



# Long-Run & Downturn Estimates for Basel IRB Using UK Bureau Data

Credit Scoring and Credit Control XVI, Aug 2019



Vassilis Ioannou  
Regulatory Analytics, Equifax

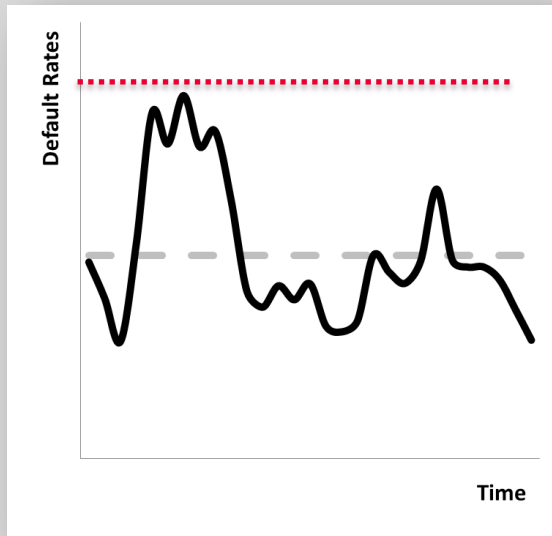
# Agenda



- **Introduction**  
*Basel Credit Risk Model, Regulatory Guidelines for IRB estimates*
- **Long-Run Probability of Default**  
*Methodological approach & key learnings using UK Industry Data*
- **Conclusions**

# The Basel credit risk model

**Lenders are required to provide sufficient capital to cover future losses limiting the overall systemic risk under stressed conditions.** Basel Credit Risk model makes use of “average” input PDs and “downturn” LGD to estimate unexpected losses under adverse economic conditions.



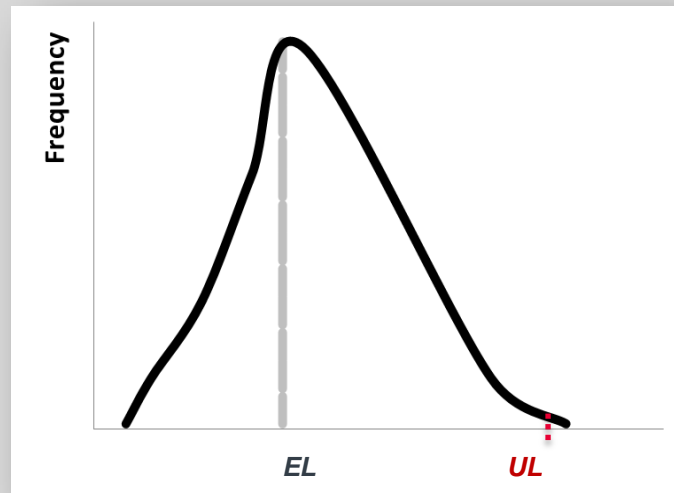
## Conditional PD

- ASRF Model
- Conservative

## “Average” PD

- Long-run
- Normal business conditions

## UL: “Downturn” LGD & Conditional PD



EL: Expected Losses

# Regulatory Guidelines for IRB estimates

**Even though the nature of the underlying model and the required estimates have not changed since 2008**, when European banks were first permitted to apply the IRB approach for Credit Risk, regulatory guidelines have been evolving towards more specific requirements for the data and techniques used to quantify PD and LGD.

## **CRR, Structure of Rating Systems, Article 170**

...(a) obligor and transaction risk...  
...(b) risk differentiation ...meaningful quantification

## **CRR, Risk Quantification Article 179**

...(d) the population of exposures ...comparable with those of the institution's exposures and standards...  
**The economic or market conditions ...relevant to current and foreseeable conditions...The number of exposures ...used for quantification shall be sufficient ...and robustness of its estimates:**  
...(f) margin of conservatism ...

## **CRR, Risk Quantification Article 180**

...(a) institutions shall estimate PDs by obligor grade from **long run averages of one-year default rates**...  
...(d) institutions shall use PD estimation techniques only with supporting analysis...  
...(h) irrespective of whether an institution is using external, internal, or pooled data sources, or a combination of the three, for its PD estimation, **the length of the underlying historical observation period used shall be at least five years** for at least one source

## **CRR, Use of models, Article 174**

...(c) the data used to build the model shall be **representative of the population of the institution's actual obligors or exposures;**

## **CRR, Competent authorities' assessment of an application to use an IRB Approach, Article 174**

...2. **EBA shall develop draft regulatory technical standards** to specify the assessment methodology competent authorities shall follow in assessing the compliance of an institution with the requirements to use the IRB Approach.

# Regulatory Guidelines for IRB estimates

**Even though the nature of the underlying model and the required estimates have not changed since 2008**, when European banks were first permitted to apply the IRB approach for Credit Risk, regulatory guidelines have been evolving towards more specific requirements for the data and techniques used to quantify PD and LGD.

## **Evolving Guidelines, Data for Modelling:**

- *Use of external data alongside internal data for Mortgages (UK PRA)*
  - Scorecard Development
  - PD Quantification
  - Elements of LGD Quantification
- *Clearer definition of data representativeness (EBA)*
  - Scope of application
  - Default Definition
  - Distribution of risk characteristics
  - Current and foreseeable economic / market conditions
  - Lending standards & recovery policies

## **Evolving Guidelines, Risk Quantification:**

- *PD and Economic Conditions*
  - *Mortgages Long-run PD reflecting conditions since 1990+, representative mix of good and bad years (UK PRA)*
  - *Direct PD estimation still not preferred approach vs averages, but external data & statistical methods can be considered for back casting to include historical periods into the final estimates.*
- *LGD and Economic Conditions*
  - *Specification of nature, severity and duration of economic downturns (EBA)*
  - *Direct modelling of LGD possible when not sufficient internal data, with increased Margins of Conservatism (EBA)*

## **Evolving Guidelines, Risk Quantification:**

- *Rating Systems Philosophy (UK PRA)*
  - *Cyclicality% Measure*
  - *Risk Drivers selection*
  - *Balance PiT vs TTC*
  - *Internal Awareness*
  - *Empirical evidence*
- *Final IRB estimates (EBA)*
  - *Deficiencies identification*
  - *Required Adjustments*
  - *Margin of Conservatism*

# The value of UK bureau data for IRB purposes

**PRA's Supervisory Statement S11/13** permits firms with limited internal experience of defaults to use **external alongside internal data to develop rating systems for mortgages with improved rank-ordering**.

For the estimation of the long-run PD of mortgage loans, firms are also permitted to **use external arrears data to infer default rates per PD rating / pool reflecting economic conditions** experienced historically since 1990.

When using external data, the margins of conservatism applied should also consider evidence about the **comparability** of the third-party data, potential **differences in the default** definition and the **overall representativeness** of the modelling samples.

**Bureau data in the UK reach back to 1980s** and **can be used by its closed group members to develop scorecards**, subject to the closed group rules, GDPR and other data related regulations.

The **arrears reported directly from the closed group members abide to common rules** described by the industry document 'Principles for the Reporting of Arrears, Arrangements and Defaults at CRAs'.

Among the **permitted uses** and with the constraints described in these regulations, **UK CRA data can be used for:**

- **Statistical analysis** (not to be shared outside the closed group)
- **Benchmarking** against selected set of peers
- **Scorecard Development** on pooled data

# The value of UK bureau data for IRB purposes

As previously mentioned, **statistical analysis with CRA data is permitted but cannot be shared outside the closed group.**

We will instead discuss our **learnings** applying a methodology that attempts to capture the relevance of Credit Bureau data when estimating **Mortgages long-run PD.**

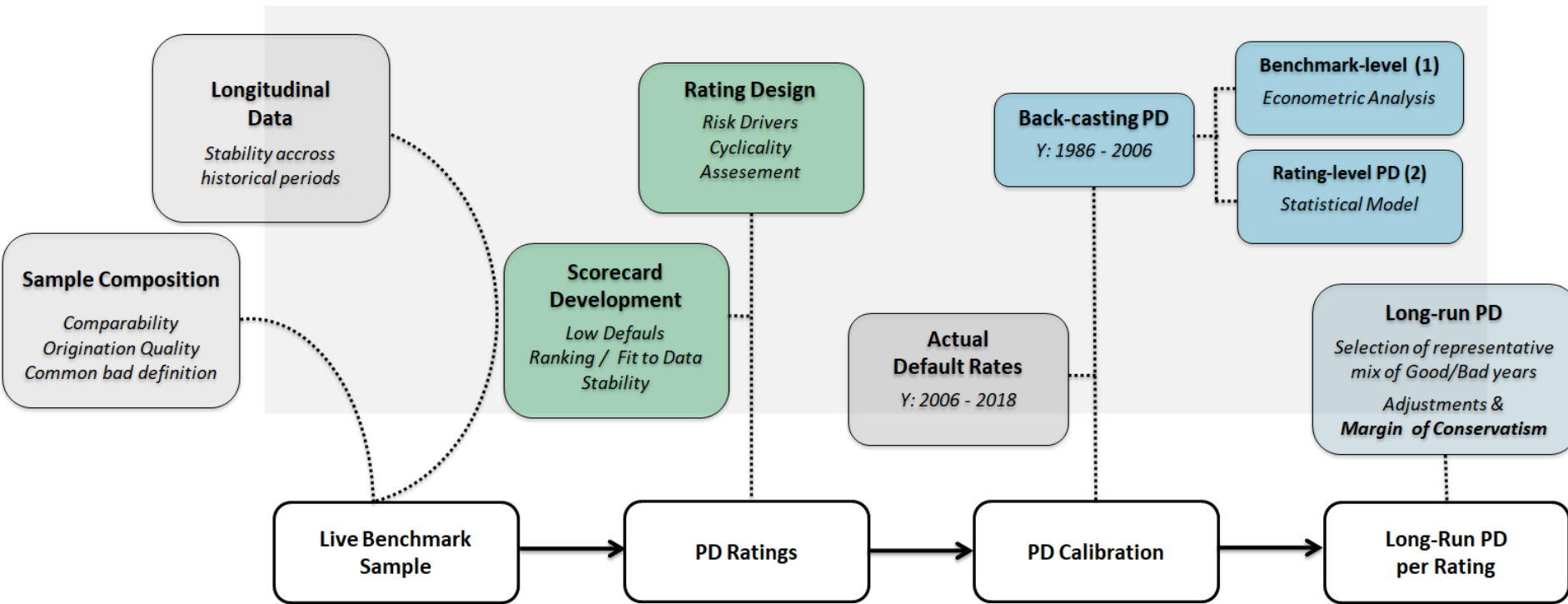
This case study makes use of anonymised data compliant to all applicable data regulations, to create a mortgages benchmark sample with relatively stable composition across periods.

**For this benchmark sample we have created a PD rating system and implemented a statistical methodology to back-cast PD per rating class back to 1986.**

The methodology followed and key learnings are summarised in the following slides.

The steps described in this presentation aim to highlight the value of bureau data. As part of an IRB project many additional considerations would have to be addressed including the ones necessary to justify the use of internal data as a primary source (PRA SS11/13, 12.38), comparability to external data, proposed adjustments and margins of conservatism.

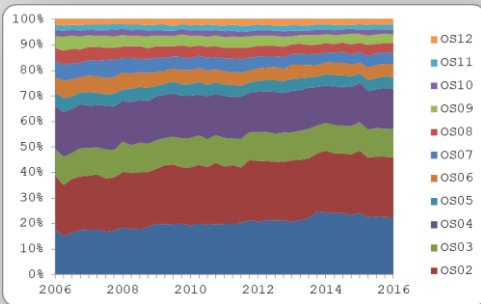
# Methodological Approach



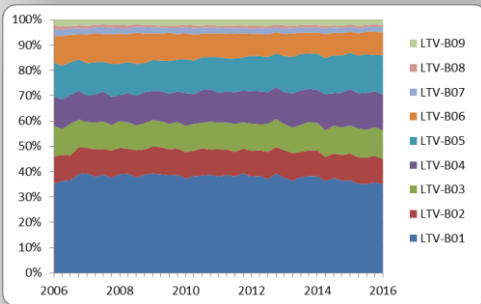
# Benchmark Sample

## Sample composition vs time

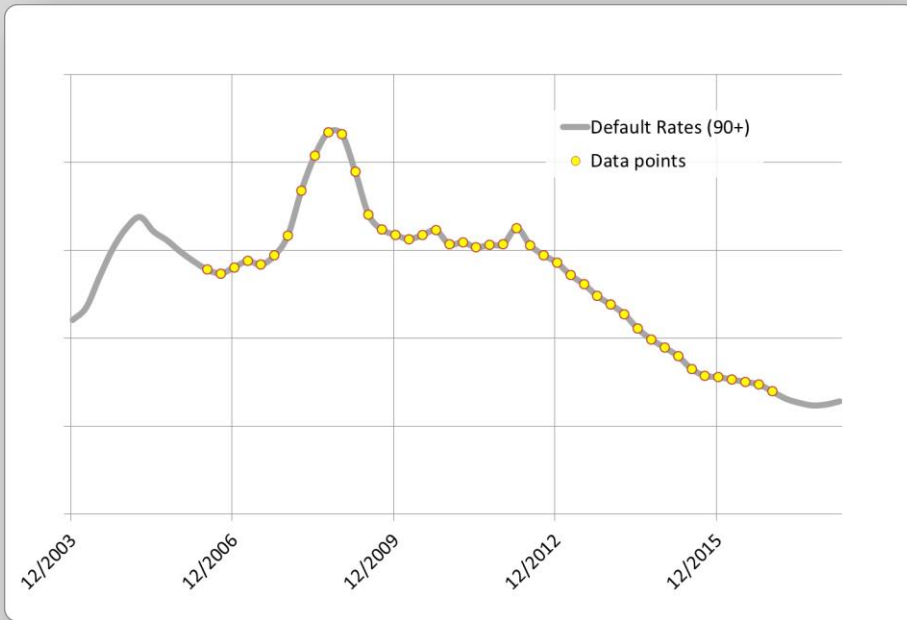
Obligor Score Pop%



LTV proxy Pop%



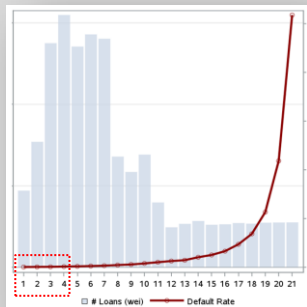
## Account level data



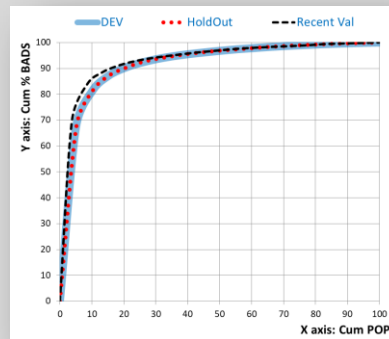
## Extended time series data

# PD Ratings – Model Development

## Final Risk Grades

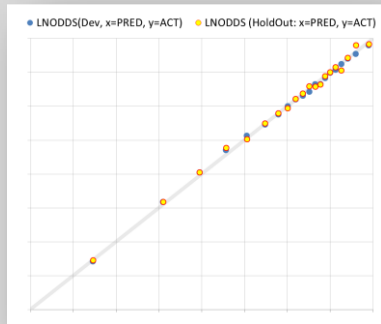


## Strong Ranking Performance



**Gini > 80**  
**KS > 68**

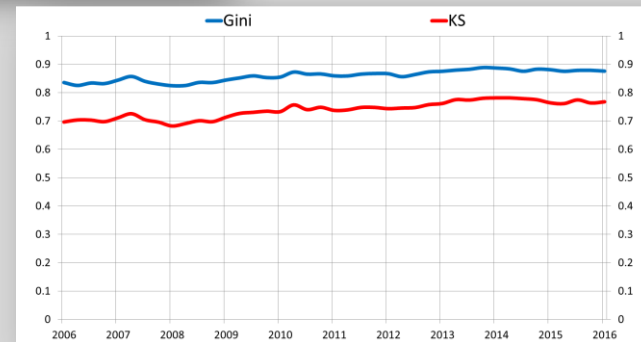
## Scorecard fit to the Data



**RMSE**  
**<0.12%**

*\*By scoreband  
@ Holdout*

## Across periods



# Backcasting Mortgages PD

Our backcasting approach, used to estimate conditional PD per rating across historical periods, is implemented in two steps:

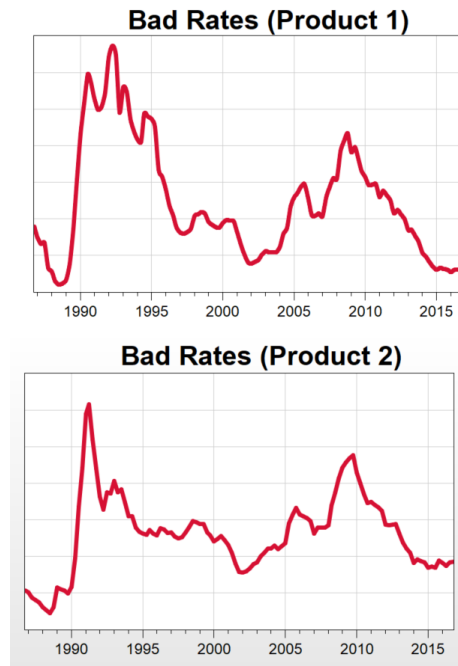
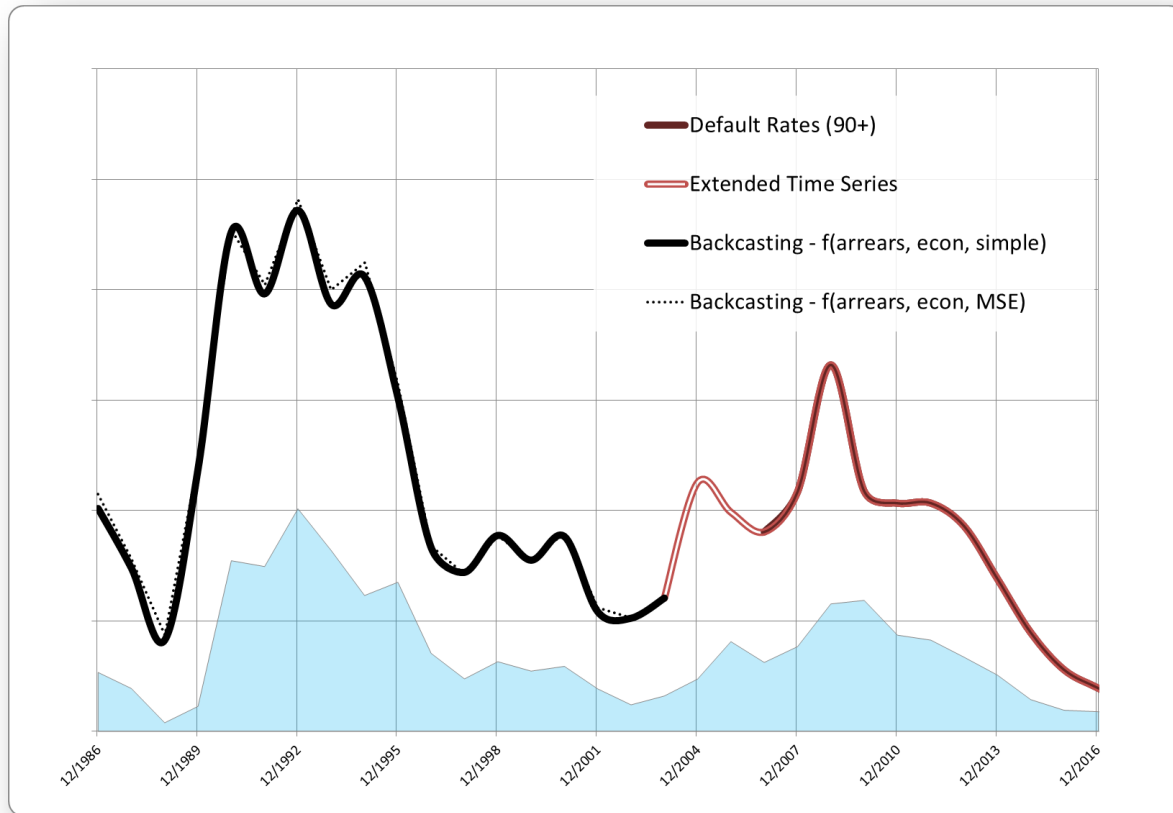
## Step (1): Time-series analysis

- A collection of linear regression models is developed, linking Benchmark portfolio DR to other bureau arrears data and economic variables
- The produced forecasts (best model, simple avg, MSE weighted) are used in step (2)

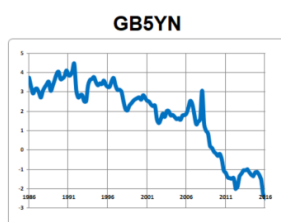
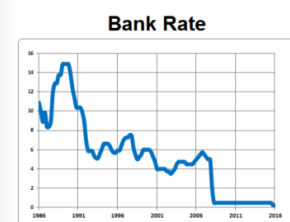
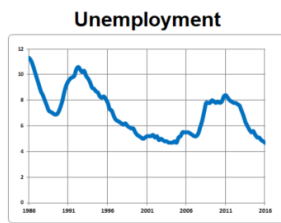
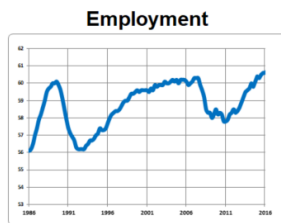
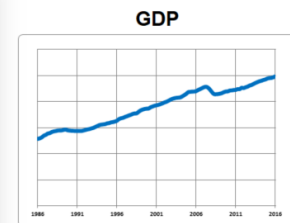
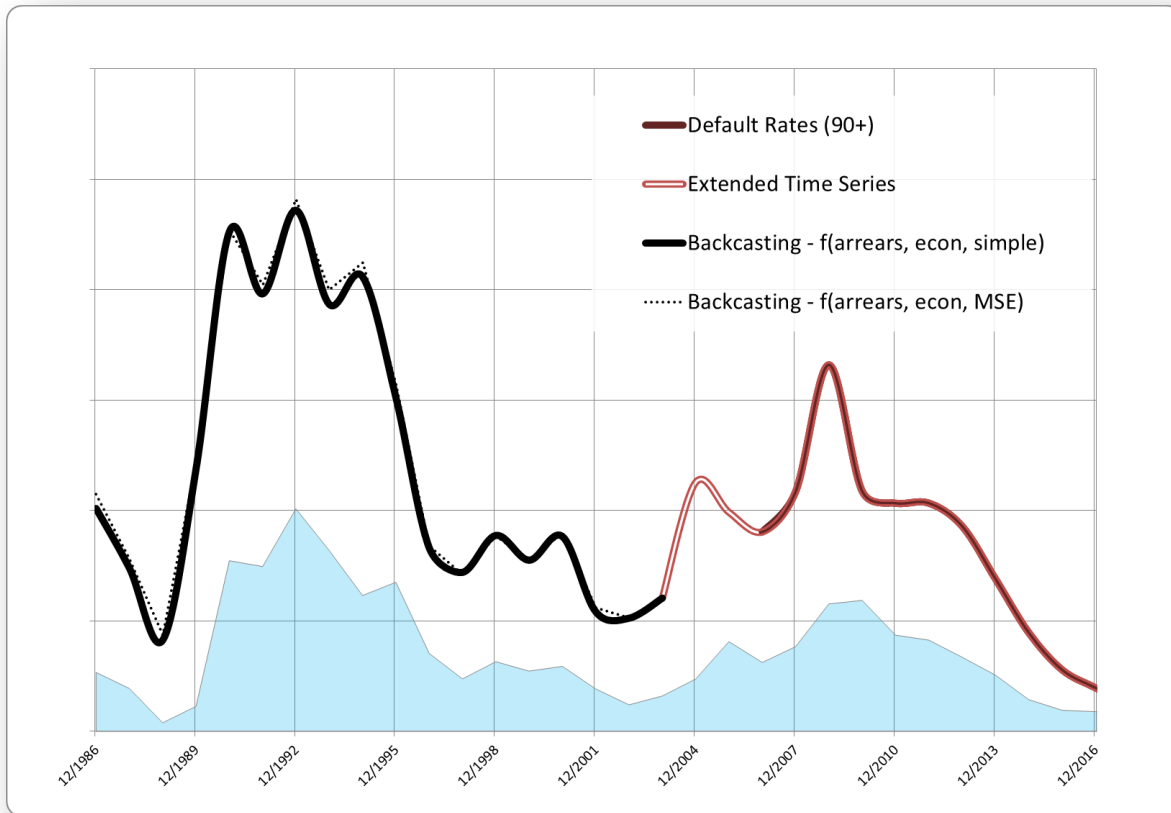
## Step (2): Logistic Regression

- A logistic regression model is fitted to the data, with the rating and the benchmark PD from step (1) as independent variables, to predict default per rating and historical period.
- PD per rating is estimated for all historical periods since 1986

# Backcasting Mortgages PD – step 1



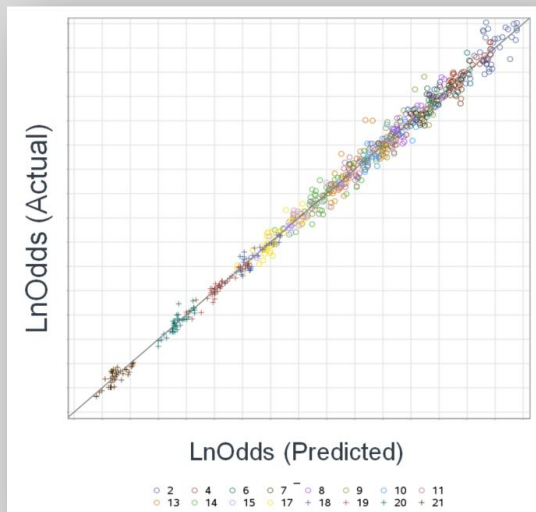
# Backcasting Mortgages PD – step 1



# Backcasting Mortgages PD – step 2

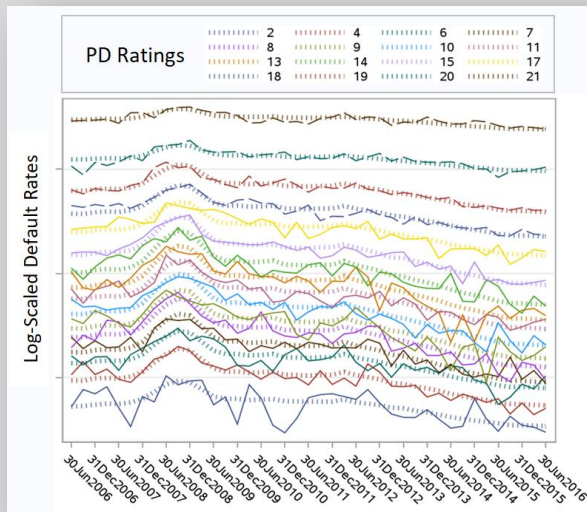
Step 2: A logistic regression model is fitted to the data, with the PD rating and the benchmark PD from step (1) used as independent variables.

## Model fit to the data (LnOdds)



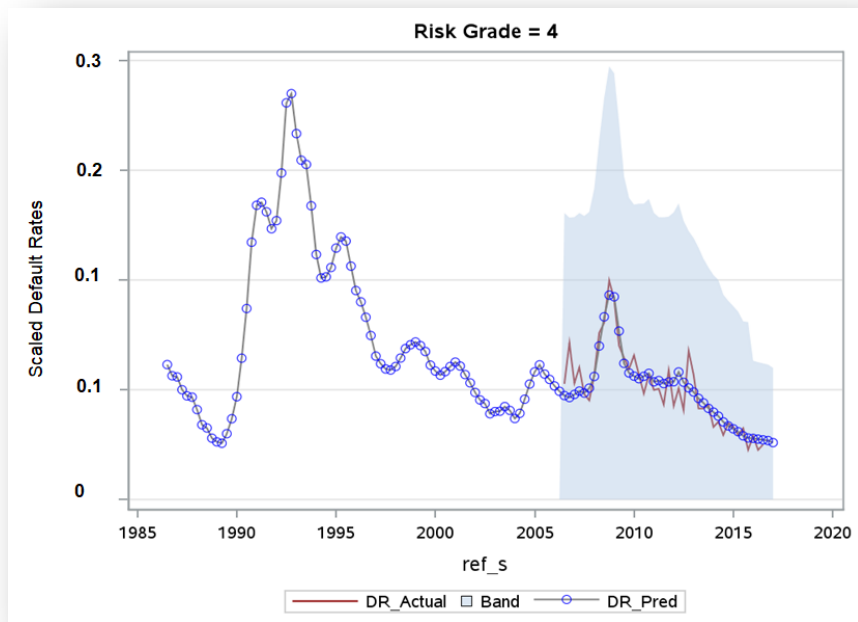
By Rating  
/ Quarter

## Actual vs Predicted (DR)

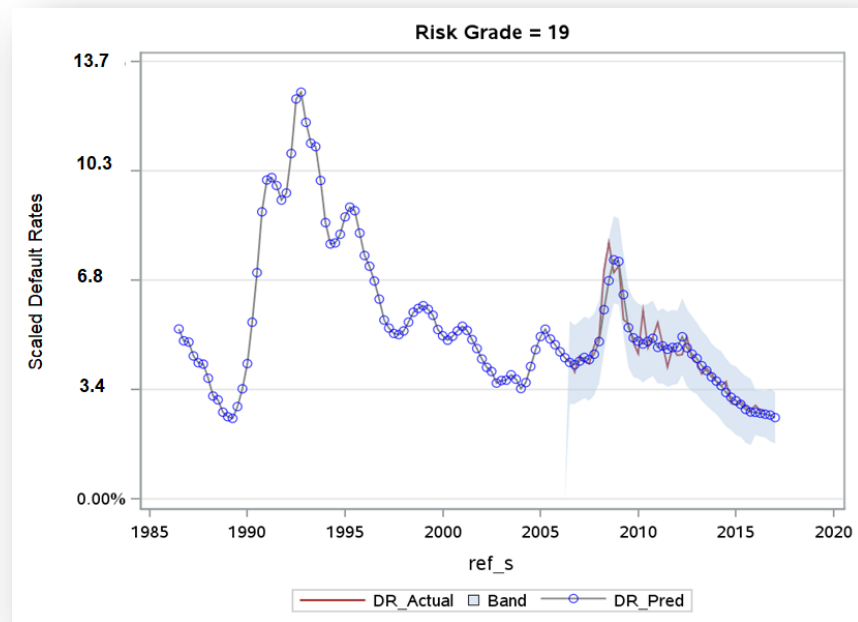


# Backcasting Mortgages PD – step 2

## Backcasting results for PD Rating (4)



## Backcasting results for PD Rating (19)



\*Scaled Default Rates  
\*\* Confidence intervals 5%

# Conclusions

- ✓ **UK Bureau data** can be used for conditional credit risk estimates and backcasting historical default rates.
- ✓ To support this hypothesis, **we created a benchmark portfolio** with relatively stable composition and relevant for a wide range of mortgage lenders.
- ✓ **We developed a rating system** combining comprehensive obligor risk characteristics and selected transaction risk drivers (LTV, BTL proxies)
- ✓ Achieved **stable performance with high ranking** across periods
- ✓ **Implemented a two-stage backcasting approach to estimate conditional PD for periods 1986 – 2006 using new Industry-level data assets**
- ✓ Created empirical material that can be incorporated in the long-run PD estimation process.

