Forecasting with Judgment

Simone Manganelli
DG-Research
European Central Bank
(Frankfurt am Main, Germany)

Disclaimer: The views expressed in this paper are our own and do not necessarily reflect the views of the ECB or the Eurosystem
Classical Theory of Forecasting

• Forecasts serve to make decisions about the future.
• Forecast errors impose costs on the decision-maker.
• Agents want to minimize expected loss associated to forecast errors.
• Classical forecast is the minimizer of the sample equivalent of the expected loss.
• See, e.g., Haavelmo (1944), Granger and Newbold (1986), Granger and Machina (2005).
Optimal Point Forecast

Simplest case:
• i.i.d. normally distributed observations with variance 1
• quadratic loss function

\[
\{y_t\}_{t=1}^T \quad \text{Want to forecast } y_{T+1}
\]

\[
\min_{\theta} Q_0(\theta) = E[(y_{T+1} - \theta)^2]
\]

FOC: \[-2E[y_{T+1} - \theta] = 0 \quad \Rightarrow \quad \hat{\theta}_T = T^{-1} \sum_{t=1}^T y_t
\]

\[
\theta^* = E[y_{T+1}]
\]
The Problem

\( \hat{\theta}_T \) is the minimiser of:

\[
\min_{\theta} Q_T(\theta) \equiv T^{-1} \sum_{t=1}^{T} [(y_t - \theta)^2]
\]

\[
\Rightarrow \min_{\theta} \{E[(y_{T+1} - \theta)^2] + T^{-1} \sum_{t=1}^{T} [(y_t - \theta)^2] - E[(y_{T+1} - \theta)^2]\} = \varepsilon_T(\theta)
\]

\[
\Rightarrow \min_{\theta} \{E[(y_{T+1} - \theta)^2] + \varepsilon_T(\theta)\}
\]
Outline

Introducing Judgment
Risk Analysis of New Estimator
Example 1: Asset Allocation
Example 2: Forecasting US GDP
Relationship with Bayesian
Conclusion
Start from a Subjective Guess

\[ \min_{\theta} Q_0(\theta) \equiv E[(y_{T+1} - \theta)^2] \]

\[ \nabla Q_0(\theta) \equiv E[y_{T+1} - \theta] = 0 \]

The sample equivalent of the FOC is:

\[ \nabla \hat{Q}_T(\tilde{\theta}) \equiv T^{-1} \sum_{t=1}^{T} [y_t - \tilde{\theta}] = \hat{y}_T - \tilde{\theta} \]

where \( \tilde{\theta} \) is a subjective guess of the decision maker.

**NOTE**: This is a random variable, which may be different from zero just because of statistical error.
Then Do Hypothesis Testing

\[ H_0 : \tilde{\theta} \text{ is the true mean} \]

\[ \hat{y}_T - \tilde{\theta} \sim N(0,1/T) \]

• Choose a confidence level \( \alpha \)
• \( \eta_{\alpha/2} \) is the critical value

If null is rejected:
→ First derivatives statistically \( \neq 0 \)
→ Can be confident to decrease the true objective function
→ Only up to the point \( \hat{\theta}_T^* \) where new \( H_0 \) cannot be rejected
Graphical Illustration

Distribution of FOC under $H_0$

$\nabla \hat{Q}_T(\tilde{\theta}) \rightarrow \nabla \hat{Q}_T(\hat{\theta}^*_T)$

$\alpha/2$
Intuition

First derivatives not statistically different from zero

Introducing Judgment
Risk Analysis of New Estimator
Example 1: Asset Allocation
Example 2: Forecasting US GDP
Relationship with Bayesian
Conclusion
New estimator:

\[ \theta_T^* = \begin{cases} 
\hat{y}_T - \eta_{\alpha/2} & \text{if } \hat{y}_T - \tilde{\theta} > \eta_{\alpha/2} \\
\tilde{\theta} & \text{if } |\hat{y}_T - \tilde{\theta}| < \eta_{\alpha/2} \\
\hat{y}_T + \eta_{\alpha/2} & \text{if } \hat{y}_T - \tilde{\theta} < -\eta_{\alpha/2}
\end{cases} \]
Interpretation

• $\alpha$ is the probability of committing type I errors, i.e. of rejecting the null when it is true.

• Choose low $\alpha$ when confident in subjective guess or if the cost of type I errors is high.

• Classical paradigm sets $\alpha = 1$:
  – Always FOC equal to zero;
  – No room for subjective guess;
  – It commits type I errors with probability 1.
Problems with Pretest Estimators

Test the following null hypothesis, for given confidence level $\alpha$:

$$H_0 : \hat{\theta}_T = \tilde{\theta}$$

Then:

• If do not reject, keep the subjective guess
• If reject, take the maximum likelihood estimator

This is wrong!
Introducing Judgment
Risk Analysis of New Estimator
Example 1: Asset Allocation
Example 2: Forecasting US GDP
Relationship with Bayesian
Conclusion
Introducing Judgment

**Risk Analysis of New Estimator**

Example 1: Asset Allocation

Example 2: Forecasting US GDP

Relationship with Bayesian

Conclusion
Let \( y_i \sim i.i.d. N(\theta^0, 1) \)

Estimate the mean, given a single observation \( y_1 \) (Magnus 02)

Compare the risk properties of two estimators:

1) Subjective classical estimator (coincides with Magnus 02)

\[
\theta_1^* = \begin{cases} 
  y_1 - \kappa_{\alpha/2} & \text{if } y_1 - \tilde{\theta} > \kappa_{\alpha/2} \\
  \tilde{\theta} & \text{if } |y_1 - \tilde{\theta}| < \kappa_{\alpha/2} \\
  y_1 + \kappa_{\alpha/2} & \text{if } y_1 - \tilde{\theta} < -\kappa_{\alpha/2}
\end{cases}
\]

2) Pretest estimator

\[
\hat{\theta}^p = \begin{cases} 
  \tilde{\theta} & \text{if } |y_1 - \tilde{\theta}| < \kappa_{\alpha/2} \\
  y_1 & \text{if } |y_1 - \tilde{\theta}| > \kappa_{\alpha/2}
\end{cases}
\]
Risk Associated to $f(y)$

$$E_{\theta^0}[(f(y) - \theta^0)^2], \; \tilde{\theta} = 0$$
Monte Carlo Simulation

• Random draws from standard normal
• Sample sizes = 5, 20, 60, 120, 240, 1000
• Quadratic loss function
• Two estimators: classical, new (α=0.10)
• Evaluated expected loss with MC simulation
• Repeat 5000 times and average
Introducing Judgment

Risk Analysis of New Estimator

Example 1: Asset Allocation
Example 2: Forecasting US GDP

Relationship with Bayesian

Conclusion
Implications

• Good guess (gut feelings) are as important as good econometric models.

• Organization of forecasting process:
  – Subjective guess based on maximum likelihood estimates can never be rejected by construction;
  – Clear separation b/w:
    Who provides guess, based on judgment;
    Who tests the guess, based on econometric models.

• Shared responsibility for the quality of the forecasts:
  – High confidence in bad judgment results in bad forecast.
Introducing Judgment
Risk Analysis of New Estimator
Example 1: Asset Allocation
Example 2: Forecasting US GDP
Relationship with Bayesian
Conclusion
Generalization

\[
\sqrt{T} \nabla_{\theta} \hat{U}_T(\theta^0) \overset{d}{\rightarrow} N(0, \Sigma)
\]

\[
\theta^*_T(\lambda) = \lambda \hat{\theta}_T + (1 - \lambda) \tilde{\theta} \quad \lambda \in [0,1]
\]

Under the null \( H_0 : \theta^*_T(\lambda) = \theta^0 \)

\[
\hat{z}_T(\theta^*_T(\lambda)) \equiv T \nabla_{\theta} \hat{U}_T'(\theta^*_T(\lambda)) \cdot \hat{\Sigma}^{-1} \cdot \nabla_{\theta} \hat{U}_T(\theta^*_T(\lambda))
\]

if \( \hat{z}_T(\theta^*_T(0)) \leq \eta_{\alpha,k} \) \( \Rightarrow \) \( \hat{\lambda} = 0 \)

if \( \hat{z}_T(\theta^*_T(0)) > \eta_{\alpha,k} \) \( \Rightarrow \) \( \hat{\lambda} = \arg \max_{\lambda \in [0,1]} \hat{U}_T(\theta^*_T(\lambda)) \)

s.t. \( \hat{z}_T(\theta^*_T(\lambda)) = \eta_{\alpha,k} \)
Mean-Variance Asset Allocation

• Mean-variance optimizers tend to produce portfolio allocations with little or negative investment value.

• “[They] overuse statistically estimated information and magnify the impact of estimation errors. It is not simply a matter of garbage in, garbage out, but, rather, a molehill of garbage in, a mountain of garbage out” (Michaud 1998)
Set up

- $N+1$ assets
- $y_{t,i}$ return of asset $i$
- $\theta_i$ weight of asset $i$ (they sum to 1)

$$y_t(\theta) = \sum_{i=1}^{N+1} \theta_i y_{t,i} \quad \text{portfolio return}$$
Optimization

\[
U[\theta; y_t] = E[y_t(\theta)] - \lambda V[y_t(\theta)]
\]

\[
= \theta' E[y_t] - \lambda \{\theta' V[y_t]\theta\}
\]

\[
= E[y_t(\theta)] - \lambda \{E[y_t^2(\theta)] - E[y_t(\theta)]^2\}
\]

\[
\hat{U}[\theta; y_t] = T^{-1} \sum_{t=1}^{T} y_t(\theta) - \lambda \{T^{-1} \sum_{t=1}^{T} y_t^2(\theta) - [T^{-1} \sum_{t=1}^{T} y_t(\theta)]^2\}
\]
Implementation Details

- Monthly log returns from DJIA
- From January 1987 to July 2005
- $\lambda = 4$ (coefficient of risk aversion)
- Rolling windows ($M=60$ and $120$ as in DeMiguel et al. 2007)
- Benchmark portfolio: equally weighted
- Report out of sample differences in average utilities between optimal and benchmark portfolios
Out of Sample Evaluation

\[ U_t(\theta) = y_t(\theta) - \lambda \{ y_t^2(\theta) - [(T - M)^{-1}] \sum_{t=M+1}^{T} y_t(\theta) \} y_t(\theta) \]

\[ Z_{T-M} \equiv (T - M)^{-1} \sum_{t=M+1}^{T} [U_t(\theta_t^{*}(\hat{\lambda})) - U_t(\widetilde{\theta})] \]

Test for statistical significance using Giacomini and White (2006)
## Introducing Judgment

Risk Analysis of New Estimator

Example 1: Asset Allocation

Example 2: Forecasting US GDP

Relationship with Bayesian

Conclusion

### Results

<table>
<thead>
<tr>
<th></th>
<th>M=60</th>
<th></th>
<th>M=120</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N=4</td>
<td>N=16</td>
<td>N=30</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>α=1</td>
<td>-0.30</td>
<td>1.00</td>
<td>-8.40***</td>
</tr>
<tr>
<td>α=0.10</td>
<td>0.48*</td>
<td>0.09</td>
<td>0</td>
</tr>
<tr>
<td>α=0.01</td>
<td>0.02</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Introducing Judgment
Risk Analysis of New Estimator
Example 1: Asset Allocation
Example 2: Forecasting US GDP
Relationship with Bayesian
Conclusion
Implementation Issues

Difficult to formulate guess on abstract model parameters.
Assume the decision-maker can formulate a guess on the variable to be forecast (U.S. GDP, in this example). This guess can be mapped into a guess for model parameters as follows:

\[ \tilde{\theta} = \arg \max_{\theta} \hat{U}_T(\theta) \]

\[ \text{s.t. } \hat{y}_{T+1}(\theta) = \tilde{y}_{T+1} \]

Subjective guess

Model’s forecast
Implementation Details

The model, AR(4):

\[ y_t = \theta_0 + \sum_{i=1}^{4} \theta_i y_{t-i} + \varepsilon_t \]

The objective function (quadratic):

\[ \hat{U}_T(\theta) = -T^{-1} \sum_{t=1}^{T} [y_t - \hat{y}_t(\theta)]^2 \]

The data:

• Quarterly U.S. real GDP growth rates
• FRED data base, seasonally adjusted
• From Q1 1983 to Q3 2005
## Results

<table>
<thead>
<tr>
<th>$\hat{y}_{T+1}$</th>
<th>$\hat{\theta}_T$</th>
<th>$\hat{\theta}$</th>
<th>$\theta_T^*(\hat{\lambda})$</th>
<th>$\tilde{y}_{T+1} = 3%$</th>
<th>$\tilde{y}_{T+1} = 5%$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta_0$</td>
<td>1.65</td>
<td>1.54</td>
<td>1.54</td>
<td>2.73</td>
<td>1.99</td>
</tr>
<tr>
<td>$\theta_1$</td>
<td>0.23</td>
<td>0.21</td>
<td>0.21</td>
<td>0.42</td>
<td>0.29</td>
</tr>
<tr>
<td>$\theta_2$</td>
<td>0.36</td>
<td>0.37</td>
<td>0.37</td>
<td>0.30</td>
<td>0.34</td>
</tr>
<tr>
<td>$\theta_3$</td>
<td>-0.16</td>
<td>-0.18</td>
<td>-0.18</td>
<td>-0.03</td>
<td>-0.12</td>
</tr>
<tr>
<td>$\theta_4$</td>
<td>0.04</td>
<td>0.05</td>
<td>0.05</td>
<td>-0.04</td>
<td>0.01</td>
</tr>
<tr>
<td>$\hat{y}_{T+1}(\theta)$</td>
<td>3.19%</td>
<td>3%</td>
<td>3%</td>
<td>5%</td>
<td>3.77%</td>
</tr>
</tbody>
</table>
Introducing Judgment
Risk Analysis of New Estimator
Example 1: Asset Allocation
Example 2: Forecasting US GDP
Relationship with Bayesian
Conclusion
Non sample information is summarised by:

- priors, in Bayesian econometrics
- subjective guess and confidence associated to it, in our case

Special cases:

1. No uncertainty about optimal value of $\theta^0$:
   - Bayesian: prior is degenerate distribution with total mass on $\theta^0$
   - Classical: subjective guess = $\theta^0$, $\alpha = 0$ (i.e. never reject the null)

2. No information, besides the sample:
   - Bayesian: diffuse prior
   - Classical: $\alpha = 1$ (i.e. set FOC equal to zero)

In other intermediate cases, no clear mapping b/w the two.

When non sample info if formulated:

- via prior distributions, be Bayesian
- via subjective guess and confidence, use subjective classical estimator

In general, the choice is dictated by the decision maker, through the format in which s/he provides the non sample info.
Conclusion

- Ignoring non sample information and estimation error are connected problems in the classical theory of forecasting.
- Forecasts should maximize the objective function in a stochastic sense, not deterministic.
- Start from subjective guess and construct estimator (instead of the other way round)
- Test if FOC are statistically different from zero:
  - If not, subjective guess is retained as forecast
  - If yes, objective function is increased as long as FOC are \( \sim= 0 \)