Designing and Implementing a Basel II Compliant PIT-TTC Ratings Framework

Aguais, Scott

27 January 2008

10

Designing and Implementing a Basel II Compliant PIT–TTC Ratings Framework

Scott D. Aguais, Lawrence R. Forest Jr, Martin King, Marie Claire Lennon, Brola Lordkipanidze[1]

Barclays Capital

INTRODUCTION

Since the first edition of *The Basel Handbook* was published in early 2004, major internationally active banks around the world have continued to engage in substantial projects for designing and implementing the extensive Basel II framework. To achieve the advanced internal ratings based (AIRB) status, banks need to develop a variety of credit models that estimate, for each obligor, probability of default (PD) and, for each credit exposure or facility, loss given default (LGD) and exposure at default (EAD). In developing the required PD models, many banks have had to redesign or refine their risk-rating approaches. In this process, banks have found it necessary to determine whether various PD measures are "point-in-time" (PIT), "through-the-cycle" (TTC) or a hybrid, somewhere between PIT and TTC.[2]

In the first edition of *The Basel Handbook*, we contributed a chapter that introduced concepts for thinking about PIT–TTC PD issues and presented preliminary empirical tests for measuring risk rating accuracy. In extending our original PIT–TTC discussion, we

describe here a framework for calculating both PIT and TTC PDs. To achieve Basel II compliance at an advanced level, we believe that a bank needs to implement this or a comparable framework. The framework described here reflects broadly the one implemented globally in May 2005 at Barclays Capital.[3]

We begin by reviewing the PIT–TTC principles presented in our earlier contribution and highlighting the role played by measures of overall, credit cycle conditions. We then describe a PIT–TTC PD framework encompassing the following components:

❏ development of Basel II compliant PD models;
❏ design of a PD master scale that defines each internal risk-rating as a PD range;
❏ calculation of a comprehensive set of monthly regional and industry credit indices (regional and industry Zs);
❏ determination of monthly translations of agency ratings to one-year PDs as controlled by a single factor for each set of agency ratings (agency Zs);
❏ creation of an approach, drawing on the regional and industry Zs, for converting the PDs from various, primary models into pure PIT and TTC PDs; and
❏ estimation of PD term structures extending out five years by averaging across Monte Carlo simulations of the Z factors and the associated Z-conditional PDs.

We also outline broadly the implementation process that supports the ongoing, monthly updating of the credit cycle and agency Zs, the forecasting of the credit cycle Zs and the estimation of the related PDs and PD term structures.

A REVIEW OF KEY PIT–TTC CONCEPTS
PIT–TTC overview
In January 2001, the Basel consultative document on the proposed IRB approach published by the Basel Committee on Banking Supervision (BCBS) offered the first formal distinction between PIT and TTC PD estimates (see also BCBS 1999; 2000). Up to that point PIT and TTC terminology for risk ratings had been used only informally within the credit ratings and risk literature. While the Basel Committee at that time did not explicitly define the terminology, they did start the process of identifying the detailed credit rating

requirements that would ultimately be needed for satisfying the Basel II framework.

To set the context for describing the design characteristics of an integrated PIT–TTC PD approach, we summarise here some of the key PIT–TTC principles articulated in our first chapter.

Over the last eight to ten years, various studies of bank credit rating systems by the Federal Reserve (see Treacy and Carey 1998) and other regulators and credit risk researchers have uncovered an important distinction between PIT and TTC. Furthermore, they recognise that the industry does not seem to have migrated towards a consistent rating PD approach, as some banks contend that they use mostly PIT ratings while others suggest they use mostly TTC ones. Other banks use some kind of muddled combination of PIT and TTC.[4] We see this lack of consistency as being caused by the fact that, first, not all PD models produce consistent PIT or TTC measures, and, second, not all of the key credit risk objectives that drive a bank's risk management processes can be supported by one type of PD measure – be it PIT or TTC. For a recent discussion of some of the design characteristics of successful PD estimates, which includes the role of PIT–TTC distinctions, see Ranson (2005).

In 2004 we argued that, ultimately, the type of PD estimates required to support various credit risk measures needed to be linked explicitly to the specific objective and its related time horizon. For example, estimates of one-year expected credit losses would most likely be more accurate if they used one-year PIT PDs that fully reflected the current credit conditions prevailing at that time. In contrast, assessing the risk/reward characteristics of a ten-year primary exposure would, if assessed correctly, utilise a ten-year PD term structure. Finally, the implied requirements established by Basel II seem, by our interpretation, to be more focused on TTC PDs, adding further impetus to the need for an integrated PIT–TTC approach. However, the Basel II requirements are still subject to interpretation as we discuss below.

Evidence on credit cycles motivates PIT–TTC distinctions

To start, observe that in distinguishing between PIT and TTC PDs, one presumes the existence of predictable macro level credit fluctuations – that is, the existence of a "credit cycle". By a credit cycle,

we mean that, if PDs in a broad region or industry are unusually high, we can expect them to fall generally. Alternatively, if PDs are unusually low, we can expect them to rise, although this may occur more slowly and less predictably than the cyclical recovery. This discussion assumes that one can identify an equilibrium or "normal" state, which one might estimate with a historical average.
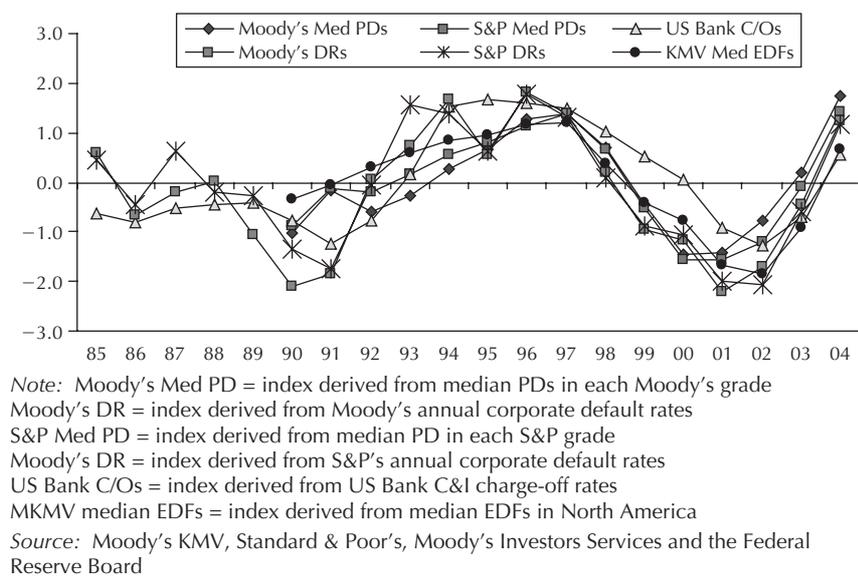
This view departs from the prevailing 1990s credit model, which assumes that systematic credit factors evolve as random walks. Under this legacy model, the credit economy exhibits no predictable cycles, only unpredictable fluctuations, and any cyclical patterns seen historically are accidental, not indicative of future patterns. In this case, looking forward, one does not distinguish between PIT and TTC PDs. In any future year, one expects that the PDs in a broad representative population will be the same as now. The PIT PDs are the only relevant ones for managing risk generally.

While the existence of credit cycles is not by any means proven, we find the intuitive notion of a credit cycle plausible for the following reasons:

❑ Public policy such as the Unemployment Act in the US implies that monetary authorities will act to curtail recessions and, thus, cause default rates to move predictably back towards normal levels following an increase during an economic downturn.
❑ Rates series, including unemployment rates, inflation rates, relative commodity prices, relative currency values and interest rates, are often found to exhibit mean reversion, which is tantamount to a predictable cycle; thus, it wouldn't be surprising to find a cycle in default rates.
❑ Recent research finds evidence that stock price indices exhibit mean reversion, which implies a similar pattern for credit – credit indices themselves display cyclical historical patterns.

On this last point, we observe that the latent credit factors that we have derived from various default and loss series are highly correlated, and they exhibit historical cyclical patterns (see Figure 1).[5] Using notation consistent with our earlier work (see Belkin, Suchower and Forest 1998a, 1998b; Aguais *et al* 2004), we refer to

DESIGNING AND IMPLEMENTING A BASEL II COMPLIANT PIT–TTC RATINGS FRAMEWORK

Figure 1 Normalised, latent credit factors derived from various default, Rating and loss series

Note: Moody's Med PD = index derived from median PDs in each Moody's grade
Moody's DR = index derived from Moody's annual corporate default rates
S&P Med PD = index derived from median PD in each S&P grade
Moody's DR = index derived from S&P's annual corporate default rates
US Bank C/Os = index derived from US Bank C&I charge-off rates
MKMV median EDFs = index derived from median EDFs in North America
Source: Moody's KMV, Standard & Poor's, Moody's Investors Services and the Federal Reserve Board

these as Z credit cycle factors. These Z factors exhibit a zero mean and unit variance. That is, when Zs are positive, credit conditions are "better than historical average" and, therefore, PIT PDs are "lower than their historical averages". When Zs are negative the reverse is true.[6]

In our research we have estimated time series models for forecasting the Z factors describing various sectors. The results indicate that systematic credit factors exhibit statistically significant mean reversion and momentum. By mean reversion, we mean that if a Z factor deviates substantially from a long-run mean, it will tend to revert towards that mean. By momentum, we mean that a Z factor tends to move in the same direction as it has been moving. The combination of mean reversion and momentum produces measurable cyclical patterns.

While, as noted earlier, no one has yet definitively established the existence of predictable credit cycles, we find the evidence substantial enough to pursue models that allow for such phenomena

271

as well as the limiting case of no predictability. In all cases but the limiting one, the PIT–TTC distinction has important meaning.

Distinct PIT and TTC measures help achieve multiple objectives

We now define the terms *PIT* and *TTC* formally. A PIT PD assesses the likelihood of default over a future period, most often the period one year from now and sometimes the next two, three, five or even ten years forward. An accurate PIT PD describes an expectation of the future, starting from the current situation and integrating all relevant cyclical changes and all values of the obligor idiosyncratic effect with appropriate probabilities. Thus, a PIT PD corresponds to the usual meaning of "probability of default" and is, in fact, *unconditional* with respect to unpredictable factors. We attach the "PIT" modifier for clarity in situations that also involve the TTC concept. To estimate default losses accurately, a bank needs PIT measures.

TTC PDs, in contrast, reflect circumstances anticipated over an extremely long period in which effects of the credit cycle would average close to zero. Some analysts report that they approximate this result by using models in determining PDs over the next year under the assumption that credit conditions over that period will correspond to the average observed historically. Others report that they determine TTC PDs and ratings by assuming a particular stress scenario. In any case, such TTC PDs typically assume that credit conditions in broad sectors will differ from those actually expected. Starting from the current situation, such indicators are *conditional* in the two cases above on credit conditions reverting to the historical average or to a particular level of stress.

In practice, TTC PDs may largely represent ordinal measures. Ordinal ratings display the same central tendency regardless of whether overall credit conditions are strong or weak. Ordinal ratings are analogous to those that teachers assign when "grading on a curve". Such ratings provide a ranking within a population at a point in time, but they may prove misleading in comparing across populations or across time. Because of that, one needs a calibration equivalent to standardising test scores. In assessing default risk, PIT PDs embody this calibration.

Analysts sometimes derive TTC PDs by assigning to each company with a particular ordinal rating the long-run, historical average default rate of past companies with that rating. This approach essentially assumes that the future default rate will correspond to the historical average. The resulting PDs will describe the future well, only, if both the ordinal grading process has remained the same over time, and current and expected future economic conditions are close to their historical average.

PIT PDs are essential to the management of credit risk. A bank with only TTC PDs would not be able to quantify accurately its actual risk looking forward. But, for some purposes, banks have found it convenient to have both TTC and PIT PDs.

In deriving the PDs for use in calculating regulatory capital, Basel II calls for the use of a "long-run average of one-year default rates . . ." (see BCBS 2005, Paragraph 447). The FSA's approach in the UK calls for banks to "estimate PDs by obligor grade from long run averages of default rates" (see Financial Services Authority 2006, Paragraph 4.4.24). Some interpret these passages as requiring banks to use TTC PDs. Actually, the words lend themselves to different interpretations, since the estimation of both PIT and TTC PDs involves averaging over many years of data.

In estimating a PIT PD model, for example, one typically calibrates to data on many obligors across many years. One pools across time, as well as obligors, as a way of increasing the sample size and reducing sampling error. Even with such pooling that involves averaging over time, the model will be PIT if it includes explanatory variables that track and therefore control fully for the state of the credit cycle. Such a model would explain the temporal fluctuations in default rates as arising from broad shifts in the ratings and PDs of many obligors. In other words, cyclical changes would show up as large-scale ratings migrations.
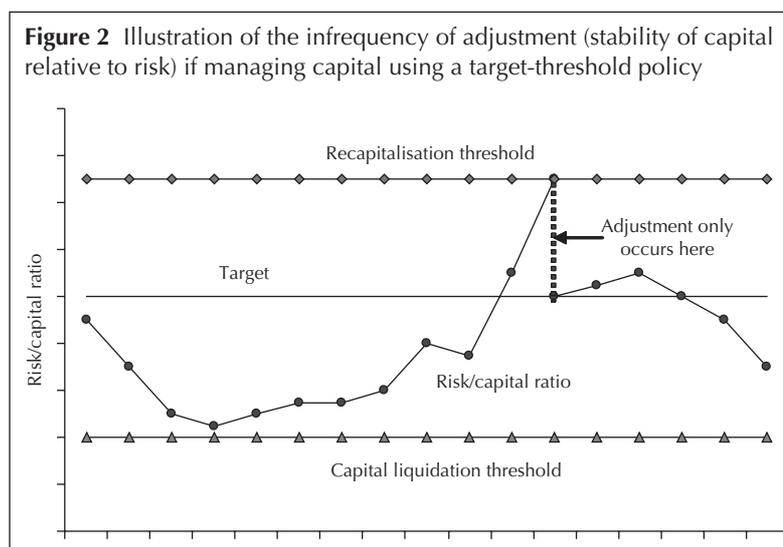
In contrast, one could also calibrate a TTC PD model to the same default data. But the model would account for the temporal fluctuations differently. Due to the assumed constancy of aggregate credit conditions, ratings from the model would display close to a fixed distribution over time. Thus, migrations would explain little, if any, of the broad-based cyclical changes. Instead, such fluctuations would appear, *ex post*, as wide variations in the realised

default rates of each of the ratings. Thus, "averaging over many years" doesn't distinguish a TTC PD from a PIT PD, but a more stable ratings distribution (and less stable default rates for each rating) does.

Still, many view Basel II as calling for the use of TTC PDs, since this implies stable estimates of capital requirements. Stability is seen as a desirable attribute of a strategic capital reserve and some regulators have expressed concerns that if banks were to use PIT measures, they might overload their balance sheets during peak periods prior to an economic downturn. (For a related discussion, see Heitfield 2004).

Not surprisingly, bankers believe that wide fluctuations in capital involving large liquidations followed by substantial recapitalisations (or alternatively, large cycles in the amount of risk taking) would prove highly inefficient and most likely impossible to manage. Of course, one could apply optimal inventory methods in explaining the advantages of a stable capital reserve in the presence of substantial fluctuations in portfolio risk (see Figure 2).

Figure 2 Illustration of the infrequency of adjustment (stability of capital relative to risk) if managing capital using a target-threshold policy

Under such an inventory-theoretic approach, a bank would gauge its credit portfolio risk properly using PIT measures such as the 99.9th percentile, next year value-at-risk, and the cyclical fluctuations in those risk measures would change the ratio of risk to capital. Over time, however, a bank would not act to control that ratio unless it moved outside of an acceptable range. At that point, the bank would bring the ratio back to a desired long-run target and it could accomplish this by adjusting its amount of capital or risk. However, if such adjustments entail large costs, the range of acceptable values of risk per unit of capital would be large and so adjustments would occur rarely. The use of TTC PDs seems an expedient way of achieving the same objective.

However, in using the TTC solution, one must keep in mind that the resulting calculations do not measure actual risk. One must never lose sight of the actual risks, and, to paraphrase some remarks by Gordy (2003), "one may stabilise the outputs, but not the inputs". We see that the optimal-inventory approach above complies with this advice – the risk measure (the input) is PIT and volatile, but nonetheless the bank might find it optimal to maintain capital (the output) stable.

Banks may also find the TTC PD measures useful in guiding some other risk management activities that rely on estimates of long-run aggregate risks. Some banks, for example, use TTC PDs in determining discretions, which are rules on the level of authority required of people approving different amounts of exposure in different ratings categories.

Note, however, that while PIT PDs fluctuate much more than TTC PDs, when averaged across a large portfolio, they need not be much more volatile for the individual obligor. Most PD volatility is specific to the obligor (idiosyncratic), and this volatility will remain in a good TTC PD indicator.

Nonetheless, we observe in practice that individual obligor PDs that arise from ratings or other information viewed as TTC often are more stable than those that arise from ratings or information viewed as PIT. This seems mainly a result of the differing frequencies of the underlying model inputs. Many TTC indicators are derived from low frequency information such as annual financial reports. Many PIT ones involve much higher frequency information such as stock prices. The higher frequency of the information

contributes to the greater volatility of PIT PDs. But the stability in the TTC indicators is artificial, caused by much of the information on the current situation being hidden until the belated release of an annual statement.

Despite the primary emphasis that we put on PIT PDs, we believe that Basel II compliant banks need an integrated PIT/TTC risk-rating and PD approach. We now turn to the design of this dual PIT–TTC credit rating system. This overall approach encompasses the individual PD models of various types and the overall PIT–TTC framework that is used to make various PD models consistently comparable. Here, we emphasise the key role played by the various PD models as we see credit ratings arising from PDs – not the reverse. The framework provides one-year PIT and TTC PDs for each borrower and assesses credit factor adjusted, forward PD term structures.

DESIGNING A PIT–TTC DEFAULT RATING SYSTEM
Overview: key components of the rating system

To design a credit rating system that satisfies the requirements of both Basel II and a firm's own credit risk management objectives, there are a number of key steps that need to be followed in developing the integrated components of the approach. Generally, banks need to:

❑ develop Basel II compliant PD models covering all obligor types or accounts and assess each model's degree of "PIT-ness";[7]
❑ define the overall rating master scale (ie, PD bins) as having a number and spacing of ratings satisfying both regulatory and management reporting requirements;
❑ formulate conversions of all PD models from their PIT–TTC starting point into "pure" PIT and TTC PDs; and
❑ calculate accurate PD term structures.

These overall requirements reflect a solution that supports both advanced risk management and advanced IRB Basel II objectives. We discuss each component in turn below and then explain how the overall framework comes together.

We see the PD models and ratings being developed and used in accordance with an overall credit policy and protocol (see Panel 1).

PANEL 1: ILLUSTRATIVE RATING PROTOCOL
Rating protocol
Rating occurs entirely through the use of approved PD models developed largely by the credit risk management function. Each PD model is clearly designated as producing a PIT, TTC or hybrid one-year PD output. Such models have clearly defined:

❑ inputs, including adjustments to vendor supplied models;
❑ formulas, embodying calibrations; and
❑ overrides to outputs.

Credit officers (COs) use these models in determining PDs and ratings. In each application of a model, either the CO enters particular values for the inputs manually or these are supplied by external third-party vendors. Inputs may also be subject to validation rules. The formulas then uniquely determine the one-year PD (or multi-year PD term structure). Thus, only one PD (or term structure) is possible for any choice of inputs. The CO may, at that point, override the model-determined PD, given senior management approval. In the last step, the rating master scale translates the PD into a rating.

To substantiate any override rationale, the CO may draw on secondary approved models, such as ones that convert between TTC and PIT PDs and ones that forecast aggregate credit conditions in selected segments defined by asset classes, regions and industries. Adjustments and overrides are subject to regular monitoring and senior management review.

In some asset classes, particularly within retail, the exposures to individual obligors or accounts are small but extremely numerous, and so the determination of PDs and ratings involves automated methods. COs intervene in exceptional cases and in reviewing pools of exposures.

The approval of a model prior to their use in making risk and business decisions includes a formal process: preparation of descriptive documents; conduct of an external review; and presentation of the documents and review to a technical committee for official sign-off. Each model development document follows a standard format. Among other things, the document must present the conceptual and statistical evidence indicating that the model would predict accurately and better than available alternative approaches.

In rating, COs must use an approved PD model, if available. If more than one is applicable, a model-hierarchy policy determines the one to choose. Models usually remain unchanged over annual intervals. Most credit models are on an annual cycle of review, revision, and reapproval.

Developing Basel II valid PD models:

PD models form the foundation of any rating system. If a bank has a credible PD model for each of its obligors, deriving the rating becomes a simple process of classifying obligors into bins (ratings) defined by PD ranges.

We see mostly four types of PD models, namely:

❑ single obligor statistical ones, in which one obtains a large representative sample of company (or account) default and no default outcomes and fits a model based on earlier values of company (or account) credit indicators, that offer the best explanation of the observed outcomes;

❑ approaches based on agency ratings, in which one translates each agency rating to the PD that it currently implies;

❑ scorecard (expert-system) models in which one starts with often subjective, ordinal measures of an obligor's creditworthiness and applies a conventional, low default portfolio (LDP) algorithm in establishing a calibration based on a small sample of default and no default observations (see Pluto and Tasche 2005); and

❑ derivative credit risk models in which one typically uses simulation or stress methods in evaluating the likelihood of default and loss on a structured position affected by the performance of an underlying asset pool involving many obligors.

We could write extensively on the proper design and calibration of each of these PD model types. However, we provide only a few comments here, focusing on validation.

A bank may acquire models externally from vendors or develop its own, using a combination of internal or external data. In any case, the standards for validity remain the same. In our view, a valid model must be both plausible conceptually and reliable statistically. A plausible PD model will derive from a logical theory of the default process. For larger companies, this usually means a Merton-type approach, which might draw on variables beyond those in the original Merton formulation. A reliable model's parameters will:

❑ have correct arithmetic signs $(+/-)$; and
❑ differ from zero at conventional confidence levels (eg, 95%).

Additionally, the PD model should outperform alternative, plausible models in terms of goodness-of-fit. We find that PD model

developers sometimes pay little attention to the plausibility of their formulations. Whenever possible, we avoid such models. Experience from a wide range of scientific fields indicates that empirical approaches that search arbitrarily for close statistical fits to a particular data sample produce models that may perform poorly in practice (see, eg, Jefferys and Berger 1991; Findley 1993; Clark 2000).

In their application, PD models often need to work under conditions that diverge from those prevalent in a historical sample. Highly calibrated, purely empirical approaches can lock onto relationships that hold only under a narrow set of conditions. With a conceptually plausible model, one has a better chance of controlling for changes in the fundamental factors.

In evaluating a PD model's goodness-of-fit, analysts commonly examine rank-ordering performance. While this represents an important aspect of performance, it does not tell the whole story. A rating system could rank-order perfectly, yet, due to a poor calibration, cause a bank to fail.

We evaluate goodness-of-fit primarily using likelihood measures, which when formulated correctly, include complex adjustments for correlation. Likelihood measures consider both rank-ordering and calibration accuracy together as discussed in our previous *Basel Handbook* (Chapter 7).

By "calibration accuracy" we refer not only to the average PD for an entire population but also to averages for various subpopulations of higher and lower risk and to averages over different time periods of broadly higher and lower risk. Indeed, estimation of the long-run PD population average is comparatively straightforward. The difficulty compounds when one turns to determining the amount of PD change implied by various moves up and down an ordinal default risk ranking as well as changes in the population average over time. Properly formulated likelihood measures combine all of these aspects of model performance. The measures impose penalties if a model:

❑ ranks obligors poorly;
❑ fails to differentiate clearly periods of broadly higher and lower risk; or
❑ under- or overstates the long-run population average PD.

279

In managing the development of PD models in cases in which limited data preclude one from satisfying the usual reliability standards, banks will need to set conventions for determining calibrations with less precision. Adapting the Basel II language, we refer to such procedures as *low default portfolio* (LDP) calibrations. We've experimented so far with two LDP approaches that involve:

❑ accepting parametric calibrations if the parameter estimates have intuitively correct, arithmetic signs and plausible magnitudes, even if those estimates have less precision than, for example, the usually required 95% confidence of being different from 0; or
❑ applying the non-parametric, Pluto–Tasche algorithm, which provides a calibration even in the most difficult case of no observed defaults and entirely subjective, scorecard-based, risk differentiation.

In each case, we require that, in backtests, the model indicates that the default count would be equal to or greater than that actually observed with a probability equal to a threshold value of over 50% (eg, 60%). This condition usually implies that we shift the "best-fit" calibration upwards – to broadly higher PDs. We justify this adjustment by noting that, with only a small number of defaults detected, a default undercount seems more likely than an overcount. Today's calibrations often rely on default counts obtained by searching historical records not designed for distinguishing defaulters from non-defaulters. Considering the reluctance for defaulters or their creditors to advertise such failures, one realises that such searches could miss some defaulters and misclassify them as non-defaulters.

Defining a default rating master scale

A default rating master scale at each selected horizon corresponds to a set of disjoint PD bins covering the entire range from 0 to 100%. In establishing a master scale (set of PD bins), one would likely consider such objectives as:

❑ satisfying business and regulatory needs for risk differentiation by ratings;
❑ reconciling with external sources of credit information;
❑ providing similar differentiation throughout the risk spectrum; and
❑ aligning with legacy rating systems.

Today's large corporate bond and loan markets typically use Moody's, S&P and Fitch alphanumeric ratings for sorting obligors into different risk classes. Further, Basel II often endorses the use of agency ratings as a basis for determining regulatory capital. Thus, to achieve the first two objectives above, one would likely need a count and spacing of ratings similar to or more refined than the alphanumeric scales of the ratings agencies. This would probably entail 15 or more PD bins spaced like the agency, alphanumeric ratings. If one wants more ratings, one would likely structure them to roll up to the agency ratings.

We interpret the third objective as implying PD bins with close to the same widths as measured by default distance (DD). Defined as in the CreditMetrics (CM) model, DD has no drift and an annual volatility of one (see Gupton, Finger Bhatia 1987). Thus, with ratings bins having a constant width as measured by CM DD, one would expect that companies in the different rating bins would exhibit similar one-year probabilities of migrating one-rating, two-ratings, and so on. This uniformity in credit migration simplifies the development of ratings-based credit models with limited numbers of parameters. While one might desire greater differentiation in the higher risk ratings (since as the PD rises, equal DD width translates into wider PD width), the increasing PD volatility in those ratings usually precludes this.

Sometimes historical information includes only old credit ratings and not the underlying PDs or other information that provides for a more granular differentiation. Thus, to make it easy to compare or consolidate new and old information, an institution may want any new ratings to reconcile with older, less granular master scales. By this, we mean that the new rating bins or combinations of the new bins should map one-to-one to the older ones.

Default rating master scale – an illustration
We now illustrate the design of a hypothetical default rating master scale that:

❑ aligns closely with the 19 agency alphanumeric ratings;
❑ provides close to the same risk discrimination (measured by DD) across the entire risk spectrum; and

Table 1 Mapping of old to new rating scale and to S&P ratings

Old ratings	S&P*	DD bins			PD bins (%)			New ratings
		Max	**Mid**	**Min**	**Min**	**Mid**	**Max**	
1	AAA	∞	3.72	3.58	0.000	0.010	0.017	1
2	AA	3.58	3.50	3.43	0.017	0.023	0.031	2
3	A+/A	3.43	3.35	3.28	0.031	0.040	0.053	3
	A/A−	3.28	3.20	3.13	0.053	0.069	0.089	4
4	BBB+	3.13	3.05	2.98	0.089	0.114	0.146	5
	BBB	2.98	2.90	2.83	0.146	0.186	0.236	6
		2.83	2.75	2.68	0.236	0.298	0.373	7
	BBB−	2.68	2.60	2.53	0.373	0.466	0.578	8
5	BB+	2.53	2.45	2.38	0.578	0.714	0.877	9
	BB	2.38	2.30	2.23	0.877	1.071	1.303	10
		2.23	2.15	2.08	1.303	1.576	1.898	11
	BB	2.08	2.00	1.93	1.898	2.273	2.709	12
6	B+	1.93	1.85	1.78	2.709	3.213	3.792	13
		1.78	1.70	1.63	3.792	4.453	5.204	14
	B	1.63	1.55	1.48	5.204	6.053	7.006	15
		1.48	1.40	1.33	7.006	8.070	9.252	16
	B−	1.33	1.25	1.18	9.252	10.558	11.992	17
7	CCC+	1.18	1.03	0.88	11.992	15.259	19.069	18
	CCC	0.88	0.73	0.58	19.069	23.411	28.252	19
	CCC−	0.58	−1.84	−4.27	28.252	37.070	99.999	20
D	D	−4.27	−∞	−∞	100.00	100.00	100.00	D

*Average historical relationship.
This example of a master scale illustrates the reasoning and trade-offs involved in designing a rating system.

❑ offers a simple (non-dynamic) mapping to a legacy rating system assumed to coincide with the seven agency alpha ratings for simple illustration as in the PIT/TTC approach proposed here, we use a dynamic agency approach as described in more detail below.

We use DD width as the main measure for assuring similar risk differentiation across the risk spectrum (see Table 1). We search for a DD bin width small enough so that integer-number multiples of it line up closely with the legacy ratings. In this example, we find that a width of 0.15 works most of the time. However, we observe that this produces sparsely populated ratings at the extremely high-risk end (CCC range). Thus, in that range, we use bins with a DD width of 0.30.
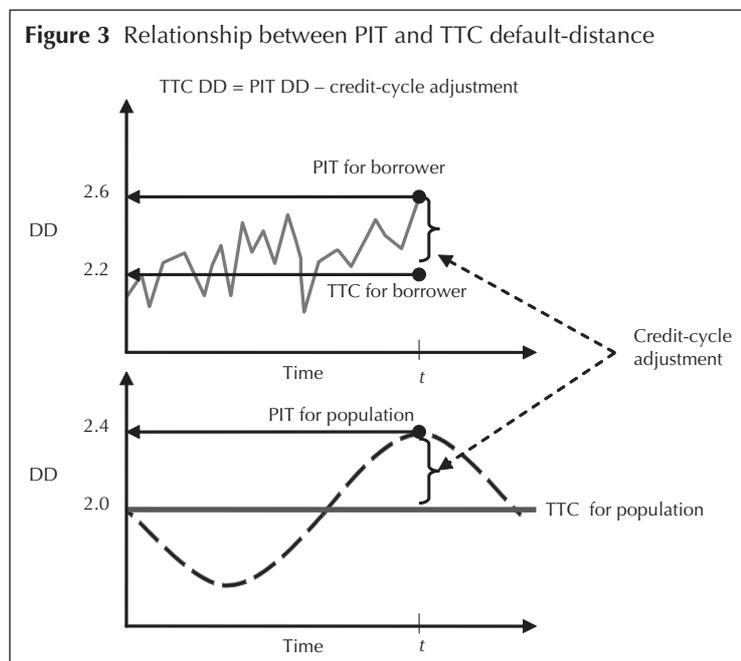
Formulating PIT–TTC conversions

This subsection describes an approach for converting PIT PDs to TTC, TTC PDs to PIT, or even hybrid cases (PDs intermediate to PIT or TTC) to pure PIT and TTC PDs. As discussed earlier, we define a PIT PD as the unconditional probability of an obligor's defaulting in a future period, often the next year. As explained in more detail above, these estimates reflect all information relevant to the determination of that probability, including indicators of the current state and outlook for the credit cycle.

We define a TTC PD as the expected PD, assuming that credit conditions in the obligor's sector (eg, region/industry/asset class grouping) return to and stay in a cyclically neutral, normal state. By "normal", we mean "implying long-run average default rates". Since aggregate default rates exhibit skewed distributions, default rates in a normal year will exceed the default rate in an average year.

In formulating the PIT/TTC conversions, we work with DD measures. In our experience, every PD model draws on risk indicators that imply a spot measure of DD. For large corporations, the best models use indicators that, in combination, form a measure closely related to the Merton DD concept. In other cases, the relationship is less close, but we still find a measure, intrinsic to the PD model, that corresponds to spot DD. For example, one can always derive a "synthetic" DD by applying the inverse normal function to the shortest term PD yielded by the model. The degree of abstraction rises as one deals with asset classes increasingly distant from large corporate borrowers, but, as indicated by its pervasive use in the Basel II Accord, the concept still works.

In some cases, the spot DD measure reflects current economic conditions and we regard it as point-in-time (DD^{PIT}). In other cases, the measure seems basically ordinal or it assumes that current conditions are at a historical norm. We regard such measures as through-the-cycle (DD^{TTC}). Finally, the DD measure may partly account for current conditions and we view it therefore as an intermediate case between "pure" PIT and TTC.

We illustrate the full conversion from PIT to TTC in Figure 3. We subtract the current cyclical component from DD^{PIT} and obtain DD^{TTC}. The resulting DD^{TTC} and the related TTC PD will still move over time as a result of obligor idiosyncratic (but not systematic) factors. Alternatively, if one starts with a DD^{TTC}, one gets DD^{PIT} by

Figure 3 Relationship between PIT and TTC default-distance

TTC DD = PIT DD – credit-cycle adjustment

adding the cyclical component. The resulting DD^{PIT} and PIT PD will move over time as a result of both idiosyncratic and systematic factors.

The average of PIT PDs across a broad portfolio will have considerably higher volatility than the average of TTC PDs. For individual obligors, however, the idiosyncratic component of credit risk typically dominates and so TTC PDs will exhibit nearly as much volatility as the PIT ones.

Assuming the existence of either DD^{PIT} or DD^{TTC} or a known intermediate measure for an obligor or account, the derivation of complementary PIT and TTC measures involves:

❑ identifying the current state of credit conditions in the obligor's economic sector (or sectors) relative to the historical normal state (s); and
❑ using that information in adjusting the DD indicator to PIT or TTC or both.

We outline these two steps below.

Deriving credit indices for economic sectors

Sector credit indices measure average credit conditions for a group of obligors or accounts that one identifies as representing an important component of systematic risk. In developing these indices, one typically partitions risk by asset class, region and, on the wholesale credit side, industry. To construct such indices, one may:

❑ collect time series data on DD^{PIT} for the largest, possible, representative sample of obligors or accounts within the sector; and
❑ form a time series, summary measure (eg, mean or median) of the individual obligor or account DD^{PIT} measures.

This is a standard approach for deriving latent risk factors. One obtains performance measures for a large, representative sample of obligors or accounts and summarises them.

One might scale the summary results to reconcile with the CreditMetrics model's assumption that annual changes in systematic risk factors have unit variance. One achieves this scaling by introducing a correlation parameter, $\rho(s)$, into the index construction, as follows:

$$Z(s,t) = \frac{DD^{PIT}(s,t)}{\sqrt{\rho(s)}}$$

$$DD^{PIT}(s,t) = \underset{f \in S(s,t)}{SUMM}(DD^{PIT}(f,t)) \qquad \textbf{(1)}$$

$$\rho(s) = \underset{t}{Var}(DD(s,t) - DD(s,t-1))$$

$Z(s,t)$ = unit – variance, credit index for sector s
$\rho(s)$ = correlation or scaling factor for sector s
$DD^{PIT}(s,t)$ = summary PIT DD for sector s
$SUMM$ = summarisation operator (eg, median, mean, weighted mean)
$S(s,t)$ = set of indices of obligors in sector s at time point t
f = obligor index
$DD^{PIT}(f,t)$ = PIT DD for obligor f at time point t
$\underset{t}{Var}$ = variance computed across all history.

One needs a "normal" Z value for a sector ($Z_n(s)$) to be used in estimating the TTC PD (PD^{TTC}) for each related obligor. One might consider using the past average value of $Z(s, t)$. But, in most PD

models, the relationship between PD and DD (and therefore Z) is non-linear and convex. Thus, this choice of $Z_n(s)$, when applied in determining a PD^{TTC} for each of the past sample of obligors, typically implies an average PD^{TTC} that falls slightly below the historical average PD^{PIT} and the historical average, realised, default rate (DR = default count/total cases).

One therefore might instead define $Z_n(s)$ as the value usually slightly below the $Z(s, t)$ average that, when applied in calculating PD^{TTC} for each of the historical obligor samples, produces an average that reconciles closely with the past average PD^{PIT}. This seems compatible with the Basel II focus on reconciling with long-run average default rates. One might formulate this approach as follows:

$$Z_n(s) = \frac{F^{-1}(PD(s))}{\sqrt{\rho(s)}} \tag{2}$$

F^{-1} = inverse of the PD function for the sector
$PD(s)$ = long – run, past average default rate for sector s.

While these formulas may seem abstract, in practice they usually become clear. Assume, for example, that one's primary credit indicators are one-year PDs computed under the assumption that credit conditions evolve as Brownian motion processes. In that case, one derives a DD^{PIT} by applying the inverse normal function to a PIT PD. Then, assuming that the sector has never experienced a structural shift to permanently higher or lower PDs, one obtains the sector's normal DD by applying the inverse normal function to the long-run, historical average PD or DR for the sector. In other words, in formula (2) above, F^{-1} corresponds to the inverse normal function. Using that approach applied to MKMV EDFs, which incorporate the Brownian motion assumption, one can derive credit indices for listed companies grouped by region or industry (see Figures 4 and 5). The correlation parameters used to derive these Z index examples are derived as shown in equation (1) above.

In addition to the Z factors measuring overall credit conditions in sectors (regions and industries), we also construct another index, which we use in deriving PIT PDs for agency ratings. We call this index the "agency Z".
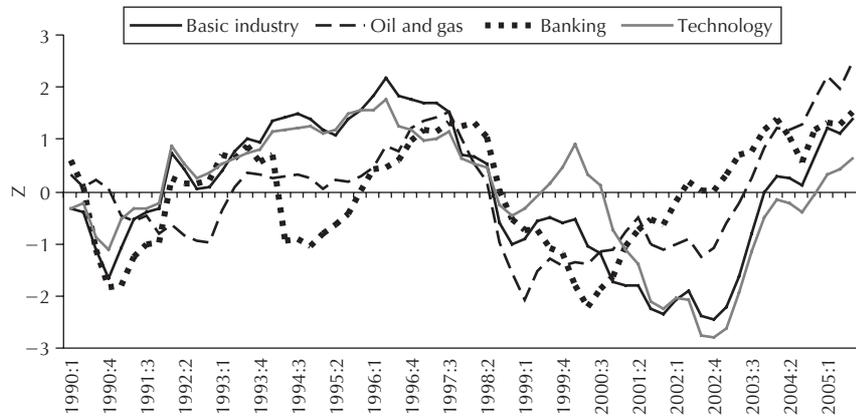
Figure 4 Normalised Z indices for regional groupings as derived from MKMV EDFs
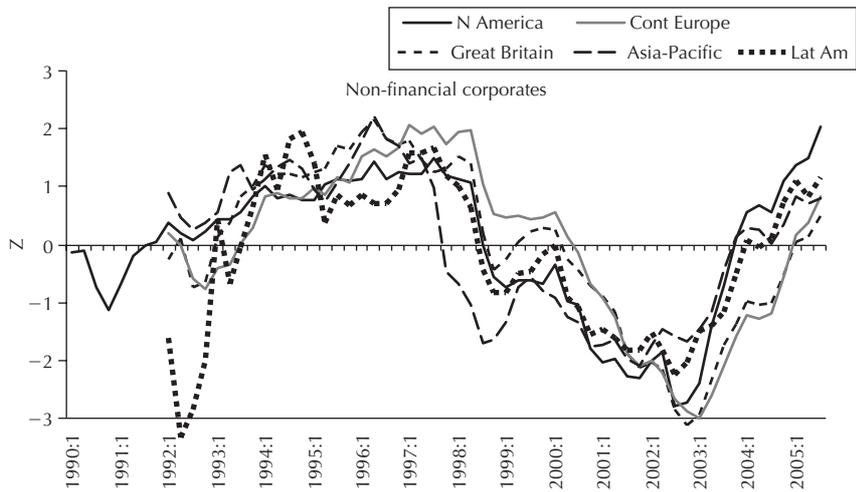
Source: Moody's KMV

Figure 5 Normalised Z indices for industry groupings as derived from MKMV EDFs

Source: Moody's KMV

Agency ratings are used widely in gauging the creditworthiness of wholesale counterparties. Many view these ratings as TTC indicators, since the agencies describe their rating approach as "looking across the cycle". While we don't find agency ratings to

be pure TTC measures, they move much less than pure PIT ones as overall credit conditions change. Thus, to get accurate PIT PDs from agency ratings, one needs a modelling or mapping procedure that adds more movement. The agency Z is the index that measures the extent to which each agency rating's current one-year PD stands higher or lower than its idealised, long-run average, one-year PD.

In our previous *Basel Handbook* chapter, we described this so called dynamic mapping approach for translating agency ratings to one-year PIT PDs. We've also tested the performance of these dynamically mapped agency PDs in comparison with MKMV EDFs on both a one- and five-year basis.

We currently derive a single agency Z that we apply across all agency ratings for a given agency in estimating the current one-year PIT PDs. While one can imagine distinct agency Zs for different ratings or other different categories of obligors, we currently find that sampling variations make more detailed adjustments unfeasible. We define the agency Z as below:

$$Z_{\text{Agency}}(t) = A \underset{r}{VG} \left(\frac{\Phi^{-1}(EDF(r,t)) - \Phi^{-1}(EDF(r)) + \Phi^{-1}(PD(r))}{\sqrt{\rho}} \right)$$

$$(3)$$

AVG = average across all alphabetic agency ratings
Φ^{-1} = inverse standard – normal distribution function
$EDF(r, t)$ = median MKMV EDF for obligors with rating r at time t
$EDF(r)$ = historical average of median EDFs for the rating r
$PD(r)$ = idealised, agency, historical average default rate for rating r.

We apply this agency Z in translating each rating's long-run average PD to an estimate of its current PIT PD.

Be clear that this index is different from the sector Zs. Sector Zs measure general credit conditions in, for example, a geographic region or global industry. The agency Z measures the average creditworthiness of companies within each agency rating, relative to its respective, long-run, historical average. Thus, if migrations in agency ratings were to track closely the overall credit cycle, agency Zs would remain nearly constant. We find, however, that agency Zs fluctuate widely (see Figure 6). This indicates that agency ratings

Figure 6 Normalised agency *Z* factor derived from S&P-rated firms

Source: Moody's KMV and Standard and Poor's
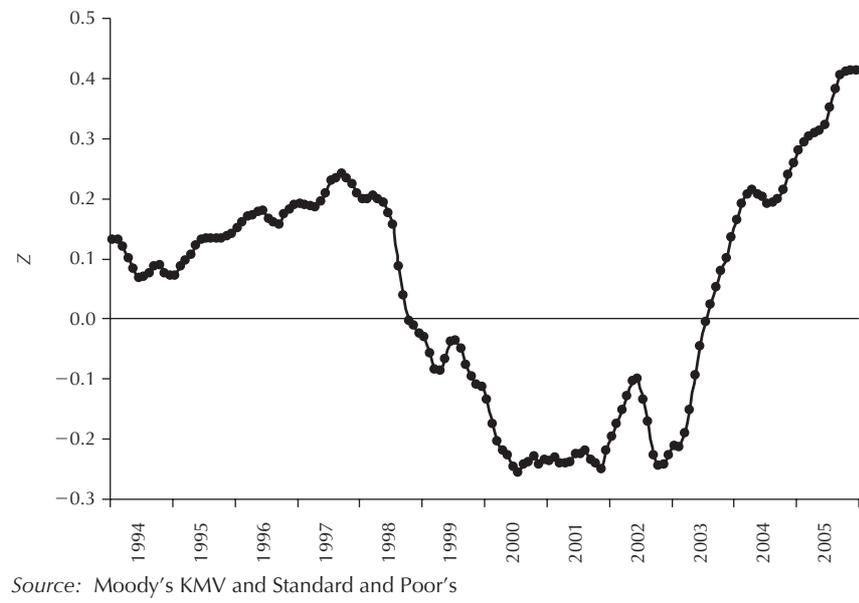
migrations explain a minority share of temporal changes in credit conditions, with a much larger proportion picked up by changes in the average creditworthiness of obligors within each ratings grade.

Using sector credit indices for making PIT–TTC conversions

Given a comprehensive set of sector credit indices obtained as explained above and beta (β) coefficients measuring the loading of each index on each obligor or account and factors, $\delta(f)$ indicating the degree from 0 to 100% that an obligor's or account's DD measure is PIT, we may convert between PIT and TTC PDs using the formula below:

$$DD^{PIT}(i, t) = DD(i, t) + (1 - \delta(f))\sum_{j} \beta(i, s)(Z(s, t) - Z_n(s))$$

$$DD^{TTC}(i, t) = DD(i, t) - \delta(f)\sum_{j} \beta(i, s)(Z(s, t) - Z_n(s))$$

(4)

We indicate in (4) above that the DDs that derive from an existing PD model may fall somewhere between the extremes of 100% PIT

or 100% TTC. In practice, working with legacy PD models, the calibration of "PIT-ness" as measured by the δ factor may largely involve judgement informed by scattered empirical results on other, comparable models. For example, we've judged the extent to which some financial ratio based models are PIT by comparing the varying cyclical responsiveness of other financial ratio based models with and without the MKMV based DD gap as an explanatory variable. Additionally, if a model purports to mimic an agency rating approach, we might assign it the 30% PIT weight consistent with the historical performance of agency ratings.

Note, however, formula (4) offers a recipe for translating a legacy PD model that isn't 100% PIT into a 100% PIT one. One need only re-estimate the model with the relevant credit indices obtained from another PIT PD model (such as MKMV's) included as additional explanatory variables. Since 100% PIT models provide the best predictors of default risk, we expect that most PD models will in time become 100% PIT. In this case, one would need only the second of the two formulas in (4) and the parameter δ would always be one.

Consider now estimation of the betas. If one starts with PIT DDs, one may determine the betas by regressing a relevant sample of obligor or account DDs on the corresponding credit indices. If one starts other than with PIT DDs, one might initially use the betas obtained by regression for an otherwise comparable, PIT model. Or, better yet, one would include the credit indices in re-estimating the PD model and obtain by means of that estimation both a 100% PIT model and the betas for translating PIT DDs to TTC ones.

The task of implementing this approach across an entire bank to develop consistent PIT–TTC PDs is formidable. One must develop a large set of credit indices and determine credible estimates of beta coefficients and indicators of the degree to which existing DD indicators are PIT or TTC.

Developing PD term structures

To complete this section's discussion of the design of the overall framework, we briefly describe an approach for estimating obligor-level PD term structures reflecting the anticipated credit outlook and the risks in that outlook. The approach involves:

❑ estimating second-order (mean reversion, momentum) time series models of the stochastic evolution of the sector Zs and then applying those models in deriving Monte Carlo simulations of future Z paths;

❑ translating those sector Z simulations into PD simulations over one-year and longer horizons by entering the simulated Z changes (z) over a year as conditioning factors in a model of annual DD^{PIT} transitions;

❑ multiplying the simulated sequences of Z conditional transition matrixes and, from the default columns of those multiplied matrixes, obtaining sector PD scenarios over various horizons for the different DD^{PIT} bins; and

❑ averaging the PD scenarios over each horizon and thereby obtaining PD term structures for obligors in each sector within each DD^{PIT} bin.

Long-run, average transition matrixes provide the calibration of the default and DD transition barriers inherent to the transition model. We combine agency default and MKMV EDF transition data in deriving the barrier calibrations. We view the non-default agency ratings transitions as underestimates of DD^{PIT} transitions. We use the MKMV EDF transition data for getting better estimates of those non-default transition rates.

Banks often use such long-run average transition matrices to project PDs over various horizons. This conventional approach assumes that the credit outlook never varies and that any deviations observed after the fact reflect entirely random events. Here, we generalise this approach by allowing that the outlook could vary at least a little in predictable ways as indicated by the various Z models. In particular, those models generally anticipate mean reversion in credit conditions. Thus, if conditions within a sector are far below (above) average, the associated model would on balance project a recovery (deterioration). However, these same models foresee a wide range of possible scenarios around these central tendencies.

As noted, the sector PD simulations provide term structures for a selection of DD^{PIT} bins. To avoid a loss of resolution, we treat those term structures as applicable to the numerical mid-point DD^{PIT} in that bin. We then use interpolation in deriving PD term

Figure 7 Region/sector five-year term structures
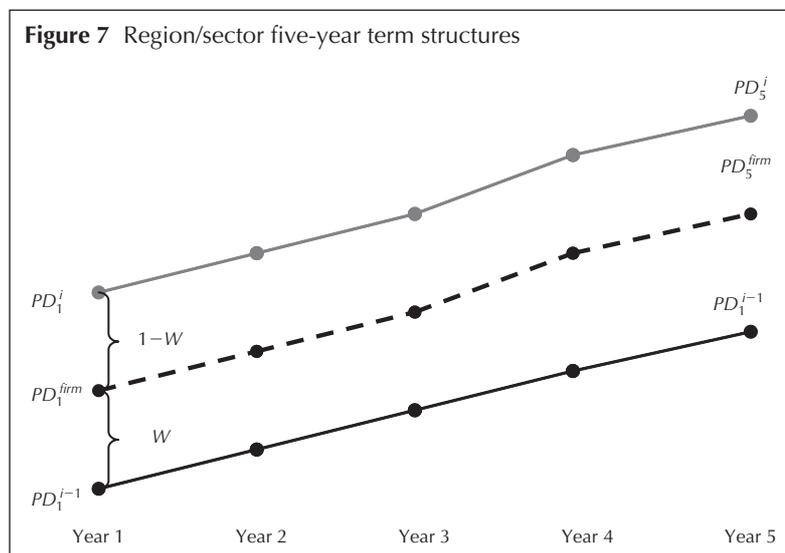
structures for each obligor with a DD^{PIT} value intermediate to a pair of mid-points (see Figure 7).

IMPLEMENTING PIT–TTC DEFAULT RATING SYSTEM

Having described the components and design approach of an integrated PIT–TTC PD approach, we now turn to a brief discussion of some aspects of a successful implementation.

As we have highlighted in both this chapter and our previous work, to be successful in both managing credit risk and satisfying Basel II, banks require a rating and PD approach that provides two distinct views of PDs to support the multiple objectives banks must satisfy. Once the framework is understood, the real test comes about during implementation when an organisation needs in a Kuhnian sense to change its risk rating paradigm in a substantial way.[9] By this we mean that an organisation needs a major shift in its overall ratings perspective and its language of ratings and it must apply one consistent overall framework. This is about substantial change, not improvements around the edges.

Figure 8 Monthly *Z* credit-factor estimation process
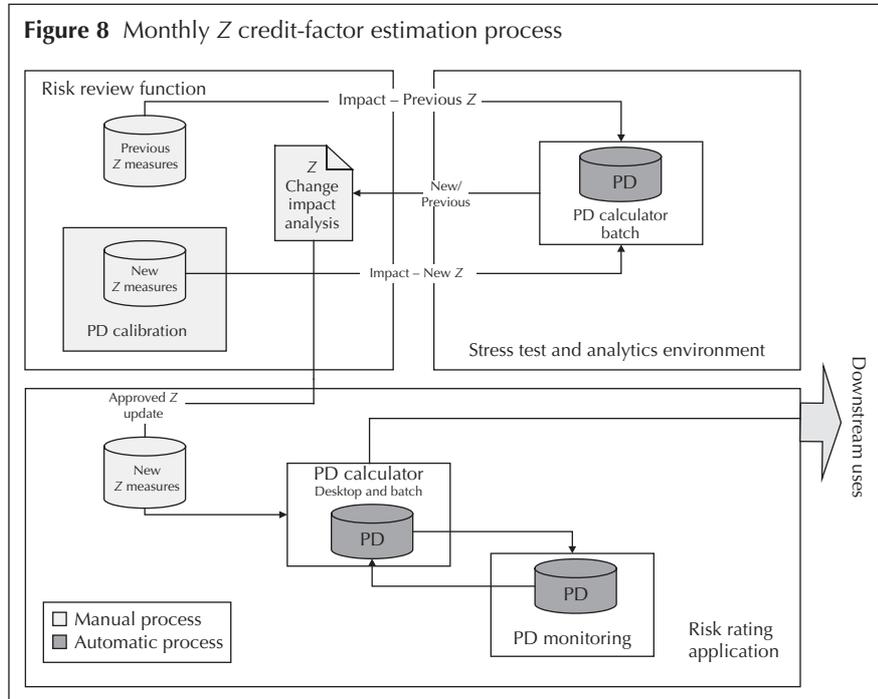
Implementation

We view banks with a model history in assessing economic capital consumption as those that will find the transition to using two PDs easier. Only by using models to manage credit risk over several years will it become clear that PIT and TTC models behave very differently, and no single PIT and TTC rating indicator provides either the breadth of portfolio coverage or the required level of accuracy.

The Basel II mandatory use of "all relevant and available information" in particular helps to clarify the path to implementing two PDs per obligor. Without a ratings framework that allows the consistent comparison of credit risk indicators on both a PIT and TTC basis – a credit officer is essentially comparing apples with oranges in attempting to derive an accurate measure of the client PD.

Unique client identification, the management of client hierarchies and supporting reference data are critical to the implementation of a PIT/TTC ratings framework. These steps are required to successfully link desktop, batch and monitoring applications so

that PDs for many thousands of clients can change automatically on a frequent basis without manual intervention.

Calculating PIT and TTC PDs across the client portfolio starts with the estimation of various region, sector and agency Z factors. This is represented in Figure 8 as the manual PD calibration batch process conducted within the risk review function.

New Z values are then applied to a static portfolio to assess the impact on client PIT and TTC PD values. Results are forwarded to senior management for approval, and desktop and batch applications are tested with new Z parameters prior to implementation within the live production environment.

Finally, changes in PIT and PD values are subject to regular monitoring to ensure any continuation of overrides or significant movements are reviewed and approved by appropriate personnel in line with internal policy guidelines.

SUMMARY

In this follow-up chapter to our first *Basel Handbook* contribution, we have extended the discussion of PIT/TTC concepts and related issues to include the specification and design of an integrated PIT/TTC PD approach. Discussed in the context of wholesale credit risk specifically, this approach can be adapted more generally across all of the various obligors and portfolio types within a large, internationally active bank, albeit with differences in data and application for the retail, SME and wholesale worlds.

Our conclusions remain the same – any bank looking to satisfy multiple objectives across both internal credit risk management and Basel II requires a consistent multi-PD solution. The main evidence motivating this need for both PIT and TTC PDs rests on the empirical existence of statistically measurable credit cycles. Not only do latent credit factors derived from various default and loss series show high correlation over the last 20 years, additional research shows that the phenomena of both mean reversion and momentum are statistically observable in forecast models for these factors.

Once it is understood that credit cycles exist at some measurable level, we observe that PIT and TTC PDs are not differentiated simply by a random walk process as legacy credit factor models assume. Therefore, observable systematic behaviour leads us directly to the key conclusion that PIT and TTC PDs are separated

by a measurable difference – statistical credit cycles. Ultimately, the integrated PIT–TTC system described here is motivated by both the need to support multiple objectives and the existence of measurable differences in PIT and TTC PDs.

The rest of the discussion has provided what we believe are the foundations of an advanced approach, which includes appropriate validation of PD models, a well-designed master scale, a focus on PDs but not ratings, the overall conversion apparatus using credit factors to ensure PIT–TTC consistency, and, finally, an approach for extending PD term structures beyond one-year to incorporate the cyclical nature of credit risk. In the end, the world is always evolving, and what we presented in 2003 as a discussion and overall taxonomy for thinking about PIT–TTC issues has now evolved into a fully integrated PIT–TTC approach that has actually been implemented on a global basis.

1 The authors would like to acknowledge the contributions of a number of Barclays colleagues who have contributed to the overall debate and implementation of a PIT–TTC ratings framework, including, David Williams, Keith Ho, Tim Thompson, Ian Wilson, Peng Cheng and Yen-Ting Hu. In addition, others who have provided helpful ideas especially include Brian Ranson, who remains forever happy to engage in active dialogue on credit risk issues. We would also like to thank Moody's KMV for the use of their global EDF data. Finally, we thank the editor, Michael Ong for his helpful comments during the editorial process. All errors remain the responsibility of the authors. The views and opinions expressed in this chapter are those of the authors and do not necessarily reflect the views and opinions of Barclays Bank PLC.

2 For the purposes of this chapter, we will generally use PD to refer to both "credit ratings" and probabilities of default. Further below, we explain some more subtle differences between these two concepts.

3 Overall, this approach and discussion is oriented towards corporate counterparties; however, the overall framework is just as applicable to retail and SME customers.

4 For a further discussion of bank rating systems and PIT–TTC issues, see also BCBS (2005).

5 We derive the Z series for a sector by: (i) computing a summary PD measure, often a median or imputed average; (ii) applying the inverse normal function to the summary PD; and (iii) normalising the resulting series so that annual changes have a mean of 0 and standard deviation of 1. The summarisation creates latent measures of systematic risk. The inverse normal transformation creates series with close to normal rather than skewed distributions. The normalisation produces series with properties compatible with the CreditMetrics model (see Gupton *et al* 1997) of conditional PDs and ratings transitions.

   The series that derive from median MKMV EDFs, US bank charge-off (C/O) rates, and agency yearly default rates provide measures of the credit cycle in various sectors. The series based on median MKMV EDFs for each agency rating are used in translating these ratings to PDs. While not, strictly speaking, credit cycle indices, agency rating based series have been correlated historically with the true credit cycle indices. In particular, they reflect the part of the credit cycle not picked up by agency ratings migrations and, instead, tracked by changes in the PDs of each agency grade.

6 We generally refer to this historic comparison of various latent credit factors as a "credit cycle Rorschach test" (the Rorschach test is used by psychologists and involves the interpretation of ink-blot images by subjects).

7 Assessing the "degree of PIT-ness" for hybrid indicators that are combinations of PIT–TTC is key to being able to estimate consistently both fully PIT and TTC PDs. Our approach is to define MKMV EDFs as the PIT benchmark in the sense that these one-year EDFs reflect nearly continuous updates of information on current credit conditions. In contrast, scorecard types of models may use expert judgement assessments and possibly annual financial data that are not updated frequently. In designing this framework we assess each PD model as lying along a continuous PIT–TTC spectrum, ranging from MKMV EDFs as the PIT benchmark to internal scorecard types of model, which we define usually as TTC, if the information content inherent in these is updated only infrequently.

8 The historical Z factors are used to make the adjustments in the likelihood calculation to achieve the required statistical independence.

9 The move to an integrated PIT–TTC approach as discussed in this chapter represents a paradigm shift in the way most banks implement their PDs and credit ratings (see Kuhn 1962). Therefore, as Kuhn describes, it requires a substantial change relative to current thinking.

10 And, in the case of deal evaluation beyond one year, PD term structures need to be considered.

REFERENCES

Aguais, S. D., *et al*, 2004, "Point-in-Time versus Through-the-Cycle Ratings", in M. Ong (ed), *The Basel Handbook: A Guide for Financial Practitioners* (London: Risk Books).

BCBS, 1999, "Principles for the Management of Credit Risk", consultative document, July.

BCBS, 2000, "Range of Practice in Banks' Internal Ratings Systems", discussion paper, January.

BCBS, 2001, "The New Basel Capital Accord, Consultative Document", the Internal Ratings-Based Approach, Bank for International Settlements, January.

BCBS, 2005, "International Convergence of Capital Measurement and Capital Standards", A Revised Framework, Bank for International Settlements, November.

BCBS, 2005, "Studies on the Validation of Internal Rating Systems", Working Paper 14, Bank for International Settlements, May.

Belkin, B., S. Suchower, and L. Forest, 1998a, "The Effect of Systematic Credit Risk on Loan Portfolios and Loan Pricing", *Credit-Metrics Monitor*, April, pp 17–28.

Belkin, B., S. Suchower, and L. Forest, 1998b, "A One Parameter Representation of Credit Risk and Transition Matrices", *Credit-Metrics Monitor*, October, pp 45–56.

Clark, T., 2000, "Can Out-of-Sample Forecast Comparisons Help Prevent Overfitting?", Federal Reserve Bank of Kansas City, research working paper.

Financial Services Authority, 2006, "Strengthening Capital Standards 2", Appendix: Volume 1, draft handbook text, February.

Findley, D, 1993, "The Overfitting Principles Supporting AIC", Bureau of the Census, Washington, DC.

Gordy, M., 2003, "A Risk-Factor Model Foundation for Ratings-Based Bank Capital Rules", *Journal of Financial Intermediation*, **12**, pp 199–232.

Gupton, G. M., C. C. Finger, and M. Bhatia, 1997, "CreditMetrics" technical document, Morgan Guaranty Trust Co., New York, April.

Heitfield, E., 2004, "Rating System Dynamics and Bank-reported Default Probabilities under the New Basel Capital Accord", research paper, Board of Governors of the Federal Reserve System, March.

Jefferys, W. H., and J. O. Berger, 1991, "Ockham's Razor and Bayesian Statistics" (preprint available as "Sharpening Occam's Razor on a Bayesian Strop"), *American Scientist,* **80**, pp 64–72.

Kuhn, T. S., 1962, *The Structure of Scientific Revolutions* (Chicago: University of Chicago Press).

Pluto, K. and D. Tasche, 2005, "Thinking Positively", *Risk*, August.

Ranson, B., 2005, *Credit Risk Management* (Sheshunoff).

Treacy, W. and M. Carey, 1998, "Credit Risk Rating at Large US Banks", *Federal Reserve Bulletin*, **xx(cc)**, November, pp 897–921.