategy in terms of the RMSE is to employ the bias-corrected average forecast (BCAF) – for aluminium, lead and zinc, followed by average forecast (AF) – best for tin. Thus, for 4 out of 6 metal prices, forecast combinations were the best forecast strategy out of sample; see Table 11. In terms of out-of-sample R^2 statistics, these best models outperformed the random walk with and without drift: $R^2 = 8.59\%$ for aluminium, $R^2 = 15.90\%$ for lead, $R^2 = 8.98\%$ for tin, and $R^2 = 23.8\%$ for zinc. Here, contrary to the evidence for monthly frequency, the best forecasts strategies are significantly different (better) than the random walk using Clark and West's test.

5.4.2 Best Predictors and Models used in Forecast Combinations

Tables 12 and 13 report, respectively, the best models and predictors for monthly and annual forecasts, considering out-of-sample results across all horizons. For each out-of-sample observation, and each forecast horizon, we compared models using squared forecast errors. The model with the smallest squared error is considered the best. Tables 12 and 13 report the percentage which each model type is considered best across all horizons and out-of-sample observations. Regarding predictors, notice that the models being combined are all autoregressive, so the lags of metal prices are used as predictors in them. The category *no extra predictors* includes only the these lags. For some models, we use extra predictors as well, which are also reported

[13] Notice that, even when the average forecast was not the best strategy, it has beaten the random walk.

[14] Alquist, Kilian, and Vigfusson (2012) criticize Clark and West's test, arguing that it rejects the null of equal out of sample accuracy too often in favor of the nested model.

separately.

For monthly frequency, Table 12, the multivariate models – restricted vector error-correction models (VECM) described in Section 5.3 – with all metal prices and industrial production (U.S.), have the best forecasting performance overall. This connects "understanding" and "forecasting," both in the title of this paper. Moreover, it serves as empirical validation of the theoretical model discussed in Section 2 above, where demand for metals commodities are modelled as a *derived demand* in producing industrial output and the supply of metal commodities are supposed to be fixed in the short run, due to the long maturity of metal-commodity projects and to their high capital intensity. The unrestricted VECMs also perform well, but not nearly as well as their restricted counterparts. Restricted and unrestricted VECMs highlight the importance of investigating short- and long-run relationships as done above.

The best predictors in monthly frequency are the lagged dependent-variables. This is expected, since most time-series models used here predict the future using the past and present. What is really informative is to look which *extra predictors* did well out of sample. For monthly frequency, U.S. industrial production overwhelmingly outperforms all other predictors, with the S&P 500 and measures of realized volatility coming a far second for different metals.

For the annual frequency, we did not did not forecast using restricted *common-cycle* models (restricted VECMs), since U.S. industrial production has no serial correlation at that frequency. We were left with AR, VAR, and VECMs for forecasting metal prices. In this context, AR models were the best for aluminium and copper, VARs were best for lead and zinc, and VECMs were best for nickel and tin. For predictors, the best are lagged dependent-variables, as expected. In the class of *extra predictors*, U.S. industrial production and volatility measures performed well for different metals.

6 Conclusion and Further Research

The objective of this paper was to study (*understand* and *forecast*) spot metal price levels and changes at monthly, quarterly, and annual horizons. The data used consist of metal-commodity prices at a monthly and quarterly frequencies, from 1957 to 2012, extracted from the IFS, and USGS annual data from 1900 through 2010. We also employed the (relatively large) list of co-variates used in Welch and Goyal (2008) and in Hong and Yogo (2009) , which are available for download.

Regarding the *understanding* part of the paper, we were able to show theoretically that there must be a positive correlation between metal-price variation and industrial-production variation if metal supply is held fixed in the short run when its demand is optimally chosen taking into account optimal production for the industrial sector. This is simply a consequence of the derived-demand model for cost-minimizing firms, which is paramount in microeconomics (Section 2). Our empirical evidence (monthly and quarterly data) fully supports this theoretical result. Indeed, we have shown overwhelming evidence that cycles in metal prices are synchronized to those in industrial production. This evidence is stronger regarding the global economy but holds as well for the U.S. economy to a lesser degree. As far as we know, we were the first authors to investigate and find common cycles in this way, accounting for theory and empirics, and not just describing a stylized fact. This is one of the main contributions of this paper.

The second objective of the paper was to *forecast* metal prices in short horizons (monthly data) and in long horizons (annual data). We propose a novel technique which views the optimal forecast in the MSE sense as a *common feature* (latent variable), which can be identified by using a cross-sectional average of a diverse group of forecasts, once we estimate a mean-bias term. There are several ways to combine forecasts optimally, but these combinations usually beat individual models in this context. This was indeed our finding, where combinations have beaten not only individual models most of the time but also the random walk model in several instances – which is usually not true for individual models. In predicting metal prices we connected the *forecasting* and *understanding* parts of this paper by using several forecast models that imposed *common-cycle* restrictions found in the *understanding* part[15]. We combined a variety of models (linear and non-linear, single equation and multivariate) and a variety of co-variates to forecast the returns and prices of six metal commodities. We found that the best performances in terms of RMSE were achieved by the average forecast (AF), the bias-corrected average forecast (BCAF), and the weighted average forecast (WAF). These are all forecast-combination schemes, which achieve optimality by eliminating individual-model forecast errors by the use of a weak law of large numbers. These empirical results are true for most metal prices, frequencies, and horizons, although some individual models performed well on occasion.

Finally, we were able to identify which models and predictors had the best forecast performance for different metal prices. For monthly frequency, the multivariate models – restricted vector error-correction models (VECM) – with all metal prices and industrial production (U.S.), have the best forecasting performance overall. Best predictors are the U.S. industrial production, followed by the S&P 500 and measures of realized volatility coming a far second. For the annual frequency, the AR model, the VAR and the VECM performed well for different metal prices. The best predictors were the U.S. industrial production and volatility measures.

References

Alquist, Ron, Lutz Kilian, and Robert J. Vigfusson (2012). Forecasting the Price of Oil. *Working Paper: University of Michigan*.

Athanasopoulos, G., O. T. de Carvalho Guillén, J. V. Issler, and F. Vahid (2011). Model selection, estimation and forecasting in var models with short-run and long-run restrictions. *Journal of Econometrics 164*(1), 116 – 129.

Batchelor, R. (2007). Bias in macroeconomic forecasts. *International Journal of Forecasting 23*(2), 189 – 203.

[15]This is true for short-run forecasts at the monthly frequency.

Bates, J. M. and C. W. J. Granger (1969). The combination of forecasts. *OR 20*(4), pp. 451–468.

Beveridge, S. and C. Nelson (1981). A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the 'business cycle'. *Journal of Monetary economics 7*(2), 151–174.

Cashin, P., H. Liang, and C. McDermott (2000). How persistent are shocks to world commodity prices? *Working Paper: International Monetary Fund*.

Cashin, P., C. McDermott, and A. Scott (1999). The myth of comoving commodity prices. *Working Paper International Monetary Fund 99*.

Cashin, P., C. McDermott, and A. Scott (2002). Booms and slumps in world commodity prices. *Journal of Development Economics 69*(1), 277–296.

Chambers, M. J. and R. E. Bailey (1996). A theory of commodity price fluctuations. *Journal of Political Economy 104*(5), pp. 924–957.

Clark, Todd E. and Kenneth D. West (2007). Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics 138*(5), pp. 291–311.

Clements, M. P. and D. F. Hendry (2006). Chapter 12 forecasting with breaks. Volume 1 of *Handbook of Economic Forecasting*, pp. 605 – 657. Elsevier.

Cuddington, J. (1992). Long-run trends in 26 primary commodity prices: A disaggregated look at the prebisch-singer hypothesis. *Journal of Development Economics 39*(2), 207–227.

Cuddington, J. and C. Urzúa (1989). Trends and cycles in the net barter terms of trade: a new approach. *The Economic Journal 99*(396), 426–442.

Davies, A. and K. Lahiri (1995). A new framework for analyzing survey forecasts using three-dimensional panel data. *Journal of Econometrics 68*(1), 205 – 227.

Deaton, A. (1999). Commodity prices and growth in africa. *The Journal of Economic Perspectives 13*(3), 23–40.

Deaton, A. and G. Laroque (1992). On the behaviour of commodity prices. *The Review of Economic Studies 59*(1), 1–23.

Deaton, A. and G. Laroque (1996). Competitive storage and commodity price dynamics. *Journal of Political Economy 104*(5), pp. 896–923.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica 50*(4), pp. 987–1007.

Engle, R. F. and C. W. J. Granger (1987). Co-integration and error correction: Representation, estimation, and testing. *Econometrica 55*(2), pp. 251–276.

Engle, R. F. and J. V. Issler (1995). Estimating common sectoral cycles. *Journal of Monetary Economics 35*(1), 83 – 113.

Engle, R. F. and S. Kozicki (1993). Testing for common features. *Journal of Business and Economic Statistics 11*(4), pp. 369–380.

Fuller, W. A. and G. E. Battese (1974). Estimation of linear models with crossed-error structure. *Journal of Econometrics 2*(1), 67 – 78.

Gargano, A. and A. Timmermann (2012). Predictive dynamics in commodity prices. *Working Paper*.

Grilli, E. and M. Yang (1988). Primary commodity prices, manufactured goods prices, and the terms of trade of developing countries: what the long run shows. *The World Bank Economic Review 2*(1), 1–47.

Hamilton, J. D. (2011). Nonlinearities and the macroeconomic effects of oil prices. *Macroeconomic Dynamics 15*, 364–378.

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica 50*(4), pp. 1029–1054.

Hecq, A., F. Palm, and J. Urbain (2006). Common cyclical features analysis in var models with cointegration. *Journal of Econometrics 132*(1), 117–141.

Hendry, D. F. and M. P. Clements (2004). Pooling of forecasts. *Econometrics Journal 7*(1), 1–31.

Hong, H. and M. Yogo (2009). Digging into commodities. *Unpublished Working Paper, Princeton University and Wharton of University of Pennsylvania*.

Hong, H. and M. Yogo (2012). What does futures market interest tell us about the macroeconomy and asset prices? *Journal of Financial Economics 105*(3), 473 – 490.

IMF (2012, April). *World Economic Outlook: Growth resuming, dangers remain*, Chapter 1, pp. 1–47. International Monetary Fund.

Issler, J. V. and L. R. Lima (2009). A panel data approach to economic forecasting: The bias-corrected average forecast. *Journal of Econometrics 152*(2), 153 – 164.

Issler, J. V. and F. Vahid (2001). Common cycles and the importance of transitory shocks to macroeconomic aggregates. *Journal of Monetary Economics 47*(3), 449 – 475.

Issler, J. V. and F. Vahid (2006). The missing link: using the nber recession indicator to construct coincident and leading indices of economic activity. *Journal of Econometrics 132*(1), 281 – 303.

Jerrett, D. and J. Cuddington (2008). Broadening the statistical search for metal price super cycles to steel and related metals. *Resources Policy 33*(4), 188–195.

Johansen, S. (1991). Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models. *Econometrica 59*(6), pp. 1551–1580.

Labys, W., A. Achouch, and M. Terraza (1999). Metal prices and the business cycle. *Resources Policy 25*(4), 229–238.

Laster, D., P. Bennett, and I. Geoum (1999). Rational bias in macroeconomic forecasts. *The Quarterly Journal of Economics 114*(1), 293–318.

Lettau, M. and S. Ludvigson (2001). Consumption, aggregate wealth, and expected stock returns. *The Journal of Finance 56*(3), 815–849.

Lettau, M. and S. C. Ludvigson (2004). Understanding trend and cycle in asset values: Reevaluating the wealth effect on consumption. *The American Economic Review 94*(1), pp. 276–299.

Newey, W. and K. West (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica: Journal of the Econometric Society*, 703–708.

Palm, F. C. and A. Zellner (1992). To combine or not to combine? issues of combining forecasts. *Journal of Forecasting 11*(8), 687–701.

Patton, A. and A. Timmermann (2007). Testing forecast optimality under unknown loss. *Journal of the American Statistical Association 102*(480), 1172–1184.

Phillips, P. C. B. and H. R. Moon (1999). Linear regression limit theory for nonstationary panel data. *Econometrica 67*(5), pp. 1057–1111.

Pindyck, R. S. and J. J. Rotemberg (1990). The excess co-movement of commodity prices. *The Economic Journal 100*(403), pp. 1173–1189.

Stock, J. and M. Watson (2006). Forecasting with many predictors. *Handbook of economic forecasting 1*, 515–554.

Timmermann, A. (2006). Forecast combinations. In C. G. G. Elliott and A. Timmermann (Eds.), *Handbook of Economic Forecasting*, Volume 1 of *Handbook of Economic Forecasting*, pp. 135 – 196. Elsevier.

Vahid, F. and R. F. Engle (1993). Common trends and common cycles. *Journal of Applied Econometrics 8*(4), pp. 341–360.

Vahid, F. and R. F. Engle (1997). Codependent cycles. *Journal of Econometrics 80*(2), 199 – 221.

Vahid, F. and J. V. Issler (2002). The importance of common cyclical features in var analysis: a monte-carlo study. *Journal of Econometrics 109*(2), 341 – 363.

Wallace, T. D. and A. Hussain (1969). The use of error components models in combining cross section with time series data. *Econometrica 37*(1), pp. 55–72.

Welch, I. and A. Goyal (2008). A comprehensive look at the empirical performance of equity premium prediction. *The Review of Financial Studies 21*(4), pp. 1455–1508.

West, K. D. (1996). Asymptotic inference about predictive ability. *Econometrica 64*(5), pp. 1067–1084.

Table 1: Common-Cycle Tests - Metal Prices (Monthly)

		Strong-form SCCF			Weak-form SCCF	
$\Delta y_{1,t}$	$\Delta y_{2,t}$	$\tilde{\alpha}^*$	J-statistic	Cointegration	$\tilde{\alpha}^*$	J-statistic
Aluminum	Lead	-0.193*	0.0118	-	-	-
		(0.106)	[0.05]			
Copper	Aluminum	-1.15***	0.0353	-	-	-
		(0.029)	[0.000]			
Aluminum	Tin	-0.168	0.0130	-	-	-
		(0.146)	[0.035]			
Nickel	Aluminum	-0.735***	0.1407	-	-	-
		(0.186)	[0.000]			
Zinc	Aluminum	-0.283	0.0344	(1,-0.463)	-0.295	0.0344
		(0.188)	[0.012]		(0.19)	[0.007]
Lead	Copper	-0.241**	0.0275	(1,-0.99)	-0.274***	0.0262
		(0.097)	[0.001]		(0.096)	[0.001]
Tin	Lead	-0.38***	0.0320	(1,-1.625)	-0.411***	0.0303
		(0.094)	[0.000]		(0.097)	[0.000]
Nickel	Lead	-0.23	0.0265	(1,-0.673)	-0.32	0.0257
		(0.189)	[0.000]		(0.194)	[0.000]
Zinc	Lead	-0.383***	0.0255	(1,-0.488)	-0.467***	0.0256
		(0.146)***	[0.000]		(0.156)	[0.000]
Copper	Tin	-0.841***	0.0403	-	-	-
		(0.238)	[0.000]			
Copper	Nickel	-0.319**	0.0358	(1,-1.67)	-0.365***	0.0355
		(0.128)	[0.000]		(0.133)	[0.000]
Zinc	Copper	-0.442***	0.0479	(1,-0.281)	-0.418***	0.0431
		(0.094)	[0.000]		(0.093)	[0.000]
Tin	Nickel	-0.122*	0.0361	(1,-5.198)	-0.165**	0.0379
		(0.07)	[0.001]		(0.077)	[0.000]
Tin	Zinc	-0.284***	0.0377	(1,-4.035)		0.0362
		(0.095)	[0.006]			[0.005]
Zinc	Nickel	-0.287***	0.0233	(1,-0.922)	-0.278***	0.0238
		(0.092)	[0.000]		(0.086)	[0.000]

Notes: GMM estimation using equation (9) for Strong-form SCCF and the analogue equation for Weak-form SCCF.

Robust Standard Errors (HAC) are in parentheses and p-values are in brackets.

Table 2: Common-Cycle Tests - Metal Prices and Industrial Production (Monthly)

		Strong-form SCCF			Weak-form SCCF	
$\Delta y_{1,t}$	$\Delta y_{2,t}$	$\tilde{\alpha}^*$	J-statistic	Cointegration	$\tilde{\alpha}^*$	J-statistic
Panel A - Sample: 1992M1-2012M3						
Aluminum	Global Industrial Production	-5.316***	0.0427	(1,-0.201)	-3.784***	0.0275
		(0.969)	[0.036]		(0.949)	[0.085]
Lead	Global Industrial Production	-4.052*	0.0331	-	-	-
		(2.101)	[0.047]			
Copper	Global Industrial Production	-7.523***	0.0310	-	-	-
		(1.504)	[0.189]			
Nickel	Global Industrial Production	-5.23***	0.0096	-	-	-
		(1.603)	[0.512]			
Tin	Global Industrial Production	-6.034***	0.0292	-	-	-
		(1.728)	[0.219]			
Zinc	Global Industrial Production	-5.827***	0.0337	-	-	-
		(1.601)	[0.329]			
Aluminum	US Industrial Production	-2.683***	0.0558	-	-	-
		(0.897)	[0.434]			
Lead	US Industrial Production	0.839	0.0577	-	-	-
		(1.799)	[0.056]			
Copper	US Industrial Production	-3.033	0.0513	-	-	-
		(2.018)	[0.094]			
Nickel	US Industrial Production	-2.622	0.0650	-	-	-
		(1.683)	[0.03]			
Tin	US Industrial Production	-2.524*	0.0429	-	-	-
		(1.301)	[0.176]			
Zinc	US Industrial Production	-1.923	0.0619	-	-	-
		(1.357)	[0.484]			
Panel B - Sample: 1957M1-2012M3						
Aluminum	US Industrial Production	0.833	0.0310	(1,0.537)	0.679	0.0270
		(0.603)	[0.000]		(0.658)	[0.000]
Lead	US Industrial Production	-0.993	0.0400	-	-	-
		(0.743)	[0.000]			
Copper	US Industrial Production	-1.782**	0.0390	-	-	-
		(0.878)	[0.000]			
Nickel	US Industrial Production	-0.257	0.0270	-	-	-
		(0.995)	[0.000]			
Tin	US Industrial Production	-1.277*	0.0320	-	-	-
		(0.658)	[0.000]			
Zinc	US Industrial Production	-2.301***	0.0300	(1,0.24)	-2.053**	0.0260
		(0.774)	[0.012]		(0.811)	[0.016]

Notes: GMM estimation using equation (9) for Strong-form SCCF and the analogue equation for Weak-form SCCF.

Robust Standard Errors (HAC) are in parentheses and p-values are in brackets.

Table 3: Common-Cycle Tests - Metal Prices (Quarterly)

		Strong-form SCCF			Weak-form SCCF	
$\Delta y_{1,t}$	$\Delta y_{2,t}$	$\tilde{\alpha}^*$	J-statistic	Cointegration	$\tilde{\alpha}^*$	J-statistic
Aluminum	Lead	0.241**	0.0604	(1,-0.278)	0.144	0.0543
		(0.114)	[0.225]		(0.123)	[0.232]
Copper	Aluminum	0.433***	0.0647	(1,-0.134)	0.355***	0.0658
		(0.128)	[0.177]		(0.136)	[0.117]
Aluminum	Tin	0.377	0.0473	(1,-0.203)	0.429*	0.0430
		(0.241)	[0.004]		(0.244)	[0.002]
Aluminum	Nickel	0.391***	0.0326	(1,9.784)	0.366***	0.0319
		(0.091)	[0.314]		(0.108)	[0.227]
Zinc	Aluminum	0.67**	0.0356	(1,-0.836)	0.467	0.0313
		(0.275)	[0.02]		(0.317)	[0.009]
Lead	Copper	0.562***	0.0447	(1,-1.013)	0.7***	0.0364
		(0.169)	[0.007]		(0.214)	[0.005]
Lead	Tin	0.701***	0.0583	(1,-0.548)	0.904***	0.0456
		(0.137)	[0.049]		(0.16)	[0.079]
Nickel	Lead	-0.126	0.0594	(1,-0.677)	0.078	0.0542
		(0.176)	[0.045]		(0.186)	[0.038]
Zinc	Lead	0.477***	0.0528	(1,-2.593)	0.497***	0.0573
		(0.136)	[0.076]		(0.141)	[0.029]
Copper	Tin	0.782***	0.0520	-	-	-
		(0.199)	[0.01]			
Copper	Nickel	0.681***	0.0525	(1,-0.412)	0.618***	0.0408
		(0.247)	[0.077]		(0.237)	[0.115]
Zinc	Copper	0.904***	0.0517	(1,-0.269)	0.732***	0.0435
		(0.142)	[0.024]		(0.161)	[0.023]
Tin	Nickel	0.604***	0.0573	(1,-0.192)	0.962***	0.0451
		(0.189)	[0.053]		(0.246)	[0.081]
Zinc	Tin	0.479**	0.0505	(1,-0.249)	0.714**	0.0459
		(0.232)	[0.004]		(0.292)	[0.002]
Zinc	Nickel	0.604***	0.0303	(1,-0.517)	0.552***	0.0234
		(0.18)	[0.036]		(0.186)	[0.024]

Notes: GMM estimation using equation (9) for Strong-form SCCF and the analogue equation for Weak-form SCCF.

Robust Standard Errors (HAC) are in parentheses and p-values are in brackets.

Table 4: Common-Cycle Tests - Metal Prices and Industrial Production (Quarterly)

		Strong-form SCCF			Weak-form SCCF	
$\Delta y_{1,t}$	$\Delta y_{2,t}$	$\tilde{\alpha}^*$	J-statistic	Cointegration	$\tilde{\alpha}^*$	J-statistic
Panel A - Sample: 1992Q1-2012Q1						
Aluminum	Global Industrial Production	-4.328***	0.0211	-	-	-
		(0.46)	[0.65]			
Lead	Global Industrial Production	-3.088**	0.0762	-	-	-
		(1.22)	[0.014]			
Copper	Global Industrial Production	-3.735***	0.1223	-	-	-
		(0.729)	[0.232]			
Nickel	Global Industrial Production	-4.792**	0.0623	-	-	-
		(1.886)	[0.027]			
Tin	Global Industrial Production	-3.407*	0.0628	-	-	-
		(0.697)	[0.18]			
Zinc	Global Industrial Production	-2.213	0.0448	-	-	-
		(1.571)	[0.06]			
Aluminum	US Industrial Production	-2.557***	0.0943	-	-	-
		(0.912)	[0.161]			
Lead	US Industrial Production	-2.457	0.0717	-	-	-
		(1.863)	[0.014]			
Copper	US Industrial Production	-4.226**	0.0422	-	-	-
		(1.705)	[0.06]			
Nickel	US Industrial Production	-3.576*	0.0638	-	-	-
		(2.156)	[0.021]			
Tin	US Industrial Production	-2.178*	0.0843	-	-	-
		(1.23)	[0.069]			
Zinc	US Industrial Production	-2.299	0.0396	-	-	-
		(1.675)	[0.067]			
Panel B - Sample: 1957Q1-2012Q1						
Aluminum	US Industrial Production	-0.896*	0.0653	(1,0.675)	-0.635	0.0633
		(0.515)	[0.117]		(0.496)	[0.089]
Lead	US Industrial Production	-2.083***	0.0745	(1,1.573)	-2.149***	0.0747
		(0.676)	[0.063]		(0.717)	[0.04]
Copper	US Industrial Production	-1.918***	0.0567	-	-	-
		(0.667)	[0.091]			
Nickel	US Industrial Production	-0.416	0.0552	(1,0.346)	-0.334	0.0524
		(0.74)	[0.214]		(0.692)	[0.182]
Tin	US Industrial Production	-1.606***	0.0459	-	-	-
		(0.582)	[0.018]			
Zinc	US Industrial Production	-2.39***	0.0625	(1,0.22)	-1.441*	0.0362
		(0.677)	[0.018]		(0.737)	[0.094]

Notes: GMM estimation using equation (9) for Strong-form SCCF and the analogue equation for Weak-form SCCF.
Robust Standard Errors (HAC) are in parentheses and p-values are in brackets.

Table 5: Common-Cycle Tests - Metal Prices (Annual)

		Strong-form SCCF			Weak-form SCCF	
$\Delta y_{1,t}$	$\Delta y_{2,t}$	$\tilde{\alpha}^*$	J-statistic	Cointegration	$\tilde{\alpha}^*$	J-statistic
Aluminum	Lead	-0.621***	0.0112	(1,3.63)	-0.601**	0.0111
		(0.116)	[0.876]		(0.239)	[0.752]
Aluminum	Copper	0.552***	0.0515	-	-	-
		(0.21)	[0.135]			
Aluminum	Tin	-0.02**	0.0454	(1,0.122)	-0.012	0.0301
		(0.008)	[0.778]		(0.008)	[0.867]
Aluminum	Nickel	0.54***	0.0638	(1,-0.399)	-0.043	0.0394
		(0.188)	[0.142]		(0.251)	[0.235]
Aluminum	Zinc	0.264*	0.0730	(1,-0.298)	-0.121	0.0429
		(0.151)	[0.096]		(0.132)	[0.201]
Copper	Lead	0.286**	0.0116	(1,2.557)	0.241	0.0115
		(0.117)	[0.533]		(0.363)	[0.262]
Tin	Lead	-	-	-	-	-
		-	-		-	-
Lead	Nickel	0.618***	0.0836	(1,0.311)	-0.134	0.0283
		(0.167)	[0.06]		(0.177)	[0.382]
Lead	Zinc	0.095	0.0297	-	-	-
		(0.1)	[0.36]		-	-
Copper	Tin	0.033	0.0579	(1,2.818)	0.095	0.0535
		(0.075)	[0.402]		(0.132)	[0.334]
Copper	Nickel	-0.223*	0.0367	-	-	-
		(0.134)	[0.265]		-	-
Copper	Zinc	-0.009	0.0408	(1,-0.973)	-0.193	0.0260
		(0.077)	[0.354]		(0.132)	[0.422]
Nickel	Tin	0.564***	0.0569	(1,2.95)	0.634**	0.0548
		(0.099)	[0.189]		(0.248)	[0.116]
Tin	Zinc	0.758***	0.0901	(1,0.372)	0.465***	0.0301
		(0.111)	[0.045]		(0.118)	[0.355]
Nickel	Zinc	-0.498***	0.0561	-	-	-
		(0.084)	[0.116]		-	-

Notes: GMM estimation using equation (9) for Strong-form SCCF and the analogue equation for Weak-form SCCF.

Robust Standard Errors (HAC) are in parentheses and p-values are in brackets.

Table 6: Multivariate Common-Cycle Test - Metal Prices and Industrial Production

Industrial Production/Sample starts	Global / 1992	USA / 1992	USA / 1957
Null Hypotheses: Number of cofeature vectors (s)	J-test	J-test	J-test
$s = 1$	0.278	18.85	19.722
	[0.598]	[0.016]	[0.183]
$s = 2$	7.386	19.404	34.854
	[0.117]	[0.054]	[0.091]
$s = 3$	14.221	18.894	46.301
	[0.115]	[0.274]	[0.141]
$s = 4$	20.57	16.848	54.368
	[0.196]	[0.817]	[0.347]
$s = 5$	27.165	27.745	63.214
	[0.348]	[0.682]	[0.609]
$s = 6$	30.288	30.769	82.73
	[0.737]	[0.919]	[0.549]

Notes: The p-values are in brackets.

Table 7: Correlations Between the Metal Prices and Co-variates

	Aluminum	Copper	Lead	Nickel	Tin	Zinc	BASISM	ISRETM	LRETB10	LTY	RV	SP500	TBL	USIP	VSRETM
Aluminum	1	0.427	0.671	0.605	0.420	0.110	-0.080	0.026	-0.041	-0.216	-0.193	-0.461	-0.095	-0.540	-0.210
Copper	-	1	0.561	0.217	0.249	0.302	0.257	0.043	-0.079	0.135	-0.062	-0.508	0.257	-0.548	-0.084
Lead	-	-	1	0.333	0.672	0.310	-0.091	0.028	-0.045	0.157	-0.128	-0.518	0.209	-0.540	-0.105
Nickel	-	-	-	1	0.165	0.378	-0.117	-0.007	-0.050	0.007	-0.024	-0.073	0.021	-0.087	-0.068
Tin	-	-	-	-	1	0.130	-0.052	-0.022	0.026	0.427	-0.157	-0.585	0.303	-0.498	-0.039
Zinc	-	-	-	-	-	1	-0.165	-0.051	0.032	0.238	0.001	0.019	0.231	0.000	0.041
BASISM	-	-	-	-	-	-	1	0.142	-0.141	0.015	0.075	-0.033	0.166	-0.009	0.116
ISRETM	-	-	-	-	-	-	-	1	-0.090	-0.086	-0.156	0.025	-0.074	0.010	0.113
LRETB10	-	-	-	-	-	-	-	-	1	-0.053	0.126	0.034	-0.065	0.059	0.022
LTY	-	-	-	-	-	-	-	-	-	1	-0.071	-0.505	0.838	-0.354	0.131
RV	-	-	-	-	-	-	-	-	-	-	1	0.295	-0.096	0.320	0.164
S&P500	-	-	-	-	-	-	-	-	-	-	-	1	-0.505	0.955	0.121
TBL	-	-	-	-	-	-	-	-	-	-	-	-	1	-0.435	0.089
USIP	-	-	-	-	-	-	-	-	-	-	-	-	-	1	0.173
VSRETM	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1

Notes: BASISM - Metals Prices Index Basis; ISRETM - 1 month excess spot returns of Metal Prices Index; LRETB10 - 10 years Treasury bond: 1 month excess returns; LTY - Long term yield; RV - Realized volatility on S&P 500; TBL - Treasury-bill rates 3 months; USIP - Industrial Production (USA); VSRETM - Volatility of 1 month excess spot returns of Metal Prices Index; SP500 - return on the SP 500 index.

Table 8: Mean Bias Significance Test (Monthly)

Aluminum

	Bias	t-statistic	p-value
1 step-ahead	0.012	0.048	0.481
2 step-ahead	0.041	0.081	0.468
3 step-ahead	0.070	0.091	0.464
4 step-ahead	0.111	0.108	0.457
5 step-ahead	0.156	0.122	0.451
6 step-ahead	0.200	0.131	0.448

Copper

	Bias	t-statistic	p-value
1 step-ahead	0.341	0.285	0.388
2 step-ahead	0.808	0.297	0.383
3 step-ahead	1.299	0.304	0.380
4 step-ahead	1.818	0.319	0.375
5 step-ahead	2.299	0.329	0.371
6 step-ahead	2.742	0.338	0.368

Lead

	Bias	t-statistic	p-value
1 step-ahead	-0.088	-0.424	0.336
2 step-ahead	-0.184	-0.401	0.344
3 step-ahead	-0.278	-0.387	0.349
4 step-ahead	-0.355	-0.365	0.358
5 step-ahead	-0.424	-0.346	0.365
6 step-ahead	-0.486	-0.329	0.371

Nickel

	Bias	t-statistic	p-value
1 step-ahead	0.257	0.387	0.350
2 step-ahead	0.648	0.432	0.333
3 step-ahead	1.040	0.439	0.330
4 step-ahead	1.628	0.506	0.306
5 step-ahead	2.244	0.562	0.287
6 step-ahead	2.838	0.605	0.273

Tin

	Bias	t-statistic	p-value
1 step-ahead	-1.254	-0.471	0.319
2 step-ahead	-2.641	-0.466	0.321
3 step-ahead	-4.059	-0.466	0.320
4 step-ahead	-5.276	-0.451	0.326
5 step-ahead	-6.406	-0.439	0.330
6 step-ahead	-7.367	-0.423	0.336

Table 9: Mean Bias Significance Test (Annual)

Aluminum			
	Bias	t-statistic	p-value
1 step-ahead	-276.842	-0.525	0.300
2 step-ahead	-444.360	-0.524	0.300
3 step-ahead	-486.603	-0.474	0.318
4 step-ahead	-487.179	-0.417	0.338
5 step-ahead	-450.445	-0.355	0.361

Copper			
	Bias	t-statistic	p-value
1 step-ahead	-233.249	-1.114	0.133
2 step-ahead	-400.706	-1.114	0.133
3 step-ahead	-525.299	-1.212	0.113
4 step-ahead	-619.110	-1.309	0.095
5 step-ahead	-689.353	-1.349	0.089

Lead			
	Bias	t-statistic	p-value
1 step-ahead	263.951	0.435	0.332
2 step-ahead	602.373	0.454	0.325
3 step-ahead	783.477	0.458	0.323
4 step-ahead	744.257	0.434	0.332
5 step-ahead	666.113	0.384	0.351

Nickel			
	Bias	t-statistic	p-value
1 step-ahead	-1140.074	-1.073	0.142
2 step-ahead	-1834.030	-1.220	0.111
3 step-ahead	-2542.124	-1.500	0.067
4 step-ahead	-3141.645	-1.793	0.037
5 step-ahead	-3495.002	-1.909	0.028

Tin			
	Bias	t-statistic	p-value
1 step-ahead	16049.680	0.845	0.199
2 step-ahead	37310.520	0.751	0.226
3 step-ahead	55645.350	0.743	0.229
4 step-ahead	53266.960	0.818	0.207
5 step-ahead	42927.230	0.983	0.163

Zinc			
	Bias	t-statistic	p-value
1 step-ahead	-276.644	-1.277	0.101
2 step-ahead	-458.376	-1.662	0.048
3 step-ahead	-586.994	-2.044	0.021
4 step-ahead	-702.797	-2.652	0.004
5 step-ahead	-749.982	-2.879	0.002

Table 10: Forecast Root-Mean-Squared-Error (Monthly)

Aluminium	BCAF	Weighted Average (MSE)	Average Forecast	Median	Best Model (BIC)	5 Best Models	10 Best Models
1 step-ahead	0.4898	0.5912	0.4858	0.4816	0.4711	0.4764	0.4838
2 step-ahead	0.7681	1.2979	0.7580	0.7555	0.7568	0.7567	0.7690
3 step-ahead	0.9963	2.2290	0.9747	0.9755	0.9932	0.9845	0.9919
4 step-ahead	1.1491	3.1522	1.1313	1.1362	1.1651	1.1475	1.1540
5 step-ahead	1.2666	4.3410	1.259	1.2652	1.2639	1.2725	1.2844
6 step-ahead	1.3415	5.4414	1.3449	1.3499	1.3853	1.3742	1.3721
R^2 with drift[1]	1.137%	-44.013%	2.751%	4.43%	8.535%	6.477%	3.561%
R^2 without drift[2]	1.022%	-44.182%	2.637%	4.318%	-3417.154%	-5957.989%	-4542.11%

Lead	BCAF	Weighted Average (MSE)	Average Forecast	Median	Best Model (BIC)	5 Best Models	10 Best Models
1 step-ahead	0.6239	1.0408	0.6245	0.6266	0.6198	0.6189	0.6209
2 step-ahead	1.0730	2.6032	1.0721	1.0810	1.0506	1.0524	1.0551
3 step-ahead	1.3457	4.3603	1.3259	1.3408	1.278	1.2782	1.2838
4 step-ahead	1.5483	6.1930	1.5137	1.5311	1.4504	1.4516	1.4604
5 step-ahead	1.7696	8.0726	1.7201	1.7417	1.6056	1.6702	1.6683
6 step-ahead	1.9877	9.9262	1.9242	1.9516	1.7911	1.8775	1.8617
R^2 with drift[1]	11.149%	-147.265%	10.983%	10.39%	12.32%	12.571%	11.998%
R^2 without drift[2]	10.952%	-147.813%	10.785%	10.192%	-690.959%	-667.622%	-759.259%

Copper	BCAF	Weighted Average (MSE)	Average Forecast	Median	Best Model (BIC)	5 Best Models	10 Best Models
1 step-ahead	1.6483	2.8374	1.6276	1.6355	1.6820	1.6489	1.6532
2 step-ahead	2.8148	6.9168	2.7387	2.7431	2.9217	2.8337	2.8182
3 step-ahead	3.7865	11.7815	3.6352	3.6507	4.0372	3.8192	3.7823
4 step-ahead	4.5263	16.9421	4.2829	4.3012	4.8954	4.5132	4.4303
5 step-ahead	5.1093	22.0737	4.7402	4.7686	5.3125	4.9581	4.8823
6 step-ahead	5.6178	26.9173	5.1021	5.1250	5.4809	5.2830	5.2055
R^2 with drift[1]	12.099%	-160.474%	14.295%	13.462%	8.469%	12.041%	11.58%
R^2 without drift[2]	11.866%	-161.166%	14.067%	13.232%	-514.624%	-591.317%	-634.495%

Nickel	BCAF	Weighted Average (MSE)	Average Forecast	Median	Best Model (BIC)	5 Best Models	10 Best Models
1 step-ahead	8.0858	8.0375	8.1756	8.0614	8.0352	8.1274	8.1199
2 step-ahead	14.3789	14.1544	14.7694	14.5915	14.7195	14.7049	14.6288
3 step-ahead	19.6148	19.1494	20.4887	20.2466	20.4566	20.3554	20.3190
4 step-ahead	23.2824	22.552	24.5965	24.2389	24.4405	24.4277	24.3303
5 step-ahead	26.2060	25.1953	27.8735	27.4466	27.8046	27.9348	27.7182
6 step-ahead	29.0008	27.7018	31.0609	30.5918	31.2149	31.3089	30.9118
R^2 with drift[1]	18.847%**	19.812%**	17.035%**	19.335%**	19.859%**	18.01%**	18.161%**
R^2 without drift[2]	18.685%**	19.652%**	16.869%**	19.174%**	-386.86%	-402.051%	-424.178%

Tin	BCAF	Weighted Average (MSE)	Average Forecast	Median	Best Model (BIC)	5 Best Models	10 Best Models
1 step-ahead	3.0223	10.8160	2.9422	2.9463	3.0542	3.0009	2.9936
2 step-ahead	4.8095	23.9791	4.5809	4.6004	4.7888	4.7871	4.7493
3 step-ahead	7.1636	37.6754	6.7924	6.8712	6.7907	7.0074	6.9465
4 step-ahead	8.7257	51.4525	8.194	8.3404	8.4795	8.3109	8.3409
5 step-ahead	10.2284	65.4156	9.5326	9.7679	9.7563	9.7035	9.6948
6 step-ahead	10.9799	79.2277	9.982	10.3178	10.5761	10.0995	10.1324
R^2 with drift[1]	10.283%**	-1049.052%**	14.976%**	14.74%**	8.378%**	11.546%**	11.976%**
R^2 without drift[2]	9.672%**	-1056.874%*	14.397%**	14.16%**	-776.537%	-780.975%	-926.804%

$T = 528(2008M12)$, $T_1 = 204(1981M12)$, $T_2 = 312(1990M12)$

BCAF: Bias-Corrected Average Forecast; MSE: Mean Squared Error; The best model, the 5 best models, and 10 best models were chosen using the in-sample BIC criterium.

1 - The Benchmark Model is a Random Walk with drift.

2 - The Benchmark Model is a Random Walk without drift. * and ** mean 10% and 5% of significance respectively.

Table 11: Forecast Root-Mean-Squarred-Error (Annual)

Aluminum	BCAF	Weighted Average (MSE)	Average Forecast	Median	Best Model (BIC)	5 Best Models	10 Best Models
1 step-ahead	37.206	40.657	38.358	40.861	37.900	36.715	36.116
2 step-ahead	59.374	70.052	63.762	66.324	67.485	65.130	63.431
3 step-ahead	65.961	82.476	70.684	75.429	77.922	74.876	72.269
4 step-ahead	69.016	93.859	75.082	81.393	83.052	79.635	76.129
5 step-ahead	69.13	99.694	75.681	84.285	87.665	83.348	79.164
R^2 with drift[1]	8.589%**	0.11%**	5.759%**	-0.39%**	6.884%**	9.795%**	11.267%**
R^2 without drift[2]	1.022%	-44.182%	2.637%	4.318%	-3417.154%	-5957.989%	-4542.11%

Lead	BCAF	Weighted Average (MSE)	Average Forecast	Median	Best Model (BIC)	5 Best Models	10 Best Models
1 step-ahead	87.215	88.310	95.419	102.963	98.490	96.390	96.919
2 step-ahead	164.499	161.849	182.150	199.458	188.621	180.429	180.717
3 step-ahead	194.459	199.206	231.145	247.535	225.020	218.144	219.826
4 step-ahead	197.945	217.125	261.908	280.107	231.141	228.625	232.594
5 step-ahead	204.004	223.123	274.640	288.786	226.445	230.735	237.098
R^2 with drift[1]	15.895%**	14.838%**	7.982%**	0.708%**	5.021%**	7.046%**	6.536%**
R^2 without drift[2]	13.59%**	12.505%**	5.461%**	-2.013%**	-184.711%	-336.515%	-322.209%

Copper	BCAF	Weighted Average (MSE)	Average Forecast	Median	Best Model (BIC)	5 Best Models	10 Best Models
1 step-ahead	17.665	17.128	13.821	12.165	12.414	15.894	15.021
2 step-ahead	44.467	46.675	35.102	34.609	31.649	32.732	32.260
3 step-ahead	75.580	83.605	59.676	62.879	55.337	55.212	55.486
4 step-ahead	106.519	121.296	100.515	100.425	81.601	82.790	84.118
5 step-ahead	128.250	150.484	128.840	128.84	100.425	102.647	105.021
R^2 with drift[1]	-25.879%	-22.051%	1.512%**	13.311%**	11.54%**	-13.257%**	-7.04%**
R^2 without drift[2]	-18.152%	-14.559%	7.557%**	18.632%**	-644.108%	-569.998%	-698.1%

Nickel	BCAF	Weighted Average (MSE)	Average Forecast	Median	Best Model (BIC)	5 Best Models	10 Best Models
1 step-ahead	780.106	791.635	784.654	760.113	728.444	739.280	747.302
2 step-ahead	1727.239	1840.379	1825.805	1898.553	1475.888	1421.853	1441.278
3 step-ahead	1967.854	2148.107	2060.348	2160.290	1696.273	1587.752	1607.805
4 step-ahead	2263.931	2663.476	2317.876	2342.069	2065.623	1971.324	2000.096
5 step-ahead	2678.195	3306.581	2872.925	2839.961	2335.586	2306.738	2352.873
R^2 with drift[1]	7.105%**	5.733%**	6.564%**	9.486%**	13.257%**	11.967%**	11.012%**
R^2 without drift[2]	6.85%**	5.473%**	6.307%**	9.237%**	-836.53%	-1776.489%	-2341.942%

Tin	BCAF	Weighted Average (MSE)	Average Forecast	Median	Best Model (BIC)	5 Best Models	10 Best Models
1 step-ahead	64.457	72.979	64.111	67.698	74.204	72.228	72.420
2 step-ahead	159.763	168.255	141.708	147.906	161.558	148.089	148.736
3 step-ahead	221.401	216.056	176.578	183.942	192.313	179.855	179.121
4 step-ahead	203.960	258.841	204.587	210.945	201.923	199.095	197.25
5 step-ahead	201.165	313.839	226.033	229.935	204.341	214.522	211.547
R^2 with drift[1]	8.486%**	-3.614%**	8.978%**	3.884%**	-5.353%**	-2.548%**	-2.82%**
R^2 without drift[2]	7.432%**	-4.808%**	7.929%**	2.776%**	-83.473%**	-200.372%	-177.335%

Zinc	BCAF	Weighted Average (MSE)	Average Forecast	Median	Best Model (BIC)	5 Best Models	10 Best Models
1 step-ahead	141.114	143.832	195.797	161.538	150.502	149.399	153.070
2 step-ahead	244.172	251.799	479.166	292.524	254.273	255.094	267.004
3 step-ahead	245.351	266.658	875.957	291.309	247.958	248.965	260.206
4 step-ahead	242.105	279.292	250.617	253.082	233.096	236.309	241.427
5 step-ahead	250.851	301.163	264.693	269.355	235.248	242.127	248.668
R^2 with drift[1]	23.813%**	22.346%**	-5.711%**	12.786%**	18.744%**	19.34%**	17.358%**
R^2 without drift[2]	23.123%**	21.643%**	-6.667%**	11.997%**	-1223.213%	-1638.582%	-2269.327%

Notes: $T = 111(2011)$, $T_1 = 35(1935)$, $T_2 = 70(1970)$

BCAF: Bias-Corrected Average Forecast; MSE: Mean Squared Error; The best model, the 5 best models, 10 best models were chosen using the in-sample BIC criterium.

1 - The Benchmark Model is a Random Walk with drift.

2 - The Benchmark Model is a Random Walk without drift. * and ** mean 10% and 5% of significance respectively.

Table 12: Best Models and Predictors (Monthly)

Aluminum					
Best Models		Best Predictor			
AR	0.79%	RV	0.00%	S&P 500 TBL	0.23%
VAR	2.03%	SP500	0.45%	RV SP500 TBL	0.00%
VECM	0.23%	TBL	0.00%	USIP	15.03%
VECM Restricted (Common-Cycles)	96.95%	RV SP500	0.00%	No Extra Predictors	84.29%
		RV TBL	0.00%		
Lead					
Best Models		Best Predictor			
AR	0.34%	RV	0.11%	SP500 TBL	0.00%
VAR	3.05%	SP500	0.00%	RV SP500 TBL	0.00%
VECM	18.42%	TBL	0.00%	USIP	24.75%
VECM Restricted (Common-Cycles)	78.19%	RV SP500	0.11%	No Extra Predictors	75.03%
		RV TBL	0.00%		
Copper					
Best Models		Best Predictor			
AR	9.27%	RV	2.82%	SP500 TBL	0.00%
VAR	0.23%	SP500	0.00%	RV SP500 TBL	0.00%
VECM	13.33%	TBL	1.24%	USIP	3.28%
VECM Restricted (Common-Cycles)	77.18%	RV SP500	2.49%	No Extra Predictors	87.80%
		RV TBL	2.37%		
Nickel					
Best Models		Best Predictor			
AR	7.80%	RV	0.00%	SP500 TBL	0.00%
VAR	4.86%	SP500	0.00%	RV SP500 TBL	0.00%
VECM	19.10%	TBL	0.00%	USIP	33.90%
VECM Restricted (Common-Cycles)	68.25%	RV SP500	0.00%	No Extra Predictors	66.10%
		RV TBL	0.00%		
Tin					
Best Models		Best Predictor			
AR	2.03%	RV	0.11%	SP500 TBL	0.11%
VAR	1.92%	SP500	0.90%	RV SP500 TBL	0.45%
VECM	17.85%	TBL	0.11%	USIP	60.34%
VECM Restricted (Common-Cycles)	78.19%	RV SP500	0.00%	No Extra Predictors	37.63%
		RV TBL	0.00%		

RV - Realized volatility on S&P 500; TBL - Treasury-bill rates 3 months; USIP - Industrial Production (USA); SP500 - return on the SP 500 index.

Table 13: Best Models and Predictors (Annual)

Aluminum					
Best Models		Best Predictor			
AR	63.11%	SP500	0.00%	SVAR USIP	4.44%
VAR	11.56%	SVAR	0.00%	SP500 SVAR	0.00%
VECM	25.33%	USIP	11.56%	No Extra Predictors	84.00%
Lead					
Best Models		Best Predictor			
AR	44.44%	SP500	0.00%	SVAR USIP	0.00%
VAR	46.67%	SVAR	27.11%	SP500 SVAR	0.00%
VECM	8.89%	USIP	6.22%	No Extra Predictors	66.67%
Copper					
Best Models		Best Predictor			
AR	55.56%	SP500	2.22%	SVAR USIP	0.00%
VAR	28.44%	SVAR	29.78%	SP500 USIP	0.00%
VECM	16.00%	USIP	8.00%	No Extra Predictors	62.22%
Nickel					
Best Models		Best Predictor			
AR	12.89%	SVAR	8.00%	No Extra Predictors	84.44%
VAR	35.11%	USIP	3.11%		
VECM	52.00%	SVAR USIP	4.44%		
Tin					
Best Models		Best Predictor			
AR	10.22%	SVAR	6.67%	SP500 SVAR	0.00%
VAR	40.44%	USIP	18.22%	No Extra Predictors	73.33%
VECM	49.33%	SVAR USIP	1.78%		
Zinc					
Best Models		Best Predictor			
AR	12.44%	SP500	0.00%	SVAR USIP	0.00%
VAR	87.11%	SVAR	0.00%	SP500 SVAR	0.00%
VECM	0.44%	USIP	0.89%	No Extra Predictors	99.11%

Notes: SVAR: Stock Variance - average squared daily returns on the S&P 500; USIP - Industrial Production (USA); SP500 - return on the SP 500 index.