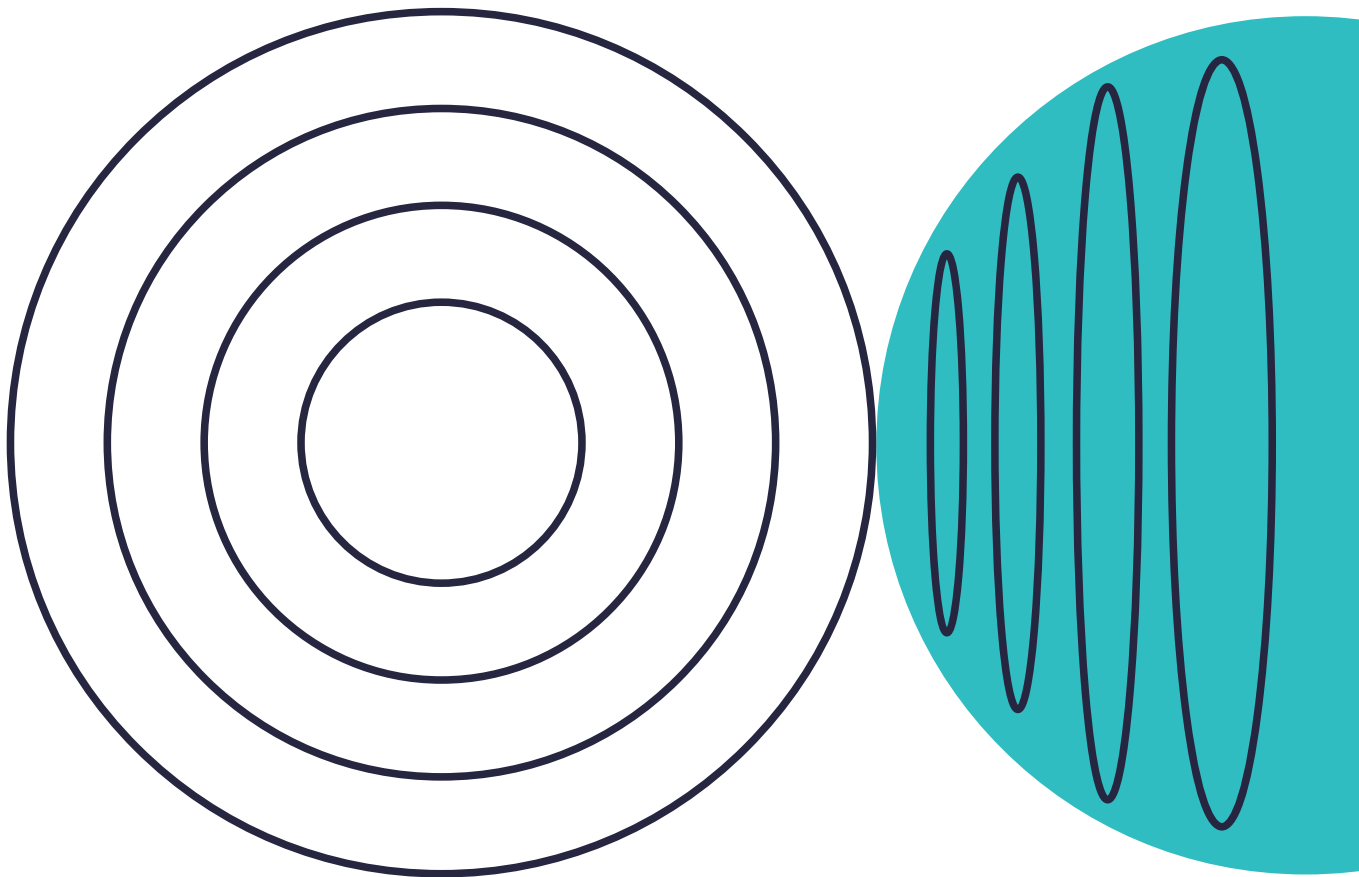


Monitoring hybrid rating systems and the role of Cyclicalality

13th September 2022



Sumit Dutta
Consultant

sumit.dutta@4-most.co.uk



Chris Warhurst
Technical Director

christopher.warhurst@4-most.co.uk

Monitoring of a Hybrid PD Rating System

Firms are continuing to work through the initial and ongoing PRA feedback regarding the development of Mortgage IRB PD models to meet the new regulatory requirements contained in SS 11/13 following the release of PS 13/17. However, an area that has seen less focus to date has been that of monitoring, although we are aware of recent challenges from regulators on how this will be addressed.

The monitoring of a hybrid rating system introduces new considerations to IRB monitoring, as the expectation of accuracy is that the model does not predict perfectly but does track the trend in the observed default rates to an extent. This is defined as the cyclical nature of the rating system.

This deviates from most existing approaches whereby a simple comparison of actual vs expected is required for Point in Time (PiT) approaches or there is limited consideration of short-term accuracy with Through the Cycle (TtC)/long run approaches. Even with existing hybrid type approaches, often seen in IRB models for commercial lending, the focus on 'accuracy' and 'cyclical nature' is not at the level anticipated for mortgages given the PRA feedback during developments.

This new requirement introduces the need to have an expectation of the movements in observed default rate, the PD produced by the rating system and the associated drivers. This paper explores approaches and challenges with regards to assessing whether the overall modelled PD is in line with expectations and therefore focuses heavily on the cyclical nature of the rating system as a key driver.

It should be noted that monitoring principles of the other key aspects – discrimination and stability – is not expected to fundamentally change from current practices, although there may be some additional considerations. For example, confidence intervals for risk grade default rates and continuous score vs pooled discrimination amongst others.

Overall Model Accuracy Assessment

To support our thinking in the area of overall accuracy monitoring, we have developed an illustrative hybrid model using publicly available information contained in the Freddie Mac Single Family Loan Level Dataset. This model was developed utilising data from 1999 to 2015 to allow a monitoring period post 2015 to be considered, simulating a real-life historical deployment. It is acknowledged that the final model is not necessarily compliant with all PRA requirements but was sufficient to explore the concept of accuracy monitoring.

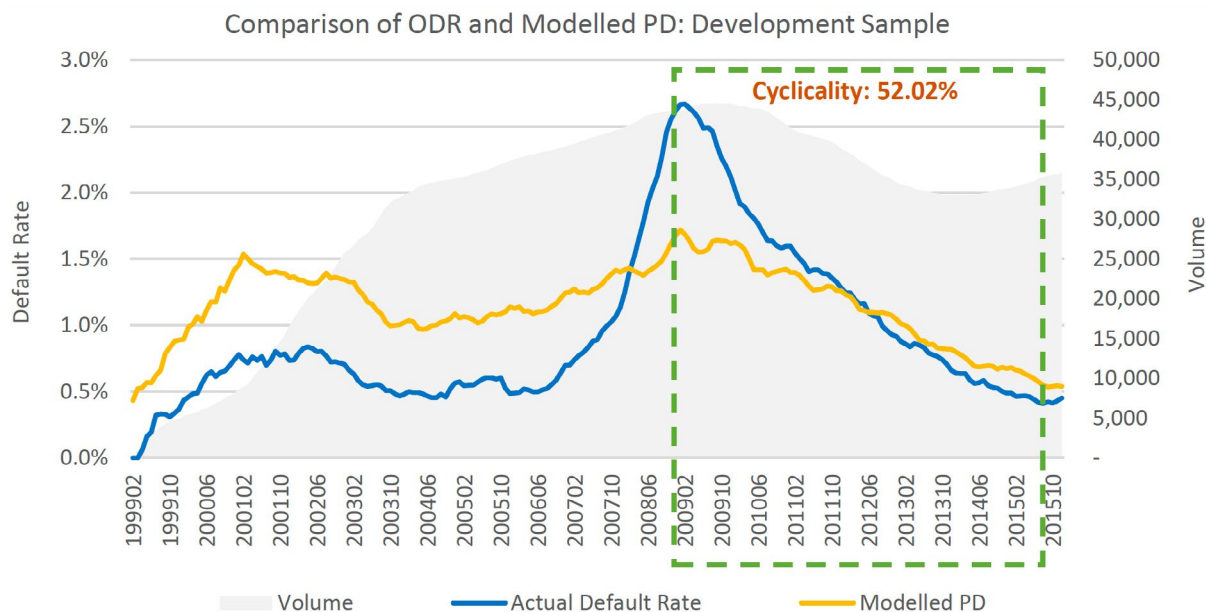
The main area of interest is understanding the level of cyclical nature that the model could be expected to exhibit over the monitoring period. As is currently common practice with the ongoing hybrid model builds, during development a single view of cyclical nature was defined based on a peak to 'local trough' observed default rate (ODR) movement (although it is understood that this view is likely to need to be expanded to consider other periods before final submission). However, as it is unlikely the monitoring period would contain a similar period, an acceptable variation in cyclical nature in recent times needs to be understood.

On this basis, several options were explored to quantify the expected level of cyclical nature based on the movement in the actual default rate. The aim is that, if an acceptable level of cyclical nature can be understood, this can then be used to derive the boundaries that the modelled PD should sit within based on the ODR movement. The model accuracy can then be assessed by comparing the modelled PD (from rating system) with the 'expected PD' (based on the cyclical nature).

As would occur with live monitoring, all parameters were developed using data within the development window and then applied to the monitoring period.

Initial Approach: Single Cyclicity Based on the Latest Cycle (Latest Peak to Trough: Jan 2009 – August 2015)

In this approach, cyclicity is calculated based on the latest peak to trough period (in this example, Jan 2009 – Aug 2015), which produces a cyclicity measure of 52.02%.



This cyclicity is then used to create an expected PD. Two approaches are considered in the application of the single cyclicity to create an expected PD.

Approach-1a

In this approach, for each monitoring period the 'Expected PD' is derived from the previous monitoring period based on the movement in the observed default rate (ODR) and the quantified level of cyclicity.

$$\text{Expected PD}_t = \text{Expected PD}_{t-1} + (\text{Cyclicity} * (\text{ODR}_t - \text{ODR}_{t-1}))$$

As this approach always uses the Expected PD, it does not correct itself based on how the modelled PD has tracked the ODR since development. Therefore all deviation between modelled PD and ODR since development is reflected into the latest monitoring values.

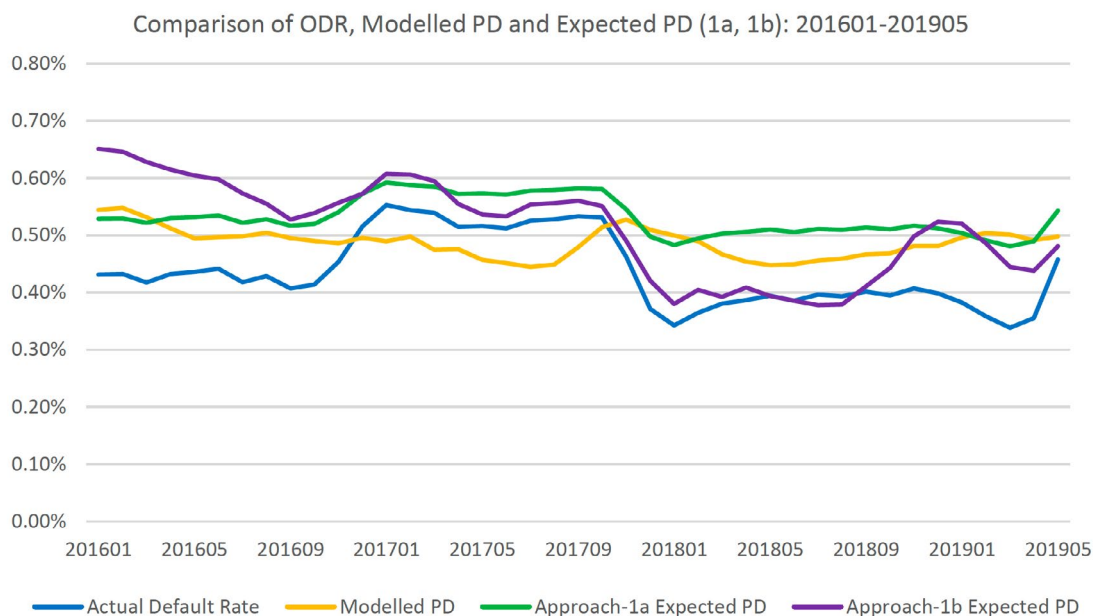
Approach-1b

Alternatively, an approach could only consider more recent deviations between ODR and modelled PD. This could adjust the modelled PD of 12-month prior (instead of from deployment), based on the movement in the observed default rate and the quantified level of cyclicity (note that alternate periods could be used but 12 months is considered reasonable for this example).

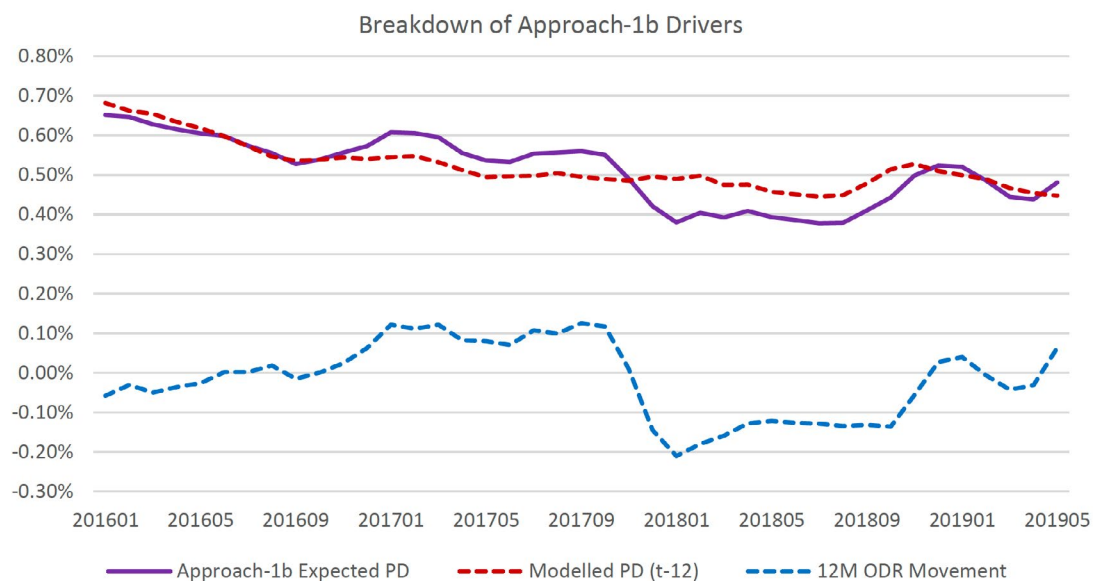
$$\text{Expected PD}_t = \text{Modelled_PD}_{t-12} + (\text{Cyclicity} * (\text{ODR}_t - \text{ODR}_{t-12}))$$

Comparison

The two approaches each have advantages and disadvantages, dependent upon the usage. Approach 1a will continue to show deviation in the current estimates due to the historical path of the ODR vs the modelled PD regardless of the time elapsed since the occurrence. This could be seen as advantageous when compared to approach 1b as it ensures that monitoring does not have the ability to trend back to green if prior unexplained variances are not addressed, as would happen after 12 months with approach 1b. However, if an organisation is confident of its ability to remediate any genuine concerns in a timely manner, it would be assumed that any unaddressed deviation has been analysed and comfort given that it is not of concern. In this instance approach 1b would be preferable as it would consider recent deviations. The charts below show the impact of applying these methods to the theoretical monitoring period.



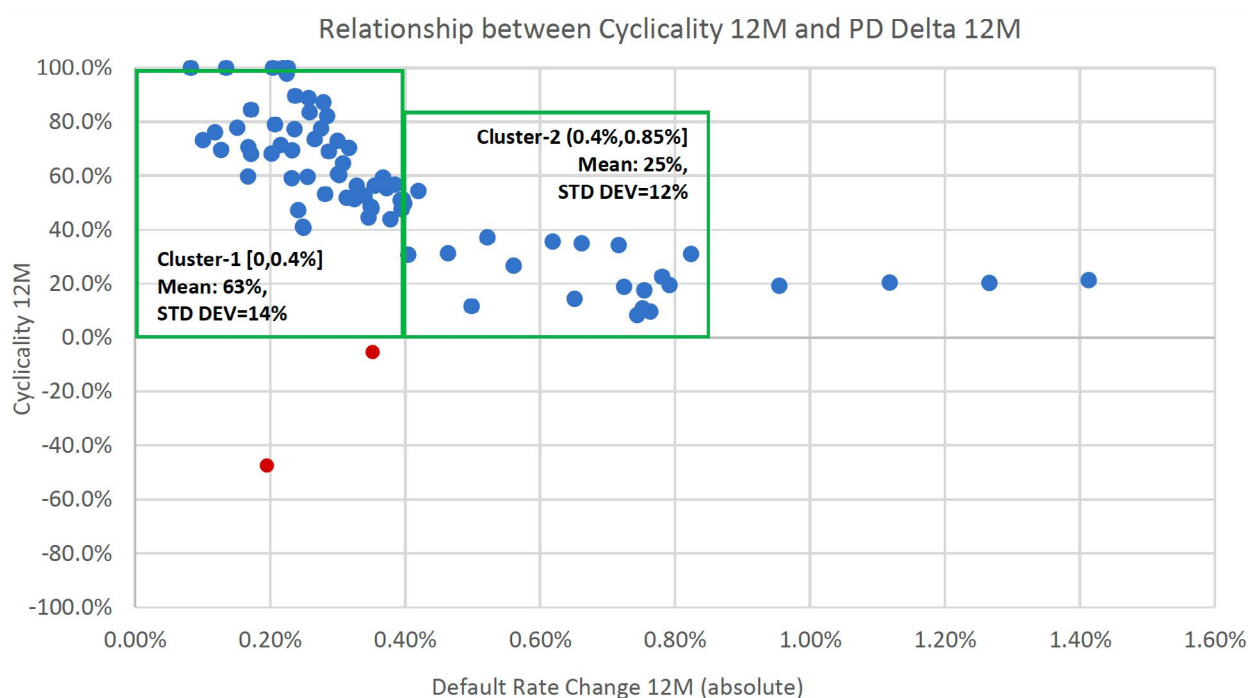
Due to the reliance of the modelled PD 12-month's prior and the short shocks to the ODR, Approach 1b is significantly more volatile and shows a greater 'error' between the modelled PD. This is particularly at the start of the series, which is driven by a large fall in the modelled PD over the prior 12-months and also from mid-2018 where there is a large increase in the expected PD. The impact of the movement in the expected PD from 12m prior can be seen when the drivers are isolated.



As can be seen in certain periods, e.g. Q1-Q3 2016, the movement in expected PD is heavily driven by the previous movements in the modelled PD (at t-12) rather than the shift in ODR. However, in periods of large ODR movement, e.g. Q4 2017, it does react well to the movement in ODR. This is different to approach 1a, which has a much smaller reaction to the change.

Tolerances: Lower and upper bounds for cyclicity based on the observed trend during development

The initial approaches considered each give a singular view of the expected PD. However, for monitoring purposes, there needs to be a level of tolerance as the modelled PD will likely never align exactly with expectations. To consider the assessment of tolerances, the development data is used to calculate cyclicity for 12 month intervals (based on the latest cycle, i.e. latest peak to trough 2009Q1-2015Q8) and capped/floored between -100% and +100%. Further, to get rid of the negative cyclicity, the trend is analysed only for cases with cyclicity greater than 0% (omitted points are shown in red in the chart below). The quantified level of cyclicity for each period can be compared to the movement in ODR over the same period. This is an important consideration as the monitoring period may have significantly different ODR movements to that seen over the development window. From the scatter plot below, it can be observed that cyclicity is not static, but does follow a welldefined trend. It is generally observed that it is higher and very volatile when the magnitude of default rate change is lower, and lower and less volatile when the magnitude of default rate change is significantly higher. Two clusters have been identified from the scatter plot, based on the mean and standard deviation of cyclicity within the clusters. It should be noted that alternate approaches are available whereby regression lines are fitted through the data points (or log transformed points) and upper lower bounds estimated from this. However, for simplicity and explainability we have taken a more basic approach here.



Using these clusters, upper and lower bounds of cyclicity have been assessed using 2-standard deviations to represent a c95% confidence interval.

| | Cluster-1 [0,0.4%] | Cluster-2 (0.4%,0.85%] |
|--|--------------------|------------------------|
| Mean (Cyclicity) 63% 25% | 63% | 25% |
| Standard Deviation (STD)- Cyclicity | 14% | 12% |
| Mean+2*STD (Cyclicity)- Upper bound | 90% | 49% |
| Mean-2*STD (Cyclicity)- Lower bound | 36% | 0% |

In this example, we arrive at a 0% lower bound for cluster 2, which would need to be assessed in more detail and likely overridden in a full monitoring development. As with Approach 1, an “expected PD range” can be derived by using these lower and upper bounds of cyclicity. The modelled PD will be assessed against this “range” during the ongoing monitoring of the model. In this example no RAG threshold has been set. However, this approach can be used with varying confidence intervals to set green, amber and red tolerances.

Approach-1a

As an extension to Approach 1a, this approach considers the change in lower/upper limits for the 'Expected PD' from 12-month prior, as this would depend on the movement in the ODR and the quantified level of cyclical. Additionally, it was considered that a near flat ODR should not necessarily be expected to result in a near flat modelled PD (as would occur with the underlying approach) due to general inherent model volatility. Therefore, it is proposed that a buffer needs to be considered in the lower and upper limits of the expected PD to arrive at a reasonable range. Otherwise, if there is no or very little change in the observed default rate between t and $t-12$, then both the lower and upper limits of the expected PD will be almost the same with little tolerance applied to the expected PD. Results with no buffer options are presented in [Appendix-D](#).

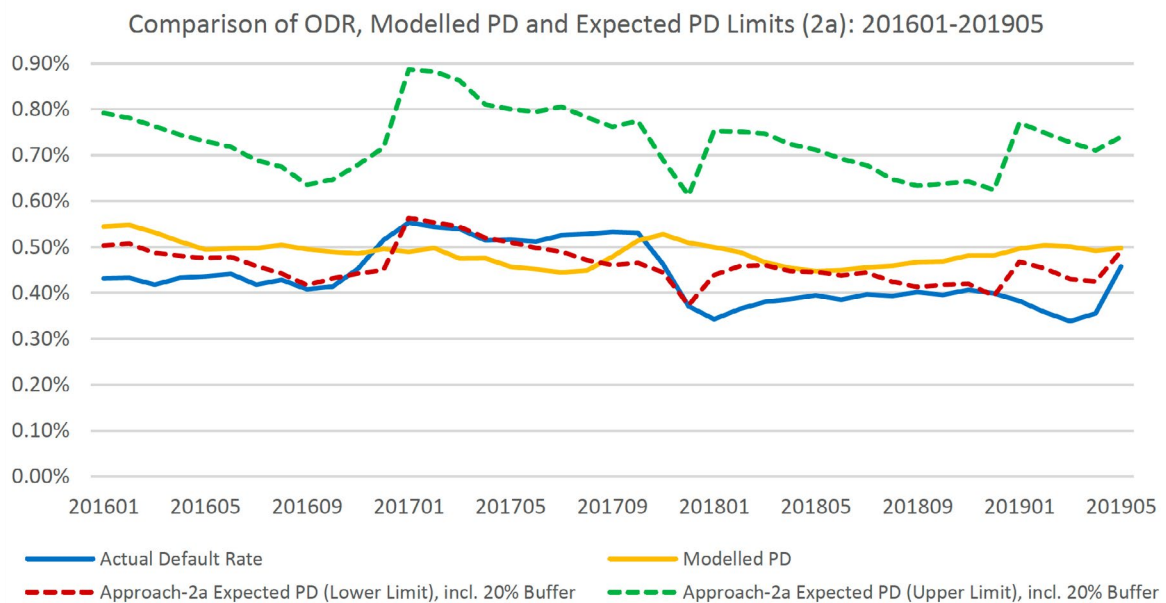
$$\text{Lower Cyc Exp PD}_t = \text{Lower Cyc Exp PD}_{t-12} + (\text{Lower Cyclical} * (\text{ODR}_t - \text{ODR}_{t-12}))$$

$$\text{Upper Cyc Exp PD}_t = \text{Upper Cyc Exp PD}_{t-12} + (\text{Upper Cyclical} * (\text{ODR}_t - \text{ODR}_{t-12}))$$

$$\text{Lower Limit of Expected PD}_t = \min(\text{Lower Cyc Exp PD}_t, \text{Upper Cyc Exp PD}_t) * (1 - \text{Buffer (20\%)})$$

$$\text{Upper Limit of Expected PD}_t = \max(\text{Lower Cyc Exp PD}_t, \text{Upper Cyc Exp PD}_t) * (1 + \text{Buffer (20\%)})$$

Note: Lower and Upper Cyclical will be applied based on the magnitude in default rate change. If $\Delta \leq 0.4\%$, then lower and upper limit of Cluster-1 will be used, else for Cluster-2.



Approach-1b

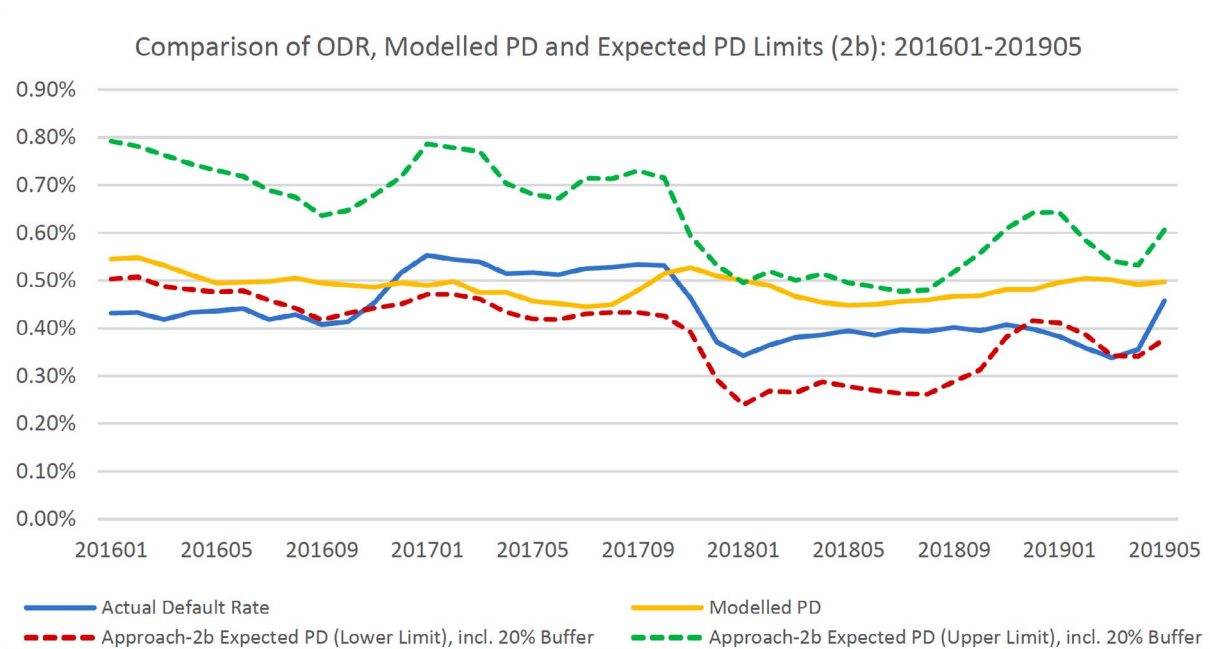
As an extension to Approach 1b, expected PD limits are derived in a similar fashion as 2a, but adjusting the Modelled PD instead of the expected PD of 12-month prior.

$$\text{Lower Cyc Exp PD}_t = \text{Modelled_PD}_{t-12} + (\text{Lower Cyclical} * (\text{ODR}_t - \text{ODR}_{t-12}))$$

$$\text{Upper Cyc Exp PD}_t = \text{Modelled_PD}_{t-12} + (\text{Upper Cyclical} * (\text{ODR}_t - \text{ODR}_{t-12}))$$

$$\text{Lower Limit of Expected PD}_t = \min(\text{Lower Cyc Exp PD}_t, \text{Upper Cyc Exp PD}_t) * (1 - \text{Buffer (20\%)})$$

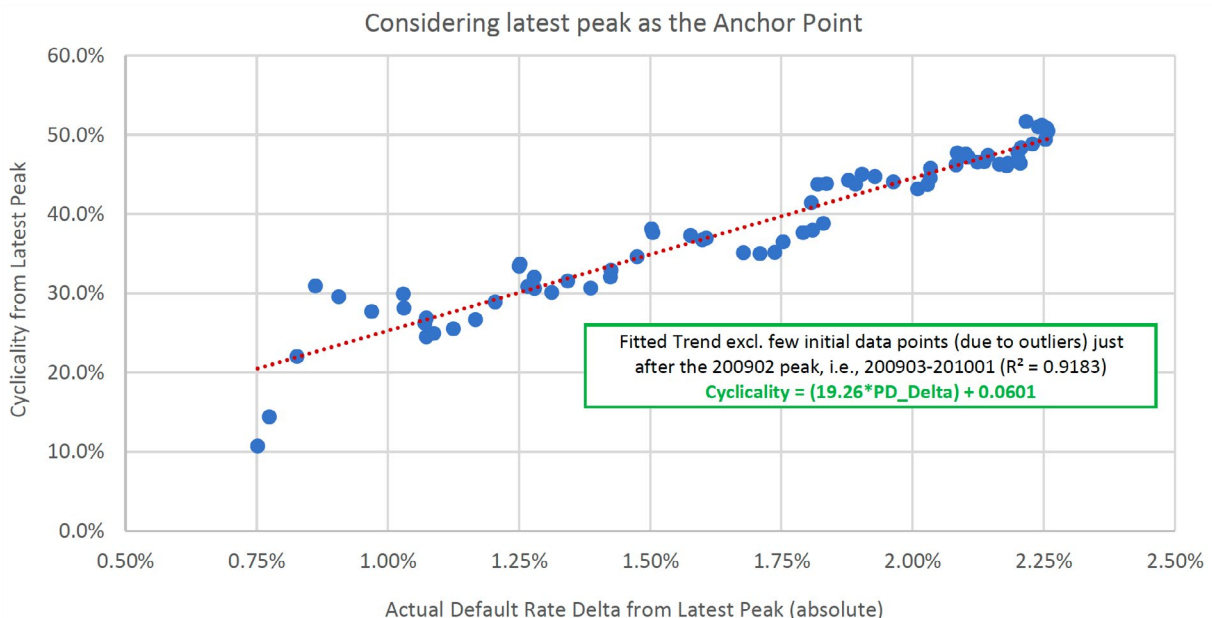
$$\text{Upper Limit of Expected PD}_t = \max(\text{Lower Cyc Exp PD}_t, \text{Upper Cyc Exp PD}_t) * (1 + \text{Buffer (20\%)})$$



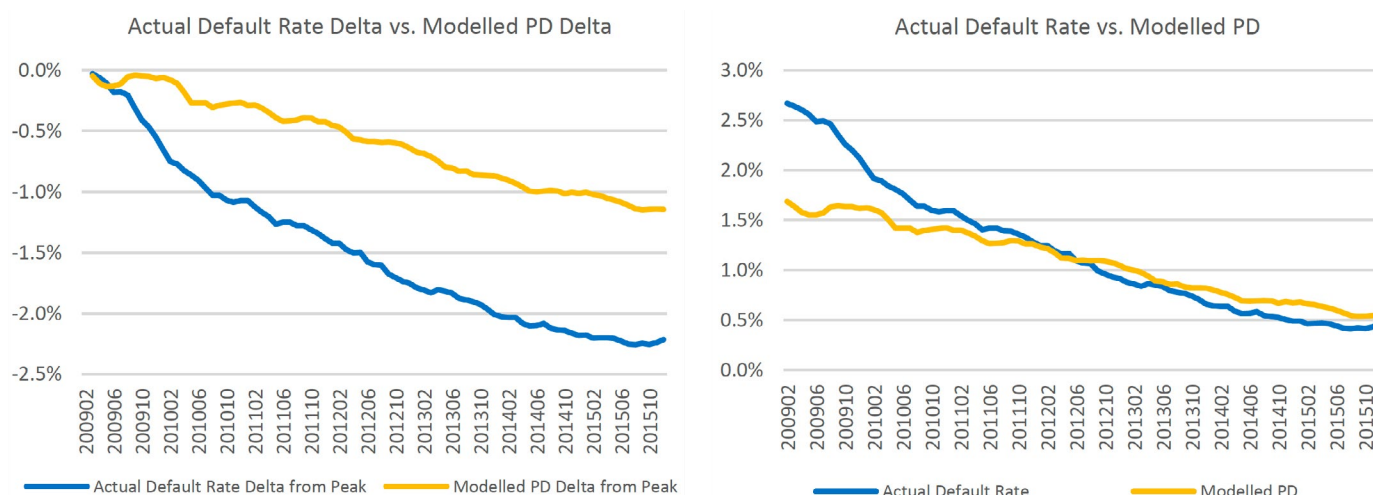
As would be expected, similar to Approach-1, in periods of large ODR movement (for example Q4 2017), Approach-2b reacts more compared to Approach-2a. Approach-2a has a relatively smaller reaction to the change due to the fall in ODR being partially offset by a rise in ODR during Q4 2016.

Further Analysis: Relationship Between Cyclicity and Default Rate Change from the Latest Peak (February 2009)

Another approach considered when assessing the expected level of cyclicity was to set the latest peak (February 2009) as the anchor point and derive the relationship between cyclicity and default rate change of these observations. This approach is more aligned to the PRA guidance on the assessment of cyclicity. As with the other approaches a strong linear relationship can be seen between the cyclicity and absolute default rate change from the latest peak.



However, although we are getting a strong linear relationship between PD Delta and Cyclicalit, the trend is opposite compared to that seen in other analysis (present in [Appendix-C](#)), where there isn't the anchoring of one of the points. In other options seen earlier, a negative relationship is observed between cyclicalit and the magnitude of PD Delta. But, in this approach the cyclicalit reduces with a decrease in magnitude of PD Delta. However, upon investigation the reason for this difference in trend is due to the behaviour over differing time intervals within the period, rather than a fundamental reversal of previous findings. As the default rate is falling over the entirety of the period of interest, the data points relating to the larger deltas in ODR are, by definition, where t is further away from t-1 (i.e. t is later in the period, as shown in the left-hand chart below). However, the movement (slope) of the ODR delta begins to decrease in the later months, and this coincides with the ODR and model PD moving in greater sync, as can be seen in the right-hand chart (i.e. the cyclicalit begins to increase as the short term deltas reduce, as was observed in the earlier analysis). This results in the overall cyclicalit increasing for the data points where t is in the later months and the marginal deltas are low but the cumulative delta from peak is still increasing, albeit at a lower rate.



Therefore, this apparent change in relationship between cyclicalit and ODR delta can be explained as being due to the time period they are taken over and change in the cyclicalit observed over differing rates of ODR change. It does however raise a consideration which is with regards to the impact of time interval on the level of cyclicalit. Some initial analysis has been undertaken to assess the results on both a 12m and 24m period (contained in [Appendix-C](#)) and it does show that the 24m interval is more stable and delivers higher quantified cyclicalit levels for smaller ODR deltas.

Further Analysis: Relationship between Cyclicalit and Default rate change from the latest Peak (February 2009)

Another approach tested is based on the 'Theoretical PD', by applying the cyclicalit to the difference between the observed default rate and the central tendency, adding that value back onto the central tendency, and comparing that to the modelled PD.

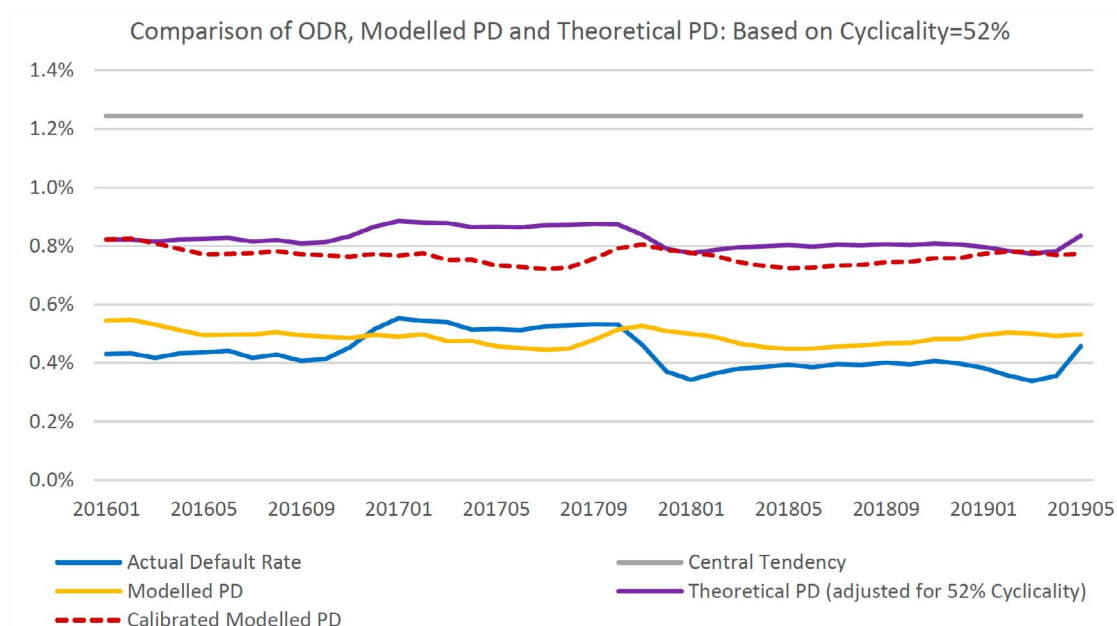
$$\text{Theoretical PD}_t = \text{Central Tendency} + (\text{Cyclicalit} \times (\text{ODR}_t - \text{Central Tendency}))$$

Central Tendency (1.24%) is the long run average PD, i.e. the time-weighted average of actual default rate calculated over the economic cycle (i.e. trough to trough period: 200606-201509).

As can be seen in the charts below, for both options the modelled PD is well below the theoretical PD throughout the monitoring period. This is because the actual default rate has reduced significantly in the recent period and is very low compared to the central tendency (1.24%) calculated over the economic cycle (200606-201509). Even at the point of deployment (i.e. the first monitoring month 201601), a big difference can be observed between the theoretical PD and the modelled PD. This has been raised as a concern previously by the PRA and also makes any attempt to apply go forward monitoring using this approach very difficult given the constant large deviation.

One simple solution (for the purposes of this example) could be to derive an adjustment factor at the point of deployment and use that factor to calibrate the modelled PD for the monitoring purposes.

$$\text{Calibrated Modelled PD}_t = \text{Modelled PD}_t + (\text{Theoretical PD}_{\text{deployment point}} - \text{Modelled PD}_{\text{deployment point}})$$



Once calibrated, this approach delivers very similar results to Approach-1a outlined earlier. This is intuitive as both ignore any prior movements in the modelled PD and embed all movements and deviations from the point of deployment.

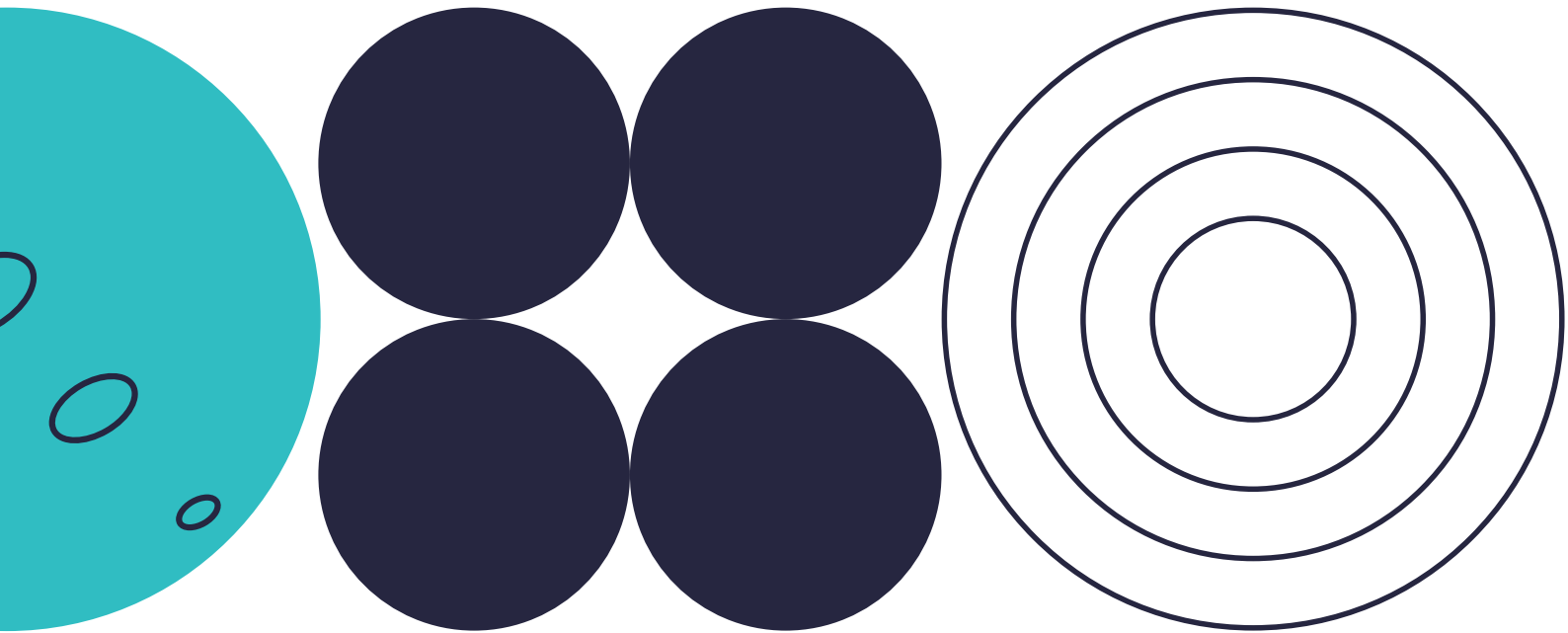
Conclusion

As outlined above, this is a focused investigation of the impact that cyclicalit has on the monitoring of hybrid PD, which has tried to outline where thinking needs to go in order to address the complexity of the challenges to be faced. The consideration regarding the variability of cyclicalit and how this is impacted by factors such as ODR delta and time period have been the focus of PRA discussions, but not necessarily with regards to how this impacts monitoring. As with any monitoring framework, the proposed solution will need to ensure that it can provide timely insights into the deterioration of the rating system, but also protect against unnecessary responses to simple variation and volatility. It is felt that the analysis presented here provides firms with starting guidance as to approaches that will fulfil this need.

It is acknowledged that the approaches detailed here could also be addressed by monitoring the cyclicalit directly rather than its impact on an 'expected PD'. However, it is felt that PD accuracy is a known measure in monitoring and that the analysis presented would form part of a wider investigation that may result in the adjustment to the monitoring ODR, based on observed changes in ODR outside of the cycle such as acquisition strategy or collections practices. If these adjustments were to take place it is felt that it would be easier for business understanding as to the impact on the 'accuracy' of the model.

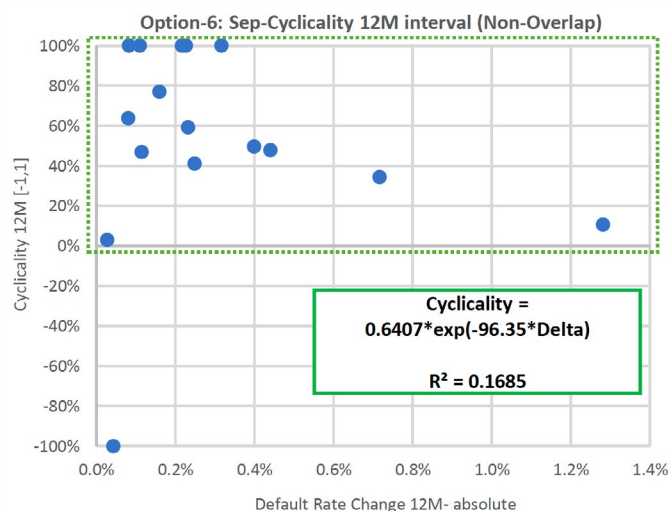
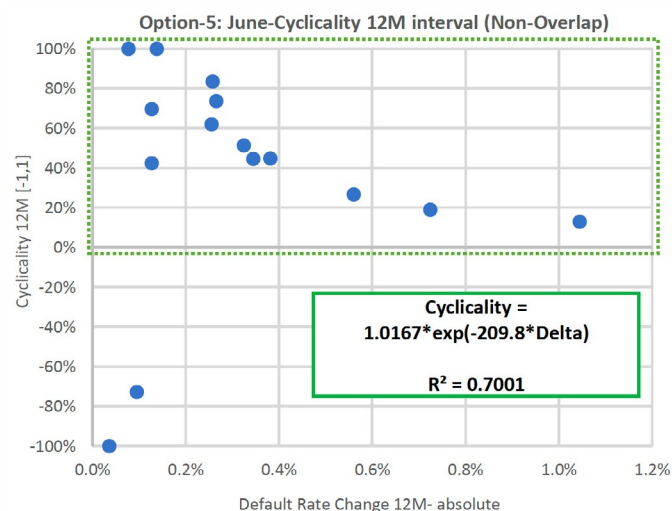
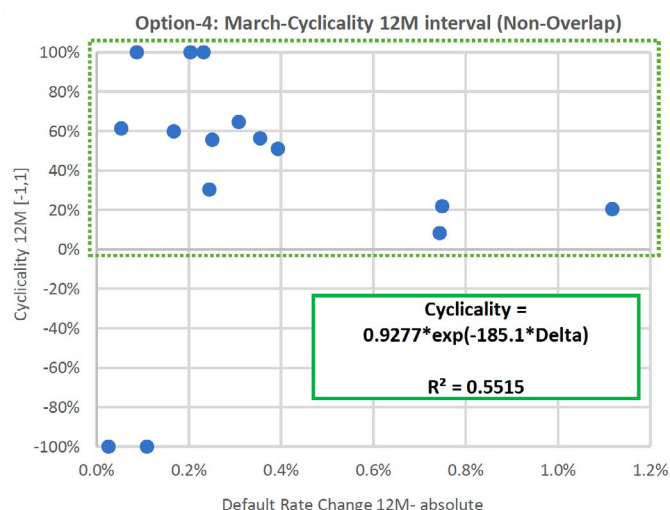
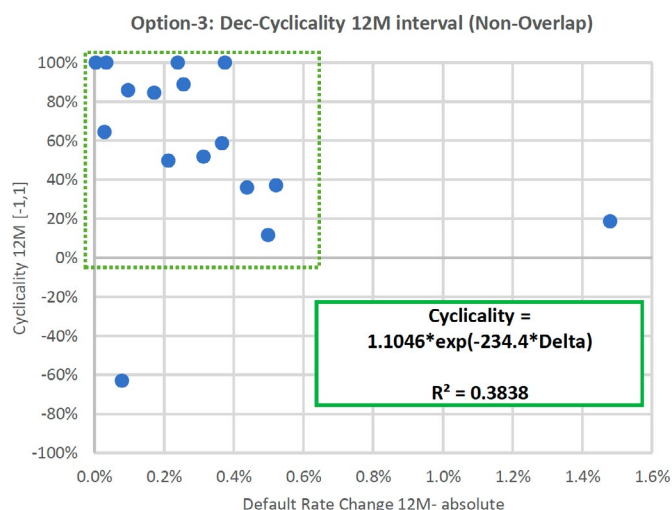
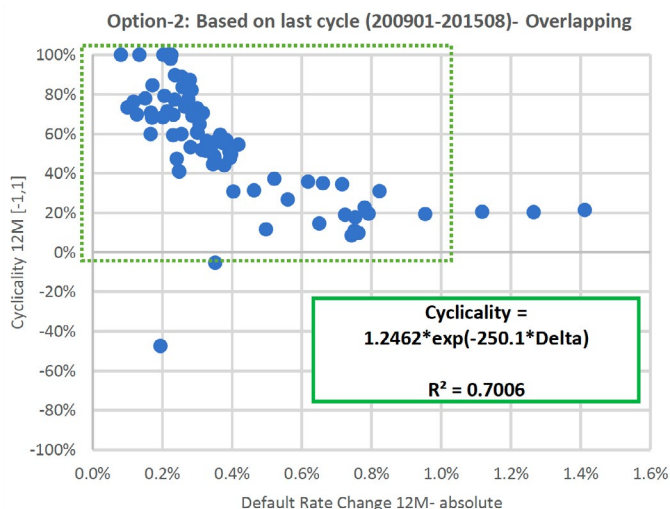
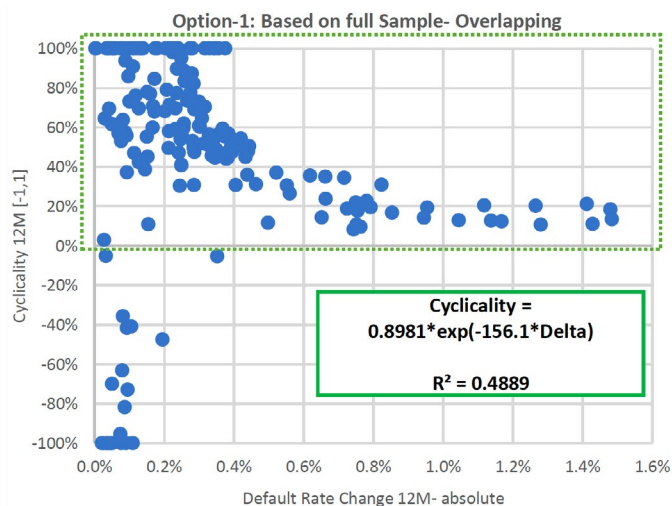
Additionally, to reiterate, it should not be forgotten that a full monitoring solution will require further metrics, both standard and new, alongside the overall 'accuracy' of the rating system. These other considerations should also support with the consideration of the accuracy performance and will likely include (but not be limited to) understanding of the ODR performance with risk grades, overall discrimination of the model and portfolio ODR performance compared to the industry.

Appendix



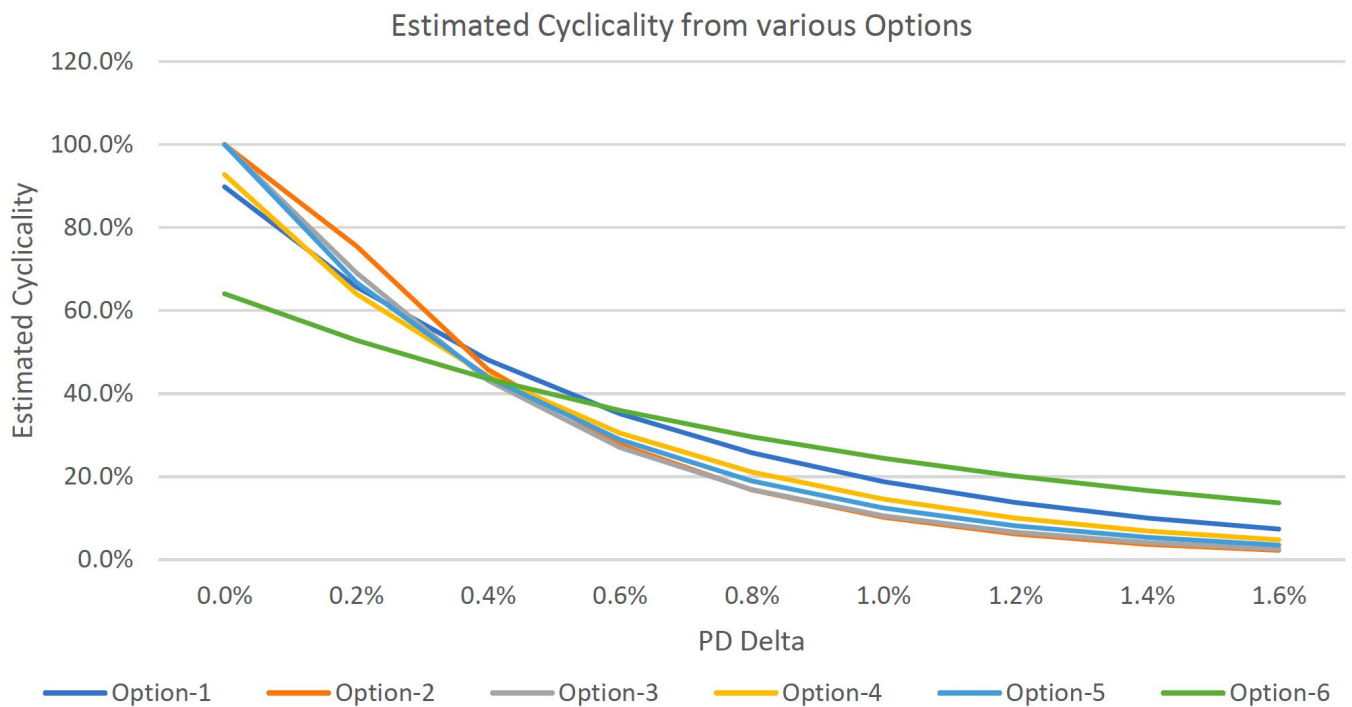
Appendix-A: Cyclicity trend based on the Overlap and Non-Overlap data points

Due to the very strong relationships seen in the analysis presented, upon review concern was raised with regards to whether the large clustering was due to overlapping data periods containing very similar cyclicity and ODR deltas. Hence, analysis was undertaken to understand if this trend changes drastically if it is derived based on overlapping or non-overlapping data points. Whilst the relationship is not as strong, as would be expected due to the independence of the points and lower volumes, the overall trend is still aligned to prior results.



The similarity of the trend is further evidenced when, based on these scatter plots (shown above), exponential trend lines are fitted.

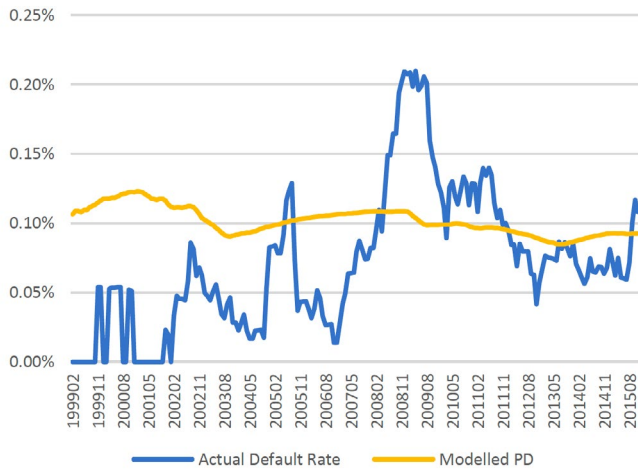
As can be observed in the chart below, the estimated cyclicality is very similar across all these options, except option-6 (i.e. based on the September trend). It can be noted that for option-6, the R-square is also very low (17%) compared to the other options.



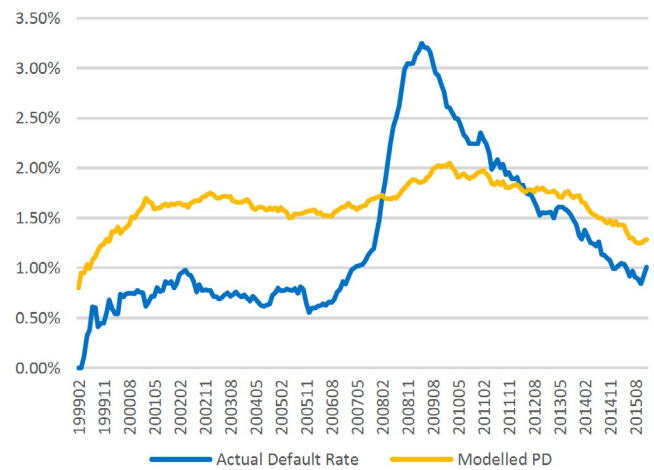
Appendix-B: Cyclicity trend by Low, Medium & High buckets

As part of the overall monitoring process, it will be necessary to consider more granular views and in particular underlying risk levels. Therefore, as an initial investigation, analysis was undertaken to understand if the cyclicity trend differs by aggregated risk buckets (the buckets are defined based on the PD grades). It is observed that the cyclicity is relatively low and less volatile for the Low (Grade=1,2) and High (Grade=11,12) PD buckets compared to the Medium (Grade=3 to 10) bucket. This analysis, alongside individual risk grade default rates, can be used for more detailed monitoring of the rating system

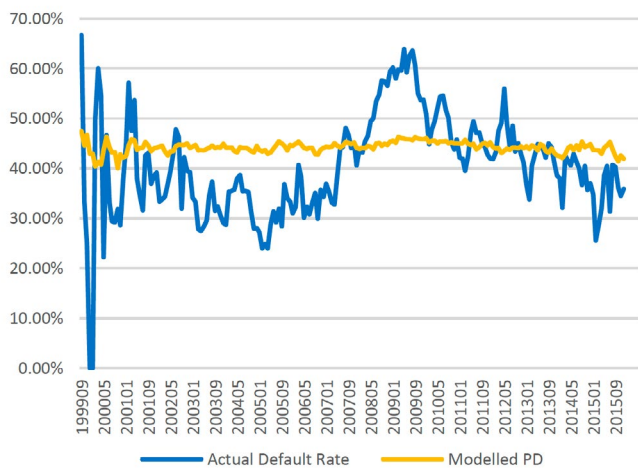
ODR vs. Modelled PD (Low Bucket: Grade 1,2)



ODR vs. Modelled PD (Medium Bucket: Grade 3-10)



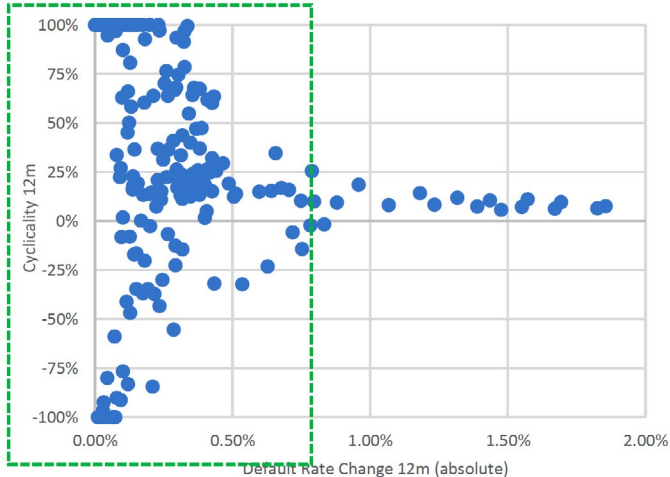
ODR vs. Modelled PD (High Bucket: Grade 11,12)



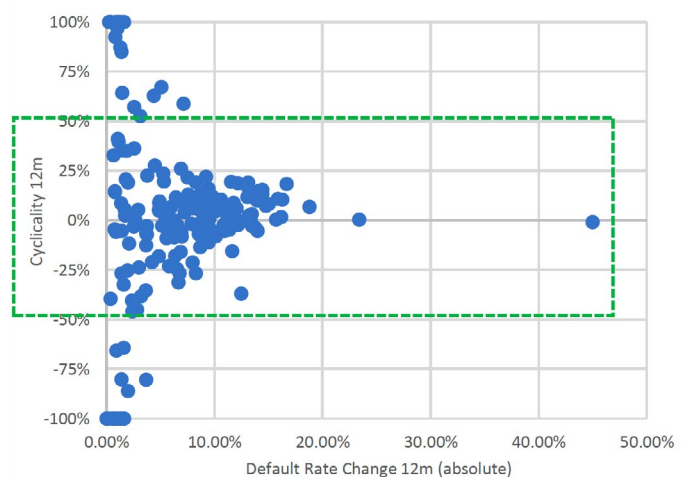
Cyclicity vs. PD Delta (Low Bucket: Grade 1-2)



Cyclicity vs. PD Delta (Medium Bucket: Grade 3-10)

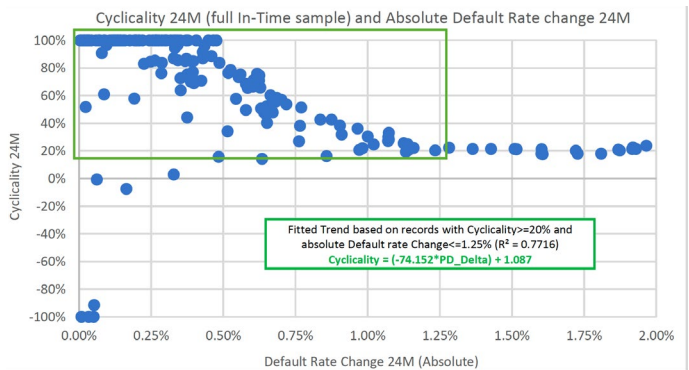
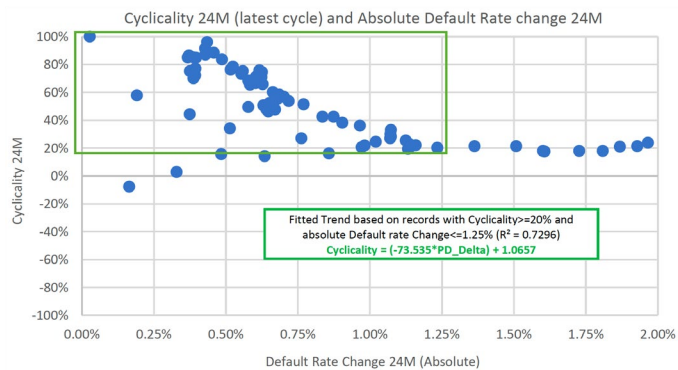
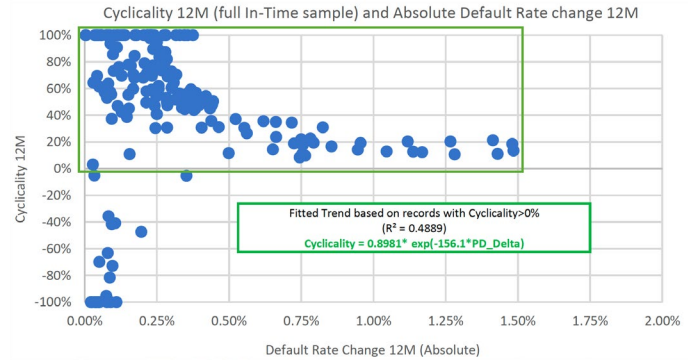
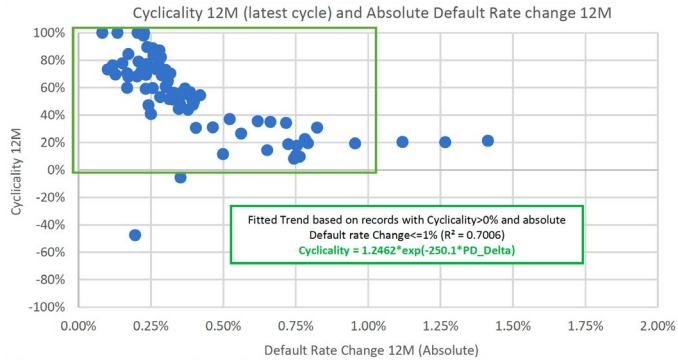


Cyclicity vs. PD Delta (High Bucket: Grade 11,12)



Appendix C: Relationship between Cyclicity over 12M and 24M intervals

Consideration was given to the assessment of cyclicity over both 12m (as per the original assessment) and 24m. Similar trends have been observed and the higher default rates indicate a similar overall level of cyclicity (circa 20%), however the 24m period does give a higher level of stability, as evidenced by the more stable R-square. Additionally, for lower ODR deltas, the 24m period does on average give a higher level of cyclicity. For example, for ODR deltas less than 0.5%, the average cyclicity for the 24m period is 81% compared to 52% based on the 12m period (considering a full in-time sample).



Appendix D: Results for Approach-2 with no buffer options

In Approach-2 (as discussed previously), a buffer is considered in the lower and upper limit of the expected PD to arrive at a reasonable range. Otherwise, if there is no or very little change in the observed default rate between t and $t-12$, then both the lower and upper limits of the expected PD will be almost the same with little tolerance applied to the expected PD. Results for both Approach-w2a and 2b with no-buffer options are presented below. As can be seen, the lower and upper limits are almost the same in 2016. Q1-Q4, as the change in the observed default rate compared to 12m prior is very low.

