Journal of International Money and Finance 42 (2014) 289–309

Contents lists available at ScienceDirect

Journal of International Money and Finance

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

A factor model for co-movements of commodity prices

Kenneth D. West [a,*], Ka-Fu Wong [b]

[a] University of Wisconsin, USA
[b] University of Hong Kong, Hong Kong

Keywords:
Factor model
Commodity price
Out-of-sample performance
Forecast
Co-movement

A B S T R A C T

We fit a factor model to two monthly panels of deflated prices of energy, metals and agricultural commodities. Prices consistently display a tendency to revert towards the factor, though the speed of reversion to the factor is slow. Using both in- and out-of-sample metrics, we compare the factor model to that of a "no change" model and to two simple models that tie changes in commodity prices to percentage change in either global industrial production or the U.S. dollar. The factor model does relatively well at long (12 month) horizons. In terms of commodities, the factor model's performance is best for energy prices, worst for metals, with agricultural prices falling in between.

1. Introduction

In this paper we explore modeling co-movements of real commodity prices via a static factor model. We use in- and out-of-sample methods to evaluate the explanatory power of the factor model relative to some simple alternatives.

A substantial literature views co-movements as a central and distinctive characteristic of commodity prices. This is evident in much popular discussion of commodity prices, e.g. July 2012 articles in the *Economist* and the *Financial Times*.[1] Scholarly studies referencing "supercycles" that may have

* Corresponding author.
E-mail address: kdwest@wisc.edu (K.D. West).

[1] *Economist* July 28, 2012, "Downhill cycling: A peak may be in sight for commodity prices," www.economist.com/node/21559647; *Financial Times* July 17, 2012, "Supercycle runs out of steam – for now", www.ft.com/intl/cms/s/0/ba9b6d96-cb3d-11e1-916f-00144feabdc0.html#axzz27mmmjhV0.

0261-5606/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.jimonfin.2013.08.016

characterized real commodity price movements as long ago as the 19th century include Cuddington John and Jerrett (2008) and Erten and Ocampo (2012). And scholarly literature focusing, as we do, on recent decades has used real co-movements, or nominal co-movements conditioned on inflation, in a variety of ways: to consider whether such co-movements are excessive in a precise sense (Ohashi and Okimoto (2012)); to forecast inflation (Gospodinov and Ng (2013)); and to evaluate the extent to which oil price movements translate into movements in a wide range of commodity prices (Baffes (2007)).

In recent years, factor models have been widely used by economists.[2] Recent papers that have applied factor models to a panel of nominal commodity prices include Byrne et al. (2011) and Chen et al. (2012). Using a factor augmented vector autoregression, Byrne et al. (2011) find that after deflating by CPI, the deflated factor responds in a sensible way to shocks to interest rates and to risk. Chen et al. (2012) successfully use the factor to forecast the US dollar. We interpret these papers as indicating that a factor extracted from a panel of commodity prices reflects some central characteristics of commodity prices.

In this paper, we also fit a factor model to a panel of commodity prices. In contrast to the cited papers, our focus is not on how this factor interacts with macroeconomic variables. Nor do we attempt to model dynamics through lagged factors. Rather, we aim to evaluate the extent to which a simple static factor model captures co-movements of commodity prices. By "static" we mean that we do not use lags or attempt to model dynamics in the factor. We focus on static models for logical simplicity and also on the presumption that simpler models work relatively well in forecasting. We use in- and out-of-sample evidence to gauge the extent to which commodity prices revert towards a factor. We posit that commodity prices that are above the factor will, on average, tend to fall, while those below the factor will tend to rise.

We use a monthly panel of US dollar commodity prices deflated by lagged US CPI. We deflate to control for what we view as a likely but uninteresting tendency for general increases in the overall price level to be reflected in general upward movement of commodity prices. (See Gillman and Nakov (2009) for empirical evidence on such a tendency for oil prices.) We use two samples, with beginning points chosen so that we could include a cross-section of commodity prices. A 1989–2012 sample includes prices of 13 commodities: four energy products, six industrial products, two agricultural products and gold. 1996–2012 sample includes nine additional commodity prices not available in 1989. The 22 commodity prices in this sample include five energy products, seven industrial products, nine agricultural products and gold.

As in Byrne et al. (2011) and Chen et al. (2012), the first principal component from each panel correlates positively with a measure of real activity (global industrial production–IP for short) and negatively with the US dollar (ER as in "exchange rate" for short). But the correlation is also low, meaning the factor either impounds useful information not in those two macro variables or it omits useful information found in those variables (or, most likely, both). We document that reversion towards the factor is a robust characteristic of the data. For the whole sample and subsamples as short as six years, fixed effects regressions with the factor as an explanatory variable imply reversion towards the factor. This applies whether or not IP or ER is included in the regression along with the factor. Hence it appears that the factor has some information not also present in these two macro variables.

Beginning with samples of size six years, we estimate a sequence of regressions in which we use the factor to forecast subsequent percentage changes in the prices of each of the commodities in our sample. The correlation between the factor model prediction and out-of-sample realization of subsequent commodity prices changes is positive for the vast majority of commodities: the reversion towards the factor found in sample also holds out-of-sample.

We compare the predictive ability of our factor model to that of the two macro variables and to a "no change" forecast. A no change forecast is optimal if the (log) level of the commodity price follows a random walk. For comparison with the macro variables, we forecast changes in deflated commodity prices. For comparison with a no change forecast, we forecast changes in both real and nominal commodity prices. (We deflate by lagged rather than contemporaneous CPI so that we can use our

[2] See Stock and Watson (2006) for a review. We use "factor" model as shorthand for the class of models that include what, in a more nuanced discussion, would instead be called index models or principal components models.

model to forecast either deflated or nominal changes in commodity prices.). While no model dominated every other by every measure we look at, the general pattern of results is as follows. (1)On balance, the factor model, IP and no change models are about comparable. The model with the exchange rate performs worse than the other models. (2)Overall, the factor model does better than any of the other models at 12 month ahead predictions. It does worse than does the model with IP, or the no change model, at a 1 month horizon. It does about comparably to those two models at a 3 month horizon. (3)In terms of commodities, the factor model's performance is best for energy prices, worst for metals, with agricultural prices falling in between.

We conclude that the factor, which is constructed as a weighted average of commodity prices, has information about the evolution of commodity prices not impounded in current values of industrial production or exchange rates. We leave for future research the task of identifying whether that information could be captured by other macro variables or reflects objects such a risk premia that are not easily measured.

We close this introduction by acknowledging some limitations that we also hope to address in future research. First, we do not attempt to use industry data such as inventories nor financial market data such as futures prices, as is done in the comprehensive analysis for oil in Alquist et al. (2013) and for many commodities in Baffes (2013). Second, there are many other macro variables that one could consider in addition to industrial production and exchange rates. So we acknowledge that our examination is hardly definitive. Finally, and related, we do not attempt to use structural macroeconomic models to interpret our factor or, more generally, our results. See Harris et al. (2010) for one of many papers with a macro model that includes oil prices.

Section 2 describes our empirical models, Section 3 our data and forecast evaluation techniques. Section 4 presents factor model results, Section 5 forecasting comparisons of the factor model to other models, Section 6 a summary of some additional results. Section 7 concludes. An appendix contains detailed empirical results that are omitted from the paper to save space.

2. The factor model and forecasting evaluation

Our baseline model uses two factors, estimated by principal components.[3] We repeated all work with one factor, with slightly inferior results that are summarized below. Let there be n commodities, where $n = 22$ in our shorter sample (1996–2012) and $n = 13$ in our longer sample (1989–2012). Our first step is to estimate the factor and factor loadings from the levels of the real commodity prices. Let P_{it} be the US dollar price of commodity i. Let $P_{\mathrm{CPI},t}$ be the US CPI. For a lag h, let $p_{\mathrm{it}} = 100 \times \ln(P_{\mathrm{it}}/P_{\mathrm{CPI},t\text{-}h})$ be the real price of commodity i (the reason for lagging P_{CPI} by h months is given below). For commodity i, $i = 1, ...,n$, the factor model is

$$\begin{aligned} p_{\mathrm{it}} &= \text{constant} + \delta_{1i} f_{1t} + \delta_{2i} f_{2t} + v_{\mathrm{it}} \\ &= \text{constant} + F_{\mathrm{it}} + v_{\mathrm{it}}. \end{aligned} \tag{2.1}$$

The factors f_{1t} and f_{2t} are unobserved. The constant is chosen so that v_{it} has mean zero.

In contrast to research such as Groen and Pesenti (2010), the factor is not constructed from a set of fundamental variables reflecting the state of the economy or of supply and demand in a particular industry. Instead, like Byrne et al. (2011) and Chen et al. (2012), we construct the factor from the commodity prices themselves.

Some previous research (e.g., Dahl and Iglesias (2009)) has found real commodity prices to be I(1), and the informal examination of our data presented in the next section suggests that our data are also I(1) or nearly so. Our procedures are applicable with or without unit roots. If prices are I(1), then f_1, f_2 and F_i are I(1); if prices are stationary, then f_1, f_2 and F_i are stationary. If the data are I(1), we assume that p_{it}-F_{it} is stationary, i.e., prices are cointegrated with the factors. Here and throughout, we do not attempt to test for unit roots or cointegration. See Bai (2004) and Stock and Watson (2006) on estimation of factor models with unit root or with stationary data.

[3] Parts of the discussion in this section are based on Engel et al. (2012).

We posit that F_{it} is a central tendency towards which commodity prices revert. After controlling for means via a fixed effect, we hypothesize that when (demeaned) p_{it} is above F_{it}, p_{it} subsequently tends to fall; when p_{it} is below F_{it}, p_{it} subsequently tends to rise.[4] Thus if the data are I(1), F_{it}-p_{it} is an error correction term.

Factors were constructed by principal components, with normalizations as follow: the factors $\widehat{f}_{1t}$ and $\widehat{f}_{2t}$ have mean zero; the factor loadings $\widehat{\delta}_{1i}$ and $\widehat{\delta}_{2i}$ satisfy $\sum_{i=1}^{n}\widehat{\delta}_{1i}^2 = \sum_{i=1}^{n}\widehat{\delta}_{2i}^2 = 1$ (n = number of commodities) with $\widehat{\delta}_{11} > 0$ and $\widehat{\delta}_{21} > 0$ (i.e., the sign of the loading on the first commodity in our list, which happens to be aluminum, is normalized to be positive). The factors and factor loadings were used to construct $\widehat{F}_{it} = \widehat{\delta}_{1i}\widehat{f}_{1t} + \widehat{\delta}_{2i}\widehat{f}_{2t}$.

We used $\widehat{F}_{it}$ to model h period changes in commodity prices, $h = 1$, 3 or 12 months. We estimate

$$p_{it+h} - p_{it} = \alpha_i + \beta\left(\widehat{F}_{it} - p_{it}\right) + u_{it+h}, \tag{2.2}$$

where α_i is a fixed effect for commodity i. Our presumption that commodity prices revert towards rather than away from F implies that we expect the estimates of $\widehat{\beta}$ to be positive.

For each of two possible measures of aggregate economic activity x_t, we also estimate

$$p_{it+h} - p_{it} = \alpha_i + \beta\left(\widehat{F}_{it} - p_{it}\right) + \gamma x_t + u_{it+h}, \tag{2.3}$$

and

$$p_{it+h} - p_{it} = \alpha_i + \gamma x_t + u_{it+h}. \tag{2.4}$$

The two alternative measures of x_t are:(1)monthly percentage change in industrial production in the OECD and 6 emerging markets (Brazil, China, India, Indonesia, the Russian Federation and South Africa), as measured by the OECD,[5] and (2)the monthly percentage change in the trade-weighted US dollar, obtained from FRED (series DTWEXM). We use these variables because measures of aggregate output or exchange rates have appeared in earlier studies of commodity prices (e.g., Gilbert (1989) and Borensztein and Reinhart (1994) as two relatively early papers, Chen et al. (2012) and Baffes and Dennis (2013) as two relatively recent papers). We employ these two alternative measures to gauge the explanatory power of a factor model.

Our analysis begins with estimates of (2.2) and (2.3) for horizons of 1, 3 and 12 months, using all available data. Those comparisons, while useful, gave limited insight into the value added of our factor model. So we used partial sample estimates and forecasts to evaluate our factor model and compare it to some alternatives. We did so in part because some literature finds that co-movements across commodity prices change over time,[6] suggesting the importance of trying various samples to insure that results are not specific to a particular sample. We also did so in part because financial prices are difficult to forecast and we consider prediction a strong test of our or any other model for commodity prices. We used the sequence of estimates to informally consider robustness of our full sample estimates of β and of our presumption that commodity prices would tend to fall/rise when (demeaned) p_{it} was above/below F_{it}. As well, we compared forecasts from the factor model (2.2) to those that rely on a measure of aggregate economic activity (2.4) and to a "no change" forecast.

We arbitrarily began our partial samples 72 months from the start of the sample. Let us illustrate the mechanics of our estimation and forecasting algorithm using a three month horizon ($h = 3$), for the first estimate for the 13 commodity sample. For a three month percentage change in real commodity price,

[4] Such a pattern has been found for aggregate CPI indices in OECD countries by Ciccarelli and Mojon (2010): CPI tends to rise in countries whose inflation rate is higher than the cross-country average, fall in countries whose inflation rate is lower than average.

[5] This series has been discontinued by the OECD. Christiane Baumeister kindly supplied data through 2011:9. Hence all analysis with industrial production data end in 2011:9.

[6] See, e.g., Tang and Xiong (2010) and Ohashi and Okimoto (2012), who work in terms of correlations or conditional correlations, or Baffes (2013), who offers an interesting non-parametric measure that considers up and down movements.

we define the real price as $p_{it} = 100 \times \ln(P_{it}/P_{CPI,t-3})$. As shown in (2.5) below, we use 72 months of data from 1989:6 to 1995:5 to estimate the factor and factor loadings, and construct $\widehat{F}_{it}$ for $i = 1,...,13$. The estimation technique is principal components.

```
|_____________________________________|_____|_____|
89:6                                  95:2  95:5  95:8
|—data used to estimate panel regression —|
|————————data used to estimate factor ——————/
```
(2.5)

We then use right hand side data from 1989:6 to 1995:2 to estimate a fixed effects panel data regression

$$p_{it+3} - p_{it} = \alpha_i + \beta\left(\widehat{F}_{it} - p_{it}\right) + u_{it+3}, t = 1989{:}6, \ldots, 1995{:}2. \tag{2.6}$$

We use 1995:5 data to predict the 3 month change in p:

$$\text{Prediction of}\left(p_{i,1995:8} - p_{i,1995:5}\right) = \widehat{\alpha}_i + \widehat{\beta}\left(\widehat{F}_{i,1995:5} - p_{i,1995:5}\right). \tag{2.7}$$

We then add an observation to the end of the sample, and repeat. For a 3 month horizon, this resulted in 208 predictions (first prediction is for $t+3 = 1995{:}8$, last is for 2012:11). The comparable figures for $h = 1$ month and $h = 12$ month horizons are 210 (first prediction is for $t+1 = 1995{:}6$, last = 2012:11) and 199 (first prediction is for $t+12 = 1996{:}5$, last = 2012:11). For the shorter sample with $n = 22$ commodities, the base month for the first prediction was 2002:6 (instead of 1995:5), with number of predictions 125, 123 and 114 predictions for $h = 1$, 3 and 12.

As is indicated by this discussion, the recursive method is used to generate partial sample estimates and predictions: observations are added to the end of the estimation sample, so that the sample size used to estimate factors and panel data regressions grows. The direct (as opposed to iterated) method is used to make multiperiod predictions.

One measure of how robustly prices revert towards $\widehat{F}_{it}$ is whether $\widehat{\beta}$ is consistently positive as we generate partial sample estimates. A second, and perhaps stronger measure, is whether realized prices tended to move towards $\widehat{F}_{it}$. The literature on speculative and financial prices is replete with examples in which in-sample estimates accorded with an investigator's presumption (in our case, $\widehat{\beta} > 0$) but out-of-sample predictions did not. So we compute and present the correlation between prediction and out-of-sample realization for each commodity and horizon and each sample.

As noted above, we compare our factor model to some alternatives. One alternative is a "no change" model, that is, a model for growth rates consistent with the log level of either the real or nominal price following a random walk without drift. (Of course, a no change model for real prices implies a different data generating process than does a no change model for nominal prices. See below on how we used our factor model to generate predictions of nominal commodity price changes.) In a no change model, the forecasted change in the commodity price is zero. In our discussion, we focus on the no change model for real prices, concisely summarizing the very similar results we obtained when we compared the factor model to a no change model for nominal prices.

One measure of forecast performance is root mean squared prediction error (RMSPE). For a no change model for deflated commodity prices, for example, MSPE for $h = 3$ is computed as

$$\frac{1}{208}\sum_t (p_{it+3} - p_{it})^2, \tag{2.8}$$

where, in the $n = 13$ commodity sample, the sum runs over the 208 predictions from $t+3 = 1995{:}8$ to $t+3 = 2012{:}11$. A second measure of forecast performance is directional accuracy, i.e., the fraction of predictions that were the correct sign. If commodity price changes follow a no change model and are symmetric around a zero mean, a coin flip is expected to produce correct signs half the time. Hence good performance in directional accuracy means getting the correct sign in more than half the predictions.

Additional forecasts are obtained in straightforward fashion from (2.4), for the two measures of x_t listed above: percentage change in worldwide industrial production and trade weighted US exchange rate.

We present relative RMSPE's: Theil's U-statistic, the ratio of the RMSPE from our factor model to the alternative model. In terms of point estimates, our measure of success for the factor model is to produce a U-statistic less than one. This means that the RMSPE is lower for the factor model than for the alternative model, i.e., the no change model or model relying in industrial production or exchange rates. A U-statistic of exactly 1 indicates that the sample RMSPEs from the factor model and the alternative model are the same.

When comparing our factor model to another model, we computed t-statistics on equality of RMSPEs. These were rarely significant at traditional levels, so we do not always report them. How we computed them depends on whether the alternative model is nested in the factor model. When the alternative involves another aggregate measure, or a no change model for nominal exchange rates, we compute t-statistics by computing the Diebold-Mariano-West (DMW) test for equality of MSPEs. West (1996, 2006) shows that because the loss function is MSPE, and because the models are not nested, the fact that models rely on estimated regression parameters is irrelevant asymptotically and hence DMW is asymptotically valid. A no change model the real commodity price is nested in our factor model (it results when $\alpha_i = \beta = 0$). So t-statistics for equality of MSPEs from the factor model and a no change model for real commodity price was computed as in Clark and West (2006). For both nested and non-nested comparisons, we compute standard errors following Newey and West (1994).

For the nested comparison, with 13 (or 22) commodities rather than just a single commodity, it is possible that one or more test statistics will be significant even if in population the factor model's RMSPE is not smaller than that of the no change model for any of the 13 commodity prices. We guarded against this possibility using bootstrapping and the logic of Hubrich and West (2010), on occasion reporting the p-value for the "max t-statistic."

We mentioned above that we use our factor model to predict not only real but nominal commodity price changes. Here are some details. Because we lagged the CPI by h periods in construction of the left hand side variable in (2.2), the only datum not already realized at the date of the prediction is the h month ahead value of the nominal commodity price. Hence we can and do use our estimates to analyze changes in nominal commodity prices, conditional on the overall price level. For example, the prediction of the nominal price change between 1995:5 and 1995:8 (i.e., of $100 \times [\ln(P_{i,1995:8}) - \ln(P_{i,1995:5})]$) is constructed by adding the percentage change in CPI between 1995:2 and 1995:5 to the left hand side of (2.7). Note that this change in CPI has already been realized in 1995:5 and thus we can use the term "prediction" without abuse of language.

3. Commodity price data

Our commodity price data are from Bloomberg. The daily Bloomberg data were converted to monthly by sampling the last day of the month. We began with the list of commodities in the IMF's commodity price database. With exceptions given in a footnote,[7] we used all commodity price data from the list that we were able to locate in Bloomberg that were available at least from 1996 forward. New commodity prices were added to Bloomberg at various dates. By our 1996 date we had the 22 commodities listed in Table 1. A subset of these commodities were available in 1989. These 13 commodities are noted with a "y" in the "longer sample" column in Table 1. In the end, our longer sample, covering 13 commodities, is 1989:6-2012:11. Our shorter sample, covering 22 commodities, is 1996:7-2012:11. Both panels are balanced. That is, we did not attempt to combine longer and shorter data series in a single panel. With two exceptions, the Bloomberg data were in US dollars. We used daily exchange rates obtained from the IMF web site to convert to US dollars those two exceptions, which were coco (denominated in SDR) and palm oil (denominated in Malaysian ringitt).

Basic statistics for the resulting levels and first differences are in Table 2, when real prices are computed by deflating by the previous month's US CPI and the longer sample is used when data from the

[7] The exceptions are: rice, cobalt, poultry, fine wool, coarse wool and uranium. We thank John Baffes for advising us that the prices for these commodities are not set in liquid markets.

Table 1
List of commodities and Bloomberg ticker.

			Longer sample?
(1)	Aluninum	LMAHDY	y
(2)	Brent Crude	EUCRBRDT	y
(3)	Coco	COCOCPSD	
(4)	Coffee arabica	COFECPAR	
(5)	Coffee robusta	COFECPRB	
(6)	Copper	LMCADY	y
(7)	Corn	CORNLA2Y	
(8)	Cotton	COTLOOKA	
(9)	Gasoline	MOIGC87P	y
(10)	Gold	GOLDLNPM	y
(11)	Heating Oil	NO2INYPR	y
(12)	Natural Gas	NGUSHHUB	
(13)	Lead	LMPBDY	y
(14)	Nickel	LMNIDY	y
(15)	Palm oil	KO1	
(16)	Rubber	RG1	
(17)	Soybean meal	SM1	y
(18)	Soybean oil	BO1	y
(19)	Tin	LMSNDY	y
(20)	Wheat	WEATHO1H	
(21)	WTI Crude	USCRWTIC	y
(22)	Zinc	LMZSDY	y

Notes: 1. The data are from Bloomberg. The 13 series with a "y" in the "longer sample" column include 282 monthly observations, 1989:6-2012:11. Other data include 197 monthly observations, 1996:7-2012:11.

Table 2
Basic commodity price statistics.

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
		Levels			Growth rates					
		ρ_1	ρ_2	ρ_3	Mean	s.d.	Median	ρ_1	ρ_2	ρ_3
(1)	Aluminum	0.94	0.88	0.82	−0.18	5.95	−0.59	0.08	0.03	0.09
(2)	Brent crude	0.98	0.95	0.92	0.43	10.58	1.54	0.06	−0.02	0.03
(3)	Coco	0.96	0.93	0.89	−0.04	8.81	−0.03	−0.14	0.01	−0.06
(4)	Coffee arabica	0.97	0.95	0.92	−0.06	9.74	−0.07	−0.12	0.09	0.02
(5)	Coffee robusta	0.98	0.96	0.94	−0.04	8.64	−0.78	−0.19	0.10	0.02
(6)	Copper	0.98	0.97	0.94	0.19	7.64	0.24	0.15	0.09	0.04
(7)	Corn	0.95	0.91	0.86	0.08	8.28	−0.01	−0.06	0.06	0.06
(8)	Cotton	0.97	0.92	0.86	−0.18	6.77	−0.20	0.25	0.12	0.07
(9)	Gasoline	0.96	0.92	0.90	0.31	13.04	0.89	0.00	−0.22	−0.04
(10)	Gold	0.98	0.97	0.96	0.32	4.52	0.18	−0.13	−0.07	0.10
(11)	Heating oil	0.97	0.94	0.92	0.42	11.51	0.27	−0.08	−0.04	−0.02
(12)	Lead	0.98	0.96	0.94	0.21	8.39	0.27	0.07	0.05	0.07
(13)	Natural gas	0.90	0.83	0.79	0.06	20.34	1.48	−0.19	−0.13	0.12
(14)	Nickel	0.98	0.95	0.92	−0.10	10.06	−0.41	0.09	0.01	0.05
(15)	Palm oil	0.97	0.93	0.89	0.18	9.80	−0.17	0.13	−0.03	0.00
(16)	Rubber	0.99	0.97	0.95	0.20	8.63	0.16	0.24	0.02	0.06
(17)	Soybean meal	0.93	0.87	0.81	0.03	8.25	0.12	−0.05	0.06	−0.09
(18)	Soybean oil	0.97	0.94	0.90	0.09	7.30	0.26	−0.09	0.11	−0.08
(19)	Tin	0.98	0.96	0.94	0.05	6.66	−0.30	0.10	0.11	0.20
(20)	Wheat	0.95	0.90	0.86	0.09	7.91	−0.22	−0.03	−0.01	0.08
(21)	WTI crude	0.98	0.95	0.93	0.30	9.37	0.73	0.14	−0.03	0.04
(22)	Zinc	0.97	0.94	0.91	−0.17	7.91	−0.36	−0.01	0.09	0.08

Notes: 1. All data are monthly. For the commodities with "y" in the "longer sample?" column in Table 1 (i.e., rows 1, 2, 6, 9-11, 14, 17-19, 21 and 22), data run over 282 (281, for growth rates) observations from 1989:6 to 2012:11. In the remaining nine rows, data run over 197 (196, for growth rates) observations from 1996:7-2012:11.
2. The data are constructed from Bloomberg daily data by sampling the last day of the month and deflating by the previous month's US CPI. Growth rates are 100 × log differences.
3. The symbols ρ_1, ρ_2 and ρ_3 denote the first, second and third autocorrelations; s.d. refers to standard deviation.

longer sample are available. As explained above, our modeling strategy is robust to the presence or absence of unit roots. We do note that the first order autocorrelations of the levels presented in the "ρ_1" column are all at or above 0.9 and all but three are at or above 0.95. This is suggestive of near unit root or unit root behavior. In columns (4)–(5), the values for mean and s.d. show that the differenced series have little drift: for all series, the absolute value of the point estimate of the mean is far below $(s.d./\sqrt{T})$, where $T = 281$ for the 13 prices whose start data is 1989:6, $T = 196$ for the remaining 9 prices. Hence the growth rate of the nominal commodity prices has tracked that of overall CPI. An approximate $1/\sqrt{T} \approx 0.06$ or 0.07 standard error (0.06 or 0.07 depending on whether $T = 281$ or $T = 195$) for the first order autocorrelation ρ_1 suggests first order serial correlation significant at the five percent level for a half-dozen of the series. But the point estimates are small, maxing out at 0.25 (for cotton). Hence from the point of view of first and second moments, the univariate series by and large look like driftless random walks.

As stated above, for analysis of the h month change in commodity prices, $h = 1$, 3 or 12, we deflate by CPI lagged h months. Thus real price p_{it} is different in regressions for different h because a different month's CPI was used to deflate. This means that factors and $\widehat{F}_{it}$ also vary with h for a given t. To keep notation relatively uncluttered, we do not index p_{it} or $\widehat{F}_{it}$ by h. In practice, however, the time series properties of p_{it} and of $\widehat{F}_{it}$ are very similar across horizons and so we only reported basic statistics for $h = 1$ in Table 2 and only report factors and factor loadings for $h = 1$ in tables and figures below.

The regression coefficients estimated in (2.2)–(2.4) ($\widehat{\alpha}_i$, $\widehat{\beta}$ and $\widehat{\gamma}$) vary with h (mainly because the left hand side variable varies, to a minor extent because of the variation in p_{it} and $\widehat{F}_{it}$ just noted).

4. Factor model results

4.1. Full sample estimates: factor loadings and factors

We begin with full sample results. Table 3 presents factor loadings for each of our two commodity price panels, for $h = 1$. All commodities load positively on the first factor. In the 22 commodity panel,

Table 3
Factor loadings.

		A. 1996:7-2012:11, $n = 22$		B. 1989:6-2012:11, $n = 13$	
		$\widehat{\delta}_{1i}$	$\widehat{\delta}_{2i}$	$\widehat{\delta}_{1i}$	$\widehat{\delta}_{2i}$
(1)	Aluminum	0.07	0.05	0.07	0.15
(2)	Brent Crude	0.31	0.21	0.37	−0.35
(3)	Coco	0.10	−0.23		
(4)	Coffee arabica	0.12	−0.36		
(5)	Coffee robusta	0.13	−0.35		
(6)	Copper	0.32	0.02	0.33	0.26
(7)	Corn	0.14	−0.18		
(8)	Cotton	0.06	−0.22		
(9)	Gasoline	0.27	0.22	0.33	−0.32
(10)	Gold	0.27	−0.13	0.27	0.23
(11)	Heating Oil	0.29	0.22	0.35	−0.32
(12)	Natural Gas	0.06	0.46		
(13)	Lead	0.30	−0.04	0.33	0.26
(14)	Nickel	0.25	0.23	0.30	0.03
(15)	Palm oil	0.16	−0.23		
(16)	Rubber	0.31	−0.07		
(17)	Soybean meal	0.10	−0.15	0.08	0.24
(18)	Soybean oil	0.15	−0.18	0.13	0.35
(19)	Tin	0.27	−0.13	0.28	0.29
(20)	Wheat	0.13	−0.09		
(21)	WTI Crude	0.28	0.22	0.34	−0.32
(22)	Zinc	0.16	0.02	0.17	0.30
% of variance explained:		71	14	80	9

Notes: 1. Define the real commodity price $p_{it} = 100 \times \ln(P_{it}/P_{CPI,t-1})$ where P_{it} is the level of commodity price i. The fitted model is $p_{it} = \text{const.} + \widehat{\delta}_{1i}\widehat{f}_{1t} + \widehat{\delta}_{2i}\widehat{f}_{2t} + \widehat{v}_{it}$, where $\widehat{f}_{1t}$ and $\widehat{f}_{2t}$ are the first and second principal components of the real commodity prices. In a one factor model whose results are briefly summarized in Table 9 below, $\widehat{F}_{it} = \widehat{\delta}_{1i}\widehat{f}_{1t}$; in a 2 factor model whose results are reported in detail in the following tables, $\widehat{F}_{it} = \widehat{\delta}_{1i}\widehat{f}_{1t} + \widehat{\delta}_{2i}\widehat{f}_{2t}$.

factor loadings are particularly large for oil products (Brent, gasoline, heating oil and WTI), certain industrial commodities (copper, lead, rubber, tin) and gold. Loadings on the second factor are negative for all agricultural products (coco, both coffees, corn, cotton, palm oil, both soybeans and wheat) as well as a few other commodities (gold, lead, tin, rubber). In the 13 commodity panel, there are only two agricultural commodities (soybean meal and soybean oil). Here, oil products load negatively on the second factor. Note that differences across the two panels result not only from breadth of commodity coverage, but also in sample. The bottom row of the table reports that the first factor explains the bulk of the variance of the series, the second factor about 10%. We took the 10% figure, and the fact that there was a systematic pattern to factor loadings, to indicate that a two factor model is a reasonable starting point. All our detailed results are from two factor models.

Fig. 1A and B plot the two factors (PC1 and PC2) for each of our panels. By construction, the factors have mean zero, and, within each sample, PC1 and PC2 have zero correlation. The first principal component PC1 looks very similar across panels, for example tracking the commodity boom/collapse/boom of 2002–2012. Indeed, over the 1996:7-2012:11 period that the two samples overlap, the correlation of the growth rates of the two PC1's is 0.97. PC2 has a pattern that is sample specific. The correlation of the growth rates of the two PC2's is −0.69. The negative sign of the correlation reflects a normalization (namely, $\delta_{2i} > 0$ for i = aluminum) and hence has no economic substance. That the

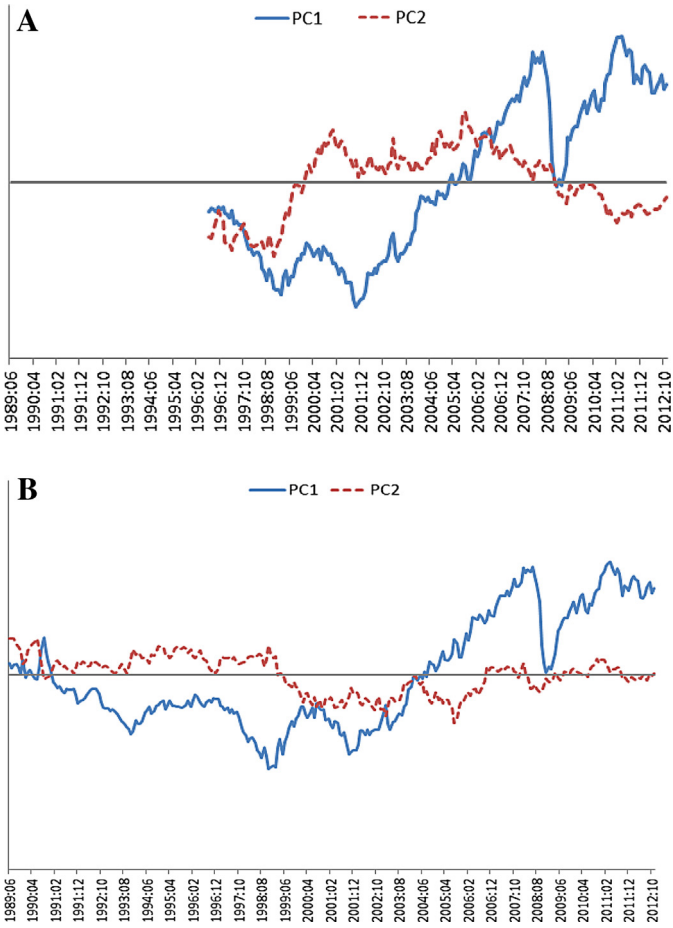

Fig. 1. A) Principal components, $n = 22$ commodity panel. B) Principal components, $n = 13$ commodity panel.

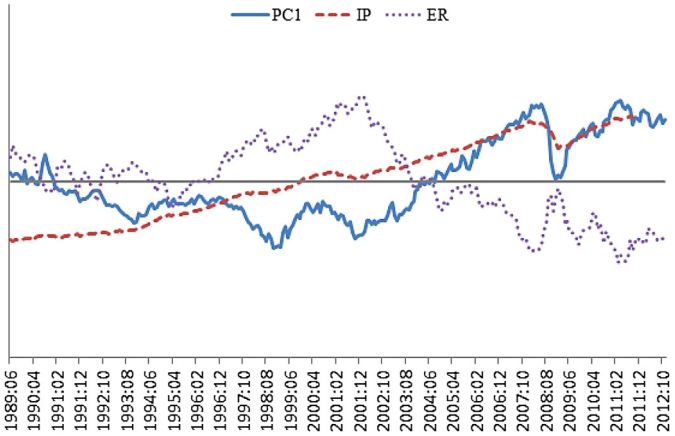

Fig. 2. First principal component, industrial production and exchange rate.

correlation is smaller in absolute value than is the correlation for the first principal component does have substance and reveals some sensitivity to sample.

Other factor model studies of commodity prices have noted a correlation between factors and aggregate data (Byrne et al. (2011), Chen et al. (2012)). Fig. 2 plots the factor along with the logarithm of our industrial production (IP) and exchange rate (ER) data. The data have been standardized to have zero mean and identical standard deviations. In the last 10 years of the sample, one can see that PC1 is positively related to IP and negatively related to ER. The positive sign for co-movement with IP and the negative sign for co-movement with ER are sensible. The fact that the loadings on PC1 are positive ($\delta_{1i}>0$ in Table 3) suggests that variables whose movements are positively related to commodity price movements will also be positively related to PC1. Higher IP of course is associated with higher commodity prices and hence the positive correlation with PC1 is sensible. As well, the fact that our commodity prices are denominated in US dollars means that ceteris paribus a weaker dollar will mean higher commodity prices. Hence the negative correlation between PC1 and ER is sensible. These patterns are consistent with those depicted for a principal component and US GDP in Byrne et al. (2011) and for a principal component and trade weighted US dollar in Chen et al. (2012).

To get a rough gauge of the magnitude of the relationship between the series, we compute the correlation between the growth rates of the series (with samples for correlations with IP ending in 2011:9 rather than 2012:11):

	A. 1996:7-2012:11, n=22		B. 1989:6-2012:11, n=13		
	IP	ER	IP	ER	
PC1	0.27	-0.13	0.27	-0.14	(4.1)
PC2	0.12	-0.09	-0.03	0.08	

The signs of the correlations with PC1 are as expected. For PC2, the fact that signs of the correlation are opposite for IP and ER again is consistent with the logic just noted. We note that the magnitude of the correlations is quite small, raising the possibility that the factors have information useful for understanding commodity movements that is not in either aggregate measure or conversely (or both).

4.2. Full sample results, regression estimates

Table 4 presents full sample estimates of $p_{\mathrm{it}+h} - p_{\mathrm{it}} = \alpha_i + \beta(\widehat{F}_{\mathrm{it}} - p_{\mathrm{it}}) + \gamma x_t + u_{\mathrm{it}+h}$, where x_t is an indicator of aggregate activity. This and subsequent tables are structured so that the left half of the table presents results for the 22 commodity panel using data from 1996:7 to 2012:11, while the right

Table 4
Full Sample Estimates of β and γ

	A. 1996:7-2012:11, $n = 22$						B. 1989:6-2012:11, $n = 13$					
	— $h = 1$ —		— $h = 3$ —		— $h = 12$ —		— $h = 1$ —		— $h = 3$ —		— $h = 12$ —	
	$\widehat{\beta}$	$\widehat{\gamma}$	$\widehat{\beta}$	$\widehat{\gamma}$	$\widehat{\beta}$	$\widehat{\gamma}$	$\widehat{\beta}$	$\widehat{\gamma}$	$\widehat{\beta}$	$\widehat{\gamma}$	$\widehat{\beta}$	$\widehat{\gamma}$
(1)F	0.10		0.25		0.86		0.08		0.22		0.78	
	(0.01)		(0.03)		(0.08)		(0.01)		(0.03)		(0.10)	
(2)IP		1.82		4.12		0.55		1.88		3.85		1.2
		(0.27)		(0.53)		(1.65)		(0.25)		(0.53)		(1.48)
(3)F, IP	0.11	1.85	0.27	3.77	0.88	0.91	0.09	1.80	0.25	4.08	0.80	0.36
	(0.01)	(0.26)	(0.03)	(0.57)	(0.09)	(1.26)	(0.02)	(0.29)	(0.04)	(0.62)	(0.13)	(1.47)
(4)ER		−0.13		−0.62		−0.57		−0.19		−0.54		−0.36
		(0.08)		(0.16)		(0.25)		(0.09)		(0.17)		(0.26)
(5)F, ER	0.10	−0.14	0.25	−0.63	0.86	−0.57	0.08	−0.19	0.23	−0.54	0.74	−0.35
	(0.01)	(0.09)	(0.03)	(0.16)	(0.08)	(0.19)	(0.01)	(0.09)	(0.04)	(0.17)	(0.12)	(0.22)

Notes: 1. Data are monthly; n is the number of commodities in the panel.
2. Let p_{it} be the log real price of commodity i when it is deflated by US CPI lagged h months for $h = 1$, 3 or 12, $p_{it} = 100 \times \ln(P_{it}/P_{CPI,t-h})$. Let $\widehat{F}_{it}$ be the fit from a two factor model estimated by principal components to the panel of real commodity prices. We do not index p_{it} or F_{it} by "h" to keep notation relatively uncluttered.
3. The table presents point estimates and standard errors from the fixed effects panel regression $p_{it+h} - p_{it} = \alpha_i + \beta(\widehat{F}_{it} - p_{it}) + \gamma x_t + u_{it+h}$. Lines with $\widehat{F}_{it} - p_{it}$ included are labeled "F". The regressor x_t has two definitions: the monthly percentage change in: industrial production in the OECD and six emerging markets (lines 2 and 3, denoted "IP") or the trade weighted US dollar (lines 4 and 5, denoted "ER" as in "exchange rate; " an increase means appreciation of the dollar). Estimates of the fixed effects are not reported to save space.
4. For industrial production (lines 2 and 3) both samples end in 2011:9 rather than 2012:11.

half presents results for the 13 commodity panel using data from 1989:6 to 2012:11. In each half of the table, we present results for $h = 1$, 3 and 12 month changes in real commodity prices.

Line (1) in Table 4 presents estimates of (2.2), the regression in which the only stochastic regressor is that of the factor model. For $h = 1$, the estimates of $\widehat{\beta}$ are 0.10 (22 commodity panel) and 0.08 (13 commodity panel). That is, if $\widehat{F}_{it}$ is (say) one percent below demeaned p_{it}, then, next month's Δp_{it+1} falls, on average, by 0.10 or 0.08 percent. The estimates for $h = 3$ and $h = 12$ tell an equivalent story about slow reversion towards the factor over 3 and 12 month horizons, with estimates that are about 3 or 12 times as large as that for $h = 1$. At 12 months, about 80% of the gap between F and p is expected to be closed. The estimates are highly statistically significant, though the standard errors should be taken with a grain of salt because $\widehat{F}_{it}$ is a generated regressor.

Lines (2) and (3) introduce monthly percentage changes in global industrial production (IP). The positive sign of the estimates of $\widehat{\gamma}$, the coefficient on IP, is sensible. The estimates indicate that a 1% increase in global industrial production is expected to lead to a 1.8%–1.9% increase in commodity prices over $h = 1$ month, 3.9%–4.1% over $h = 3$ months, and 0.6–1.3% over $h = 12$ months. The values at the $h = 12$ horizon are not significant at conventional levels. Since, as well, the point estimates for $h = 12$ are smaller than those for $h = 1$ and $h = 3$, one possible interpretation is that mean reversion in IP is such that effects on commodity prices have disappeared within a year. This may be consistent with Akram (2009, p843), whose pictures of VAR impulse responses seem to suggest approximately equal impact and four quarter ahead responses of the level commodity prices to an industrial production shock. This would imply a one year ahead response of the change that is near zero.

Lines (4) and (5) introduce monthly percentage change in the trade weighted US dollar (ER). The negative sign of the estimates of $\widehat{\gamma}$, the coefficient on ER, is also sensible: given that most of our commodity prices are denominated in dollars, an appreciation of the dollar (an increase in ER), will, ceteris paribus, dampen demand (though of course the ceteris paribus assumption is questionable when one thinks about equilibrium in commodity and exchange rate markets). The estimates for $h = 12$ are relatively small both numerically and statistically, with possible interpretation parallel to that just given for IP.[8]

[8] This is, however, inconsistent with the cited pictures in Akram (2009, p843), in which the commodity price response to a (real) exchange rate shock seems to us to be larger 4 quarters out than on impact.

Comparison of lines (1) and (2) to line (3), and of lines (1) and (4) to line (5), indicate that estimates of both β and γ are quite insensitive to inclusion of the other variable. This suggests that there is imperfect overlap of the information captured by the factor on the one hand and by IP or ER on the other.

4.3. Partial sample results

Recall that we begin partial sample estimation 72 months into each of our two samples, adding one month at a time to the sample used to estimate (2.2) and generate predictions. The number of estimates of β and predictions is given in the notes to Table 5. In this subsection, we discuss patterns produced by the sequences of estimates of β and of forecasts. The next section compares factor model forecasts to forecasts from alternative models.

Begin with the estimates of β. As noted above, our presumption that commodity prices revert towards F implies that in (2.2) we expect the estimates of β to be positive. Indeed every single one of our estimates of β was positive, in each of the three horizons and two panels. Median values were as follows:

A. 1996:7-2012:11, n=22			B. 1989:6-2012:11, n=13			
h=1	h=3	h=12	h=1	h=3	h=12	(4.2)
0.14	0.32	0.98	0.15	0.36	0.79	

The values in (4.2) underscore two points revealed in Table 4: First, p_{it} reverts towards F_{it}. Second, such reversion is slow: the median values indicate that over one month, about 14%–15% of the gap between p_{it} towards F_{it} is closed.

Table 5
Correlation of prediction and realization.

		A. 1996:7-2012:11, $n = 22$			B. 1989:6-2012:11, $n = 13$		
		$h = 1$	$h = 3$	$h = 12$	$h = 1$	$h = 3$	$h = 12$
(1)	Aluminum	**0.10**	**0.17**	**0.26**	**0.12**	**0.22**	**0.29**
(2)	Brent Crude	**0.13**	**0.23**	**0.49**	**0.09**	**0.05**	**0.08**
(3)	Coco	**0.11**	**0.20**	**0.37**			
(4)	Coffee arabica	**0.05**	**0.09**	**0.26**			
(5)	Coffee robusta	**0.12**	**0.17**	**0.06**			
(6)	Copper	−0.10	−0.15	−0.42	−0.01	**0.04**	−0.19
(7)	Corn	**0.18**	**0.27**	**0.59**			
(8)	Cotton	**0.01**	**0.07**	**0.47**			
(9)	Gasoline	**0.29**	**0.43**	**0.36**	**0.31**	**0.40**	**0.21**
(10)	Gold	**0.04**	**0.01**	−0.05	−0.08	−0.20	−0.42
(11)	Heating Oil	**0.19**	**0.26**	**0.57**	**0.23**	**0.28**	**0.30**
(12)	Natural Gas	**0.21**	**0.28**	**0.37**	**0.11**	**0.16**	**0.41**
(13)	Lead	−0.02	−0.09	**0.08**	−0.08	−0.12	−0.01
(14)	Nickel	−0.11	−0.19	**0.00**			
(15)	Palm oil	**0.02**	**0.09**	**0.32**			
(16)	Rubber	**0.08**	**0.06**	−0.07			
(17)	Soybean meal	**0.12**	**0.19**	**0.35**	**0.14**	**0.24**	**0.45**
(18)	Soybean oil	**0.14**	**0.22**	**0.52**	**0.11**	**0.17**	**0.28**
(19)	Tin	**0.10**	**0.13**	**0.44**	**0.20**	**0.31**	**0.45**
(20)	Wheat	**0.10**	**0.22**	**0.65**			
(21)	WTI Crude	**0.14**	**0.27**	**0.53**	**0.06**	**0.06**	**0.15**
(22)	Zinc	−0.17	−0.29	−0.19	−0.06	−0.10	**0.02**
Median		**0.10**	**0.17**	**0.36**	**0.11**	**0.16**	**0.21**
# > 0.0		18	18	17	9	10	10

Notes: 1. The table presents the correlation between out of sample prediction and realization for the factor model whose whole-sample estimates are presented in line 1 of the previous table. See notes to the previous table. Positive values are in bold.
2. The base month for the first forecast is 72 months into the sample, which is 1995:5 for the 13 commodity sample (column B) and 2002:6 for the 22 commodity sample (column A). In both samples, the last forecast was for data realized in 2012:11. For the 13 commodity, the first 1, 3 and 12 month ahead predictions were for t+1 = 1995:6, t+3 = 1995:8 and t+12 = 1996:5, resulting in 210, 208 and 199 predictions. For the 22 commodity sample, the comparable dates and number of predictions are 2002:7, 2002:8 and 2003:6; 125, 123 and 114.

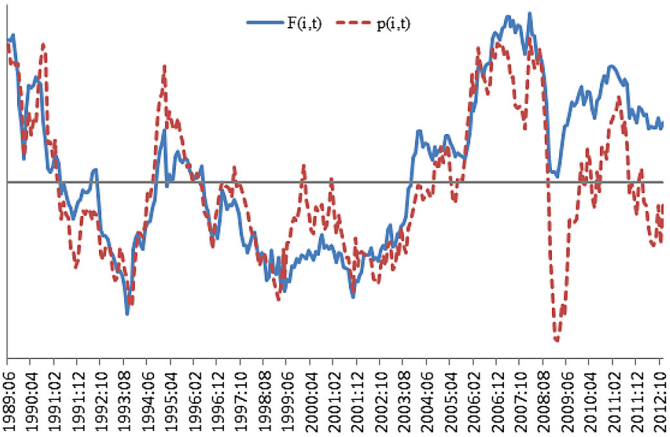

Fig. 3. Demeaned p_{it} and F_{it}, i = aluminum, $n = 13$ commodity panel.

Did the out-of-sample value of p_{it+h} actually move towards F_{it}? Table 5 presents the correlation between prediction and realization. The answer to the question just posed is yes: the correlation is positive in roughly 80% of the entries, with exceptions in metals (copper, gold, lead, nickel and zinc). Consistent with the estimates of β just discussed, the typical value of the correlation is small, around 0.1 to 0.2.

While the evidence in Table 5 is supportive of reversion towards the factor, the correlations are not particularly large: commodity price data are noisy. To illustrate this noisiness, along with a broad tendency of reversion towards F_{it}, Fig. 3 presents a plot for i = aluminum, $n = 13$ commodity sample.[9] This commodity and sample have been chosen because most measures the factor model's performance for aluminum was roughly middle of the pack when compared to other commodities. What Fig. 3 plots is a pasted together sequence of estimates of $\widehat{F}_{it} \equiv \widehat{\delta}_{1i}\widehat{f}_{1t} + \widehat{\delta}_{2i}\widehat{f}_{2t}$ along with price p_{it} that has been demeaned so that whether we predict a commodity price to rise or to fall essentially depends on whether $\widehat{F}_{it}$ is above p_{it}. Details are in a footnote.[10] If one computes the difference between p_{it} and $\widehat{F}_{it}$, the units are percentage. For example, at the first observation, $p = 40$, $\widehat{F} = 38$, so $\widehat{F}$ is 2 percent below p.

Let us use Fig. 3 to illustrate our forecasting presumption, which was described above as "when (demeaned) p_{it} is above F_{it}, p_{it} subsequently tends to fall; when p_{it} is below F_{it}, p_{it} subsequently tends to rise." We will point to periods when our presumption was broadly supported, and others where it was not. Supported: from mid-2000 to mid-2001, p_{it} was consistently above $\widehat{F}_{it}$ and, per our presumption, p_{it} generally fell during that period (with some upward wiggles, which contradict our presumption); from mid-2002 to mid-2004, p_{it} was consistently below $\widehat{F}_{it}$ and, per our presumption, p_{it} generally rose during that period (with some downward wiggles, which contradict our presumption). Not supported: In the last two years of our sample, p_{it} is below $\widehat{F}_{it}$, but p_{it} nonetheless shows a downward trend. And a counterexample that looms large in performance of our model in terms of RMSPE: From 2006 forward, p_{it} was below $\widehat{F}_{it}$, implying, we posit, that p_{it} will rise. Consider $h = 12$ month ahead predictions, from a

[9] The figure is for $h = 1$, i.e., p_{it} has been constructed by deflating nominal price by the previous month's CPI. But since the picture hardly changes for deflation for other h, we will use the figure for illustration for all h.

[10] Construction of $\widehat{F}_{it}$: the first (1989:6-1995:5) 72 observations are the 72 fitted values of $\widehat{\delta}_{1i}\widehat{f}_{1t} + \widehat{\delta}_{2i}\widehat{f}_{2t}$ from the first sample used to estimate the factor model. Each successive observation is the value used to predict, i.e., the last fitted value from the sample that ends on that date; for example, the observation at 1995:6 is the 1995:6 fitted value for estimation using 1989:6-1995:6 data. Construction of demeaned p_{it}: the first 72 observations adjust p_{it} by subtracting the 1989:6-1995:5 mean of p_{it}. Each successive observation subtracts the mean up to that date. For example, for 1995:6, we subtract from $p_{i,1995:6}$ the 1989:6-1995:6 mean of p_{it}. The algebra of fixed effects is such that whether we predict a commodity price to rise or to fall essentially depends on whether $\widehat{F}_{it}$ is above p_{it} when p_{it} is demeaned as just described. The reason for the qualifier "essentially" is that the sample used for fixed effects estimation is slightly different from that used for factor estimation and does not include the observation used for prediction. (See (2.5)–(2.7).).

base of mid-2007 to mid-2008. Not only did p_{it} fall 12 months later, but it fell dramatically. We will return to this period and forecast horizon below.

One indicator of the strength of reversion towards the factor is persistence of deviations of p_{it} from F_{it}. One can see in Fig. 3 that prices stay on one side of $\widehat{F}_{it}$ for periods as long as several years. Thus, reversion towards the factor is slow, as was indicated by the estimated values of β. On the other hand, one can that see that p_{it} crosses $\widehat{F}_{it}$ more often than the zero line.

We summarize results so far as: the factors correlate in sensible ways with observable macro data. Reversion of prices towards the factors is robust, though slow. For quantitative evidence in addition to that already presented in the tables, the next section turns to comparisons with other models.

5. Comparisons to other models

We compare our factor model forecasts first to those of a "no change" model, and then to those of models that rely on an aggregate predictor as in (2.4).

5.1. Comparison to "no-change" model

Table 6 presents a forecast comparison to a no change forecast, i.e., a model that posits that the log real commodity price follows a random walk. The metric of comparison the ratio of root mean squared prediction errors (RMSPE). In the "median U" line we see values ranging from 1.00 to 1.02, indicating that for the median commodity, the RMSPE factor model was about 0–2 percent larger than that of the no change model.[11] While the median is about the same across horizons, the extremes of the distribution spread out as the prediction horizon h increases. In the 22 commodity panel, for example, the minimum and maximum U-statistics for $h = 1$ are 0.97 (gasoline) and 1.13 (gold). The comparable figures for $h = 12$ are 0.78 (wheat) and 1.74 (gold). In the 22 commodity panel, the entries in the "$\#U < 1$" line reveal a tendency for improvement as the horizon increases: 11 of the U-statistics are below 1 for $h = 12$, 6 for $h = 1$. No such tendency is apparent in the 13 commodity panel. Overall, in each panel about 40 percent of the U-statistics are below 1. When a U-statistic was below 1, it usually but not always was significant at the 10 percent level (i.e., t-statistic > 1.282). However, the p-value for the maximum t-statistic was never close to significant in the 22 commodity panel, though it is significant at the 10 percent level for both $h = 1$ and $h = 3$ in the 13 commodity panel.

In terms of specific commodities: U-statistics strictly below one are in bold. Because of rounding to two digits, some values of 1.00 are in bold (e.g., Brent, $h = 1$, panel A, U-statistic to 4 digits $= 0.9952$) while others are not (e.g., coffee robustica, $h = 1$, panel A, U-statistic to 4 digits $= 1.0046$). Commodities that appear in both panels perform comparably in both panels. Factor model forecasts of energy products worked well in both panels. By and large, and consistent with the correlation values in Table 5, factor model forecasts of metals did not do well. Factor model forecasts of agricultural products fell somewhere between energy and metals. Factor model gold forecasts performed so poorly as to fall into a class by themselves, with U-statistics for $h = 12$ indicating a percentage increase in RMSPE that is three times a large as that of the next worst performing factor model forecast (zinc).

As an alternative to RMSPE as a measure of predictive performance, we evaluated the directional accuracy of the factor model forecast. For each forecast, we checked whether the sign of the forecast of the change in real commodity price matched the sign of the realization. We report the fraction of signs that were correct. Confidence intervals are not available. In this measure of predictive accuracy, values lie between 0 and 1. Larger values are better from the point of view of the factor model: a value of "1.0" is ideal and means every sign was correct. If commodity price changes are symmetric around a zero mean, then a coin flip would yield a value of 0.5 by this measure. Hence the factor model beats a no change model with symmetric innovations if the fraction is above 0.5.

One can see in Table 7 that this happens about two-thirds of the time. Unsurprisingly, entries in Table 6 with U-statistics below 1 also yield fractions greater than 0.5 in Table 7, with occasional exceptions (Brent oil, $h = 1$, both panels, and WTI, $h = 12$, panel B). Agricultural commodities tend to have

[11] The median values are strictly above 1.00; the value of 1.00 for $h = 3$, panel B is rounded down from 1.004.

Table 6
Comparison to RMSPE of "No Change" forecast.

		A. 1996:7-2012:11, $n = 22$						B. 1989:6-2012:11, $n = 13$					
		$h = 1$		$h = 3$		$h = 12$		$h = 1$		$h = 3$		$h = 12$	
		U	t	U	t	U	t	U	t	U	t	U	t
(1)	Aluminum	1.02	0.65	1.01	0.54	1.03	0.37	1.01	1.36	1.00	1.21	1.02	0.59
(2)	Brent Crude	**1.00**	0.89	**0.98**	0.93	**0.90**	1.97	**1.00**	0.93	1.00	0.13	1.01	0.05
(3)	Coco	1.03	1.40	1.06	1.15	1.22	1.00						
(4)	Coffee arabica	1.03	0.97	1.03	0.89	**0.98**	2.30						
(5)	Coffee robusta	1.00	1.57	1.01	1.27	1.19	0.35						
(6)	Copper	1.04	−0.68	1.05	−0.33	1.18	−0.64	1.02	0.00	1.02	0.36	1.10	−0.58
(7)	Corn	**1.00**	1.42	**1.00**	1.02	**0.90**	1.81						
(8)	Cotton	1.05	0.10	1.04	0.27	**0.89**	0.70						
(9)	Gasoline	**0.97**	2.62	**0.95**	2.96	**0.91**	2.60	**0.98**	3.11	**0.95**	3.15	**0.98**	1.38
(10)	Gold	1.13	−2.06	1.36	−2.09	1.74	−1.32	1.14	−2.00	1.43	−2.03	1.74	−1.63
(11)	Heating Oil	**0.98**	2.04	**0.97**	1.52	**0.85**	2.12	**0.98**	2.61	**0.97**	2.58	**0.96**	1.19
(12)	Natural Gas	**0.99**	1.88	**0.98**	2.30	1.08	1.75						
(13)	Lead	1.04	−0.39	1.08	−1.08	1.07	0.16	1.01	1.08	1.03	1.31	**0.99**	0.67
(14)	Nickel	1.07	−0.76	1.12	−0.46	1.12	0.08	1.06	−0.87	1.12	−0.58	1.11	−0.01
(15)	Palm oil	1.08	0.06	1.07	0.59	1.12	0.58						
(16)	Rubber	1.02	0.87	1.04	0.32	1.18	−0.45						
(17)	Soybean meal	1.00	2.74	**0.99**	1.81	**1.00**	2.90	1.00	3.34	**0.99**	2.90	**0.92**	2.44
(18)	Soybean oil	1.00	1.17	**1.00**	0.90	**0.92**	1.29	1.01	1.19	1.02	0.92	1.01	0.80
(19)	Tin	1.02	0.57	1.02	0.29	**0.98**	0.93	**0.99**	1.53	**0.98**	0.88	**0.96**	0.85
(20)	Wheat	1.01	1.15	**0.99**	1.35	**0.78**	2.06						
(21)	WTI Crude	**0.99**	1.03	**0.97**	0.96	**0.88**	1.78	**1.00**	0.58	1.00	0.23	**1.00**	0.31
(22)	Zinc	1.13	−1.10	1.24	−0.82	1.28	−0.19	1.08	−0.61	1.16	−0.45	1.14	0.06
Median U		1.02		1.02		1.01		1.01		1.00		1.01	
$\#U < 1/\#t > 1.282$		6	7	9	5	11	10	5	5	4	4	6	2
p-value max-t			[0.46]		[0.35]		[0.20]		[0.08]		[0.10]		[0.17]

Notes to Table 6: 1. Let "RMSPE" denote "root mean squared prediction error," i.e., the standard deviation of the out of sample prediction error. The "U" columns present the U-statistic: RMSPE Model/RMSPE no change. When $U < 1$ the factor model had a smaller RMSPE than did a no change model. The number of commodities for which this was less than 1 is given in the $\#U < 1$ entries. U-statistics strictly less than 1 are in bold. Some entries with value 1.00 are in bold. This indicates that he U-statistic was between 0.995 and 1.0, and so was rounded to 1.00 when results were rounded to two digits.
2. The "t" columns present a t-test of H_0: $U = 1$ (equality of RMSPEs) against one-sided H_A: $U < 1$ (RMSPE factor model is smaller), using the Clark and West (2006) procedure. The number of commodities for which this test rejected equality at the 10 percent level is given in the $t > 1.282$ entry.
3. "p-value max-t" gives the fraction of 1000 bootstrap samples in which the maximum t-statistic exceeded the maximum of the 13 t-statistics in the sample for a given horizon. For example, the entry [0.46] for $h = 1$, panel A, reflects the fact that 461 of the 1000 bootstrap samples, the maximum of the 22 $h = 1$ Clark-West (2006) t-statistics for H_0: $U = 1$ exceeded 2.74; 2.74 is the critical value because it is the maximum of the 22 sample t-statistics for $h = 1$.
4. See notes to previous tables for sample periods and number of predictions.

entries in Table 7 less than 1 even when the associated U-statistics in Table 6 are above 1. So, too, do some metals, for some horizons (aluminum and copper). Finally, directional accuracy for the remaining metals is poor, consistent with the poor performance in Tables 5 and 6

There is a discrepancy between Tables 5 and 7 on the one hand (correlation of prediction and realization, and directional accuracy) and Table 6 on the other (RMSPE). Overall, Tables 5 and 7 suggest the factor model helps explain future changes in commodity prices while Table 6 does not. To better understand why these tables have different overall patterns, we again depict some results for aluminum, $n = 13$ commodity panel.

Fig. 4 presents scatterplots of actual vs. predicted change in commodity price for aluminum, $n = 13$ sample. Note that the scale on the three graphs is different, with predictions and realizations becoming increasingly spread out as one moves from $h = 1$ to $h = 12$ month ahead predictions and realizations.

If predictions were perfect, all points would lie on the 45° line that is drawn in each plot. Of course, in practice, realizations are more variable than predictions, so the dots do not lie on the 45° line but instead are closer to the vertical than horizontal axis. The positive correlation quantified in the aluminum row of panel B of Table 5 is evident, with increasingly positive correlation as one moves from

Table 7
Directional accuracy.

		A. 1996:7-2012:11, $n = 22$			B. 1989:6-2012:11, $n = 13$		
		$h = 1$	$h = 3$	$h = 12$	$h = 1$	$h = 3$	$h = 12$
(1)	Aluminum	0.46	**0.55**	**0.58**	**0.54**	**0.58**	**0.59**
(2)	Brent Crude	0.46	**0.52**	**0.54**	0.49	0.50	0.48
(3)	Coco	**0.54**	**0.56**	**0.54**			
(4)	Coffee arabica	**0.52**	**0.57**	**0.69**			
(5)	Coffee robusta	0.48	**0.59**	**0.56**			
(6)	Copper	0.47	0.49	0.37	**0.55**	**0.53**	0.47
(7)	Corn	**0.54**	**0.57**	**0.69**			
(8)	Cotton	**0.50**	**0.53**	**0.61**			
(9)	Gasoline	**0.57**	**0.67**	**0.68**	**0.57**	**0.65**	**0.59**
(10)	Gold	0.46	0.41	0.30	0.40	0.33	0.27
(11)	Heating Oil	**0.54**	**0.54**	**0.71**	**0.61**	**0.63**	**0.65**
(12)	Natural Gas	**0.50**	**0.54**	**0.54**			
(13)	Lead	0.45	0.46	0.46	0.50	0.50	**0.51**
(14)	Nickel	**0.56**	0.49	**0.54**	**0.54**	0.49	**0.58**
(15)	Palm oil	**0.54**	**0.53**	**0.57**			
(16)	Rubber	**0.50**	**0.56**	0.46			
(17)	Soybean meal	**0.52**	**0.56**	**0.68**	0.50	**0.53**	**0.63**
(18)	Soybean oil	0.49	**0.50**	**0.64**	**0.53**	**0.54**	0.49
(19)	Tin	0.46	0.43	**0.59**	**0.55**	**0.58**	**0.69**
(20)	Wheat	**0.57**	**0.54**	**0.72**			
(21)	WTI Crude	**0.51**	**0.55**	**0.56**	**0.53**	0.45	0.45
(22)	Zinc	0.49	0.46	0.50	0.48	0.50	**0.59**
Median fraction		**0.50**	**0.54**	**0.57**	**0.53**	**0.53**	**0.58**
#>0.5		13	16	17	8	7	8

Notes: 1. The table presents the fraction of predictions in which the sign of the factor model's prediction matched the sign of the realization. Fractions strictly greater than 0.5 are in bold.
2. See notes to previous tables for sample periods and number of observations.

$h = 1$ to $h = 3$ to $h = 12$ months. That directional accuracy figures are above 0.5 in the aluminum row of panel B of Table 7 means that in each of the three graphs, there are more dots in the northeast and southwest quadrants than in the northwest and southeast quadrants, though that fact might not be readily apparent at a glance. Finally, the fact that U-statistics are not particularly favorable is consistent with the entries in the southeast quadrant. These are entries in which the prediction error (=actual change – predicted change) is negative, and the predicted change is positive. This implies that for these entries, there is a negative correlation between the prediction error and prediction. In the $h = 12$ graphs, the outlier points towards the bottom are for predictions of mid-2007 through mid-2008 and realizations one year later. This negative correlation is sharp enough to produce a negative correlation for the whole sample, with values that happen to be -0.15 ($h = 1$), -0.13 ($h = 3$) and -0.23 ($h = 12$).

Aluminum, $n = 13$ commodity panel, is not the only series for which prediction and realization tend to be positively correlated and directional accuracy is good (Tables 5 and 7), but RMSPEs are larger than a "no change" model (i.e., $U > 1$ in Table 6). The no change model, of course, by construction produces zero correlation between prediction and realization. Some algebra can be used to show that a positive correlation for the factor model is guaranteed to lead to an RMSPE smaller than that of the no change model if there is zero correlation between factor model prediction and prediction error (=difference between realization and prediction). Evidently, this is not the case. Indeed, aluminum is representative in that the correlations between prediction and prediction error are negative and in absolute value are comparable to the values in Table 5. For the $n = 22$ commodity panel, the median correlations between prediction and prediction error were -0.17, -0.20 and -0.28 for horizons $h = 1$, 3 and 12; the values for the $n = 13$ commodity panel were -0.16, -0.20 and -0.24 (not reported in any table).

We interpret Fig. 4 and Tables 5–7 as underscoring two points. First, performance differences between the factor and no change forecasts are small by either an MSPE or directional accuracy criterion, with the no change forecast doing modestly better by an MSPE criterion and the factor model doing better by a directional accuracy criterion. Second, when Table 6 indicates that there is a relatively large

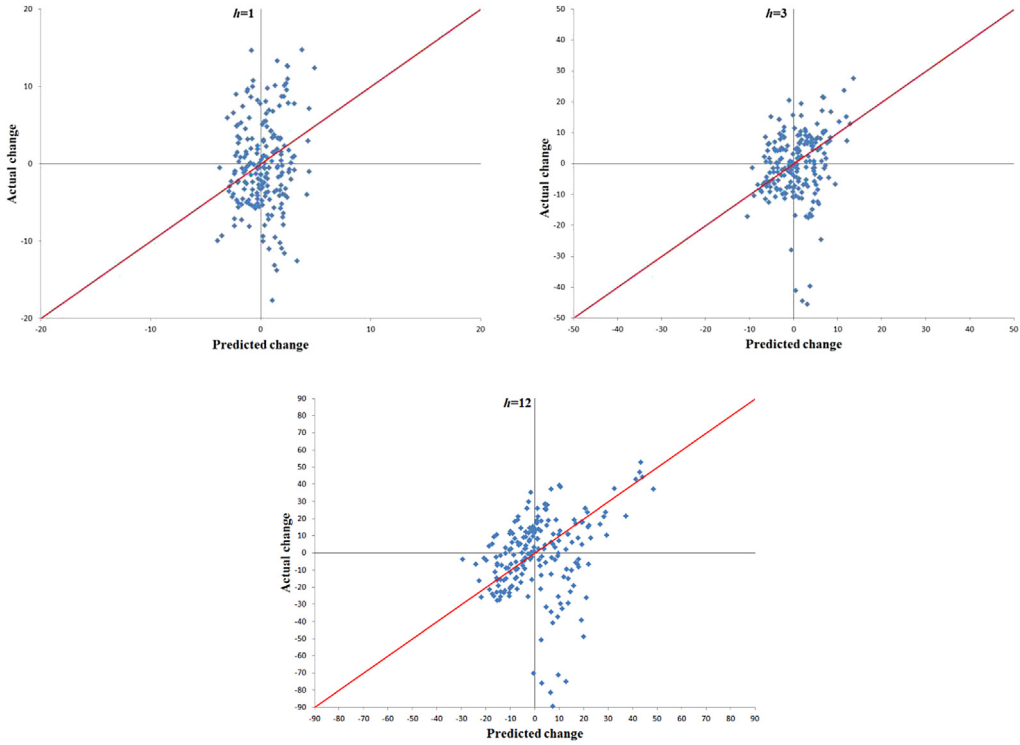

Fig. 4. Predicted and actual price change, aluminum, $n = 13$ commodity panel.

difference in MSPE performance that difference often but not always shows up when performance is measured by directional accuracy.

5.2. Comparison to other models

Let us now consider commodity price forecasts that rely on global industrial production IP or the trade weighted US dollar ER. As in (2.3) and (2.4), we estimate fixed effects models where IP or ER either joins or replaces our factor model. Here, and in the remainder of the paper, we present summaries of results. Commodity by commodity results, like the ones presented in Tables 5–7, are available in an unpublished appendix.

Table 8 presents results for the measures reported in Tables 5–7: correlation between prediction and realization (panel A), U-statistics relative to a no change forecast (panel B) and directional accuracy (panel C). The three panels have consistent implications. We start with a comparison of results relying on a single stochastic predictor, presented in lines 1, 2 and 4 in the table. We see in line (2) in each panel that IP performed better than the factor model (line (1)) at a $h = 1$ month horizon, worse at $h = 12$, about comparably at $h = 3$. We see in line (4) in each panel that ER performed worse at all three horizons. The results for IP are consistent with the full sample regression results in lines 2 and 3 of Table 4, in which the $h = 12$ coefficient on IP was small and statistically insignificant. The results for ER for $h = 1$ and $h = 3$ perhaps are a bit surprising, in that lines 4 and 5 of Table 4 seemed to indicate that ER has information not already impounded in the factor.

One can refine the comparison of the factor and IP predictions in terms of RMSPE by presenting U-statistics with the factor model RMSPE in the numerator and the IP RMSPE in the denominator. The results:

	1996:7-2012:11, n=22			1989:6-2012:11, n=13		
	h=1	h=3	h=12	h=1	h=3	h=12
Med. U (#<1)	1.02 (6)	0.99 (12)	0.95 (15)	1.01 (4)	1.01 (5)	0.97 (9)

(5.1)

Numbers less than one indicate that the factor model is preferred in terms of RMSPE. The values in (5.1) corroborate the F and IP lines (lines (1) and (2)) in the panel B of Table 8: IP performs better for $h = 1$. The factor model performs better for $h = 12$. The two are roughly equivalent for $h = 3$.

To return to Table 8: Line (5) in each panel indicates that combining ER with the factor delivers results that for the most part are not much different from the factor alone (line (1)). Line (3) in each panel indicates that, relative to either factor or IP alone, combining IP with factor strengthens correlation of prediction and realization (panel A), has ambiguous effects on directional accuracy (panel C), and seems to lead to RMSPE results that fall in between those of factor and IP (panel B).

Even though the regressions in Table 4 and the correlations between potential predictors presented in (4.1) above suggest that the factor and the macro variables contain information that only partially overlaps, our forecasting comparison finds that the "principle of parsimony" dominates the results: using two predictors generally does not improve on using just one. With hindsight of Table 8, then, it seems that for $h = 1$ month ahead predictions, the model with just IP is best; for $h = 12$ month ahead predictions, the model with just F is best; there is no clearly preferable set of predictors for $h = 3$.

All these models are static. The literature includes forecasts with dynamic models. Two examples: Alquist et al. (2013) provide a comprehensive study of models to forecast both real and nominal oil prices, using a set of ARMA and VAR models for real oil prices (results in their Tables 12–16) and forward rates and a number of other variables to forecast nominal oil prices (results in their Tables 4, 5 and 8). Chinn and Coibion (2012) consider predictions of nominal values of a number of commodities, using futures and a simple ARIMA model (results in their Table 2). In light

Table 8
Comparison of forecasts.

	1996:7-2012:11, $n = 22$			1989:6-2012:11, $n = 13$		
A. Prediction and realization: median correlation (# correlations > 0)						
Model	$h = 1$	$h = 3$	$h = 12$	$h = 1$	$h = 3$	$h = 12$
(1)F	0.10 (18)	0.17 (18)	0.36 (17)	0.11 (9)	0.16 (10)	0.21 (10)
(2)IP	0.11 (19)	0.06 (15)	−0.26 (4)	0.11 (11)	0.06 (12)	−0.15 (2)
(3)F, IP	0.14 (19)	0.21 (17)	0.35 (18)	0.16 (10)	0.18 (10)	0.19 (11)
(4)ER	−0.10 (1)	−0.12 (1)	−0.23 (1)	−0.03 (4)	−0.08 (3)	−0.18 (1)
(5)F,ER	0.09 (16)	0.17 (16)	0.35 (17)	0.10 (9)	0.17 (9)	0.20 (10)
B. Comparison to RMSPE of "no change" Forecast: median U (#$U < 1$)						
Model	$h = 1$	$h = 3$	$h = 12$	$h = 1$	$h = 3$	$h = 12$
(1)F	1.02 (6)	1.02 (11)	1.01 (10)	1.01 (5)	1.00 (4)	1.01 (6)
(2)IP	1.00 (13)	1.03 (4)	1.04 (5)	0.99 (10)	1.01 (5)	1.03 (0)
(3)F, IP	1.01 (6)	1.03 (6)	1.02 (9)	1.00 (6)	1.03 (6)	1.00 (6)
(4)ER	1.01 (2)	1.01 (1)	1.05 (5)	1.00 (1)	1.01 (1)	1.03 (0)
(5)F, ER	1.00 (6)	1.03 (6)	1.00 (6)	1.01 (4)	1.00 (5)	1.01 (6)
C. Directional accuracy: median fraction (# fractions > 0.5)						
Model	$h = 1$	$h = 3$	$h = 12$	$h = 1$	$h = 3$	$h = 12$
(1)F	0.50 (13)	0.54 (16)	0.57 (17)	0.53 (8)	0.53 (7)	0.58 (8)
(2)IP	0.55 (19)	0.56 (18)	0.46 (7)	0.54 (13)	0.56 (11)	0.49 (6)
(3)F, IP	0.55 (17)	0.58 (17)	0.60 (18)	0.53 (10)	0.53 (9)	0.58 (8)
(4)ER	0.51 (12)	0.46 (5)	0.42 (6)	0.50 (6)	0.49 (4)	0.50 (8)
(5)F, ER	0.50 (13)	0.53 (15)	0.57 (18)	0.53 (10)	0.53 (8)	0.56 (8)

Notes: 1. Panels A, B and C are analogous to Tables 5–7, respectively. For convenience of comparison, the "F" lines repeat information for the factor model presented in those tables. For example, the entry "0.10 (18)" in line (1), first $h = 1$ column in panel A, repeats the values of 0.10 and 18 found in the last two rows of the first $h = 1$ column in Table 5. See notes to Tables 5–7 for further discussion.

2. The "IP and "ER" lines present results in which the factor model term $\widehat{F}_{it} - p_{it}$ is replaced by percentage change in global industrial production or percentage change in the trade weighted US dollar, as in equation (2.4) in the text. The "F, IP" and "F, ER" report results when IP or ER joins $\widehat{F}_{it} - p_{it}$ in the model, as in equation (2.3) in the text.

Table 9
Summary of additional results.

	1996:7-2012:11, $n = 22$			1989:6-2012:11 or 1996:7-2012:11, $n = 13$		
A. Prediction and realization: median correlation (# correlations > 0)						
Model	$h = 1$	$h = 3$	$h = 12$	$h = 1$	$h = 3$	$h = 12$
(1)F	0.10 (18)	0.17 (18)	0.36 (17)	0.11 (9)	0.16 (10)	0.21 (10)
(2)1 factor	0.08 (17)	0.10 (17)	0.03 (12)	0.08 (11)	0.11 (11)	0.19 (11)
(3)Nominal	0.09 (17)	0.16 (18)	0.33 (17)	0.11 (10)	0.13 (10)	0.14 (9)
(4)96:7-12:11				0.05 (10)	0.09 (10)	0.30 (10)
B. Comparison to RMSPE of "no change" Forecast: median U (#$U < 1$)						
Model	$h = 1$	$h = 3$	$h = 12$	$h = 1$	$h = 3$	$h = 12$
(1)F	1.02 (6)	1.02 (11)	1.01 (10)	1.01 (5)	1.00 (4)	1.01 (6)
(2)1 factor	1.02 (4)	1.04 (4)	1.17 (3)	1.01 (3)	1.03 (3)	1.04 (4)
(3)Nominal	1.02 (6)	1.02 (8)	1.01 (11)	1.01 (5)	1.00 (4)	1.01 (6)
(4)96:7-12:11				1.01 (4)	1.01 (6)	1.01 (6)
C. Directional accuracy: median fraction (# fractions > 0.5)						
Model	$h = 1$	$h = 3$	$h = 12$	$h = 1$	$h = 3$	$h = 12$
(1)F	0.50 (13)	0.54 (16)	0.57 (17)	0.53 (8)	0.53 (7)	0.58 (8)
(2)1 factor	0.49 (8)	0.48 (6)	0.46 (4)	0.51 (9)	0.50 (7)	0.55 (12)
(3)Nominal	0.52 (13)	0.55 (17)	0.61 (18)	0.52 (9)	0.53 (7)	0.54 (11)
(4)96:7-12:11				0.53 (10)	0.53 (8)	0.60 (9)

Notes: 1. Panels A, B and C are analogous to Tables 5–7, respectively. For convenience of comparison, the "F" lines repeat information for the factor model presented in those tables. See notes to Tables 5–7.
2. The other three lines in each panel make a single change to the baseline specification. The "1 factor" lines uses one rather than two factors. The "nominal" lines report results when predictions are for nominal rather than deflated changes in commodity prices. The "96:7-12:11" lines report results when the $n = 13$ commodity panel is estimated on data running from 1996:7-2012:11 rather than 1989:6-2012:11.

of evidence below that our results for nominal forecasts are very similar to those for real forecasts, we compare the general tenor of results in Tables 5–7 with nominal and real results those in these two papers–"general tenor" because differences in samples and data mean we cannot make an exact comparison.

Overall, directional accuracy in our Table 7 and U-statistics in our Table 6 are in the ballpark of the numbers in the referenced tables in Alquist et al. (2013) [12] and Chinn and Coibion (2012). Our results for Brent and WTI are better than some of the entries in the Alquist et al. tables but not the best except for $h = 12$ when the measure is RMSPE and we look at the $n = 22$ commodity sample. Compared to Chinn and Coibion, we do better than their ARIMA models. Compared to futures, we do sometimes better (in particular, for $h = 12$ and using RMSPE) and sometimes worse. We conclude that our simple static model is competitive with dynamic models, especially at longer horizons.

6. Additional results

Table 9 presents three sets of additional results. All repeat the computations presented in Tables 5–7 (1)We rely on 1 rather than 2 factor models. (2)We forecast changes in nominal rather than deflated commodity prices. (3)We estimate the $n = 13$ commodity panel using data using 1996:7-2012:11 data. In Table 9, panel A reports results for correlation of prediction and realization, panel B reports U-statistics relative to a no change model, and panel C reports directional accuracy.

6.1. One factor model

The single factor in a one factor model is identical to the first factor in a two factor model. So factor loadings are those in the δ_{i1} columns in Table 3 and the factor is PC1 plotted in Figs. 1 and 2.

Line (2) in the three panels in Table 9 shows that for the most part, performance degrades by all three measures with the one factor model. The degradation relative to the two factor model

[12] We take the square root of the MSPE ratios in Alquist et al. (2013) before making the comparison.

summarized in line (1) is marked for the $n = 22$ commodity panel (left columns), mild for the $n = 13$ commodity panel. This is consistent with the fact that the many agricultural commodities in the $n = 22$ commodity panel loaded negatively on a second factor (Table 3). Omitting the second factor lessened the ability of these commodities to move in a different direction from the other commodities. We conclude that in a broad panel, we need at least two factors to well capture the co-movements of commodities.

6.2. Nominal changes

Line (3) in the three panels shows that results for nominal price changes are very similar to those for the deflated price changes presented in line (1) of the three panels. Inspection of predictions for individual commodities (not reported in the table) indicate that from the point of view of RMSPE, factor model forecasts of nominal changes in energy products perform well, those for agricultural products okay, those for metals perform poorly, a result also seen in Table 6.

6.3. Shorter sample

Comparing line (4) to the right half of line (1) shows little sensitivity to sample. Forecasts of energy products (not reported in the table) continue to work relatively well.

7. Conclusions

We explore using a static factor model to model a panel of commodity prices. We find that commodity prices tend to revert towards the factor. Such reversion is sufficiently strong and precise that by measures of correlation of prediction and realization, and directional accuracy, the factor model compared favorably to a “no change” model; by a mean squared error criterion, the factor model performed comparably. Compared to a model that uses growth in global industrial production to predict factor prices, we find that the factor model does better at longer (12 month) horizons, the industrial production model better at shorter (1 month) horizons, with the two about equivalent at 3 month horizons. Compared to a model that uses changes in the US dollar to predict factor prices, we find that the factor model does better at all horizons. The performance of the factor model is especially good for energy products, especially poor for metals, with agricultural products falling somewhere in between.

Tasks for future research include expanding the set of commodities, efficiently combining factor information with aggregate variables, using industry data, and tying the common factor to macroeconomic variables such as aggregate demand or supply shocks.

Acknowledgments

West thanks the National Science Foundation for financial support. We thank John Baffes, James Hamilton, Lutz Kilian, other conference participants and anonymous reviewers for helpful comments. We also thank Christiane Baumeister for supplying data and Yi Zhang for research assistance.

References

Akram, Q. Farooq, 2009. Commodity prices, interest rates and the dollar. Energy Econ. 31, 838–851.
Alquist, Ron, Kilian, Lutz, Vigfusson, Robert J., 2013. Forecasting the Price of Oil. forthcoming. In: Elliott, G., Timmermann, A. (Eds.), Handbook of Economic Forecasting, vol. 2A. Elsevier, Amsterdam.
Baffes, John, 2007. Oil spills on other commodities. Resour. Policy 32, 126–134.
Baffes, John, 2013. Commodity Market Outlook, January 2013. working paper. World Bank.
Baffes, John, Dennis, Allen, 2013. Long Term Drivers of Food Prices. working paper. World Bank.
Bai, Jushan, 2004. Estimating cross-section common stochastic trends in nonstationary panel data. J. Econom., 137–183.
Borensztein, Eduardo, Reinhart, Carmen M., 1994. The macroeconomic determinants of commodity prices. IMF Staff Pap. 41, 236–261.
Byrne, Joseph P., Fazio, Giorgio, Fiess, Norbert, 2011. Primary Commodity Prices: Co-movements, Common Factors and Fundamentals. World Bank Policy Research Working Paper 5578.

Chen, Shu-Ling, Jackson, John D., Kim, Hyeongwoo, Resiandini, Pramesti, 2012. What Drives Commodity Prices?. working paper, MPRA Paper No. 40711.
Chinn, Menzie D., Coibion, Olivier, 2012. The Predictive Content of Commodity Futures. manuscript. University of Wisconsin.
Ciccarelli, Matteo, Mojon, Benoit, 2010. Global inflation. Rev. Econ. Stat. 92 (3), 524–535.
Clark, Todd E., West, Kenneth D., 2006. Using out-of-sample mean squared prediction errors to test the martingale difference hypothesis. J. Econom. 135 (1–2), 155–186.
Cuddington John, T., Jerrett, Daniel, 2008. Super cycles in real metals prices? IMF Staff Pap. 55 (4), 541–565.
Dahl, Christian M., Iglesias, Emma M., 2009. Volatility spill-overs in commodity spot prices: new empirical results. Econ. Model. 26 (3), 601–607.
Engel, Charles, Mark, Nelson M., West, Kenneth D., 2012. Factor Model Forecasts of Exchange Rates. NBER Working Paper No. 18382.
Erten, Bilge, Ocampo, José Antonio, 2012. Super-cycles of Commodity Prices since the Mid-nineteenth Century. DESA Working Paper No. 110.
Gilbert, Christopher L., 1989. The impact of exchange rates and developing country debt on commodity prices. Econ. J. 99, 773–783.
Gillman, Max, Nakov, Anton, 2009. Monetary effects on nominal oil prices. North Am. J. Econ. Finance 20, 239–254.
Groen, Jan J.J., Pesenti, Paolo A., 2010. Commodity Prices, Commodity Currencies, and Global Economic Developments. NBER Working Paper 15743.
Gospodinov, Nikolay, Ng, Serena, 2013. Commodity prices, convenience yields and inflation. Rev. Econ. Stat. 95, 206–219.
Harris, Ethan S., Kasman, Bruce C., Shapiro, Matthew D., West, Kenneth D., 2010. Oil and the macroeconomy: lessons for monetary policy. In: Proceedings of the US Monetary Policy Forum, 2009, pp. 3–73.
Hubrich, Kirstin, West, Kenneth D., 2010. Forecast comparisons for small nested model sets. J. Appl. Econom. 25, 574–594.
Newey, Whitney K., West, Kenneth D., 1994. Automatic lag selection in covariance matrix estimation. Rev. Econ. Stud. 61, 631–654.
Ohashi, Kazuhiko, Okimoto, Tatsuyoshi, 2012. Increasing Trends in the Excess Comovement of Commodity Prices. working paper. Hitotsubashi University.
Stock, James H., Watson, Mark W., 2006. Forecasting with many predictors. In: Elliott, G., Granger, C.W.J., Timmermann, A. (Eds.), Handbook of Economic Forecasting, vol 1. Elsevier, Amsterdam, pp. 515–550.
Tang, Ke, Xiong, Wei, 2010. Index Investment and Financialization of Commodities. NBER Working Paper No. 16385.
West, Kenneth D., 1996. Asymptotic inference about predictive ability. Econometrica 64, 1067–1084.
West, Kenneth D., 2006. Forecast evaluation. In: Elliott, G., Granger, C.W.J., Timmerman, A. (Eds.), Handbook of Economic Forecasting, vol. 1. Elsevier, Amsterdam, pp. 100–134.