DP/02/2012

ธนาคารแห่งประเทศไทย

BANK OF THAILAND
DISCUSSION PAPER

Challenges on the Validation of PD Models for Low Default Portfolios (LDPs) and Regulatory Policy Implications

By

Rungporn Roengpitya and Pratabjai Nilla-or*

August 23, 2012

E-Mail Address: rungporr@bot.or.th , Pratabjn@bot.or.th

Challenges on the Validation of PD Models for Low Default Portfolios (LDPs) and Regulatory Policy Implications

Rungporn Roengpitya and Pratabjai Nilla-or*

August 23, 2012

Abstract

This paper is the first of its kind to compare the probability of default (PD) estimates for low default portfolios (LDPs) from various methods–notably Pluto and Tasche (2006), Van Der Burgt (2007), Benjamin, Cathcart and Ryan (2006) and Roengpitya (2012)–using the historical data of sovereign borrowers from the years 1975-2009. The comparison results give insightful information to bank supervisors and banks regarding the PD model validation and possible underestimation of PD values. We found that the most conservative approaches tend to be that of Pluto and Tasche (2006) and Roengpitya (2012) while Van Der Burgt (2007) seemed to yield the least conservative estimates. Moreover, for prudent supervisory purposes, we suggested that the accuracy ratio (AR) in the Van Der Burgt (2007) CAP curve method should be restricted to be between 40% and 80% to prevent a possible underestimation of credit risk. Finally, we presented the necessary and sufficient conditions to ensure that the rank ordering of PD estimates from Pluto and Tasche (2006)'s *most prudent* approach is satisfied.

*Financial Institution Policy Group and Supervision Group, the Bank of Thailand. Email: rungporr@bot.or.th, pratabjain@bot.or.th. We would like to thank Pongsan Yodmuangcharoen for his assistance in this project. We are also grateful for the comments received by the executives, colleagues and fellow participants in the presentation at the Bank of Thailand and at the First International Conference on Credit Analysis and Risk Management. The views in this study are truly of the authors and do not necessarily represent the views of the Bank of Thailand.

Introduction

This paper aims at providing necessary information regarding the probability of default (PD) model validation on low-default portfolios (LDPs)[1] to bank supervisors as well as banks. While there are a few theoretical PD estimation models, comparing the level of conservatism between PD estimates from these models has not been addressed. Hence, we provided empirical comparisons of PD estimates obtained from employing different theoretical approaches so as to give information to bank supervisors and banks regarding the level of conservatism of estimated PDs from each model. Moreover, we highlighted the conditions needed to improve the performance of PD estimates from Pluto and Tasche (2006)'s approach and noted our concern on the possible underestimation of credit risk when using Van Der Burgt (2007)'s method.

Although both regulators and internal rating-based (IRB) banks encounter challenges regarding the estimations of IRB parameters for LDPs, both will deal with it from different aspects. For regulators, the most important concern has to do with the underestimation of credit risk associated with LDPs as a result of data scarcity. The rarity of default data leads to the difficulty in backtesting the PDs estimated from the risk models against the true historical default rates. For banks, the challenging task of LDP risk parameter estimation urges them to exclude these LDPs from the IRB calculation, though such treatment is not supported by the industry, as claimed in the joint industry working group discussion paper between the British Banker Association, London Investment Banking Association and International Swaps and Derivatives Association (BBA-LIBA-ISDA paper). Moreover, banks may choose to employ either the simple historical average approach or expert judgment to estimate the PDs instead, which can possibly progress to the underestimation of credit risk previously mentioned.

To obtain the insight on the performance of existing PD estimation models for LDPs, we considered four existing theoretical measures under consideration: (1) the *most prudent estimation* by Pluto and Tasche (2006); (2) *CAP curve calibration* by Van Der Burgt (2007); (3) *the margin of conservatism* by Benjamin, Cathcart and Ryan (2006); and (4) *hybrid models* by Roengpitya (2012).[2] We then calculated the PD estimates using each method on the virtual sovereign portfolio. For the sovereign portfolio, we used the S&P sovereign rating defaults from the years 1975-2009, with each country's rating status determined as of 2009.

From the comparisons, we found that the PD estimates from the *hybrid models* were comparable to, and in some cases were more conservative than the estimates from the *most prudent* method. When compared between

[1]The general definition of low default portfolios (LDPs) varies very much across concerned parties, it can be loosely defined as portfolios with too limited default events to obtain a robust probability of default (PD) estimation according to the principles outline in the Basel II or internal risk management objectives. The examples of borrowers that are regarded as having these LDPs are sovereigns, banks, highly-rated corporations and special forms of lending such as project finance.

[2]The *likelihood approach* by Forrest (2005) was omitted due to its calculation intensity from having too many rating grades.

the *most prudent* and the *margin of conservatism* approach, the results were mixed, depending on how the rating grades were segmented. Among all, the *CAP curve calibration* method yielded the *least* conservative estimates.

In addition, we discovered that the optimal k used to fit the exponential function to the actual CAP curve outlined in Van Der Burgt (2007) was very sensitive to the granularity of how one computed the cumulative number of borrowers. Therefore, when this method is employed, we suggested that, when validating PD estimates for this model, bank regulators should suggest the accuracy ratio (AR) of the fitted curve to be between $40\% - 80\%$, as seen in most best practices. Finally, since the *most prudent* method does not guarantee that the PD estimates will be rank-ordered, we then presented the necessary and sufficient conditions to ensure that the rank-order property of PD estimates from Pluto and Tasche (2006)'s model holds.

This paper has four sections. Section 1 briefly states the causes and existing practices regarding the estimation of the probability of default (PD) for low default portfolios. Section 2 provides the comparison of PD estimates using various approaches on the sovereign portfolio. We presented the necessary and sufficient conditions to ensure rank ordering for the *most prudent* method and the in-depth analysis of the *CAP curve* approach. In Section 4, we offer our policy recommendations regarding the PD estimation methods for LDPs. Concluding remarks complete this paper.

Section 1: Causes of LDPs and Existing Practices

This section briefly summarizes the discussion on the causes and the existing practices used by banks to deal with the group obligors whose default events rarely happen but yet need to be forecasted. Generally, one can group the causes of LDPs into two broad categories: i) LDPs from data segmentation or data collection and ii) LDPs from rare default events. The former case generally happens when banks segment a broad portfolio with sufficient default history into too small data pool (to be used in pricing, for example) and consequently create "involuntary" LDPs. This kind of LDP problem can be solved by using the whole non-segmented data pool to develop the PD model. Should this approach is chosen, then banks will need to prove that this *pooled data model* can reflect well the risk of borrowers in the pool, since banks may have to pull more than one types of borrowers in order to make possible the model development. For the issue related to data collection, LDPs may arise from having been collecting the data on both obligors and default events covering only a short time period. Banks can employ expert judgment during the first stage of the rating process for the group of portfolios in question and, once the data becomes more abundant (i.e. longer time horizons covered), then a statistical model can be developed with constant refinement of the models as banks gather longer time-series data.

The latter case is more problematic. For these types of obligors, the default history is rare or virtually non-existent. This lack of data can come from several sources, for example, low default events or insufficient number of obligors in a particular group. This means that, no matter how one is trying to segment or put together a broader set of obligors, the problem of rare or no default data remains unsolved. Generally, there are two approaches to deal with this–qualitatively or quantitatively. The common qualitative approach is by using the expert judgment. The quantitative approach involves increasing the significance, predictability and the ability to validate the model statistically by various instruments; for example, by increasing data points or by developing new quantitative estimation approaches around the non-default or low-default environment. These approaches are discussed in the appendix.

Section 2: Comparison of PD Estimates Using Sovereign Portfolio

This section presents details regarding the data of the virtual portfolios used in our empirical comparison. It is our intention to choose sovereign borrowers to test the PD estimates from both the existing and our models, due to their low-default nature as well as data availability to perform the empirical analysis. This section contains both the information on how our virtual sovereign portfolio is constructed and also the calculation of the PD estimates from various theoretical approaches.

2.1 Sovereign Portfolio Construction

To create the virtual sovereign portfolio in this study, we need to classify both the current sovereign rating and the history of defaults for all the countries in our sample. For the sovereign data, we used the S&P sovereign rating as of January 2009 to classify the rating of each country. The sovereign debt crisis data between 1975-2002 is from Savona and Vezzoli (2008) with the crisis episodes for the years 2003-2008 coming from *S&P Sovereign Ratings History* as of December 2009. The table containing information on sovereign debt crisis is presented in the appendix.

After identifying all the crisis episodes, we then created the population of countries to be included in our virtual portfolio. We used all of the countries having sovereign ratings as of January 2009. The list of countries and their January 2009 sovereign ratings is also presented in the appendix. Table 1 below shows the distribution of countries in our sample across S&P rating grades and the number of cumulative defaults[3] in each rating grade.

[3]The cumulative defaults are accumulated across the years 1975-2009 and across countries within the same rating grade. Therefore,

TABLE 1: DISTRIBUTION OF COUNTRIES ACROSS RATINGS AND NUMBER OF CUMULATIVE DEFAULTS

S&P Rating Grades of the Crisis	Number of Countries	Number of Obligor-year	Number of Defaults
AAA	19	646	0
AA+	3	102	0
AA	3	102	0
AA-	5	170	0
A+	5	170	1
A	10	340	0
A-	4	136	1
BBB+	5	170	1
BBB	5	170	3
BBB-	7	238	5
BB+	8	272	7
BB	5	170	2
BB-	9	306	19
B+	9	306	1
B	13	442	18
B-	4	136	7
CCC+	1	34	1
CCC	0	0	0
CCC-	0	0	0
CC	0	0	0
C	0	0	0
D	2	68	4
Total	117	3978	70

This table summarizes the number of countries, the obligor years, and the default (in obligor-years) for each S&P rating used in the study. The S&P rating shown is as of January 2009.

Since the cumulative defaults are calculated in obligor-years, Table 1 also presents the number of obligor-years for each rating grade.[4] To better track the performance of countries before their default experience, we excluded the two countries with rating "D" from the sample. Therefore, there are 3910 obligor-years and 66 defaults for the sovereign portfolio in the calculation.

TABLE 2: INFORMATION OF BOTH SOVEREIGN PORTFOLIO TYPES

Portfolio Type 1: Grouping by SA Rating Guideline

SA Grade	S&P Ratings	Ob-years	Defaults	ODR	Cum Ob-years	Cum Defaults	Cum ODR
1	AA-, AA, AA+, AAA	1020	0	0%	3910	66	1.688%
2	A-, A, A+	646	2	0.310%	2890	66	2.284%
3	BBB-, BBB, BBB+	578	9	1.557%	2244	64	2.852%
4	BB-, BB, BB+	748	28	3.743%	1666	55	3.301%
5	B-, B, B+	884	26	2.941%	918	27	2.941%
6	CCC+ or below	34	1	2.941%	34	1	2.941%

Portfolio Type 2: Grouping by SA Rating Guideline With Adjustments

Modified Grade	S&P Ratings	Ob-years	Defaults	ODR	Cum Ob-years	Cum Defaults	Cum ODR
1	AA-, AA, AA+, AAA	1020	0	0%	3910	66	1.688%
2	A, A+	510	1	0.196%	2890	66	2.284%
3	A-	136	1	0.735%	2380	65	2.731%
4	BBB, BBB+	340	4	1.176%	2244	64	2.852%
5	BBB-	238	5	2.101%	1904	60	3.151%
6	BB, BB+	442	9	2.036%	1666	55	3.301%
7	BB- or below	1224	46	3.758%	1224	46	3.758%

This table presents the information on both types of sovereign low-default portfolios, using the S&P ratings as of January 2009.

From Table 1, two types of sovereign portfolio are constructed by grouping the sovereign obligors into fewer

it can be thought of as the accumulation across obligor-years.

[4]Since there are 34 years in the study, the obligor-year for each rating grade is calculated as the multiplication of total obligors (or countries) in each rating grade (as of January 2009) and the number of years which is 34.

ratings. The first type of the sovereign portfolio groups the S&P ratings into Ratings 1-6, using the Bank of Thailand's mapping of ECAIs for banks employing standardized approaches (SA). The other type of sovereign portfolio differentiates the S&P ratings into 7 different grades (which conveniently satisfy the internal rating-based (IRB) requirement), using the SA mapping and then adjusting so that the observed default rates between rating grades rank-order more appropriately. Table 2 above then presents the information on both sovereign portfolio types, including the number of obligors, defaults, the observed default rate (ODR) and cumulative defaults (from the worst rating grade to the best rating grade) by rating.[5]

2.2 Empirical Results from Different Estimation Approaches

This section presents the empirical results when comparing PD estimates on LDPs using various methods, namely the *most prudent*, the CAP curve, the *margin of conservatism* and our hybrid models.

2.2.1 Results From the Most Prudent Method

From Table 2, we have in total 3,910 obligor-years with 66 defaults. For the approach used in Pluto and Tasche (2006), after collapsing the ratings into two cases–6 and 7 grades, we then computed the *most prudent* estimated PD for the 6-grade and 7-grade cases as shown in Table 2. For each case, we performed 3 different scenarios: (1) few defaults but independent default events; (2) few defaults with asset correlation of 4%; and (3) few defaults with asset correlation of 12%. Also, we did the calculation based on different confidence intervals γ: $50\%, 75\%, 90\%, 95\%, 99\%, 99.90\%$. Table 3A below presents the estimation results on the 6-grade portfolio.

TABLE 3A: RESULTS OF *MOST PRUDENT* PD ESTIMATION FOR 6-GRADE SOVEREIGN PORTFOLIO

Few defaults but independent default events

Grades/γ	50%	75%	90%	95%	99%	99.90%
1	1.70%	1.85%	1.98%	2.07%	2.23%	2.42%
2	2.31%	2.50%	2.68%	2.80%	3.02%	3.28%
3	2.88%	3.13%	3.36%	3.50%	3.78%	4.11%
4	3.34%	3.65%	3.94%	4.12%	4.47%	4.88%
5	**3.01%**	**3.41%**	**3.79%**	**4.03%**	4.51%	5.09%
6	4.88%	7.72%	10.95%	13.20%	17.98%	24.08%

Few defaults with asset correlation of 4%

Grades/γ	50%	75%	90%	95%	99%	99.90%
1	1.90%	2.60%	3.46%	4.09%	5.63%	7.58%
2	2.55%	3.43%	4.51%	5.29%	7.16%	9.51%
3	3.15%	4.21%	5.48%	6.40%	8.55%	11.24%
4	3.64%	4.84%	6.25%	7.27%	9.65%	12.61%
5	**3.17%**	**4.31%**	**5.64%**	**6.62%**	**8.89%**	**11.84%**
6	5.18%	8.65%	12.85%	15.87%	22.52%	31.32%

[5] This cumulative default information is used for the PD estimation according to Pluto and Tasche (2006).

TABLE 3A: RESULTS OF *MOST PRUDENT* PD ESTIMATION FOR 6-GRADE SOVEREIGN PORTFOLIO

Few defaults with asset correlation of 12%

Grades/γ	50%	75%	90%	95%	99%	99.90%
1	2.37%	3.89%	6.09%	7.83%	12.43%	18.41%
2	3.10%	4.98%	7.63%	9.69%	15.01%	21.72%
3	3.79%	5.98%	9.01%	11.33%	17.22%	24.49%
4	4.32%	6.76%	10.07%	12.58%	18.88%	26.57%
5	**3.80%**	**6.07%**	**9.16%**	**11.54%**	**17.55%**	**25.17%**
6	5.90%	10.67%	16.80%	21.38%	31.70%	44.99%

It can be seen from Table 3A above that the PD estimates shown in bold font fail to meet the rank ordering property, as mentioned in Pluto and Tasche (2006). In other words, the estimated PD of the worse rating grade is lower than the PD of the better one. One can test, for example, that the rank-order failure of Grade 5 in the zero-correlation case came from it failing the third condition in Proposition 1 stated in the upcoming Section 3.1.

An interesting fact can be drawn from the behavior of PD estimates in Table 3A. That is, *the severity of the rank-order failure increases with asset correlation.* As the asset correlation increases, the number of rank-order failures increases and expands to higher confidence intervals also. For example, when default events are independent, there are only three failures. However, when the asset correlation increases to 4% or 12%, the number of rank-order failures grows to be five. Therefore, as one allows the asset to be more highly correlated, the rank ordering of PD estimates becomes even more unstable. To cope with this rank-order failure, the portfolio then is re-grouped into 7 ratings instead to meet the conditions proposed in Proposition 1.

TABLE 3B: RESULTS OF *MOST PRUDENT* PD ESTIMATION FOR 7-GRADE SOVEREIGN PORTFOLIO

Few defaults but independent default events

Grades/γ	50%	75%	90%	95%	99%	99.90%
1	1.70%	1.85%	1.98%	2.07%	2.23%	2.42%
2	2.31%	2.50%	2.68%	2.80%	3.02%	3.28%
3	2.76%	2.99%	3.21%	3.35%	3.61%	3.92%
4	2.88%	3.13%	3.36%	3.50%	3.78%	4.11%
5	3.19%	3.46%	3.73%	3.89%	4.21%	4.59%
6	3.34%	3.65%	3.94%	4.12%	4.47%	4.88%
7	3.81%	4.19%	4.55%	4.78%	5.22%	5.75%

Few defaults with asset correlation of 4%

Grades/γ	50%	75%	90%	95%	99%	99.90%
1	1.90%	2.60%	3.46%	4.09%	5.63%	7.58%
2	2.55%	3.43%	4.51%	5.29%	7.16%	9.51%
3	3.02%	4.05%	5.27%	6.16%	8.26%	10.87%
4	3.15%	4.21%	5.48%	6.40%	8.55%	11.24%
5	3.47%	4.63%	5.99%	6.97%	9.28%	12.13%
6	3.64%	4.84%	6.25%	7.27%	9.65%	12.61%
7	4.13%	5.47%	7.04%	8.17%	10.76%	14.00%

Few defaults with asset correlation of 12%

Grades/γ	50%	75%	90%	95%	99%	99.90%
1	2.37%	3.89%	6.09%	7.83%	12.43%	18.41%
2	3.10%	4.98%	7.63%	9.69%	15.01%	21.72%
3	3.64%	5.77%	8.72%	10.99%	16.76%	23.92%
4	3.79%	5.98%	9.01%	11.33%	17.22%	24.49%
5	4.14%	6.49%	9.71%	12.16%	18.32%	25.86%
6	4.32%	6.76%	10.07%	12.58%	18.88%	26.57%
7	4.86%	7.54%	11.12%	13.82%	20.50%	28.58%

The *most prudent* PD estimates of the 7-grade portfolio are presented in Table 3B above. It can be seen from Table 3B that the PD estimates are now ranked, as the rating system now satisfies the necessary and sufficient conditions outlined in Proposition 1, which ensures the rank ordering of PD estimates using this approach. We also obtained that the higher the asset correlation was, the more conservative PD estimates became, sharing a similar finding as in Pluto and Tasche (2006).

2.2.2 Results From the CAP Curve Method

With the same sovereign portfolio, we computed the estimated PDs for the 6-grade and 7-grade cases using the same exponential function and PD calibration methodology as defined in Van Der Burgt (2007). The results are summarized in Table 4.

TABLE 4: RESULTS OF THE CAP CURVE PD ESTIMATION FOR SOVEREIGN PORTFOLIO

SA Grade	Obligor-years	Defaults	Cum Ob-years	Cum Defaults	ODR	CAP PD
1	1020	0	100.00%	100.00%	0.000%	0.361%
2	646	2	73.91%	100.00%	0.310%	0.704%
3	578	9	57.39%	96.97%	1.557%	1.152%
4	748	28	42.61%	83.33%	3.743%	1.963%
5	884	26	23.48%	40.91%	2.941%	3.782%
6	34	1	0.87%	1.52%	2.941%	5.469%
Average					1.688%	1.659%
Estimated k						3.1426

Modified SA Grade	Obligor-years	Defaults	Cum Ob-years	Cum Defaults	ODR	CAP
1	1020	0	100.00%	100.00%	0.000%	0.178%
2	510	1	73.91%	100.00%	0.196%	0.411%
3	136	1	60.87%	98.48%	0.735%	0.585%
4	340	4	57.39%	96.97%	1.176%	0.758%
5	238	5	48.70%	90.91%	2.101%	1.040%
6	442	9	42.61%	83.33%	2.036%	1.508%
7	1224	46	31.30%	69.70%	3.758%	3.747%
Average					1.688%	1.593%
Estimated k						4.2726

From the table, the estimated PD in each grade is quite in line with the observed default rate. However, the estimated PD for the fourth rating in the 6-grade case is being underestimated. Van Der Burgt (2007) concluded that this methodology was sensitive to the obligor distribution, especially both default and non-default obligors. Hence, a big change in the obligor number in a particular rating can significantly gives a poor estimated PD in that grade. This also corresponds to our mathematical proof in Proposition 2 of Section 3.2 that the shape of the CAP curve depends solely on the actual cumulative defaults and cumulative obligors. Consequently, to make sure that such PD estimation is still appropriate for our portfolio, we recommended checking the stability of the portfolio distribution through time. Once the portfolio changes, the review of the CAP curve PD calibration is required.

2.2.3 Results From the Margin of Conservatism Method

Benjamin et al (2006) proposed *the margin of conservatism* approach which, in part, employed the *most prudent* estimation methodology in Pluto and Tasche (2006). However, rather than estimating the upper bound PD for each grade, this method begins with using the best grade estimated PD as the conservative portfolio PD. Then, one calculates the corresponding scaling factor to scale-up the initial estimated PD for each rating grade. As a result, factors that drive the *the margin of conservatism* PDs come from both the obligor distribution as well as the choice of method used in the initial PD estimation.

Using again our sovereign portfolio, we simulated our look-up table first for various levels of asset correlation– $0\%, 4\%$ and 12%–and with different confidence levels–$50\%, 75\%, 90\%, 95\%, 99\%, 99.9\%$. The results of the similation yield the *conservative portfolio PD*. The following table presents our conservative portfolio PD.

TABLE 5A: PD LOOK-UP TABLE FOR VIRTUAL SOVEREIGN PORTFOLIO

Correlation	50%	75%	90%	95%	99%	99.90%
0%	1.70%	1.85%	1.98%	2.07%	2.23%	2.42%
4%	1.90%	2.60%	3.46%	4.09%	5.63%	7.58%
12%	2.37%	3.89%	6.09%	7.83%	12.43%	18.41%

This table presents PD look-Up table for virtual sovereign portfolio under the parameter assumption: number of obligor-years of 3,910 and number of default 66.

In order to estimate the initial PD for each rating grade, we chose to fit a curve to the actual historical default data. Since the actual default rate does not guarantee the rank ordering across grades, an appropriate curve fitting technique is employed to solve this problem. The exponential function is used for our initial PD estimation, as it explains better the relationship between the estimated PDs and the rating grades and is in accordance with the common assumption that the probability of default is of exponential form. Therefore, we modeled the PDs of our virtual portfolio by minimizing the root-mean-square error (RMSE) of the exponential equation between the long run default rate and the rating grade, as follows.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - [a \cdot e^{b\cdot(i-1)}])^2}, \tag{1}$$

where $f(x_i) = a \cdot e^{b\cdot(i-1)}$ is our exponential function choice, i is the rating grade and y_i is the long run observed default rate for rating grade i. In our exponential curve fitting, a is an intercept (i.e. the PD of the best rating grade) while b represents the estimated coefficient.[6]

[6] One note regarding this technique. The fitted PD obtained from the RMSE approach is very sensitive to the outliers of i for the rating system with a small number of rating grades. If there is a low degree of an outlier deviation or if there is a large number of obligors in the portfolio, we can fit our exponential curve without constraints. Nonetheless, if there is a high degree of the outlier deviation or a low number of obligors in each grade, it is inevitable to set some constraints on the fitting, or else the estimated PD of the high-risk grade possibly can exceed one.

Tables 5B and 5C present the parameter estimates and the initial PD for each rating grade resulted from this exponential curve fitting. Also, the account-weighted portfolio PD is calculated to get the initial portfolio PD.

TABLE 5B: RESULTS OF THE EXPONENTIAL PARAMETER ESTIMATES AND RMSE WITHOUT PD CONSTRAINTS

Sovereign Rating System	a	b	RMSE
6-grade case	0.009	0.285	0.90%
7-grade case	0.003	0.419	0.31%

This table presents results of the exponential parameter estimate and root mean square error for 6-GRADE and 7-GRADE sovereign portfolio without any PD constraints.

TABLE 5C: RESULTS OF THE FINAL PDS FROM EXPONENTIAL CURVE FITTING AND MARGIN OF CONSERVATISM ADJUSTMENTS

6-rating Grade Case

Grade	Initial PD	No of Obligor-years	Weight	Weight × initial PD
1	0.87%	1020	26.09%	0.23%
2	1.15%	646	16.52%	0.19%
3	1.53%	578	14.78%	0.23%
4	2.04%	748	19.13%	0.39%
5	2.71%	884	22.61%	0.61%
6	3.61%	34	0.87%	0.03%
Account-weighted initial portfolio PD				1.69%

7-rating Grade Case

Grade	Initial PD	No of Obligor-years	Weight	Weight × initial PD
1	0.30%	1020	26.09%	0.08%
2	0.46%	510	13.04%	0.06%
3	0.69%	136	3.48%	0.02%
4	1.05%	340	8.70%	0.09%
5	1.06%	238	6.09%	0.10%
6	2.43%	442	11.30%	0.28%
7	3.70%	1224	31.30%	1.16%
Account-weighted initial portfolio PD				1.78%

From the tables above, for the 6-grade rating system, we observed that there was a much higher PD value in the first rating grade, which is equal to 0.87% as compared to the estimated PD of 0.3% in the 7-grade rating case. The underlying reason is that the default rate does not distributed rank-orderly and exponentially like the 7-grade case, as seen in Table 2. This disorder nature of default rates also leads to a higher RMSE for the 6-rating case (0.90% versus 0.31%).

After obtaining both the conservative portfolio PDs and the initial portfolio PDs, we then computed the scaling factor. The scaling factor is calculated from the following equation

$$scaling\, factor = \frac{conservative\, portfolio\, PD}{initial\, portfolio\, PD}. \quad (2)$$

Using Equation (2), we obtained that, for example, the scaling factor of the 6-rating grade case at 50% confidence level and 12% asset correlation was $\frac{2.37\%}{1.69\%} = 1.40$.

To get the final PD for each grade, the initial PD estimates, listed in the second column of Table 5C, are multiplied by the scaling factor calculated using Equation (2). For example, at the 50% confidence level and 12% asset correlation, the 6-grade case initial PDs should all be multiplied by 1.40, as calculated in Equation (2).

Also, note that the scaling factor should be used for the initial PD adjustment in a conservative way only. For example, since our initial portfolio PDs are higher than the conservative portfolio PDs for the 7-rating grade case at the confidence level of 50% with zero asset correlation, the intital PDs for this case need no adjustment. The calculation results for both the 6-grade and 7-grade cases are presented in tables 5D and 5E respectively.

TABLE 5D: PD FROM THE MARGIN OF CONSERVATISM METHOD FOR 6-GRADE CASE

ZERO CORRELATION

Grade/γ	50%	75%	90%	95%	99%	99.9%
1	0.88%	0.95%	1.02%	1.07%	1.15%	1.25%
2	1.16%	1.26%	1.35%	1.41%	1.52%	1.65%
3	1.54%	1.68%	1.79%	1.88%	2.02%	2.19%
4	2.05%	2.24%	2.39%	2.50%	2.70%	2.92%
5	2.73%	2.97%	3.18%	3.32%	3.58%	3.89%
6	3.64%	3.96%	4.23%	4.43%	4.77%	5.18%

4% CORRELATION

Grade/γ	50%	75%	90%	95%	99%	99.9%
1	0.98%	1.34%	1.78%	2.11%	2.90%	3.91%
2	1.29%	1.77%	2.36%	2.79%	3.84%	5.16%
3	1.72%	2.36%	3.14%	3.71%	5.10%	6.87%
4	2.30%	3.14%	4.18%	4.94%	6.80%	9.16%
5	3.05%	4.17%	5.55%	6.57%	9.04%	12.17%
6	4.06%	5.56%	7.40%	8.75%	12.04%	16.21%

12% CORRELATION

Grade/γ	50%	75%	90%	95%	99%	99.9%
1	1.22%	2.00%	3.14%	4.04%	6.41%	9.49%
2	1.61%	2.65%	4.15%	5.33%	8.47%	12.54%
3	2.15%	3.53%	5.52%	7.10%	11.27%	16.69%
4	2.86%	4.70%	7.36%	9.46%	15.02%	22.25%
5	3.80%	6.25%	9.78%	12.57%	19.96%	29.56%
6	5.07%	8.32%	13.02%	16.75%	26.58%	39.37%

TABLE 5E: PD FROM THE MARGIN OF CONSERVATISM METHOD FOR 7-GRADE CASE

ZERO CORRELATION

Grade	50%	75%	90%	95%	99%	99.9%
1	0.30%	0.33%	0.35%	0.37%	0.40%	0.43%
2	0.46%	0.50%	0.54%	0.56%	0.61%	0.66%
3	0.69%	0.76%	0.81%	0.85%	0.91%	0.99%
4	1.06%	1.15%	1.23%	1.29%	1.39%	1.51%
5	1.60%	1.75%	1.88%	1.96%	2.11%	2.29%
6	2.45%	2.66%	2.85%	2.98%	3.21%	3.48%
7	3.73%	4.06%	4.34%	4.54%	4.89%	5.30%

4% CORRELATION

Grade	50%	75%	90%	95%	99%	99.9%
1	0.34%	0.46%	0.61%	0.73%	1.00%	1.35%
2	0.52%	0.71%	0.94%	1.11%	1.53%	2.07%
3	0.78%	1.06%	1.41%	1.67%	2.30%	3.10%
4	1.18%	1.62%	2.15%	2.54%	3.50%	4.72%
5	1.80%	2.46%	3.28%	3.88%	5.34%	7.18%
6	2.74%	3.74%	4.98%	5.89%	8.10%	10.91%
7	4.16%	5.70%	7.58%	8.97%	12.34%	16.62%

12% CORRELATION

Grade	50%	75%	90%	95%	99%	99.9%
1	0.42%	0.69%	1.08%	1.39%	2.21%	3.27%
2	0.65%	1.06%	1.66%	2.13%	3.39%	5.02%
3	0.97%	1.59%	2.49%	3.20%	5.08%	7.53%
4	1.47%	2.42%	3.79%	4.87%	7.73%	11.45%
5	2.25%	3.69%	5.77%	7.42%	11.78%	17.45%
6	3.41%	5.60%	8.77%	11.27%	17.89%	26.50%
7	5.19%	8.53%	13.35%	17.16%	27.25%	40.35%

It can be seen from the table that a similar pattern of PD estimates arises when compared to the *most prudent* method. That is, the estimated PD increases with the level of asset correlation as well as with the level of confidence used.

2.2.4 Results From the Hybrid Models

In Roengpitya (2012), there are *three* types of hybrid models–*hybrid MLE, hybrid forward and hybrid backward*. We then estimated the PDs for both the 6 and 7-rating grade cases, using each type of hybrid models. First, the results for the *hybrid MLE method*, assuming first the zero correlation and then the correlation of 4% and 12%, with different confidence levels–50%, 75%, 90%, 95%, 99%, and 99.90%, are presented in Table 6A.

TABLE 6A: RESULTS OF PD ESTIMATION USING *HYBRID MLE METHOD* FOR 6 AND 7-GRADE SOVEREIGN PORTFOLIOS

Independent default events (asset correlation=0)

Grade	50%	75%	90%	95%	99%	99.90%	Grade	50%	75%	90%	95%	99%	99.90%
1	1.83%	1.94%	2.05%	2.12%	2.27%	2.46%	1	1.83%	1.94%	2.05%	2.12%	2.27%	2.46%
2	2.48%	2.62%	2.77%	2.87%	3.07%	3.32%	2	2.48%	2.62%	2.77%	2.87%	3.07%	3.32%
3	3.10%	3.28%	3.47%	3.60%	3.85%	4.16%	3	2.96%	3.13%	3.32%	3.44%	3.68%	3.98%
4	3.61%	3.83%	4.07%	4.23%	4.55%	4.95%	4	3.10%	3.28%	3.47%	3.60%	3.85%	4.16%
5	**3.33%**	**3.63%**	**3.95%**	**4.17%**	**4.55%**	5.16%	5	3.43%	3.63%	3.85%	4.00%	4.29%	4.65%
6	5.33%	7.55%	10.32%	12.32%	16.73%	22.55%	6	3.61%	3.83%	4.07%	4.23%	4.55%	4.95%
							7	4.14%	4.42%	4.72%	4.92%	5.33%	5.83%

Asset correlation=4%

Grade	50%	75%	90%	95%	99%	99.90%	Grade	50%	75%	90%	95%	99%	99.90%
1	2.57%	3.12%	3.96%	4.54%	6.13%	8.14%	1	2.57%	3.12%	3.96%	4.54%	6.13%	8.14%
2	3.40%	4.10%	5.13%	5.85%	7.77%	10.18%	2	3.40%	4.10%	5.13%	5.85%	7.77%	10.18%
3	4.17%	5.00%	6.20%	7.04%	9.24%	11.99%	3	4.01%	4.81%	5.98%	6.79%	8.93%	11.61%
4	4.78%	5.72%	7.05%	7.99%	10.40%	13.41%	4	4.17%	5.00%	6.20%	7.04%	9.24%	11.99%
5	**4.34%**	**5.28%**	**6.54%**	**7.48%**	**9.79%**	**12.83%**	5	4.57%	5.47%	6.76%	7.66%	10.01%	12.92%
6	5.84%	8.58%	12.15%	14.85%	20.95%	29.30%	6	4.78%	5.72%	7.05%	7.99%	10.40%	13.41%
							7	5.40%	6.45%	7.90%	8.95%	11.56%	14.84%

Asset correlation=12%

Grade	50%	75%	90%	95%	99%	99.90%	Grade	50%	75%	90%	95%	99%	99.90%
1	3.83%	5.19%	7.52%	9.06%	13.96%	19.81%	1	3.83%	5.19%	7.52%	9.06%	13.96%	19.81%
2	4.91%	6.56%	9.33%	11.13%	16.73%	23.30%	2	4.91%	6.56%	9.33%	11.13%	16.73%	23.30%
3	5.89%	7.79%	10.92%	12.95%	19.09%	26.23%	3	5.69%	7.53%	10.59%	12.57%	18.61%	25.62%
4	6.65%	8.73%	12.11%	14.33%	20.85%	28.47%	4	5.89%	7.79%	10.92%	12.95%	19.09%	26.23%
5	**6.11%**	**8.09%**	**11.26%**	**13.51%**	**19.81%**	**27.66%**	5	6.40%	8.41%	11.72%	13.86%	20.26%	27.70%
6	7.16%	11.01%	16.27%	20.32%	29.81%	41.30%	6	6.65%	8.73%	12.11%	14.33%	20.85%	28.47%
							7	7.41%	9.67%	13.28%	15.69%	22.57%	30.66%

The results from Table 6A indicate that the *hybrid MLE method* share a similar rank-ordering problem as seen in the *most prudent* method for the 6-grade case, as shown in the bold font. This is because it fails to meet the necessary and sufficient conditions outlined in Roengpitya (2012).[7] Also, the rank-order problem appears more frequently as the asset correlation increases. However, for the 7-grade case, the rank order problem disappears–the results which are similar to the case of the *most prudent* estimation.

Table 6B exhibits the PD estimates using the *hybrid forward method* and various asset correlation levels for both the 6-grade and 7-grade cases.

[7] As an example, consider a case between rating grades 4 and 5 with zero correlation. The ratio $\frac{\mathcal{L}(\vec{p}_4^{MLE})}{\mathcal{L}(\vec{p}_5^{MLE})} = 8.75 \times 10^{-53}$ while $\frac{1}{(p_4)^{28}(1-p_4)^{720}} = 7.71 \times 10^{51}$. Therefore, the multiplication of both terms, $\frac{\mathcal{L}(\vec{p}_4^{MLE})}{\mathcal{L}(\vec{p}_5^{MLE})} \times \frac{1}{(p_4)^{28}(1-p_4)^{720}}$, yields $8.75 \times 10^{-53} \times 7.71 \times 10^{51} = 0.674 < 1$. Hence, the required condition for the zero correlation case is violated, leading to the rank order failure.

TABLE 6B: RESULTS OF PD ESTIMATION USING *HYBRID FORWARD METHOD* FOR 6 AND 7-GRADE SOVEREIGN PORTFOLIOS

Independent default events (asset correlation=0)

Grade	50%	75%	90%	95%	99%	99.90%	Grade	50%	75%	90%	95%	99%	99.90%
1	1.69%	1.69%	1.69%	1.69%	1.69%	1.69%	1	1.69%	1.69%	1.69%	1.69%	1.69%	1.69%
2	2.28%	2.28%	2.28%	2.28%	2.28%	2.28%	2	2.28%	2.28%	2.28%	2.28%	2.28%	2.28%
3	2.85%	3.33%	3.66%	3.84%	4.15%	4.51%	3	2.73%	2.73%	2.73%	3.18%	3.64%	4.05%
4	4.32%	4.59%	4.86%	5.03%	5.36%	5.75%	4	2.85%	2.85%	3.45%	3.68%	4.04%	4.42%
5	4.42%	4.78%	5.14%	5.36%	5.82%	6.38%	5	3.74%	4.09%	4.38%	4.56%	4.89%	5.27%
6	14.77%	18.48%	22.32%	24.82%	29.87%	35.98%	6	4.12%	4.46%	4.76%	4.94%	5.30%	5.71%
							7	5.29%	5.62%	5.95%	6.16%	6.57%	7.07%

Asset correlation=4%

Grade	50%	75%	90%	95%	99%	99.90%	Grade	50%	75%	90%	95%	99%	99.90%
1	2.03%	2.03%	2.03%	2.03%	2.03%	2.03%	1	2.03%	2.03%	2.03%	2.03%	2.03%	2.03%
2	2.70%	2.70%	2.70%	2.70%	2.70%	2.70%	2	2.70%	2.70%	2.70%	2.70%	2.70%	2.70%
3	5.30%	7.26%	9.48%	10.87%	13.19%	15.02%	3	3.19%	3.19%	5.54%	7.05%	10.35%	13.34%
4	9.82%	11.86%	13.67%	14.75%	16.43%	17.91%	4	3.32%	5.54%	7.84%	9.48%	12.27%	14.62%
5	**8.86%**	**10.87%**	**12.81%**	**13.97%**	**16.02%**	18.06%	5	7.45%	9.69%	11.95%	13.19%	15.26%	16.85%
6	18.21%	23.44%	28.95%	32.57%	39.84%	48.34%	6	8.09%	10.42%	12.67%	13.92%	15.98%	17.65%
							7	11.29%	13.59%	15.60%	16.75%	18.58%	20.29%

Asset correlation=12%

Grade	50%	75%	90%	95%	99%	99.90%	Grade	50%	75%	90%	95%	99%	99.90%
1	2.81%	2.81%	2.81%	2.81%	2.81%	2.81%	1	2.81%	2.81%	2.81%	2.81%	2.81%	2.81%
2	3.64%	3.64%	3.64%	3.64%	3.64%	3.64%	2	3.64%	3.64%	3.64%	3.64%	3.64%	3.64%
3	13.04%	18.58%	24.39%	26.63%	31.87%	33.87%	3	4.23%	9.14%	15.16%	19.50%	26.66%	32.00%
4	22.39%	27.16%	32.24%	33.84%	35.83%	37.65%	4	8.62%	13.86%	20.62%	24.63%	31.04%	33.46%
5	**18.05%**	**23.58%**	**28.16%**	**31.24%**	**34.87%**	37.69%	5	17.25%	23.52%	27.82%	31.93%	34.37%	36.28%
6	25.52%	33.70%	42.12%	47.37%	56.97%	66.58%	6	**16.67%**	**23.11%**	28.09%	32.24%	35.12%	37.20%
							7	21.90%	28.05%	33.19%	35.75%	38.31%	40.49%

The PD estimates from the *hybrid forward method* are more conservative than the PD estimates from the *hybrid MLE method* as well as from the *most prudent* method, assessed at the same confidence level $1-\gamma$. Since the PD estimates are more conservative than the *most prudent* principle, this implies that the severity of the rank-order problem present in the *most prudent* method can be lessen, as seen in the well-ranked PD estimates for the 6-grade case here. Similar to the case seen before, the more conservative the PD estimates are, the chance that they will retain their rank ordering increases, which is what happens here.[8]

Finally, Table 6C presents the PD estimates using the *hybrid backward method*.

TABLE 6C: RESULTS OF PD ESTIMATION USING *HYBRID BACKWARD METHOD* FOR 6 AND 7-GRADE SOVEREIGN PORTFOLIOS

Independent default events (asset correlation=0)

Grade	50%	75%	90%	95%	99%	99.90%	Grade	50%	75%	90%	95%	99%	99.90%
1	1.69%	1.69%	1.69%	1.69%	1.69%	1.69%	1	1.69%	1.69%	1.69%	1.69%	1.69%	1.69%
2	2.28%	2.28%	2.28%	2.28%	2.28%	2.28%	2	2.28%	2.28%	2.28%	2.28%	2.28%	2.28%
3	2.85%	2.85%	3.32%	3.54%	3.89%	4.26%	3	2.73%	2.73%	2.73%	2.73%	3.31%	3.79%
4	3.87%	4.16%	4.42%	4.59%	4.92%	5.32%	4	2.85%	2.85%	2.85%	2.94%	3.60%	4.04%
5	**3.65%**	**3.97%**	**4.33%**	**4.52%**	4.96%	5.50%	5	3.15%	3.15%	3.72%	3.95%	4.32%	4.73%
6	5.33%	7.55%	10.32%	12.32%	16.73%	22.55%	6	3.30%	3.30%	3.82%	4.07%	4.48%	4.92%
							7	4.14%	4.42%	4.72%	4.92%	5.33%	5.83%

[8] In addition, shown as underlined numbers, the PD estimates for all confidence levels of rating grades 1 and 2 and some lower confidence levels of grades 3 and 4 are the same. This is because the constrained PD estimates fail to yield the likelihood required, as mentioned in the final note in Roengpitya (2012). Hence, the PD estimates taken here are the PDs that yield the highest likelihood possible, even though it will be lower than the required maximum likelihood.

TABLE 6C: RESULTS OF PD ESTIMATION USING *HYBRID BACKWARD METHOD* FOR 6 AND 7-GRADE SOVEREIGN PORTFOLIOS

Asset correlation=4%

Grade	50%	75%	90%	95%	99%	99.90%	Grade	50%	75%	90%	95%	99%	99.90%
1	2.03%	2.03%	2.03%	2.03%	2.03%	2.03%	1	2.03%	2.03%	2.03%	2.03%	2.03%	2.03%
2	2.70%	2.70%	2.70%	2.70%	2.70%	2.70%	2	2.70%	2.70%	2.70%	2.70%	2.70%	2.70%
3	3.32%	5.33%	7.21%	8.57%	11.28%	13.87%	3	3.19%	3.19%	3.19%	4.70%	7.84%	11.51%
4	7.28%	8.72%	10.59%	11.84%	14.08%	16.25%	4	3.32%	3.32%	4.36%	5.93%	8.87%	12.27%
5	**5.50%**	**6.73%**	**8.24%**	**9.39%**	**11.87%**	**14.62%**	5	3.64%	5.48%	7.37%	8.64%	11.50%	14.36%
6	5.84%	8.58%	12.15%	14.85%	20.95%	29.30%	6	3.79%	**4.70%**	**6.62%**	**7.80%**	**10.56%**	**13.77%**
							7	5.40%	6.45%	7.90%	8.95%	11.56%	14.84%

Asset correlation=12%

Grade	50%	75%	90%	95%	99%	99.90%	Grade	50%	75%	90%	95%	99%	99.90%
1	2.81%	2.81%	2.81%	2.81%	2.81%	2.81%	1	2.81%	2.81%	2.81%	2.81%	2.81%	2.81%
2	3.64%	3.64%	3.64%	3.64%	3.64%	3.64%	2	3.64%	3.64%	3.64%	3.64%	3.64%	3.64%
3	9.15%	13.10%	18.46%	22.52%	29.01%	32.57%	3	4.23%	4.23%	9.57%	12.95%	22.11%	30.23%
4	15.24%	19.60%	24.71%	27.13%	32.94%	35.62%	4	4.39%	6.36%	11.70%	14.95%	23.55%	31.04%
5	**9.43%**	**12.40%**	**16.29%**	**19.48%**	**25.93%**	**32.60%**	5	8.25%	12.36%	16.99%	20.43%	27.00%	33.38%
6	**7.16%**	**11.01%**	**16.27%**	20.32%	29.81%	42.64%	6	**4.93%**	**8.65%**	**13.01%**	**15.71%**	**23.59%**	**31.91%**
							7	7.41%	9.67%	13.28%	15.69%	22.57%	30.66%

The PD estimates from the *hybrid backward model* are less conservative, if not equal, than the PD estimates from the *hybrid forward model*, as proven in Roengpitya (2012). Also, we can see that the rank-order problem still appears under this method, as shown in bold font.

2.2.5 Comparison of PD Estimates From Various Methods

The comparison between the hybrid models and the *most prudent* method yields that, for most rating grades, the *hybrid MLE* and the *hybrid backward* produce higher PDs, except for the highest risk grade with confidence intervals greater than 75% as well as for the constrained PD case of the *hybrid backward* method. However, for the *hybrid forward* method, the PDs are mostly more conservative, except for the constrained PD case also. These comparison results are applicable for all levels of asset correlation tested.

Next, we considered the *margin of conservatism* approach. Putting the estimates side-by-side, it can be seen that the three *hybrid models* give more conservative PDs than those calculated from the *margin of conservatism* method in all rating grades and confidence levels. In addition, the differences can be huge for the highest risk grade and confidence level. For example, for the 6-grade case with zero correlation, the PD estimates for the 6th grade at 99.90% confidence using the *hybrid MLE* and *hybrid backward* differ by at least 4 times and about 5 times for the *hybrid foward* case. This is a result of using the *most prudent* principle in our hybrid calculation. On the contrary, the *margin of conservatism* approach will not have the rank-order problem as in the hybrid models as long as the initial estimated PDs rank-order properly.

As discussed in Roengpitya (2012), the PD estimates from *hybrid models* cannot be compared directly to the original maximum likelihood introduced by Forrest (2005) due to the intensive computational requirement. However, since we apply the *most prudent* method to the maximum likelihood estimation, hereby assuming that the PD of rating grade i behaves such that $p_i = p_{i+1} = ... = p_N$ with N being the riskiest rating grade, we do make a conjecture that the PD estimates from the hybrid models should be more conservative, especially for the *hybrid forward method*, than the PDs from the original maximum likelihood. This conjecture needs further proving, nevertheless. In addition, since we have not calculated the actual PDs from the original likelihood method, we do not know the magnitude of the conservatism of our hybrid PD estimates when compared to Forrest's calculation.

Finally, the PD estimates from Van Der Burgt (2007)'s *CAP curve* are relatively close to the *margin of conservatism* estimates at 99.90% with zero correlation and are much less conservative than the *most prudent* estimation. Consequently, they are also lower than the hybrid models. This is because the CAP curve method is based on the principle that the estimates should also reflect the long-term average default rate. So the estimated PDs are calibrated toward the central tendency of the portfolio–the effect which has yet to be taken into account in the hybrid models. Other than comparing in this fashion, nothing more can be said about the effect of applying different values of asset correlation, since the *CAP curve* method does not take into account explicitly the asset correlation in the calculation of its PD estimates.

Section 3: Improvements on Existing Theoretical Approaches Regarding PD Estimation of LDPs

In this section, we provide our theoretical improvements and cautions on *two* of the existing PD estimation of low-default portfolios–the *most prudent estimation* and the *CAP curve* methods.

3.1 Necessary and Sufficient Conditions Ensuring the Rank Ordering of the *Most Prudent Estimation*

As pointed out in Pluto and Tasche (2006), the *most prudent* PD estimation method outlined in the paper does not guarantee the rank ordering of the PD estimates if "the relative number of defaults in one of the better rating grades is significantly higher than those in lower rating grades." We encountered this problem while using this approach to estimate the PDs on our sovereign portfolio for the 6-grade case shown in Section 2. With this in mind, our main task is trying to *identify the necessary and sufficient conditions* that will ensure the *rank ordering* of the PD estimates from the *most prudent* approach. The summary of the *most prudent* calculation method is presented in Section A3 in the appendix. The proposition of the necessary and sufficient conditions is as follows.

Proposition 1 Let n be the number of *cumulative* obligors in a specific rating where the cumulation is taken over the current and all of the worse (more risky) rating grades. Let k be the number of *cumulative* defaults taken over the same rating grade horizon. Let p be the probability of default estimated for this rating grade. The numerical solution of Pluto and Tasche (2006)'s *most prudent* PD estimates will retain the rank-ordering property within a specified level of confidence γ going from the worse rating grade w to the better rating grade b if the following conditions are met.

1. For all $i = 0, 1, ..., k_w$, $i < n_w \cdot p_w$
2. $\Delta f(n, k, \bar{p})|_{i=0}^{k_w} = \sum_{i=0}^{k_w} \left[\binom{n_w + x}{i} \bar{p}^i (1-\bar{p})^{(n_w+x)-i} - \binom{n_w}{i} \bar{p}^i (1-\bar{p})^{n_w - i} \right]$ must be strictly negative, which can be guaranteed through the condition $\left[\frac{\prod_{j=1}^{x}(n_w + j)}{\prod_{l=1}^{x}(n_w - i + l)} \cdot (1-\bar{p})^x - 1 \right] < 0$ for $i = k_w$
3. $\Delta f(n, k, \bar{p})|_{i=0}^{k_w} + \sum_{i=k_w+1}^{k_w+y} \binom{n_w + x}{i} \bar{p}^i (1-\bar{p})^{(n_w+x)-i} < 0$,

where k_w and k_b are the cumulative defaults of the worse and better rating grades consecutively and n_w and n_b are cumulative obligors for each of those grades. $x = n_b - n_w$ represents additional obligors added when going from the rating w to b and $y = k_b - k_w$ is the additional defaults going from the rating w to b. $\bar{p}$ is a fixed PD estimate. The proof of this proposition is in Section A6 in the appendix.

There is one point worth noting. Generally, the estimated PDs will become more conservative as the level of the asset correlation increases and at a higher level of confidence. Therefore, it is not surprising that the rank-ordering property will fail more often at a higher confidence level, if the asset correlation is taken into account, because when such asset correlation is present, the upper bound of PD estimates will need to be increased, especially at a higher level of confidence (like 99.9%). For this case, the direction and speed of an decrease in PD estimates, going from one worse rating grade w to a better rating grade b, become less uniform and more unstable. Therefore, it is the case where the PD estimates of the better rating grade with a higher cumulative observed default rate is much higher than the PD estimates of a worse rating grade with lower cumulative observed default rate, making the rank-ordering problem worse.

3.2 Caution on Using the CAP Curve Calibration Method

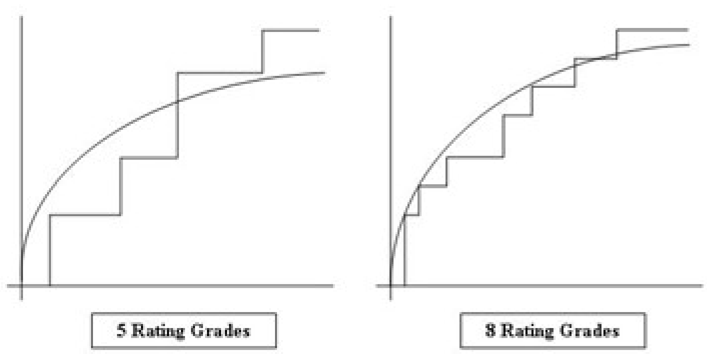
In Van Der Burgt (2007), the author proposed an alternative approach for estimating PDs by calibrating the CAP curve using the following function:

$$y_i(x) = \frac{1 - e^{-kx_i}}{1 - e^{-k}}, \tag{3}$$

where y_i is the cumulative default percentage and x_i is the cumulative percentage of obligors for rating grade i. The curvature of the function in Equation (3) depends on the value k, which is determined by minimizing the following root-mean-squared error (RMSE) for a given value of y_i and x_i for N rating grades:

$$RMSE = F(k, y_i, x_i) = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left\{y_i - \frac{1 - e^{-kx_i}}{1 - e^{-k}}\right\}^2}. \tag{4}$$

One very important observation regarding the minimization of RMSE in Equation (4) is that the value of k that minimizes RMSE is *very sensitive to the number of rating grades N*. In other words, the optimal k, k^*, varies with the granularity of the rating grades or scores. This can become problematic when the modeler has to make a choice between fitting the curve using raw ordinal scores of borrowers or using less-refined rating grades, as the optimal k obtained from each case will be different, leading to different estimates of the PDs. Consequently, the choices of how to fit the curve can affect the value of k, which in turn influences the predicted PDs.



For example, from the above figure, it can be seen that an increase in the number of rating grades from 5 to 8 grades also increases the optimal value of k, as the fitted CAP curve becomes more curvy. Also, note that the actual CAP curve for a few number of ratings generally exhibits a step-function nature.

From the picture, we can also see that the determination of k depends on minimizing the difference between the actual step-function CAP curve and the fitted CAP curve at each point of an increase in the step-function (so-called the *minimization point*). Therefore, intuitively, an increase in the granularity of the rating grades does not guarantee an increase in k. This will have to depend on how the difference at each minimization point can be traded off across all ratings. We therefore propose the following (the proof can be found in the appendix).

Proposition 2 Let y_i be the cumulative default percentage and x_i be the cumulative percentage of obligors for rating grade i. The effect of the change in the granularity of the rating grades used to fit the function $\frac{1-e^{-kx_i}}{1-e^{-k}}$ to the actual step-function CAP curve on the optimal level of k, k^*, depends *solely* on the difference between the fitted curve and the actual CAP curve, $\left\{y_i - \frac{1-e^{-kx_i}}{1-e^{-k}}\right\}$, at each minimization point.

Since different granularity of x, i.e. number of rating grades, can affect the optimal k^*, the next question to ask is how the change in the optimal k^* coming from the change in x can affect the PD estimates using this CAP curve method. We therefore propose the following.

Proposition 3 An increase in the optimal level of k, k^*, will have no effect on the CAP-curve PD estimate for the rating grade that has the cumulative total obligors at $\hat{x}$ which solves $(1 - k^*\hat{x}) = (1 - k^*\hat{x} + k^*)e^{-k^*}$ for a given level of k^*. However, such increase in k^* will lead to higher PD estimates for higher risk obligors (where x is low or near 0) whose rating grades cumulated to $x < \hat{x}$. For lower-risk borrowers (where x is high or near 1) whose rating grades cumulated to $x > \hat{x}$, an increase in k will lower PD estimates for these borrowers.

From the propositions above, it can be seen that the optimal k^* is very sensitive to the level of granularity in calculating the cumulative obligors (i.e. the values of x). Overfitting the actual CAP curve can lead to a high k^* value and, consequently, too concave function, where risky borrowers may be punished too much and good ones too little. On the other hand, a low value of k^* means that the accuracy ratio may be too low and the risky obligors' estimated PDs may be too low. Given these challenges in using the CAP-curve method in estimating the PDs for low-default portfolios, we then propose our views as bank supervisors regarding the usage of this method in Section 4.

To illustrate our claim further, we used our sovereign portfolio as an example. To modify the estimated PD in the high risk grades 4 through 6, we decide to relax the RMSE minimization condition, i.e. increasing the degree of concavity k. However as shown in Proposition 3, we need to derive the critical cumulative obligor, $\hat{x}$, which represents the cutoff value that will determine the effect of a change in k on the estimated PD. That is, for each rating grade, if the level of cumulative obligors is less than $\hat{x}$, an increase in k will increase the estimated PD and vice versa. Based on our empirical result for the 6-grade portfolio, given the estimated k at 3.1426, we derived $\hat{x}$ from equation $(1 - k^*\hat{x}) = e^{-k}(1 - k^*\hat{x} + k^*)$ and obtained $\hat{x}$ equal to 27.31%. According to Table 4 previously shown in Secion 2.2.2, an increase in k will yield a higher estimated PD only for grades 5 and 6. In addition, the increase in k will severely lower the estimated PD for Grade 4. So if we intend to increase the estimated PD for the 4th rating grade, we should decrease k instead.

TABLE 7: RESULTS OF CAP PD ESTIMATION WHEN INCREASING K

SA Grade	ODR	CAP PD (k = 2)	CAP PD (k* = 3.1426)	CAP PD (k = 4)
1	0.000%	0.686%	0.361%	0.212%
2	0.310%	1.050%	0.704%	0.498%
3	1.557%	1.436%	1.152%	0.931%
4	3.743%	2.016%	1.963%	1.834%
5	2.941%	3.061%	3.782%	4.226%
6	2.941%	3.871%	5.469%	6.759%
Average	1.688%	1.676%	1.659%	1.640%

Modified SA Grade	ODR	CAP PD (k = 3)	CAP PD (k* = 4.2726)	CAP PD (k = 5)
1	0.000%	0.290%	0.178%	0.110%
2	0.196%	0.576%	0.411%	0.292%
3	0.735%	0.769%	0.585%	0.442%
4	1.176%	0.952%	0.758%	0.599%
5	2.101%	1.233%	1.040%	0.867%
6	2.036%	1.671%	1.508%	1.339%
7	3.758%	3.522%	3.747%	3.885%
Average	1.688%	1.627%	1.593%	1.555%

Similar to the 6-grade case, the 7-grade estimated PD is quite low compared to the observed default rate, especially for the higher risk grades. Using the optimal k^* for this case, we get that $\hat{x} = 21.75\%$. This means that an increase in k leads to a higher PD estimate for the 7th grade only and yields lower PDs for the others. Table 7 above presents the comparison of estimated results using different values of k for both 6-grade and 7-grade cases.

Even though a decrease in k seems to provide better PD estimates, as shown in Table 7, but we have to be concern about the implied predictive power of the fitted CAP curve as well. According to the relationship between the concavity, k and the accuracy ratio (AR) derived from the fitted CAP curve, we can approximate the AR for each type of our portfolio from the equation $AUC = \int_0^\infty \left(\frac{1-e^{-kx}}{1-e^{-k}}\right) dx = \frac{1}{1-e^{-k}} - \frac{1}{k}$, where $AR = 2\left[\frac{1}{1-e^{-k}} - \frac{1}{k}\right] - 1$.

Practically for non-retail portfolios, AR should lie within the range of $40 - 80\%$, which corresponds to the boundary of k approximately between $2.67 - 10$ from the above equation. As a result, setting k lower than 2.67 might not be appropriate for the PD estimation via the CAP curve method. In case of the 6-grade portfolio, for example, we try to reduce k from the optimal level of 3.1426 to 2 in order to increase the estimated PD for the 4th grade. However, when lowering k to be less than 2.67, the calculated AR stands at only 31.31%, which is lower than the minimum acceptable threshold of 40% so the estimated PD from $k = 2$, in our opinion, should not be used.

To summarize, **in employing the CAP curve method, one needs to think about the trade off between *two* aspects. On one hand, there is the accuracy of the estimated PD coming from both the curvature and the level of the fitted CAP curve.** The curvature of the curve relies on the k value, which can be adjusted to increase the accuracy as previously discussed. The level of the fitted curve represents how well the average estimated PD reflects the central tendency of the portfolio, denoted as $\langle D \rangle$ in the PD estimation formula previously mentioned in Section 2. The adjustment on the value of $\langle D \rangle$ will lead to the shift of the fitted CAP curve, not its curvature. **On the other hand, there is an issue regarding the AR sensitivity of the fitted CAP curve** previously mentioned and proved in the propositions. If this trade off does not yield a fruitful outcome, it can be the case that the existing rating methodology may not be effective enough and there is a need to fully review the rating methodology, such as factors used in rating model, the credibility of risk factors, etc., instead of trying to manipulate k, $\langle D \rangle$ or AR values.

Section 4: Regulatory Recommendations for Validating PD Estimates

First and foremost, in order to verify that the PD estimates are sufficiently conservative, bank supervisors need to understand well the theoretical foundation underlying each estimation method, its suitability to the credit portfolio in consideration, and the possible channel where the shortfall might occur. This understanding is crucial for identifying the advantages and downsides of each method.

For instance, the *most prudent* method yieds conservative PD estimates which can beneficially vary with the level of asset correlation and can be calculated despite having no default history but however cannot guarantee rank ordering. The *hybrid models* possess similar characteristics but have an add-on feature of the maximum likelihood concept with less computation intensity. The *margin of conservatism* provides a logical way to scale up the PD estimates but banks will need to have their initial PD estimates to begin with. Finally, the CAP curve has a comprehensive way of estimating the PDs with few default data but cannot be used if there is no default history and the estimates can be sensitive to the level of granularity.

More specifically, from what we observed in Propositions 2 and 3 of Section 3.2, there can be a case where the different granularity used to fit the exponential function to the actual CAP curve can lead to different estimates of k and consequently the PDs in Van Der Burgt (2007). If the fitting leads to an increase in the optimal k, k^*, then there is an increase in the PD estimates of higher risk obligors and a decrease in PD estimates for the lower risk obligors. Given these conditions, the accuracy ratio (AR) of the fitted curve should be higher for a higher optimal k^*.

Therefore, when it comes to validating whether the financial institutions which choose the CAP curve method have set an appropriate granularity (for example using raw scores versus score bands) in fitting the function to the actual CAP curve, regulators need to consider both the optimal level of k^* obtained, the PD estimates as compared to the historical default rates, and the accuracy ratio. To prevent the underfitting or overfitting of the CAP curve from differences in the granularity levels, **we recommend that the accuracy ratio of the fitted CAP curve should be between** 40% **and** 80%, as targeted by best-practice standards.

By restricting the AR value in this fashion, we can ensure that the curve will not be underfitted too much that it cannot differentiate between good and bad borrowers whereas will not be overfitted so that it punishes good obligors too little and risky obligors too much. In fact, the upper bound of the AR should ensure that the PDs for the low-risk borrowers will not be too low such that one default happening in a lower risk grade will lead to a backtesting failure as the observed default rate suddenly surges and becomes higher than the estimated PD.

In summary, this paper simply puts in perspective for bank supervisors how each existing estimation method for LDPs yields different PD estimates and how they measure up to each other. This information should be used as a part of the evaluation of PD models which should follow by constant PD backtesting to make certain that the estimated PDs obtained from the model still reflect well the risk characteristics of the borrowers and bank capital for absorbing the credit risk of LDPs is deemed sufficient.

Concluding Remark

The essence of this paper is comparison of the estimated PD values from various methods on low-default portfolios. The models used in the comparison are Pluto and Tasche (2006), Van Der Burgt (2007), Benjamin, Cathcart and Ryan (2006) and Roengpitya (2012), using the historical data of sovereign borrowers from the years 1975-2009. Under some circumstances outlined in Roengpitya (2012) paper, the PD estimates of the hybrid models can be *more conservative* than the PD estimates from the *most prudent* method, while Van Der Burgt's *CAP Curve* method seemed to yield the *least conservative* PD estimates. In addition, we also provided the necessary and sufficient conditions in order to ensure the rank order of the PD estimates from the *most prudent* method, as well as raising an important caution in employing Van Der Burgt (2007)'s CAP curve calibration, as the choice of granularity of cumulative borrowers can affect the PD estimates. We then recommend that the upper and lower bound of the fitted CAP curve should be between the best practice range of 40% to 80%.

The issue of PD estimates for low-default portfolios (LDPs) has been a challenge to financial institutions and regulators alike, both in terms of coming up with appropriate PD estimates with sufficient conservatism and in terms of validating the estimates. Therefore, both financial institutions and regulators should be aware of the consequence of employing these methods for PD estimations and should be able to identify incidences where the PD estimates may be too low to cover the default risk. In all, these theories are the existing estimation tools, which should be used in conjunction with other prudent qualitative judgments in order to ensure that the risk will not be underestimated for this specific type of borrowers.

References

Basel Committee on Banking Supervision (BCBS). (2005) "Validation of Low-default Portfolios in the Basel II Framework." *Basel Committee Newsletter.* Bank of International Settlements. No. 6. September.

Benjamin, Nathanaël, Alan Cathcart and Kevin Ryan. (2006) "Low Default Portfolios: A Proposal for Conservative Estimation of Default Probabilities." *Financial Services Authority.*

Duncan, Edward, John Phipps and Katherine Seal. (2005) "Low Default Portfolios." *Joint Industry Working Group Discussion Paper.* British Banker Association, London Investment Banking Association and International Swaps and Derivatives Association.

Dwyer, Douglas W. (2006) "The Distribution of Defaults and Bayesian Model Validation." *Working Paper Series.* Moody's KMV Company.

Efron, Bradley. (1994) "Missing Data, Imputation, and the Bootstrap." *Journal of the American Statistical Association.* Vol. 89, No. 426, pp. 463-475.

Financial Services Authority. (2005) "CRSG Expert Group paper on Low Default Portfolios." *Financial Services Authority Publication.*

Forrest, Alan. (2005) "Likelihood Approaches to Low Default Portfolios." *Working Paper.* Dunfermline Building Society.

Ieraci, Rocky and Bogie Ozdemir. (2008) "Basel II Validation - Part II: Conceptual Soundness and Operations." *Standard & Poor's Presentation.*

Lucas, Alan. (2007) "Basel Compliant Modeling with Little or No Data." *Paper for the Credit Research Center (CRC) Conference 2007.* pp. 1-21.

Papanastasopoulos, George. (2007) "Using Option Theory and Fundamentals to Assessing Default Risk of Listed Firms." *Munich Personal RePEc Archive Paper Series.* MPRA Paper No. 453.

Pluto, Katja and Dirk Tasche. (2006) "Estimating Probabilities of Default for Low Default Portfolios." *The Basel II Risk Parameters: Estimation, Validation and Stress Testing.* Springer Publishing. pp. 79-103.

Ricke, Marcus and Georg Von Pföstl. (2007) "Quantitative Validation of Rating Models for Low Default Portfolios through Benchmarking." *Financial Stability Report of the Oesterreichische Nationalbank.* Vol. 14, pp. 117-125.

Roengpitya, Rungporn. (2012) "Proposal of New Hybrid Models for PD Estimates on Low Default Portfolios (LDPs) and theoretical PD comparisons." *Bank of Thailand.* Discussion Paper (Forthcoming).

Savona, Roberto and Marika Vezzoli. (2008) "Multidimensional Distance to Collapse Point and Sovereign Default Prediction." *Working Paper.* Carefin No. 12/08, Milano.

Schuermann, Til and Samuel Hanson. (2004) "Estimating Probabilities of Default." *Federal Reserve Bank of New York Staff Reports.* Federal Reserve Bank of New York. No. 190.

Standard & Poor's Credit Risk Solutions. (2005) "Using Expert Judgment to Create LGD Estimates Helps Overcome Lack of Data." *S&P's Credit Risk Solution Publication.* Issue II.

Tasche, Dirk. (2008) "Validation of Internal Rating Systems and PD Estimates." *The Analytics of Risk Model Validation.* Elsevier Finance Quantitative Finance Series. pp. 169-196.

Van der Burgt, Marco J. (2007) "Calibrating Low-Default Portfolios using the Cumulative Accuracy Profile." *The Journal of Risk Model Validation.* Vol. 1, No. 4, pp. 17-33.

Appendix

A1: Qualitative Approach to LDPs: Expert Judgment

Since developing and backtesting the PD model for low-default portfolios can be difficult, some banks have chosen to employ the expert judgment in this case. The ratings for obligors will be assigned "subjectivity" or based on the judgment of the experts involved in the rating assignment process. However, the downside of this approach is the inconsistency of the final rating of obligors when rated by different officers, since the process itself is not very transparent. To minimize this inconsistency problem, a few things can be done. In some cases, experts will follow a loose guideline before making the final decision on the rating. In other cases, the final rating is benchmarked against the ECAI's ratings or, if possible, ratings of other peer banks.

The use of the expert judgment can range from using pure expert judgment alone in determining the rating or using the judgment together with or as a part of the quantitative approach for the rating assignment. This approach, though difficult to verify, is preferable for the LDP setting where the data is scarce; for example, for the portfolios of large corporations, sovereigns, or financial institutions. The most important aspect of the expert judgment model is to test, validate and update it consistently. Even though the data on validation is limited, there may be enough to perform some basic analysis that provides updated weights given to factors in the model and hence retain the claim of its sufficient discriminatory power (BBA-LIBA-ISDA paper). S&P also provides overall criteria in validating the expert judgment model, which consist of the validation of risk drivers and their weights, calibration to historical default rates and other fundamental credit knowledge (Ieraci and Ozdemir (2008)). They also emphasized that, for the LDP, the conceptual validation is the most crucial among the three axioms used in validation, notably conceptual soundness, confirmation of model operations, and outcomes analysis.

A2: Quantitative Approaches to LDPs: Non-theoretical Approaches

Apart from dealing with LDPs using the expert-judgment approach, one can choose to cope with the rare default data quantitatively as follows.

1. **Simple Average Approach**

   The simple average approach here is just using the average of default events throughout the years of data obtained by banks as a proxy for the probability of default. This approach is quantitatively simple but, by the characteristics of LDPs, may not truly reflect the probability of default. The characteristics of LDPs are that the default events rarely happen for this group of good-rating obligors and, once the default do occur, the losses and consequences tend to be severe. Due to these characteristics, the distribution of risk in this group of portfolios is not fully known and should be different from other portfolios. Therefore, using just the simple average alone to assess the PD may be insufficient and misleading.

To make the matter more clear, consider an example of the recent subprime crisis. Prior to the final deal to be acquired by JP Morgan Chase, Bear Stern's long-term rating slipped from 'A' to 'BBB' on March 14. While most U.S. banks weathered well through the Asian Crisis in 1997, using the average of the default history during the past decade (covering the Asian Crisis also) as a proxy for the PD would therefore yield a close-to-zero PD. This level of predicted PD would have been useless when it comes to assessing the performance of Bear Stern amid the subprime crisis, which consequently led to a sudden change in the PD before being bailed out.

2. **Benchmarking with ECAI Ratings or Bond/CDS Prices**

Using external ratings from various rating agencies or market prices can help banks assess the probability of default also. Benchmarking is one of the two quantitative validation methodologies used by regulators– notably backtesting and benchmarking. Whereas backtesting relies on the *internal* data, benchmarking is based on the *external* data.[9]

There are two ways one can perform benchmarking. One is to use the ECAI ratings, which are claimed to be through-the-cycle (TTC) by rating agencies, while the agencies focus mainly on the *stability* of ratings through time. The other way is to use external data derived from the market prices, notably the bond spreads or the credit default swap (CDS) spreads.[10] The use of either bond or CDS spreads is classified to be a point-in-time (PIT) approach, as these prices rapidly respond to changes in the economic environment.

The application of benchmarking by banks can range from applying it directly as a rating associated with a certain obligor or using it as a benchmark for an internal rating system. Since obligors classified as being in

[9] It is worth noting that the BBA-LIBA-ISDA paper provided an alternative methodology related to benchmarking can be done to obtain the rating of obligors under LDP. The logic of this method is that a central tendency (or long run average PD) from a model with similar or comparable characteristics can be used to assign a PD for an obligor on the LDP. For example, a large corporate model are used as a basis for assigning PDs for a project finance. This is not exactly like benchmarking, as it uses the internal model in existence, but the method of comparison shares a similar logic.

[10] A bond spread is the difference in yield between a risky bond and a risk-free bond with the same maturity. The riskier the issuer is, the higher the difference will be. CDS is a financial obligation coming from hedging for credit risks. The seller provides the buyer with a certain amount of payment in the event of defaults in exchange for the premium in purchasing that swap. High premium simply reflects a higher probability of default.

LDPs are large, well-performed corporations or government organizations, the ratings or bond/CDS prices of these entities are largely present. Though convenient as it may be, the users of benchmarking techniques should be careful when it comes to the following concerns.

(a) Not all entities have ECAI ratings or bond/CDS prices. Since firms rated by leading external rating agencies are generally large international firms, some local but good performing firms may not possess the ECAI ratings, especially the decent firms in developing countries. Also, some firms may not issue bond or have CDS traded so these prices may not be available. Banks then will need to find alternative ways to come up with ratings for these firms.

(b) ECAI ratings or bond prices do not translate directly to accurate risk characteristics of banks' portfolios. As mentioned previously, the benchmarking can be done through either TTC or PIT approaches. Therefore, the banks need to know first the principles of their internal rating systems whether it's concentrated on being TTC or the hybrid of TTC and PIT and then choose the right benchmarking tool so that it will reflect the true risk characteristics of the portfolios.

    For the TTC application, the ECAI ratings may be an appropriate choice. However, different external rating agencies use various ways to calculate their ratings. The differences may lie in several areas, ranging from the basic concept such as default definitions to very complex methodologies like cohort approaches (factoring in rating migration) versus original approaches or how the agencies view their 'central tendency' concepts, not to mention various criteria used for qualitative analysis. If banks choose to employ the ECAI ratings directly as internal ratings, they need to keep in mind the fact that calibrating the model and obtaining the rating internally yield an advantage in a sense that the internal models usually capture bank-specific risk characteristics better than external rating calibration, which considers rating factors in a more general level. Realizing this gap enables banks to understand the nature of risks they are facing and also considering limits on how and how much they can apply ECAI ratings to their rating assignments. If the difference is small, the application of ECAI ratings should closely approximate the true risk characteristics of banks. This is the case where the application is most beneficial. As mentioned previously, different external rating agencies may have different approaches in assigning ratings.

    For the application of the PIT approach, one must first note that the internal rating system cannot rely on the PIT rating alone. Under Basel II Capital Accord, the internal rating should also reflect somewhat the long-term tendency (or TTC rating) of counterparties. Therefore, that leaves us with the possible "hybrid" models of the PIT and TTC. If the rating systems of banks are in fact hybrids and should banks decided to use both approaches to mimic the credit risk approximation by these models, then the proper application of benchmarking relies on the right weights given to the ECAI ratings and the bond/CDS prices. For example, if the internal rating system's PD is estimated with a fraction of α being PIT and $(1-\alpha)$ being TTC, then the application of the external PIT and TTC benchmarks should be somewhat similar to the following

$$PD_{use} = \alpha PD^B_{PIT} + (1-\alpha) PD^B_{TTC},$$

where PD^B denotes the PD from the external sources to be used as benchmarks. The weight α can either be determined from the total weights given to the PIT and TTC factors used in the calibration of models or be assigned using expert judgment.

In Ricke and Von Pföstl (2007) compared the results of the Deutsche Bank's internal rating with various ECAI ratings and the CDS spreads of 9 sovereigns. Using the modified calculation of Kendall's τ,[11] the authors found that the correlations τ between the bank's internal ratings and S&P, Moody's, Fitch and CDS spreads are 0.81, 0.86, 0.83 and 0.89 respectively. This result implies the *association* of the internal rating system with those of the ECAIs and the spread. In other words, it can only be concluded that the internal rating model has a high discriminatory power if the benchmark also possesses a high discriminatory power. The reverse causal effect should still be examined.

In summary, benchmarking can be used to fill in the missing ratings of obligors belong in the LDPs or benchmarking against the internal ratings. However, the application of ECAI ratings and bond/CDS prices should be executed carefully, as the external ratings may not always reflect the bank-specific risk nature.

[11] There are several ways to measure the correlation (or correspondence) between *two* ordinally scaled variables, notably Spearman's rank correlation, Somer's D or Kendall's τ. Spearman's rank correlation, or Spearman's ρ, measures the correlation *non-parametrically*. In this case, the raw scores are converted to ranks and then the Spearman's ρ is therefore $\rho = 1 - \frac{6\sum d_i^2}{n(n^2-1)}$, where d_i is the distance between the ranks and n is the number of values in each data set. The other two approaches, Somer's D and Kendall's τ, are somewhat related. Both approaches rely on the calculation of the number of pairs that are concordant, P, the number of pairs that are not concordant, Q, the number of ties on variable x, x_0 and the number of ties on variable y, y_0. Kendall's τ is defined as $\tau_K = \frac{P-Q}{\binom{n}{2}}$, where n is the number of observations. When the sample size is large, the denominator $\binom{n}{2}$ needs to be replaced by $\sqrt{\left[\binom{n}{2} - x_0\right] \times \left[\binom{n}{2} - y_0\right]}$. Finally, Somer's D is defined as $d_{yx} = \frac{P-Q}{P+Q+y_0}$, with the assumption that x causes or predicts y. All of these measurements yield the values between $[-1, 1]$ with -1 being perfectly negatively correlated and 1 being perfectly positively correlated.

3. **Enhancing Data Points**

Bootstrapping. One of the frequently-used statistical approaches to deal with the scarcity of data points is the bootstrapping method. Bootstrapping is a way of estimating properties of estimators (mean, variance, etc.) using an approximated distribution. Ideally, if the distribution is identically and independently distributed (*i.i.d.*), bootstrapping involves creating resampled data of the observed dataset by sampling randomly (in some cases with a replacement) from the original dataset. This method usually works well with a large sample size but becomes less reliable in a small sample setting.

The application of bootstrapping to LDPs is mainly to create the otherwise-scarce default event data in order to gain statistical significance in the calibrated models. There are two possible concerns regarding the application of bootstrapping method to LDPs.

(a) Sample size and the number of resampling. In theory, as the sample size (that is the number of data points in a resample) N is approaching infinity, the bootstrap sampling distribution should be quite similar to the sampling distribution. Especially, for a sample size of $N \approx 30 - 50$, the standard error obtained from bootstrapping, σ_{boot}, is (almost) indistinguishable from the true standard error, provided that the original sample is *truly a random one*. However, for the bootstrap mean, μ_{boot}, the larger sample size N is needed.

The observations above lead to the question of how large should the sample size be for the application to LDPs? This surely depends on the availability and the nature of the LDP default data. If the data points are too scarce or the sample is not completely random (which is usually the case for LDPs where the distribution of defaults is skewed), then bootstrapping results can be unreliable. If the such requirements on the data are not a problem, then the next question is how many times should we resample from the sample distribution? Efron and Tibshirani (1993) suggested that, to obtain the reliable bootstrap standard error, one needs to resample about 200 times and about 1000 resampling for the confidence interval which is quite calculation-intensive. Therefore, in order to apply bootstrapping to the LDP case, one should keep in mind the limitations and requirements of the bootstrap method and make sure that sample size and number of resampling are large enough.

(b) Ensuring the predictive power–backtesting. Gaining sufficient data points from bootstrapping does not ensure that the model based on this bootstrapped data will yield a good predictive power and hence good backtesting results. There are two possible reasons why the bootstrap-based model may not predict the PD well. First, as discussed above, specific requirements on the availability and nature of the data need to be met for the bootstrap estimation to be reliable. The failure to control for such factors may eventually lead to unimpressive predictive power due to biased estimations. Second, even if the methodology is well-controlled, there might not be sufficient data to backtest the results of the model or there is a chance that the predicted PD still does not reflect the true PD.

However, bootstrapping provides a mean to generate the non-existing data points but the application of this statistical tool will need to be executed carefully due to the limitation of the methodology and the nature and availability of the data.

Creating multiple observations. It is worth mentioning another issue relating to obtaining additional data points, which is to enhance the data set by performing multiple observation sampling. Lucas (2007) pointed out that enhancing data points by sampling the same obligor at different time periods created the 100% correlation between these data points, since the data points came from the same obligor. The problem is that the usual linear or logistic regression assumes that the data points are independent. Therefore, the model experts will need to employ their own judgment to whether the regression is over-fitted or not from using the highly-correlated sampled data points.

Changing definitions. Another way to obtain additional default data is by expanding the time periods covered and, in some cases, relaxing the default definition, although the calibration will follow the required time period and default definition. Bias can arise from expanding the sample period due to new cohorts entering the data pool. For example, if the time period is expanded from one to two years, then the latest cohort entered will only have one year of history versus other cohorts with two, leading to the bias since one-year data may be quite different from the two-year data. Usually, the difference between the one and two-year cohort is that the one-year cohort will have a better credit history. Lucas (2007) suggested that there are a few ways one can minimize this problem. The better way is to weight the observations or use the survival analysis. The less popular way is to include only the full time period data, which is not preferred since the latest cohort data is excluded unnecessarily.

4. **Modifying Existing Models or Creating New Approaches**

Because of the scarcity of data for LDPs, applying the existing models to this group of obligors directly and obtaining unbiased estimations of parameters may not be possible. Some theory-based models have been developed to cope with these LDPs. Although the application of such models may not have been executed by most banks at present, it is worth mentioning these models as it is possible that they may be used in the future. Some of the models involve modifying the existing credit risk modeling concepts, such as calculating the conservative upper bounds for the PD range (Pluto and Tasche (2006)) or modifying the Bayesian model to fit the LDPs (Dwyer (2006)). Other new approaches have to do with applying the statistical concepts to LDPs, such as the likelihood estimation (Forrest (2005)) or calibrating using cumulative accuracy profile (Van Der Burgt (2007)). These are the models employed in this paper.

A3: Pluto and Tasche's *Most Prudent Estimation* Equations

Pluto and Tasche (2006) discussed in detail one of the best ways to cope with LDPs and gave suggestions to the *most prudent estimation principle*. The authors aimed at providing the upper confidence bounds for the PD estimation, while preserving the rank-order requirement of the PD estimation. This piece of literature provides a possible solution to the problem of estimating the PDs for LDPs and also for portfolios with *no default event*. The model has the following assumptions.

1. **Assumption #1**. There are *three* rating grades A, B, and C, with each rating having n_A, n_B and n_C number of obligors. A is the highest rating while C is the lowest. Let p_A, p_B and p_C be the probability of default for each rating grade.

2. **Assumption #2**. The three ratings above satisfy the rank order requirement. That is, using the fact that A is the highest rating, we must have

$$p_A \leq p_B \leq p_C. \tag{A-1}$$

The inequality above implies that the rank order is correct in a sense that p_A cannot be greater than p_B and p_C. For the case where there are a few default events but no asset correlation, we need to find the upper bound for p_A. Recall that the *most prudent estimation* implies that the condition is $p_A = p_B = p_C$ and therefore we have a homogeneous sample size of $n_A + n_B + n_C$. Next, we need to find first the probability of observing *not more than three* default events, which is simply expressed by the binomial distribution,

$$\sum_{i=0}^{3} \binom{n_A + n_B + n_C}{i} (p_A)^i (1 - p_A)^{(n_A+n_B+n_C)-i}. \tag{A-2}$$

Therefore, the upper bound of p_A can be calculated by solving the following inequality

$$1-\gamma \leq \sum_{i=0}^{3} \binom{n_A+n_B+n_C}{i} (p_A)^i (1-p_A)^{(n_A+n_B+n_C)-i}. \tag{A-3}$$

Similarly, the upper bound for p_B with the sample size $n_B + n_C$ can be obtained by solving

$$1-\gamma \leq \sum_{i=0}^{3} \binom{n_B+n_C}{i} (p_B)^i (1-p_B)^{(n_B+n_C)-i}. \tag{A-4}$$

Finally, for p_C, we will only consider the rating grade C. Therefore, there is only one default event and consequently we need to find the probability of observing *no more than one* default event. So, the upper bound is calculated from

$$1-\gamma \leq \sum_{i=0}^{1} \binom{n_C}{i} (p_C)^i (1-p_C)^{n_C} = (1-p_C)^{n_C} + n_C p_C (1-p_C)^{n_C-1}. \tag{A-5}$$

A4: One Factor Model

One-factor Binomial model. The one-factor mentioned in the correlated default event case is derived from the model used in the BCBS' *Studies on the Validation of Internal Rating Systems*. The idea behind this is to first realize that the log-return of the asset value between period T and period 0, $r_i = \log\left(\frac{A_T^i}{A_0^i}\right)$, depends on the composite factor of firm i, Φ_i, and the residual or the *idiosyncratic* part of the firm's asset value log-return, ε_i.[12] Therefore, we have the following relationship:

$$r_i = \beta_i \Phi_i + \varepsilon_i \quad i = 1, ..., m. \tag{A-6}$$

Here, β_i measures the linear correlation between r_i and Φ_i. Alternatively, one can also look at the *coefficient of determination* of the regression above by looking at the *R-squared*, $R^2 = \frac{\beta_i^2 Var(\Phi_i)}{Var(r_i)}$, of this regression. This R-squared value measures how much the variability of r_i can be explained by Φ_i. Therefore, the relationship if r_i is often written as

$$r_i = R_i \Phi_i + \varepsilon_i \quad i = 1, ..., m. \tag{A-7}$$

Generally, $r_i \sim N(0,1)$, $\Phi_i \sim N(0,1)$ and $\varepsilon_i \sim N(0, 1-R_i^2)$.

Bernoulli loss statistics

Using the framework above, we can use the asset value A_T^i as a latent variable driving the default event. Let $L_i \sim B(1; p_i)$ be the ith element of the vector $\mathbb{L}$=($L_1, ..., L_m$), where $L_i = 1$ with probability p_i and $L_i = 0$ with

[12] Usually, the composite factor Φ_i is the weighted sum of both the industry and country factors and therefore is a proxy for the *systematic risk* of the counterparty i.

probability $1-p_i$. The vector $\mathbb{L}$=$(L_1, ..., L_m)$ is called the *Bernoulli loss statistics.* Using the concept of Bernoulli loss statistics, we can use A_T^i to define the default event as:

$$L_i = \mathbf{1}_{A_T^i < C_i} \sim B(1; \mathbb{P}[A_T^i < C_i]) \quad i = 1, ..., m, \tag{A-8}$$

where C_i is the critical threshold of firm i such that the firm defaults *if and only if* $A_T^i < C_i$.

Recall that $r_i = \log\left(\frac{A_T^i}{A_0^i}\right)$ and A_0^i is some fixed value of firm i's assets at time 0, A_T^i hence has a one-to-one relationship with r_i and we can rewrite the loss statistics as

$$L_i = \mathbf{1}_{r_i < c_i} \sim B(1; \mathbb{P}[r_i < c_i]) \quad i = 1, ..., m, \tag{A-9}$$

where c_i is the threshold C_i after transforming A_T^i to r_i. In addition, we can apply equation (A-7) to the default threshold $r_i < c_i$ to and rewrite it as

$$\varepsilon_i < c_i - R_i\Phi_i \quad i = 1, ..., m. \tag{A-10}$$

<u>Calculating for the PD</u>

Since the probability of default is calculated forward-looking for one year, we are interested in the case where $T = 1$. Let p_i denote the probability of *default* for obligor i. Since the obligor will default if the asset value slips below the threshold, then $p_i = \mathbb{P}(r_i < c_i)$ where $\mathbb{P}(\cdot)$ denotes the probability function. Since $r_i \sim N(0, 1)$, it must be that $p_i = \mathbb{N}(c_i)$, where $\mathbb{N}$ is the *cumulative* normal distribution. Rewriting yields

$$c_i = \mathbb{N}^{-1}(p_i) \quad i = 1, ..., m. \tag{A-11}$$

Substitute equation (A-11) into equation (A-10) and normalize by the standard deviation of ε_i, $\sqrt{1-R^2}$, equation (A-10) is therefore,

$$\tilde{\varepsilon_i} < \frac{\mathbb{N}^{-1}(p_i) - R_i\Phi_i}{\sqrt{1-R_i^2}}, \quad \tilde{\varepsilon_i} \sim N(0, 1). \tag{A-12}$$

Recall from equation (A-10) that obligor i will default if $\varepsilon_i < c_i - R_i\Phi_i$ (which comes from the default condition $r_i < c_i$). However, we have transformed ε_i to be $\tilde{\varepsilon_i}$ by means of normalization. Hence, the obligor i will default if $\tilde{\varepsilon_i} < \frac{\mathbb{N}^{-1}(p_i) - R_i\Phi_i}{\sqrt{1-R_i^2}}$. Since, $\tilde{\varepsilon_i} \sim N(0, 1)$, the probability of default for obligor i is then,

$$p_i(\Phi_i) = \mathbb{N}\left(\frac{\mathbb{N}^{-1}(p_i) - R_i\Phi_i}{\sqrt{1-R_i^2}}\right) \quad i = 1, ..., m, \tag{A-13}$$

where again $\mathbb{N}(\cdot)$ is the cumulative normal distribution. The PD $p_i(\Phi_i)$ depends on the composite factor Φ_i, which is assumed to have some randomness. Equation (A-13) can also be calculated *conditional on* $\Phi_i = z$:

$$p_i(z) = \mathbb{N}\left(\frac{\mathbb{N}^{-1}(p_i) - R_i z}{\sqrt{1-R_i^2}}\right) \quad i = 1, ..., m, \tag{A-14}$$

One factor model

The one factor model assumes that there is only one single factor common to all obligors and consequently the asset correlation between obligors is uniform. Therefore, the composite factors Φ_i of all obligors depend on one single factor, denoted by $Y \sim N(0,1)$. Instead of r_i being defined as in equation (A-7), r_i can be formulated as

$$r_i = \sqrt{\rho}Y + \sqrt{1-\rho}Z_i \quad i = 1, ..., m, \tag{A-15}$$

where r_i can be thought of as a weighted average of the single common factor and the residual component $Z_i \sim N(0,1)$ specific to obligor i. The uniform asset correlation ρ captures the extent of how much the common factor affects the asset return r_i. If assets are perfectly correlated among firms, then $\rho = 1$ and the return on assets will depend solely on the common factor and not the firm-specific residual at all.

The one-factor model treats the asset correlation ρ to be equal to the *R-squared* defined from the regression in equation (A-6), $\rho = R_i$. With the assumption of r_i in equation (A-15), we can modify the PD calculation accordingly. Here, $r_i \sim N(0,1)$ still and the threshold $r_i < c_i$ still applies. So, equation (A-11) still holds. The modification of equation (A-10) comes from $r_i = \sqrt{\rho}Y + \sqrt{1-\rho}Z_i < c_i$ and so

$$Z_i < \frac{c_i - \sqrt{\rho}Y}{\sqrt{1-\rho}} = \frac{\mathbb{N}^{-1}(p_i) - \sqrt{\rho}Y}{\sqrt{1-\rho}}. \tag{A-16}$$

Since Z_i is already normally distributed with $N(0,1)$, we can calculate directly the PD as in equation (A-13) as

$$p_i(Y) = \mathbb{N}\left(\frac{\mathbb{N}^{-1}(p_i) - \sqrt{\rho}Y}{\sqrt{1-\rho}}\right) \quad i = 1, ..., m, \tag{A-17}$$

Calculating the probability for the group of obligors

Using the PD from equation (A-17), we can apply it and calculate the number of defaults in a group of obligors by using the concept of probability. The probability of observing k defaults from a group of n obligors is

$$\binom{n}{k} p_g^k (1-p_g)^{n-k},$$

where p_g is the probability of default for this group of obligors.

Using the the probability defined in (A-17), if one wants to find the probability of observing 1 default ($D = 1$) in a group of n obligors for a specific level of $Y = \bar{y}$, the probability is then

$$\mathbb{P}(D=1) = \binom{n}{1} p_g(Y=\bar{y})(1-p_g(Y=\bar{y}))^{n-1} = \binom{n}{1} \mathbb{N}\left(\frac{\mathbb{N}^{-1}(p_g) - \sqrt{\rho}\bar{y}}{\sqrt{1-\rho}}\right)\left(1 - \mathbb{N}\left(\frac{\mathbb{N}^{-1}(p_g) - \sqrt{\rho}\bar{y}}{\sqrt{1-\rho}}\right)\right)^{n-1} \tag{A-18}$$

Note that the probability of observing no default in a group of n obligors is then

$$\mathbb{P}(D=0) = \left(1 - \mathbb{N}\left(\frac{\mathbb{N}^{-1}(p_g) - \sqrt{\rho}\bar{y}}{\sqrt{1-\rho}}\right)\right)^{n} \tag{A-19}$$

However, if one wants to figure out the probability of observing *no more than 3* defaults in this group of obligors for $Y = \bar{y}$, the probability is then

$$\mathbb{P}(D \leq 3) = \sum_{k=0}^{3} \binom{n}{k} \mathbb{N}\left(\frac{\mathbb{N}^{-1}(p_g) - \sqrt{\rho}\bar{y}}{\sqrt{1-\rho}}\right)^k \left(1 - \mathbb{N}\left(\frac{\mathbb{N}^{-1}(p_g) - \sqrt{\rho}\bar{y}}{\sqrt{1-\rho}}\right)\right)^{n-k} \tag{A-20}$$

Finally, we need to factor in the fact that $Y \sim N(0,1)$ so the probability of observing no more than 3 defaults in this group of obligors for *all possible values of* Y is then

$$\mathbb{P}(D \leq 3) = \int_{-\infty}^{\infty} \sum_{k=0}^{3} \binom{n}{k} \mathbb{N}\left(\frac{\mathbb{N}^{-1}(p_g) - \sqrt{\rho}y}{\sqrt{1-\rho}}\right)^k \left(1 - \mathbb{N}\left(\frac{\mathbb{N}^{-1}(p_g) - \sqrt{\rho}y}{\sqrt{1-\rho}}\right)\right)^{n-k} \varphi(y)dy, \tag{A-21}$$

where equation (A-21) is the integration of equation (A-20) over all possible values of Y and $\varphi(y)$ is the standard normal density function.

A5: Sovereign Debt Crisis Table

Table A1 presents the sovereign debt crisis episodes between the years 1975-2008 used in this study.

TABLE A1: SOVEREIGN DEBT CRISIS EPISODES DURING THE YEARS 1975-2008

Year of the Crisis	Number of Defaults	Countries
1975	1	Zimbabwe
1976	2	Congo, Peru
1978	3	Jamaica, Peru, Turkey
1979	2	Nicaragua, Sudan
1980	2	Bolivia, Peru
1981	8	Costa Rica, Dominican Republic, El Salvador, Honduras, Jamaica, Madagascar, Poland, Romania
1982	7	Argentina, Ecuador, Haiti, Malawi, Mexico, Nigeria, Turkey
1983	13	Brazil, Burkina Faso, Chile, Ivory Coast, Morocco, Niger, Panama, Peru, Philippines, Sierra Leone, Uruguay, Venezuela, Zambia
1984	1	Egypt
1985	1	Cameroon
1986	7	Bolivia, Gabon, Madagascar, Morocco, Paraguay, Romania, Sierra Leone
1987	2	Jamaica, Uruguay
1988	1	Malawi
1989	1	Jordan
1990	2	Uruguay, Venezuela
1991	3	Algeria, Ethiopia, Russia
1994	1	Kenya
1995	1	Venezuela
1997	2	Sierra Leone, Sri Lanka
1998	3	Indonesia, Pakistan, Ukraine
1999	2	Ecuador, Gabon
2000	2	Ivory Coast, Zimbabwe

TABLE A1: SOVEREIGN DEBT CRISIS EPISODES DURING THE YEARS 1975-2008

Year of the Crisis	Number of Defaults	Countries
2001	1	Argentina
2002	2	Gabon, Indonesia
2003	2	Paraguay, Uruguay
2004	3	Cameroon, Grenada, Paraguay
2005	3	Dominican Republic, Grenada, Venezuela
2006	2	Belize, Grenada
2008	2	Ecuador, Seychelles

This table presents the sovereign debt crisis episodes between 1975-2008 used in the study. The data for the years 1975-2002 comes from Savona and Vezzoli (2008) and the data for the years 2003-2008 is from *S&P Sovereign Ratings History* as of December 2009.

A6: Mathematical Proofs of the Necessary and Sufficient Conditions

The mathematical proofs and proposition of the necessary and sufficient conditions for the rank ordering of PD estimates are as follows. Let n be the number of *cumulative* obligors in a specific rating where the cumulation is taken over the current and worse rating grades. Let k be the number of *cumulative* defaults taken over the same rating grade horizon. Let p be the probability of default estimated for this rating grade.

To begin, we will prove the claim that, as long as the number of defaults k is such that $k < n \cdot p$, then as γ increases, so does p. First, recall the equation needed to solve for the estimated p for a specific level of γ,

$$f(n,k,p) = \sum_{i=0}^{k} \binom{n}{i} p^i (1-p)^{n-i} = 1-\gamma. \tag{A-22}$$

As γ increases and $1-\gamma$ decreases, what is left to show is that $f(n,k,p)$ also decreases as p increases to ensure the positive correlation in the relationship of γ and p. Consequently, we will show that $\frac{\partial f(n,k,p)}{\partial p} < 0$ when $k < n \cdot p$.

$$\begin{aligned}
\frac{\partial f(n,k,p)}{\partial p} &= \frac{\partial \left(\sum_{i=0}^{k} \binom{n}{i} p^i (1-p)^{n-i} \right)}{\partial p} \\
&= \sum_{i=0}^{k} \binom{n}{i} i \cdot p^{i-1}(1-p)^{n-i} + \sum_{i=0}^{k} \binom{n}{i} p^i (n-i)(1-p)^{n-i-1}(-1) \\
&= \sum_{i=0}^{k} \binom{n}{i} p^i (1-p)^{n-i} \left(\frac{i}{p}\right) - \sum_{i=0}^{k} \binom{n}{i} p^i (1-p)^{n-i} \left(\frac{n-i}{1-p}\right) \\
&= \sum_{i=0}^{k} \binom{n}{i} p^i (1-p)^{n-i} \left(\frac{i}{p} - \frac{n-i}{1-p}\right) \\
&= \sum_{i=0}^{k} \binom{n}{i} p^i (1-p)^{n-i} \left(\frac{i-np}{p(1-p)}\right)
\end{aligned} \tag{A-23}$$

Clearly, the condition $\frac{\partial f(n,k,p)}{\partial p} < 0$ can be guaranteed when the maximum of the cumulative defaults i, notably k, is less than $n \cdot p$. This is what happens in our *most prudent estimation* for the sovereign portfolio and therefore we observe the monotone increase in the estimates of p as γ increases for each rating grade and for all rating grades.

As for the condition to ensure the *rank ordering* of *the most prudent* PD estimates across rating grades with the same value of γ, we need to calculate the change in the PD estimates, p, if $f(n,k,p)$ changes from the change in n, k, and p, leaving γ unchanged. From Equation (A-22) and with γ being held constant, the total differentiation of this equation yields:

$$\begin{aligned}\frac{\partial f}{\partial n}dn + \frac{\partial f}{\partial i}di + \frac{\partial f}{\partial p}dp &= 0\\ dp &= -\left[\frac{\frac{\partial f}{\partial n}dn + \frac{\partial f}{\partial i}di}{\frac{\partial f}{\partial p}}\right]\\ dp &= \left[\frac{\frac{\partial f}{\partial n}dn + \frac{\partial f}{\partial i}di}{-\frac{\partial f}{\partial p}}\right] \qquad \text{(A-24)}\end{aligned}$$

with $i = 1, 2, ...k$ being the variable reflecting each observed default level. Next, what is left to find is the partial derivatives of all the terms needed in Equation (A-24). Note that the denominator of Equation (A-24) is positive if $k < n \cdot p$, since $\frac{\partial f}{\partial p} < 0$ under such condition.

The partial derivative with respect to p is already calculated in Equation (A-23). Note that the top part of Equation (A-24) measures the change in $f(n,k,p)$ if there is going to be a change in n and $i = 1, 2, ..., k$, holding p constant. The fact that both n and i are positive integers and therefore are discrete variables, measuring the rate of change cannot come simply from differentiating Equation (A-22) with respect to the desired variables. Since the concept of estimating *the most prudent* PD is based on the level of the *cumulative* default rate from the worst to best rating grades, when jumping from the worse rating grade to the better rating grade, both n and k increase at the same time but at different proportions as one incorporates additional obligors while the other accumulates defaulted obligors both from the better rating and the former worse rating(s). This proves to be the key as to why the rank ordering can possibly fail.

Using the concept of the incorporation of obligors from a better rating grade, we can determine the change in $f(n,k,p)$ as a result of the change in both n and k from such addition. Without loss of generality, let n_w be the number of cumulative obligors in an arbitrary worse rating grade with k_w cumulative defaults and let n_b be the number of cumulative obligors in the next better-rated grade with k_b number of cumulative defaults. Define the additional obligors added when going from the rating w to b as $x = n_b - n_w$ and define the additional defaults as $y = k_b - k_w$. To find out how $f(n,k,p)$ has changed due to the additional number of obligors and defaults going from rating grade w to b, holding p constant at $\bar{p}$, we need to compute the following

$$\Delta f(n,k,\bar{p}) = f(n_b,k_b,\bar{p}) - f(n_w,k_w,\bar{p}) = \sum_{i=0}^{k_b}\binom{n_b}{i}\bar{p}^i(1-\bar{p})^{n_b-i} - \sum_{i=0}^{k_w}\binom{n_w}{i}\bar{p}^i(1-\bar{p})^{n_w-i}, \tag{A-25}$$

or writing it in terms of x and y, we have

$$\Delta f(n,k,\bar{p}) = f(n_b,k_b,\bar{p}) - f(n_w,k_w,\bar{p}) = \sum_{i=0}^{k_w+y}\binom{n_w+x}{i}\bar{p}^i(1-\bar{p})^{(n_w+x)-i} - \sum_{i=0}^{k_w}\binom{n_w}{i}\bar{p}^i(1-\bar{p})^{n_w-i}. \tag{A-26}$$

Note that we can rewrite $\sum_{i=0}^{k_w+y}\binom{n_w+x}{i}\bar{p}^i(1-\bar{p})^{(n_w+x)-i}$ as:

$$\sum_{i=0}^{k_w+y}\binom{n_w+x}{i}\bar{p}^i(1-\bar{p})^{(n_w+x)-i} = \sum_{i=0}^{k_w}\binom{n_w+x}{i}\bar{p}^i(1-\bar{p})^{(n_w+x)-i} + \sum_{i=k_w+1}^{k_w+y}\binom{n_w+x}{i}\bar{p}^i(1-\bar{p})^{(n_w+x)-i}. \tag{A-27}$$

Substituting Equation (A-27) into Equation (A-25), we have

$$\begin{aligned}
\Delta f(n,k,\bar{p}) &= \sum_{i=0}^{k_w+y}\binom{n_w+x}{i}\bar{p}^i(1-\bar{p})^{(n_w+x)-i} - \sum_{i=0}^{k_w}\binom{n_w}{i}\bar{p}^i(1-\bar{p})^{n_w-i} \qquad \text{(A-28)}\\
&= \sum_{i=0}^{k_w}\binom{n_w+x}{i}\bar{p}^i(1-\bar{p})^{(n_w+x)-i} + \sum_{i=k_w+1}^{k_w+y}\binom{n_w+x}{i}\bar{p}^i(1-\bar{p})^{(n_w+x)-i} - \sum_{i=0}^{k_w}\binom{n_w}{i}\bar{p}^i(1-\bar{p})^{n_w-i}\\
&= \sum_{i=0}^{k_w}\left[\binom{n_w+x}{i}\bar{p}^i(1-\bar{p})^{(n_w+x)-i} - \binom{n_w}{i}\bar{p}^i(1-\bar{p})^{n_w-i}\right] + \sum_{i=k_w+1}^{k_w+y}\binom{n_w+x}{i}\bar{p}^i(1-\bar{p})^{(n_w+x)-i}
\end{aligned}$$

When considering the first term on the right-hand side of Equation (A-28), i.e. $f(n_b,k_b,\bar{p})$ for all $i=0,1,2,...,k_w$ compared to $f(n_w,k_w,\bar{p})$, the change in $f(n,k,\bar{p})$ will come only from the change in n. This is because, for the first $i=0,1,...,k_w$ terms, we calculate the following:

$$\sum_{i=0}^{k_w}\binom{n_w+x}{i}\bar{p}^i(1-\bar{p})^{(n_w+x)-i} - \sum_{i=0}^{k_w}\binom{n_w}{i}\bar{p}^i(1-\bar{p})^{n_w-i}, \tag{A-29}$$

with the only difference coming from n. However, for the second term on the right-hand side of Equation (A-28), i.e. $f(n_b,k_b,\bar{p})$ for $i=k_w+1,...,k_b$, the change in $f(n,k,\bar{p})$ will come from both the changes in n and k. Note that this additional term is not present before in $f(n_w,k_w,\bar{p})$ and therefore will contribute wholly to the change in $f(n,k,p)$.

We first look at Equation (A-29). The first step is to calculate the binomial formula if n_w increases by x and compare it to the original binomial with n_w. Using the fact that $\binom{n}{i}$ can be written as $\binom{n}{i} = \frac{n\cdot(n-1)\cdots(n-(i-1))}{i\cdot(i-1)\cdots 1}$, such comparison for a specific level of i results in

$$\begin{aligned}
\binom{n_w+x}{i}\bar{p}^i(1-\bar{p})^{(n_w+x)-i} - \binom{n_w}{i}\bar{p}^i(1-\bar{p})^{n_w-i} &= \frac{\prod_{j=1}^{x}(n_w+j)}{\prod_{l=1}^{x}(n_w-i+l)}\binom{n_w}{i}\bar{p}^i(1-\bar{p})^{(n_w+x)-i} - \binom{n_w}{i}\bar{p}^i(1-\bar{p})^{n_w-i}\\
&= \binom{n_w}{i}\bar{p}^i(1-\bar{p})^{n_w-i}\left[\frac{\prod_{j=1}^{x}(n_w+j)}{\prod_{l=1}^{x}(n_w-i+l)}\cdot(1-\bar{p})^x - 1\right]. \qquad \text{(A-30)}
\end{aligned}$$

Since $\binom{n_w}{i}\bar{p}^i(1-\bar{p})^{n_w-i} > 0$, if the term in the parenthesis is less than zero, then the left-hand side of Equation (A-30) will be less than zero also for a specific level of i. Summing the terms across i in Equation (A-30), we have

$$\begin{aligned}\Delta f(n,k,\bar{p})|_{i=0}^{k_w} &= f(n_b,k_b,\bar{p}) - f(n_w,k_w,\bar{p})|_{i=0}^{k_w} \\ &= \sum_{i=0}^{k_w}\binom{n_w+x}{i}\bar{p}^i(1-\bar{p})^{(n_w+x)-i} - \sum_{i=0}^{k_w}\binom{n_w}{i}\bar{p}^i(1-\bar{p})^{n_w-i} \\ \Delta f(n,k,\bar{p})|_{i=0}^{k_w} &= \sum_{i=0}^{k_w}\binom{n_w}{i}\bar{p}^i(1-\bar{p})^{n_w-i}\left[\frac{\prod_{j=1}^{x}(n_w+j)}{\prod_{l=1}^{x}(n_w-i+l)}\cdot(1-\bar{p})^x - 1\right]. \qquad \text{(A-31)}\end{aligned}$$

Therefore, Equation (A-31) represents the contribution to the change in $f(n,k,\bar{p})$ coming from the first $i = 0,1,...,k_w$ terms of Equation (A-28). If the term in the parenthesis in Equation (A-31) is less than zero for $i = k_w$, then it is guaranteed that $\Delta f(n,k,\bar{p})|_{i=0}^{k_w} < 0$ because $\prod_{j=1}^{x}(n_w+j)/\prod_{l=1}^{x}(n_w-i+l)$ is increasing in i, making the term in the parenthesis more positive as i increases. If such condition applies, for all $i = 0,1,...,k_w$, the terms in the parentheses must be less than zero, making the final summation less than zero also.

Next, we consider the second term in Equation (A-28). Since this term is always positive and does not exist before in the calculation for $\sum_{i=0}^{k_w}\binom{n_w}{i}\bar{p}^i(1-\bar{p})^{n_w-i}$, this term will contribute positively to $\Delta f(n,k,\bar{p})|_{i=k_w+1}^{k_w+y}$, making $\Delta f(n,k,\bar{p})$ positive. From equations (A-24), (A-28) and (A-31) as well as using the fact that $-\frac{\partial f}{\partial p} < 0$ is positive if $i < n \cdot p$, we propose the following.

Proposition 1 The numerical solution of the *most prudent* PD estimates will retain the rank-ordering property within a specified level of γ going from the worse rating grade w to the better rating grade b if the following conditions are met.

1. For all $i = 0,1,...,k_w$, $i < n_w \cdot p_w$
2. $\Delta f(n,k,\bar{p})|_{i=0}^{k_w} = \sum_{i=0}^{k_w}\left[\binom{n_w+x}{i}\bar{p}^i(1-\bar{p})^{(n_w+x)-i} - \binom{n_w}{i}\bar{p}^i(1-\bar{p})^{n_w-i}\right]$ must be strictly negative, which can be guaranteed through the condition $\left[\frac{\prod_{j=1}^{x}(n_w+j)}{\prod_{l=1}^{x}(n_w-i+l)}\cdot(1-\bar{p})^x - 1\right] < 0$ for $i = k_w$
3. $\Delta f(n,k,\bar{p})|_{i=0}^{k_w} + \sum_{i=k_w+1}^{k_w+y}\binom{n_w+x}{i}\bar{p}^i(1-\bar{p})^{(n_w+x)-i} < 0$

A7: Mathematical Proofs of the CAP Curve Sensitivity

This section gives detailed mathematical proofs of the propositions in Section 3.2. Recall that, in Van Der Burgt (2007), the CAP curve is calibrated using the following function:

$$y_i(x) = \frac{1 - e^{-kx_i}}{1 - e^{-k}}, \tag{A-32}$$

where y_i is the cumulative default percentage and x_i is the cumulative percentage of obligors for rating grade i. The curvature of the function in Equation (A-32) depends on the value k, which is determined by minimizing the following root-mean-squared error (RMSE) for a given value of y_i and x_i for N rating grades:

$$RMSE = F(k, y_i, x_i) = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left\{y_i - \frac{1 - e^{-kx_i}}{1 - e^{-k}}\right\}^2}. \tag{A-33}$$

To prove our claims in Section 3.2, we first proposed the following.

Proposition 2 The effect of the change in the granularity of the rating grades used to fit the function $\frac{1-e^{-kx_i}}{1-e^{-k}}$ to the actual step-function CAP curve on the optimal level of k, k^*, depends *solely* on the difference between the fitted curve and the actual CAP curve, $\left\{y_i - \frac{1-e^{-kx_i}}{1-e^{-k}}\right\}$, at each minimization point.

Mathematically, we can assess the impact of the sensitivity of the k value by first differentiating Equation (A-33) with respect to k to get the optimal k, k^*, and then see how k^* can possibly change with the change in x_i, whose value reflects the change in the granularity of the ratings. Dropping the subscript i, the differentiation of Equation (A-33) with respect to k yields the following:

$$\begin{aligned}
\frac{\partial F}{\partial k} &= \frac{1}{2}\left[\frac{1}{N}\sum_{allgrades}\left\{y - \frac{1 - e^{-kx}}{1 - e^{-k}}\right\}^2\right]^{-\frac{1}{2}} \\
&\cdot \left[\frac{1}{N}\sum_{allgrades} 2\cdot\left\{y - \frac{1 - e^{-kx}}{1 - e^{-k}}\right\}\left\{-\frac{(xe^{-kx}(1 - e^{-k}) - (1 - e^{-kx})e^{-k}}{(1 - e^{-k})^2}\right\}\right] \\
&= \frac{1}{2}\left[\frac{1}{N}\sum_{allgrades}\left\{y - \frac{1 - e^{-kx}}{1 - e^{-k}}\right\}^2\right]^{-\frac{1}{2}} \\
&\cdot \left[\frac{1}{N}\sum_{allgrades} 2\cdot\left\{y - \frac{1 - e^{-kx}}{1 - e^{-k}}\right\}\left\{\frac{e^{-k} + e^{-kx}(xe^{-k} - e^{-k} - x)}{(1 - e^{-k})^2}\right\}\right].
\end{aligned} \tag{A-34}$$

The optimal value k^* is the value of k that sets Equation (A-34) to zero. Hence, to see how the change in the number of ratings (and hence the change in x) affects k^*, we need to assess how Equation (A-34) will change with the change in x.

From the equation, the first term, $\frac{1}{2}\left[\frac{1}{N}\sum_{allgrades}\left\{y_i - \frac{1-e^{-kx_i}}{1-e^{-k}}\right\}^2\right]^{-\frac{1}{2}}$, is always positive regardless of the change in x as long as $y_i \neq \frac{1-e^{-kx_i}}{1-e^{-k}}$ for all rating grades i. Therefore, the way Equation (A-34) will change with x depends on either the term $\left\{y - \frac{1-e^{-kx}}{1-e^{-k}}\right\}$ or the term $e^{-k} + e^{-kx}(xe^{-k} - e^{-k} - x)$, as $(1-e^{-k})^2$ is always positive.

We claim that the term $f(x) = e^{-k} + e^{-kx}(xe^{-k} - e^{-k} - x)$ is always negative for $x \in [0,1]$ and $k \in [0,\infty)$. We will show this in two steps. First, we show that, for the extreme values of $x = 0$ and $x = 1$, $f(x) = 0$ for all $k \in [0,\infty)$:

$$\begin{aligned} x = 0 \quad &: \quad f(x) = \frac{1}{e^k} - \frac{1}{e^k} = 0 \\ x = 1 \quad &: \quad f(x) = \frac{1}{e^k} + \frac{1}{e^k}\left(\frac{1}{e^k} - \frac{1}{e^k} - 1\right) = 0. \end{aligned}$$

Secondly, we prove that, for all values of $k \in [0,\infty)$, the *slope* of $f(x)$ with respect to x is decreasing where $x \in [0,\bar{x})$ and increasing where $x \in (\bar{x},1]$ for some $\bar{x} = \frac{1}{k} - \frac{1}{e^k-1}$ and for all values of $k \in [0,\infty)$. We begin by calculating the slope of $f(x) = e^{-k} + e^{-kx}(xe^{-k} - e^{-k} - x)$ with respect to x, which is

$$\frac{\partial f(x)}{\partial x} = -ke^{-kx}(xe^{-k} - e^{-k} - x) + e^{-kx}(e^{-k} - 1). \tag{A-35}$$

Since $\frac{\partial f(x)}{\partial x}$ represents the slope of $f(x)$, we now show that $\frac{\partial f(x)}{\partial x} < 0$ for some $x < \bar{x}$. Also, note that $e^{-k} - 1 < 0$. So, $\frac{\partial f(x)}{\partial x} < 0$ when

$$\begin{aligned} -ke^{-kx}(xe^{-k} - e^{-k} - x) + e^{-kx}(e^{-k} - 1) \quad &< \quad 0 \\ -k(xe^{-k} - e^{-k} - x) + (e^{-k} - 1) \quad &< \quad 0 \\ -k\left(x - \frac{e^{-k}}{e^{-k} - 1}\right) + 1 \quad &> \quad 0 \\ -k\left(x + \frac{1}{e^k - 1}\right) + 1 \quad &> \quad 0 \\ x \quad &< \quad \frac{1}{k} - \frac{1}{e^k - 1} = \bar{x}. \end{aligned} \tag{A-36}$$

From Equation (A-36), $\frac{\partial f(x)}{\partial x} < 0$ for $x \in [0,\bar{x})$, $\frac{\partial f(x)}{\partial x} > 0$ for $x \in (\bar{x},1]$ and $\frac{\partial f(x)}{\partial x} = 0$ at $x = \bar{x}$, where $\bar{x} = \frac{1}{k} - \frac{1}{e^k-1}$. Since $f(x) = 0$ where $x = 0$ and $x = 1$ and the slope of $f(x)$ decreases and then increases within the boundary $x \in [0,1]$, $f(x) = e^{-k} + e^{-kx}(xe^{-k} - e^{-k} - x) \leq 0$ for $x \in [0,1]$ and for all $k \in [0,\infty)$.

What is left to show is that $\bar{x} > 0$. With $\bar{x} = \frac{1}{k} - \frac{1}{e^k-1}$, it is obvious that $e^k - 1 > k$ for all $k > 0$. This is because $e^k - 1 = k = 0$ when $k = 0$ and $\frac{\partial e^k - 1}{\partial k} = e^k > \frac{\partial k}{\partial k} = 1$ for $k > 0$ so the function $e^k - 1$ always has a positive slope with the value higher than the slope of k for all $k > 0$. Therefore, $\bar{x} = \frac{1}{k} - \frac{1}{e^k-1} > 0$ for all $k > 0$.

Therefore, using Equation (A-34), how $\frac{\partial F}{\partial k}$ varies with x will depends solely on the term $\left\{y - \frac{1-e^{-kx}}{1-e^{-k}}\right\}$. So, if

x changes, the optimal k, k^*, that sets $\frac{\partial F}{\partial k} = 0$ will be an outcome of the trade off between under and over-fitting of the function $\frac{1-e^{-kx_i}}{1-e^{-k}}$ at each minimization point. Proposition 2 is now proved.

Since different granularity of x can affect the optimal k^*, the next question to ask is how the change in the optimal k^* coming from the change in x can affect the PD estimates using this CAP curve method. We can then assess this by differentiating the following PD equation in Van Der Burgt (2007).

$$PD(R) = \langle D \rangle \frac{ke^{-kx}}{1-e^{-k}}, \tag{A-37}$$

where $\langle D \rangle$ is the observed default rate of rating grade R. Differentiating Equation (A-37) yields

$$\begin{aligned} \frac{\partial PD(R)}{\partial k} &= \langle D \rangle \left[\frac{(e^{-kx} - kxe^{-kx})(1-e^{-k}) - ke^{-kx}e^{-k}}{(1-e^{-k})^2} \right] \\ &= \langle D \rangle\, e^{-kx} \left[\frac{(1-kx)(1-e^{-k}) - ke^{-k}}{(1-e^{-k})^2} \right] \\ &= \langle D \rangle\, e^{-kx} \left[\frac{(1-kx) - (1-kx+k)e^{-k}}{(1-e^{-k})^2} \right]. \end{aligned} \tag{A-38}$$

Since $\langle D \rangle\, e^{-kx}$ and $(1-e^{-k})^2$ are positive, we direct our attention to determining the sign of $(1-kx)-(1-kx+k)e^{-k}$ in order to see whether $\frac{\partial PD(R)}{\partial k}$ is positive or negative.

Let $g(k,x) = (1-kx) - (1-kx+k)e^{-k}$. Consider the extreme values of x. When $x = 0$, $g(k,0) = 1-(1+k)e^{-k} > 0$, since $e^k - 1 > k$ for all $k > 0$. When $x = 1$, $g(k,1) = (1-k) - e^{-k} = 1-e^{-k}-k < 0$. This statement can be proved by using the fact that when $k = 0$, $1-e^{-k} = k = 0$. Then, if we differentiate both $1-e^{-k}$ and k with respect to k, the slope $\frac{\partial(1-e^{-k})}{\partial k} = e^{-k} < 1$ for any $k > 0$. Since the slope of the function k is 1, then the slope of function k exceeds that of the function $(1-e^{-k})$. Therefore, $(1-e^{-k}) < k$ always, given the same starting point $1-e^{-k} = k = 0$ when $k = 0$. Finally, an increase in k will have no effect on the PD estimate of the rating grade at $x = \hat{x}$ where $\hat{x}$ is the value of x that solves $(1-k\hat{x}) = (1-k\hat{x}+k)e^{-k}$ for a given level of k. We therefore propose the following.

Proposition 3 An increase in the optimal level of k, k^*, will have no effect on the CAP-curve PD estimate for the rating grade that has the cumulative total obligors at $\hat{x}$ which solves $(1-k^*\hat{x}) = (1-k^*\hat{x}+k^*)e^{-k^*}$ for a given level of k^*. However, such increase in k^* will lead to higher PD estimates for higher risk obligors (where x is low or near 0) whose rating grades cumulated to $x < \hat{x}$. For lower-risk borrowers (where x is high or near 1) whose rating grades cumulated to $x > \hat{x}$, an increase in k will lower PD estimates for these borrowers.