USING CO-MOVEMENTS TO FORECAST INDUSTRIAL COMMODITY PRICES

Kenneth D. West
University of Wisconsin

Ka-Fu Wong
University of Hong Kong

March 2013

I. Introduction

II. Data and factor model

III. Results

IV. Conclusion

I. Introduction

"Co-movement" widely perceived as a central characteristic of commodity prices

- Popular press

  "The past decade has been a remarkable one for metals and bulk commodities – iron ore and coal...Many analysts talked of a "supercycle", a long-term surge in prices...." (*Economist* July 2012)

- Scholarly work

  - Documentation of co-movment
  - Explanation of co-movement

- Scholarly work: documentation of co-movment
  - "Supercycles" (Cuddington and Jerrett (2008), Erten and Ocampo (2012))
  - Co-integration across nominal or real commodity prices (Chauduri (2001), Baffles (2007))
  - Factor model (Byrne et al. (2011)

- Scholarly work: explanation of co-movement
  - Attempts to tie movements to “fundamentals” have met with limited success
    - Pindyck and Rotemberg (1990) find co-movements far exceed what can be explained by industrial production and inflation
    - Byrne et al.’s (2011) factor is barely correlated with US GDP, real interest rates, etc.
  - In forecasting competitions, simple models such as a random walk often do as well or better than models that rely on futures prices or macroeconomic data

- Our paper

- We conjecture that a factor or factors constructed from a panel of commodity prices form a point of attraction towards which those prices revert

- Algebraic statement, for single factor model

  - f_t = factor f_t, a weighted average of commodity prices
  - δ_i = factor loading for i'th commodity price
  - $F_{it} \equiv \delta_i f_t$
  - p_{it} = real price of commodity i

  - After accounting for means, we conjecture that

$$F_{it} > p_{it} \Rightarrow \text{expect future } p_{it} \text{ to rise}$$

$$F_{it} < p_{it} \Rightarrow \text{expect future } p_{it} \text{ to fall}$$

- We evaluate the conjecture in part via pseudo out of sample forecasts, applied to a panel of 10 real commodity prices for oil, coal and metals (see Table 1), and for 3 horizons (1, 4 and 8 quarters)

- We find that the conjectured mean reversion is present

    - Our recursive set of forecasts rely on a sequence of 260 in-sample estimates of the correlation between future changes in p_{it} and F_{it}-p_{it}. All are positive, though numerically small.

    - We generate 30 time series of predictions (30 = 10 commodities × 3 horizons). In 29 of the 30, predictions are positively correlated with the realization, usually mildly so, occasionally strongly so.

•We also evaluate our forecasts by a root mean squared error (RMSPE) criterion.

•RMSPE for our model is better than a random walk in about half the comparisons (for example, 14 of 30 comparisons in our baseline model).

  •Magnitude of improvement typically is small, less than 5%

•Most though not all of the comparisons are significant at traditional levels.

•A test that accounts for the correlation across our 30 comparisons also finds that our model significantly improves relative to a random walk by our RMSPE criterion.

II. Data and factor model

- Real dollar prices of 10 commodities listed in Table 1, deflated by US CPI all consumers.
    - Quarterly data, 1980.1-2012.2. Nominal commodity price is average of last month of quarter.
- Basic statistics on levels and differences in Table 2.

- Estimation technique = principal components
    - Baseline model = 1 factor
- Illustrate mechanics for 1 quarter horizon.

- Using data from 1980.1-1989.4

  - Extract first principal component $\hat{f}_t$ and factor weights $\hat{\delta}_i$, i=1,...,10

  - Define $\hat{F}_{it} = \hat{\delta}_i \hat{f}_t$.

  - Using data from 1980.1-1989.3, do fixed effects regression

    $$\Delta p_{it+1} = \alpha_i + \beta_1(\hat{F}_{it} - p_{it}) + u_{it+1}$$

  - Prediction of $\Delta p_{i,1990.1} = \hat{\alpha}_i + \hat{\beta}_1(\hat{F}_{i,1989.4} - p_{i,1989.4})$. Compute and save prediction error

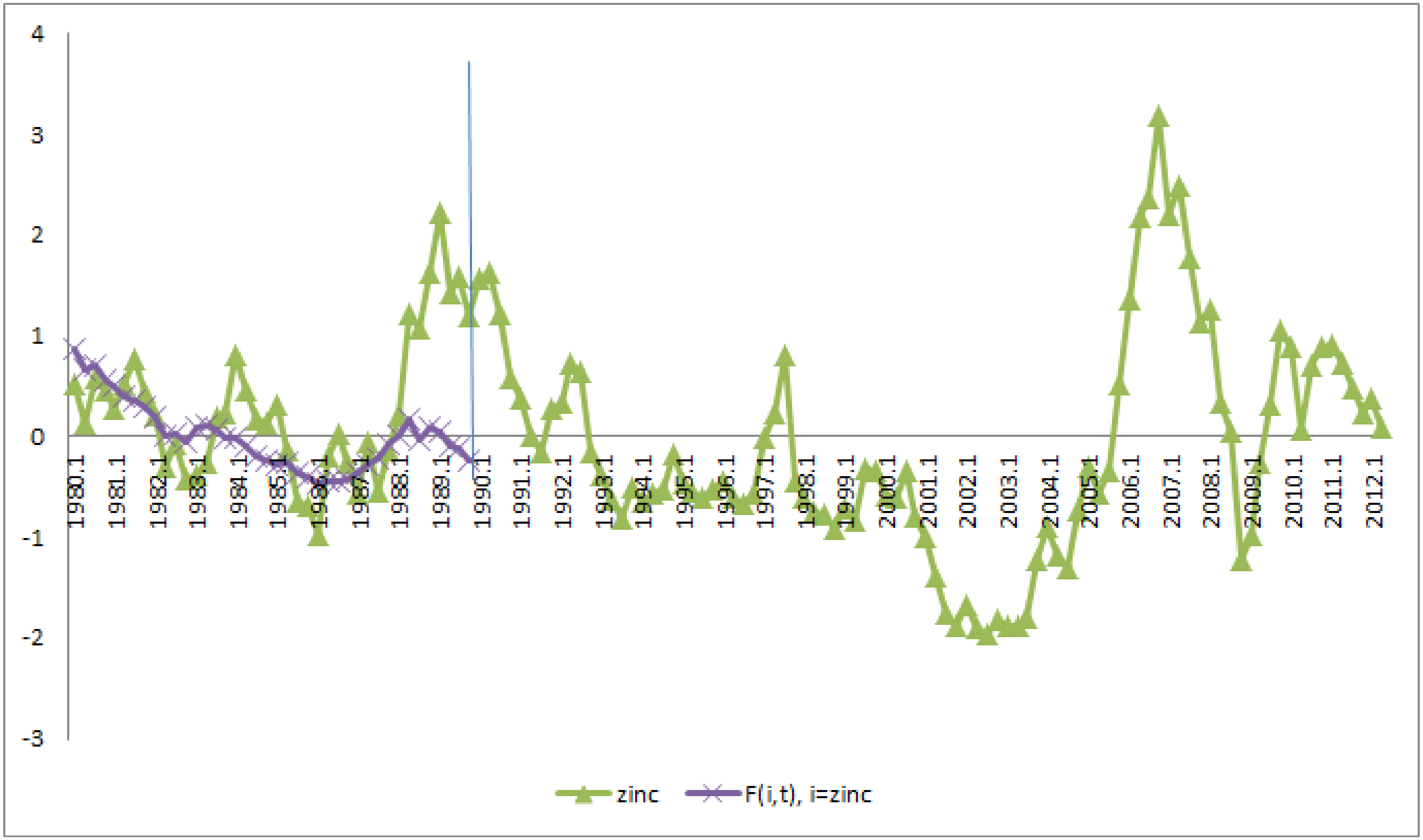
- Repeat, using data from 1980.1-1990.1; 1980.1-1990.2; ...; 1980.1-2012.1

•4 and 8 quarter predictions use same regressor $\hat{F}_{it}$-p_{it} but have 4 or 8 quarter changes in p_{it} on the left hand side (and thus use estimation samples that are 3 or 7 observations smaller than is the 1 quarter sample)

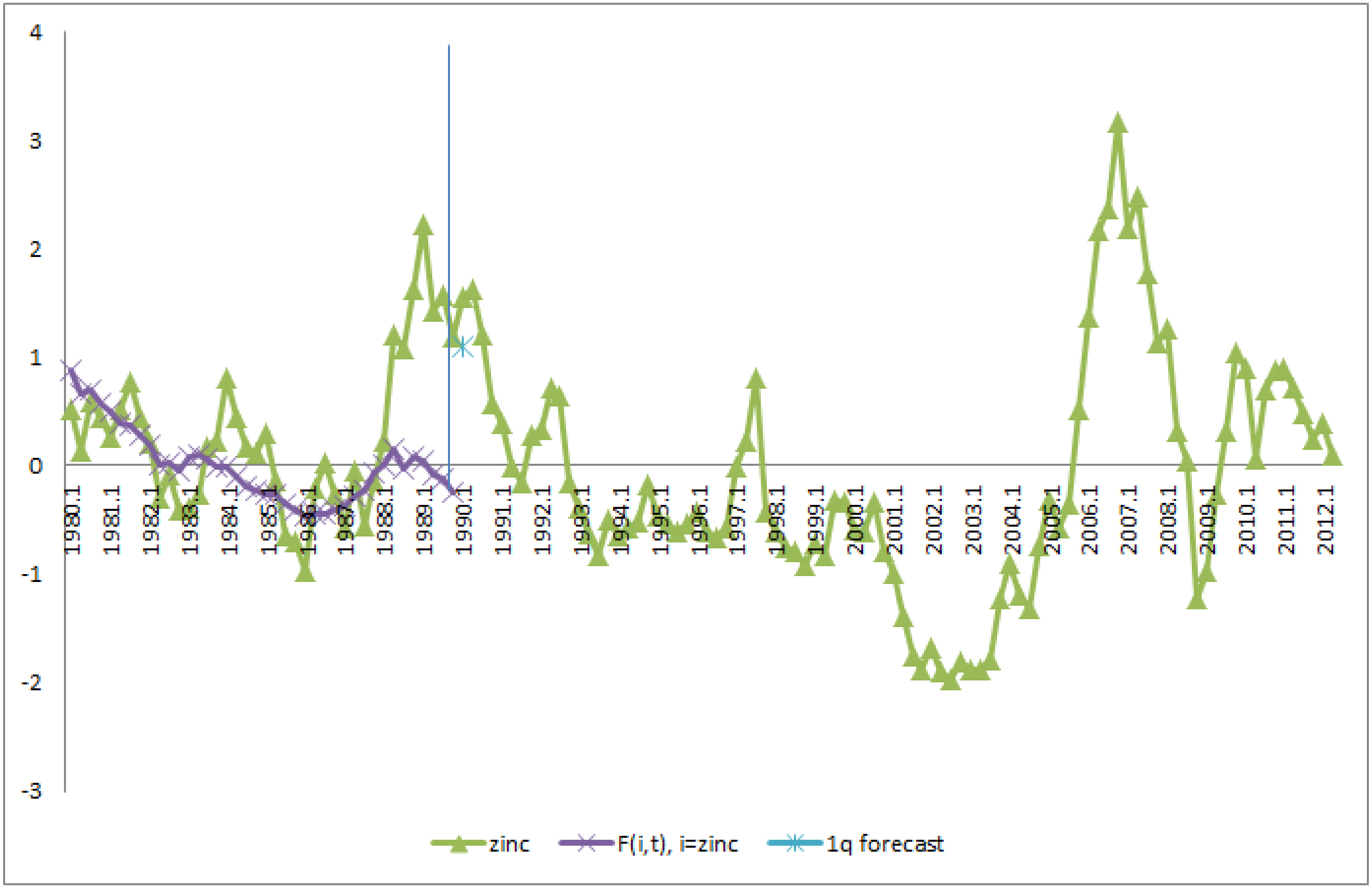
•Direct method used for 4 and 8 quarter predictions, recursive method for sequence of samples used in estimation.

III. Results

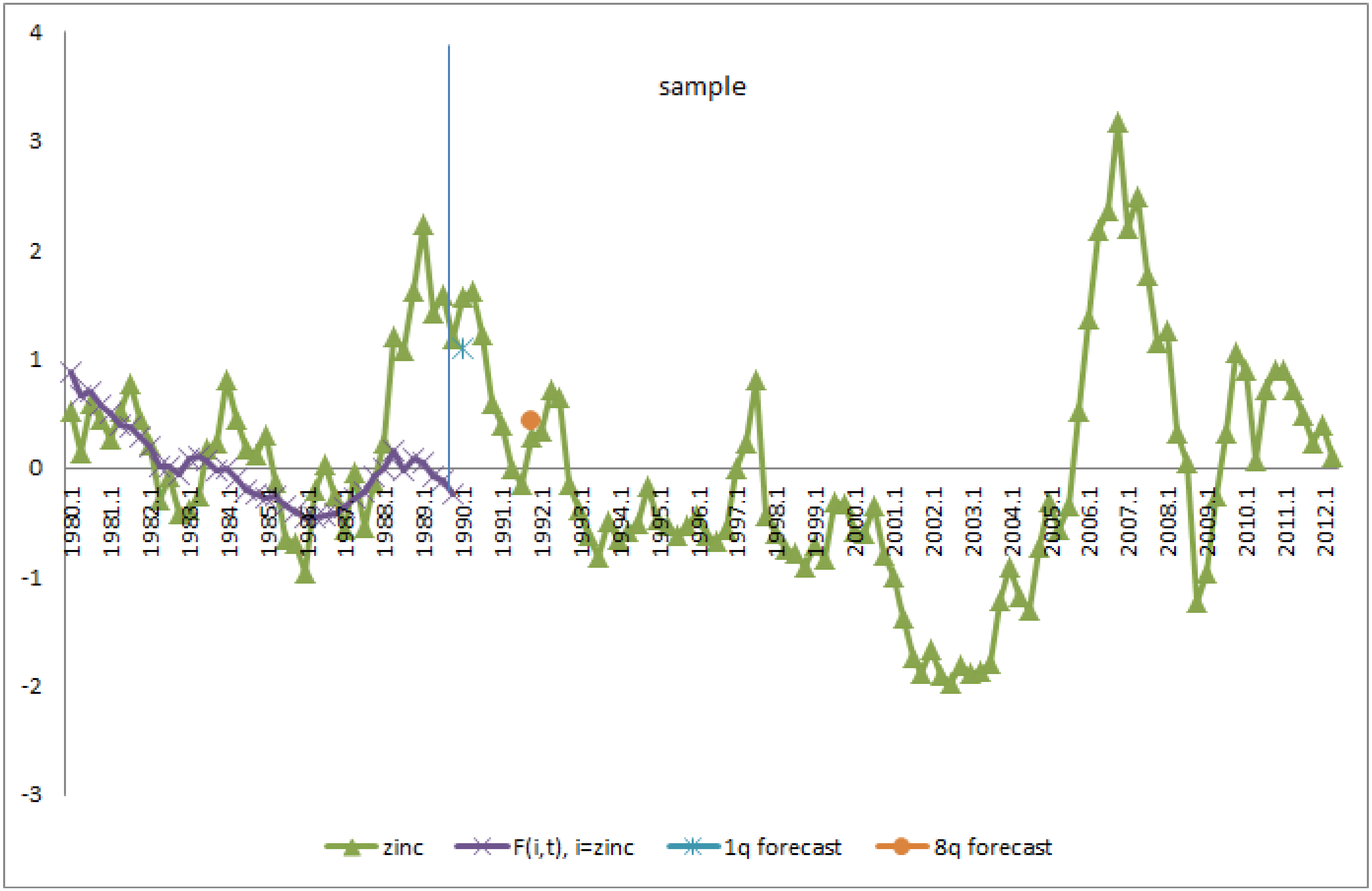
Illustrate with zinc, first couple of predictions.



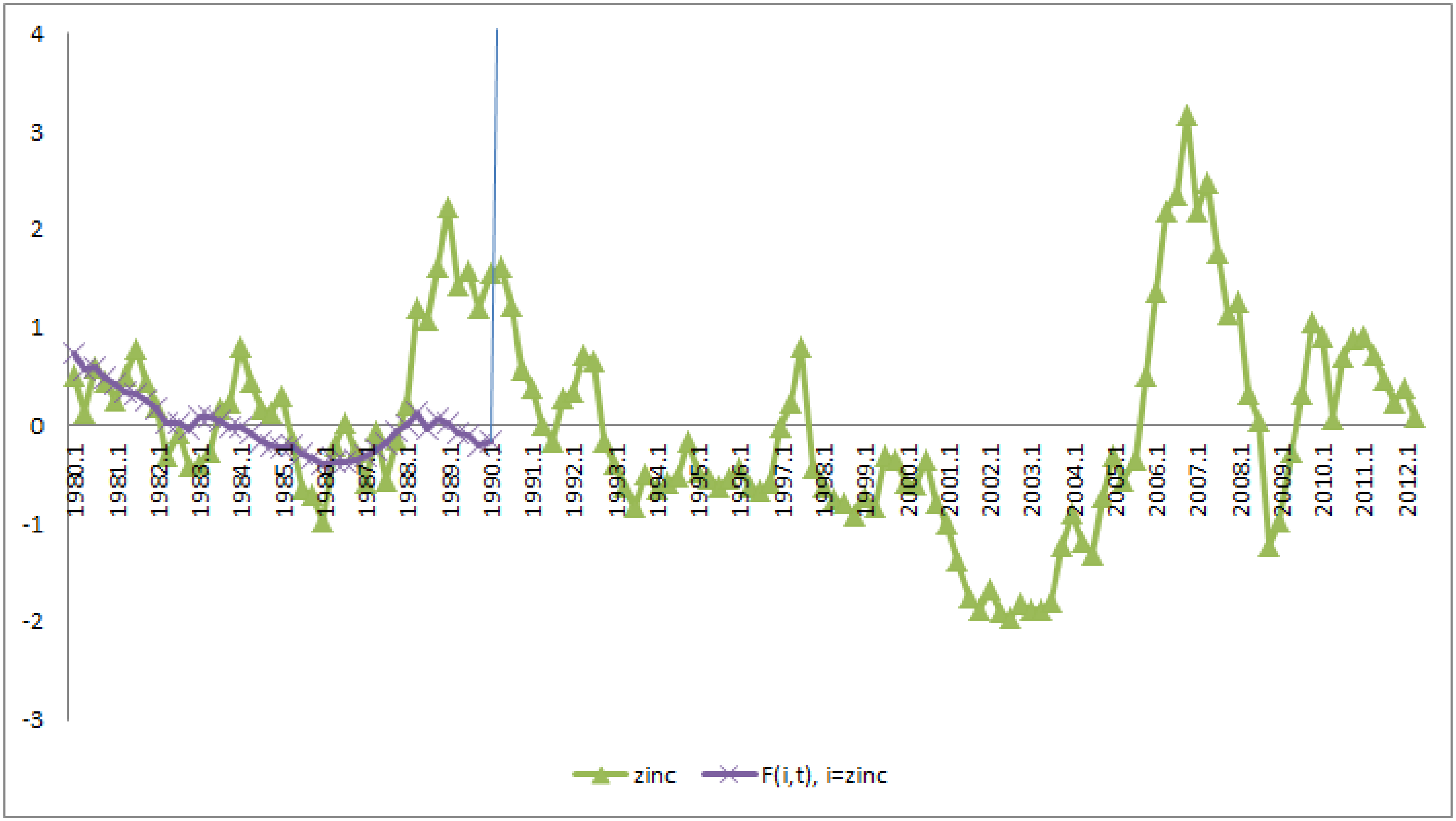
4
3
2
1
0
-1
-2
-3
1980.1
1981.1
1982.1
1983.1
1984.1
1985.1
1986.1
1987.1
1988.1
1989.1
1990.1
1991.1
1992.1
1993.1
1994.1
1995.1
1996.1
1997.1
1998.1
1999.1
2000.1
2001.1
2002.1
2003.1
2004.1
2005.1
2006.1
2007.1
2008.1
2009.1
2010.1
2011.1
2012.1
zinc
F(i,t), i=zinc



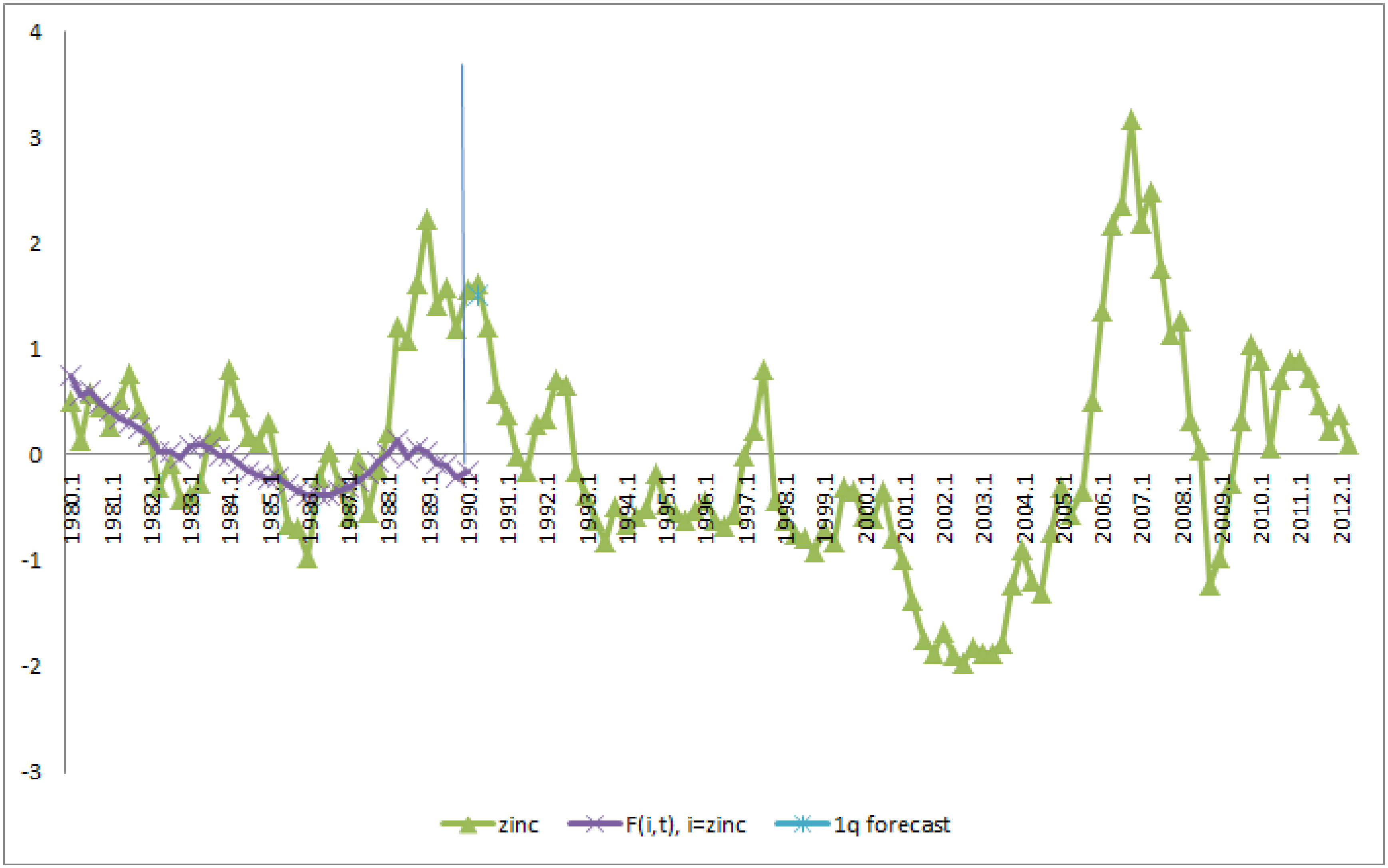
4
3
2
1
0
-1
-2
-3
1980.1
1981.1
1982.1
1983.1
1984.1
1985.1
1986.1
1987.1
1988.1
1989.1
1990.1
1991.1
1992.1
1993.1
1994.1
1995.1
1996.1
1997.1
1998.1
1999.1
2000.1
2001.1
2002.1
2003.1
2004.1
2005.1
2006.1
2007.1
2008.1
2009.1
2010.1
2011.1
2012.1
zinc
F(i,t), i=zinc
1q forecast

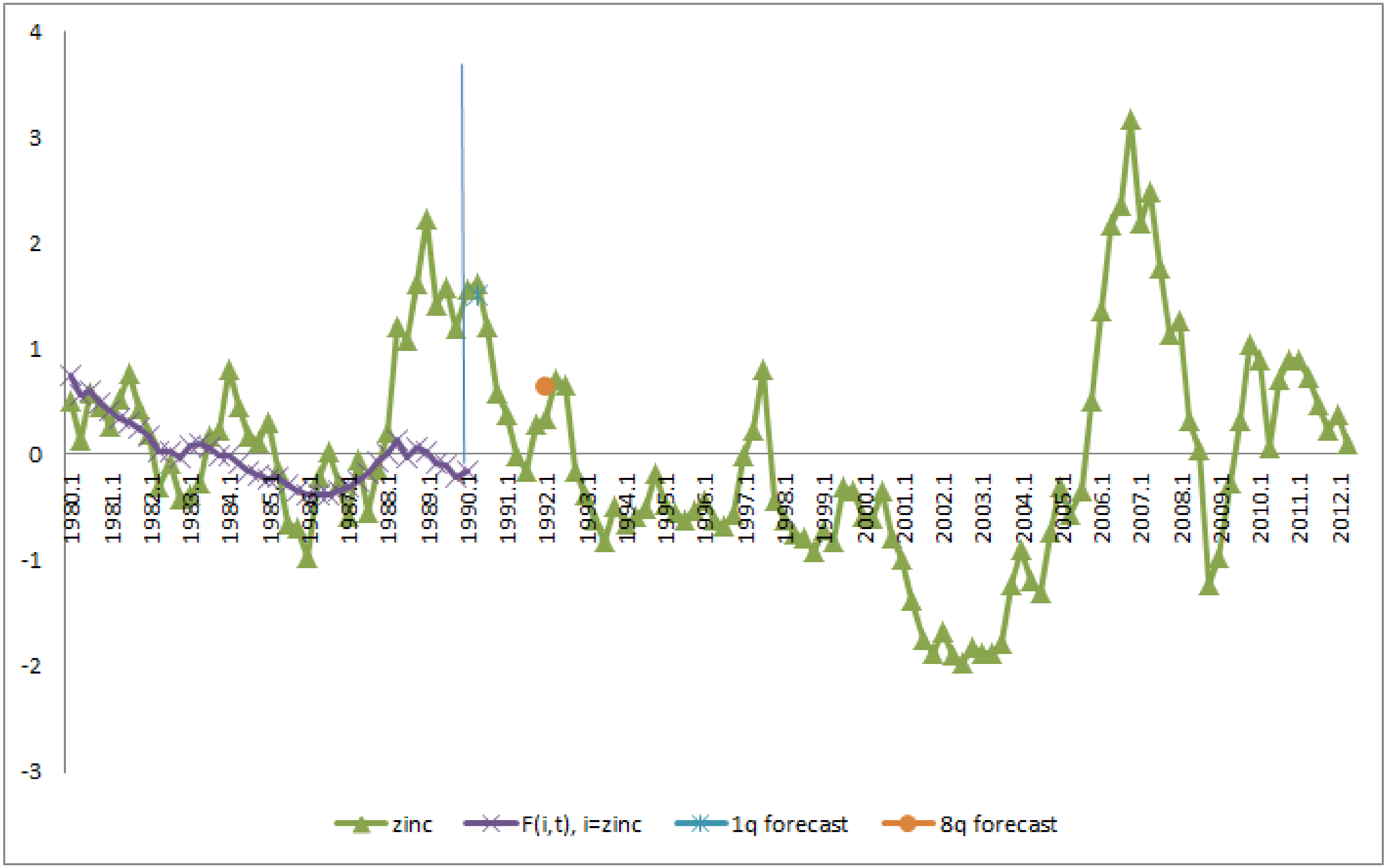


sample
zinc
F(i,t), i=zinc
1q forecast
8q forecast



4
3
2
1
0
-1
-2
-3
1980.1
1981.1
1982.1
1983.1
1984.1
1985.1
1986.1
1987.1
1988.1
1989.1
1990.1
1991.1
1992.1
1993.1
1994.1
1995.1
1996.1
1997.1
1998.1
1999.1
2000.1
2001.1
2002.1
2003.1
2004.1
2005.1
2006.1
2007.1
2008.1
2009.1
2010.1
2011.1
2012.1
zinc
F(i,t), i=zinc


4
3
2
1
0
-1
-2
-3
1980.1
1981.1
1982.1
1983.1
1984.1
1985.1
1986.1
1987.1
1988.1
1989.1
1990.1
1991.1
1992.1
1993.1
1994.1
1995.1
1996.1
1997.1
1998.1
1999.1
2000.1
2001.1
2002.1
2003.1
2004.1
2005.1
2006.1
2007.1
2008.1
2009.1
2010.1
2011.1
2012.1
zinc
F(i,t), i=zinc
1q forecast



4
3
2
1
0
-1
-2
-3
1980.1
1981.1
1982.1
1983.1
1984.1
1985.1
1986.1
1987.1
1988.1
1989.1
1990.1
1991.1
1992.1
1993.1
1994.1
1995.1
1996.1
1997.1
1998.1
1999.1
2000.1
2001.1
2002.1
2003.1
2004.1
2005.1
2006.1
2007.1
2008.1
2009.1
2010.1
2011.1
2012.1
zinc
F(i,t), i=zinc
1q forecast
8q forecast

Summary of forecasting results for baseline specification in Table 3

$$\text{“U”} = \frac{\text{RMSPE(factor model)}}{\text{RMSPE(random walk)}}\text{; U<1 means factor model “wins”}$$

30 comparisons (30 = 10 commodities × 3 horizons)

U<1 in 14, more at h=1 than h=4 or h=8
Of those 14, reject H_0:U=1 against H_A:U<1 in 9 cases

Results especially good for aluminum, nickel, zinc; especially bad for tin, uranium oil; coal, copper, lead, rubber in between “p value max t”: at the usual significance level, reject

H_0: U=1 for all commodities

against

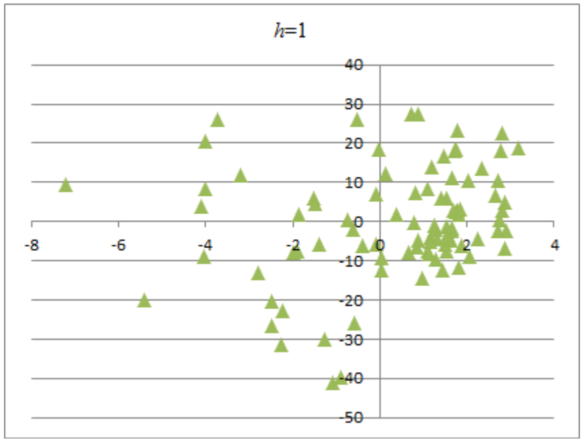
H_A: U<1 for at least one commodity.

Table 4: results not sensitive to number of factors or number of terms in $\hat{F}_{it}$-p_{it} in the regression

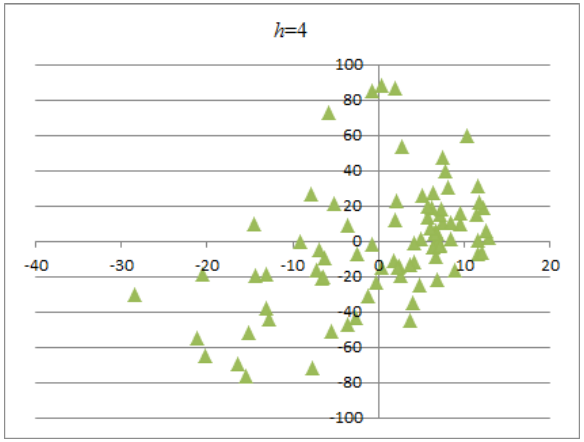
Directional accuracy: get the sign right of the commodity price change in 15 of 30 comparisons (30 = 10 commodities × 3 horizons)

Actual vs. Predicted Change in Real Zinc Prices

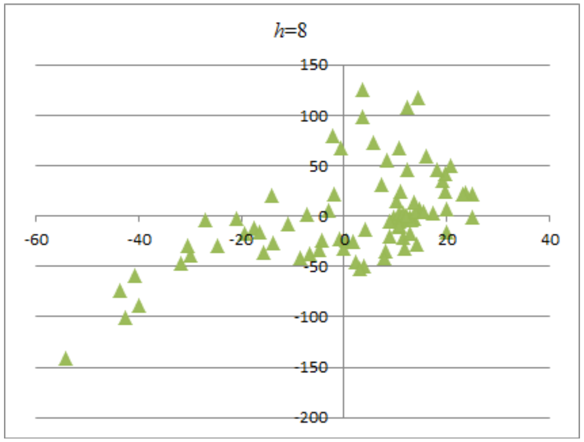
h=1



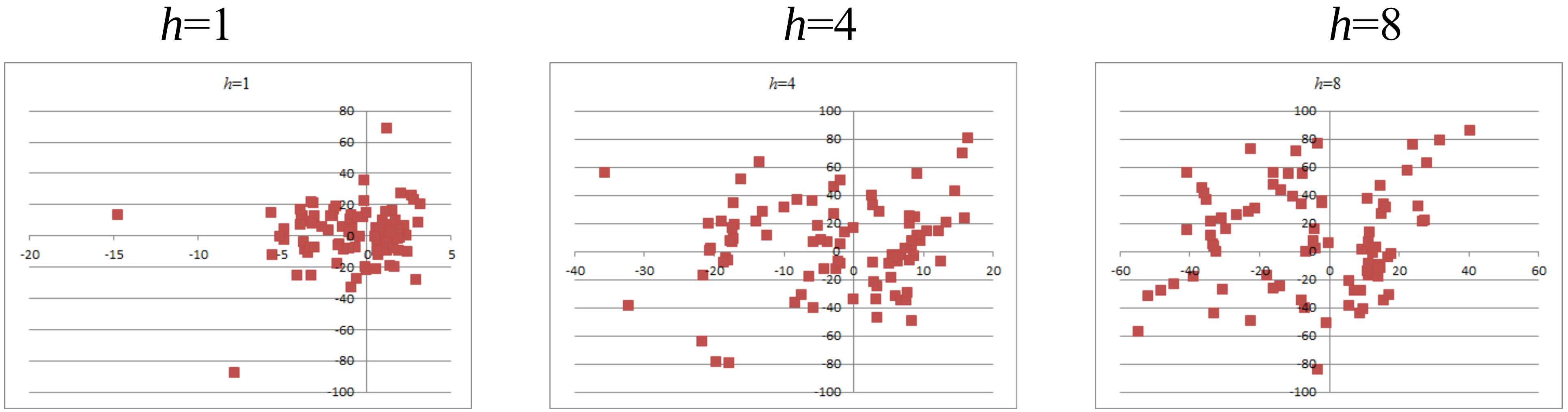
h=4



h=8



Actual vs. Predicted Change in Real Oil Prices



•Oil: U>1 (i.e., factor model RMSPE > random walk RMSPE) for h=4, 8, but

corr(prediction, realization) = 0.10 (h=4), 0.12 (h=8)

•Indeed, 29 of 30 such correlations were positive (exception: tin, h=8).

•U>1 nonetheless because

corr(prediction, prediction error) $\neq$ 0 for factor model

IV. Conclusion

- Commodity prices tend to revert towards a weighted average of commodity prices (a.k.a. factor).

- Mean reversion is slow, but reliable enough that exploiting this mean reversion sometimes results in forecasts that beat a random walk by a mean squared error criterion

- Possible extensions:

    - expanded data set
    - use of industry and macro data