Conference on

Forecasting and Monetary Policy

Berlin, 23-24 March 2009

Katrin Assenmacher-Wesche
Swiss National Bank

Discussion of „Efficient Estimation of Forecast Uncertainty“

Discussion of “Efficient Estimation of Forecast Uncertainty” by Malte Knüppel

Katrin Assenmacher
Swiss National Bank

March 23, 2009

The paper

- Estimate the forecast uncertainty of multi-step ahead forecasts.
  - Typically we report the MSFE (or MAFE) for each forecast horizon.
  - Using information from different horizons together can improve our estimate of the forecast uncertainty.
- Great paper.
  - Topic that is becoming increasingly noticed but not much investigated yet.
  - Focuses on forecast errors only, independent from the underlying model.
  - In small samples efficiency gains can be obtained.

General comments

- The paper contains too much.
    - Different data structures: RT, FM, FR.
    - Different estimators: GLS, SUR, FGLS.
    - Different data generating processes: AR(1), AR(2), VAR(2), Markov-Switching, non-invertible MA(19), ARCH(1,1), AR(1) with break in σ^2.
- This makes 71 pages, 28 tables and 5 appendices -

General comments

- The paper contains too much.
  - Different data structures: RT, FM, FR.
  - Different estimators: GLS, SUR, FGLS.
  - Different data generating processes: AR(1), AR(2), VAR(2), Markov-Switching, non-invertible MA(19), ARCH(1,1), AR(1) with break in σ^2.
- This makes 71 pages, 28 tables and 5 appendices -

A nightmare for the discussant!

General comments

- The paper contains too much.
  - Different data structures: RT, FM, FR.
  - Different estimators: GLS, SUR, FGLS.
  - Different data generating processes: AR(1), AR(2), VAR(2), Markov-Switching, non-invertible MA(19), ARCH(1,1), AR(1) with break in σ^2.
- This makes 71 pages, 28 tables and 5 appendices -

A nightmare for the discussant!

- Split or shorten the paper.
  - Improve notation: Two different notations for the RT structure, with the second one being much easier to understand.
  - Give the reader more intuition.
  - Focus on applicability of the method.

Data structures

- Why are the efficiency gains for different data structures so different?
- Example with H=2:
  - h = 1:
    - no efficiency gain for RT,
    - increasing efficiency gain with $|b_1|$ for FR,
    - decreasing efficiency gain with $|b_1|$ for FM.
  - h = 2:
    - Non-monotonous efficiency gain for RT,
    - increasing efficiency gain with $|b_1|$ for FR,
    - decreasing efficiency gain with $|b_1|$ for FM.
- Intuition?

Efficiency gains for AR(1)

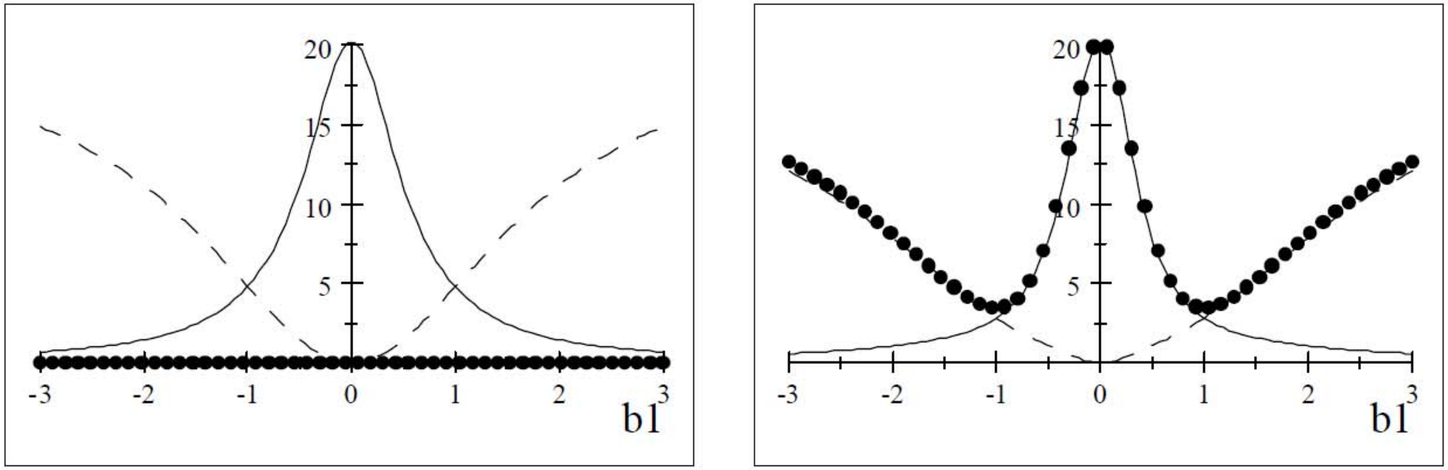


Figure 1: Efficiency gains φ_1 (left panel) and φ_2 (right panel) as functions of b_1 for the FR data structure ($N = 2, H = 2$, dashed line), the FM data structure ($N = 2, H = 2$, solid line) and the RT data structure ($N = 3, H = 2$, dots), always with $\alpha = 3$.

Data structures for H=2

From/for period	1	2	3	4	5
0	e_{10}	e_{20}			
1		e_{21}	e_{31}		
2			e_{32}	e_{42}	
3				e_{43}	e_{53}
4					e_{54}

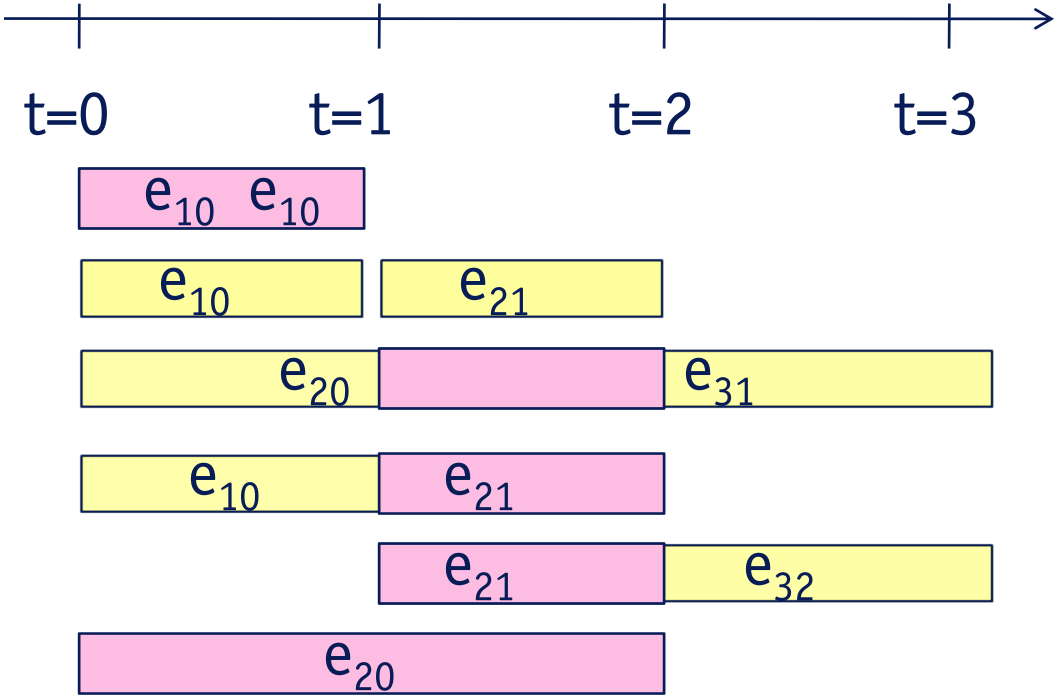
RT: All forecast errors: 5 observations for h=1, 4 for h=2.

FM: Forecast errors **from** a certain time span (without e_{54}).

FR: Forecast errors **for** a certain time span (without e_{10}).

Using all available information is always most efficient.

The covariances (ignoring $2\sigma^4$)



$\mathrm{cov} = 1$

$\mathrm{cov} = 0$

$\mathrm{cov} = b_1^2$

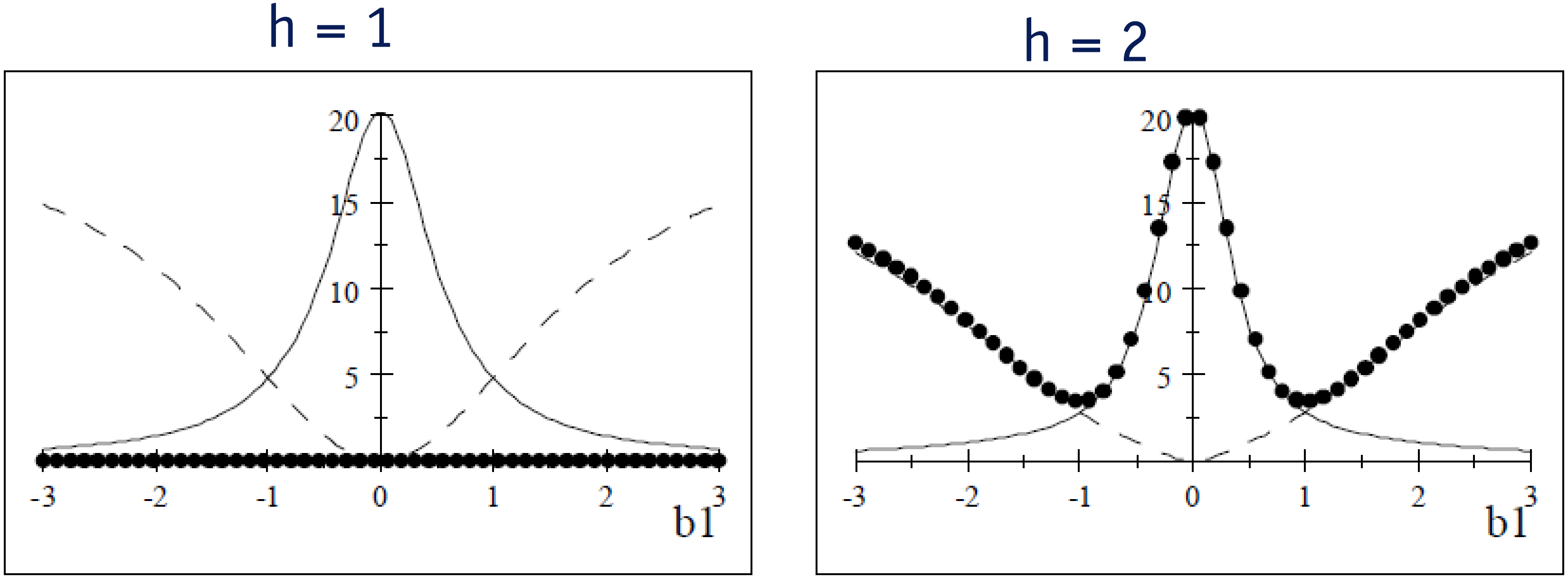
$\mathrm{cov} = b_1^2$

$\mathrm{cov} = b_1^2$

$\mathrm{cov} = (b_1^2+1)^2$

Overlap between forecasts determines the covariance.

Efficiency gains for AR(1)



Intuition: Substitute one observation with cov = 1 in FR structure by one observation with cov = b_1^2 in FM structure.

Intuition

$$\mathbf{\Omega}_{FR,2,2} = 2\sigma^4 \begin{array}{c} \begin{array}{cccc} e_{21} & e_{20} & e_{32} & e_{31} \end{array} \\ \left[\begin{array}{cccc} 1 & 1 & 0 & b_1^2 \\ 1 & (b_1^2+1)^2 & 0 & b_1^2 \\ 0 & 0 & 1 & \enclose{circle}{1} \\ b_1^2 & b_1^2 & \enclose{circle}{1} & (b_1^2+1)^2 \end{array}\right] \end{array} \begin{array}{c} \\ e_{21} \\ e_{20} \\ e_{32} \\ e_{31} \end{array}$$

$$\mathbf{\Omega}_{FM,2,2} = 2\sigma^4 \begin{array}{c} \begin{array}{cccc} e_{10} & e_{21} & e_{20} & e_{31} \end{array} \\ \left[\begin{array}{cccc} 1 & 0 & \enclose{circle}{b_1^2} & 0 \\ 0 & 1 & 1 & b_1^2 \\ \enclose{circle}{b_1^2} & 1 & (b_1^2+1)^2 & b_1^2 \\ 0 & b_1^2 & b_1^2 & (b_1^2+1)^2 \end{array}\right] \end{array} \begin{array}{c} \\ e_{10} \\ e_{21} \\ e_{20} \\ e_{31} \end{array}$$

What do we learn from this?

- Small sample aspect of a small sample problem – take any observation you can get.
- For RT structure: Efficiency gains only at longer horizons which have less observations (known from SUR literature).
- **RT is still most efficient at h = 1!**
- But note:
  - Differences between data structures will increase with H because more observations are cut off.
  - Adjusting *N* to have same number of forecasts to evaluate is possible in simulations but not quite realistic.

Feasible GLS

- True SUR (block-diagonal) covariance matrix instead of true covariance matrix still more efficient than OLS.
  - Remark:
    - $\rho = 1.5$ and $\rho = 2.0$ do not seem to be relevant in practice.
    - Derivation of covariance matrices assumes stationarity at the beginning.
- For SUR estimation with RT a certain structure of the covariance matrix is sufficient: No need to estimate any parameters.
- Intuition: We exploit the fact that forecast errors are MA(h-1) and that we have more observations at $h < H$.

A muddle of results for non-optimal forecasts

- Almost optimal forecasts with N=4, H = 2:
    - Estimating **b** by the Kalman filter is almost as efficient as GLS.
    - Direct estimation of Ω less efficient than OLS.
    - **Why?** Neglecting information about the structure of Ω?
- Strongly non-optimal forecasts
    - „The GLS estimator used here is *not* the true GLS estimator, but it is based on the AR(1)-model used for forecasting.“
    - Seems to be a deviation from the original intention of presenting a method that relies on the forecast errors only.
- Is this relevant if we can use SUR that does not depend on any estimated parameters?

What should the applied researcher do?

- Use all forecast errors you can get (RT structure).
- Don‘t care about the underlying model.
- Use SUR which takes account of the MA(h-1) structure of forecast errrors.