

Modeling Ultimate Loss Given Default on Corporate Debt

MICHAEL JACOBS AND AHMET K. KARAGOZOGLU

MICHAEL JACOBS, JR., is senior financial economist in the Credit Risk Modeling Division of the Office of the Comptroller of the Currency, U.S. Department of the Treasury in Washington, DC. michael.jacobs@occ.treas.gov

AHMET K. KARAGOZOGLU is an associate professor at the Zarb School of Business at Hofstra University in Hempstead, NY. finakk@hofstra.edu

Loss given default (LGD), the loss severity on defaulted debt obligations, is a critical component of risk management, pricing, and portfolio models of credit.¹ LGD is among the three primary determinants of credit risk, the other two being probability of default (PD) and exposure at default (EAD). However, LGD has not been as extensively studied and is considered a much greater modeling challenge than PD. Traditional credit models such as PD have focused on systematic components of credit risk that attract risk premiums. Unlike PD, determinants of LGD have typically been ascribed to idiosyncratic, borrower-specific factors. However, there is now an ongoing debate about whether the risk premium on defaulted debt should reflect systematic risk and, in particular, whether the intuition that LGDs would rise in worse states of the world is correct; and how this could be refuted empirically given limited and noisy data. This heightened focus on LGD has been motivated by the large number of defaults and nearly simultaneous decline in recovery values observed through the last credit cycle as well as the current credit crisis, regulatory developments such as Basel II (Basel Committee of Banking Supervision [2005]), and the continued growth in credit markets. However, obstacles to better understanding and predicting LGD include a dearth of relevant data and the lack of a

coherent theoretical underpinning, a continuing challenge to researchers.

This study contributes to the research on LGD on several fronts. The review of literature considers recent contributions and combines many elements into a unified empirical framework. The methodology builds an internally consistent model of LGD that corresponds to a priori expectations and empirical findings, which is amenable to rigorous validation and represents an advance in econometric methodology. In particular, estimation of a two-equation system models LGD simultaneously at the obligor and instrument levels, using an extensive sample of corporate bond and loan defaults. In addition to answering the many academic questions regarding LGD, we provide a practical tool for risk managers, traders, and regulators in the field of credit. For example, these players in the credit markets can use our model to forecast ultimate LGD, which can serve as input into credit models for value at risk (VaR), distressed debt pricing, or regulatory capital.

LGD can be defined variously depending upon the institutional setting, the type of instrument (e.g., traded bonds or bank loans), or the credit risk model (e.g., pricing debt instruments subject to the risk of default, expected loss calculation or credit risk capital). The ultimate LGD represents eventual discounted loss per dollar of outstanding balance at default. When considering loans that

may not be traded, and taking into consideration when cash was received as well as other losses incurred in the collection process, ultimate LGD is the relevant measure for an input into a regulatory or economic credit capital model. In the case of bonds or marketable loans, one can measure the prices of traded debt at the initial credit event, or the discounted market values of instruments received at the resolution of default. The latter is potentially a proxy for the ultimate LGD, the focus of our study for two purposes. Our primary objective is to provide results of use to agents invested in defaulted securities having time horizons that span the resolution period, who wish to assess expected value upon emergence relative to some benchmark available at default, such as trading or model-based prices. Agents who would benefit include bank workout specialists, risk managers, or vulture hedge fund investors. Second, our results would be relevant for financial institutions attempting to quantify economic LGD for purposes of the Basel II Advanced IRB approach to regulatory capital, which requires estimation of the ultimate LGD.

Aspects of this modeling exercise deserving of special attention include the distributional properties of LGD. While the available theory and empirical evidence suggests it to be stochastic and predictable with respect to other variables, in most extant credit models LGD has been treated as either deterministic or as an exogenous stochastic process. The quest for tractability gives rise to such assumptions, but in practical applications this results in understated capital, mispricing, and unrealistic dynamics of model outputs. Our research helps to resolve such deficiencies by modeling *ex ante* the distribution of LGD as a function of empirical determinants such as contractual features, firm capital structure, borrower characteristics, debt and equity market variables, and systematic factors. In order to empirically investigate the determinants of, and build predictive econometric models for, ultimate LGD, we use an extensive sample of 871 defaulted firms (covering 1985 to 2008), a dataset containing the complete capital structures of each obligor. Our sample is highly representative of the U.S. large corporate loss experience over the last two decades.

LGD has been a relatively neglected aspect of credit risk research.² Starting with the seminal work by Altman [1968], modeling of PD is currently in a relatively mature stage as compared to LGD. The heightened focus on LGD is evidenced by the recent flurry of research into

the application of credit models and estimation of LGD from available data. This literature ranges from simple quantification of LGD, to calibration of credit models embedding LGD assumptions, to empirical or vendor models of LGD. The Zeta model of Altman, Haldeman, and Narayanan [1977] was a second generation of the Altman Z-score PD estimation model. In this model, loan LGD estimates were based on a workout department survey (1971–1975), which yielded conclusions regarding the magnitude of discounted post-restructuring recoveries on unsecured bank loans. Bank studies focusing on internal loan data included research by JP Morgan Chase (Araten, Jacobs, Jr., and Varshney [2004]), where the authors studied ultimate workout LGD for wholesale loans during 1982 through 2000.

Among early studies relying almost exclusively on secondary market prices of bonds or loans soon after the time of default, Altman and Kishore [1996] estimated LGDs for defaulted senior secured and senior unsecured bonds from 1978 to 1995, yielding estimates that could be statistically distinguished among various industry groups. Altman and Eberhart [1994] and Fridson, Garman, and Okashima [2000] provided evidence that the more senior bonds significantly outperformed the more junior bonds in the post-default period, results confirmed by Hamilton, Gupton, and Berthault [2001] for secondary market loan prices a month after default. Emery [2003] and Altman and Fanjul [2004] compared LGDs on bank loans and bonds, respectively, as inferred from the prices of the traded instruments at default in a Moody's database, revealing that loans experience lower loss severities when controlling for seniority. Cantor, Hamilton, and Varma [2003] showed similar findings for corporate bonds as Altman and Fanjul and additionally found differential LGD by rating at origination, such that "fallen angels" of the same seniority had significantly lower LGDs.³

Among studies that looked at ultimate LGD, Standard and Poor's (Keisman and van de Castle [2000]) presented empirical results from the LossStatsTM in the 1987–1996 period for marketable bonds and loans. This study also showed that the influence of position in the capital structure (i.e., the proportion of debt above or below a claimant in bankruptcy) was independent of collateral and seniority in determining loss severities. A more recent rating agency study by Moody's (Cantor and Varma [2004]) examined the determinants of ultimate LGD for North American corporate issuers over a

period of 21 years (1983–2003), looking at many of the variables considered herein (e.g., seniority and security; firm/industry-specific and macroeconomic factors).

Several recent empirical studies of LGD attempting to measure LGD–PD correlation have put more structure around this exercise, either by building predictive econometric models or by attempting to directly test models. Frye [2000a, 2000b, 2000c] examined the LGD–PD correlation in extensions of the Merton [1974] framework allowing for systematic recovery risk and found a significant negative relationship at various levels of aggregation. Other studies examined this by looking at LGD as implied from the prices of traded debt at or prior to default, as opposed to ultimate LGD, or the “reduced-form” approach. Among these, Jarrow [2001] developed a hybrid structural–reduced model, in which PDs and LGDs were functions of the underlying state of the macro-economy. Hu and Perraudin [2002] also examined this relationship and found LGD–PD correlations on the order of 20%. Jokivuolle and Peura [2003] presented an option theoretic model for bank loans and were able to produce a positive correlation between PD and LGD. Bakshi et al. [2001] extended the reduced-form approach by allowing a flexible correlation between PD, LGD, and the risk-free rate. Imposing a negative correlation between PD and LGD, they found that a 4% increase in the (risk-neutral) hazard rate was associated with a 1% increase in the expected LGD. In related work on the resolution of default focusing on high-yield debt portfolios, Parnes [2009] developed a theoretical model that explicitly incorporated LGD assumptions.

Several studies showed that LGDs tend to rise more in periods of recession than they fall during expansions, suggesting that more is at play than a macroeconomic factor influencing the value of collateral. Keisman and van de Castle [2000] found that during the earlier stress period at the beginning of the previous decade, LGDs of all seniorities rose in the S&P LossStats™ database for the 1982–1999 period. Altman, Resti, and Sironi [2001, 2003] also found that LGDs increase as the credit cycle worsens and as default rates increase above the cycle’s long-run average.⁴ Araten et al. [2004] related unsecured U.S. large corporate borrower-level LGDs to the average Moody’s All-Corporate default rate and reported similar behavior. However, Altman, Brooks, Resti, and Sironi [2005] found that a systematic variable had no effect on LGD when controlling for bond market conditions (e.g., supply–demand imbalances). Acharya, Bharath,

and Srinivasan [2007] examined the same data and time period as Keisman and van de Castle [2000], and while they verified that seniority and security are key determinants of LGD, in addition they found industry-specific factors influencing LGD independently of the macroeconomic state and bond market conditions analyzed in Altman et al. [2005]. In particular, after controlling for firm-specific, contractual, and systematic factors they found elevated LGDs in distressed industries, defined as those sectors having significantly lower profitability than the overall market. They argued that in these cases fewer re-deployable assets, greater leverage, and lower liquidity are driving lower average recoveries and that their results support a test of the Schleifer and Vishny [1992] “fire-sale” hypothesis, an industry equilibrium phenomenon in which macro and bond market variables are spuriously significant due to omitting an industry factor.

Among studies similar to ours, Carey and Gordy [2007] argued for a two-stage approach to measuring LGD, first estimating an “estate LGD” at the obligor level and then treating instrument-level LGDs according to a contingent claims approach, as under APR such recoveries can be viewed as collar options on residual value of the firm. However, they argued that the endogeneity of the bankruptcy decision would result in a measurement problem in the first-stage borrower level. Furthermore, an extensive literature on violations of APR suggested a similar problem in the second-stage instrument level (Eberhart, Moore, and Rosenfeldt [1989]; Hotchkiss [1993]; Weiss [1990]). The authors also addressed the issue of whether systematic variation in LGD could be refuted given data limitations, in particular the large proportion of unexplained variation in the cross-section of recovery cash-flows.

Finally, we make note of this evidence regarding the PD–LGD correlation influencing the Basel II guidelines: paragraph 468 on downturn LGD in the Bank for International Settlements (BIS) Accord (Basel Committee of Banking Supervision [2003, 2004]) and the additional guidance offered by the BIS (Basel Committee of Banking Supervision [2005]). Basel II requires advanced internal ratings based (IRB) banks to capture all relevant risks regarding possible cyclical variability in LGD, and at the same time it states that bank estimates of long-term ultimate LGDs having no such systematic variations may be acceptable. Miu and Ozdemir [2006] argued that banks can incorporate conservatism into cyclical LGDs estimated in a point-in-time framework

without an LGD–PD correlation; however, they estimated commensurate increases in credit capital to compensate for this.

Our approach extends the existing research along several dimensions. First, we contribute to prior work by modeling LGD jointly at the firm and instrument levels. In particular, a simultaneous equation estimation of LGD at the instrument and obligor levels is advantageous in that we can model enterprise value coherently, which is understood by market participants in the large-corporate sector of the distressed and defaulted debt market to be a key determinant of recoveries. That is, under the assumption of a strict absolute priority rule (APR), recovery on an instrument residing somewhere in the capital structure of a firm can be likened to a collar option—there is positive recovery only if it is “in the money” or when enterprise value is sufficient to satisfy all superior claims. Second, as compared to extant work, we integrate new variables with those previously considered into a unified framework; including the cumulative-abnormal returns (CARs) on borrower’s equity prior to default, which decreases the ultimate LGD, size of the defaulted firm, and market price of distressed debt at default. Finally, in addition to these we confirm many of the findings of the literature in regard to determinants of LGD such as contractual, capital structure, firm-specific features, industry characteristics, and macroeconomic considerations.

ECONOMETRIC MODELING

To predict and estimate distributional characteristics of LGD, in our empirical analysis we utilize the beta-link generalized linear model (BLGLM), which is in the class of generalized linear models (GLMs) that encompass the logistic, truncated normal, and Tobit models. These models have been used mainly in understanding and quantifying economic relationships, as well as having been employed in prediction, both of which are of importance in the modeling of LGD. However, much of the literature has focused upon qualitative dependent variables, into which the case of PD estimation naturally falls. Maddala [1983, 1991] introduced, discussed, and formally compared the different models.⁵

The i th observation of the dependent (or response) variable, the ultimate LGD, is denoted by y_i . The vector of independent variables corresponding to y_i is denoted by $x_i = (x_{1i}, \dots, x_{pi})^T$. We assume that the conditional

expectation of y_i depends upon a linear function η of the x_i only through a smooth, invertible function m :

$$E_p[y_i | x_i] = \mu = \int_0^1 p(y_i | x_i) y_i d\mathbf{v}(y_i) = m(\eta) \quad (1)$$

$$\eta = \beta^T x_i = m^{-1}(\mu) \quad (2)$$

where $m^{-1}(\cdot)$ is defined as the *link function* that maps from the conditional mean of the response variable μ to the linear function $\eta(\cdot)$. Then the distribution of y_i resides in the exponential family, which implies a probability distribution function of the following form:

$$p(y_i | x_i, \beta, A_i, \zeta) = \exp \left[\frac{A_i}{\zeta} \{y_i \theta(x_i | \beta) - \gamma(x_i | \beta)\} + \tau \left(y_i, \frac{\zeta}{A_i} \right) \right] \quad (3)$$

where $\gamma(\cdot)$ and $\tau(\cdot)$ are smooth functions (satisfying certain regularity conditions), A_i is a known prior weight, $\zeta \in \mathbb{R}^+$ is a scale parameter (possibly known), and the location function $\theta(\cdot)$ is related to the linear predictor according to:

$$\theta(x_i | \beta) = (\gamma')^{-1}(\mu(x_i)) = (\gamma')^{-1}(m(\beta^T x_i)) \quad (4)$$

This framework subsumes many of the models in the literature on the classical linear regression and limited/qualitative dependent variables framework. Here we consider the case most relevant for LGD estimation and least pursued in the GLM literature. In this context, we are dealing with a random variable in a bounded region, the unit interval. This is most conveniently modeled through employing a beta distribution, in which case we denote the response percent loss rate as $l_i \in [0,1]$. As in Mallick and Gelfand [1994], we take the link function to be a mixture of cumulative beta distributions:

$$\eta(x_i | \beta, \phi, a, b) = \beta^T x_i = \sum_{j=1}^k \phi_j \int_{u=0}^{l_i} \frac{u^{a_j-1} (1-u)^{b_j-1}}{B[a_j, b_j]} du \quad (5)$$

where $B[x, y] = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)} = \int_0^1 u^{x-1} (1-u)^{y-1} du$ in Equation (5) is the standard beta function, $\Gamma(x) = x!$ is the gamma function, and the parameters ϕ, a, b are chosen to match features of the data.⁶ While in most cases we do not have a closed-form solution, we can always estimate

the underlying parameters β consistently and efficiently by maximizing the log-likelihood function:

$$\begin{aligned}
 & l(\theta(\beta | x_i), \zeta(\beta | x_i), \beta | x_i, y_i, A_i) \\
 &= \sum_{i=1}^n \left[\frac{A_i}{\zeta(\beta | x_i)} \{y_i \theta_i(\beta | x_i) - \gamma(\beta | x_i)\} \right. \\
 & \quad \left. + \tau \left(y_i, \frac{\zeta[\beta | x_i]}{A_i} \right) \right] \quad (6)
 \end{aligned}$$

We are able to obtain global convergence through optimizing Equation 6 across observations in our dataset, as discussed in subsequent sections.⁷ This approach has the benefit of the capability to model the highly bimodal nature of the distribution rather accurately, without recourse to more complex, less transparent (or replicable), more computationally demanding, and less stable alternatives such as non-parametric approaches (Renault and Scaillet [2003]) or maximum-entropy (Friedman and Sandow [2003]).

DATA AND SUMMARY STATISTICS

We have built a database of defaulted firms representative of the U.S. large corporate loss experience (bankruptcies and out-of-court settlements), all having Moody's rated instruments at some point prior to default. Our database is constructed through merging the December 2008 release of the Moody's Ultimate Recovery Database™ (MURD) with information from various sources such as www.Bankruptcy.com, Edgar SEC filings, LexisNexis, Bloomberg, Compustat, and the Center for Research in Security Prices (CRSP). It contains data on 3,902 defaulted instruments from 1986 to 2006 for 871 borrowers, for which there is information on all classes of debt in the capital structure at the time of default. Despite our sample selection being driven mainly by data availability in the matching of the MURD database to other databases, the final dataset is largely representative of the U.S. large-corporate default experience over the last 20 years.⁸ All instruments are detailed by debt type, seniority, collateral type, position in the capital structure, original and defaulted amount, resolution outcome, and instrument price or value of securities at the resolution of default (emergence from Chapter 11 bankruptcy as an independent entity or acquisition by a third party, Chapter 7 liquidation or

out-of-court settlement of a distressed exchange). It includes either the prices of pre-petition instruments at the time of emergence from, or prices of new instruments received in settlement of, bankruptcy or other distressed restructuring, respectively. For a subset of observations, we can obtain the prices of traded debt, the equity prices, or financial statement data at around the time of default.⁹ We calculate economic LGD by discounting nominal LGD by the coupon rate on the instrument prevailing just prior to default.¹⁰

Exhibit 1 presents the counts and sample averages of various key variables: trading instrument inferred LGD at default, discounted LGD, number of major creditor classes, principal at default, and time to final resolution. These are presented at the instrument level and the obligor level, and then further broken down among bankruptcies and out-of-court settlements. Most defaults are bankruptcies as opposed to out-of-court settlements, with 3,273 bankruptcies versus 629 settlements at the instrument level and 728 bankruptcies versus 143 out-of-court settlements at the obligor level.

Our dataset contains the prices of traded debt available at the time of default for only 1,118 out of 3,902 (or 28.6%) of the instruments and for 460 out of 871 (or 52.8%) of the obligors. Average LGDs across the entire sample as inferred from the prices of traded instruments at default are significantly higher than ultimate LGDs, 61.04% versus 43.21% at the instrument level, and 63.46% versus 46.88% at the obligor level, consistent with previous research (Cantor, Emery, Keisman, and Ou [2007]).¹¹ Discounted LGD is much higher for bankruptcies as compared with out-of-court settlements, 48.38% versus 16.31% at the instrument level, and 51.41% versus 23.80% at the obligor levels.¹² We also see the marked non-normality of the LGD estimates, which is accentuated for the ultimate LGDs as compared to the trading prices, with standard deviations at the instrument level of 40.4% versus 28.7% and at the obligor level of 31.8% versus 23.9%.

Firms in the database tend to have about two major creditor classes in a range from 1 to 6, with an average of 2.44 at the instrument level and 2.20 at the obligor level. Furthermore, and perhaps surprisingly, if we take this as a measure of the complexity of the capital structure, out-of-court settlements do not tend to have fewer creditor classes than bankruptcies, an average of 2.55 versus 2.42 at the instrument level, and 2.31 versus 2.18 at the obligor level. Average time to resolution from first instrument

