Applied Time Series
Model building with multiple goals

Kamil Kovar

CERGE-EI - Summer Semester 2019

Outline

Introduction: Automatic model selection

Outline

Introduction

In practice we are often faced with problem of choosing final model from large number of candidate models.

Introduction

In practice we are often faced with problem of choosing final model from large number of candidate models.

Common solution is to use automatic selection algorithms.

Best subset selection

Consider that you have variable y and set of potential regressors $X = \{x_, , x_2, \ldots, x_k\}$.

Best subset selection

Consider that you have variable y and set of potential regressors $X = \{x_,, x_2, \ldots, x_k\}$.

Best subset selection algorithm will estimate all the possible models and choose the best model.

- Best model: information criteria (C_p, AIC, BIC) or cross-validation performance.

Best subset selection

Consider that you have variable y and set of potential regressors $X = \{x_,, x_2, \ldots, x_k\}$.

Best subset selection algorithm will estimate all the possible models and choose the best model.

- Best model: information criteria (C_p, AIC, BIC) or cross-validation performance.

Number of models grows very fast $\Rightarrow$ limits on number of regressors often imposed.

Stepwise selection

Start from initial model and add/remove one regressor at time, yielding set of potential models; choose best model as starting point for next iteration.

- Forward selection: start with no regressors and add one regressor at a time.
- Backward selection: start with all regerssors and remove one regressor at a time.

Limits of automatic model selection

The various methods of model selection have their limits.

Limits of automatic model selection

The various methods of model selection have their limits.

- They focus on single measure summarizing overall forecasting performance, but single measure can mask important heterogeneity.

Limits of automatic model selection

The various methods of model selection have their limits.

- They focus on single measure summarizing overall forecasting performance, but single measure can mask important heterogeneity.
- Sensitivity to shocks is not among selection factors.

Limits of automatic model selection

The various methods of model selection have their limits.

- They focus on single measure summarizing overall forecasting performance, but single measure can mask important heterogeneity.
- Sensitivity to shocks is not among selection factors.
- Important modifications of the model structure might be unexplored due to a-priori limits on model structures (e.g. using spreads, ratios etc.).

Limits of automatic model selection

The various methods of model selection have their limits.

- They focus on single measure summarizing overall forecasting performance, but single measure can mask important heterogeneity.
- Sensitivity to shocks is not among selection factors.
- Important modifications of the model structure might be unexplored due to a-priori limits on model structures (e.g. using spreads, ratios etc.).
- Model selection does not reflect theoretical (long-run) restrictions, which can be often violated.

Limits of automatic model selection

The various methods of model selection have their limits.

- They focus on single measure summarizing overall forecasting performance, but single measure can mask important heterogeneity.
- Sensitivity to shocks is not among selection factors.
- Important modifications of the model structure might be unexplored due to a-priori limits on model structures (e.g. using spreads, ratios etc.).
- Model selection does not reflect theoretical (long-run) restrictions, which can be often violated.
- The resulting model structure cannot be explained beyond referring to its optimality with respect to selection criteria.

Model building

Outline

1. Introduction: Automatic model selection
2. Model building
   - Estimation output
   - Forecast performance
   - Sensitivity to shocks
   - Tips&tricks
3. Applications

Introduction

Model building is process of trial and error that leads to improvement of model based on understanding of sources of previous models' shortcomings.

Introduction

Model building is process of trial and error that leads to improvement of model based on understanding of sources of previous models' shortcomings.

In process of model building the builder should learn to understand the nature of the data and hence be able to explain the choice of final model.

Estimation output

Introduction

The starting point of investigating model is checking the estimation output.

Introduction

The starting point of investigating model is checking the estimation output.

There are two categories of information we are interested in estimation output:

1. Coefficient estimates.
2. Regression statistics.

Coefficient sign

Theory often provides clear guidance on the signs of coefficients $\Rightarrow$ model that does not satisfy sign restrictions can be ruled out.

Coefficient sign

Theory often provides clear guidance on the signs of coefficients $\Rightarrow$ model that does not satisfy sign restrictions can be ruled out.

Why would regressor have wrong coefficient sign?

Coefficient sign

Theory often provides clear guidance on the signs of coefficients $\Rightarrow$ model that does not satisfy sign restrictions can be ruled out.

Why would regressor have wrong coefficient sign?

- Timing: sign can change with lags of regressors.

Coefficient sign

Theory often provides clear guidance on the signs of coefficients $\Rightarrow$ model that does not satisfy sign restrictions can be ruled out.

Why would regressor have wrong coefficient sign?

- Timing: sign can change with lags of regressors.
- Omitted variables: coefficients can be biased due to regressor being correlated with omitted factor.

Coefficient sign

Theory often provides clear guidance on the signs of coefficients $\Rightarrow$ model that does not satisfy sign restrictions can be ruled out.

Why would regressor have wrong coefficient sign?

- Timing: sign can change with lags of regressors.
- Omitted variables: coefficients can be biased due to regressor being correlated with omitted factor.

Importantly, we are interested in sign of effects: if multiple lags are used then wrong coefficient signs on some lags might not be an issue.

Coefficient size

Coefficient size is the effect of change in regressor by one unit on the dependent variable.

Coefficient size

Coefficient size is the effect of change in regressor by one unit on the dependent variable.

We would like to have a measure of relative importance of regressors.

Coefficient size

Coefficient size is the effect of change in regressor by one unit on the dependent variable.

We would like to have a measure of relative importance of regressors.

Key problem: Coefficient size depends on units of regressor so they are not directly comparable.

Coefficient size

Coefficient size is the effect of change in regressor by one unit on the dependent variable.

We would like to have a measure of relative importance of regressors.

Key problem: Coefficient size depends on units of regressor so they are not directly comparable.

Solution: Use **standardized** coefficients.

Coefficient size

Coefficient size is the effect of change in regressor by one unit on the dependent variable.

We would like to have a measure of relative importance of regressors.

Key problem: Coefficient size depends on units of regressor so they are not directly comparable.

Solution: Use **standardized** coefficients.

- Standardized coefficients are normal coefficients multiplied by relative standard deviation of regressor and dependent variable: $\beta^* = \beta \frac{\sigma_y}{\sigma_x}$

Coefficient size

Coefficient size is the effect of change in regressor by one unit on the dependent variable.

We would like to have a measure of relative importance of regressors.

Key problem: Coefficient size depends on units of regressor so they are not directly comparable.

Solution: Use **standardized** coefficients.

- Standardized coefficients are normal coefficients multiplied by relative standard deviation of regressor and dependent variable: $\beta^* = \beta\frac{\sigma_y}{\sigma_x}$
- Interpretation: By how many standard deviations does y increase if we increase x by one standard deviation.

Coefficient size

Coefficient size is the effect of change in regressor by one unit on the dependent variable.

We would like to have a measure of relative importance of regressors.

Key problem: Coefficient size depends on units of regressor so they are not directly comparable.

Solution: Use **standardized** coefficients.

- Standardized coefficients are normal coefficients multiplied by relative standard deviation of regressor and dependent variable: $\beta^* = \beta \frac{\sigma_y}{\sigma_x}$
- Interpretation: By how many standard deviations does y increase if we increase x by one standard deviation.

  $\Rightarrow$ **Replacing natural units with statistical units comparable across variables.**

Coefficient significance

T-statistic and p-value are measures of coefficient significance: degree of confidence that regressor has non-zero effect on dependent variable.

Coefficient significance

T-statistic and p-value are measures of coefficient significance: degree of confidence that regressor has non-zero effect on dependent variable.

What does low significance mean?

Coefficient significance

T-statistic and p-value are measures of coefficient significance: degree of confidence that regressor has non-zero effect on dependent variable.

What does low significance mean?

- Small coefficient/influence of regressor on dependent variable.

Coefficient significance

T-statistic and p-value are measures of coefficient significance: degree of confidence that regressor has non-zero effect on dependent variable.

What does low significance mean?

- Small coefficient/influence of regressor on dependent variable.
- Large uncertainty about coefficient value.

Coefficient significance

T-statistic and p-value are measures of coefficient significance: degree of confidence that regressor has non-zero effect on dependent variable.

What does low significance mean?

- Small coefficient/influence of regressor on dependent variable.
- Large uncertainty about coefficient value.

Crucially, significance is imperfect measure of influence of regressor: even regressor with large influence can have insignificant coefficient.

Regression statistics

Main measure of model fit is R-squared: share of variance in dependent variable explained by the model.

Regression statistics

Main measure of model fit is R-squared: share of variance in dependent variable explained by the model.

Standard error of regression is the standard deviation of the residuals of the regression: it is measure of how far on average are observations from the predicted values.

- It is measured in units of dependent variable.

Regression statistics

Main measure of model fit is R-squared: share of variance in dependent variable explained by the model.

Standard error of regression is the standard deviation of the residuals of the regression: it is measure of how far on average are observations from the predicted values.

- It is measured in units of dependent variable.

Durbin-Watson statistic is measure of autocorrelation in residuals.

- Autocorrelation of residuals suggests omitted autocorrelated factors $\Rightarrow$ possible indication of model misspecification.

Illustration

Tip: Color coding estimation output can speed up understanding.

Dependent Variable: FRMP_I_IEUZN-FRMP_T_IEUZN
Method: Least Squares
Date: 09/26/18 Time: 17:37
Sample (adjusted): 2004Q4 2017Q3
Included observations: 52 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.	Standardized coefficients
FRMP_I_IEUZN(-1)-FRMP_T_IEUZN(-1)	0.926804	0.039861	23.25069	0.0000	0.920298
D(FRFED_US(-1)-FRMP_I_IEUZN(-1))	0.001570	0.183569	0.008550	0.9932	0.000620
D(DUM_FINCRISIS_IEUZN)	-0.255427	0.312935	-0.816229	0.4185	-0.051709
DLOG(FSPI_IEUZN)-@MOVAV(DLOG(FSPI_IEUZN),4)	2.633736	1.044464	2.521615	0.0151	0.190296
DLOG(FTFXIUSAQ_IEUZN)*DUM_FINCRISIS_IEUZN	-3.013924	2.876196	-1.047885	0.3001	-0.067799

R-squared	0.915809	Mean dependent var	-0.576243
Adjusted R-squared	0.908644	S.D. dependent var	1.429578
S.E. of regression	0.432093	Akaike info criterion	1.250859
Sum squared resid	8.775098	Schwarz criterion	1.438479
Log likelihood	-27.52233	Hannan-Quinn criter.	1.322788
Durbin-Watson stat	1.603717		

Forecast performance

Introduction

Main purpose of models is forecasting $\Rightarrow$ forecasting performance is main factor in choosing model.

Introduction

Main purpose of models is forecasting $\Rightarrow$ forecasting performance is main factor in choosing model.

Forecasting performance is useful for three reasons:

Introduction

Main purpose of models is forecasting $\Rightarrow$ forecasting performance is main factor in choosing model.

Forecasting performance is useful for three reasons:

1. Instrumental: Good historical forecasting performance is indication of good future forecasting performance.

Introduction

Main purpose of models is forecasting $\Rightarrow$ forecasting performance is main factor in choosing model.

Forecasting performance is useful for three reasons:

1. Instrumental: Good historical forecasting performance is indication of good future forecasting performance.
2. Decision-making: More "correct" models are likely to have better forecasting performance $\Rightarrow$ relative forecasting performance is guide during model building.

Introduction

Main purpose of models is forecasting ⇒ forecasting performance is main factor in choosing model.

Forecasting performance is useful for three reasons:

1. Instrumental: Good historical forecasting performance is indication of good future forecasting performance.
2. Decision-making: More "correct" models are likely to have better forecasting performance ⇒ relative forecasting performance is guide during model building.
3. Indicative: Bad forecasting performance during particular period can suggest factors missing from model.

Introduction

Main purpose of models is forecasting ⇒ forecasting performance is main factor in choosing model.

Forecasting performance is useful for three reasons:

1. Instrumental: Good historical forecasting performance is indication of good future forecasting performance.
2. Decision-making: More "correct" models are likely to have better forecasting performance ⇒ relative forecasting performance is guide during model building.
3. Indicative: Bad forecasting performance during particular period can suggest factors missing from model.

The process of evaluating forecast from model on historical data is called **backtesting**.

- After model has been developed we can test it in true out-of-sample testing as new data are released.

One-step ahead forecasts

Introduction

Starting point of analyzing forecasting performance are one-step ahead forecasts.

Introduction

Starting point of analyzing forecasting performance are one-step ahead forecasts.

Summary of forecast performance in terms of one-step-ahead forecasts is readily available as output of any regression.

Summary statistics

The main regression statistics can be interpreted as measures of one-step-ahead forecasting.

Summary statistics

The main regression statistics can be interpreted as measures of one-step-ahead forecasting.

- R-squared is share of explained variance $\Rightarrow$ high R-squared suggests good one-step-ahead forecasting performance.

Summary statistics

The main regression statistics can be interpreted as measures of one-step-ahead forecasting.

- R-squared is share of explained variance $\Rightarrow$ high R-squared suggests good one-step-ahead forecasting performance.
- Standard error of regression is RMSE of residuals aka one-step-ahead forecasts $\Rightarrow$ average forecast mistake in units of dependent variable.

Summary statistics: Drawbacks I

R-squared measure of explained variance in dependent variable $\Rightarrow$ it is sensitive to transformation of dependent variable.

- Levels are typically more volatile and easier to forecast $\Rightarrow$ equations in levels have higher R-squared than equations in differences.

Summary statistics: Drawbacks I

R-squared measure of explained variance in dependent variable $\Rightarrow$ it is sensitive to transformation of dependent variable.

- Levels are typically more volatile and easier to forecast $\Rightarrow$ equations in levels have higher R-squared than equations in differences.
- Models with identical forecasts for level of dependent variable can have very different R-squared $\Leftrightarrow$ lower R-squared does not necessarily correspond to worse forecasting performance.

Summary statistics: Drawbacks I

R-squared measure of explained variance in dependent variable ⇒ it is sensitive to transformation of dependent variable.

- Levels are typically more volatile and easier to forecast ⇒ equations in levels have higher R-squared than equations in differences.
- Models with identical forecasts for level of dependent variable can have very different R-squared ⇔ lower R-squared does not necessarily correspond to worse forecasting performance.
- Log-likelihood, SSR and standard error of regression share first but not the second problem.

Dependent Variable: LBR_US
Method: Least Squares
Date: 06/10/19 Time: 10:02
Sample (adjusted): 1948M02 2019M03
Included observations: 854 after adjustments
LBR_US = LBR_US(-1) +C(1)

	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	0.000468	0.007144	0.065562	0.9477

R-squared	0.983762	Mean dependent var	5.759133
Adjusted R-squared	0.983762	S.D. dependent var	1.638360
S.E. of regression	0.208775	Akaike info criterion	-0.293946
Sum squared resid	37.17981	Schwarz criterion	-0.288384
Log likelihood	126.5151	Hannan-Quinn criter.	-0.291816
Durbin-Watson stat	1.761709		

Dependent Variable: D(LBR_US)
Method: Least Squares
Date: 06/10/19 Time: 10:02
Sample (adjusted): 1948M02 2019M03
Included observations: 854 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.000468	0.007144	0.065562	0.9477

R-squared	0.000000	Mean dependent var	0.000468
Adjusted R-squared	0.000000	S.D. dependent var	0.208775
S.E. of regression	0.208775	Akaike info criterion	-0.293946
Sum squared resid	37.17981	Schwarz criterion	-0.288384
Log likelihood	126.5151	Hannan-Quinn criter.	-0.291816
Durbin-Watson stat	1.761709		

Summary statistics: Drawbacks II

R-squared is also very sensitive to inclusion of dynamics terms.

- When LDV or ARMA terms are included then forecasts are anchored by last observed value $\Rightarrow$ R-squared will be typically much higher with dynamic terms.

Summary statistics: Drawbacks II

R-squared is also very sensitive to inclusion of dynamics terms.

- When LDV or ARMA terms are included then forecasts are anchored by last observed value $\Rightarrow$ R-squared will be typically much higher with dynamic terms.
- The overall improved forecasting performance might mask worsening of forecast performance in stress periods.

Summary statistics: Drawbacks II

R-squared is also very sensitive to inclusion of dynamics terms.

- When LDV or ARMA terms are included then forecasts are anchored by last observed value $\Rightarrow$ R-squared will be typically much higher with dynamic terms.
- The overall improved forecasting performance might mask worsening of forecast performance in stress periods.
- This is shared by other summary statistics.

Fitted values and residuals

The forecasting performance can be readily visualized by graphing:

Fitted values and residuals

The forecasting performance can be readily visualized by graphing:

1. Actual vs. fitted values.

Fitted values and residuals

The forecasting performance can be readily visualized by graphing:

1. Actual vs. fitted values.
2. Residuals (aka one-step-ahead forecast errors).

Fitted values and residuals

The forecasting performance can be readily visualized by graphing:

1. Actual vs. fitted values.
2. Residuals (aka one-step-ahead forecast errors).

Not very informative in presence of dynamic terms.

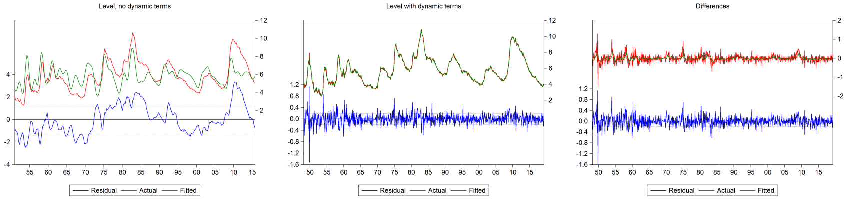
Fitted values and residuals

The forecasting performance can be readily visualized by graphing:

1. Actual vs. fitted values.
2. Residuals (aka one-step-ahead forecast errors).

Not very informative in presence of dynamic terms.

Note that the graphs are in units of dependent variable ⇒ limited value when levels are of interest but differences are modelled.



Adjusted measures of fit

Standard measures of fit are monotonically increasing in number of regressors $\Rightarrow$ they cannot be used for comparison of models.

Adjusted measures of fit

Standard measures of fit are monotonically increasing in number of regressors $\Rightarrow$ they cannot be used for comparison of models.

Adjusted measures of fit measure the fit of the model after accounting for number of regressors.

Adjusted measures of fit

Standard measures of fit are monotonically increasing in number of regressors ⇒ they cannot be used for comparison of models.

Adjusted measures of fit measure the fit of the model after accounting for number of regressors.

- Adjusted R-squared: $R^2_{adj} = 1 - (1 - R^2)\frac{T-1}{T-k}$

Adjusted measures of fit

Standard measures of fit are monotonically increasing in number of regressors ⇒ they cannot be used for comparison of models.

Adjusted measures of fit measure the fit of the model after accounting for number of regressors.

- Adjusted R-squared: $R^2_{adj} = 1 - (1 - R^2)\frac{T-1}{T-k}$
- Information criteria: AIC, BIC, HQ,...

Adjusted measures of fit

Standard measures of fit are monotonically increasing in number of regressors ⇒ they cannot be used for comparison of models.

Adjusted measures of fit measure the fit of the model after accounting for number of regressors.

- Adjusted R-squared: $R^2_{adj} = 1 - (1 - R^2)\frac{T-1}{T-k}$
- Information criteria: AIC, BIC, HQ,...
  - The criteria are adjusted log-likelihoods ⇒ they share limits of log-likelihood.

Adjusted measures of fit

Standard measures of fit are monotonically increasing in number of regressors ⇒ they cannot be used for comparison of models.

Adjusted measures of fit measure the fit of the model after accounting for number of regressors.

- Adjusted R-squared: $R^2_{adj} = 1 - (1 - R^2)\frac{T-1}{T-k}$
- Information criteria: AIC, BIC, HQ,...
  - The criteria are adjusted log-likelihoods ⇒ they share limits of log-likelihood.
  - The criteria are derived from optimality principles ⇒ preferable to adjusted R-squared.

Multi-period forecasts

Introduction

In most applications we are interested in forecasts for multiple periods $\Rightarrow$ one-step-ahead forecasts have limited informational value for practical purposes.

Introduction

In most applications we are interested in forecasts for multiple periods $\Rightarrow$ one-step-ahead forecasts have limited informational value for practical purposes.

Regressions are estimated with aim of minimizing one-step ahead forecasts $\Rightarrow$ multi-period forecasts are more informative about the quality of the model.

Recursive forecasts

Starting point is creating forecasts. We will focus on recursive forecasts.

Recursive forecasts

Starting point is creating forecasts. We will focus on recursive forecasts.

1. Use observations $1, 2, \ldots, t$ to estimate your model coefficients.

Recursive forecasts

Starting point is creating forecasts. We will focus on recursive forecasts.

1. Use observations $1, 2, \ldots, t$ to estimate your model coefficients.
2. Make forecast starting from $t + 1$ and ending in $t + h$.
   - Use actual values for dependent variable as initial conditions.
   - Use period t errors as initial conditions for ARMA terms.
   - Use forecast values for dependent variable for future periods (dynamic forecast).
   - Use actual values for independent variables (conditional forecast).

Recursive forecasts

Starting point is creating forecasts. We will focus on recursive forecasts.

1. Use observations $1, 2, \ldots, t$ to estimate your model coefficients.
2. Make forecast starting from $t+1$ and ending in $t+h$.
   - Use actual values for dependent variable as initial conditions.
   - Use period t errors as initial conditions for ARMA terms.
   - Use forecast values for dependent variable for future periods (dynamic forecast).
   - Use actual values for independent variables (conditional forecast).
3. Add observation $t+1$ to your sample and re-do (1-2).

Recursive forecasts

Starting point is creating forecasts. We will focus on recursive forecasts.

1. Use observations $1, 2, \ldots, t$ to estimate your model coefficients.
2. Make forecast starting from $t + 1$ and ending in $t + h$.
   - Use actual values for dependent variable as initial conditions.
   - Use period t errors as initial conditions for ARMA terms.
   - Use forecast values for dependent variable for future periods (dynamic forecast).
   - Use actual values for independent variables (conditional forecast).
3. Add observation $t + 1$ to your sample and re-do (1-2).
4. Proceed until $T - h$.

Recursive forecasts

Starting point is creating forecasts. We will focus on recursive forecasts.

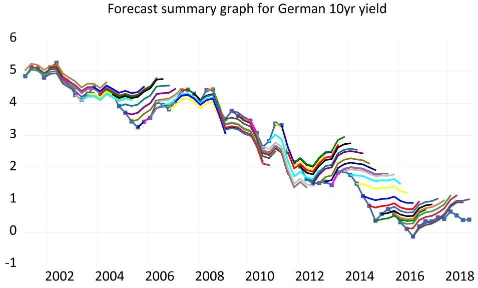
1. Use observations $1, 2, \ldots, t$ to estimate your model coefficients.
2. Make forecast starting from $t+1$ and ending in $t+h$.
    - Use actual values for dependent variable as initial conditions.
    - Use period t errors as initial conditions for ARMA terms.
    - Use forecast values for dependent variable for future periods (dynamic forecast).
    - Use actual values for independent variables (conditional forecast).
3. Add observation $t+1$ to your sample and re-do (1-2).
4. Proceed until $T-h$.

Alternatives: Rolling forecasts, k-fold forecasts.

Forecast summary graphs

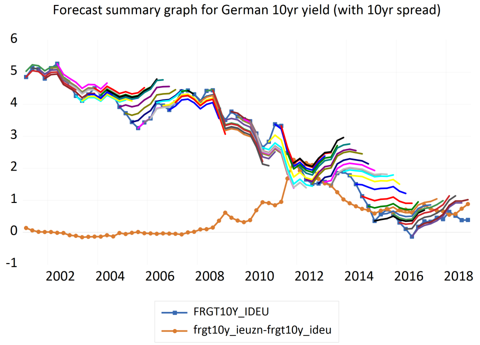
The best way to evaluate recursive forecasts is to graph them together with actuals.

- Graphs allow us to identify periods with large forecast errors, or with systematically positive/negative forecast errors.



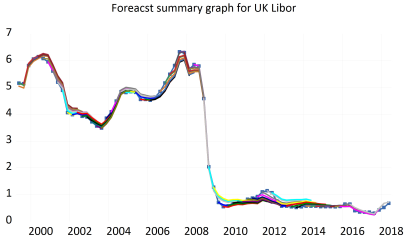
Forecast summary graphs: Comparison variables

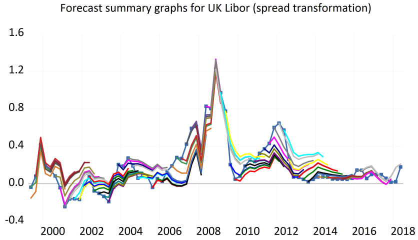
Often developments (or bad forecasts) can be explained by comparing actuals with potential drivers $\Rightarrow$ add them to the graph!



Forecast summary graphs: Transformations

Often we are interested in transformation of given variable such as growth ate or spread $\Rightarrow$ graph the transformation.





Measures of forecast precision

The overall forecast performance can be summarized in terms of summary measures.

Measures of forecast precision

The overall forecast performance can be summarized in terms of summary measures.

1. Forecast performance in particular period can be calculated as difference between forecast and actual value (=forecast error).

Measures of forecast precision

The overall forecast performance can be summarized in terms of summary measures.

1. Forecast performance in particular period can be calculated as difference between forecast and actual value (=forecast error).
2. We want to sum forecast errors over multiple forecasts $\Rightarrow$ use absolute or square function to avoid cancelling.
    - Square gives bigger weight to larger deviations, what is often desirable.
    - If square function is used then reverse it with root function to return to values of given variable.

Measures of forecast precision

The overall forecast performance can be summarized in terms of summary measures.

1. Forecast performance in particular period can be calculated as difference between forecast and actual value (=forecast error).
2. We want to sum forecast errors over multiple forecasts $\Rightarrow$ use absolute or square function to avoid cancelling.
    - Square gives bigger weight to larger deviations, what is often desirable.
    - If square function is used then reverse it with root function to return to values of given variable.

Mean absolute error: $MAE = \frac{1}{H}\sum_{h=1}^{H}|f_{t+h} - a_{t+h}|$

Root mean squared error: $RMSE = \sqrt{\frac{1}{H}\sum_{h=1}^{H}(f_{t+h} - a_{t+h})^2}$

Measures of forecast precision: Additional features

Above definitions use single forecast lasting H periods. What to do when we have multiple forecasts?

> $\Rightarrow$ Take average over all forecasts: for T periods with horizon H we have $T-H$ forecasts yielding e.g. $\frac{1}{T-H}\sum_{t=1}^{T-H} RMSE(t)$.

Measures of forecast precision: Additional features

Above definitions use single forecast lasting H periods. What to do when we have multiple forecasts?

$\Rightarrow$ Take average over all forecasts: for T periods with horizon H we have $T-H$ forecasts yielding e.g. $\frac{1}{T-H}\sum_{t=1}^{T-H} RMSE(t)$.

If dependent variable is transformation, what should be your f_t and a_t?

$\Rightarrow$ Use level of dependent variable, unless transformation is of interest (e.g. growth rate, spread).

Measures of forecast precision: Additional features

Above definitions use single forecast lasting H periods. What to do when we have multiple forecasts?

$\Rightarrow$ Take average over all forecasts: for T periods with horizon H we have $T-H$ forecasts yielding e.g. $\frac{1}{T-H}\sum_{t=1}^{T-H} RMSE(t)$.

If dependent variable is transformation, what should be your f_t and a_t?

$\Rightarrow$ Use level of dependent variable, unless transformation is of interest (e.g. growth rate, spread).

When dependent variable is trending then same percentage errors in different period translate into different absolute errors.

$\Rightarrow$ Divide the errors by actual values to obtain percentage errors:
$MAPE = \frac{1}{H}\sum_{h=1}^{H} |(f_{t+h} - a_{t+h})/a_{t+h}|$
$RMSPE = \sqrt{\frac{1}{H}\sum_{h=1}^{H}((f_{t+h} - a_{t+h})/a_{t+h})^2}$

Forecast bias

Forecast precision statistics measure how much on average are we wrong. Sometimes we are also interested in **direction** of forecast errors: **does model has tendency to over-predict or under-predict on average?**

Forecast bias

Forecast precision statistics measure how much on average are we wrong. Sometimes we are also interested in **direction** of forecast errors: **does model has tendency to over-predict or under-predict on average?**

Forecast bias measures indicate consistent tendency for the model to make positive or negative forecast errors.

Forecast bias

Forecast precision statistics measure how much on average are we wrong. Sometimes we are also interested in **direction** of forecast errors: **does model has tendency to over-predict or under-predict on average?**

Forecast bias measures indicate consistent tendency for the model to make positive or negative forecast errors.

As simplest statistics we can use the mean error, i.e. the average of forecast errors.

- Positive sign will indicate tendency for over-predicting, and negative for under-predicting.

In-sample vs. out-of-sample

The recursive forecasting included re-estimation step: coefficients were estimated on sub-sample before the start of forecasts.

In-sample vs. out-of-sample

The recursive forecasting included re-estimation step: coefficients were estimated on sub-sample before the start of forecasts.

- This is called (pseudo) out-of-sample forecasting: coefficients are based only on data before the start of forecasts.

In-sample vs. out-of-sample

The recursive forecasting included re-estimation step: coefficients were estimated on sub-sample before the start of forecasts.

- This is called (pseudo) out-of-sample forecasting: coefficients are based only on data before the start of forecasts.
- Aim: How well would the **model** forecast given variable if we used it at given historical period? $\Leftrightarrow$ Is your model good model? Which model is the best model?

In-sample vs. out-of-sample

The recursive forecasting included re-estimation step: coefficients were estimated on sub-sample before the start of forecasts.

- This is called (pseudo) out-of-sample forecasting: coefficients are based only on data before the start of forecasts.
- Aim: How well would the **model** forecast given variable if we used it at given historical period? $\Leftrightarrow$ Is your model good model? Which model is the best model?

We can also evaluate model in-sample: skip estimation step and instead use coefficient estimates based on full sample.

In-sample vs. out-of-sample

The recursive forecasting included re-estimation step: coefficients were estimated on sub-sample before the start of forecasts.

- This is called (pseudo) out-of-sample forecasting: coefficients are based only on data before the start of forecasts.
- Aim: How well would the **model** forecast given variable if we used it at given historical period? $\Leftrightarrow$ Is your model good model? Which model is the best model?

We can also evaluate model in-sample: skip estimation step and instead use coefficient estimates based on full sample.

- Aim: How well would the **estimated equation** forecast given variable if we used it at given historical period? $\Leftrightarrow$ Does your equation work well on historical data? Which *estimated* model is working best?

In-sample vs. out-of-sample

Estimation methods optimize one-step-ahead forecasts, potentially leading to over-fitting $\Rightarrow$ if one-step-ahead forecast are of interest focus on out-of-sample.

In-sample vs. out-of-sample

Estimation methods optimize one-step-ahead forecasts, potentially leading to over-fitting $\Rightarrow$ if one-step-ahead forecast are of interest focus on out-of-sample.

The distinction between in-sample and out-of-sample is less important when multi-period forecasts are used since multi-period forecast performance is not being optimized.

- In most applications ranking of in-sample and out-of-sample for multi-period forecasting is going to be fairly similar.

In-sample vs. out-of-sample

Estimation methods optimize one-step-ahead forecasts, potentially leading to over-fitting $\Rightarrow$ if one-step-ahead forecast are of interest focus on out-of-sample.

The distinction between in-sample and out-of-sample is less important when multi-period forecasts are used since multi-period forecast performance is not being optimized.

- In most applications ranking of in-sample and out-of-sample for multi-period forecasting is going to be fairly similar.

When you want to know how well does your estimated equation work under different conditions use in-sample forecasts.

In-sample vs. out-of-sample

Estimation methods optimize one-step-ahead forecasts, potentially leading to over-fitting $\Rightarrow$ if one-step-ahead forecast are of interest focus on out-of-sample.

The distinction between in-sample and out-of-sample is less important when multi-period forecasts are used since multi-period forecast performance is not being optimized.

- In most applications ranking of in-sample and out-of-sample for multi-period forecasting is going to be fairly similar.

When you want to know how well does your estimated equation work under different conditions use in-sample forecasts.

When you want to know how well does your model account for the data generating proces then use out-of-sample forecasts.

Unconditional forecasts

Introduction

Previous sections discussed forecast performance for single equation.

- When single equation has to be evaluated then we use actual historical data for RHS variables to be able to create any forecasts.

Introduction

Previous sections discussed forecast performance for single equation.

- When single equation has to be evaluated then we use actual historical data for RHS variables to be able to create any forecasts.
- Forecasts are based on perfect knowledge of some relevant "future" information.

Introduction

Previous sections discussed forecast performance for single equation.

- When single equation has to be evaluated then we use actual historical data for RHS variables to be able to create any forecasts.
- Forecasts are based on perfect knowledge of some relevant "future" information.

An alternative is to use equations for RHS to forecast their values.

- However, we need to have equation for all model variables $\Leftrightarrow$ multiple equations models (VAR, SEM,DSGE,...).

Long-term forecasts

The forecast evaluation tools for unconditional forecasts are identical as for conditional forecasts.

Long-term forecasts

The forecast evaluation tools for unconditional forecasts are identical as for conditional forecasts.

One additional tool are long-term forecast that can identify problems with model structures (i.e. violation of restrictions or equilibrium conditions).

- In unconditional forecasts we can identify unintended feedback mechanisms.

Terminology: Conditional and unconditional forecasts

Forecasts of series from multiple equations model are called unconditional forecasts.

Terminology: Conditional and unconditional forecasts

Forecasts of series from multiple equations model are called unconditional forecasts.

- In single equation environment we forecast given series based on the historical values for RHS variables $\Leftrightarrow$ the forecasts are conditional on the values of RHS.

Terminology: Conditional and unconditional forecasts

Forecasts of series from multiple equations model are called unconditional forecasts.

- In single equation environment we forecast given series based on the historical values for RHS variables $\Leftrightarrow$ the forecasts are conditional on the values of RHS.
- In unconditional forecasts we are forecasting values of RHS variables $\Leftrightarrow$ they are unconditional on any "future" information.

Why do we care about conditional forecasts?

Only unconditional forecasts are *true* forecasts. Why do we study conditional forecasts?

- Forecast error for particular variable in unconditional forecasts can originate in:
  1. Equation for given variable.
  2. Forecasts for RHS variables.

Why do we care about conditional forecasts?

Only unconditional forecasts are *true* forecasts. Why do we study conditional forecasts?

- Forecast error for particular variable in unconditional forecasts can originate in:
  1. Equation for given variable.
  2. Forecasts for RHS variables.

  $\Rightarrow$ It is hard to identify the source of forecast error in unconditional forecasts.

Why do we care about conditional forecasts?

Only unconditional forecasts are *true* forecasts. Why do we study conditional forecasts?

- Forecast error for particular variable in unconditional forecasts can originate in:
  1. Equation for given variable.
  2. Forecasts for RHS variables.

  $\Rightarrow$ It is hard to identify the source of forecast error in unconditional forecasts.

  $\Rightarrow$ **In unconditional forecasts we are evaluating all model equations, not just particular equation.**

Why do we care about conditional forecasts?

Only unconditional forecasts are *true* forecasts. Why do we study conditional forecasts?

- Forecast error for particular variable in unconditional forecasts can originate in:
  1. Equation for given variable.
  2. Forecasts for RHS variables.

  $\Rightarrow$ It is hard to identify the source of forecast error in unconditional forecasts.

  $\Rightarrow$ **In unconditional forecasts we are evaluating all model equations, not just particular equation.**

Good conditional forecast performance is necessary condition for good unconditional forecasting performance $\Rightarrow$ use it as primary selection tool.

Why do we care about conditional forecasts?

Only unconditional forecasts are *true* forecasts. Why do we study conditional forecasts?

- Forecast error for particular variable in unconditional forecasts can originate in:
  1. Equation for given variable.
  2. Forecasts for RHS variables.

  $\Rightarrow$ It is hard to identify the source of forecast error in unconditional forecasts.

  $\Rightarrow$ **In unconditional forecasts we are evaluating all model equations, not just particular equation.**

Good conditional forecast performance is necessary condition for good unconditional forecasting performance $\Rightarrow$ use it as primary selection tool.

It is often impossible to perform pseudo out-of-sample backtesting for multiple equation models due to large number of coefficients to be re-estimated.

Sensitivity to shocks

Why do we care about sensitivity to shocks?

Forecasting performance measures (R-squared, RMSE, information criteria etc.) are summary statistics: they are averaged across all the the periods.

Why do we care about sensitivity to shocks?

Forecasting performance measures (R-squared, RMSE, information criteria etc.) are summary statistics: they are averaged across all the the periods.

However, we might be especially interested in forecasting performance in stress (unusual) periods.

Why do we care about sensitivity to shocks?

Forecasting performance measures (R-squared, RMSE, information criteria etc.) are summary statistics: they are averaged across all the the periods.

However, we might be especially interested in forecasting performance in stress (unusual) periods.

- This is especially true when models are used for scenario forecasting: scenarios are collections of shocks, and we need these **shocks to be transmitted properly to all model variables.**
- The value of scenario forecast often depends on consistency of the forecasts.

Why do we care about sensitivity to shocks?

Forecasting performance measures (R-squared, RMSE, information criteria etc.) are summary statistics: they are averaged across all the the periods.

However, we might be especially interested in forecasting performance in stress (unusual) periods.

- This is especially true when models are used for scenario forecasting: scenarios are collections of shocks, and we need these **shocks to be transmitted properly to all model variables.**
- The value of scenario forecast often depends on consistency of the forecasts.

Unusual periods are by definition rare $\Rightarrow$ even good average forecasting performance does not ensure good performance in unusual periods.

Why do we care about sensitivity to shocks?

Forecasting performance measures (R-squared, RMSE, information criteria etc.) are summary statistics: they are averaged across all the the periods.

However, we might be especially interested in forecasting performance in stress (unusual) periods.

- This is especially true when models are used for scenario forecasting: scenarios are collections of shocks, and we need these **shocks to be transmitted properly to all model variables.**
- The value of scenario forecast often depends on consistency of the forecasts.

Unusual periods are by definition rare $\Rightarrow$ even good average forecasting performance does not ensure good performance in unusual periods.

We want a **way to ensure that model responds sufficiently to shocks**.

Standardized shocks

Standardized coefficients

In principle we have already discussed one measure of sensitivity: standardized coefficients.

- Standardized coefficients measure the response of dependent variable to **standard shock** to independent variable.
- Standardized coefficients tells us whether the independent variable influence the forecast for independent variable.

Drawbacks: Size

Standard deviation is a summary statistic $\Rightarrow$ it might not be representative of extreme shocks, which are the shocks we are most interested in!

Drawbacks: Size

Standard deviation is a summary statistic $\Rightarrow$ it might not be representative of extreme shocks, which are the shocks we are most interested in!

What if the variables do not have standard deviation?

Drawbacks: Size

Standard deviation is a summary statistic $\Rightarrow$ it might not be representative of extreme shocks, which are the shocks we are most interested in!

What if the variables do not have standard deviation?

- Co-integration regressions can include non-stationary varaible $\Leftrightarrow$ standard deviation changes over time.

Drawbacks: Size

Standard deviation is a summary statistic $\Rightarrow$ it might not be representative of extreme shocks, which are the shocks we are most interested in!

What if the variables do not have standard deviation?

- Co-integration regressions can include non-stationary varaible $\Leftrightarrow$ standard deviation changes over time.
- Co-integration regressions can include trending variables $\Rightarrow$ standard deviation is non-nonsensical.

Drawbacks: Size

Standard deviation is a summary statistic $\Rightarrow$ it might not be representative of extreme shocks, which are the shocks we are most interested in!

What if the variables do not have standard deviation?

- Co-integration regressions can include non-stationary varaible $\Leftrightarrow$ standard deviation changes over time.
- Co-integration regressions can include trending variables $\Rightarrow$ standard deviation is non-nonsensical.
- Dummy variables do not have meaningful standard deviation.

Drawbacks: Correlation

Regressors are often correlated $\Leftrightarrow$ upside/downside shocks to regressors occur at the same time.

Drawbacks: Correlation

Regressors are often correlated $\Leftrightarrow$ upside/downside shocks to regressors occur at the same time.

- Response to single shock has limited informational value.

Drawbacks: Correlation

Regressors are often correlated $\Leftrightarrow$ upside/downside shocks to regressors occur at the same time.

- Response to single shock has limited informational value.
- Constructing multiple standardized shocks requires determining correlation between the shocks.

Drawbacks: Correlation

Regressors are often correlated $\Leftrightarrow$ upside/downside shocks to regressors occur at the same time.

- Response to single shock has limited informational value.
- Constructing multiple standardized shocks requires determining correlation between the shocks.
- Correlation between shocks is possibly different for upside and downside shocks, and for shocks of different sizes.

Historical shocks

Introduction

The problem with standardized shocks is that they are standardized, i.e. their size (and correlation) is constructed and hence might not correspond to historical shocks.

Introduction

The problem with standardized shocks is that they are standardized, i.e. their size (and correlation) is constructed and hence might not correspond to historical shocks.

Alternative approach: use actual observed historical shocks.

Introduction

The problem with standardized shocks is that they are standardized, i.e. their size (and correlation) is constructed and hence might not correspond to historical shocks.

Alternative approach: use actual observed historical shocks.

- History provides us with examples of unusual stress periods, which can serve us as natural laboratory for analyzing responses of model to shocks.

Introduction

The problem with standardized shocks is that they are standardized, i.e. their size (and correlation) is constructed and hence might not correspond to historical shocks.

Alternative approach: use actual observed historical shocks.

- History provides us with examples of unusual stress periods, which can serve us as natural laboratory for analyzing responses of model to shocks.
- Main example of such historical stress period is 2007-2010 period (Global financial crisis and Great recession).

Introduction

The problem with standardized shocks is that they are standardized, i.e. their size (and correlation) is constructed and hence might not correspond to historical shocks.

Alternative approach: use actual observed historical shocks.

- History provides us with examples of unusual stress periods, which can serve us as natural laboratory for analyzing responses of model to shocks.
- Main example of such historical stress period is 2007-2010 period (Global financial crisis and Great recession).
- For specific countries/variables other periods are available.

Introduction

The problem with standardized shocks is that they are standardized, i.e. their size (and correlation) is constructed and hence might not correspond to historical shocks.

Alternative approach: use actual observed historical shocks.

- History provides us with examples of unusual stress periods, which can serve us as natural laboratory for analyzing responses of model to shocks.
- Main example of such historical stress period is 2007-2010 period (Global financial crisis and Great recession).
- For specific countries/variables other periods are available.

The use of historical periods is also justified by the fact that macroeconomic scenarios are typically based on historical stress periods.

Illustration

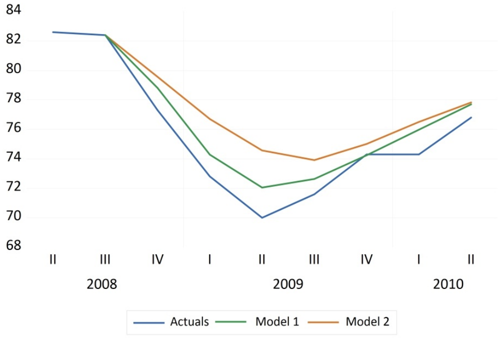
Despite better fit, model 1 is less responsive to shocks than model 2.

Dependent Variable: D(FCU_IGBR)
Method: Least Squares
Date: 06/06/19 Time: 18:29
Sample (adjusted): 1988Q3 2018Q4
Included observations: 122 after adjustments

Variable	Coefficient
C	19.29401
FCU_IGBR(-1)	-0.247059
@MOVAV(D(FGAP_IGBR),4)	1.060054
@MOVAV(DLOG(FGDP$EWQ_IGBR),2)	119.7069
R-squared	0.358519
Adjusted R-squared	0.342210

Dependent Variable: D(FCU_IGBR)
Method: Least Squares
Date: 26/04/19 Time: 20:12
Sample (adjusted): 1980Q2 2018Q4
Included observations: 155 after adjustments

Variable	Coefficient
C	23.19348
FCU_IGBR(-1)	-0.289693
D(FLBR_IGBR)	-1.190657
D(FLBR_IGBR(-1))	-1.137764
@PCY(FIP_IGBR)	0.146978
R-squared	0.378972
Adjusted R-squared	0.362411



Drawbacks

The main advantage of using historical shocks is also source of main drawback - the historical period might be unique, rather than informative about future stress periods.

Drawbacks

The main advantage of using historical shocks is also source of main drawback - the historical period might be unique, rather than informative about future stress periods.

- We might develop model that captures well the behaviour of given variable *only* during the specific stress period, but not general(future) stress periods.

Drawbacks

The main advantage of using historical shocks is also source of main drawback - the historical period might be unique, rather than informative about future stress periods.

- We might develop model that captures well the behaviour of given variable *only* during the specific stress period, but not general(future) stress periods.
- The point is the same as general point in over-fitting: by focusing on particular historical realizations we might be over-fitting our model.

Scenario shocks

Introduction

Alternative to using observed historical shocks is to use **shocks specified in particular macroeconomic scenario**.

- Macroeconomic scenarios are collection of shocks (and responses to shocks) that have undergone analytical scrutiny.

Introduction

Alternative to using observed historical shocks is to use **shocks specified in particular macroeconomic scenario**.

- Macroeconomic scenarios are collection of shocks (and responses to shocks) that have undergone analytical scrutiny.

Testing models for given variable on existing scenarios is testing whether the models will be able to perform as we expect in such scenarios.

Introduction

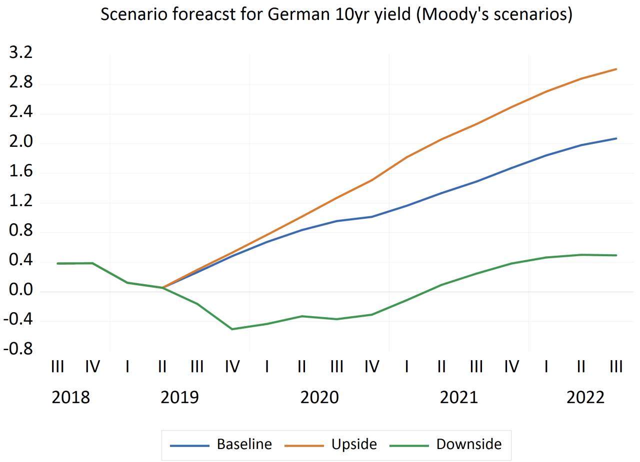
Alternative to using observed historical shocks is to use **shocks specified in particular macroeconomic scenario**.

- Macroeconomic scenarios are collection of shocks (and responses to shocks) that have undergone analytical scrutiny.

Testing models for given variable on existing scenarios is testing whether the models will be able to perform as we expect in such scenarios.

Main issue: There is nothing to compare the forecast with, so the performance is **evaluated in judgemental fashion**.

Illustration



Unconditional sensitivity

Motivation

All previous methods of analyzing sensitivity focused on single equation.

- We shocked single or multiple independent variables and studied the effect on dependent variable.

Motivation

All previous methods of analyzing sensitivity focused on single equation.

- We shocked single or multiple independent variables and studied the effect on dependent variable.

In such analysis there is only one direction of effects. The independent variables are treated as independent of (exogenous to)...

Motivation

All previous methods of analyzing sensitivity focused on single equation.

- We shocked single or multiple independent variables and studied the effect on dependent variable.

In such analysis there is only one direction of effects. The independent variables are treated as independent of (exogenous to)...

- ...the dependent variable (no feedback).

Motivation

All previous methods of analyzing sensitivity focused on single equation.

- We shocked single or multiple independent variables and studied the effect on dependent variable.

In such analysis there is only one direction of effects. The independent variables are treated as independent of (exogenous to)...

- ...the dependent variable (no feedback).
- ...shocks in other independent variables (no cross influences for RHS variables).

Motivation

All previous methods of analyzing sensitivity focused on single equation.

- We shocked single or multiple independent variables and studied the effect on dependent variable.

In such analysis there is only one direction of effects. The independent variables are treated as independent of (exogenous to)...

- ...the dependent variable (no feedback).
- ...shocks in other independent variables (no cross influences for RHS variables).

In case of multiple equation models we want to allow for all possible channels.

Responses to unconditional shocks

The key idea of using unconditional shocks is that the shocks can propagate across the system and also have feedback effects.

Responses to unconditional shocks

The key idea of using unconditional shocks is that the shocks can propagate across the system and also have feedback effects.

A shock to single variable affects the forecast in multiple ways.

Responses to unconditional shocks

The key idea of using unconditional shocks is that the shocks can propagate across the system and also have feedback effects.

A shock to single variable affects the forecast in multiple ways.

1. Shock increases current value of given variable.

Responses to unconditional shocks

The key idea of using unconditional shocks is that the shocks can propagate across the system and also have feedback effects.

A shock to single variable affects the forecast in multiple ways.

1. Shock increases current value of given variable.
2. Shocks increases future values of given variable if its equation features persistence.

Responses to unconditional shocks

The key idea of using unconditional shocks is that the shocks can propagate across the system and also have feedback effects.

A shock to single variable affects the forecast in multiple ways.

1. Shock increases current value of given variable.
2. Shocks increases future values of given variable if its equation features persistence.
3. It affects the current and/or future values of other variables through presence of given variable in equation for other variables.

Responses to unconditional shocks

The key idea of using unconditional shocks is that the shocks can propagate across the system and also have feedback effects.

A shock to single variable affects the forecast in multiple ways.

1. Shock increases current value of given variable.
2. Shocks increases future values of given variable if its equation features persistence.
3. It affects the current and/or future values of other variables through presence of given variable in equation for other variables.
4. It affects future values of other variables through *second-order effects*.

Responses to unconditional shocks

The key idea of using unconditional shocks is that the shocks can propagate across the system and also have feedback effects.

A shock to single variable affects the forecast in multiple ways.

1. Shock increases current value of given variable.
2. Shocks increases future values of given variable if its equation features persistence.
3. It affects the current and/or future values of other variables through presence of given variable in equation for other variables.
4. It affects future values of other variables through *second-order effects*.
5. It affects its *own* future values through second-order effects.

What shocks?

In unconditional sensitivity we typically introduce single shock and study the response of few target variables.

What shocks?

In unconditional sensitivity we typically introduce single shock and study the response of few target variables.

There are two types of shocks we typically use:

What shocks?

In unconditional sensitivity we typically introduce single shock and study the response of few target variables.

There are two types of shocks we typically use:

1. Economically meaningful shocks: increase in interest rates by 1%, decrease in oil/house/stock prices by 10%, increase in exchange rate by 5%,...

What shocks?

In unconditional sensitivity we typically introduce single shock and study the response of few target variables.

There are two types of shocks we typically use:

1. Economically meaningful shocks: increase in interest rates by 1%, decrease in oil/house/stock prices by 10%, increase in exchange rate by 5%,...
2. Residual-based shocks: use residual corresponding to chosen percentile of residuals.

What shocks?

In unconditional sensitivity we typically introduce single shock and study the response of few target variables.

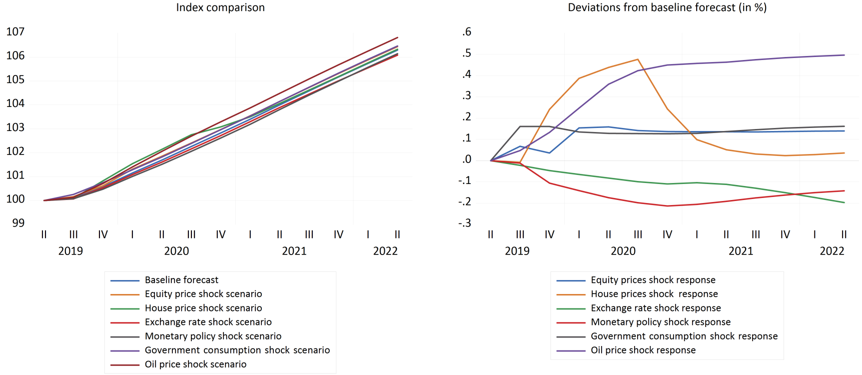
There are two types of shocks we typically use:

1. Economically meaningful shocks: increase in interest rates by 1%, decrease in oil/house/stock prices by 10%, increase in exchange rate by 5%,...
2. Residual-based shocks: use residual corresponding to chosen percentile of residuals.

The key reason for using economically meaningful shocks is that economic theory provides us with guidance about the direction and magnitudes of responses.

Illustration

UK GDP under unconditonal shocks



Forecast drivers

Motivation

Sensitivity methods based on multiple shocks pose a problem: we know only the overall shock response.

Motivation

Sensitivity methods based on multiple shocks pose a problem: we know only the overall shock response.

We would like to decompose the overall effect into parts attributable to each variable.

Decomposing forecast

In general, in any linear model $y_t = \beta_0 + \beta_1 x_t + \beta_2 z_t$ we can decompose the forecast for y_t into three components: β_0, $\beta_1 x_t$ and $\beta_2 z_t$.

- Coefficients β_1 and β_2 translate the movements in x and z into movements of y.

Decomposing forecast

In general, in any linear model $y_t = \beta_0 + \beta_1 x_t + \beta_2 z_t$ we can decompose the forecast for y_t into three components: β_0, $\beta_1 x_t$ and $\beta_2 z_t$.

- Coefficients β_1 and β_2 translate the movements in x and z into movements of y.

Logically, decomposition means that sum of the three component will be equal to y_t.

Decomposing forecast

In general, in any linear model $y_t = \beta_0 + \beta_1 x_t + \beta_2 z_t$ we can decompose the forecast for y_t into three components: β_0, $\beta_1 x_t$ and $\beta_2 z_t$.

- Coefficients β_1 and β_2 translate the movements in x and z into movements of y.

Logically, decomposition means that sum of the three component will be equal to y_t.

- Each component is the contribution of given variable to forecast for y.

Decomposing forecast

In general, in any linear model $y_t = \beta_0 + \beta_1 x_t + \beta_2 z_t$ we can decompose the forecast for y_t into three components: β_0, $\beta_1 x_t$ and $\beta_2 z_t$.

- Coefficients β_1 and β_2 translate the movements in x and z into movements of y.

Logically, decomposition means that sum of the three component will be equal to y_t.

- Each component is the contribution of given variable to forecast for y.

The evolution of forecast for y in time can be decomposed into evolution of x and z, adjusted for the coefficients

Decomposing forecast

In general, in any linear model $y_t = \beta_0 + \beta_1 x_t + \beta_2 z_t$ we can decompose the forecast for y_t into three components: β_0, $\beta_1 x_t$ and $\beta_2 z_t$.

- Coefficients β_1 and β_2 translate the movements in x and z into movements of y.

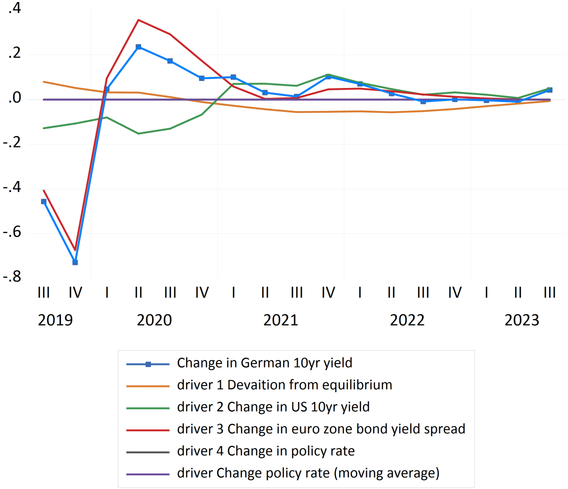
Logically, decomposition means that sum of the three component will be equal to y_t.

- Each component is the contribution of given variable to forecast for y.

The evolution of forecast for y in time can be decomposed into evolution of x and z, adjusted for the coefficients

- The forecasted changes in y are driven by changes in x and $z \Rightarrow \beta_1 x_t$ and $\beta_2 z_t$ are the forecast drivers ($FD_t(x) \equiv \beta_1 x_t$ and $FD_t(z) \equiv \beta_1 z_t$).

Illustration I



Forecast drivers across scenarios

The simple forecast drivers graph decomposes single forecast into its drivers. However, our goal was to decompose the effect of shocks, not the actual forecast.

Forecast drivers across scenarios

The simple forecast drivers graph decomposes single forecast into its drivers. However, our goal was to decompose the effect of shocks, not the actual forecast.

Solution: We can decompose both the forecast with and without shock into their drivers, and subtract the no-shock drivers from shock drivers:

$$FD_t^D(x) = FD_t^S(x) - FD_t^{BL}(x)$$

Forecast drivers across scenarios

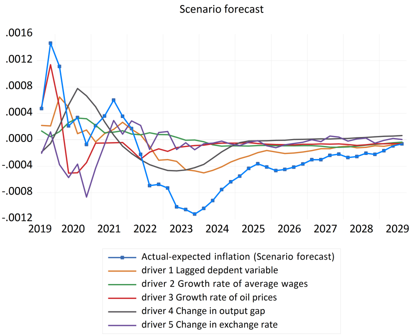
The simple forecast drivers graph decomposes single forecast into its drivers. However, our goal was to decompose the effect of shocks, not the actual forecast.

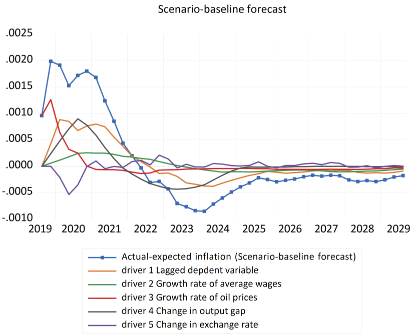
Solution: We can decompose both the forecast with and without shock into their drivers, and subtract the no-shock drivers from shock drivers:

$$FD_t^D(x) = FD_t^S(x) - FD_t^{BL}(x)$$

Resulting series capture the effect of *shocks* in independent on forecast for dependent variable.

Illustration II





Tips&tricks

Introduction

Apart from approach to model building, there are also several practical considerations that is good to keep in mind.

Introduction

Apart from approach to model building, there are also several practical considerations that is good to keep in mind.

We will discuss.

Introduction

Apart from approach to model building, there are also several practical considerations that is good to keep in mind.

We will discuss.

1. Forecast jump offs produced by level specifications...

Introduction

Apart from approach to model building, there are also several practical considerations that is good to keep in mind.

We will discuss.

1. Forecast jump offs produced by level specifications...
2. ...and whether to deal with them by using LDV or ARMA terms...

Introduction

Apart from approach to model building, there are also several practical considerations that is good to keep in mind.

We will discuss.

1. Forecast jump offs produced by level specifications...
2. ...and whether to deal with them by using LDV or ARMA terms...
3. ...and why it is typically not a good idea to solve it by using differences.

Introduction

Apart from approach to model building, there are also several practical considerations that is good to keep in mind.

We will discuss.

1. Forecast jump offs produced by level specifications...
2. ...and whether to deal with them by using LDV or ARMA terms...
3. ...and why it is typically not a good idea to solve it by using differences.
4. Whether to use level or spread specifications...

Introduction

Apart from approach to model building, there are also several practical considerations that is good to keep in mind.

We will discuss.

1. Forecast jump offs produced by level specifications...
2. ...and whether to deal with them by using LDV or ARMA terms...
3. ...and why it is typically not a good idea to solve it by using differences.
4. Whether to use level or spread specifications...
5. ...and how that relates to error-correction models.

Introduction

Apart from approach to model building, there are also several practical considerations that is good to keep in mind.

We will discuss.

1. Forecast jump offs produced by level specifications...
2. ...and whether to deal with them by using LDV or ARMA terms...
3. ...and why it is typically not a good idea to solve it by using differences.
4. Whether to use level or spread specifications...
5. ...and how that relates to error-correction models.
6. Whether to use log-differences or percentage changes when modelling growth.

Introduction

Apart from approach to model building, there are also several practical considerations that is good to keep in mind.

We will discuss.

1. Forecast jump offs produced by level specifications...
2. ...and whether to deal with them by using LDV or ARMA terms...
3. ...and why it is typically not a good idea to solve it by using differences.
4. Whether to use level or spread specifications...
5. ...and how that relates to error-correction models.
6. Whether to use log-differences or percentage changes when modelling growth.
7. And why not to (almost) ever use moving average of y/y growth rates as dependent variable.

Applications

Outline

Modelling money market rate

Introduction

Money market rates are interest rates for which banks can borrow from other banks.

- Current generation of rates is based on self-assessed answers from panel of large banks.
- New generation will be based on actual transactions.

Introduction

Money market rates are interest rates for which banks can borrow from other banks.

- Current generation of rates is based on self-assessed answers from panel of large banks.
- New generation will be based on actual transactions.

How do we model such rates in multivariate environment?

Introduction

Money market rates are interest rates for which banks can borrow from other banks.

- Current generation of rates is based on self-assessed answers from panel of large banks.
- New generation will be based on actual transactions.

How do we model such rates in multivariate environment?

Key idea: Borrowing from central bank is alternative source of funding to borrowing from other banks.

Introduction

Money market rates are interest rates for which banks can borrow from other banks.

- Current generation of rates is based on self-assessed answers from panel of large banks.
- New generation will be based on actual transactions.

How do we model such rates in multivariate environment?

Key idea: Borrowing from central bank is alternative source of funding to borrowing from other banks.

$\Rightarrow$ Rates should be anchored by key policy rates.

Introduction

Money market rates are interest rates for which banks can borrow from other banks.

- Current generation of rates is based on self-assessed answers from panel of large banks.
- New generation will be based on actual transactions.

How do we model such rates in multivariate environment?

Key idea: Borrowing from central bank is alternative source of funding to borrowing from other banks.

$\Rightarrow$ Rates should be anchored by key policy rates.

We will look at specific example of 3m Euribor, and our policy rates will be using main refinancing rate (RMP) and deposit rate (RMDEP).

Level specification

Estimation output

Our starting point could be a level specification: $MMR_t = \beta_0 + \beta_1 RMP_t$

Estimation output

Our starting point could be a level specification: $MMR_t = \beta_0 + \beta_1 RMP_t$

The specification looks good in terms of coefficient significance and R-squared.

- One warning sign is the main coefficient which implies more than proportional effect of change in RMP.

Dependent Variable: FREURIBOR3M_IEUZN
Method: Least Squares
Date: 06/17/19 Time: 08:53
Sample (adjusted): 1999Q1 2019Q1
Included observations: 81 after adjustments

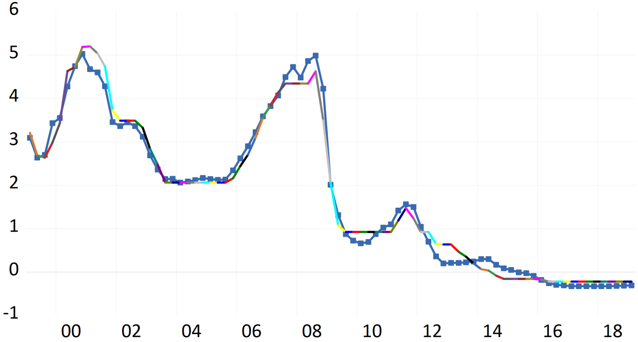
Variable	Coefficient	Std. Error	t-Statistic	Prob.	Standardized
C	-0.219456	0.038867	-5.646284	0.0000	
FRMP_IEUZN	1.140468	0.016645	68.51659	0.0000	0.994209

R-squared	0.983450	Mean dependent var	1.825443
Adjusted R-squared	0.983241	S.D. dependent var	1.730968
S.E. of regression	0.224086	Akaike info criterion	-0.129193
Sum squared resid	3.966944	Schwarz criterion	-0.070071
Log likelihood	7.232334	Hannan-Quinn criter.	-0.105473
F-statistic	4694.523	Durbin-Watson stat	0.663553
Prob(F-statistic)	0.000000		

Performance

In terms of level the model seems to perform well.

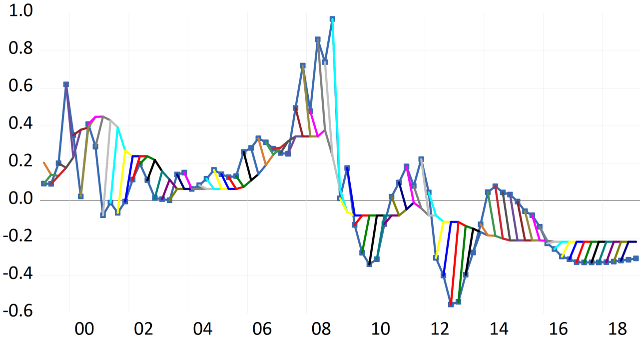
- RMSE at 2-year horizon is 0.205.



Performance: Spread

First indication of problem is the fit of the model in terms of spread.

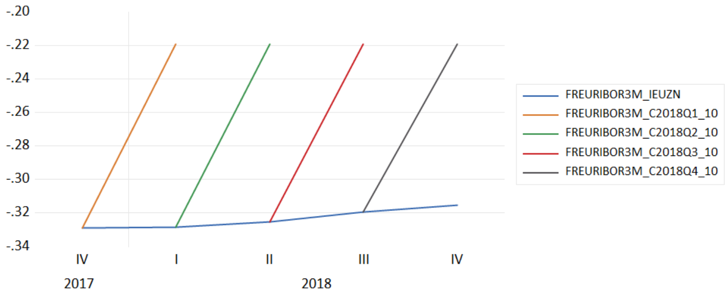
- The model implies **immediate** return to *implied equilibrium* spread.



Level specification: Forecast jump-offs

The main problem of the specification is the implied forecast jump-offs.

- Source: absence of any dynamic terms, despite persistence of the spread between the two series.



Dynamic specifications

Motivation

There are three reasons to consider dynamic specifications:

Motivation

There are three reasons to consider dynamic specifications:

1. Empirical: Persistence of spread.

Motivation

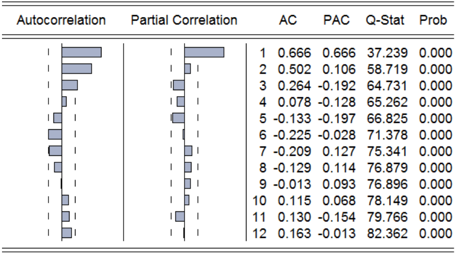
There are three reasons to consider dynamic specifications:

1. Empirical: Persistence of spread.
2. Practical: Forecast-jump-offs.

Motivation

There are three reasons to consider dynamic specifications:

1. Empirical: Persistence of spread.
2. Practical: Forecast-jump-offs.
3. Econometrical: High autocorrelation in residuals.



Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.666	0.666	37.239	0.000
		2	0.502	0.106	58.719	0.000
		3	0.264	-0.192	64.731	0.000
		4	0.078	-0.128	65.262	0.000
		5	-0.133	-0.197	66.825	0.000
		6	-0.225	-0.028	71.378	0.000
		7	-0.209	0.127	75.341	0.000
		8	-0.129	0.114	76.879	0.000
		9	-0.013	0.093	76.896	0.000
		10	0.115	0.068	78.149	0.000
		11	0.130	-0.154	79.766	0.000
		12	0.163	-0.013	82.362	0.000

Two possible approaches: Lagged dependent variable or ARMA errors.

LDV model

Specification with LDV looks still good in terms of main statistics.

- Coefficient on RMP looks problematic.
- Coefficient on LDV is quite low.
- DW statistic still low.

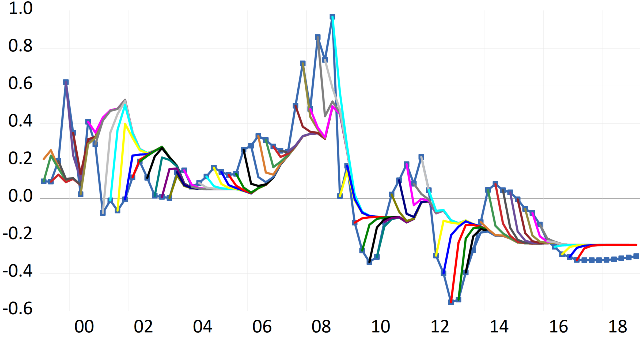
Variable	Coefficient	Std. Error	t-Statistic	Prob.	Standardized
C	-0.189694	0.036289	-5.227238	0.0000	
FREURIBOR3M_IEUZN(-1)	0.235480	0.058008	4.059434	0.0001	0.234155
FRMP_IEUZN	0.878369	0.066335	13.24136	0.0000	0.763783

R-squared	0.986337	Mean dependent var	1.825443
Adjusted R-squared	0.985987	S.D. dependent var	1.730968
S.E. of regression	0.204909	Akaike info criterion	-0.296171
Sum squared resid	3.275030	Schwarz criterion	-0.207488
Log likelihood	14.99492	Hannan-Quinn criter.	-0.260590
F-statistic	2815.416	Durbin-Watson stat	0.859809
Prob(F-statistic)	0.000000		

LDV performance

While the performance seems improved, the spreads is still not sufficiently persistent.

- 2-year RMSE is actually higher (0.217).



ARMA model

Model with ARMA(1,0) errors has much bigger persistence than model with LDV.

- DW statistic no longer indicates problems with residuals.

Variable	Coefficient	Std. Error	t-Statistic	Prob.	Standardized
C	-0.196031	0.100864	-1.943524	0.0556	
FRMP_IEUZN	1.124652	0.027849	40.38449	0.0000	0.977938
AR(1)	0.667100	0.067241	9.921062	0.0000	
SIGMASQ	0.027079	0.003039	8.910386	0.0000	

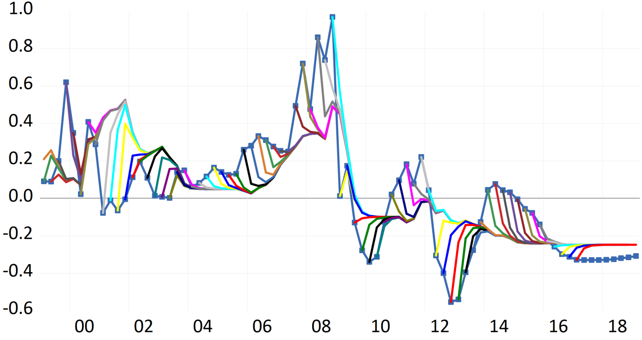
R-squared	0.990849	Mean dependent var	1.825443
Adjusted R-squared	0.990493	S.D. dependent var	1.730968
S.E. of regression	0.168777	Akaike info criterion	-0.665092
Sum squared resid	2.193384	Schwarz criterion	-0.546847
Log likelihood	30.93621	Hannan-Quinn criter.	-0.617650
F-statistic	2779.268	Durbin-Watson stat	2.124675
Prob(F-statistic)	0.000000		

Inverted AR Roots	.67

ARMA model performance

The greater persistence substantially improves the fit in terms of spread.

- 2-year RMSE is slightly lower than for level specification (0.196 vs 0.205).



LDV vs AMRA

LDV model and ARMA model seems very similar. Why is there such a large difference in them?

LDV vs AMRA

LDV model and ARMA model seems very similar. Why is there such a large difference in them?

While the two dynamic model structures look very similar, they actually have very different meaning.

LDV vs AMRA

LDV model and ARMA model seems very similar. Why is there such a large difference in them?

While the two dynamic model structures look very similar, they actually have very different meaning.

- LDV: Current MMR is equal to (fraction of) current RMP and fraction of previous value of MMR.

LDV vs AMRA

LDV model and ARMA model seems very similar. Why is there such a large difference in them?

While the two dynamic model structures look very similar, they actually have very different meaning.

- LDV: Current MMR is equal to (fraction of) current RMP and fraction of previous value of MMR.
- ARMA: Current MMR is equal to (multiple of) current RMP and fraction of the difference between last forecast value and last observed value.

LDV vs AMRA

LDV model and ARMA model seems very similar. Why is there such a large difference in them?

While the two dynamic model structures look very similar, they actually have very different meaning.

- LDV: Current MMR is equal to (fraction of) current RMP and fraction of previous value of MMR.
- ARMA: Current MMR is equal to (multiple of) current RMP and fraction of the difference between last forecast value and last observed value.

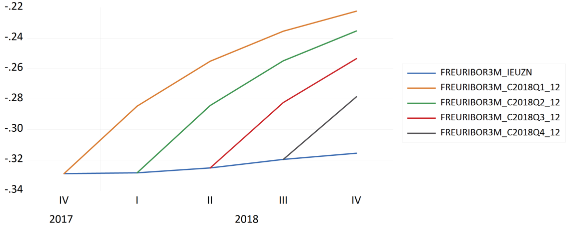
The LDV structure makes the *actual dependent variable* persistent, while the ARMA structure makes the *shocks* persistent.

- The latter is more appropriate description of reality, since Euribor can move a lot from period to period.

Specification in differences

Motivation

While the ARMA model alleviates the forecast jump-offs problem, it does not elimiate it completely.



This seems to be indicative of non-stationarity: specifications in levels will return to average relationship with RMP, but this return does not seem to correspond to actual data.

Model

Using differences instead of levels will eliminate tendency to return to historical average relationship.

Model

Using differences instead of levels will eliminate tendency to return to historical average relationship.

The model has high R-squared despite using differences, the single coefficient is significant and DW is close to 2.

Dependent Variable: @D(FREURIBOR3M_IEUZN)
Method: Least Squares
Date: 06/17/19 Time: 10:10
Sample (adjusted): 1999Q2 2019Q1
Included observations: 80 after adjustments

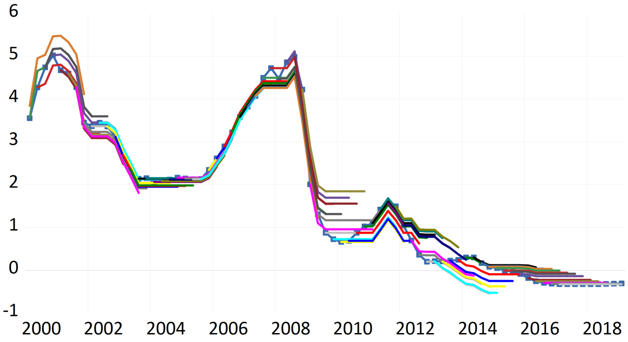
Variable	Coefficient	Std. Error	t-Statistic	Prob.	Standardized
@D(FRMP_IEUZN)	1.055987	0.066404	15.90244	0.0000	0.877069

R-squared	0.758751	Mean dependent var	-0.042495
Adjusted R-squared	0.758751	S.D. dependent var	0.367888
S.E. of regression	0.180696	Akaike info criterion	-0.571581
Sum squared resid	2.579430	Schwarz criterion	-0.541806
Log likelihood	23.86324	Hannan-Quinn criter.	-0.559643
Durbin-Watson stat	2.440453		

Performance: Level

The model seem to be working relatively well with exception of 2008-2012 period.

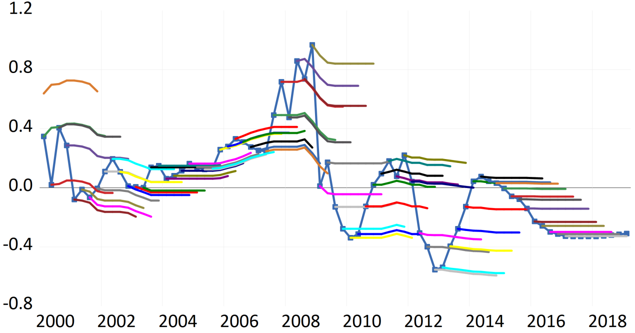
- The model fulfilled the main goal of good forecast performance in recent years.



Performance: Spread

Looking at performance in terms of spread reveals that the model is actually very bad.

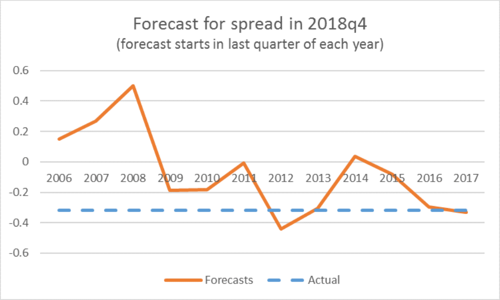
- 2-year RMSE is 0.263.



Additional problems

The model is also problematic because of instability of long-term forecasts *for spread*.

Related issue is the implication of the estimated coefficient: Coefficient of 1.05 means that spread increases by 5bps for each 100bps increase in policy rates.



[A] Solutions

It seems that adding ARMA could help: shocks to changes in spread are likely negative correlated. However, in practice the effect on forecasts is too small.

[A] Solutions

It seems that adding ARMA could help: shocks to changes in spread are likely negative correlated. However, in practice the effect on forecasts is too small.

Another potential solution is include regressoring capturing stress component such as U.S. Libor spread.

- While the forecast performance is improved in in stress period, the long-term forecasts are still unstable.

Error-correction models

Motivation

Econometric testing would lead us to conclude that the two series are non-stationary.

Motivation

Econometric testing would lead us to conclude that the two series are non-stationary.

The two series are also clearly co-integrated: they move closely together and theory suggests there is equilibrium relationship between them.

Motivation

Econometric testing would lead us to conclude that the two series are non-stationary.

The two series are also clearly co-integrated: they move closely together and theory suggests there is equilibrium relationship between them.

Natural candidate model in such situation is the error-correction model:

$$d(MMR_t) = \beta(MMR_{t-1} - \gamma_0 - \gamma_1 RMP_{t-1})$$

- Note that the model is combining features of level specification and differenced specification.
  - ▷ Differences as dependent variable ensure that forecasts will be smooth.
  - ▷ Level anchor ensures that long-term spread is stable.

Model

The model implies that 23% of the disequilibrium is eliminated each quarter.

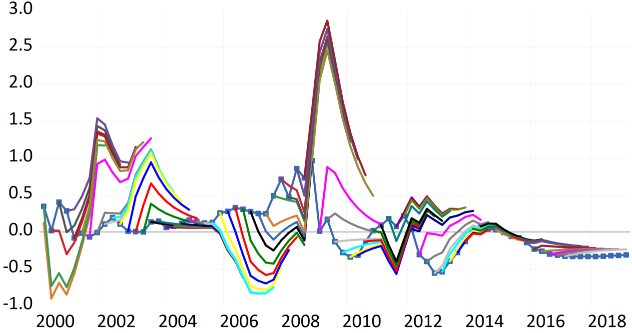
Dependent Variable: @D(FREURIBOR3M_IEUZN)
Method: Least Squares
Date: 06/18/19 Time: 10:24
Sample (adjusted): 1999Q2 2019Q1
Included observations: 80 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.	Standardized
FREURIBOR3M_IEUZN(-1)--0.22-1.14*FRMP_IEUZN(-1)	-0.229213	0.184344	-1.243401	0.2174	-0.139476

R-squared	0.005942	Mean dependent var	-0.042495
Adjusted R-squared	0.005942	S.D. dependent var	0.367888
S.E. of regression	0.366794	Akaike info criterion	0.844386
Sum squared resid	10.62847	Schwarz criterion	0.874162
Log likelihood	-32.77545	Hannan-Quinn criter.	0.856324
Durbin-Watson stat	0.834789		

Performance

The model performance is actually bad: MMR reacts to changes in RMP with delay, causing increase/decrease in spread when RMP decreases/increases.



Adjusted model

Potential solution to the problem of basic ECM is adding changes of RMP as regressor.

- To address the problem we need the additional term to have current timing.

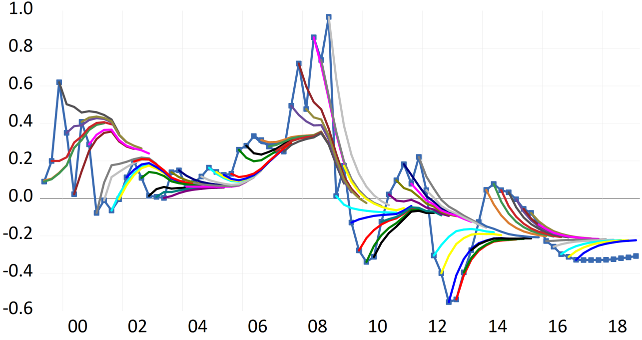
Variable	Coefficient	Std. Error	t-Statistic	Prob.	Standardized
FREURIBOR3M_IEUZN(-1)--0.219456323991033-1.140468...	-0.326752	0.083793	-3.899488	0.0002	-0.198827
@D(FRMP_IEUZN)	1.071884	0.061270	17.49436	0.0000	0.885212

R-squared	0.798110	Mean dependent var	-0.042495
Adjusted R-squared	0.795521	S.D. dependent var	0.367888
S.E. of regression	0.166357	Akaike info criterion	-0.724684
Sum squared resid	2.158611	Schwarz criterion	-0.665134
Log likelihood	30.98738	Hannan-Quinn criter.	-0.700809
Durbin-Watson stat	2.090817		

- The coefficient implies that increase in RMP by 100bps leads to increase in MMR by 107bps.

Adjusted model performance

The forecasts are very similar to model in levels with ARMA error. The adjusted ECM has slightly more persistence.



Interpreting adjusted model

When is it appropriate to adjust the standard ECM with current changes of driving variable?

Interpreting adjusted model

When is it appropriate to adjust the standard ECM with current changes of driving variable?

Crucial question: **Does the dependent variable react to changes in driving variable with a lag?**

- Lag reaction is reasonable in slow-moving macroeconomic variables (e.g. consumption).
- It is less appropriate in financial variables which do not have slow-moving nature (e.g. prices).

Interpreting adjusted model

When is it appropriate to adjust the standard ECM with current changes of driving variable?

Crucial question: **Does the dependent variable react to changes in driving variable with a lag?**

- Lag reaction is reasonable in slow-moving macroeconomic variables (e.g. consumption).
- It is less appropriate in financial variables which do not have slow-moving nature (e.g. prices).

This is also related to relationship between the two variables: current timing is appropriate when one variable is the driving the other variable but not vice versa.

Specifications in spread

Motivation: Practical considerations

All the previous models had spread vary with level of policy rates, reflecting specific feature of data which we would not want to replicate.

Motivation: Practical considerations

All the previous models had spread vary with level of policy rates, reflecting specific feature of data which we would not want to replicate.

Moreover, the ECM model showed that we need to ensure that MMR responds to changes in policy rates **immediately**.

Motivation: Practical considerations

All the previous models had spread vary with level of policy rates, reflecting specific feature of data which we would not want to replicate.

Moreover, the ECM model showed that we need to ensure that MMR responds to changes in policy rates **immediately**.

It seems that it might be better idea to model MMR in terms of spread. Modelling in terms of spread will ensure...

Motivation: Practical considerations

All the previous models had spread vary with level of policy rates, reflecting specific feature of data which we would not want to replicate.

Moreover, the ECM model showed that we need to ensure that MMR responds to changes in policy rates **immediately**.

It seems that it might be better idea to model MMR in terms of spread. Modelling in terms of spread will ensure...

- ...stability of spread in long-run.

Motivation: Practical considerations

All the previous models had spread vary with level of policy rates, reflecting specific feature of data which we would not want to replicate.

Moreover, the ECM model showed that we need to ensure that MMR responds to changes in policy rates **immediately**.

It seems that it might be better idea to model MMR in terms of spread. Modelling in terms of spread will ensure...

- ...stability of spread in long-run.
- ...immediate response of MMR to changes in policy rates.

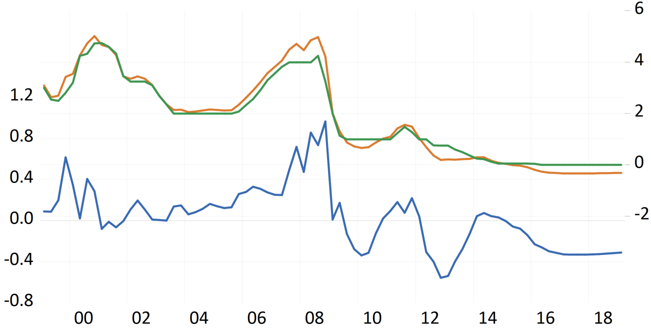
Motivation: Theoretical considerations

Modelling spread is especially appropriate since the Euribor is likely determined in terms of spread: banks choose the rate at which they are willing to lend *as function of* their own funding costs (i.e. policy rates).

Motivation: Theoretical considerations

Modelling spread is especially appropriate since the Euribor is likely determined in terms of spread: banks choose the rate at which they are willing to lend *as function of* their own funding costs (i.e. policy rates).

- While level of the two series is quite variable, the spread seems to be relatively stable.



Model

Starting point is simple ARMA model in spread.

- The shocks to spread are very persistent.
- Despite simplicity model explain large share of variation in spread.

Dependent Variable: FREURIBOR3M_IEUZN-FRMP_IEUZN
Method: ARMA Maximum Likelihood (BFGS)
Date: 06/18/19 Time: 12:19
Sample: 1999Q1 2019Q1
Included observations: 81
Convergence achieved after 5 iterations
Coefficient covariance computed using outer product of gradients

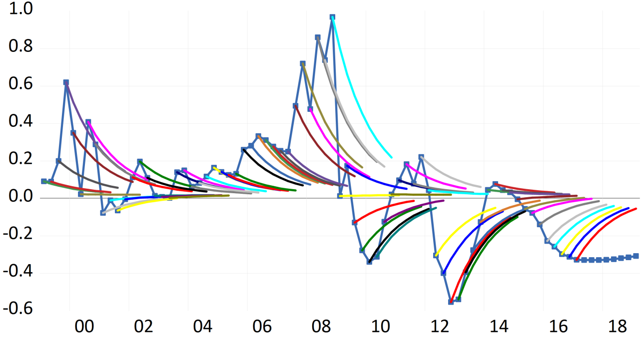
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.017899	0.148245	0.120738	0.9042
AR(1)	0.822690	0.055584	14.80088	0.0000
SIGMASQ	0.029570	0.003142	9.410780	0.0000

R-squared	0.682466	Mean dependent var	0.032408
Adjusted R-squared	0.674324	S.D. dependent var	0.307064
S.E. of regression	0.175235	Akaike info criterion	-0.595098
Sum squared resid	2.395172	Schwarz criterion	-0.506415
Log likelihood	27.10147	Hannan-Quinn criter.	-0.559517
F-statistic	83.82151	Durbin-Watson stat	2.142462
Prob(F-statistic)	0.000000		

Performance

The model performance is not bad, but is worse than previous models.

- 2-year RMSE is 0.243.



Model with stress regressors

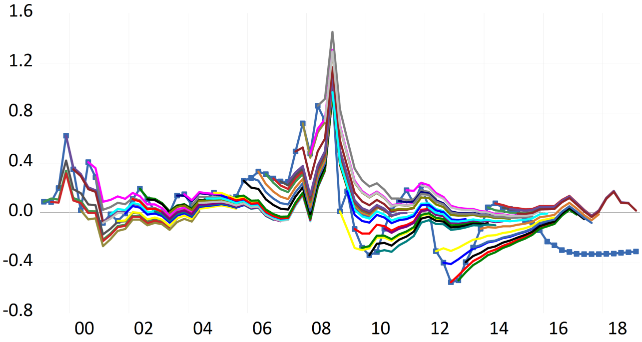
We can stress regressors to improve the model.

Variable	Coefficient	Std. Error	t-Statistic	Prob.	Standardized
C	-0.121424	0.121701	-0.997722	0.3215	
FRILIBOR3M_US-FRFED_US	0.470103	0.030670	15.32774	0.0000	0.462339
AR(1)	0.881182	0.062339	14.13526	0.0000	
SIGMASQ	0.016911	0.001979	8.544712	0.0000	

R-squared	0.818407	Mean dependent var	0.032408
Adjusted R-squared	0.811332	S.D. dependent var	0.307064
S.E. of regression	0.133376	Akaike info criterion	-1.124672
Sum squared resid	1.369762	Schwarz criterion	-1.006428
Log likelihood	49.54923	Hannan-Quinn criter.	-1.077231
F-statistic	115.6752	Durbin-Watson stat	1.708918
Prob(F-statistic)	0.000000		
Inverted AR Roots	.88		

Performance with stress regressors

Stress regressors help with the forecasting performance, but lead to worsening of forecast performance in recent periods.



Transformation of stress regressor

One issue with using level of the stress regressor is that if there are long-term changes in the level of the regressor then they will change the long-term forecasts for our dependent variable.

- For example, the natural level of the Libor spread seems to have changed with change in conduct of monetary policy.

Transformation of stress regressor

One issue with using level of the stress regressor is that if there are long-term changes in the level of the regressor then they will change the long-term forecasts for our dependent variable.

- For example, the natural level of the Libor spread seems to have changed with change in conduct of monetary policy.

  $\Rightarrow$ It is typically **better to use changes of the stress regressor.**

Transformation of stress regressor

One issue with using level of the stress regressor is that if there are long-term changes in the level of the regressor then they will change the long-term forecasts for our dependent variable.

- For example, the natural level of the Libor spread seems to have changed with change in conduct of monetary policy.

  $\Rightarrow$ It is typically **better to use changes of the stress regressor.**

In principle we could use changes of the stress regressor in our equation with level of spread between MMR and MPR. However, this will lead to bad shock profiles.

Transformation of stress regressor

One issue with using level of the stress regressor is that if there are long-term changes in the level of the regressor then they will change the long-term forecasts for our dependent variable.

- For example, the natural level of the Libor spread seems to have changed with change in conduct of monetary policy.

  $\Rightarrow$ It is typically **better to use changes of the stress regressor.**

In principle we could use changes of the stress regressor in our equation with level of spread between MMR and MPR. However, this will lead to bad shock profiles.

- Since shocks to stress regressors are mean reverting, an increase is followed by decrease.

Transformation of stress regressor

One issue with using level of the stress regressor is that if there are long-term changes in the level of the regressor then they will change the long-term forecasts for our dependent variable.

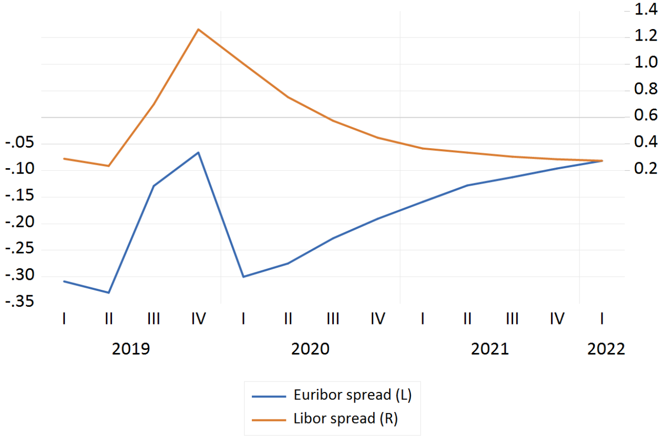
- For example, the natural level of the Libor spread seems to have changed with change in conduct of monetary policy.

  $\Rightarrow$ It is typically **better to use changes of the stress regressor.**

In principle we could use changes of the stress regressor in our equation with level of spread between MMR and MPR. However, this will lead to bad shock profiles.

- Since shocks to stress regressors are mean reverting, an increase is followed by decrease.
- This profile will translate into increase in dependent varaible followed by rapid decrease, much faster than that of shock regressor.

Transformation of stress regressor: Illustration



Transformation of stress regressor: Solution

A solution to our problem is changing the model structure from ARMA in level of spread to ECM in spread:

$$d(MMR_t - RMP_t) = \beta(MMR_{t-1} - RMP_{t-1} - \gamma_0) \tag{1}$$

Note: γ_0 will be our estimate of natural level of spread, which is better to estimate rather than to force to be equal to average.

Transformation of stress regressor: Solution

A solution to our problem is changing the model structure from ARMA in level of spread to ECM in spread:

$$d(MMR_t - RMP_t) = \beta(MMR_{t-1} - RMP_{t-1} - \gamma_0) \tag{1}$$

Note: γ_0 will be our estimate of natural level of spread, which is better to estimate rather than to force to be equal to average.

Equation (1) is basically equivalent to ARMA(1,0) in level of spread.

- The only difference is in estimation due to non-linearity in coefficients.

Transformation of stress regressor: Solution

A solution to our problem is changing the model structure from ARMA in level of spread to ECM in spread:

$$d(MMR_t - RMP_t) = \beta(MMR_{t-1} - RMP_{t-1} - \gamma_0) \quad (1)$$

Note: γ_0 will be our estimate of natural level of spread, which is better to estimate rather than to force to be equal to average.

Equation (1) is basically equivalent to ARMA(1,0) in level of spread.

- The only difference is in estimation due to non-linearity in coefficients.

Equation (1) has one main advantage over ARMA model in levels of spread: we can use the changes in stress regressors.

ECM in spread

This model has following properties:

- The spread does not vary with level of policy rate.
- The spread comoves with Libor spread in short run.
- The spread does not depend on Libor spread in long-run.

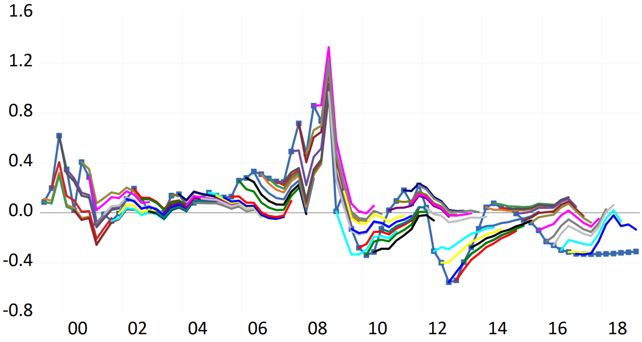
Dependent Variable: @D(FREURIBOR3M_IEUZN-FRMP_IEUZN)
Method: Least Squares (Gauss-Newton / Marquardt steps)
Date: 06/19/19 Time: 09:21
Sample (adjusted): 1999Q2 2019Q1
Included observations: 80 after adjustments
Convergence achieved after 3 iterations
Coefficient covariance computed using outer product of gradients
@D(FREURIBOR3M_IEUZN-FRMP_IEUZN)= C(2)*(FREURIBOR3M_IEU
ZN(-1)-FRMP_IEUZN(-1)-C(1)) + C(3)*@D(FRILIBOR3M_US
-FRFED_US)

	Coefficient	Std. Error	t-Statistic	Prob.
C(2)	-0.151860	0.047121	-3.222790	0.0019
C(1)	0.001868	0.095074	0.019644	0.9844
C(3)	0.464009	0.056127	8.267204	0.0000

R-squared	0.512633	Mean dependent var	-0.004995
Adjusted R-squared	0.499975	S.D. dependent var	0.181437
S.E. of regression	0.128299	Akaike info criterion	-1.232129
Sum squared resid	1.267467	Schwarz criterion	-1.142803
Log likelihood	52.28516	Hannan-Quinn criter.	-1.196316
Durbin-Watson stat	1.693214		

Performance

The model is best out of models in terms of spread. While its overall performance is not perfect, it works very well during crisis periods.



Stationarity of spread

Is the spread really stationary?

Stationarity of spread

Is the spread really stationary?

- Crisis periods show clear sign of stationarity, with spread returning towards average value after period of deviation.

Stationarity of spread

Is the spread really stationary?

- Crisis periods show clear sign of stationarity, with spread returning towards average value after period of deviation.
- There is a clear break between pre-2008 and post-2008 average level.

Stationarity of spread

Is the spread really stationary?

- Crisis periods show clear sign of stationarity, with spread returning towards average value after period of deviation.
- There is a clear break between pre-2008 and post-2008 average level.
- There does not seem to be any 'normal' value after 2008.

Stationarity of spread

Is the spread really stationary?

- Crisis periods show clear sign of stationarity, with spread returning towards average value after period of deviation.
- There is a clear break between pre-2008 and post-2008 average level.
- There does not seem to be any 'normal' value after 2008.

Formal tests are inconclusive: ADF suggested stationarity but KPSS suggests non-stationarity.

Stationarity of spread

Is the spread really stationary?

- Crisis periods show clear sign of stationarity, with spread returning towards average value after period of deviation.
- There is a clear break between pre-2008 and post-2008 average level.
- There does not seem to be any 'normal' value after 2008.

Formal tests are inconclusive: ADF suggested stationarity but KPSS suggests non-stationarity.

Possible interpretation: Spread was stationary before 2008 but not after.

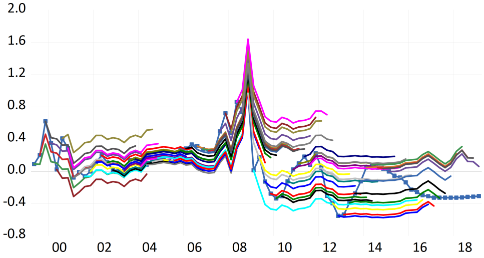
Non-stationary model for spread

Given the apparent non-stationarity we could consider modelling spread in differences without error-correction term.

Non-stationary model for spread

Given the apparent non-stationarity we could consider modelling spread in differences without error-correction term.

In practice this will have the same drawbacks of non-stationary models in level of Euribor: the long-run spread will be unstable.



Model with excess reserves

Motivations

None of the models so far was fully satisfactory. From the perspective of behaviour of spread the models were either...

Motivations

None of the models so far was fully satisfactory. From the perspective of behaviour of spread the models were either...

1. ...causing spread to revert to historical average value when it was not appropriate (e.g. 2016-2018)

Motivations

None of the models so far was fully satisfactory. From the perspective of behaviour of spread the models were either...

1. ...causing spread to revert to historical average value when it was not appropriate (e.g. 2016-2018)
2. ...causing natural level of spread to change with shocks when it was not appropriate (e.g. 2007-2008)

Motivations

None of the models so far was fully satisfactory. From the perspective of behaviour of spread the models were either...

1. ...causing spread to revert to historical average value when it was not appropriate (e.g. 2016-2018)
2. ...causing natural level of spread to change with shocks when it was not appropriate (e.g. 2007-2008)

In econometric terminology the models were either imposing too much or not enough stationarity. In reality, there seem to be shocks that are purely transitory, but also small changes in the natural level of spread.

Motivations

None of the models so far was fully satisfactory. From the perspective of behaviour of spread the models were either...

1. ...causing spread to revert to historical average value when it was not appropriate (e.g. 2016-2018)
2. ...causing natural level of spread to change with shocks when it was not appropriate (e.g. 2007-2008)

In econometric terminology the models were either imposing too much or not enough stationarity. In reality, there seem to be shocks that are purely transitory, but also small changes in the natural level of spread.

Key idea: Could we somehow account for the small changes in natural level of spread?

Role of excess reserves

A prime candidate for factor driving the changes in natural level of spread is amount of excess reserves.

Role of excess reserves

A prime candidate for factor driving the changes in natural level of spread is amount of excess reserves.

- Amount of excess reserves is controlled by ECB.

Role of excess reserves

A prime candidate for factor driving the changes in natural level of spread is amount of excess reserves.

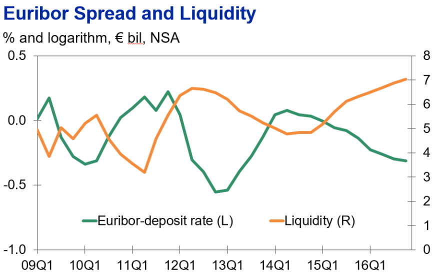
- Amount of excess reserves is controlled by ECB.
- Changes in amount of excess reserves change the supply/demand in the interbank lending market, influencing MMR.

Role of excess reserves

A prime candidate for factor driving the changes in natural level of spread is amount of excess reserves.

- Amount of excess reserves is controlled by ECB.
- Changes in amount of excess reserves change the supply/demand in the interbank lending market, influencing MMR.

Simple data analysis confirms this intuition.



Multiple regimes

There are two additional problems we need to address.

Multiple regimes

There are two additional problems we need to address.

1. Prior to 2008 there were no excess reserves in the interbank system and hence they did not play role in determining MMR.

Multiple regimes

There are two additional problems we need to address.

1. Prior to 2008 there were no excess reserves in the interbank system and hence they did not play role in determining MMR.
2. Prior to 2008 the MMR was always above RMP, but after 2008 it was mostly below RMP. Moreover, after 2008 it was not responding to change in RMP, only to changes in DR.

Multiple regimes

There are two additional problems we need to address.

1. Prior to 2008 there were no excess reserves in the interbank system and hence they did not play role in determining MMR.
2. Prior to 2008 the MMR was always above RMP, but after 2008 it was mostly below RMP. Moreover, after 2008 it was not responding to change in RMP, only to changes in DR.

These observations suggest that there are two regimes:

Multiple regimes

There are two additional problems we need to address.

1. Prior to 2008 there were no excess reserves in the interbank system and hence they did not play role in determining MMR.
2. Prior to 2008 the MMR was always above RMP, but after 2008 it was mostly below RMP. Moreover, after 2008 it was not responding to change in RMP, only to changes in DR.

These observations suggest that there are two regimes:

1. Normal regime: No excess reserves and RMP is anchored by main refinancing rate.

Multiple regimes

There are two additional problems we need to address.

1. Prior to 2008 there were no excess reserves in the interbank system and hence they did not play role in determining MMR.
2. Prior to 2008 the MMR was always above RMP, but after 2008 it was mostly below RMP. Moreover, after 2008 it was not responding to change in RMP, only to changes in DR.

These observations suggest that there are two regimes:

1. Normal regime: No excess reserves and RMP is anchored by main refinancing rate.
2. Excess reserves regime: Significant excess reserves and MMR anchored by deposit rate.

Model: Equilibrium component

We will first consider equilibrium component (i.e. no-stress component).

$$MMR^{eq} = \begin{cases} \beta_{01} + RMP_t + \gamma_0 & \text{if } ER < \tau \\ \beta_{02} + DR_t + \beta_{22} log(ER_t) & \text{if } ER > \tau \end{cases}$$

Model: Equilibrium component

We will first consider equilibrium component (i.e. no-stress component).

$$MMR^{eq} = \begin{cases} \beta_{01} + RMP_t + \gamma_0 & \text{if } ER < \tau \\ \beta_{02} + DR_t + \beta_{22} log(ER_t) & \text{if } ER > \tau \end{cases}$$

- When excess reserves are below threshold then equilibrium spread is constant β_{01}.

Model: Equilibrium component

We will first consider equilibrium component (i.e. no-stress component).

$$MMR^{eq} = \begin{cases} \beta_{01} + RMP_t + \gamma_0 & \text{if } ER < \tau \\ \beta_{02} + DR_t + \beta_{22} log(ER_t) & \text{if } ER > \tau \end{cases}$$

- When excess reserves are below threshold then equilibrium spread is constant β_{01}.
- When excess reserves are above threshold then equilibrium spread varies with logarithm of excess reserves.

Model: Equilibrium component

We will first consider equilibrium component (i.e. no-stress component).

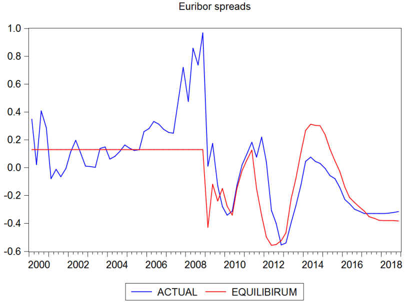
$$MMR^{eq} = \begin{cases} \beta_{01} + RMP_t + \gamma_0 & \text{if } ER < \tau \\ \beta_{02} + DR_t + \beta_{22} log(ER_t) & \text{if } ER > \tau \end{cases}$$

- When excess reserves are below threshold then equilibrium spread is constant β_{01}.
- When excess reserves are above threshold then equilibrium spread varies with logarithm of excess reserves.
- The anchor is different in the two regimes.

Performance: Equilibrium component

There are two key takeaways:

1. The equilibrium component can account for most of the variation in the spread after 2008.
2. The equilibrium component does not imply return to historical average in 2016-2018 period.



Full model

The full model can then estimated as ECM with stress regressors:

$$
\begin{aligned}
&d(MMR_t - RMP_t) = \\
&\quad \beta_1(MMR_{t-1} - RMP_{t-1} - (MMR_t^{eq} - RMP_t))+ \\
&\quad + \beta_2 d(Libor_t^{US} - RMP_t^{US}) + \beta_3 d(Libor_t^{US} - RMP_t^{US}) * DUM^{euro\_crisis} \\
&\quad + \beta_4 d(RMP_t - DR_t) * DUM^{ER}
\end{aligned}
$$

Full model

The full model can then estimated as ECM with stress regressors:

$$
\begin{aligned}
&d(MMR_t - RMP_t) = \\
&\quad \beta_1(MMR_{t-1} - RMP_{t-1} - (MMR_t^{eq} - RMP_t)) + \\
&\quad + \beta_2 d(Libor_t^{US} - RMP_t^{US}) + \beta_3 d(Libor_t^{US} - RMP_t^{US}) * DUM^{euro\_crisis} \\
&\quad + \beta_4 d(RMP_t - DR_t) * DUM^{ER}
\end{aligned}
$$

- Compared to (1) we use estimates of time-varying equilibrium spread instead of estimated constant spread.

Full model

The full model can then estimated as ECM with stress regressors:

$$
\begin{aligned}
& d(MMR_t - RMP_t) = \\
& \quad \beta_1(MMR_{t-1} - RMP_{t-1} - (MMR_t^{eq} - RMP_t)) + \\
& \quad + \beta_2 d(Libor_t^{US} - RMP_t^{US}) + \beta_3 d(Libor_t^{US} - RMP_t^{US}) * DUM^{euro\_crisis} \\
& \quad + \beta_4 d(RMP_t - DR_t) * DUM^{ER}
\end{aligned}
$$

- Compared to (1) we use estimates of time-varying equilibrium spread instead of estimated constant spread.
- We use current period equilibrium spread to allow for immediate (partial) effect of changes in policy affecting equilibrium spread.

Full model

The full model can then estimated as ECM with stress regressors:

$$
\begin{aligned}
& d(MMR_t - RMP_t) = \\
& \quad \beta_1(MMR_{t-1} - RMP_{t-1} - (MMR_t^{eq} - RMP_t)) + \\
& \quad + \beta_2 d(Libor_t^{US} - RMP_t^{US}) + \beta_3 d(Libor_t^{US} - RMP_t^{US}) * DUM^{euro\_crisis} \\
& \quad + \beta_4 d(RMP_t - DR_t) * DUM^{ER}
\end{aligned}
$$

- Compared to (1) we use estimates of time-varying equilibrium spread instead of estimated constant spread.
- We use current period equilibrium spread to allow for immediate (partial) effect of changes in policy affecting equilibrium spread.
- We include US Libor spread as regressor, allowing for effect to vary between normal times and euro zone crisis periods.

Full model

The full model can then estimated as ECM with stress regressors:

$$
\begin{aligned}
&d(MMR_t - RMP_t) = \\
&\quad \beta_1(MMR_{t-1} - RMP_{t-1} - (MMR_t^{eq} - RMP_t)) + \\
&\quad + \beta_2 d(Libor_t^{US} - RMP_t^{US}) + \beta_3 d(Libor_t^{US} - RMP_t^{US}) * DUM^{euro\_crisis} \\
&\quad + \beta_4 d(RMP_t - DR_t) * DUM^{ER}
\end{aligned}
$$

- Compared to (1) we use estimates of time-varying equilibrium spread instead of estimated constant spread.
- We use current period equilibrium spread to allow for immediate (partial) effect of changes in policy affecting equilibrium spread.
- We include US Libor spread as regressor, allowing for effect to vary between normal times and euro zone crisis periods.
- We include changes in spread between policy rates so that they have immediate effect.

Full model: Estimates

All the coefficients have expected sign and are highly significant. We are able to 78% of variation in *changes* of spread.

Dependent Variable: D(FREURIBOR3M_I_IEUZN-FRMP_IEUZN)
Method: Least Squares
Date: 03/29/19 Time: 22:06
Sample (adjusted): 2001Q1 2018Q3
Included observations: 71 after adjustments

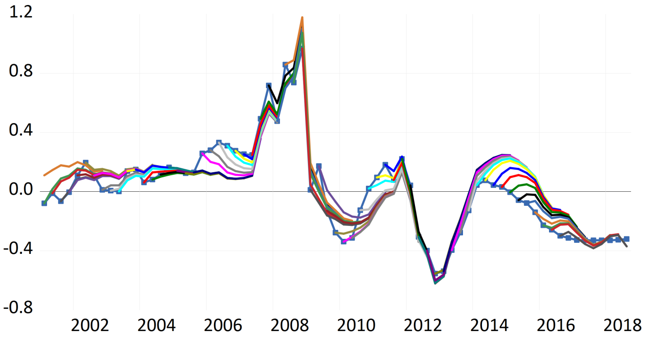
Variable	Coefficient	Std. Error	t-Statistic	Prob.	Standardized
(FREURIBOR3M_I_IEUZN(-1)-FRMP_IEUZN(-1))-(FRSPR...	-0.311205	0.049075	-6.341391	0.0000	-0.477797
D(FRILIBOR3M_US-FRFED_US)	0.268342	0.043676	6.143923	0.0000	0.405607
D(FRILIBOR3M_US-FRFED_US)*DUM_SOVCRISIS_IEUZN	1.052950	0.334296	3.149753	0.0024	0.183358
D(FRMP_IEUZN-FRMDEP_IEUZN)*DUM_EXRES_IEUZN	-0.641266	0.243916	-2.629043	0.0106	-0.178705

R-squared	0.778316	Mean dependent var	-0.008545
Adjusted R-squared	0.768390	S.D. dependent var	0.171833
S.E. of regression	0.082696	Akaike info criterion	-2.092600
Sum squared resid	0.458189	Schwarz criterion	-1.965126
Log likelihood	78.28732	Hannan-Quinn criter.	-2.041908
Durbin-Watson stat	1.549771		

Full model: Performance

The model performance is far better than alternatives (2-year RMSE=0.11). In addition,...

- ...it implies stable long-run spread.
- ...it does not lead to mean-reversion in recent period.
- ...it does not feature forecast jumps.



Lessons learned

Static vs. dynamic models

Key drawback of static models is that the foretasted variable returns to predicted relationship immediately.

- Static specifications ignore information contained in last historical observation.

Static vs. dynamic models

Key drawback of static models is that the foretasted variable returns to predicted relationship immediately.

- Static specifications ignore information contained in last historical observation.

This drawback is particularly problematic in multiple variable models where the forecast jumps will propagated throughout the model.

Static vs. dynamic models

Key drawback of static models is that the foretasted variable returns to predicted relationship immediately.

- Static specifications ignore information contained in last historical observation.

This drawback is particularly problematic in multiple variable models where the forecast jumps will propagated throughout the model.

Dynamic models offer solution to this problem.

Variations of dynamic models

There are different possible dynamic models.

Variations of dynamic models

There are different possible dynamic models.

First distinction is between models using lagged-dependent variable and models using ARMA errors.

- ARMA errors make forecast errors persistent, while LDV models make actual dependent variable persistent.

Variations of dynamic models

There are different possible dynamic models.

First distinction is between models using lagged-dependent variable and models using ARMA errors.

- ARMA errors make forecast errors persistent, while LDV models make actual dependent variable persistent.

Meanwhile, model using changes in dependent variable ensures persistence in level of given variable, but will lead to instability of long-term level.

Variations of dynamic models

There are different possible dynamic models.

First distinction is between models using lagged-dependent variable and models using ARMA errors.

- ARMA errors make forecast errors persistent, while LDV models make actual dependent variable persistent.

Meanwhile, model using changes in dependent variable ensures persistence in level of given variable, but will lead to instability of long-term level.

Error-correction model is combination of dynamic level models and models in differences.

- In simplest form it is equivalent to LDV/ARMA models.
- It allows for inclusion of regressors without long-term effects.

Error-correction model vs spread models

Error-correction models ensure that given variables return to their equilibrium relationship eventually.

- Problem: gradual return is not appropriate when variables respond to movements in driving variables immediately.

Error-correction model vs spread models

Error-correction models ensure that given variables return to their equilibrium relationship eventually.

- Problem: gradual return is not appropriate when variables respond to movements in driving variables immediately.

One possible solution is use current changes of driving variables.

Error-correction model vs spread models

Error-correction models ensure that given variables return to their equilibrium relationship eventually.

- Problem: gradual return is not appropriate when variables respond to movements in driving variables immediately.

One possible solution is use current changes of driving variables.

Alternative solution: include the driving variable on the LHS in form of spread.

- This ensures immediate effect of movements in driving variable.
- Equivalent to assuming co-integrating vector (1,-1).

Error-correction model vs spread models

Error-correction models ensure that given variables return to their equilibrium relationship eventually.

- Problem: gradual return is not appropriate when variables respond to movements in driving variables immediately.

One possible solution is use current changes of driving variables.

Alternative solution: include the driving variable on the LHS in form of spread.

- This ensures immediate effect of movements in driving variable.
- Equivalent to assuming co-integrating vector (1,-1).

The spread itself can be modelled in error-correction form.

Sources of non-stationarity

This example also showed us crucial aspect of non-stationarity: often non-stationarity in particular series results from dependence on factors that are non-stationary.

Sources of non-stationarity

This example also showed us crucial aspect of non-stationarity: often non-stationarity in particular series results from dependence on factors that are non-stationary.

- After controlling for these factors the shocks to the series do not have permanent effect.

Sources of non-stationarity

This example also showed us crucial aspect of non-stationarity: often non-stationarity in particular series results from dependence on factors that are non-stationary.

- After controlling for these factors the shocks to the series do not have permanent effect.

  $\Rightarrow$ In this view there is no longer tension between econometric and practical views of non-stationarity.

Dealing with non-stationarity I

From model building point of view there is important lesson: we should aim to control for factors that are causing non-stationarity rather than transforming the series into stationary form.

Dealing with non-stationarity I

From model building point of view there is important lesson: we should aim to control for factors that are causing non-stationarity rather than transforming the series into stationary form.

When we did not account the sources of non-stationarity we were encountering **both** of the problems of modelling series wrongly.

Dealing with non-stationarity I

From model building point of view there is important lesson: we should aim to control for factors that are causing non-stationarity rather than transforming the series into stationary form.

When we did not account the sources of non-stationarity we were encountering **both** of the problems of modelling series wrongly.

- When we modelled in levels there was tendency to return to historical average value when there should not be any.

Dealing with non-stationarity I

From model building point of view there is important lesson: we should aim to control for factors that are causing non-stationarity rather than transforming the series into stationary form.

When we did not account the sources of non-stationarity we were encountering **both** of the problems of modelling series wrongly.

- When we modelled in levels there was tendency to return to historical average value when there should not be any.
- When we modelled in differences there was not enough tendency to return to historical average.

Dealing with non-stationarity I

From model building point of view there is important lesson: we should aim to control for factors that are causing non-stationarity rather than transforming the series into stationary form.

When we did not account the sources of non-stationarity we were encountering **both** of the problems of modelling series wrongly.

- When we modelled in levels there was tendency to return to historical average value when there should not be any.
- When we modelled in differences there was not enough tendency to return to historical average.

After accounting for sources of non-stationarity we no longer faced **either** of these problems.

Dealing with non-stationarity II

When should we model in differences?

Dealing with non-stationarity II

When should we model in differences?

- Simplistic view: model in differences if series are non-stationary.

Dealing with non-stationarity II

When should we model in differences?

- Simplistic view: model in differences if series are non-stationary.
- Co-integration view: model in differences when shocks to relationship between series are permanent.

Dealing with non-stationarity II

When should we model in differences?

- Simplistic view: model in differences if series are non-stationary.
- Co-integration view: model in differences when shocks to relationship between series are permanent.

In which situation are such shocks permanent? Typically, when we do not account for influence of some other non-stationary variable.

Dealing with non-stationarity II

When should we model in differences?

- Simplistic view: model in differences if series are non-stationary.
- Co-integration view: model in differences when shocks to relationship between series are permanent.

In which situation are such shocks permanent? Typically, when we do not account for influence of some other non-stationary variable.

- Theoretical example.

Dealing with non-stationarity II

When should we model in differences?

- Simplistic view: model in differences if series are non-stationary.
- Co-integration view: model in differences when shocks to relationship between series are permanent.

In which situation are such shocks permanent? Typically, when we do not account for influence of some other non-stationary variable.

- Theoretical example.
  - ▷ y_t depends on x_t, which is stationary, and w_t, which is non-stationary.
  - ▷ Shocks to w_t will have permanent effect on **relationship between levels** of y_t and x_t.

Dealing with non-stationarity II

When should we model in differences?

- Simplistic view: model in differences if series are non-stationary.
- Co-integration view: model in differences when shocks to relationship between series are permanent.

In which situation are such shocks permanent? Typically, when we do not account for influence of some other non-stationary variable.

- Theoretical example.
  - ▷ y_t depends on x_t, which is stationary, and w_t, which is non-stationary.
  - ▷ Shocks to w_t will have permanent effect on **relationship between levels** of y_t and x_t.
- Empirical example:
  - ▷ Quantity of money and prices should be proportional to each other, but changes in nominal interest rates affect this relationship.
  - ▷ Not including interest rates will force us to "wrongly" stationarize the series.