Forecasting Aggregated Time Series Variables
A Survey

Helmut Lütkepohl
European University Institute, Florence

Issues of Interest

- Forecasting macroeconomic variables is the theme of the conference
- Macro variables are often aggregates
- Types of aggregation:
  - temporal aggregation over a number of periods
  - contemporaneous aggregation
    - over countries, e.g., EU or euro-area
    - over regions of a specific country, e.g., Bundesländer in Germany
    - over sectors of an economy

Overview

Overview

VARMA model

Variables of interest: $y_t = (y_{1t}, \dots, y_{Kt})'$

$$y_t = A_1 y_{t-1} + \cdots + A_p y_{t-p} + u_t + M_1 u_{t-1} + \cdots + M_q u_{t-q}, \quad t \in \mathbb{Z}$$

or

$$A(L) y_t = M(L) u_t, \qquad t \in \mathbb{Z}$$

VARMA model

Variables of interest: $y_t = (y_{1t}, \dots, y_{Kt})'$

$$y_t = A_1 y_{t-1} + \cdots + A_p y_{t-p} + u_t + M_1 u_{t-1} + \cdots + M_q u_{t-q}, \quad t \in \mathbb{Z}$$

or

$$A(L) y_t = M(L) u_t, \qquad t \in \mathbb{Z}$$

with

- $A(L) = I_K - A_1 L - \cdots - A_p L^p$
- $M(L) = I_K + M_1 L + \cdots + M_q L^q$
- $u_t \sim (0, \Sigma_u)$

VARMA model

Variables of interest: $y_t = (y_{1t}, \dots, y_{Kt})'$

$$y_t = A_1 y_{t-1} + \cdots + A_p y_{t-p} + u_t + M_1 u_{t-1} + \cdots + M_q u_{t-q}, \quad t \in \mathbb{Z}$$

or

$$A(L) y_t = M(L) u_t, \qquad t \in \mathbb{Z}$$

with

- $A(L) = I_K - A_1 L - \cdots - A_p L^p$
- $M(L) = I_K + M_1 L + \cdots + M_q L^q$
- $u_t \sim (0, \Sigma_u)$
- $\det A(z) \neq 0,\ |z| \leq 1$ **and** $\det M(z) \neq 0,\ |z| \leq 1$ **for** $z \in \mathbb{C}$

Alternative representations

Wold MA representation

$$y_t = \sum_{i=0}^{\infty} \Phi_i u_{t-i}$$

with $\Phi(L) = I_K + \sum_{i=1}^{\infty} \Phi_i L^i = A(L)^{-1} M(L)$

Alternative representations

Wold MA representation

$$y_t = \sum_{i=0}^{\infty} \Phi_i u_{t-i}$$

with $\Phi(L) = I_K + \sum_{i=1}^{\infty} \Phi_i L^i = A(L)^{-1} M(L)$

VAR representation

$$y_t = \sum_{i=1}^{\infty} \Xi_i y_{t-i} + u_t$$

with $\Xi(L) = I_K - \sum_{i=1}^{\infty} \Xi_i L^i = M(L)^{-1} A(L)$

Alternative representations

Wold MA representation

$$y_t = \sum_{i=0}^{\infty} \Phi_i u_{t-i}$$

with $\Phi(L) = I_K + \sum_{i=1}^{\infty} \Phi_i L^i = A(L)^{-1} M(L)$

VAR representation

$$y_t = \sum_{i=1}^{\infty} \Xi_i y_{t-i} + u_t$$

with $\Xi(L) = I_K - \sum_{i=1}^{\infty} \Xi_i L^i = M(L)^{-1} A(L)$

Deterministic terms

$$y_t = \mu_t + x_t$$

Forecasting

Minimum MSE *h*-step forecast

$$y_{\tau+h|\tau} \equiv E(y_{\tau+h}|y_\tau, y_{\tau-1}, \dots)$$

Forecasting

Minimum MSE *h*-step forecast

$$y_{\tau+h|\tau} \equiv E(y_{\tau+h}|y_\tau, y_{\tau-1}, \dots)$$

Forecast error

$$\begin{aligned} y_{\tau+h} - y_{\tau+h|\tau} &= u_{\tau+h} + \Phi_1 u_{\tau+h-1} + \cdots + \Phi_{h-1} u_{\tau+1} \\ &\sim (\mathcal{N})(0, \Sigma_y(h)) \end{aligned}$$

Forecasting

Minimum MSE h-step forecast

$$y_{\tau+h|\tau} \equiv E(y_{\tau+h}|y_\tau, y_{\tau-1}, \dots)$$

Forecast error

$$\begin{aligned} y_{\tau+h} - y_{\tau+h|\tau} &= u_{\tau+h} + \Phi_1 u_{\tau+h-1} + \cdots + \Phi_{h-1} u_{\tau+1} \\ &\sim (\mathcal{N})(0, \Sigma_y(h)) \end{aligned}$$

MSE matrix

$$\Sigma_y(h) \equiv E[(y_{\tau+h} - y_{\tau+h|\tau})(y_{\tau+h} - y_{\tau+h|\tau})'] = \sum_{j=0}^{h-1} \Phi_j \Sigma_u \Phi_j'$$

$$\Sigma_y(h) \to_{h\to\infty} \Sigma_y \equiv E(y_t y_t') \text{ for stationary process}$$

Overview

Three predictors

Transformed process of interest: $z_t = Fy_t$

e.g., $z_t = \sum_{k=1}^{K} y_{kt} = [1, \ldots, 1]y_t$

Three predictors

Transformed process of interest: $z_t = F y_t$

e.g., $z_t = \sum_{k=1}^{K} y_{kt} = [1, \dots, 1] y_t$

The predictors

Three predictors

Transformed process of interest: $z_t = Fy_t$

e.g., $z_t = \sum_{k=1}^{K} y_{kt} = [1, \dots, 1]y_t$

The predictors

Forecast based on aggregated process

$$z_{\tau+h|\tau}$$

Three predictors

Transformed process of interest: $z_t = F y_t$

e.g., $z_t = \sum_{k=1}^{K} y_{kt} = [1, \dots, 1] y_t$

The predictors

Forecast based on aggregated process

$$z_{\tau+h|\tau}$$

Aggregate forecast of disaggregated process

$$z^o_{\tau+h|\tau} \equiv F y_{\tau+h|\tau}$$

Three predictors

Transformed process of interest: $z_t = F y_t$

e.g., $z_t = \sum_{k=1}^{K} y_{kt} = [1, \ldots, 1] y_t$

The predictors

Forecast based on aggregated process

$$z_{\tau+h|\tau}$$

Aggregate forecast of disaggregated process

$$z^o_{\tau+h|\tau} \equiv F y_{\tau+h|\tau}$$

Aggregate forecasts of individual components of disaggregated process

$$z^u_{\tau+h|\tau} \equiv F y^u_{\tau+h|\tau}$$

Properties of predictors

- All three forecasts are unbiased

Properties of predictors

- All three forecasts are unbiased
- $\Sigma_z(h) \geq \Sigma_z^o(h)$
- $\Sigma_z^u(h) \geq \Sigma_z^o(h)$

  (Forecasting disaggregated process and aggregating forecasts is optimal)
- $\Sigma_z(h) \geq \Sigma_z^u(h)$ or $\Sigma_z(h) \leq \Sigma_z^u(h)$

Properties of predictors

- All three forecasts are unbiased
- $\Sigma_z(h) \geq \Sigma_z^o(h)$
- $\Sigma_z^u(h) \geq \Sigma_z^o(h)$

  (Forecasting disaggregated process and aggregating forecasts is optimal)
- $\Sigma_z(h) \geq \Sigma_z^u(h)$ or $\Sigma_z(h) \leq \Sigma_z^u(h)$
- For stationary processes:

  $$\Sigma_z(h), \Sigma_z^o(h), \Sigma_z^u(h) \to \Sigma_z \equiv E(z_t z_t') \quad \text{as} \quad h \to \infty$$

Equality of predictors

Tiao and Guttman (1980), Kohn (1982), Lütkepohl (1984) give results on equality of predictors

Examples

- $z^o_{\tau+1|\tau} = z^u_{\tau+1|\tau}$ if y_t consists of independent components
- $z^o_{\tau+1|\tau} = z_{\tau+1|\tau}$ if y_t consists of independent components generated by the same ARMA process and $F = [1, \dots, 1]$

Overview

Results for estimated processes

Forecast based on estimated parameters: $\hat{y}_{\tau+h|\tau}$

Forecast error

$$y_{\tau+h} - \hat{y}_{\tau+h|\tau} = (y_{\tau+h} - y_{\tau+h|\tau}) + (y_{\tau+h|\tau} - \hat{y}_{\tau+h|\tau})$$

Forecast MSE

$$\Sigma_{\hat{y}}(h) = \text{MSE}(y_{\tau+h|\tau}) + \text{MSE}(y_{\tau+h|\tau} - \hat{y}_{\tau+h|\tau})$$

$$= \Sigma_y(h) + E[(y_{\tau+h|\tau} - \hat{y}_{\tau+h|\tau})(y_{\tau+h|\tau} - \hat{y}_{\tau+h|\tau})']$$

$$\approx \Sigma_y(h) + \frac{1}{T}\Omega(h)$$

$$\text{where } \Omega(h) = E\left[\frac{\partial y_{\tau+h|\tau}}{\partial \theta'} \Sigma_{\tilde{\theta}} \frac{\partial y'_{\tau+h|\tau}}{\partial \theta}\right]$$

Yamamoto (1980), Baillie (1981), Lütkepohl (2005)

Implications for aggregated processes

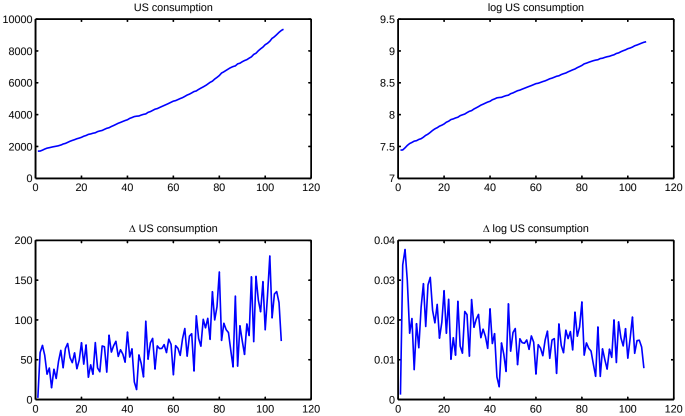
- $\hat{z}^o_{\tau+h|\tau}$ superior to $\hat{z}_{\tau+h|\tau}$ in general. There are exceptions.
- Specification uncertainty can make $\hat{z}_{\tau+h|\tau}$ superior to $\hat{z}^o_{\tau+h|\tau}$

Overview

Nonlinear transformations

The log transformation

- log transformations are common in time series analysis.
- Reasons:
  - Linearize economic relations.
  - Stabilize variance of a time series.
  - Make model residuals look normal.
- Question of interest:
  Is the log transformation useful for forecasting?


US consumption
10000
8000
6000
4000
2000
0
0 20 40 60 80 100 120
log US consumption
9.5
9
8.5
8
7.5
7
0 20 40 60 80 100 120
Δ US consumption
200
150
100
50
0
0 20 40 60 80 100 120
Δ log US consumption
0.04
0.03
0.02
0.01
0
0 20 40 60 80 100 120

Predictors

Linear predictor

$y^{lin}_{t+h|t}$ based on ARMA model for y_t.

Predictors

Linear predictor

$y^{lin}_{t+h|t}$ based on ARMA model for y_t.

Naive predictor

$y^{nai}_{t+h|t} = \exp x_{t+h|t}$,
where $x_{t+h|t}$ is based on ARMA model for $x_t = \log y_t$.

Predictors

Linear predictor

$y^{lin}_{t+h|t}$ based on ARMA model for y_t.

Naive predictor

$y^{nai}_{t+h|t} = \exp x_{t+h|t}$,
where $x_{t+h|t}$ is based on ARMA model for $x_t = \log y_t$.

Optimal predictor

$y^{opt}_{t+h|t} = \exp(x_{t+h|t} + \frac{1}{2}\sigma^2_x(h))$,
where $x_{t+h|t}$ is based on ARMA model for $x_t = \log y_t$ and $\sigma^2_x(h)$ is the corresponding forecast error variance.

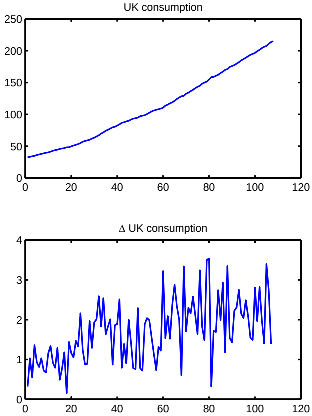
Table 1: Forecast MSEs (Naive/Linear) for Consumption Series

Country	h	Sample 1980Q1- Forecast period 2000Q1	Sample 1980Q1- Forecast period 2003Q1	Sample 1985Q1- Forecast period 2000Q1	Sample 1985Q1- Forecast period 2003Q1	Sample 1990Q1- Forecast period 2000Q1	Sample 1990Q1- Forecast period 2003Q1
US	1	1.0235	0.7678	1.0685	0.6963	1.1045	0.8015
	2	0.9416	0.4251	0.9251	0.4385	0.9587	0.5327*
	3	0.8956	0.2590	0.8348	0.3175	0.8810	0.4467
	4	0.8437	0.2879	0.7458	0.3156	0.7302	0.4228

Table 1: Forecast MSEs (Naive/Linear) for Consumption Series

		Sample 1980Q1- Forecast period		Sample 1985Q1- Forecast period		Sample 1990Q1- Forecast period	
Country	h	2000Q1	2003Q1	2000Q1	2003Q1	2000Q1	2003Q1
US	1	1.0235	0.7678	1.0685	0.6963	1.1045	0.8015
	2	0.9416	0.4251	0.9251	0.4385	0.9587	0.5327*
	3	0.8956	0.2590	0.8348	0.3175	0.8810	0.4467
	4	0.8437	0.2879	0.7458	0.3156	0.7302	0.4228
UK	1	1.8154	2.0701	1.8468	2.0323	1.5834	2.5313
	2	3.3214*	3.3397	3.5009	2.5500	2.7726*	3.1062
	3	4.0529*	5.7324*	4.1098*	3.3790*	3.0556*	3.6966
	4	5.2223*	7.8226*	5.1315*	4.6133*	3.7520*	4.6226*

Note: AR order selection based on SC with maximum lag order of 8.


UK consumption
250
200
150
100
50
0
0 20 40 60 80 100 120
log UK consumption
5.5
5
4.5
4
3.5
3
0 20 40 60 80 100 120
Δ UK consumption
4
3
2
1
0
0 20 40 60 80 100 120
Δ log UK consumption
0.04
0.03
0.02
0.01
0
0 20 40 60 80 100 120

Problems related to aggregation

- Theoretical results relate to linear aggregation
- Forecast aggregate or aggregate forecasts?
- Which forecast is optimal?

Non-Gaussian processes

- Linear forecasts may not be best (minimum MSE) anymore
- Use, e.g., bootstrap methods to compute interval forecasts

  Findley(1986), Masarotto (1990), Grigoletto (1998), Kabaila (1993), Kim (1999), Clements and Taylor (2001), Pascual, Romo and Ruiz (2004)

Aggregation with time-varying weights

EU unemployment rate

$$u_t^{EU} = \sum_{i=1}^{N} w_i u_t^{(i)}$$

Aggregation with time-varying weights

EU unemployment rate

$$u_t^{EU} = \sum_{i=1}^{N} w_i u_t^{(i)}$$

EMU growth rate (Winder 1997)

$$\Delta \log y_t^{EU} = \sum_{i=1}^{N} \frac{y_{t-1}^{(i)}/e_{TB}^{(i)}}{y_{t-1}^{EU}} \Delta \log y_t^{(i)}$$

where $e_t^{(i)}$ is the exchange rate of country i

Aggregation with time-varying weights

EU unemployment rate

$$u_t^{EU} = \sum_{i=1}^{N} w_i u_t^{(i)}$$

EMU growth rate (Winder 1997)

$$\Delta \log y_t^{EU} = \sum_{i=1}^{N} \frac{y_{t-1}^{(i)}/e_{TB}^{(i)}}{y_{t-1}^{EU}} \Delta \log y_t^{(i)}$$

where $e_t^{(i)}$ is the exchange rate of country i

EMU growth rate (Beyer, Doornik, Hendry 2001)

$$\Delta \log y_t^{EU} = \sum_{i=1}^{N} \frac{y_{t-1}^{(i)}/e_{t-1}^{(i)}}{y_{t-1}^{EU}} \Delta \log y_t^{(i)}$$

Structural change

Problem in constructing pre-EMU data:
Adjustment processes in some countries during run-up to EMU possibly induced structural changes in generation process of some variables

Possible solution:
Use German data for pre-EMU period
(Brüggemann and Lütkepohl, 2005, 2006)

Adjusting German data to EMU level

- No adjustment (GER0): West-German (70Q1-90Q4), German (91Q1-98Q4) and EMU data (99Q1-) have been joined

- Statistical adjustment (GER1): GER0 data has been increased to EMU level in 99Q1. German series has been adjusted multiplicatively using factor:

  $x = \frac{y_{99Q1}^{\text{EUR}}}{y_{99Q1}^{\text{GER0}}}$

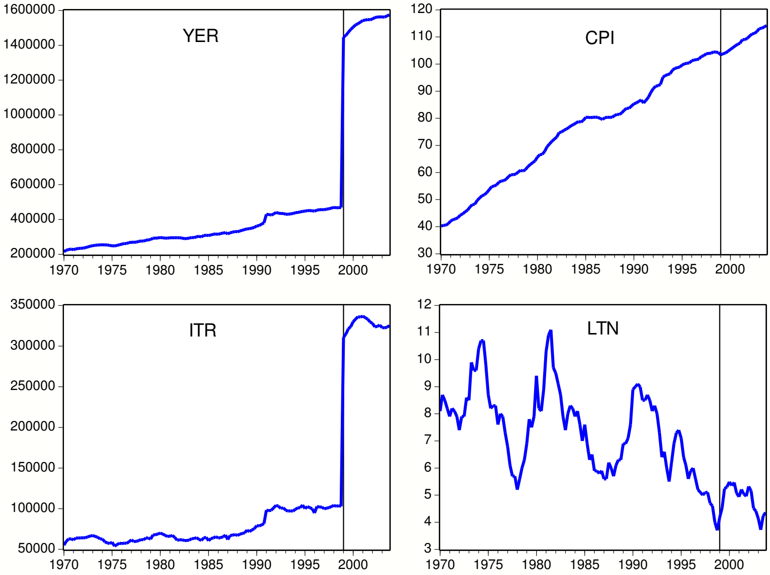


Figure 3: Time series based on German data: GER0

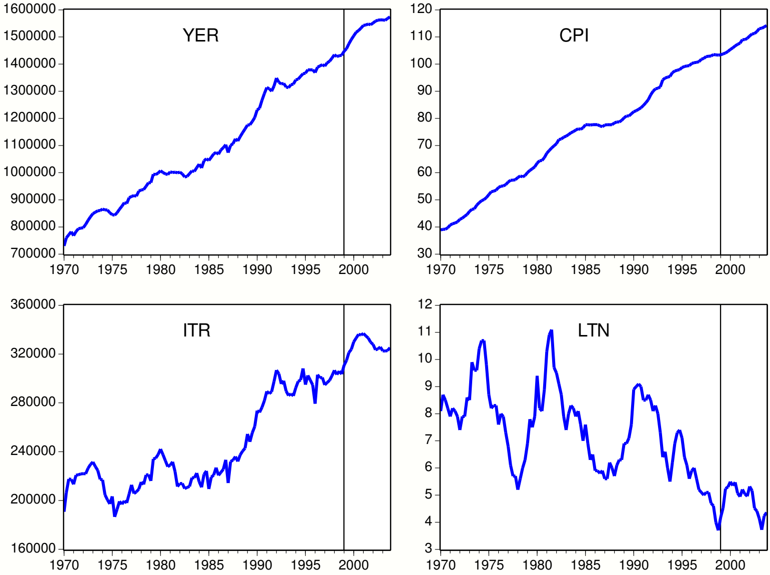


Figure 4: Time series based on German data: GER1

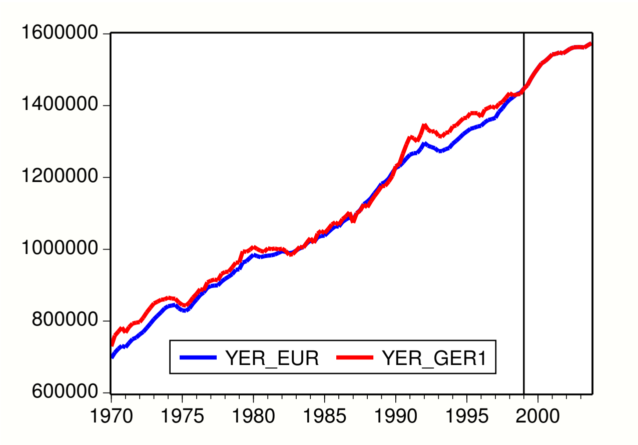


Figure 5: Real GDP based on German and European data

Forecast comparison

	Models used
AR1	AR(4) with constant, levels
AR2	AR(4) with linear trend, levels
AR3	AR(4) with constant, first differences
AR4	AR(4) with linear trend, first differences
AR5	AR(4) with constant, pretest for unit root
AR6	AR(4) with linear trend, pretest for unit root
AR7	AIC-AR with constant, levels
AR8	AIC-AR with linear trend, levels
AR9	AIC-AR with constant, first differences
AR10	AIC-AR with linear trend, first differences
AR11	AIC-AR with constant, pretest for unit root
AR12	AIC-AR with linear trend, pretest for unit root
AR13	BIC-AR with constant, levels
AR14	BIC-AR with linear trend, levels
AR15	BIC-AR with constant, first differences
AR16	BIC-AR with linear trend, first differences
AR17	BIC-AR with constant, pretest for unit root
AR18	BIC-AR with linear trend, pretest for unit root

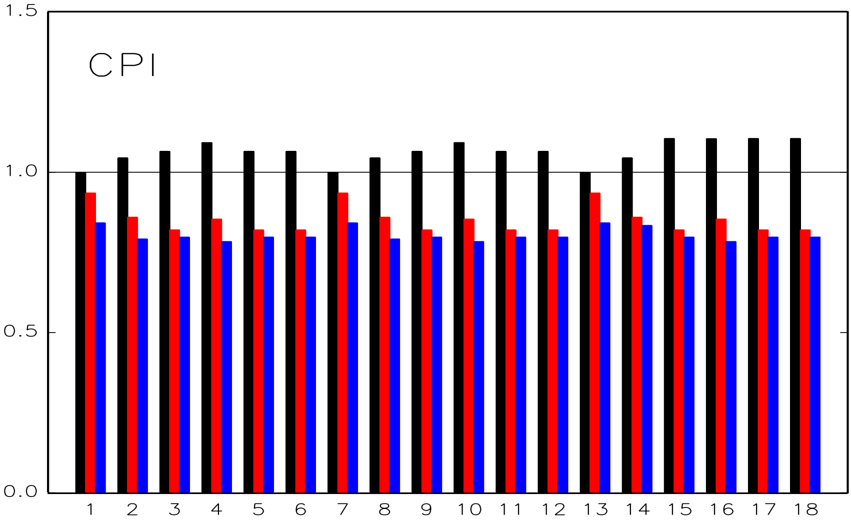


Figure 6: MSEs relative to the AR benchmark model for aggregated Euro-area data ($h = 1$). Black: EUR, Red: GER1, Blue: GER2

Overview

Conclusions

Predictors for contemporaneous aggregates:

- Direct forecast of aggregated variable
- Forecast disaggregated process and aggregate forecasts (is theoretically optimal)
- Forecast component series individually and aggregate forecasts

Many results are available for linear aggregates when the DGPs are known.

Conclusions

Practical problems for linear aggregates:

- Parameter estimation
- Model specification

More general practical problems:

- Nonlinear transformations, e.g., log
- Non-Gaussian processes
- Time-varying aggregation weights
- Structural change