Forecasting with large-scale macroeconometric models

Umidjon Abdullaev, Ulrich Gunter, Gašper Tompa

Econometric Forecasting

January 8th 2008
Introduction
The economic system and forecasting models
A taxonomy of forecast errors
Sources of forecast uncertainty

Agenda

1. Introduction
Agenda

1. Introduction
2. The economic system and forecasting models
   Taxonomy
   General formulation
Agenda

1. Introduction
2. The economic system and forecasting models
   Taxonomy
   General formulation
3. A taxonomy of forecast errors
Agenda

1. Introduction

2. The economic system and forecasting models
   Taxonomy
   General formulation

3. A taxonomy of forecast errors

4. Sources of forecast uncertainty
   Parameter change
   Model mis-specification
   Estimation uncertainty
   Variable uncertainty
   Error accumulation
Introduction

- They often consist of hundreds of equations.
- They typically share a focus on economic aggregates.
- They are also subject of critique.

"...existing Keynesian macroeconometric models are incapable of providing reliable guidance in formulating monetary, fiscal, or other types of policy. This conclusion is based in part on the spectacular recent failures of these models, and in part on their lack of a sound theoretical or econometric basis" ROBERT E. LUCAS Jr. and THOMAS J. SARGENT (1978)
Macroeconometric model building

- choosing the variables to be included in the model;
- separating variables: endogenous-exogenous;
- sketching a priori causal relationships among the variables, in the style of a flow chart;
- specification of estimable equations;
- estimation;
- forecasting.
Problems with large-scale ME forecasting

- Parameter non-constancy
- Parameters are unknown
- Forecast origin may be subject to error and revision, or may not yet be available.
- Non-modelled variables
- Error process distribution may change over time
- Extraneous information - not included in the model
- Cost of asymmetric loss - over and under prediction
Taxonomy

\[ x_t = \tau + \gamma x_{t-1} + \nu_t \]  
\[ \nu_t \approx N[0, \Omega], \text{ with expectation } E[\nu_t] = 0, \text{ and variance matrix } \]
\[ V[\nu_t] = \Omega \]

(1) satisfies \( r < n \) cointegration relation, such that:

\[ \gamma = I_n + \alpha \beta' \]  
\[ \text{where } \alpha \text{ and } \beta \text{ are } n \times r \text{ full rank matrices.} \]
\[ \Delta x_t = \tau + \alpha \beta' x_{t-1} + \nu_t \] (3)

\[ \tau = \gamma - \alpha \mu \] (4)

We define \( \gamma \) as the expected growth rate of the system: \( \gamma \equiv E[\Delta x_t] \), Taking expectation through (3):

\[ \gamma = \alpha E[\beta' x_{t-1}] + \tau \] (5)

\( \mu \) is the equilibrium mean \( \mu \equiv E[\beta' x_t] \)
\[ \Delta x_t - \gamma = \alpha (\beta' x_{t-1} - \mu) + \nu_t \] (6)

In (4) \(\tau\) is \(n \times 1\) and from: \(\beta' \gamma = E[\beta' \Delta x_t] = E[\Delta \beta' x_t] = 0\), where \(\beta\) is \(n \times r\), then \(\gamma\) consists of only \(n - r\) free parameters.

\(\gamma = \beta'_{\perp} \gamma^*\)
General formulation as \( \text{VAR}(q) \)

\[
x_t = \begin{pmatrix}
x_t \\
x_{t-1} \\
\vdots \\
x_{t-q+1}
\end{pmatrix} = \tau + \gamma x_{t-1} + \nu_t
\]
General formulation as $\text{VAR}(q)$

\[
\begin{pmatrix}
\tau \\
0 \\
\vdots \\
0
\end{pmatrix} + \begin{pmatrix}
\gamma_1 & \gamma_2 & \cdots & \gamma_{q-1} & \gamma_q \\
I_n & 0 & \cdots & 0 & 0 \\
0 & I_n & 0 & 0 & \cdots \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & \cdots & \cdots & I_n & 0
\end{pmatrix} \begin{pmatrix}
x_{t-1} \\
x_{t-2} \\
\vdots \\
x_{t-q}
\end{pmatrix} + \begin{pmatrix}
\nu_t \\
0 \\
\vdots \\
0
\end{pmatrix}
\]
General formulation as VAR(q)

\[ x_t = \tau_p + \Gamma_p x_{t-1} + \nu_t \]  

(\hat{\tau} : \hat{\Gamma} : \hat{\Omega})

\[ \tau_p \neq \tau \]

\[ \Gamma_p \neq \Gamma \]
Stationarity

If $\pi = E[x_t]$ the relationship between $\pi$ and $\tau$ is given by:

$$\tau = (I_n - \Upsilon)\pi$$

$$\Delta x_t = \gamma + \xi_t,$$

which is correctly specified when $\alpha = 0$ in (6), in which case $\xi_t = \nu_t$. 
Taxonomy is developed for a structural change over the forecast period, when the model and DGP may differ over the sample period.

The model:

\[ x_t = \tau + \gamma x_{t-1} + \nu_t \]

\( h \)-step forecast is given by:

\[ \hat{x}_{T+h} = \hat{\gamma} + \sum_{i=1}^{h-1} \gamma^i \hat{x}_{T+h-1} + \gamma^h \hat{x}_T \]

\( h = 1, \ldots, H, \)
Assume there is a single (one-step) break in the system, such that

\[(\tau : \nu) \rightarrow (\tau^* : \nu^*)\]

and the variance, autocorrelation and distribution of the error may change to

\[\nu_{T+h} \sim D_n(0, \Omega^*)\]
Thus, the data generated by the process for the next $H$ periods are given by:

$$X_{T+h} = \tau^* + \gamma^* x_{T+h-1} + \nu_{T+h}$$

$$= \sum_{i=1}^{h-1} (\gamma^*)^i \tau^* + (\gamma^*)^h x_T + \sum_{i=0}^{h-1} (\gamma^*)^i \nu_{T+h-i}$$
From above, the $h$-step forecast error is:

$$\hat{\nu}_{T+h} = \sum_{i=0}^{h-1} (\gamma^*)^i \tau^* - \sum_{i=0}^{h-1} \hat{\gamma}^i \hat{\tau} + (\hat{\gamma}^*)^h x_T$$

$$- \hat{\gamma}^h \hat{x}_T + \sum_{i=0}^{h-1} (\gamma^*)^i \nu_{T+h-i} \quad (1)$$
Let’s denote deviations between sample estimates and population parameters by $\delta_\tau = \hat{\tau} - \tau_p$ and $\delta_\Upsilon = \hat{\Upsilon} - \Upsilon_p$

Then,

$$\hat{\Upsilon}^i = (\Upsilon_p + \delta_\Upsilon)^i \approx \Upsilon_p^i + \sum_{j=0}^{i-1} \Upsilon_p^j \delta_\Upsilon \Upsilon_p^{i-j-1} = \Upsilon_p^i + C_i$$
Vectoring gives the following:

\[
\begin{align*}
\left( \hat{\gamma}^i \right)^\nu & \approx \left( \gamma_p^i \right)^\nu + \left( \sum_{j=0}^{i-1} \gamma_j p \otimes \gamma_i^{i-j-1} \right) \delta^\nu \\
\Rightarrow \hat{\gamma}^i \hat{\tau} & \approx \left( \gamma_p^i + \sum_{j=0}^{i-1} \gamma_j p \delta \tau \gamma_i^{i-j-1} \right) \left( \tau_p + \delta \tau \right) \\
& = \sum_{j=0}^{i=1} \gamma_j p \delta \tau \gamma_i^{i-j-1} \tau_p + \gamma_i^i \tau_p + \gamma_i^i \delta \tau
\end{align*}
\]
This yields

\[
\hat{\gamma}_i \hat{\tau} \approx \gamma_p \tau_p + \gamma_p \delta_{\tau} + \left( \sum_{j=0}^{i-1} \gamma_j \otimes \tau_p \gamma_{i-j-1} \right) \delta_{\tau}^{\nu}
\]

\[= \gamma_p \tau_p + \gamma_p \delta_{\tau} + D_i \delta_{\nu}^{\nu},\]

and we obtain...

\[
\sum_{i=0}^{h-1} \left[ (\gamma^*)^i \tau^* - \hat{\gamma}_i \hat{\tau} \right] \approx \sum_{i=0}^{h-1} \left[ (\gamma^*)^i \tau^* - (\gamma_p + \gamma_p \delta_{\tau} + D_i \delta_{\nu}^{\nu}) \right]
\] (2)
We decompose the term \( \sum_{i=1}^{h-1} [(\gamma^*)_i \tau^* - \hat{\gamma}_i \hat{\tau}] \) in (2) into

\[
\sum_{i=0}^{h-1} [(\gamma^*)_i (\tau^* - \tau) + (\gamma^*)_i (\tau - \tau_p) + ((\gamma^*)_i - \gamma_i) \tau_p] \\
+ \sum_{i=0}^{h-1} [(\gamma^i - \gamma^i_p) \tau_p - \gamma^i_p \delta \tau - D_i \delta \nu].
\]
Similarly, for the terms in (1) post-multiplied by $x_T$ or $\hat{x}_T$, using:

$$C_h x_T = \sum_{j=0}^{h-1} \gamma^j_p \delta_\tau \gamma^h_{-j-1} x_T = \left( \sum_{j=0}^{h-1} \gamma^k_p \otimes x_T^j \gamma^h_{-j-1} \right) \delta_\nu^{\nu} = F_h \delta_\nu^{\nu}$$

and letting $(x_T - \hat{x}_T) = \delta_x$:

$$(\gamma^*_h) x_T - \tilde{\gamma}^h \hat{x}_T \approx ( (\gamma^*_h) - \gamma^h ) x_T + (\gamma^h - \gamma^h_p) x_T - F_h \delta_\nu^{\nu} + (\gamma^h_p + C_h) \delta_x$$
Forecast error $\hat{\nu}_{T+h}$

- $\sum_{i=0}^{h-1} ((\tau^*)^i - \tau^i) \tau_p + ((\tau^*)^h - \tau^h) x_T$ - slope change

- $+ \sum_{i=0}^{h-1} (\tau^*)^i (\tau^* - \tau) - intercept change$

- $+ \sum_{i=1}^{h-1} (\tau^i - \tau^i_p) \tau_p + (\tau^h - \tau^h_p) x_T$ - slope misspecification

- $+ \sum_{i=0}^{h-1} (\tau^*)^i (\tau - \tau_p) - intercept misspecification$
• \(-\sum_{i=0}^{h-1} (D_i - F_i)\delta^\nu\) - slope estimation

• \(-\sum_{i=0}^{h-1} \gamma^i_p\delta^\tau\) - intercept estimation

• \(+(\gamma^h_p + C_h)\delta^\chi\) - initial estimation

• \(+\sum_{i=0}^{h-1} (\gamma^*^i)^\nu_{T+h-i}\) - error estimation
In general, there are five different sources of forecast errors, both in the closed system (or closed loop model) and in the open system (or open loop model), where the latter additionally comprises non-modelled (exogenous) policy variables.

Let us generalize the different parameter vectors or matrices $\tau$ and $\gamma$ from the closed system ($x_t = \tau + \gamma x_{t-1} + \nu_t$) or $\Gamma$ and $\Phi$ from the open system ($y_t = \Gamma z_t + \Phi y_{t-1} + \eta_t$) to a generic parameter vector $\theta$.

Then the error sources can be expressed as follows:
Parameter change: $\theta^* \neq \theta$

Model mis-specification: $\theta \neq \theta_p$

Estimation uncertainty: $E[(\hat{\theta} - \theta_p)(\hat{\theta} - \theta_p)'] \neq 0$

Variable uncertainty such as variable mis-measurement in the closed system ($x_T \neq \hat{x}_T$) and some other forms in the open system

Error accumulation because of $M[\sum_{i=0}^{h-1}(\nu^*)^i \nu_{T+h-i}] \neq 0$
Parameter change: $\theta^* \neq \theta$

- Probably the most important source of forecast uncertainty
- Usually: DGP is assumed to be time-invariant, which may not be the case due to structural change or regime shifts
- Parameter change and model mis-specification are often hard to distinguish

⇒ Model mis-specification and changes in relevant parameters may result in forecasting errors without structural breaks occurring in the underlying behavioral equations.
From table 7.1 in Clements and Hendry (1998, "A taxonomy of forecast errors", p. 166), the h-step ahead forecast error
\( \hat{\nu}_{T+h} = x_{T+h} - \hat{x}_{T+h} \) in case of slope and intercept change reads as follows:

\[
\hat{\nu}_{T+h} = \sum_{i=0}^{h-1} (\gamma^*)^i \tau^* - \sum_{i=0}^{h-1} \gamma^i \tau
\]

\[
+ \quad (\gamma^*)^h x_T - \gamma^h x_T + \sum_{i=0}^{h-1} (\gamma^*)^i \nu_{T+h-i}
\]
Assuming that only the intercept vector is non-constant \( \tau^* \neq \tau \), the forecast error \( \hat{\nu}_{T+h} \) is simplified to the following:

\[
\hat{\nu}_{T+h} = \sum_{i=0}^{h-1} \gamma^i (\tau^* - \tau) + \sum_{i=0}^{h-1} \gamma^i \nu_{T+h-i}
\]

This highlights a persistent (and typically increasing) bias in case \((\tau^* - \tau) > 0\).
However, changes in the equilibrium mean $\pi = E[x_t]$ are the major sources of error. As one may recall from the closed stationary system:

$$\pi = (I_n - \Upsilon)^{-1} \tau$$

Changes in $\pi$ may be "direct results" from changes in $\tau$ or "indirect results" from changes in $\Upsilon$. 
\[ x_t = \tau + \gamma x_{t-1} + \nu_t \]

With \( \pi = (I_n - \gamma)^{-1}\tau \):

\[ x_t - \pi = \gamma (x_{t-1} - \pi) + \nu_t \]

\( \Rightarrow \) h-step ahead predictor at forecast origin \( T \):

\[ \hat{x}_{T+h} - \pi = \gamma (\hat{x}_{T+h-1} - \pi) \]
\[ = \gamma^h (x_T - \pi) \]
If one assumes that \((\tau, \Upsilon)\) changes to \((\tau^*, \Upsilon^*)\) in period \(T + h\), with \(\tau^* = (I_n - \Upsilon^*)\pi^*:\)

\[
x_{T+h} - \pi^* = \Upsilon^*(x_{T+h-1} - \pi^*) + \nu_{T+h} \\
= (\Upsilon^*)^h(x_T - \pi^*) + \sum_{i=0}^{h-1}(\Upsilon^*)^i\nu_{T+h-i}
\]
The economic system and forecasting models

A taxonomy of forecast errors

Sources of forecast uncertainty

Parameter change

Model mis-specification

Estimation uncertainty

Variable uncertainty

Error accumulation

\[ \hat{\nu}_{T+h} = x_{T+h} - \hat{x}_{T+h} : \]

\[ \hat{\nu}_{T+h} = \pi^* - \pi + (\gamma^*)^h (x_T - \pi^*) - \gamma^h (x_T - \pi) \]

\[ + \sum_{i=0}^{h-1} (\gamma^*)^i \nu_{T+h-i} \]

\[ = (I_n - (\gamma^*)^h) (\pi^* - \pi) + ((\gamma^*)^h - \gamma^h) (x_T - \pi) \]

\[ + \sum_{i=0}^{h-1} (\gamma^*)^i \nu_{T+h-i} \]
Taking expectations from the last equation, with $E[x_T - \pi] = 0$, we get:

$$E[\hat{\nu}_{T+h}] = (I_n - (\Upsilon^*)^h) (\pi^* - \pi) > 0$$

Therefore, forecasts are only biased in case the long-run mean is not time-constant.

They would be unbiased if the long-run mean were time-constant ($\pi^* = \pi$), if the observed process had zero mean ($\pi^* = \pi = 0$), or if, e.g., changes in $\Upsilon^*$ could be compensated by changes in $\tau^*$. 

Umidjon Abdullaev, Ulrich Gunter, Gašper Tompa
Forecasting with large-scale macroeconometric models
⇒ Systematic mean-shift structural changes are the major source of forecast error in case of parameter change.

Changes in $\tau^*$ also affect the estimates of the forecast-error variances $M[\hat{\nu}_{T+h}]$, but these effects are secondary.
Model mis-specification: $\theta \neq \theta_p$

All models are wrong but some are useful. (G.E.P. Box)

- One usually does not know the true DGP.
- Parameterized (forecasting) models are designed to be congruent to the assumed DGP.
- Therefore, the included variables of the fitted model should be relatively orthogonal to wrongly excluded influences (omitted variable bias) to obtain consistent parameter estimates.
• To find a good parameterized model, one should apply the so-called Hendry’s method (general-to-specific approach).

• However, one should avoid over-fitting of the model because then secondary or irrelevant features of the data-set are likely to undermine forecasting performance.

• For forecasting purposes, it may also be useful to omit variables suspicious to structural breaks in case their omission has no crucial impact on the fit.
Estimation uncertainty: \[ E[(\hat{\theta} - \theta_p)(\hat{\theta} - \theta_p)'] \neq 0 \]

Estimation uncertainty depends on:

- Size and information content of the sample
- Quality of the model (goodness of fit)
- Choice of parameterization
- Selection of the estimator
- Degree of integration and existence of unit roots
Most of these problems can be improved by efficient specification, modelling and estimation strategies (Wald F tests on significance of variables, general-to-specific approach etc.).
Variable mis-measurement (closed system):
\[ x_T \neq \hat{x}_T \]

Initial condition uncertainty (open system):
\[ y_T \neq \hat{y}_T \]

A badly measured forecast origin may be an important source of short-term forecast errors.

This argument can be backed by the frequency with which provisional data is revised.

⇒ One should only use revised and reliable data.
Non-modelled variable uncertainty (open system): 

\[ z_{T+h} \neq \hat{z}_{T+h} \text{ for } h = 0, \ldots, H \]

Off-line extrapolations of policy variables can also constitute a source of forecast error to the model if not deliberately conducted.

Maybe some part of this flaw is compensated by intercept correction and other forms of ”correction” such as add factors...

⇒ Putting emphasis on reasonably predicting future developments of policy variables would be advisable.
Feedbacks on non-modelled variables (open system)

It could be the case that non-modelled variables $z$ are Granger-caused by the modelled variables $y$ as well (Granger feedback), although the $z$'s are assumed to be exogenous.

This would contradict the chosen parameterization of the model.

⇒ However, one can test for Granger feedback using Wald F tests.
Invariance to policy changes (open system)

Maybe the model is invariant to changes in the policy variables. This can be the case if the policy variables are not super exogenous, but only weakly exogenous.

Super exogeneity may be violated if agents change the way they form expectations (Lucas critique).

⇒ Then, future policy changes modelled off-line may have little or no impact in reality, contrary to what the forecast suggests.
One can control at least for part of this error source – namely possible impacts of forecasts on outcomes – by incorporating expected future values of the modelled variables into the forecasting model, e.g.:

\[ y_t = \alpha \hat{y}_{t+1} + \beta x_t + \nu_t \]
\[ \Rightarrow \hat{y}_\tau = \alpha \hat{y}_{\tau+1} + \beta x_\tau \]

with \( \hat{y}_{t+1} = E[y_{t+1} | \Omega_t] \), for \( \tau = T + 1, \ldots, T + H \).
Error accumulation: $M[\sum_{i=0}^{h-1} (\tau^*)^i \nu_{T+h-i}] \neq 0$

In the long run, the variance of a non-stationary DGP is unbounded and therefore the forecast error $\nu_{T+h}$ grows arbitrarily large.

Fortunately, only time ranges in terms of quarters or years are usually needed in applications such as economic stabilization policy or in investment strategies. In addition, taking differences, e.g., from a $I(1)$ process to obtain a $I(0)$ process imposes a boundary on the variance, which may be advantageous for forecasting.