Economic Notes by Banca Monte dei Paschi di Siena SpA,
vol. 34, no. 2-2005, pp. 185–230

Advancing Loss Given Default Prediction Models: How the Quiet Have Quickened

GREG M. GUPTON*

We describe LossCalc^{TM} version 2.0: the Moody's KMV model to predict loss given default (LGD), the equivalent of (1 – recovery rate). LossCalc is a statistical model that applies multiple predictive factors at different information levels: collateral, instrument, firm, industry, country and the macroeconomy to predict LGD. We find that distance-to-default measures (from the Moody's KMV structural model of default likelihood) compiled at both the industry and firm levels are predictive of LGD. We find that recovery rates worldwide are predictable within a common statistical framework, which suggests that the estimation of economic firm value (which is then available to allocate to claimants according to each country's bankruptcy laws) is a dominant step in LGD determination. LossCalc is built on a global dataset of 3,026 recovery observations for loans, bonds and preferred stock from 1981 to 2004. This dataset includes 1,424 defaults of both public and private firms – both rated and unrated instruments – in all industries. We demonstrate out-of-sample and out-of-time LGD model validation. The model significantly improves on the use of historical recovery averages to predict LGD.

(J.E.L.: C33, C52, G33).

* Senior Director of Research, Moody's KMV, 99 Church Street, New York, NY 10007, USA. Tel: +1 (212) 553 1493. E-mail: greg.gupton@mkmv.com (or) www.defaultrisk.com

I thank the numerous individuals who contributed to the ideas, modelling, validation, testing and ultimately writing of this paper: Roger M. Stein with whom I have had the pleasure to work closely for four years and Eduardo Ibarra who has provided invaluable research support throughout this project. Nil Demircubuk, Thom King and Jun Zheng gave enormous assistance with the data. Particular thanks go to Jody Rasch, Navneet Arora and Amnon Levy of Moody's KMV. I thank Kenneth Emery of Moody's Investors Service (for his insights into bank loans).

I have also received invaluable feedback from Moody's Academic Advisory and Research Committee during a presentation and discussion of preliminary findings, particularly from Darrell Duffie of Stanford University and Alan White of the University of Toronto.

1. Overview: Loss Given Default

1.1. Why Loss Given Default Is Important

Loss given default (1 − *recovery rate*) is essential in lending, investing, trading or pricing of bank facilities (loans, commitments, letters of credit etc.), bonds and preferred stock. Accurate LGD estimates are important for provisioning reserves for credit losses and calculating risk capital.

Accurate estimates of LGD are *fundamental* because it is a basic to knowing potential credit losses. It is *important* because any error in predicting LGD is as damaging as a proportional error in estimating the expected default frequency (EDF). Together, they are the sound of two hands clapping.

$$\text{(1)} \qquad \textit{Potential credit loss} = \textit{EDF} \cdot \textit{LGD}$$

Despite its significance, the common practice is to estimate LGD using a look-up table (of either expert opinions or historical averages). This can significantly hamper estimates of credit losses and the resulting credit risk exposures. Increasing the accuracy of LGD estimates improves the precision of both regulatory and economic capital allocation.

Look-up tables are backward looking and static. LossCalc addresses the key issues that have prevented institutions from building more accurate and forward-looking LGD models: (i) lack of recovery observations, (ii) complexity of the recovery process and (iii) lack of insightful predictive factors.

1 *Lack of recovery observations*: Few institutions have sufficient LGD datasets to fully specify and validate a statistical and predictive LGD model. Institutions are beginning to address this issue on how they collect LGD data, but it will be years before there are sufficient data for most to build internal models.[1]
2 *Complexity of the recovery process*: The bankruptcy process makes it difficult to predict how value is assigned to creditors. Even in North America, the *Absolute Priority Rule* almost never fully realizes its potential benefit (Longhofer and Carlstrom, 1995).
3 *Lack of insightful predictive factors*: Even when banks have enough LGD observations, the predictive factors are commonly (i) *static* (e.g. yes/no flags of, say, industry group), (ii) *backward looking* (e.g. default rate indices) and (iii) of unsatisfactorily *low power*.

[1] Our own group within Moody's KMV has worked with several banks to help them specify and create the database of their own LGD experience. These historical loan workout records are commonly archived paper files.

1.2. *Inaccuracies of Look-up Tables*

There is a wide variability in recovery values for instruments even when grouped by debt class and seniority. Any table-driven LGD model lacks. (i) a time-varying factor and (ii) any means of discriminating differences in recovery *within* any given cell of the look-up table.

Figure 1 shows the range of recoveries for instruments based on debt type and seniority class. It is similar to annual default studies by Moody's Investors Service[2] except that it has two additional classes: Industrial Revenue Bonds (IRBs) and Corporate Mortgage Bonds. What is striking in this figure is the wide variability of recoveries even within seniority classes.

1.3. *Use of LGD in Basel II*

Basel II allows internal models to be used in estimating LGD if an institution uses the Advanced IRB approach. Although initially a standard LGD allocation may be used (the Foundation Approach), institutions that have adopted the IRB approach for probability of default are being encouraged to use the IRB approach for LGD. The IRB gives a more accurate assessment of loss.

In order to qualify for the IRB approach:

> *... A bank must estimate an LGD for each of its internal LGD grades. ... Each estimate of LGD must be grounded in historical experience and empirical evidence. At the same time, these estimates must be forward looking ... LGD estimates that are based purely on subjective or judgmental consideration and not grounded in historical experience and data will be rejected by supervisors.* (Basel Committee on Banking Supervision § 336 & 337)

We believe that LossCalc research can assist in meeting the requirements of LGD specified under Basel II.

2. The LGD Model

2.1. *Overview*

We have fit and validated a security-level LGD model using 23 years of global data for loans, bonds and preferred stock. We distinguish six country/regions: Asia (which includes Australia and New Zealand),

[2] See Exhibit No 20 in Hamilton *et al.* (2001).

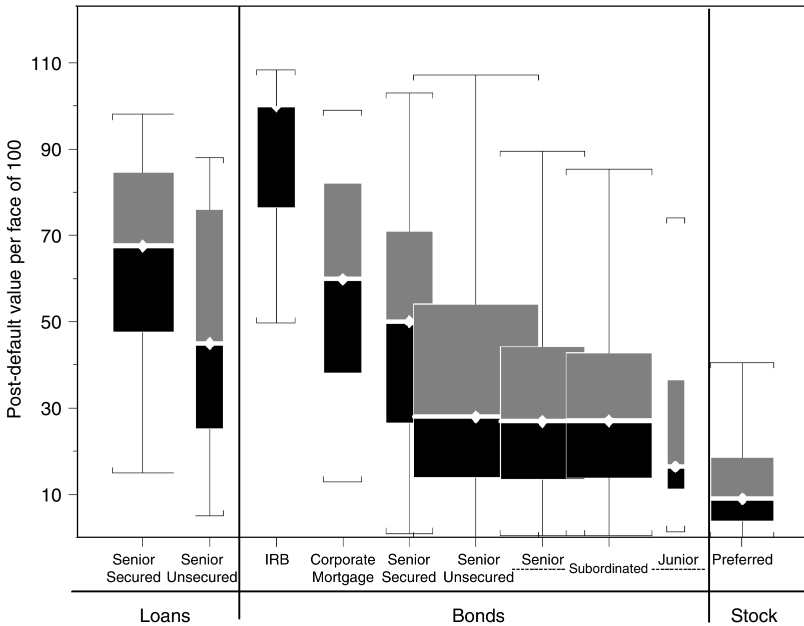

Figure 1: Recovery Value by Debt Type and Seniority Class, 1981–2003, Global
Notes: Grouping instruments shows a pattern but still leaves great variability: 1981–2003, Global. The shaded boxes cover the interquartile range (grey) extending from the 25th percentile to the median whereas the black extends from the median to the 75th percentile. White bars mark the medians. Squared brackets mark the substantive data range. The width of each box is proportional to the square root of the number of observations within each group.

Canada, Europe, Latin America, the United States and the United Kingdom. Our model forecasts the LGD for defaults occurring immediately and one year from the time of evaluation. Prediction horizons are not used in most LGD predictions because of the *static nature* of historical averages, which the prevailing market practises to estimate LGD. Averages overlook (i) the point in the credit cycle and (ii) the sensitivity of a borrower to the economic environment. Because the LossCalc dataset spans from 1981 to 2003 with 3,026 observations of recovery values and 1,424 defaulted public and private firms in all industries, it can incorporate both macroeconomic information and firm-specific cyclicalities in its recovery estimates. Our model uses predictive information on five different levels of information: collateral, instrument, firm, industry and macroeconomy/geographic.

2.2. Time Horizon

We actually implement models for our two risk horizons 'immediate' for risk horizons *up to one year* and 'one-year' for risk horizons that are over one year. Look-up table models of LGD cannot address different risk horizons because of the inherently *static nature* of a table. Our models use the same functional form and are fit and tested in the same way. The only difference is the lagging of the factors and weights assigned to each.

We find that the remaining life of an instrument (i.e. tenor the debt would have had if it had not defaulted) is *not* predictive of LGD. The relevant timeframe is the risk horizon rather than the instrument's maturity. Parenthetically, this observation is supportive of the Recovery of Par (RP) hypothesis.[3]

2.3. Predictive Factors

LossCalc v2 uses nine explanatory factors to predict LGD. We organize these into five broad groups:

1 *Collateral and backing*: this includes cash, 'all assets', property plant & equipment (PP&E) and support from subsidiaries;
2 *Debt type/seniority class*: debt types are loan, bond and preferred stock and seniority classes are secured, senior unsecured, subordinate etc.;
3 *Firm status*: cycle-adjusted leverage, relative seniority standing and (for public companies) Moody's KMV distance-to-default (D2D);[4]
4 *Industry*: historical average of industry recoveries and D2Ds across many firms aggregated at the industry (and regional) level;
5 *Macroeconomic/geographic*: regional flags (i.e. Asia, Canada, Europe, Latin America, the United Kingdom and the United States); the industry-level D2Ds mentioned earlier serve double duty as we aggregate them separately within each country/region.

These factors have little colinearity (intercorrelation), each is statistically significant both univariately and in combination, and so they join to make a more robust prediction of LGD.

2.4. Framework

LossCalc is a data-intensive, empirically driven, statistical model that adheres to economic principles. The broad steps in this framework are transformation, modelling and mapping.

[3] There are various frameworks for expressing the recovery on defaulted debt. Following nomenclature from Schönbucher (2003), they include Recovery of Treasury (RT), Recovery of Market Value (RMV) and RP. Testing these alternatives empirically, Guha (2002) found that RP was the best characterization of recovery data. Interestingly, his data source was Moody's Investors Service and so it significantly overlaps with the dataset that we use.

[4] The 'D2D' is an output of a structural (Merton-type) valuation of the debt. It is the firm's debt funding measured in units of standard deviations of asset volatility (Crosbie and Bohn, 2003).

1 *Transformation*: We transform raw data into 'mini-models' rather than taking the simple levels of factors. For example, we find that leverage is more damaging to recoveries during downturns in the credit cycle. We thus interact leverage by the 'Global All Corporate Default Rate'.
2 *Modelling*: Once we have transformed individual factors and converted them into mini-models, we aggregate these using multi-variate regression techniques.
3 *Mapping*: We statistically map the model output to historical LGD.

Each of the three steps in this process relies on the application of standard statistical techniques. We outline the details of these in section 3.2.6.

2.5. Performance

Our model is a better predictor of LGD than the traditional methodologies of historical averages segmented by debt type and seniority. By 'better', we mean that we have

1 significantly *lower error* as stated by mean squared error (MSE) or standard deviation (Figure 9);
2 significantly *more correlation* with actual LGD: this means they have better tracking of both high and low recoveries (Figure 10);
3 *better discrimination between instruments of the same type*: for example, the model provides a much better ordering (best to worst recoveries) of bank loans than historical averages (Figure 11);
4 *far fewer large errors* (Figure 12): over 10 per cent of the time, this *reduction in error is greater than 28 per cent of original par value*;
5 narrower prediction intervals (PIs) than historical LGDs and a previous version of our model.

2.6. The Dataset

Our dataset links LGD observations and company data from Moody's Investors Service, with D2D and related statistics from Moody's KMV, with financial statement data from Compustat and WorldScope. We fit our models on 3,026 observations *globally* of LGD for defaulted loans, bonds and preferred stock (total of 10 seniority grades) extending from January 1981 to December 2003. We distinguish six country/regions: Asia (which includes Australia and New Zealand), Canada, Europe, Latin America, the United States and the United Kingdom with at least seven years of data in each. The dataset includes 1,424 defaulted public and private firms in all of 62 industries. In US$ equivalents, the issue sizes range from $370 thousand to $4.6 billion, with a median size of about $125 million. The median firm size (sales at annual report before default) was \$660 million but ranged from zero to \$48 billion.[5]

2.7. *Validation*

The primary goals of validation and testing are to

1 determine how well a model performs;
2 ensure that a model has not been over-fit and that its performance is reliable and well understood;
3 confirm that the modelling approach, not just an individual model, is robust through time and credit cycles.

We apply *walk-forward* validation. It involves fitting a model on one set of data from one period and testing it on a subsequent period. We then repeat this process, stepping ahead in one-year increments until we have tested the model on all periods up to the present. Thus, we never use data to test the model that we used to fit its parameters and so minimize over-fitting. We can also assess the behaviour of the modelling *approach* over various economic cycles. Walk-forward testing is a robust methodology that accomplishes the three goals set out earlier.[6]

3. Factors

The central goal of our approach to forecasting LGD is to increase predictive power through the inclusion of multiple factors, each designed to capture specific aspects of LGD determination. We take a statistical approach using multiple explanatory factors to develop an immediate and one-year LGD.

3.1. *Definition of LGD*

LossCalc defines recovery ($1 - LGD$) on a defaulted instrument as its market value approximately one month after default.[7] The model uses security-specific bid-side market quotes.[8]

[5] Note that neither debt nor firm size is predictive of LGD in this dataset and has not been included in our models.

[6] (Sobehart *et al.* 2000a, b) describe the walk-forward methodology more in detail. Appendix B of this paper gives a brief overview of the approach.

[7] The date of default is not always well defined. As an example, bankers commonly write loan covenants with terms that are more sensitive to credit distress than those of bond debentures. Thus, different debt obligations of a single-defaulted firm may technically default on different dates. The vast majority of securities in our dataset have quotes within the range of 15–60 days after the date assigned to the initial default of the firm's public debt. Importantly, our study found no distinction in the quality or explicability of default prices across this 45-day range.

[8] Quotes are contributed by IDC, Goldman Sachs, Citibank, BDS Securities, Loan Pricing Corporation, Merrill Lynch, Lehman Brothers, LoanX and LPC. 'Matrix' prices were not accepted.

Moody's KMV chose to use price observations one month after default for three reasons. This period

1 gives the market sufficient time to assimilate new post-default corporate information;
2 is not so long after default that market quotes become too thin for reliance;
3 best aligns with the goal of many investors to trade out of newly defaulted debt.

This definition of recovery value avoids the practical difficulties associated with determining the post-default cash flows of a defaulted debt or the identification and valuation of instruments provided in the replacement of defaulted debt. The very long resolution times in a typical bankruptcy proceeding (commonly 1.25 to 5 years) compounds these problems.

Figure 2 shows the timing of price observation of recovery estimates and the ultimate resolution of the claims. Broker quotes on defaulted debt provide a more timely and objective recovery valuation vs. waiting to observe the completion of court-ordered resolution payments. A model built on resolution data (accounting LGDs) would need to be fit to defaults that were all at least two years in the past.

The Relationship of Market Pricing and Ultimate Recovery: There have been several studies of the market's ability to price defaulted debt efficiently (Eberhart and Sweeney, 1992; Ward and Griepentrog, 1993; Wagner, 1996; Altman *et al.*, 2004). These studies do not always show statistically significant results, but they consistently support the market's efficient pricing of ultimate recoveries. At different times, Moody's has studied recovery estimates derived from both bid-side market quotes and discounted estimates of resolution value. We find consistent with the

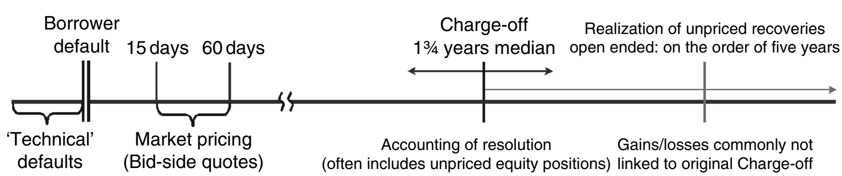

Figure 2: Timeline of Default Recovery Estimation

Notes: This diagram illustrates the timing of the possible observation of recovery estimates. Recovery from a default is an extended process rather than a culminating event. After default, the market prices the expectation of anticipated recoveries. These are most liquid 15–60 days after default. Some $1^3/_4$ years later, half of defaults have been charged-off the accounting books. Recoveries at that point include cash, new debt extensions, equity in the (emerging) borrower etc. Equities may not trade and may not have a market price. Eventually, all receipts in satisfaction of the default realize a value, but this is typically not traceable back to the original borrower.

academic research that these two tend to be unbiased estimates of each other.

3.2. Factor Descriptions

Historical averages broken out by debt type (loan, bond and preferred stock) and seniority class (secured, senior, subordinate etc.) are important factors in predicting LGD. However, as factors in our model, they account for only about 40 per cent of the influence in predicting the levels of recoveries. The distinguishing benefit of our approach is that we assemble multiple predictive factors at different information levels. This approach is powerful because (i) each factor is predictive, (ii) they are uncorrelated with each other (i.e. they speak to very different parts of the puzzle) and (iii) they aggregate within a consistent framework.

We group LossCalc's factors into five categories of predictive information (Table 1): (i) collateral, (ii) debt type and seniority class, (iii) firm-level information, (iv) industry and (v) macroeconomic/geographic information. These factors have low colinearity (little intercorrelation) and together make a significant and more accurate prediction of LGD. All factors enter both LossCalc forecast horizons (i.e. immediate and one-year) in the same direction.

All of these factors are individually highly statistically significant and with their signs in the expected direction, both population wide and within all sub-populations tested. Figure 3 shows the contributions of each broad factor category towards the prediction of the immediate and one-year LGD forecasts. Bars of each colour add up to 100 per cent.

3.2.1. Collateral and Support

We define six different types of collateral and support for secured instruments. These broad collateral types summarize the more detailed information in the Moody's Investors Services default database. We let two (of the ten) debt/seniority types link with collateral/support: Senior Secured Loans and Senior Secured Bonds. The types of collateral and support are as follows:

1 *Cash and marketable securities collateral*: This is cash, compensating balances linked balances and liquid instruments held on account;
2 *Pledge of 'all assets' of the firm*: On its face, this would appear to be the best a lender could hope for, but there arise practical questions of its absolute scope and enforceability;
3 *Generically secured, by 'unknown'*: This category is used where the specific nature of the collateral is not available. The effect is set at the average across all collateral (that *is* identified) in our dataset;

Table 1: Explanatory Factors in the LossCalc Models

Collateral and other support	
The proportion of coverage (of the exposure) by cash, 'all assets' or property, plant and equipment. Support from subsidiaries takes a yes/no flag rather than a coverage ratio	Collateral
Debt type and seniority class	
LGD, controlling for debt type (loan, bond and preferred stock) and seniority classes (senior, junior, secured, unsecured, subordinate etc.)	Historical averages
Firm-level information	
Seniority standing of debt within the firm's overall capital structure; this is the *relative* seniority of a claim. This is different from the *absolute* seniority stated in debt type and seniority class above. For example, the most senior obligation of a firm might be a subordinate note if no claim stands above it	Seniority standing
Cycle-adjusted firm leverage (gearing): All Corporate default rate interacted with the default probabilities directly implied by book leverage	Leverage
The firm's distance-to-default (applied to public obligors only)	Firm distress
Industry	
Historical normalized industry recovery averages after controlling for seniority class	Industry experience
The industry's distance-to-default (tabulated by country/region)	Industry distress
Macroeconomic and Geographic	
The country/region's distance-to-default (tabulated by industry)	Region distress
Country/region shifts in mean expectation	Shift

Notes: This is a summary of the factors applied in Moody's LossCalc model to predict LGD. The table highlights the five broad categories of predictive information: collateral, instrument, firm, industry and broad economic environment. These factors have little intercorrelation and join to make a powerful LGD prediction.

4 *PP&E*: This represents the actual means of production. In the majority of cases, it is a physical asset, but in our model it is extended to include other 'instruments of production' (e.g. airport landing rights);
5 *Subsidiary support*: This refers to any (i) guarantees by subsidiaries, (ii) pledged stock of subsidiaries or (iii) pledged and substantively sized key assets of subsidiaries;
6 *Unsecured*.

We allow combinations of different collaterals (with the exception of PP&E combined and subsidiary support which did not appear in our dataset). This reflects the typical practice of bankers to secure as much collateral as they reasonably can.[9] Moody's work on collateral and support goes back to 1998 (Hamilton, 1999).

[9] Moody's KMV has material that is available on a client's request to give guidance in mapping different collaterals into LossCalc's six categories. This defines 78 collateral definitions that reflect 15 unique codings within LossCalc.

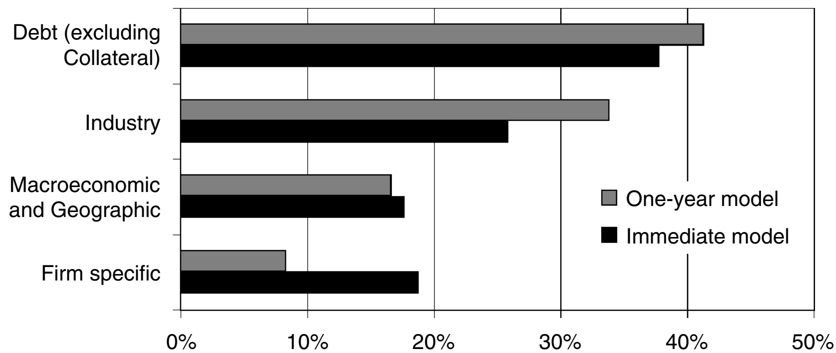

Figure 3: Relative Influence of Different Factor Groups in Predicting LGD
Notes: This figure shows the normalized marginal effects (relative influence) of each broad grouping of factors. These groupings give intuition as to the source of LossCalc's power, but there is overlap between groups. For instance, we tabulate industry distance-to-defaults over time and within individual country/regions. Therefore, it influences both the second and third groupings in this graph simultaneously. Note that the influence of collateral is substantive (Figure 4). Collateral is *not* included here and so it would represent predictive power *in addition* to the factors shown here.

Importantly, collateral's boost in predictive power is *in addition to* the power from LossCalc's other factors. In this document, *none* of the other figures and tables include the effect of collateral. So the boost in power shown in Figure 4 is additive to the power we illustrate elsewhere.

3.2.2. Debt Type and Seniority Class

Historical average recovery rates are a starting point for LossCalc. Controlling for debt type and seniority classes has several benefits. In North America, it addresses the effects of the *Absolute Priority Rule* of default resolution. This gives our model a base from which to make adjustments in other countries/regions.

3.2.3. Firm-level Information

We consider three types of information for the firm: (i) cycle-adjusted leverage ratio, (ii) the standing in the capital structure and (iii) the credit distress of the firm.

1 *Firm leverage (or gearing)* This is how much asset value is available to cover the liabilities of the firm. This notion is imperfect because asset values may be different from how they were booked and their worth may rapidly decline as a firm approaches bankruptcy. To this, we

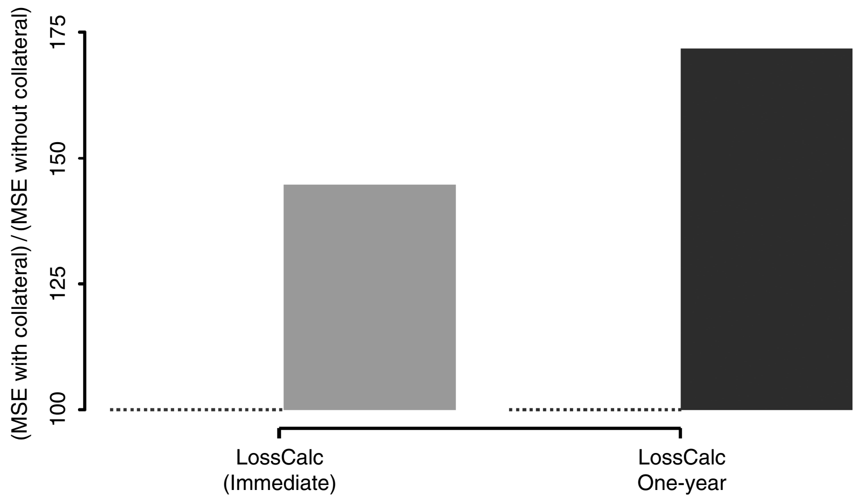

Figure 4: Relative Influence of Different Factor Groups in Predicting LGD
Notes: This figure shows relative improvement (reduced mean squared error (MSE)) of LossCalc's LGD forecasts when compared with actual LGD. We show results for both the immediate forecast (in grey) and the one-year forecast (in black). The dotted lines represent the performance of the LossCalc models where no collateral information is input. The solid bars show the relative reduction in MSE. For example, when running LossCalc with particular collateral information, its one-year forecast is 72 per cent more accurate compared with LossCalc run selecting merely 'Generically Secured by Unknown'.

apply a Heckman adjustment[10] because the leverage ratio is itself predictive of the event of default. We do not apply the leverage ratio in the case of secured debt or for financial industries.[11]

We also find that firm leverage has more of an impact on LGD during periods of economy-wide credit distress. When relatively many firms in the economy are defaulted, then a firm's leverage tends to count against it more than during less distressed times. We address this relationship by interacting the leverage ratio with the *Global All Corporates Default Rate* as published by Moody's Investors Service.

2 '*Standing*' *within the firm's capital structure* is a debt's *relative* seniority (i.e. are there claimants who stand more senior at the time of default).

[10] Heckman (1979) has described the potential bias caused by applying a factor that is itself predictive of being in the conditioned state of the world. In this case, because leverage is a strong predictor of being in default, then the leverage ratio needs to have a bias adjustment to properly predict LGD.

[11] Specifically, this applies to 'Senior Secured Loans', 'Senior Secured Bonds' and 'Corporate Mortgage Bonds' as well as Lessors, Real Estate, Banks and S&Ls, and Finance not otherwise classified. Secured claims look principally to the security for satisfaction, and the leverage of financial firms is far more problematic to judge than for corporate and industrials.

For example, preferred stock is the lowest seniority class short of common stock, but it might hold the *highest* seniority rank within a particular firm that has no funding from loans or bonds. In addition, some more senior class of debt may mature thus revealing another (lower) debt class to be 'most senior'.

The ordering that LossCalc applies for this tabulation is the same as that shown in Figure 1. Exceptions are 'Industrial Revenue Bonds' and 'Corporate Mortgage Bonds'. These two debt classes are 'outside' of our models' relative seniority scheme. They are not influenced by (and they do not influence) any other seniority class that may be in the firm's capital structure at the time of default.

It is reasonable to ask why we did not use a predictor such as 'the *amount* of debt that stands more senior' or 'the *proportion* of total liabilities that is more senior?' Although these seem intuitively more appealing, we chose the simpler indicator for two reasons:

- *Resolution Procedure*: In bankruptcy proceedings, a junior claimant's ability to extract concessions from more senior claimants is not proportional to its claim size. Junior claimants can force the full due process of a court hearing and so have a practical veto power on the *speediness* of an agreed settlement.[12]
- *Availability of Data*: Claim amounts at the time of default are not the same as original borrowing/issuance amounts. In many cases, borrowers partially pay down their debt before maturity. Amortization schedules (for loans) and sinking funds (for bonds) are examples of this. Determining the exposure at default for many obligations can be challenging, particularly for firms that pursue multiple funding channels. In many cases, these data are unavailable. Requiring such an extensive detailing of claims before being able to make *any* LGD forecast would be onerous in industry use.

3 *Firm-specific D2D* is the best measure of credit distress for public firms that we found. It is also attractive in that there is a clear economic rationale for constructing the variable. It uses the firm's capital structure as well as information from the equity markets that is a timely and efficient signal of the collective belief of the firm's future prospects. It is an economically important measure because a firm that suffers more severe distress has higher LGD. Within a Merton-type structural model of debt, the D2D is addressing the 'coverage' aspect of the

[12] We tested this on a sub-population selected to have fully populated claim amount records. The best predictor of recoveries, both univariately and in combination with a core set of LossCalc factors was a simple flag of Who-Has-Highest-Standing. We tested many alternatives, such as amount or proportion of 'cushion' and amount or proportion of 'overhead' as well as certain transformations such as logarithms.

leverage ratio we discussed earlier, but it is accounting for much more information and doing a more accurate job. This measure is only available for publicly traded firms.

3.2.4. Industry

Researchers frequently compile recovery averages broken out by industry in an attempt to refine historical estimates.[13] An example implementation would be to have a look-up table across industry categories to read off the LGD for a borrower in that industry. The assumption is that industry characteristics are persistent over time so, for example, a *hard asset* and *high franchise* industry like Gas Utilities would have low LGDs consistently across time; an opposite example would be Business Services with consistently high LGDs.

However, we find strong evidence of industry-level variability in recovery rates *across time*. Furthermore, the timings of these movements are not coincident from one industry to another. We also find that some industries enjoy periods of prolonged superior recoveries but fall well below average recoveries at other times, like the case of Telephone companies discussed below. A simple industry bump up or notch down, held constant over time, does not capture this behaviour. We see examples in Figure 5.

The two upper panels of Figure 5 show the mean recoveries for different industries and the distribution around those recoveries compared with the entire population. They show a low recovery of Business Services relative to high recoveries in Gas Utilities. In addition, industry recovery distributions (any sub-grouping really) change over time. This implies that observations of, say, the three humps in Gas Utilities, is not meaningful because we could not expect it to repeat in future periods. The Telephone industry is an example of industry shift. The bottom panels of Figure 5 show the recent decline in Telephone recoveries.

For many years, the phone industry was mature in its technology and somewhat like a utility with redeployable assets and above-average recoveries. However, as always happens, things changed. Wireless technology emerged, the asset base in the industry faced a quicker obsolescence cycle and a significant proportion of once hard assets shifted to assets like ‘air rights’. Not surprisingly, recoveries fell and the seemingly well-established industry-level distribution did not prove to be predictive.

[13] See Altman and Kishmore (1996) and Izvorski (1997) for broad recovery findings by industry and Borenstein and Rose (1995) for a single industry (airlines) case study.

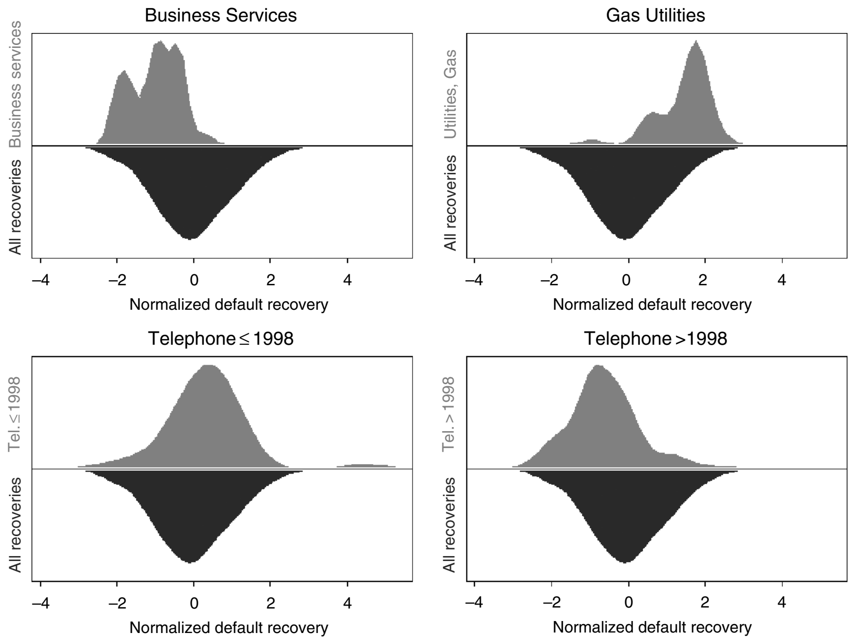

Figure 5: Industries Recover Differently and Are Not Static
Notes: Shown here in black for each panel is the full dataset of LossCalc's dependent variable, normalized default recovery. We then contrast how industries might recover quite differently in the examples of Business Services vs. Gas Utilities in the top panels. Importantly, an industry distribution (any sub-population really) cannot be fixed and forgotten because it may well change over time, see Telephone industry example on the bottom panels. In all four panels, we contrast the variable nature of an individual industry (in grey) against LossCalc's, normally distributed, recovery population. Multiple 'humps' in the industry distributions are not meaningful because they are not persistent over time. See section 4.1.

To address industry differences across time, we first organize the dataset into the 62 specific industries defined by Moody's KMV.[14] We then produce two different measures of industry behaviour:

1 the industry's historical recovery experience;
2 the aggregated D2D across all firms in that industry (and region).

We recompile both of these measures monthly so they vary over time. By construction, the D2D is the more dynamic of these two factors. These measures exhibit stable predictive behaviour across industries and country/regions (Figure 6).

[14] A mapping of SIC codes to these industry groups is available upon client request.

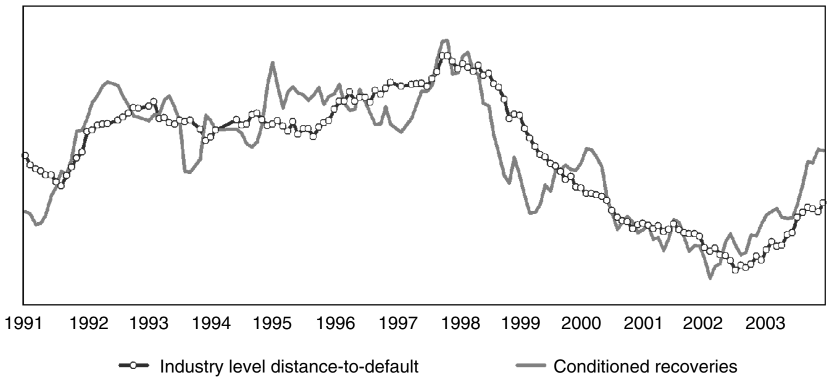

Figure 6: Recoveries are Lower in Distressed Industries
Notes: Distance-to-default measures are aggregated for each industry and region, but only those that coincide with an LGD observation are averaged and plotted here. Separately, we want to see the trend in recoveries across all seniority classes. Therefore, it is important for us to put them on a comparable basis. To do this, we first normalize them and then control for their seniority class. The data are global.

3.2.5. Macroeconomic

Our D2D measure is useful because recovery rates tend to rise and fall together rather than being fully independent. In addition, because recoveries have positive and significant intercorrelation within bands of time, this effect has material implications for portfolio calculations of Credit-VaR. A positive correlation between EDF and LGD lengthens the tail of a portfolio loss distribution, thus raising economic capital assessments.

Note that the inclusion of the D2D factor can replace a number of other traditional macro indicators that are sometimes used as proxies for the credit cycle, including those that were included in version 1 of LossCalc. These deposed indices were the RiskCalc Probability of Default index, the Moody's Bankrupt Bond Index (MBBI) (Hamilton and Berthault, 2000) and changes in Leading Economic Indicators.

Trailing 12-month All Corporate Default Rate (global) is used as a factor in LossCalc v2. It interacts with book leverage to give a credit cycle adjustment to that firm indicator. It replaces the US Speculative-grade Trailing 12-month Default Rate used in version 1. Moody's Investors Service publishes both indices monthly.

We had tested an alternative factor, which was a set of sector-level default rated indices. We rejected this because they were far too noisy. Because defaults are rare and can occur in clumps, default rate averages are volatile indicators when aggregated across small populations such as an industry sector. In contrast, our D2D statistic is a continuous measure available for 25,000 firms globally. D2D is also *forward* looking.

3.2.6. Geographic

Ours is an international model with LGD observation sourced globally. Likewise, our predictive factors are sourced separately within each of the country/regions that LossCalc addresses. Factors applied in Europe, for example, are not weighted/mixed/supplemented in any way with the factors applied in any other country/region – including the United States. Because we have detailed credit information on 25,000 firms worldwide, we have been able to construct granular and powerfully predictive indices broken out by (i) country/region, (ii) industry and (iii) across time.

Legal Differences: Although legal differences in bankruptcy play a part in recovery, we find that fundamental economics drives firm values and recoveries in a predictable way regardless of the particular legal jurisdiction.

The LGD 'process' has two distinct steps. First is to determine the aggregate firm-level value of the defaulted entity. In other words, what is the total worth available with which to satisfy all the claimants of the firm? Second is to determine how the value of the firm is divided up among the firm's debtors. Legislation differences, which differ by country, only affect the second step.

The cross-national LGD forecasting on a firm-wide level is addressed by using the D2D indices compiled by country/region. LossCalc addresses the major second-level components through 'masking off' of certain factors. An example is that the 'Most Senior Debt' factor (which addresses relative seniority) only applies in North America and is *masked off* elsewhere.[15] This adjusts for the United States' use of the *Absolute Priority Rule* (although with common practical exceptions) and the Canadian experience that closely emulates the US rules (even though written law differs), whereas outside North America, various considerations beyond an absolute rule of priority come into play.

Referring to Figure 7, we see that by properly addressing both 'steps' in the LGD process LossCalc captures a significant portion of the variability across countries and regions.

[15] For recent academic research into non-US LGD, see Davydenko and Franks (2004), Fisher and Martel (2003), Schmit and Stuyck (2002), Singh (2003) and Xu (2004).

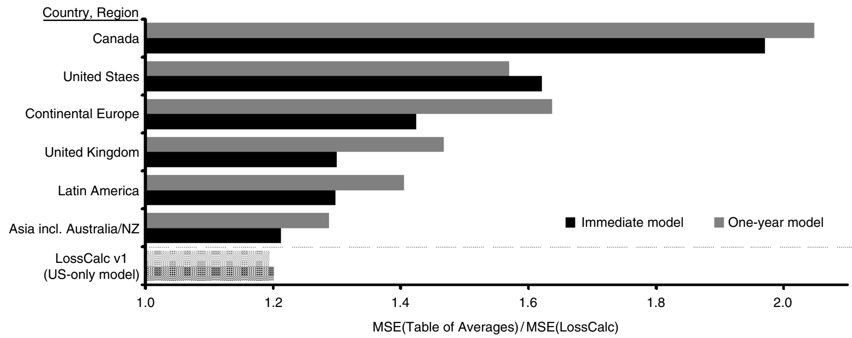

Figure 7: LossCalc Accuracy by Country/Region

Notes: Shown here is the improvement in the mean squared error (MSE) performance measure of LossCalc relative to a Table of Averages. We show two risk horizons for each of six country/regions that LossCalc covers. A horizontal bar with no length (i.e. simply marking the 1.0 value would say that LossCalc did as well as a Table of Averages. Thus, LossCalc outperformed in all cases. LGD predictions were most accurate in Canada. All country/regions outperformed the US-only version 1 of LossCalc (see ghosted bars at bottom).

4. Modelling Framework

The steps in the LossCalc methodological framework are transformation, modelling and mapping.

1 *Transformation*: We transform factors into 'mini-models'. This increases factors' predictive power univariately.
2 *Modelling*: We aggregate mini-models using regression techniques.
3 *Mapping*: We map model output to historical LGD statistically.

4.1. Establishing an LGD Measure

There are several potential sources of LGD data. Each has its strengths and weaknesses. We have organized a 'Family Tree' of LGD data sources in Figure 8.

4.1.1. Accounting Loss

Accounting measures of LGD are available typically only for loans because banks source these from their accounting records. This information is also scarce because typical bank systems have not captured recovery cash flows and tracked the collateral realizations electronically, so collecting it becomes a manual extraction from paper files. In addition, the value at resolution is often subjective because equity, rights and warrants received as payment in resolution commonly have no market price. In previous research, Hamilton (1999) found that 15 per cent of the recovery receipts to satisfy Senior Secured Loans came in the form of equity of the

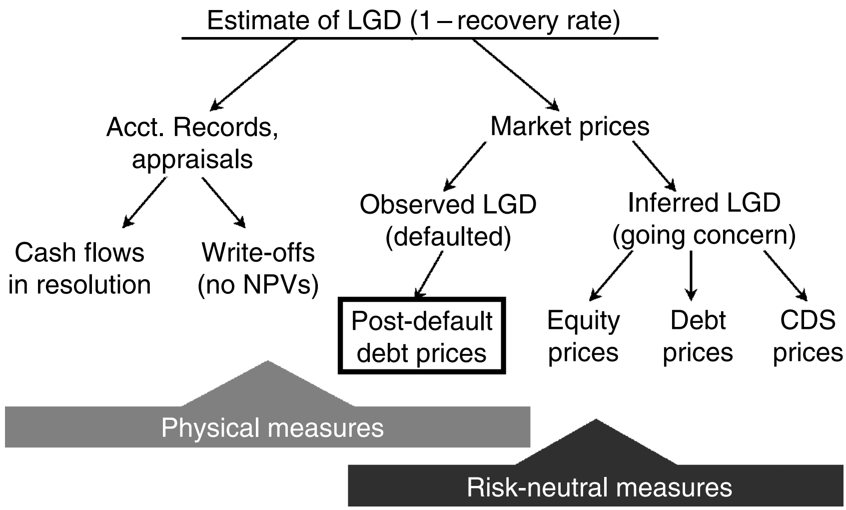

Figure 8: Sources of LGD Observations

Notes: There are several ways of observing a LGD (or recovery rate). They all have their pros and cons. The accounting measure can be directly observed and are called 'Physical Measures'. Market valuations represent fair market value and are 'Risk Neutral Measures'.[16] The dependent variable for LossCalc falls in the middle because it is both physically observable and a fair market valuation. LossCalc uses bid-side quotes about one month after default as the recovery rate.

defaulted (and presumably re-emerging) firm. Because these equity interests commonly do not trade, their value will be unrealized and unknown for years. Once the equity amount is estimated, the actual realized value of the equity is commonly *not* linked back to the original charge-off.

The following two accounting values are called 'physical measures'.

- *Cash flows from a default resolution*: Using cash flows describes the full profile of when the creditor, typically a lending bank, receives economic value. In order to calculate the Credit-VaR, the institution could calculate a net present value of recoveries.
- *Charge-off amounts*: An alternative estimate of accounting. LGD is the lender's net write-off value when recovery efforts are exhausted or when they choose to abandon a debt.

Note that *both* these approaches suffer the same uncertainties because both fall short of the ultimate realization of all payments typically given in resolution (Figure 2).

[16] In mathematical finance, a *risk-neutral measure* is today's fair (i.e. arbitrage-free) price of a security, which is equal to the discounted expected value of its future payoffs. The measure is so called because, under that measure, all financial assets in the economy have the same expected rate of return, regardless of the asset's 'riskiness'. This is in contrast to the *physical measure* – i.e. the actual prices where (typically) more risky assets (those assets with a higher price uncertainty) have a greater expected rate of return than less risky assets.

4.1.2. Market Valuation of LGD

The right-hand side of Figure 8 shows ways of inferring LGD from market prices. The market-based values represent fair pricing and are called 'risk-neutral measures'. Interestingly, because defaulted debt prices are both observed and fair valuations, they are both a *physical* and a *risk-neutral* measure.

Inferring LGD from non-defaulted security prices requires first estimating the default likelihood of the firm (given either a 'structural' or 'reduced form' model) and then picking the LGD to best reconcile the market price of the debt or CDS with its model valuation. This approach is most suited for securities valuation because it can reconcile the pricing of different securities related to one firm.

Direct observations of market value of newly defaulted debt does not require this modelling step. For investors in liquid credit, direct observation directly represents their recovery, as it is the security's realizable value. We find that market liquidity is good at about one month after default, as many investors are disallowed from holding defaulted assets and so trade their positions to a group that specializes in defaulted debt.

Interestingly, we had conducted an unintentional experiment. We found that in version 2 of our model our Industry D2D factor easily displaced the MBBI, which we had used in version 1. A criticism of LossCalc v1 had been that our use of market prices was essentially flawed under the argument that vicissitudes in supply/demand in the distressed debt market was so strong that those prices were not useful as a recovery proxy. However, these hypothetical vicissitudes are exactly what an index like the MBBI should well capture. The fact that the MBBI dropped out as insignificant when industry D2D entered is strong evidence that supposed imbalances in supply and demand do not drive market prices of defaulted debt. Thus, market prices are truly reflecting credit (LGD) status.

Regarding a related issue, we have heard concerns that recovery information from defaulted loans might be of lower quality because of the perception that loans might have less liquidity than bonds. In fact, an academic study has recently addressed this very issue (Altman *et al.*, 2004). They find that '... the loan market is informationally more efficient than the bond market around loan default dates and bond default dates'. They find that this is consistent with the monitoring role of loans.

4.1.3. Loss Distribution

We find that the distribution of LGD is not normally distributed. An alternative distribution that better approximates recoveries in our data is the Beta-distribution. We 'transform' these Beta-distributed defaulted debt prices into a normally distributed dependent variable by applying techniques such as we describe in Appendix B. Making recovery rates match to a Normal distribution makes them statistically 'well-behaved' statistical modelling. We create separate Beta-distribution transformations for each debt type: loans, bonds and preferred stock.

The Beta-distribution commonly ranges within the interval of zero to one but is not restricted to being symmetrical. This matches the recovery rate range of 0 to 100 per cent recovery. A Beta-distribution can be specified by two parameters: its 'centre' and 'shape'. This means that it has great flexibility to describe a wide variety of distributions. It can equally well represent loan recoveries averaging 60–80 per cent or recoveries averaging 30–40 per cent for bonds. It can do this while still maintaining smoothed cut-offs at the bounds of zero to one. Although a Beta-distribution could be manipulated to generate huge (even infinite) probability mass at its bounds of zero and one, such extremes are not needed (and not used) within our model.[17]

4.2. Transformation and Mini-modelling

Before the final modelling, the individual variables are assessed on a stand-alone (univariate) basis. Some of the univariate variables, however, are directly affected by other variables. Examples include (i) the relationship of leverage and corporate default rates, (ii) historical industry LGD and (iii) industry-level D2D.

Leverage and corporate default rates: Higher leverage suggests worse recoveries, but the impact is more pronounced during times when many firms in the economy are defaulting (Gupton and Steen, 2002). Across our 23-year sample, we found the influence of the default rate on leverage reaches a limit. In order to reflect that, we need to combine leverage and default rate into a mini-model where this dampening has a reduced marginal effect above a certain bound.

Historical industry LGD: There are differences across time in recoveries according to their industry group. Some industries, such as service firms, may have 'softer' assets compared with others such as a Gas Utility (top panels of Figure 5). In order to estimate industry LGD effectively, we wish to use all seniority classes. To do this, we developed a mini-model that normalizes and standardizes the LGD and then compares recovery rates by the number of standard deviations below or above the historical expectation conditioned on its debt type and seniority class.

[17] Application of a Beta distribution is robust across many LGD models and datasets, see recent research: Gordy and Jones (2002), Ivanova (2004), Onorota and Altman (2003), Pesaran *et al.* (2004), Singh (2003) and Tasche (2004).

Industry-level D2D: The industry D2D is reflects the value of industry assets and recoveries. It is forward looking, powerful and consistent across all industries, countries and time. Because the primary buyer for defaulted assets is another firm within the defaulter's industry, if those firms are in a period of credit distress, then they will have less incentive and fewer resources with which to purchase defaulted assets and the recovery on those assets will be low. A mini-model of asset values and industry D2D is constructed to account for this affect.

By building mini-models, our factors are better stand-alone predictors of default. Both mini-models and stand-alone measures are univariate measures. This level of modelling is performed before assembling an 'overall' multi-variate model.

4.2.1. Factor Inclusion

The model drops certain factors in certain cases when it would not make economic sense to apply them. For example, although *leverage* is one of the nine predictive factors in the LossCalc model, it is not included for financial institutions.[18] These are typically highly leveraged with lending and investment portfolios having very different implications than an industrial firm's plant and equipment.

Similarly, we do not consider leverage when assessing secured debt.[19] The recovery value of a secured obligation depends primarily on the value of its collateral rather than broad recourse to general corporate assets.

4.3. Modelling and Mapping: Explanation to Prediction

The modelling phase of the LossCalc methodology involves statistically determining the appropriate weights to use to combine the transformed variables and mini-models. The combination of the predictive factors is a linear weighted sum, derived using regression techniques. The model takes the additive form without a constant term:

$$\hat{r} = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \cdots + \beta_k x_k \tag{2}$$

Where the x_i are either the transformed values or mini-models, the β_i are the weights and $\hat{r}$ is the normalized recovery prediction. $\hat{r}$ is in 'normalized space' and it is not yet in 'recovery rate space'. Therefore, the final step is

[18] Industries that do not participate in the firm leverage factor include Banks and S&Ls, Finance N.E.C., Lessors and Real Estate.

[19] Seniority classes that do not participate in the firm leverage factor include Senior Secured Loans, Senior Secured Bonds and Corporate Mortgage Bonds.

to apply inverse Beta-distribution transformations. See Appendix A for more details.

4.4. PI Estimation

PIs have received little attention in the LGD literature, even though there is a high variability around the estimates of recovery rates produced by tables (Figure 1).

LossCalc provides an estimate of the PI (i.e. upper and lower bounds) on the recovery prediction. PIs provide a range around the prediction within which the actual value should fall a specified percentage of the time. This value can be used for portfolio risk assessments such as a portfolio Credit-VaR model. The *width* of this PI provides information about both the ability to realize the mean expectation of the prediction and the inherent uncertainty of the recovery rate process. It does *not* describe the precision of mean recovery forecast. Although we could describe this prediction accuracy by a *standard error of the estimate* measure, it would not be the appropriate value to enter into a portfolio Credit-VaR model.

A 90-per cent PI around the predicted mean value is the range (bounded by an upper bound and lower bound) in which the realized value will fall 90 per cent of the time. Therefore, we only expect the realized value to be below the lower bound or above the upper bound, 10 per cent of the time.

Although standard parametric tools can produce an (in-sample) estimate of the PIs, these estimates are relatively wide. We produce narrower PIs through a series of quantile regressions (Fitzenberger *et al.*, 2002) for the 5th, 25th, 50th, 75th and 95th percentiles of the expected recovery rate. Parenthetically, because many clients want this output, we also fit a Beta-distribution such that it exactly replicates the mean recovery estimate and best fits the quantile estimates. These quantile regressions use the same factors as described earlier.

4.5. Aligning LossCalc's Output with Banks' Default Definition

Bankers like to gain early *entrée* into a distressed borrower so that they can better mitigate potential losses. Loan covenants that trigger 'technical' defaults are a pervasive example of this. Banks are often successful in these efforts and we have seen institutions with as much as 50 per cent 'cure' rates. That is, the bank quickly redressed half of 'defaulted' loans with no economic loss.[20] And so, such a bank might broadly say that

[20] The lowest cure rate we have seen is 20 per cent.

its default rate was, say, 2 per cent with LGD of 15 per cent. But if half of these cure (i.e. 0 per cent LGD), then they might equivalently quote netted figures of 1 per cent default rate with 30 per cent LGD. Basel has stated that either view is acceptable so long as banks use the *same* default definition for *both* the default probability *and* the LGD.

In this same way, it is important to match LossCalc's LGD outputs to a bank's default definition. This discussion applies only to loans because bonds do not have this same sort of 'cure' process. We do this by first asking what the banks own *cure rate* is. Banks cure rates can differ and Moody's KMV cannot know ahead of time what this might be for any given institution although figures between 20 and 50 per cent are within our experience. LossCalc assumes a 100 per cent recovery on the cured portion and then applies its statistical forecast to the remaining (1 − *cure rate*) portion.

$$\text{(3)} \qquad \textit{Aligned recovery} = (\textit{cure rate} \cdot 100\%) + (1 - \textit{cure rate}) \cdot \textit{LossCalc recovery forecast}$$

5. Validation and Testing

The primary goals of validation and testing are to

1 determine how well a model performs;
2 ensure that a model has not been over-fit and that its performance is reliable and well understood;
3 confirm that the modelling *approach*, not just an individual model, is robust through time and credit cycles.

To validate the performance of LossCalc, we have used the approach adopted and refined by Moody's KMV termed *walk-forward* validation. It involves fitting a model on one set of data from one time period and testing it on the subsequent period. We then repeat this process, moving through time until we have tested the model on all periods up to the present. Thus, we never use data to test the model that we used to fit its parameters and so we achieve true out-of-sample and out-of-time testing with an efficient use of the data. We can also assess the behaviour of the modelling approach over various economic cycles. Walk-forward testing is a robust methodology that accomplishes the three goals set out. See Figure 17 for an illustration of the process.

Model validation is an essential step to credit model development. Tests must be performed in a rigorous and robust manner while guarding against unintended errors. For example, the same model may provide different performance results on different datasets, even when there is no specific selection bias in choosing the data. To facilitate comparison and avoid misleading results, we use the same dataset to evaluate LossCalc and competing models.

Sobehart *et al.* (2000a, b) describe the walk-forward methodology in detail. Appendix B of this document gives a brief overview of the approach.

5.1. Establishing a Benchmark for LossCalc

The standard practice in the industry is to estimate LGD by some historical average. There are many variations in the details of how these averages are constructed: long-term vs. moving window, by seniority class vs overall and dollar weighted vs. simple (event) weighted. We chose two of these methodologies as being both representative and broadly applied in practice. We then use these traditional approaches as benchmarks against which to measure the performance of our model.

5.1.1. Table-of-averages Method

For the majority of financial institutions, their model to estimate LGD is a look-up table. This often reflects expert opinion as to what LGD *ought* to be, but more commonly LGD look-up tables list historical average LGDs either from the institution's own experience or (not uncommonly) taken from rating agency recovery studies. A leading source of this type of agency recovery table is in Moody's annual default studies. With Moody's sizeable dataset, it represents a high-quality implementation of this 'classic look-up' approach. The Moody's published LGD averages have recoveries by debt type and seniority class and are updated annually.

5.1.2. Historical Average

Some institutions use a simple historical average recovery rate as their recovery estimate. Therefore, as a second hurdle (and we believe this represent a naive LGD model), we also tabulated the overall recovery rate across all instruments (the 'Historical Average').

5.2. The LossCalc Validation Tests

Because LossCalc produces an estimate of an *amount* (of recoveries), LossCalc seeks to fit a continuous variable. Thus, the diagnostics we use to evaluate its performance reflect this.

Because 1992 is the end of the first half of our dataset, we fit LossCalc to data from 1981 to 1992 and then forecast one year ahead (i.e. 1993). Following the walk-forward procedure, we constructed a validation *result* *set* containing 1,851 observations, representing 915 default events (some firms default more than once and are counted each time) from Moody's extensive database from 1993 to 2004. This result dataset was over 60 per cent of the total observations in the full dataset. It was a representative sampling of rated and unrated, public and private firms, in all industries and country/regions. See Appendix B for more details.

5.2.1. Prediction Error Rates

As a first measure of performance, we examined the error rate of the models. This is measured with an estimate of the MSE of each model. The MSE is calculated as

$$MSE = \frac{\sum (r_i - \hat{r}_i)^2}{n - 1} \tag{4}$$

where r_i and $\hat{r}_i$ are the actual and estimated recoveries, respectively, on security i. The variable, n, is the number of securities in the sample.

Models with lower MSE have smaller differences between the actual and predicted values and thus predict acutal recoveries more closely. Thus, better performing models in Figure 9 will have their symbols further to the left. LossCalc outperforms a Table of Averages by a large margin. For comparison, we have included the comparable accuracy measures from LossCalc v1. Technically, these version 1 statistics were tabulated on a

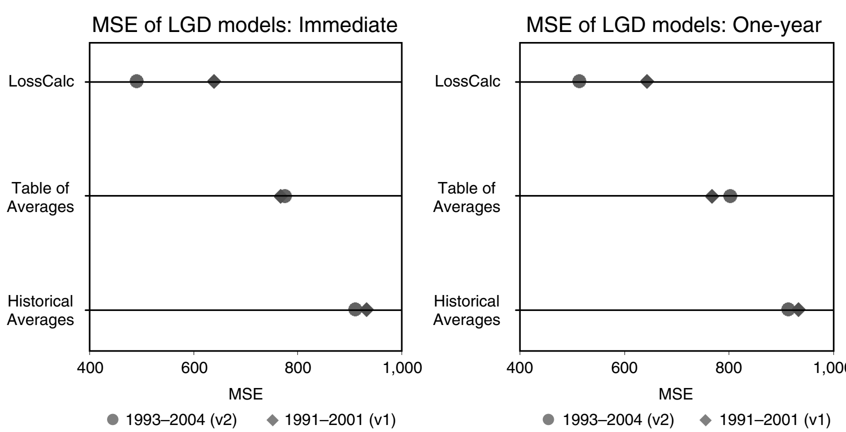

Figure 9: Mean Squared Error (MSE) of LossCalc Models and Other Alternative Models
Notes: This figure shows the out-of-sample MSE for LossCalc, the Table of Averages and the Historical Average. Note that better performance is towards the left-hand side in each panel, which is the opposite of Figure 10. It is clear that in both the immediate and one-year prediction, LossCalc has smaller error in comparison with the two alternative models.

different dataset from version 2 and so they are not directly comparable. But it is nevertheless gratifying to see that LossCalc v2 is significantly outpacing the accuracy of LossCalc v1, and it achieved this while performing against a more challenging (i.e. global) dataset (Table 2).

5.2.2. Correlation with Actual Recoveries

Next, we examined the correlation of the various models' predictions with the acutal loss experience. In this case, models with higher correlation exhibit preditions that are high when actual recoveries are high and low when actual recoveries are low more often than those that have lower correlation with the actual losses observed for defaulted securities.

Figure 10 shows the correlation of predictions vs. actuals for the three candidate models. This is also summarized in Table 3. The Historical Average, out-of-sample, actually exhibits a *negative* correlation with actual recovery experience. Therefore, for years with higher than average recoveries, it predicts lower than average recoveries and vice versa. This is because a moving average would be moving up from a previous year's economic boom just when it would do better to move down because of this year's economic bust.

Models with higher correlation have smaller differences between the actual and predicted values and thus predict acutal recoveries more closely. Better performing models shown in Figure 10 will have their symbols further to the right. LossCalc outperforms a Table of Averages by a large margin.

5.2.3. Relative Performance for Loans and Bonds

In Figure 11, we show the relative performance of LossCalc vs. the common practice of a Table of Averages with respect to three measures of relative performance: (i) MSE, (ii) correlation and (iii) area under a power curve. We show results for both loans and bonds.

Table 2: Mean Squared Error (MSE) of LGD Prediction Accuracy across Models and Horizons

	Immediate MSE		One-year MSE	
Out-of-sample	1991–2001	1993–2004	1991–2001	1993–2004
Historical Average	933.2	910.8	933.3	913.4
Table of Averages	767.3	775.8	767.3	802.4
LossCalc (v1 and v2)	639.1 (v1)	**490.6** (v2)	643.0 (v1)	**514.0** (v2)

Notes: Here, we list the specific out-of-sample MSE values illustrated in Figure 9.

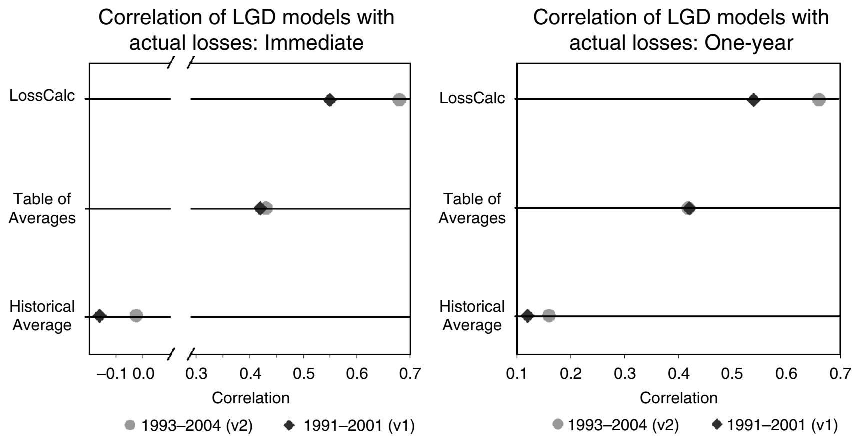

Figure 10: Correlation of LossCalc Models and Alternatives with Actual Recoveries

Notes: This figure shows the out-of-sample correlation for LossCalc, the Table of Averages and the Historical Average. Note that better performance is towards the right hand side of this graph, which is the opposite of Figure 9. It is clear that over both the immediate and one-year horizons, LossCalc has better correlation in comparison with the two alternative models.

By all three measures of performance, LossCalc increases predictive performance significantly. It is important to examine multiple dimensions of performance. Each has strengths and weaknesses:

1 *MSE* is among the most familiar test, but it lacks an intuitive sense of scale;
2 *Correlation* has an intuitive sense of scale, but it is sensitive to model construction;
3 *Area under a power curve* is the most robust of the three, but it may be less familiar to many analysts.

Although the improved performance is obvious, it is worth discussing the rather dramatic increase in performance in the case of correlation. This is somewhat counter-intuitive, because the increase in overall correlation is

Table 3: Correlation of LGD Prediction Accuracy across Models and Horizons

	Immediate correlation		One-year Correlation	
Out-of-sample	1991–2001	1993–2004	1991–2001	1993–2004
Historical Average	−0.13	−0.06	−0.13	−0.09
Table of Averages	0.42	0.43	0.42	0.42
LossCalc (v1 and v2)	0.55 (v1)	**0.68** (v2)	0.54	**0.66** (v2)

Notes: Listed here are the specific out-of-sample correlation values that we show in Figure 10.

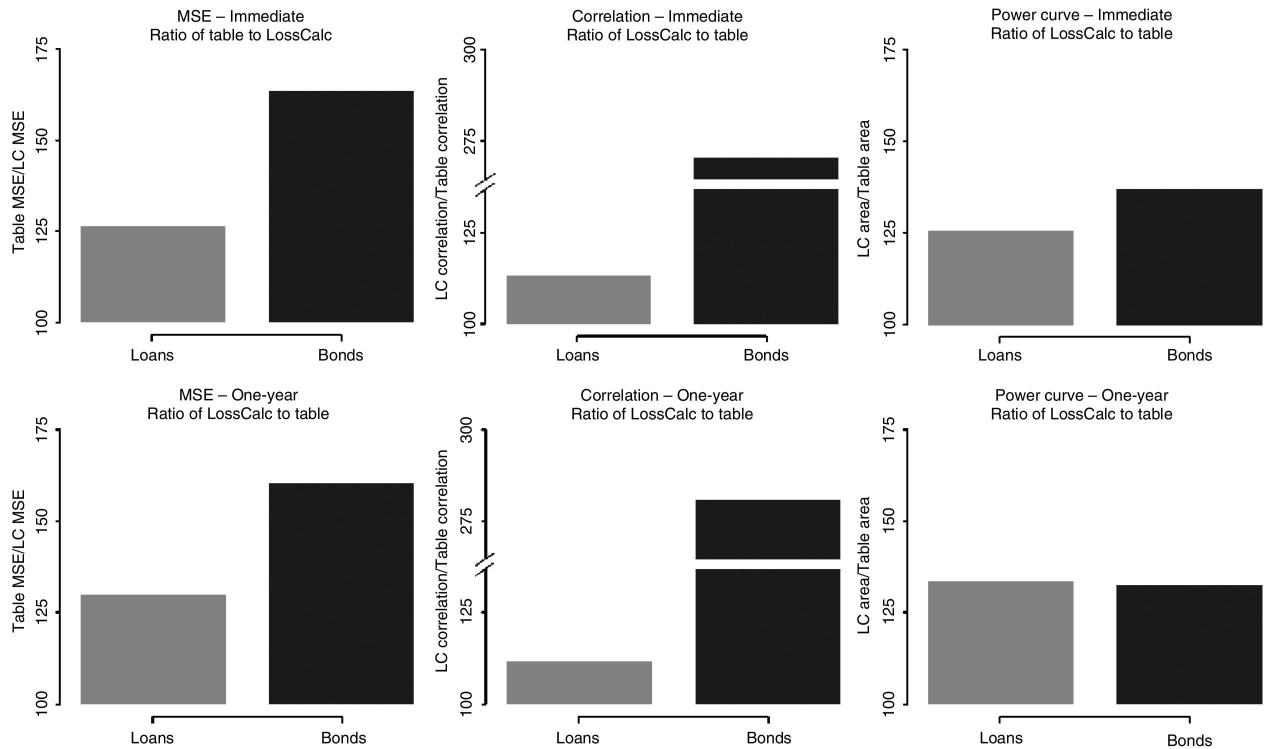

Figure 11: LossCalc Performance Increase Over a Table of Averages

Notes: This figure shows the relative increase in performance for both loans and bonds. We show three metrics of performance: (i) mean squared error (MSE), (ii) correlation and iii) area under a power curve. A bar of zero height indicates the two models performed equally well. Therefore, in every scenario, LossCalc outperforms a Table of Averages. As a specific example, we show the improvement in area under a power curve with the realized LGD between a table and LossCalc (at the one-year horizon) in the bottom right chart. Continuing this example, LossCalc is 33 per cent better than a Table of Averages with regard to loans.

in the order of 60 per cent whereas the increase in correlation for bonds achieves a relative improvement of about 275 per cent.

This is evident by how the two approaches to recovery prediction differ. The Table-of-Averages mechanism for predicting recoveries is segregation by seniority class. Across time, the only variability in the Table of Averages comes from the changes in tabulated averages from one year to the next. Thus, the Table of Averages focuses primarily on the *between-group* (debt type and seniority class) variability rather than the *within-group* (firm, industry and macroeconomic) variability in recoveries. In contrast, LossCalc uses additional information beyond simple conditioning, which allows it to incorporate both within- and between-group variability more completely. This is a fundamental weakness of *any* look-up table approach.

5.2.4. Prediction of Larger than Expected Losses

When assessing risks, it is generally true that under-estimating the severity of losses can expose an institution to more business risk than merely under-estimating positive results. To this asymmetry in risk preferences, we developed our final test to evaluate each model's ability to predict cases in which actual losses were *greater* than historical expectations.

For this test, we

1 use the most recent information available up to the time of a default, and we labelled each record to reflect if the actual loss experienced was greater or less than the historical mean loss for all instruments to date (i.e. the Historical Average first referenced in section 5.1.1);
2 ordered all out-of-sample and out-of-time predictions for each model from largest predicted loss to smallest predicted loss;
3 calculated the percentage of larger than average losses each model captured in its ordering using standard power tests.

This allowed us to convert the model performance to a binary measure which in turn allowed us to use power curves and power statistics to measure performance.

If a model was powerful at predicting larger than average losses, the largest loss predictions would be associated with the actual above-average losses and the lowest loss predictions to be associated with below-average losses. (On a power curve, this would result in the curve for a good model being bowed out towards the Northwestern corner of the chart. The random model would be a 45° line showing no difference in association between high- and low-ranked obligations.) We find that this provides a valuable non-parametric evaluation of LGD model performance.

The results of this analysis are shown in the panels of Figure 12. The figure shows both the power curve at left and the area under the curves at

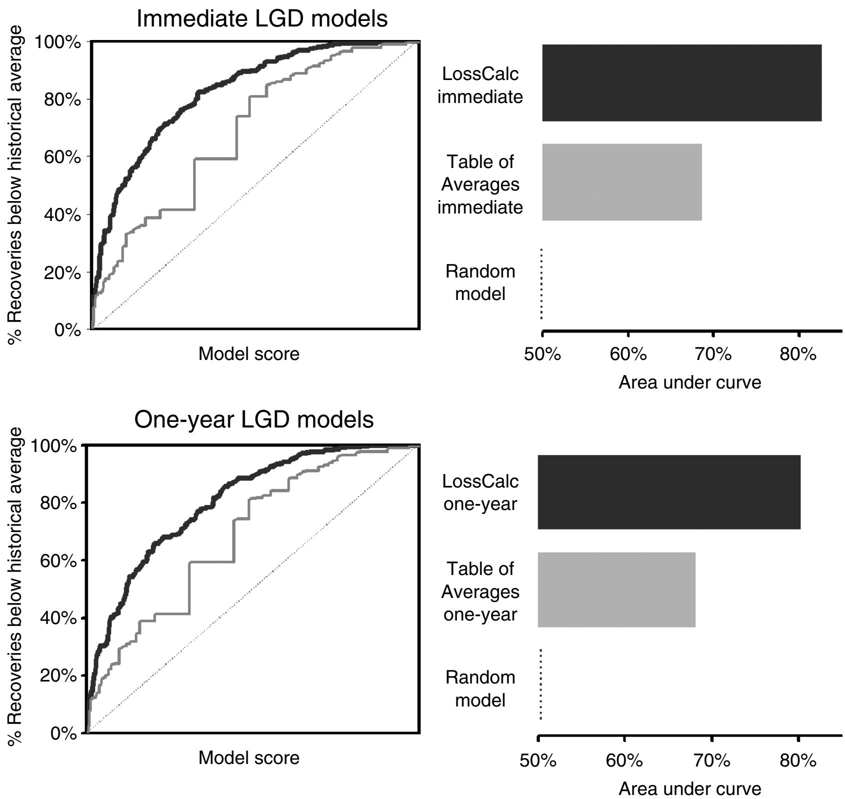

Figure 12: Power in Predicting Lower than Expected LGD

Notes: This figure shows the predictive power of LossCalc relative to the Table of Averages. It is clear that LossCalc has significantly greater power at both the immediate and one-year horizons, (i.e. its power curve is more towards the upper left corner and the area under the curve is greater).

right. The larger this area, the more accurate the model is. In this case, for a perfect model, this area would be 100 per cent.

Figure 12 shows that both the Table of Averages and LossCalc models perform better than random at differentiating high- and low-loss events, but that the LossCalc models outperform the Table of Averages by a considerable margin.[21] This relationship persists over both the immediate and one-year horizons. The comparison of areas under the curves confirms this observation.

[21] Note that 'plateaus' can occur in the power curves indicating that multiple instruments received the same LGD prediction. This mostly affects the Table of Averages model where all instruments of a particular seniority class (for a particular year) receive the identical LGD prediction. In principle, if one was a 'bad' and the other was not, the ordering could influence the power curve; although testing indicates that condition, the differences would not change the overall conclusions.

5.2.5. Reliability and Width of PIs

To test the quality of the PIs produced by the models, we examined two dimensions: width and reliability. The average *width* of a PI provides information about the precision and efficiency of the estimate. In general, a narrow PI is good because it gives less uncertainty to potential LGD realizations and allows an institution to assess losses or allocating capital more efficiently. However, a narrow model PI may not reflect the actual observed PI of the different classes if debt.

To examine these two issues, we generated PIs for each model, calibrated to in-sample data and tested on out-of-sample data. We show the average widths of these PIs in the left-hand panels of Figure 13. Then, we tested out-of-sample and out-of-time the number of cases in which the actual observed losses exceeded the predicted interval.

We examined several methods for PI prediction. Some of these required the use of actual prediction errors from previous periods. These tests were on about 500 observations for the one-year horizon and close to 600 for the immediate tests.

Figure 13 shows that LossCalc's PIs are more precise (narrower) than both the parametric (standard deviation) and quantile estimates from the table. However, uneven coverage percentages make the two Table-of-Averages PIs somewhat uncertain (i.e. the grey bars in the right-hand panels do not achieve the 10 per cent targetted coverage).

For example, for the immediate horizon version of LossCalc, the actual out-of-sample and out-of-time coverage of the historical table was higher than LossCalc signalling not only a more precise estimation, but also a more efficient one. However, the one-year version shows a slightly higher out-of-sample and out-of-time coverage than the historical table, indicating that the width of the PI could probably have been made tighter. Similarly, the width of the parametric PI for the table is likely optimistically narrow because of the higher than expected number of cases outside the PI. Unfortunately, there is no way to anticipate such variances from the expected PI *a priori*.

5.2.6. Miscellaneous Other Tests

We did many tests on our LGD model. We show a sampling of some of these tests in Figure 14. All of these tests are performed on our global dataset, out-of-sample, as of September 2004. The top left panel shows that LossCalc is nearly as accurate for private firms as for public firms even though one of the model inputs is our public-firm EDF technology (i.e. D2Ds).

The top right panel shows that LossCalc can predict LGD for new situations not included in the dataset. We did this by testing the model on

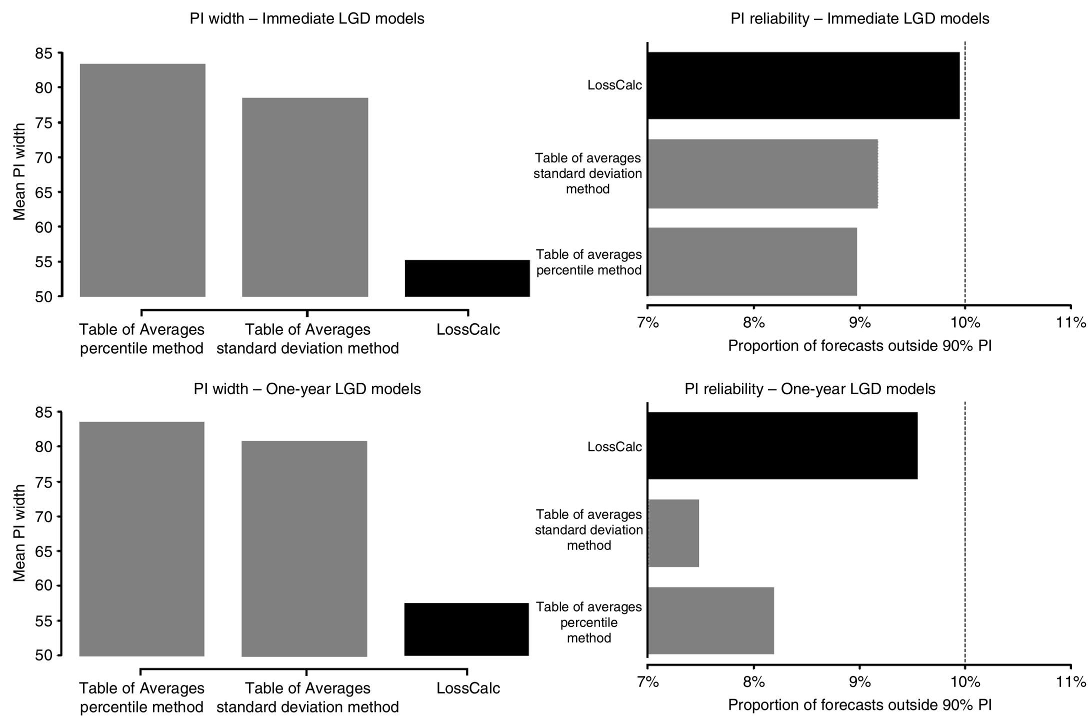

Figure 13: The Width and Reliability of Model Prediction Intervals

Notes: These panels show the prediction interval (PI) widths (left two panels) and reliabilities (right two panels) for both Immediate predictions (top two panels) and for defaults that may occur one year from now (bottom two panels). In all cases, we compare the LossCalc performance relative to an alternative Table of Averages model. We find that LossCalc's PIs are both narrower and more reliable than other alternatives.

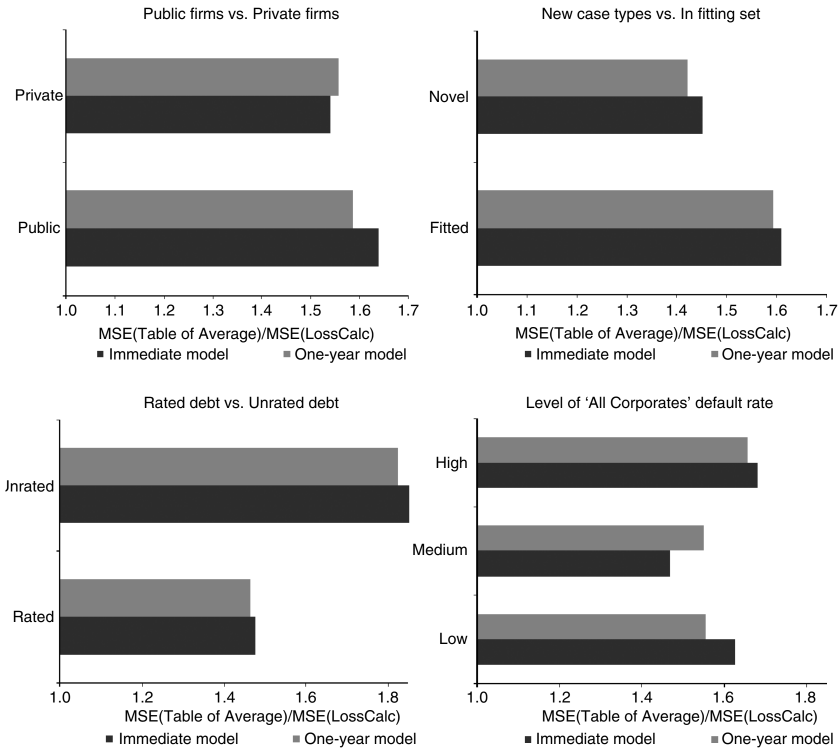

Figure 14: LossCalc has Undergone Extensive Testing

Notes: These figures examine LossCalc's predictive accuracy for four different types of divisions across our dataset. By construction, if these bars have any length to the right of the '1.0' point, then LossCalc outperforms a Table of Averages. LossCalc performs well for both public and private firms with only slightly diminished accuracy for private firms. LossCalc's performance was only somewhat reduced for cases (Industry/Seniority combinations) that were unrepresented in its fitting set. LossCalc performed relatively better during unusual time or odd cases. Shown here are two examples of this: (i) the case of unrated debt (bottom left panel) and (ii) the case of either above- or below-average default rates (bottom right panel).

industry and seniority class that were not present within the 'fitting set' of data.

The bottom two panels show that for unrated instruments (on the left) and for periods of above below-average default rate environments, LossCalc performs better than tables. These two reflect a common theme that our model performs relatively better in cases that are not 'typical'.

6. The Dataset

The dataset used to develop and test LossCalc is Moody's Investors Service proprietary default and recovery database (used in Moody's annual default studies) plus other types of data such as financial statements, credit indices and firm, industry and credit market information made available with the 2002 merger with KMV.

The default data are secondary market pricing of defaulted debt as quoted one month after the date of default. Importantly, we use debt-issue-specific market quotes that are *not* 'matrix' prices.

6.1. Historical Time Period Analysed

LossCalc uses recovery observations since January 1981 so that it covers at least two full economic cycles. We did this because our research found that the credit cycle was a strong determiner of recoveries. We also use financial statement data that only became reliably available from Compustat and WorldScope in 1981.

6.2. Scope of Geographic Coverage and Legal Domain

Bankruptcy laws vary across legal domains: for example, UK law tends to protect creditors more diligently, whereas French law contemplates the greater social good at the expense of lenders. Some domains allow *creditors* to file a petition for insolvency. There are also differences in the strength of security (Bartlett, 1999; West and de Bodard, 2000a, b, c). In order to account for major differences, we have adapted the LossCalc framework so that it makes specific country/region adjustments where necessary (Figure 15).

6.3. Scope of Firm Types and Instrument Categories

Our dataset includes three broad debt instrument types: (i) bank loans, (ii) public bonds and (iii) preferred stock. We have organized loans broadly into two seniority classes: 'senior secured', which are the more numerous, and 'senior unsecured'. Public bonds are subdivided into seven seniority classes: (i) IRBs, (ii) Corporate Mortgage Bonds, (iii) senior secured, (iv) senior unsecured, (v) senior subordinated, (vi) subordinated and 7) junior subordinated.[22]

For medium-term note programs (MTNs), which are characterized by a large number of small issues, we consolidate the many individual issues that an obligor may have into a single recovery observation per default event. Otherwise, we felt that we would over-weight MTNs. The recovery

[22] See Mann (1997) for a discussion of the common use and legal treatment of differing seniority classes and Stumpp *et al.* (1997) for a detailing of bank loan structure and collateral.

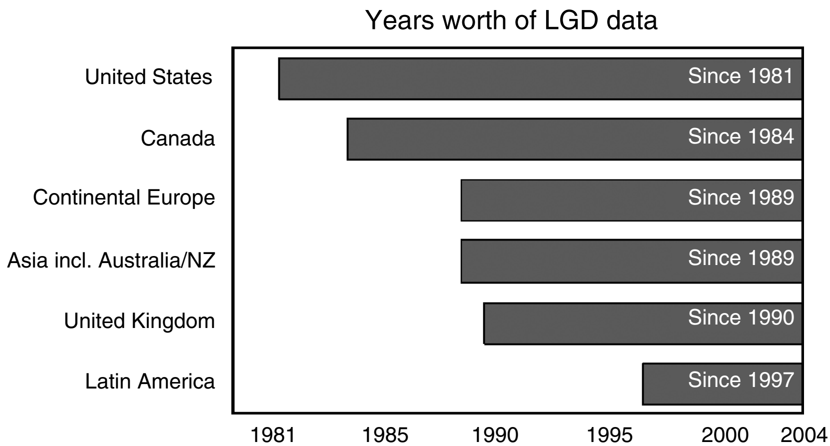

Figure 15: Depth of LGD Data by Country/Region
Notes: The LGD dataset for LossCalc extends across countries and regions. Although data collection has been in place the longest in North America, all regions have at least the seven-year minimum dataset as prescribed by the Advanced IRB requirements.

rate realized for this proxy observation is the simple average of all the individual issuances. It happens that there is never large variability in the recovery rates across issues within a single MTN program.

7. A Case Study

France Télécom is often thought of as a utility-like firm who expanded aggressively into wireless communication through a series of acquisitions. These acquisitions included British NTL (£8B; July 1999), CTE Salvador (US$275M; September 1999), Jordan Telecom (US$500M; January 2000), Germany's MobilCom (€3.7B; March 2000), British Orange (£26.5B; May 2000) as well as acquiring Telecom Argentina through a consortium. By May 2002, NTL had filed for bankruptcy, while that summer MobilCom was near bankruptcy. By the end of 2002, France Télécom's €70B of total debt was three times the market capitalization of the company.

Although France Télécom was half state owned, the French government did not move to rescue MobilCom. Europe had been seeking to privatize state-run firms. There were also the two recent American Telecom examples of WorldCom's bankruptcy filing in January 2002 and Global Crossing filing in July 2002.

Figure 16 illustrates a dramatic example of LGD zigzagging across time. More commonly, shifts in the industry/region index are what cause LossCalc changes in LGD over time both up and down. What is most

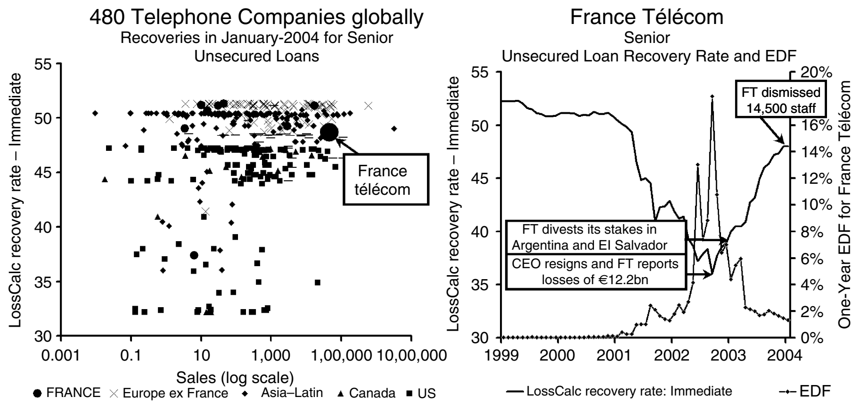

Figure 16: Telecom Recoveries and Specifically France Télécom

Notes: Shown here is the Telecom industry both at a point in time, left panel, and one of its members across time, right panel. The left panel shows recoveries (for Senior Unsecured Loans) during January 2004, for 480 firms in all country/regions. Note that even this homogeneous group can have a wide range of possible recovery expectations with the top recoveries nearly double the bottom recoveries. The right panel shows one firm, France Télécom, over a five-year period from January 1999 to January 2004. Its recovery dramatically varies during this period. This was largely driven by the extreme swing changes in default expectation of as evidenced by the Moody's KMV expected default frequency (EDF).

clear is that there is no economic reason to believe that a security's LGD should be static across time. Given our knowledge of what drives differences in LGD, it becomes only a question of how best to use that knowledge to understand, moderate and manage a portfolio's risk profile.

8. Conclusion

In this report, we described the research carried out to develop LossCalc v2, Moody's KMV LGD model. LossCalc is a multi-factor statistical model developed using a database of 3,026 defaulted instruments. It produces LGD estimates for loans, bonds and preferred stock. LossCalc assesses information on five levels of analysis, including the characteristics of collateral, the instrument, the firm, its industry and the macroeconomy/geography.

The issue of prediction horizon for LGD is one that has obtained little attention because of the largely static nature of the dominant historical average approach. This implicitly ignores the effects of the credit cycle and other time-varying environmental factors. LossCalc, by its dynamic nature, allows for a much more exact specification of LGD horizon and produces estimates on both 'immediate' and 'one-year' horizons.

We find that the measures of D2D (compiled both at the firm level and the industry level) are predictive of security-level LGD. We found that the

MBBI became insignificant when industry D2D entered the model, suggesting that defaulted debt prices truly reflect LGD rather than being buffeted by the vicissitudes of supply and demand. We found consistency in factor application and predictive power of LGD across our global dataset. Given this good statistical fit, we hypothesized that LGD is a two-step process of (i) defaulted firm economic value available to claimants and (ii) domain-dependent rules for dividing this value among claimants. We made the negative finding that, although the 12-month trailing default rate is statistically predictive of security-level LGDs, it is not economically material. This is in contrast to studies that found it to be material at the annual aggregation level.

We conducted extensive out-of-sample and out-of-time validation of this model to confirm its performance predicting LGD compared with alternative approaches. The results of this benchmarking show that the model performs better than common alternatives such as an overall historical average of LGD or a Table-of-averages LGDs in categories. LossCalc's performance is superior in both out-of-sample out-of-time prediction error and correlation of the predictions with actual recovery experience. The model was also better at identifying recoveries that were lower than historical average methods and it has fewer large errors.

LossCalc represents a robust and validated global model of LGD for the debt types it covers. We believe that this is a productive step forward in answering the call for rigour that the Bank for International Settlements has outlined in the recently proposed Basel Capital Accord.

REFERENCES

E. I. Altman – V. M. Kishmore (1996), "Almost Everything You Wanted to Know about Recoveries on Defaulted Bonds", *Financial Analysis Journal*, November/December, pp. 56–62.

E. I. Altman – A. Gande – A. Saunders (2004), "Informational Efficiency of Loans Versus Bonds: Evidence from Secondary Market Prices", Working Paper, New York University and Vanderbilt University.

F. Bartlett (1999), "Regimes, Recoveries and Loan Ratings: The Importance of Insolvency Legislation", Working Paper, Fitch/IBCA – Loan Products.

Basel Committee on Banking Supervision (2001), "The New Basel Capital Accord", Working Paper, Bank for International Settlements.

S. Borenstein – N. L. Rose (1995), "Do Airlines in Chapter 11 Harm Their Rivals?: Bankruptcy and Pricing Behavior in U.S. Airline Markets", Working Paper, National Bureau of Economic Research.

P. J. Crosbie – J. R. Bohn (2003), "Modeling Default Risk", Working Paper, Moody's KMV.

S. A. Davydenko – J. R. Franks (2004), "Do Bankruptcy Codes Matter? A Study of Defaults in France, Germany, and the UK", Working Paper, London Business School.

V. Dhar – R. Stein (1998), "Finding Robust and Usable Models with Data Mining: Examples from Finance", *PCAI*, 19(5), pp. 17–23.

A. C. Eberhart – R. J. Sweeney (1992), "Does the Bond Market Predict Bankruptcy Settlements?" *Journal of Finance*, 48(3), pp. 943–80.

T. C. G. Fisher – J. Martel (2003), "The Firm's Reorganization Decision: Empirical Evidence from Canada", Working Paper, Wilfrid Laurier University and Université de Cergy-Pontoise.

B. Fitzenberger – R. Koenker – J. A. F. Machado (2002), *Economic Applications of Quantile Regression*, New York: Physica-Verlag Heidelberg.

M. Gordy – D. Jones (2002), "Capital Allocation for Securitizations with Uncertainty in Loss Prioritization", Working Paper, Federal Reserve Board.

R. Guha (2002), "Recovery of Face Value at Default: Theory and Empirical Evidence", Working Paper, London Business School.

G. M. Gupton – R. M. Stein (2002), "LossCalc^{TM}: Model for Predicting Loss Given Default (LGD)", Moody's Investors Service.

D. T. Hamilton (1999), "Debt Recoveries for Corporate Bankruptcies", Moody's Investors Service.

D. T. Hamilton – A. Berthault (2000), "The Investment Performance of Bankrupt Corporate Debt Obligations: Moody's Bankrupt Bond Index 2000", Moody's Investors Service.

D. T. Hamilton – G. M. Gupton – A. Berhault (2001), "Default and Recovery Rates of Corporate Bond Issuers: 2000", Moody's Risk Management.

J. J. Heckman (1979), "Sample Bias as a Specification Error", *Econometrica*, 47(1), pp. 153–61.

V. Ivanova (2004), "LGD-rating for a Portfolio of Retail Loans", Working Paper, University of Oxford.

I. Izvorski (1997), "Recovery Ratios and Survival Times for Corporate Bonds", Working Paper, IMF, Washington DC.

S. D. Longhofer – C. T. Carlstrom (1995), "Absolute Priority Rule Violations in Bankruptcy", *Economic Review*, 31(4), pp. 21–30.

R. J. Mann (1997), "Explaining the Patterns of Secured Credit", *Harvard Law Review*, 110(3), pp. 625–83.

Y. M. Mensah (1984), "An Examination of the Stationarity of Multivariate Bankruptcy Prediction Models: A Methodological Study", *Journal of Accounting Research*, 22(1), pp. 380–95.

M. Onorota – E. I. Altman (2003), "An Integrated Pricing Model for Defaultable Loans and Bonds", Working Paper, City University London and New York University.

M. H. Pesaran – T. Schuermann – B. J. Treutler – S. M. Weiner (2004), "Macroeconomic Dynamics and Credit Risk: A Global Perspective", Working Paper, University of Cambridge.

M. Schmit – J. Stuyck (2002), "Recovery Rates in the Leasing Industry", Working Paper, Leaseurope.

P. J. Schönbucher (2003), *Credit Derivatives Pricing Models*, New York: John Wiley & Sons.

M. Singh (2003), "Recovery Rates from Distressed Debt – Empirical Evidence from Chapter 11 Filings, International Litigation, and Recent Sovereign Debt Restructurings", Working Paper, IMF.

J. Sobehart – S. Keenan – R. Stein (2000a), "Validation Methodologies for Default Risk Models", *Credit*, 1(4), pp. 51–6.

J. R. Sobehart – S. C. Keenan – R. M. Stein (2000b), "Benchmarking Quantitative Default Risk Models: A Validation Methodology", Moody's Risk Management.

P. Stumpp – T. Marshela – M. Mulvaney – M. Hilderman (1997), "A Sense of Security: Moody's Approach to Evaluating Bank Loan Structure and Collateral", Moody's Investors Service.

D. Tasche (2004), "The Single Risk Factor Approach to Capital Charges in Case of Correlated Loss Given Default Rates", Working Paper, Bundesbank.

H. S. Wagner III (1996), "The Pricing of Bonds in Bankruptcy and Financial Restructuring", *The Journal of Fixed Income*, June, pp. 40–7.

D. J. Ward – G. L. Griepentrog (1993), "Risk and Return in Defaulted Bonds", *Financial Analysts Journal*, May/June, pp. 61–5.

M. West – E. de Bodard (2000a), "Bankruptcy and Ratings: A Leveraged Finance Approach for Europe", Moody's Investors Service.

M. WEST – E. DE BODARD (2000b), "Bankruptcy and Ratings: A Leveraged Finance Approach for Europe – Part II: France", Moody's Investors Service.
M. WEST – E. DE BODARD (2000c), "Bankruptcy and Ratings: A Leveraged Finance Approach for Europe – Part III: Germany", Moody's Investors Service.
P. XU (2004), "Bankruptcy Resolution in Japan: Corporate Reorganization vs. Civil Rehabilitation", Working Paper, University and RIETI.

Appendix A: Beta-transformation to Normalize Loss Data

To create an approximately normally distributed dependent variable from the raw observations of recovery, we first confirmed that defaulted debt valuations were approximately Beta-distributed. There is no theoretical reason that this is the 'correct' shape of the defaulted debt prices, but previous studies have concluded that its characteristics make the Beta a reasonable description of the empirical shape.

Beta-distributions are described in this case by an upper and lower bound and by two shape parameters, α and β. Most commonly, it is naturally bounded between zero and one; its mean can be any value strictly within its range. For LossCalc, we generalize this distribution to accommodate the rare, but non-trivial cases where recoveries can range somewhat above 1.0. The conversion of the Beta-distributed recovery values to a more normally distribution-dependent variable is explicitly defined as follows:

$$(5) \quad \textit{Dependent variable} = Y_i = N^{-1}[\textit{Betadist}(\textit{RecovRt}_i, \alpha_d, \beta_d, \textit{Min}, \textit{Max}_d)]$$

where $N^{-1} \equiv$ the inverse of the Normal cumulative distribution, $RecovRt = \min\ (Max - \varepsilon$, observed recovery rate), $\varepsilon =$ some small value, $\alpha_d =$ the Beta-distribution's *center* parameter, $\beta_d =$ the Beta-distribution's *shape* parameter, $Min =$ set to zero in all cases, $Max_d =$ set to 1.1 for bonds, but otherwise is 1.0 and $d = \{$loans, bonds, preferred stock$\}$

We use the sub-notation 'd' to emphasize that LossCalc fits each debt type to its own distribution.

Thus, much of the distributions of our three separate asset classes can be captured by specifying only two shape parameter values: the α and the β of each Beta-distribution. There are various ways of fitting the distribution parameters. It is also possible, through algebraic manipulation, to specify the Beta-distribution that simply matches the mean and standard deviation, which are functions of the shape and boundary parameters.

Mathematically, a Beta-distribution is a function of Gamma distributions. With the lower bound, *Min*, fixed at zero, the distribution is as follows:

$$(6)\quad \beta(x,\alpha,\beta,Min=0,Max)=\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\left(\frac{x}{Max}\right)^{\alpha-1}\left(1-\frac{x}{Max}\right)^{\beta-1}\left(\frac{1}{Max}\right)$$

The shape parameters can be derived in various ways. For example, (7) gives them in terms of population mean and standard deviation.

$$(7)\qquad \alpha=\frac{\mu}{Max}\left[\frac{\mu\cdot(Max-\mu)}{Max\cdot\sigma^2}-1\right]\quad\text{and}\quad\beta=\left(\frac{Max}{\mu}-1\right)$$

Conversely, given Beta-distribution parameters, it is straightforward to calculate the mean and standard deviation.

$$(8)\qquad \mu=Max\cdot\left(\frac{\alpha}{\alpha+\beta}\right)\quad\text{and}\quad\sigma=Max\cdot\sqrt{\frac{\alpha\cdot\beta}{(\alpha+\beta)^2+(1+\alpha+\beta)}}$$

Appendix B: An Overview of the Validation Approach

To validate the performance of LossCalc, we have used the approach adopted and refined by Moody's KMV and used to validate the RiskCalc models of default prediction. The approach, termed *walk-forward* validation, is a robust means for ensuring that

1 models have not been 'over-fit' to the data;
2 future performance can be well understood; and
3 the modelling approach, as well as any individual model produced by it, is robust through time and credit cycles.

We give only a brief overview of the methodology here; a fuller description is detailed in Sobehart *et al*. (2000a).[23]

B.1. Controlling for 'Over-Fitting' Risk: Walk-forward Testing

In order to avoid embedding unwanted sample dependency, we have found it useful to develop and validate models using some type of out-of-sample, out-of-time and out-of-universe testing approach on panel or cross-sectional datasets.[24] However, such an approach can generate false

[23] Much of what follows was adapted from Sobehart *et al*. (2000a).
[24] A panel dataset contains observations over time on many individuals. A cross-sectional dataset contains one observation on many individuals.

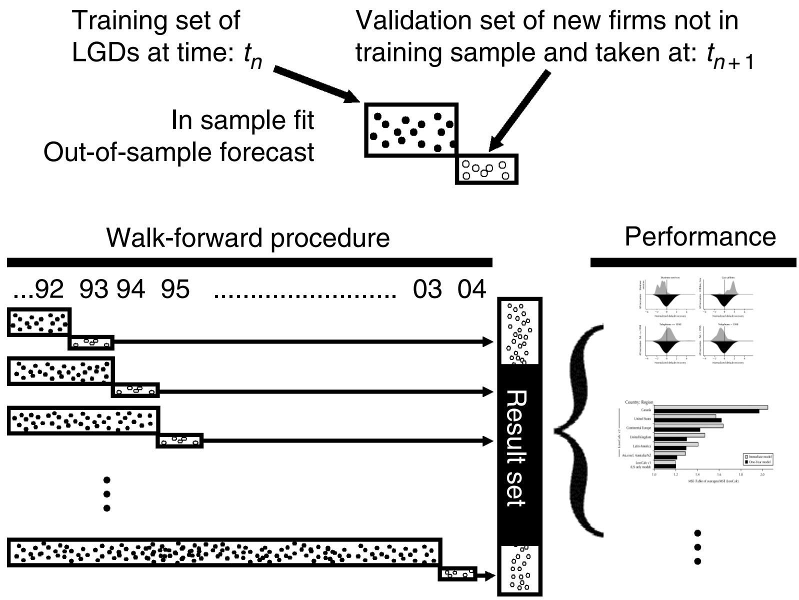

Figure 17: Validation Methodology: End-to-end

Notes: We fit a model using a sample of historical recovery data and test the model using data on new recoveries oneyear later (upper portion of exhibit). Dark circles represent data for fitting the model and white circles represent validation data. We perform 'walk-forward testing' (bottom left) by fitting the parameters of a model using data through a particular year and testing on data from the following year, and then stepping the whole process forward one year. We then resample the aggregated Result Set (lower left) to calculate particular statistics of interest.

impressions about the reliability of a model if performed incorrectly. 'Hold-out' testing can sometimes miss important model problems, particularly when processes vary over time, as credit risk does (Mensah, 1984). Dhar and Stein (1997) suggest a framework for framing these issues and providing a more detailed discussion and some examples from finance.

We designed our testing approach to test models in a realistic setting that emulates closely the practical use of these models. The trading model literature often refers to this procedure as 'walk-forward' testing.

The walk-forward procedure works as follows:

1 Select a year, for example, 1992.
2 Fit the model using all the data available on or before the selected year.
3 Once the model's form and parameters are established for the selected period, generate the model outputs for all of the instruments available during the following year (in this example 1993).

Note that these are out-of-time and generally out-of-sample (there are rare cases of firms defaulting more than once).

4 Save the prediction as part of a *result set*.
5 Now move the window up one year (e.g. 'to 1993') so that all the data through that year can be used for fitting and the data for the following year can be used for testing.
6 Repeat steps (2) to (5) adding the new predictions to the result set.

Collecting all the out-of-sample and out-of-time model predictions produces a set of model performances. We use this *result set* to rigorously validate the performance of the model in more detail.

Note that this approach closely simulates how Moody's KMV and our clients actually use LossCalc in practice. Each year, the model is refit and used to predict recoveries one year hence. We outline the walk-forward validation process in Figure 17.

Note that this approach has two significant benefits. First, it allows us to get a realistic view of how a particular model would perform over time. Second, it allows us to leverage to a higher degree the availability of data for validating models. Unless otherwise noted, all results presented in this study are from this type of out-of-sample and out-of-time walk-forward testing.

Non-technical Summary

We describe LGD research findings that lead to version 2.0 of LossCalc™ by Moody's KMV. LossCalc forecasts LGD using multiple linear regression that applies predictive factors at all relevant information levels: collateral, instrument, firm, industry, country and the macroeconomy. Our clients find it to be relevant for the Basel II Advanced IRB approach for five reasons: (i) it is built on a large representative dataset, (ii) the dataset spans 23 years (and seven years in all global regions), (iii) its LGD estimates are time varying, (iv) it is validated out-of-sample and out-of-time and (v) it is documented. Clients' internal LGD requirements include five additional issues: (i) responsiveness to the many known drivers of LGD, (ii) predictive accuracy, (iii) assured consistency across geographies, instruments and sectors, (iv) predictions at relevant risk horizons and (v) reporting of an instrument-level PI for each LGD forecast. By way of summary, we briefly touch of these ten elements:

A Large Representative Dataset

We fit our models on 3,026 observations of instrument-level LGD sourced globally. We accurately distinguish six country/regions: Asia, Canada, Europe, Latin America, the United States and the United Kingdom. The dataset includes 1,424 defaulted public and private firms in all industries. Our data sources are principally Moody's Investors Service and KMV Corporation (now Moody's KMV).

A Dataset Spanning 23 Years

Our dataset extends from 1981 to 2004. Because this covers several expansions and contractions in the economy, we well estimate the effects of economic cycles. LossCalc gives LGD forecasts distinguished by six regions of the world and so it was important to have a significant span of data in each. Of these six, Latin America is our smallest subset where we have seven years worth of data. Basels' Advanced IRB approach demands a minimum of seven years.

LGD Estimates are Time-varying

We find that D2D (compiled at both the firm level and the industry level) are predictive of security-level LGD changes across time. Intuitively, firms with a high D2D at default tend to have an intact franchise value and reorganize with minimal LGD. Similarly, 'distressed' industries (low D2D) tend not to have the capacity or inclination to buy-up and redeploy defaulted assets.

Validation Out-of-sample and Out-of-time

We apply a *walk-forward* procedure of out-of-sample and out-of-time validation. This is the same validation methodology used in other models, which our clients are implementing as Basel solutions. All reported results are out-of-sample. We report no in-sample results.

Documented

In addition to a 44-page methodology document (with more detailed disclosure and validation), we have a package of materials for clients who are implementing the methodology or talking with their regulators.

Responsiveness to the Many Known Drivers of LGD

We use nine explanatory factors organized into five broad groups: (i) collateral, (ii) debt type/seniority class, (iii) firm status, (iv) Industry and

(v) macroeconomic/geographic. The model's predictive power (explanation of variance) spans all five groups.

Predictive Accuracy

The most prevalent method for arriving at an LGD estimate is a simple look-up value in some table. The inherent problem (unless the look-up table is huge) is that there is a wide variability of realized recovery values within each cell of any table. Table-driven LGD models also lack (i) a time-varying factor and (ii) any means of discriminating differences in recovery *within* any given 'cell' of the look-up table.

Our model is easily a more accurate predictor of LGD. We measure, better, by reporting: mean squared error, correlation and non-parametric rank ordering (as graphed by a power curve).

Assured Consistency across Geographies, Instruments and Sectors

Ours is an international model with LGD observations sourced globally. Likewise, our predictive factors are sourced consistently but separately within each region. In addition to their predictive power, we find that our factors (especially the D2D factors) are consistent in their magnitude and interpretation across geographies.

Although legal differences in bankruptcy play some part in recovery, we find that *fundamental economics drives firm values (and hence recoveries) in a predictable way* regardless of the particular legal jurisdiction.

Predictions at Relevant Risk Horizons

Our model forecasts the LGD for defaults occurring immediately and one year from the time of evaluation. LossCalc, by its dynamic nature, allows for a much more exact specification of LGD horizon: recovery if default occurs tomorrow vs. recovery if default occurs in one year.

Reporting of an Instrument-level PI for Each LGD Forecast

PIs have received little attention in the LGD literature, even though there is a high variability around the estimates of recovery rates produced by tables. We provide the PI (i.e. upper and lower bounds) on our recovery prediction. We also output a full distribution with its dynamic (non-Normal) shape. Clients can use these for portfolio risk assessments such as a portfolio Credit-VaR model.