Università Commerciale Luigi Bocconi

# Paper: “Understanding Models’ Forecasting Performance”

by Barbara Rossi and Tatevik Sekhposyan

March 2011 (First version: January 2009)

## Student: Francesca Caturano

## Advanced Financial Econometrics III

February 2020

# Introduction: models' forecasting performance

This paper has two objectives:

1. Propose a new methodology for understanding why models have different forecasting performance
2. Apply the proposed methodology to study performance of models of exchange rate determination in OOS forecasting ability

- **Objective 1**: Why a model forecasts better than its competitors? Authors identify 3 possible sources of models' forecasting performance:

i. Predictive content → indicates the in-sample fit of OOS forecasting performance

ii. Over-fitting → is a situation in which a model includes irrelevant regressors, which improve the in-sample fit of the model but penalizes the model in an OOS forecasting exercise

iii. Time-varying forecasting ability (model instability) → might be caused by changes in the parameters of the models

# Introduction: models' forecasting performance

This paper has two objectives:

1. Propose a new methodology for understanding why models have different forecasting performance
2. Apply the proposed methodology to study performance of models of exchange rate determination in OOS forecasting ability

- **Objective 1**: Why a model forecasts better than its competitors? Authors identify 3 possible sources of models' forecasting performance:

i. **Predictive content** → indicates the in-sample fit of OOS forecasting performance

ii. **Over-fitting** → is a situation in which a model includes irrelevant regressors, which improve the in-sample fit of the model but penalizes the model in an OOS forecasting exercise

iii. **Time-varying forecasting ability** (model instability) → might be caused by changes in the parameters of the models

# Introduction: models' forecasting performance

- **Objective 2**: Better understand the sources of poor forecasting ability of the models, considering economic models of exchange rate determination that involve macroeconomic fundamentals
  - Empirical finding:

    exchange rate forecasts based on the random walk (RW) are superior to those of economic models on average over the OOS period
  - Reasons for inferior forecasting performance of macroeconomic fundamentals:

    - Lack of predictive content → major explanation for the lack of short-term forecasting ability

    - Instabilities → play a role especially for medium term (1y ahead) forecasts

# Introduction: models' forecasting performance

- **Objective 2**: Better understand the sources of poor forecasting ability of the models, considering economic models of exchange rate determination that involve macroeconomic fundamentals
  - Empirical finding:

    exchange rate forecasts based on the random walk (RW) are superior to those of economic models on average over the OOS period

  - Reasons for inferior forecasting performance of macroeconomic fundamentals:

    - Lack of predictive content → major explanation for the lack of short-term forecasting ability

    - Instabilities → play a role especially for medium term (1y ahead) forecasts

# Objective 1: The framework

Researcher is interested in evaluating model(s)' pseudo OOS forecasting performance, and comparing it with a measure of in-sample-fit so that the OOS forecasting performance can be decomposed into the 3 components

- It is common to compare models in a pseudo OOS forecasting environment
  - $h \geq 1$ denotes the (finite) forecasts horizon
  - We are interesting in evaluating the performance of $h$-steps ahead forecasts for the scalar variable $y_t$ using a vector of predictors $x_t$
  - Assume researcher has:

    **P** OOS predictions available,

    where the 1st OOS prediction is based on a parameter estimated using data up to time R; the 2nd prediction up to R+1, the last prediction up to $R + P - 1 = T$,

    where $R + P - 1 + h = T + h$ is the sample's size

## Objective 1: The framework

Researcher is interested in evaluating model(s)' pseudo OOS forecasting performance, and comparing it with a measure of in-sample-fit so that the OOS forecasting performance can be decomposed into the 3 components

- It is common to compare models in a pseudo OOS forecasting environment
  - $h \geq 1$ denotes the (finite) forecasts horizon
  - We are interesting in evaluating the performance of $h$-steps ahead forecasts for the scalar variable $y_t$ using a vector of predictors $x_t$
  - Assume researcher has:

    **P** OOS predictions available,

    where the 1st OOS prediction is based on a parameter estimated using data up to time R; the 2nd prediction up to R+1, the last prediction up to R + P - 1 = T,

    where R + P – 1 + $h$ = T + $h$ is the sample's size

# Objective 1: The framework

This paper focuses on measures of relative forecasting performance.

Let the two competing models be labelled as 1 and 2.

Model 1 is characterized by parameter $\alpha$ and model 2 by parameter $\gamma$

- Let $\{L_{t+h}(.)\}^{T}_{t=R}$ be a sequence of loss functions evaluating $h$-steps ahead OOS forecast errors, this framework is general enough to encompass:
  - measures of absolute forecasting performance, where $L_{t+h}(.)$ is the forecast error of a model
  - measures of relative forecasting performance, where $L_{t+h}(.)$ is the difference of the forecast error losses of the two competing models
  - Example: consider an unrestricted model $y_{t+h} = x_t'\alpha + \varepsilon_{t+h}$ and a restricted model $y_{t+h} = \varepsilon_{t+h}$, under a quadratic loss function,
  - the measure of relative performance of models would be $L_{t+h}(.) = (y_{t+h} - x_t'\hat{\alpha}_t)^2 - y_{t+h}^2$
  - Authors consider both a fixed rolling window and an expanding (recursive) window. The two differs for the data used to estimate parameters

# Objective 1: The framework

This paper focuses on measures of relative forecasting performance.

Let the two competing models be labelled as 1 and 2.

Model 1 is characterized by parameter $\alpha$ and model 2 by parameter $\gamma$

- Let $\{L_{t+h}(.)\}^{T}_{t=R}$ be a sequence of loss functions evaluating $h$-steps ahead OOS forecast errors, this framework is general enough to encompass:
  - measures of absolute forecasting performance, where $L_{t+h}(.)$ is the forecast error of a model
  - measures of relative forecasting performance, where $L_{t+h}(.)$ is the difference of the forecast error losses of the two competing models

    Example: consider an unrestricted model $y_{t+h} = x_t'\alpha + \varepsilon_{t+h}$ and a restricted model $y_{t+h} = \varepsilon_{t+h}$, under a quadratic loss function,

    the measure of relative performance of models would be $L_{t+h}(.) = (y_{t+h} - x_t'\hat{\alpha}_t)^2 - y_{t+h}^2$
  - Authors consider both a fixed rolling window and an expanding (recursive) window. The two differs for the data used to estimate parameters

## Objective 1: The framework

This paper focuses on measures of relative forecasting performance.

Let the two competing models be labelled as 1 and 2.

Model 1 is characterized by parameter $\alpha$ and model 2 by parameter $\gamma$

- Let $\{L_{t+h}(.)\}^{T}_{t=R}$ be a sequence of loss functions evaluating $h$-steps ahead OOS forecast errors, this framework is general enough to encompass:
  - measures of absolute forecasting performance, where $L_{t+h}(.)$ is the forecast error of a model
  - measures of relative forecasting performance, where $L_{t+h}(.)$ is the difference of the forecast error losses of the two competing models
  - Example: consider an unrestricted model $y_{t+h} = x_t'\alpha + \varepsilon_{t+h}$ and a restricted model $y_{t+h} = \varepsilon_{t+h}$, under a quadratic loss function,
  - the measure of relative performance of models would be $L_{t+h}(.) = (y_{t+h} - x_t'\hat{\alpha}_t)^2 - y_{t+h}^2$
  - Authors consider both a fixed rolling window and an expanding (recursive) window. The two differs for the data used to estimate parameters

# Objective 1: The framework – fixed rolling window case

In the fixed rolling window case, model's parameters are estimated using samples of R observations dated $t$ – R + 1,…, $t$ for $t$ = R, R + 1,… T

- Parameter estimates for model 1 are obtained by

  $\hat{\alpha}_{t,R} = \arg min_a \sum_{j=t-R+1}^{t} \mathcal{L}_j^{(1)}(a)$ , where $\mathcal{L}^{(1)}(.)$ denotes the in-sample loss function for model 1; similarly are found the parameter for model 2.

- At each point in time $t$, the estimation will generate a sequence of $R$ in-sample fitted errors denoted by $\{\eta_{1,j}(\hat{\alpha}_{t,R}), \eta_{2,j}(\hat{\gamma}_{t,R})\}^{t}{}_{j=t-R+1}$ among the R fitted values, we use the last in sample fitted errors at time $t$, $\{\eta_{1,t}(\hat{\alpha}_{t,R}), \eta_{2,t}(\hat{\gamma}_{t,R})\}$, to evaluate the models' in-sample fit at time $t$ $\mathcal{L}_t^{(1)}(\hat{\alpha}_{t,R})$ and $\mathcal{L}_t^{(2)}(\hat{\gamma}_{t,R})$

- As the rolling window estimation is performed for $t$ = R, R + 1,… T, we collect a series of in-sample losses $\left\{\mathcal{L}_t^{(1)}(\hat{\alpha}_{t,R}), \mathcal{L}_t^{(2)}(\hat{\gamma}_{t,R})\right\}^{T}{}_{t=R}$

# Objective 1: The framework – fixed rolling window case

- Authors consider the loss functions $L^{(1)}_{t+h}(\hat{\alpha}_{t,R})$ and $L^{(2)}_{t+h}(\hat{\gamma}_{t,R})$ to evaluate the OOS predictive ability of direct h-step ahead forecasts for models 1 and 2 made at time $t$
- As the rolling window estimation is performed for $t$ = R, R + 1,… T, we collect a series of OOS losses $\left\{L^{(1)}_{t+h}(\hat{\alpha}_{t,R}), L^{(2)}_{t+h}(\widehat{\gamma}_{t,R})\right\}^{T}_{t=R}$
- Example: unrestricted model $y_{t+h} = x'_t\alpha + \varepsilon_{t+h}$ under a quadratic loss:
  - the sequence of in-sample fitted errors at time $t$ is

    $\{\eta_{1,j}(\hat{\alpha}_{t,R})\}^{t}_{j=t-R+1} = \{y_j - x'_{j-h}\hat{\alpha}_{j,R}\}^{t}_{j=t-R+1}$ of which we use the last in-sample fitted error $\eta_{1,t}(\hat{\alpha}_{t,R}) = y_t - x'_{t-h}\hat{\alpha}_{t,R}$

    to evaluate the in-sample loss at time $t$: $\mathcal{L}^{(1)}_{t}(\hat{\alpha}_{t,R}) \equiv (y_t - x'_{t-h}\hat{\alpha}_{t,R})^2$
  - the OOS direct forecast loss at time $t$ is $L^{(1)}_{t+h}(\hat{\alpha}_{t,R}) \equiv (y_{t+h} - x'_{t-h}\hat{\alpha}_{t,R})^2$
- The loss function used for the estimation need not necessarily to be the same loss function used for forecast evaluation. In order to ensure a meaningful interpretation of the models' in-sample performance as a proxy for the OOS performance, authors require the two to be the same

# Objective 1: The framework – fixed rolling window case

- Authors consider the loss functions $L^{(1)}_{t+h}(\hat{\alpha}_{t,R})$ and $L^{(2)}_{t+h}(\hat{\gamma}_{t,R})$ to evaluate the OOS predictive ability of direct h-step ahead forecasts for models 1 and 2 made at time $t$
- As the rolling window estimation is performed for $t$ = R, R + 1,... T,
- we collect a series of OOS losses $\left\{L^{(1)}_{t+h}\left(\hat{\alpha}_{t,R}\right), L^{(2)}_{t+h}\left(\hat{\gamma}_{t,R}\right)\right\}^{T}_{t=R}$
- Example: unrestricted model $y_{t+h} = x_t'\alpha + \varepsilon_{t+h}$ under a quadratic loss:
  - the sequence of in-sample fitted errors at time $t$ is

    $\left\{\eta_{1,j}\left(\hat{\alpha}_{t,R}\right)\right\}^{t}_{j=t-R+1} = \left\{y_j - x'_{j-h}\hat{\alpha}_{j,R}\right\}^{t}_{j=t-R+1}$ of which we use the last in-sample fitted error $\eta_{1,t}\left(\hat{\alpha}_{t,R}\right) = y_t - x'_{t-h}\hat{\alpha}_{t,R}$

    to evaluate the in-sample loss at time $t$: $\mathcal{L}^{(1)}_{t}(\hat{\alpha}_{t,R}) \equiv (y_t - x'_{t-h}\hat{\alpha}_{t,R})^2$
  - the OOS direct forecast loss at time $t$ is $L^{(1)}_{t+h}(\hat{\alpha}_{t,R}) \equiv (y_{t+h} - x'_{t-h}\hat{\alpha}_{t,R})^2$
- The loss function used for the estimation need not necessarily to be the same loss function used for forecast evaluation. In order to ensure a meaningful interpretation of the models' in-sample performance as a proxy for the OOS performance, authors require the two to be the same



Paper: "Understanding Models' Forecasting Performance"

# Objective 1: The framework – fixed rolling window case

- Authors consider the loss functions $L^{(1)}_{t+h}(\hat{\alpha}_{t,R})$ and $L^{(2)}_{t+h}(\widehat{\gamma}_{t,R})$ to evaluate the OOS predictive ability of direct h-step ahead forecasts for models 1 and 2 made at time $t$
- As the rolling window estimation is performed for $t$ = R, R + 1,… T,
- we collect a series of OOS losses $\left\{L^{(1)}_{t+h}\left(\hat{\alpha}_{t,R}\right), L^{(2)}_{t+h}\left(\widehat{\gamma}_{t,R}\right)\right\}^{T}{}_{t=R}$
- Example: unrestricted model $y_{t+h} = x_t'\alpha + \varepsilon_{t+h}$ under a quadratic loss:
  - the sequence of in-sample fitted errors at time $t$ is

    $\left\{\eta_{1,j}\left(\hat{\alpha}_{t,R}\right)\right\}^{t}{}_{j=t-R+1} = \left\{y_j - x'_{j-h}\hat{\alpha}_{j,R}\right\}^{t}{}_{j=t-R+1}$ of which we use the last in-sample fitted error $\eta_{1,t}\left(\hat{\alpha}_{t,R}\right) = y_t - x'_{t-h}\hat{\alpha}_{t,R}$

    to evaluate the in-sample loss at time $t$: $\mathcal{L}^{(1)}_{t}(\hat{\alpha}_{t,R}) \equiv (y_t - x'_{t-h}\hat{\alpha}_{t,R})^2$
  - the OOS direct forecast loss at time $t$ is $L^{(1)}_{t+h}(\hat{\alpha}_{t,R}) \equiv (y_{t+h} - x'_{t-h}\hat{\alpha}_{t,R})^2$
- The loss function used for the estimation need not necessarily to be the same loss function used for forecast evaluation. In order to ensure a meaningful interpretation of the models' in-sample performance as a proxy for the OOS performance, authors require the two to be the same

# Objective 1: Sources of models' forecasting performance

The OOS forecasting performance of competing models can be attributed to: **model instability, predictive content**, and **over-fitting**

- Authors measure time variation in models' relative forecasting performance by averaging relative predictive ability over rolling windows of size *m*, where *m* < P satisfies assumption:
  $\lim_{T\to\infty} \frac{m}{P} \to \mu \in (0, \infty)\ as\ m, P \to \infty$

- They define predictive content as the correlation between the in-sample and OOS measures of fit
  - Small correlation → in-sample measure of fit have no predictive content for OOS and vice versa
  - Strong but negative correlation → in-sample predictive content is strong yet misleading for OOS

- Authors define over-fitting as a situation in which a model fits well in-sample but loses predictivity ability on OOS predictive content



Paper: "Understanding Models' Forecasting Performance"

# Objective 1: Sources of models' forecasting performance

Existing forecast comparison tests inform only about which model forecasts best and do not shed any light on why it is the case

- To capture **predictive content** and **over-fitting**, consider the regression:

$$\hat{L}_{t+h} = \beta\, \hat{\mathcal{L}}_t + u_{t+h} \text{ for } t = \text{R, R+1,..., T}$$

$\hat{\beta}\hat{\mathcal{L}}_t$ *and* $\hat{u}_{t+h}$ denote the corresponding fitted values and regression errors

Note: regression does NOT include a constant → $u_{t+h}$ measures the average OOS losses not explained by in-sample performance

- The Mean Squared Forecast Error (MSFE) can be decomposed as:

$$\frac{1}{P}\sum_{t=R}^{T} \hat{L}_{t+h} = B_P + U_P,$$

where $B_P \equiv \hat{\beta}\left(\frac{1}{P}\sum_{t=R}^{T} \hat{\mathcal{L}}_t\right)$ and $U_P = \frac{1}{P}\sum_{t=R}^{T} \hat{u}_{t+h}$

  - $B_P$ → component that was predictable on the basis of the in-sample relative fit of the model (**predictive content**)
  - $U_P$ → component that unexpected (**over-fitting**)



Paper: "Understanding Models' Forecasting Performance"

# Objective 1: Sources of models' forecasting performance

Existing forecast comparison tests inform only about which model forecasts best and do not shed any light on why it is the case

- To capture **predictive content** and **over-fitting**, consider the regression:

$$\hat{L}_{t+h} = \beta\,\hat{\mathcal{L}}_t + u_{t+h} \text{ for } t = \text{R, R+1,..., T}$$

$\hat{\beta}\hat{\mathcal{L}}_t$ *and* $\hat{u}_{t+h}$ denote the corresponding fitted values and regression errors

Note: regression does NOT include a constant → $u_{t+h}$ measures the average OOS losses not explained by in-sample performance

- The average Mean Squared Forecast Error (MSFE) can be decomposed as:

$$\frac{1}{P}\sum_{t=R}^{T}\hat{L}_{t+h} = B_P + U_P,$$

where $B_P \equiv \hat{\beta}\left(\frac{1}{P}\sum_{t=R}^{T}\hat{\mathcal{L}}_t\right)$ and $U_P \equiv \frac{1}{P}\sum_{t=R}^{T}\hat{u}_{t+h}$

  - $B_P$ → component that was predictable on the basis of the in-sample relative fit of the model (**predictive content**)
  - $U_P$ → component that was unexpected (**over-fitting**)

## Objective 1: Sources of models' forecasting performance

- Example: let data generating process $y_{t+h} = \alpha + \varepsilon_{t+h}$, $\varepsilon_{t+h} \sim iidN(0, \sigma^2)$
- We compare the forecasts of $y_{t+h}$ from two nested models' made at time $t$ based on parameter estimates obtained via the fixed rolling window approach. The first (unrestricted) models includes a constant only; the second (restricted) model sets the constant = 0
- Consider the (quadratic) forecast error loss difference

$$\hat{L}_{t+h} \equiv L_{t+h}^{(1)}\left(\hat{\alpha}_{t,R}\right) - L_{t+h}^{(2)}(0) \equiv (y_{t+h} - \hat{\alpha}_{t,R})^2 - y_{t+h}^2 \text{ and}$$

the in-sample loss difference $\hat{\mathcal{L}}_t \equiv \mathcal{L}_t^{(1)}\left(\hat{\alpha}_{t,R}\right) - \mathcal{L}_t^{(2)}(0) \equiv (y_t - \hat{\alpha}_{t,R})^2 - y_t^2$

- Let $\beta \equiv E(\hat{L}_{t+h}\,\hat{\mathcal{L}}_t)/E(\hat{\mathcal{L}}_t^2)$, it can be shown that

$$\beta = \frac{(\alpha^4 - \frac{3\sigma^2}{R^2})}{(\alpha^4 + 4\sigma^2\alpha^2 + \frac{(4\sigma^2 + 2\sigma^2\alpha^2)}{R}}$$

Note→ the numerator has 2 distinct component:

$\alpha^4 \rightarrow$ outcome of mis-specification in model 2

$\frac{3\sigma^2}{R^2} \rightarrow$ changes with sample size and captures estimation uncertainty in model 1



Paper: "Understanding Models' Forecasting Performance"

# Objective 1: Sources of models' forecasting performance

- Example: let data generating process $y_{t+h} = \alpha + \varepsilon_{t+h}, \varepsilon_{t+h} \sim iidN(0, \sigma^2)$
- We compare the forecasts of $y_{t+h}$ from two nested models' made at time $t$ based on parameter estimates obtained via the fixed rolling window approach. The first (unrestricted) models includes a constant only; the second (restricted) model sets the constant = 0
- Consider the (quadratic) forecast error loss difference
  - $\hat{L}_{t+h} \equiv L_{t+h}^{(1)}(\hat{\alpha}_{t,R}) - L_{t+h}^{(2)}(0) \equiv (y_{t+h} - \hat{\alpha}_{t,R})^2 - y_{t+h}^2$ and
  - the in-sample loss difference $\hat{\mathcal{L}}_t \equiv \mathcal{L}_t^{(1)}(\hat{\alpha}_{t,R}) - \mathcal{L}_t^{(2)}(0) \equiv (y_t - \hat{\alpha}_{t,R})^2 - y_t^2$
- Let $\beta \equiv E(\hat{L}_{t+h}\, \hat{\mathcal{L}}_t)/E(\hat{\mathcal{L}}_t^2)$, it can be shown that

$$\beta = \frac{(\alpha^4 - \frac{3\sigma^2}{R^2})}{(\alpha^4 + 4\sigma^2\alpha^2 + \frac{(4\sigma^2 + 2\sigma^2\alpha^2)}{R}}$$

Note→ the numerator has 2 distinct component:

$\alpha^4 \rightarrow$ outcome of mis-specification in model 2

$\frac{3\sigma^2}{R^2} \rightarrow$ changes with sample size and captures estimation uncertainty in model 1



Paper: "Understanding Models' Forecasting Performance"

# Objective 1: Sources of models' forecasting performance

- Given that: $\hat{L}_{t+h} = \beta\, \hat{\mathcal{L}}_t + u_{t+h}$ for $t$ = R, R+1,…, T and that average Mean Squared Forecast Error (MSFE) can be decomposed as: $\frac{1}{P}\sum_{t=R}^{T} \hat{L}_{t+h} = B_P + U_P$ ($B_P$: **predictive content**; $U_P$: **over-fitting**),
  - when $\beta$=0 → in-sample loss difference have no predictive content for the OOS
  - when mis-specification component dominates ($\alpha^4 > 3\sigma^2/R^2$ so $\beta > 0$) → in-sample loss difference provide information content for the OOS
  - when $\beta < 0$ → though in-sample fit has predictive content for OOS, it is misleading (it is driven primary by estimation uncertainty)
- For any given $\beta$, $E(B_P) = \beta E(\hat{\mathcal{L}}_t)$ and by construction
- $E(U_P) = E\left(\hat{L}_{t+h}\right) - E(B_P) = E\left(\hat{L}_{t+h}\right) - \beta E(\hat{\mathcal{L}}_t)$
- This means that also $U_P$(ie. the component designated to measure over-fitting) is affected by $\beta$, so by both mis-specification component and estimation uncertainty



Paper: "Understanding Models' Forecasting Performance"

# Objective 1: Sources of models' forecasting performance

- Given that: $\hat{L}_{t+h} = \beta\, \hat{\mathcal{L}}_t + u_{t+h}$ for $t$ = R, R+1,..., T and that Mean Squared Forecast Error (MSFE) can be decomposed as:
  $\frac{1}{P}\sum_{t=R}^{T} \hat{L}_{t+h} = B_P + U_P$ ($B_P$: **predictive content**; $U_P$:**over-fitting**),
    - when $\beta$=0 → in-sample loss difference have no predictive content for the OOS
    - when mis-specification component dominates ($\alpha^4 > 3\sigma^2/R^2$ so $\beta > 0$) → in-sample loss difference provide information content for the OOS
    - when $\beta < 0$ → though in-sample fit has predictive content for OOS, it is misleading (it is driven primary by estimation uncertainty)
- For any given $\beta$, $E(B_P) = \beta E(\hat{\mathcal{L}}_t)$ and by construction
  $E(U_P) = E\left(\hat{L}_{t+h}\right) - E(B_P) = E\left(\hat{L}_{t+h}\right) - \beta E(\hat{\mathcal{L}}_t)$
  This means that also $U_P$(ie. the component designated to measure over-fitting) is affected by $\beta$, so by both mis-specification component and estimation uncertainty

# Objective 1: Sources of models' forecasting performance

- What about **model instability?** Decompose the OOS loss function differences $\{\hat{L}_{t+h}\}^{T}{}_{t=R}$ into 3 components: not only $B_P$(**predictive content** and $U_P$(**over-fitting**) but also $A_{\tau,P}$
- $A_{\tau,P}$ reflects the extent of instabilities in the relative forecasting performance. It measures the presence of time variation in the expected relative forecasting performance
  - $A_{\tau,P} = 0$ → no time variation in expected relative forecasting performance
  - Sign of $A_{\tau,P}$ changes → the OOS predictive ability swings from favoring one model to favoring the other model
- Recall that: $B_P$ measures models' OOS relative forecasting ability reflected in the in-sample relative performance and $U_P$ measures models' OOS relative forecasting ability not reflected by in-sample fit
- Under some assumption, authors prove that $A_{\tau,P}$, $B_P$ and $U_P$ are asymptotically uncorrelated

# Objective 1: Sources of models' forecasting performance

- What about **model instability?** Decompose the OOS loss function differences $\{\hat{L}_{t+h}\}^{T}_{t=R}$ into 3 components: not only $B_P$(**predictive content** and $U_P$(**over-fitting**) but also $A_{\tau,P}$
- $A_{\tau,P}$ reflects the extent of instabilities in the relative forecasting performance. It measures the presence of time variation in the expected relative forecasting performance
  - $A_{\tau,P} = 0 \rightarrow$ no time variation in expected relative forecasting performance
  - Sign of $A_{\tau,P}$ changes, the OOS predictive ability swings from favoring one model to favoring the other model
- Recall that: $B_P$ measures models' OOS relative forecasting ability reflected in the in-sample relative performance and $U_P$ measures models' OOS relative forecasting ability not reflected by in-sample fit
- Under some assumption, authors prove that $A_{\tau,P}$, $B_P$and $U_P$ are asymptotically uncorrelated

# Objective 2: Exch. rate fluctuation and fundamentals

Older papers have shown that the RW generates the best exchange rate forecast and authors want to understand *why* the economic models' performance is poor

- Economic model: forecasting relationship btw exchange rates and economic fundamentals in a multivariate regression: $y_t = \alpha x_t + \epsilon_{1,t}$
  - LHS → growth rate of exchange rate $y_t = \ln(\frac{S_t}{S_{t-1}})$
  - RHS → growth rate differentials of the country-specific variables relative to the US counterparts: money supply, industrial production index, unemployment rate, lagged interest rate and growth rate of oil prices. In addition, authors separately consider the predictive ability of the commodity price index growth rate for the Canadian exchange rate
- The bmk is a simple RW: $y_t = \epsilon_{2,t}$ , $\epsilon_{2,t}$ is the error term
- Dataset: monthly data from Sept. 1975 to Sept. 2008 for some industrialized countries: Switzerland, UK, Canada, Japan, Germany
- Total sample size is T=396. Authors use rolling estimation window of size R (R=40,...,196), estimation window size varies → OOS forecast period P(R) is not constant



Paper: "Understanding Models' Forecasting Performance"

# Objective 2: Exch. rate fluctuation and fundamentals

If the MSFE ratio ($MSFE^{Model}/MSFE^{BMK}$) is >1, the economic model is performing worse than the bmk on average

- Considering one-step ahead forecasts for both economic model and bmk
  - Figures represent the average OOS forecasting performance of the economic model relative to the bmk. Figure 1 considers first P = 200 OOS periods following the estimation period; figure 2 considers last P = 200 OOS periods which are overlapping for the estimation windows of different size

Figure 1: Out-of-Sample Fit in Data: $MSFE^{Model}/MSE^{RW}$, P = 200 - non-overlapping

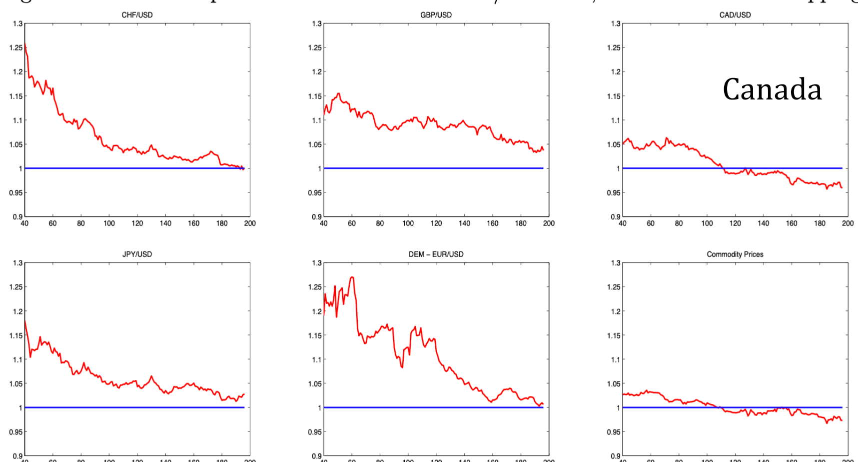
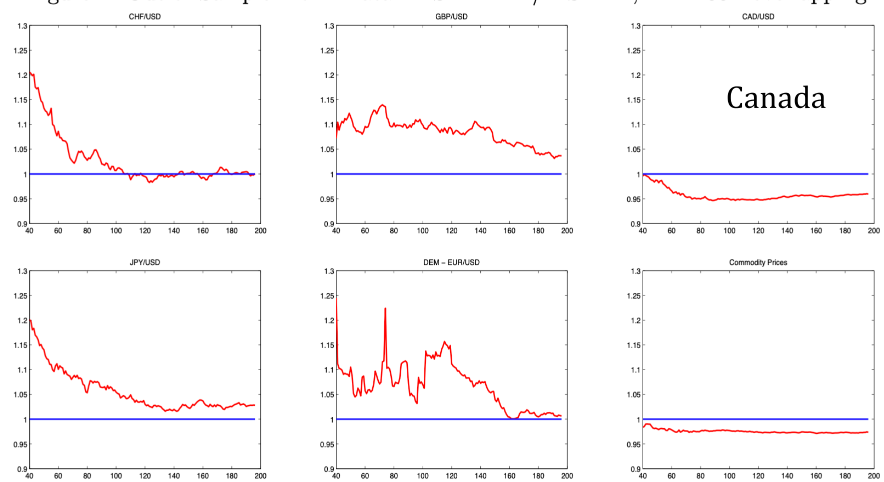



Figure 2: Out-of-Sample Fit in Data: $MSFE^{Model}/MSE^{RW}$, P = 200 - overlapping




- Y-axis→ ratio of MSFE; X-axis→ R (estimation window size)
- Figures show that the forecasting performance of the economic model is inferior to that of the RW for all countries, expect Canada
- As R↑ → forecasting ability of the model improves



Paper: "Understanding Models' Forecasting Performance"

# Objective 2: Exch. rate fluctuation and fundamentals

If the MSFE ratio ($MSFE^{Model}/MSFE^{BMK}$) is >1, the economic model is performing worse than the bmk on average

- Considering one-step ahead forecasts for both economic model and bmk
  - Figures represent the average OOS forecasting performance of the economic model relative to the bmk. Figure 1 considers first P = 200 OOS periods following the estimation period; figure 2 considers last P = 200 OOS periods which are overlapping for the estimation windows of different size

Figure 1: Out-of-Sample Fit in Data: $MSFE^{Model}/MSE^{RW}$, P = 200 - non-overlapping

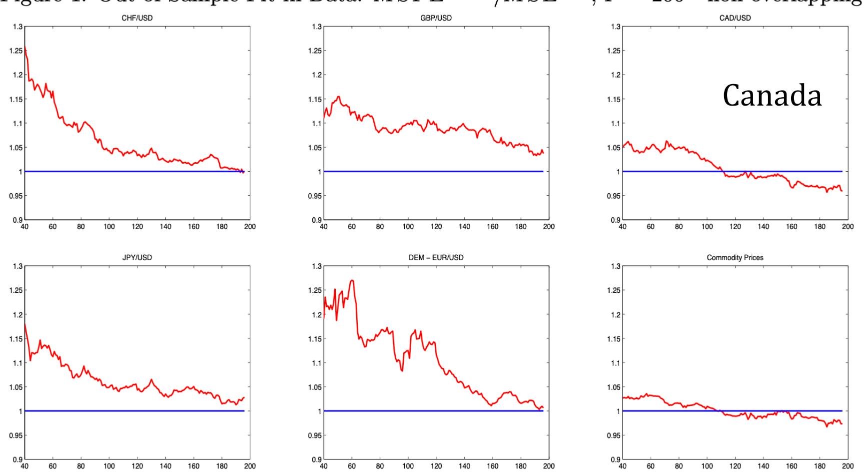
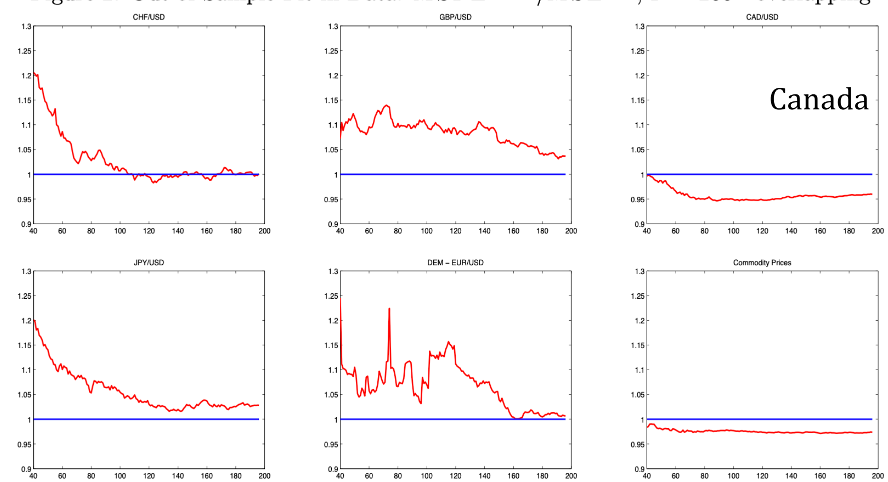



Figure 2: Out-of-Sample Fit in Data: $MSFE^{Model}/MSE^{RW}$, P = 200 - overlapping




- Y-axis→ ratio of MSFE; X-axis→ R (estimation window size)
- Figures show that the forecasting performance of the economic model is inferior to that of the RW for all countries, expect Canada
- As R↑ → forecasting ability of the model improves



Paper: "Understanding Models' Forecasting Performance"

# Objective 2: Exch. rate fluctuation and fundamentals

As the estimation window size (R) increases, the forecasting performance of model improves. The degree of improvement deteriorates when the models are compared over the same OOS period (figure 2)

- This can be seen considering the case of Japan

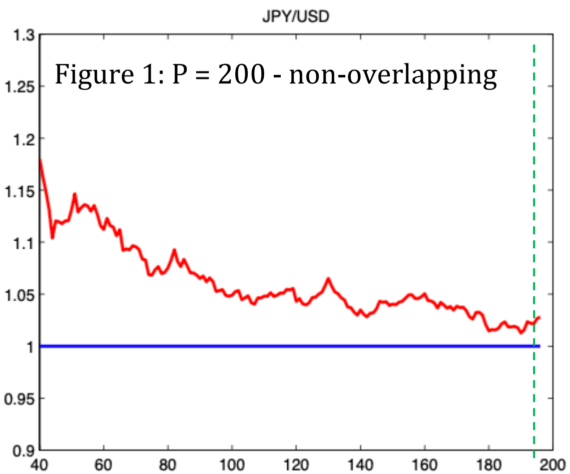


Figure 1: P = 200 - non-overlapping

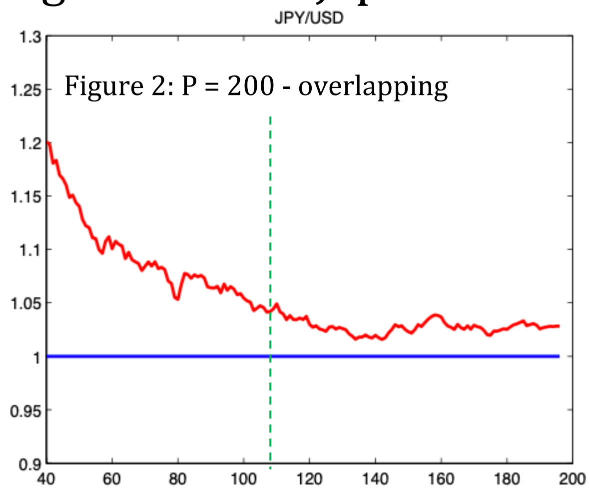


Figure 2: P = 200 - overlapping

Model's average OOS forecasting performance is similar to that of the RW starting at R=110 when compared over the same OOS forecast periods (figure 2), while otherwise (figure 1) its performance becomes similar only for R=200

- Conclusion:
  - Figure 1 suggest that poor performance of economic models is mostly due to **over-fitting**, as opposed to **parameter instability**
  - Figure 2 uncovers that the choice of the window is not crucial, so that **over-fitting** is a concern only when the window is too small

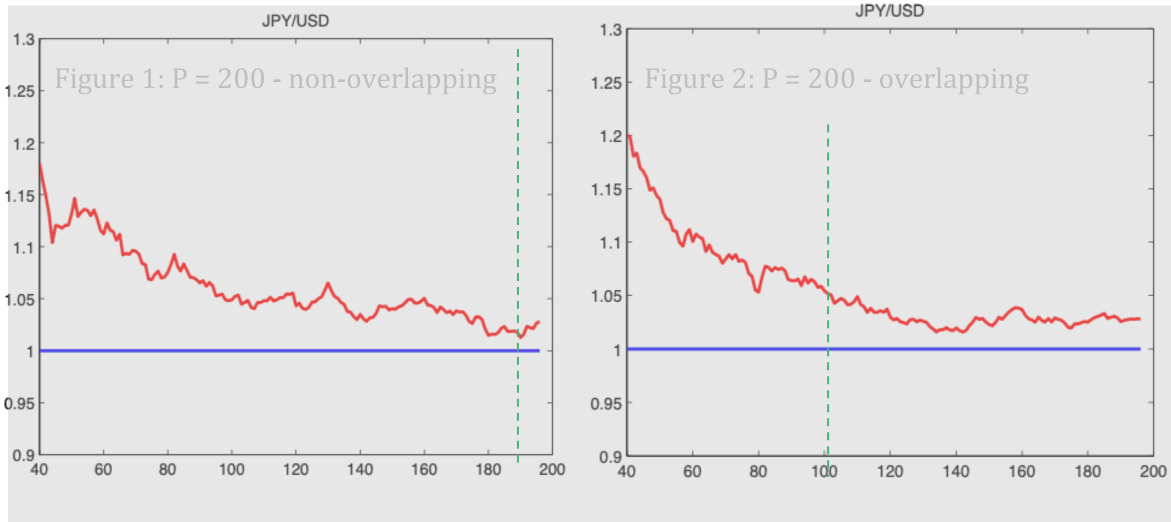


Paper: "Understanding Models' Forecasting Performance"

# Objective 2: Exch. rate fluctuation and fundamentals

As the estimation window size increases, the forecasting performance of model improves. The degree of improvement deteriorates when the models are compared over the same OOS period (figure 2)

- This can be seen considering the case of Japan




Model's average OOS forecasting performance is similar to that of the RW starting at R=110 when compared over the same OOS forecast periods (figure 2), while otherwise (figure 1) its performance becomes similar only for R=200

- Conclusion:
  - Figure 1 suggest that poor performance of economic models is mostly due to **over-fitting**, as opposed to **parameter instability**
  - Figure 2 uncovers that the choice of the window is not crucial, so that **over-fitting** is a concern only when the window is too small



Paper: "Understanding Models' Forecasting Performance"

# Objective 2: Exch. rate fluctuation and fundamentals

Decompose the difference between the MSFEs of the economic model and bmk into components $A_{\tau,P}$, $B_P$ and $U_P$. Negative MSFE differences implies that the economic model is better than bmk model

Figure 3: Decomposition for One-step Ahead Forecast.

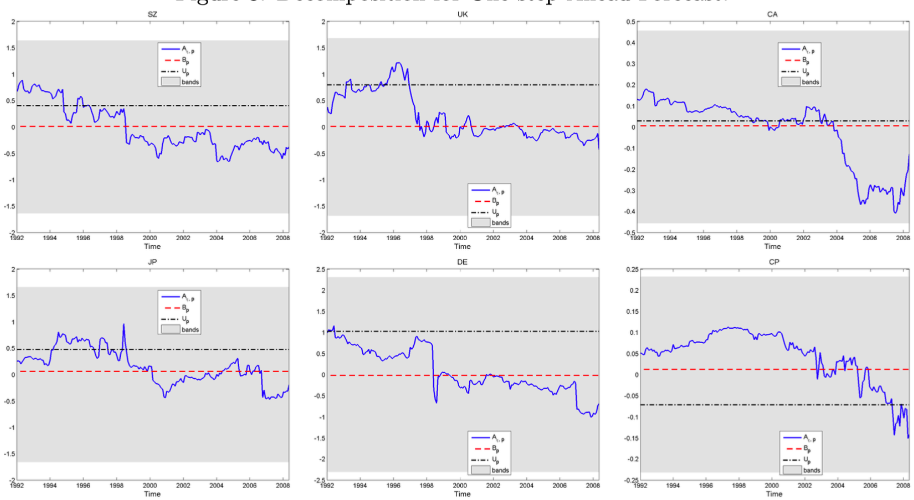
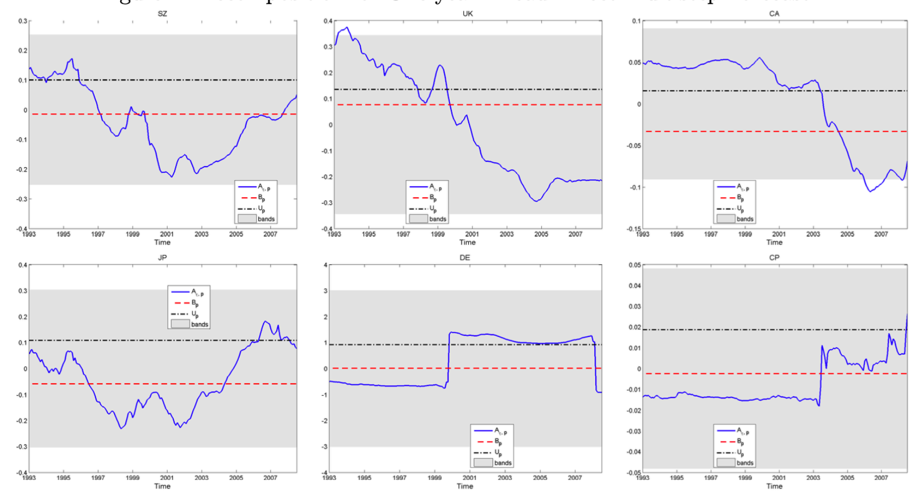



Figure 4: Decomposition for One-year Ahead Direct Multistep Forecast.




- Empirical evidence of time variation of $A_{\tau,P}$ (see blue solid line) → possible instability in the relative forecasting performance of models
- **For 1-step ahead forecast** (1m) → $B_P$ **(predictive content)** is mostly positive (see red dotted line) → lack of OOS predictive ability is related to lack of in-sample predictive content
- **For 1-y ahead forecast** → evidence of predictive content is weaker → **time variation** is mainly responsible for the lack of forecasting ability



Paper: "Understanding Models' Forecasting Performance"

# Objective 2: Exch. rate fluctuation and fundamentals

Decompose the difference between the MSFEs of the economic model and bmk into components $A_{\tau,P}$, $B_P$ and $U_P$. Negative MSFE differences implies that the economic model is better than bmk model

Figure 3: Decomposition for One-step Ahead Forecast.

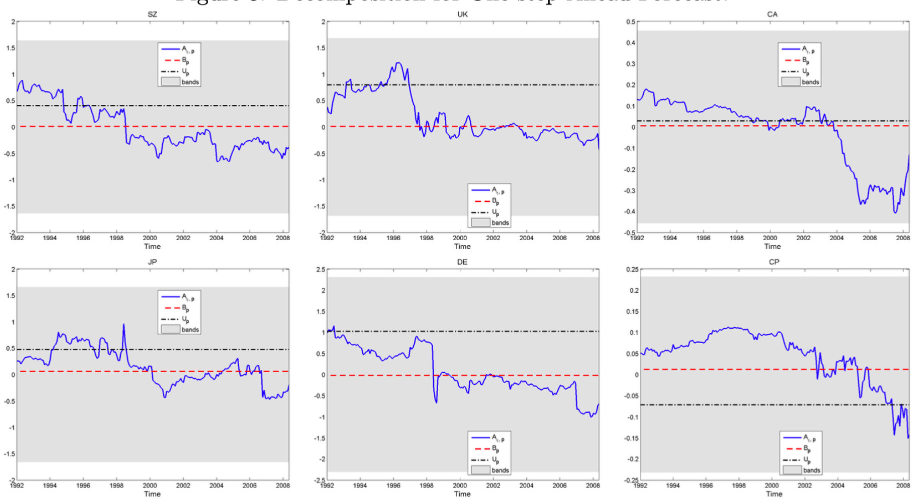
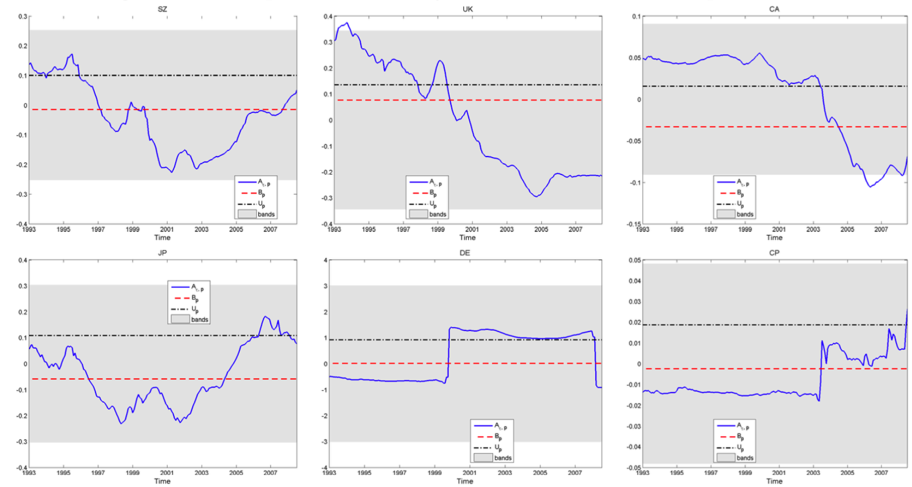



Figure 4: Decomposition for One-year Ahead Direct Multistep Forecast.




- Empirical evidence of time variation of $A_{\tau,P}$ (see blue solid line) → possible instability in the relative forecasting performance of models
- **For 1-step ahead forecast** (1m) → $B_P$ **(predictive content)** is mostly positive (see red dotted line) → lack of OOS predictive ability is related to lack of in-sample predictive content
- **For 1-y ahead forecast** → evidence of predictive content is weaker → **time variation** is mainly responsible for the lack of forecasting ability



Paper: “Understanding Models’ Forecasting Performance”

# Conclusion

The aim was to decompose the sources of the OOS forecasting performance into uncorrelated components that have meaningful economic interpretation and might provide a constructive insights to improve models' forecasts.

The OOS forecasting ability of competing models can be attributed to:

**predictive content,**

**over-fitting,**

**model instability**

THANK YOU FOR YOUR ATTENTION