

# Implied Recovery

Sanjiv R. Das  
Santa Clara University

Paul Hanouna  
Villanova University

September 2006<sup>1</sup>

<sup>1</sup>Thanks to Viral Acharya, Santhosh Bandreddi, Darrell Duffie, Lisa Goldberg, Francis Longstaff, Marco Pagano and Raghu Sundaram, as well as seminar participants at Barclays Global Investors, San Francisco, who gave us many detailed and useful suggestions. We are grateful to RiskMetrics for the data. The first author is grateful for the financial support of a Santa Clara University grant and a Breetwor fellowship. The authors may be contacted at: Sanjiv Das, Santa Clara University, 500 El Camino Real, Santa Clara, CA 95053, Email: [srdas@scu.edu](mailto:srdas@scu.edu). Paul Hanouna, Villanova University, 800 Lancaster Avenue, Villanova, PA 19085, Email: [paul.hanouna@villanova.edu](mailto:paul.hanouna@villanova.edu).

## Abstract

# Implied Recovery

The extraction of recovery rates from credit-sensitive securities is an important open problem in the default risk area. Further, the determinants of recovery levels on default are also little understood. We develop a technique for bootstrapping implied risk-neutral forward recovery rate term structures from credit default swap (CDS) spread curves at any single point in time, using only the cross-section of spreads and an additional identification condition. The model uses varied functional relationships between default and recovery to enable simultaneous identification of both hazard rate and recovery term structures under the risk-neutral measure; the underlying fixed-point algorithm is fast and convergent. No time series data is required. We then use indicative CDS spread data from 3,130 firms from January 2000 to July 2002, and extract recovery rates for each firm; these exhibit a strong negative correlation with default probabilities. A principal components analysis of the recovery rates finds one major component (the level of the risk free rate) and another minor one (implied volatility of S&P 500). In a regression framework, market variables are able to explain the variation in the first and second components very well, with  $R^2$ s of 90.7% and 35.7% respectively, suggesting that the recovery rate may be modeled on observable factors for trading purposes.

# 1 Introduction

As the market for credit derivatives expands, the opportunity to extract forward-looking information from traded securities is growing. Just as the growth of equity option markets resulted in the extraction of forward-looking implied volatilities, this paper develops a technique to extract and identify the implied forward curve of recovery rates for a given firm on any date, using the extant credit default swap spread curve. We apply our technique to 3,130 firms over the period ranging from January 2000 to July 2002, and find that recovery rates are driven by two factors, the level of risk free rates, and the implied volatility of S&P 500 (VIX). Our model makes possible what is to our knowledge the first panel data set of recovery rates extracted for the U.S. markets from credit default swaps, thereby enabling an analysis of the systematic risk in recovery levels. We also uncover evidence that credit contagion extends to recovery risk. While we provide many results of interest for financial economists, the model also keeps traders' needs in mind – it can be recalibrated daily using current information as are term structure and options models. We denote this approach the *implied recovery* model.

Whereas models for default likelihood have been explored in detail in the literature (see, amongst others, Merton (1974), Leland (1994), Jarrow and Turnbull (1995), Longstaff and Schwartz (1995), Madan and Unal (1995), Leland and Toft (1996), Jarrow, Lando and Turnbull (1997), Duffie and Singleton (1999), Sobehart, Stein, Mikityanskaya and Li (2000), and Duffie, Saita and Wang (2005)), the literature on recovery rate calibration is less developed. Academics and practitioners have often assumed that the recovery rate in their models is constant, and set it to lie mostly in the 40-50% range for U.S. corporates, and about 25% for sovereigns. The recent and rapid development of the credit default swap (CDS) market has opened up promising approaches to extracting implied default rates and recovery rates.

If the recovery rate is held fixed, then the term structure of CDS spreads may be used to extract the term structure of risk-neutral default probabilities, either using a structural model approach as in the model of **CreditGrades** (Finger, Finkelstein, Lardy, Pan, Ta and Tierney (2002)), or in a reduced-form framework, as in Jarrow, Lando and Turnbull (1997), Duffie and Singleton (1999), or Das and Sundaram (2002). However, assuming recovery rates to be constant is an impractical imposition on models, especially in the face of mounting evidence that recovery rates are quite variable over time. For instance, Altman, Brady, Resti and Sironi (2005) examined the time series of default rates and recovery levels in the U.S. corporate bond market and found both to be quite variable. They also found a statistically significant negative relation between these two variables.

Zhang (2003) shows how joint identification of the term structure of default intensities and recovery rates may be carried out in a reduced form model. He applies the model to Argentine sovereign debt and finds that recovery rates are approximately 25%, the number widely used by the market. See Christensen (2005) for similar methods. Pan and Singleton (2005), using a panel of sovereign spreads on three countries (Mexico, Russia and Turkey), also identify recovery rates and default intensities jointly assuming recovery of face value

rather than market value. Exploiting information in both the time series and cross-section they find that recovery rates may indeed be quite different from the widely adopted 25% across various process specifications.

In contrast to these times-series dependent *econometric* approaches to recovery rate extraction, this paper adopts a cross-sectional *calibration* approach. The model does not require statistical estimation, and only uses the term structure of CDS spreads, and equity prices and volatility, on a given day to extract both, the term structure of hazard rates and of recoveries, using a *fixed-point* algorithm. We illustrate different implementation approaches; specifically our empirical work uses the functional relationship from the Merton (1974) model to decouple hazard rate and recovery term structures from CDS spread curves. The main differences between this approach and the econometric ones of Zhang (2003) and Pan and Singleton (2005) are two-fold. First, only information on a given trading day is used, in much the way traders would calibrate any derivatives model in practice. In the empirical approaches, time series information is also required. Second, rather than extract a single recovery rate, the model delivers an entire *forward term structure* of recovery. This is also consistent with the empirical relationship between recovery and hazard rates - indeed, we show that the two term structures are negatively correlated (pointwise) in the model.

Other strands of the estimation literature on recovery rates are aimed at explaining recovery rates in the cross-section and times series. Acharya, Bharath and Srinivasan (2003) find that industry effects are extremely important in distinguishing levels of recovery in the cross-section of firms over a long period of time. Chava, Stefanescu and Turnbull (2006) jointly estimate recovery rates and default intensities using a large panel data set of defaults. They find many results that are complemented by the ones we obtain here.

A previous paper that adopts a calibration approach is by Chan-Lau (2003). He used a curve-fitting approach to determine the *maximal* recovery rate; this is the highest constant recovery rate that may be assumed, such that the term structure of default intensities extracted from CDS spreads admits economically acceptable values. (Setting recovery rates too high will at some point, imply unacceptably high default intensities, holding spreads fixed). This paper offers two improvements to Chan-Lau's creative idea. First, it fits an entire term structure of recovery rates, not just a single rate (i.e. flat term structure). Second, it results in exact term-dependent recovery rates, not just upper bounds.

It is also possible to tease out implied recovery rates if there are two contracts that permit separating recovery risk from the probability of default. Berd (2005) shows how this is feasible using standard CDS contracts in conjunction with digital default swaps (DDS), whose payoffs are function of only default events, not recovery (they are based on predetermined recovery rates); such pairs enable disentanglement of recovery rates from default probabilities. This forms the essential approach in valuing a class of contracts called recovery swaps, which are options on realized versus contracted recovery rates. In Berd's model, calibrated default rates and recovery rates are positively correlated (a mathematical necessity, given fixed CDS spreads and no remaining degrees of freedom); however, the empirical record shows that realized default rates and recovery are in fact negatively correlated (Altman, Brady, Resti

and Sironi (2005)). We will show that it is possible, in our functional approach here, to preserve the negative correlation of hazard rates and recovery, thereby leading to empirically supported joint dynamics.

Bakshi, Madan and Zhang (2001) (see also Karoui (2005)) develop a reduced-form model in which it is possible to extract a term structure of recovery using market prices; they show that it fits the data on BBB U.S. corporate bonds very well. They find that the recovery of face value assumption provides better fitting models than one based on recovery of Treasury. Our model is similar to their paper in that we also aim to extract a term structure of recovery. In contrast to their econometric approach, our approach is algorithmic, aimed at fast fitting for large data sets. We apply our algorithm to over three-thousand firms for which we have a time series of **CreditGrades** model CDS spreads for two and a half years. A principal components analysis of the extracted recovery rates reveals two main factors, the level of the term structure, and equity market volatility. These two factors together explain 87.1% of the variation in recovery rates.

The algorithm simultaneously extracts the term structure of hazard rates, which is also explained by these factors. We find that the correlation of hazard rates and recoveries is negative, and that the absolute size of this negative correlation grows as the aggregate level of defaults rises. This is indicative of an aspect of credit contagion not noticed before, i.e. recoveries are systematically impacted in an adverse manner when hazard rates increase.

In the ensuing sections, we develop some notation (Section 2.1) and apply it to CDS pricing (Section 2.2); we then present the bootstrapping approach for hazard rates conditional on recovery rates (Section 2.3). Section 2.4 develops the algorithm for the simultaneous identification of recovery and hazard term structures. Examples are provided and the calibration is assessed for varied input term structures of CDS spreads and risk free interest rates. The dependence of hazard rates and recovery rates is assessed under varied functional assumptions. In Section 2.4.2 we present the main algorithm of the paper. Our empirical exercise is undertaken in Section 3. Section 4 concludes and provides additional discussion.

## 2 Methodology

### 2.1 Basic Set Up

There are  $N$  periods in the model, indexed by  $j = 1, \dots, N$ . Without loss of generality, each period is of length  $h$ , designated in units of years. Thus, time intervals in the model are  $\{(0, h), (h, 2h), \dots, ((N-1)h, Nh)\}$ . The corresponding end of period maturities are  $T_j = jh$ .

Risk free forward interest rates in the model are denoted  $f((j-1)h, jh) \equiv f(T_{j-1}, T_j)$ , i.e. the rate over the  $j$ th period in the model. We write these one period forward rates in short form as  $f_j$ , the forward rate applicable to the  $j$ th time interval. The discount functions (implicitly containing zero-coupon rates) may be written as functions of the forward rates,

i.e.

$$D(T_j) = \exp \left\{ - \sum_{k=1}^j f_k h \right\}, \quad (1)$$

which is the value of \$1 received at time  $T_j$ .

For a given firm, default is likely with an intensity denoted as  $\lambda_j \equiv \lambda(T_{j-1}, T_j)$ , constant over forward period  $j$ . Given these intensities, the survival function of the firm is defined as

$$S(T_j) = \exp \left\{ - \sum_{k=1}^j \lambda_k h \right\} \quad (2)$$

We assume that at time zero, a firm is currently solvent, i.e.  $S(T_0) = S(0) = 1$ .

## 2.2 Default Swaps

For the purposes of our model, the canonical default swap is a contract contingent on the default of a bond or loan, known as the “reference” instrument. The buyer of the default swap purchases credit protection against the default of the reference security, and in return pays a periodic payment to the seller. These periodic payments continue until maturity or until the reference instrument defaults, in which event, the seller of the swap makes good to the buyer the loss on default of the reference security. Extensive details on valuing these contracts may be found in Duffie (1999).

We denote the periodic premium payments made by the buyer to the seller of the  $N$  period default swap as a “spread”  $C_N$ , stated as an annualized percentage of the nominal value of the contract. Without loss of generality, we stipulate the nominal value to be \$1. We will assume that defaults occur at the end of the period, and given this, the premiums will be paid until the end of the period. Since premium payments are made as long as the reference instrument survives, the expected present value of the premiums paid on a default swap of maturity  $N$  periods is as follows:

$$A_N = C_N h \sum_{j=1}^N S(T_{j-1}) D(T_j) \quad (3)$$

This accounts for the expected present value of payments made from the buyer to the seller.

The other possible payment on the default swap arises in the event of default, and goes from the seller to the buyer. The expected present value of this payment depends on the recovery rate in the event of default, which we will denote as  $\phi_j \equiv \phi(T_{j-1}, T_j)$ , which is the recovery rate in the event that default occurs in period  $j$ . The loss payment on default is then equal to  $(1 - \phi_j)$ , for every \$1 of notional principal. This implicitly assumes the “recovery of par” (RP) convention is being used. This is without loss of generality. Other conventions such as “recovery of Treasury” (RT) or “recovery of market value” (RMV) might be used just as well.

The expected loss payment in period  $j$  is based on the probability of default in period  $j$ , conditional on no default in a prior period. This probability is given by the probability of surviving till period  $(j - 1)$  and then defaulting in period  $j$ , given as follows:

$$S(T_{j-1}) (1 - e^{-\lambda_j h}) \quad (4)$$

Therefore, the expected present value of loss payments on a default swap of  $N$  periods equals the following:

$$B_N = \sum_{j=1}^N S(T_{j-1}) (1 - e^{-\lambda_j h}) D(T_j) (1 - \phi_j) \quad (5)$$

The fair pricing of a default swap, i.e. a fair quote of premium  $C_N$  must be such that the expected present value of payments made by buyer and seller are equal, i.e.  $A_N = B_N$ .

## 2.3 Identifying Default Intensity and Recovery

In equations (3) and (5), the premium  $C_N$  and the discount functions  $D(T_j)$  are observable in the default risk and government bond markets respectively. Default intensities  $\lambda_j$  (and the consequent survival functions  $S(T_j)$ ), as well as recovery rates  $\phi_j$  are not directly observed and need to be inferred from the observables.

Since there are  $N$  periods, we may use  $N$  default swaps of increasing maturity each with premium  $C_j$ . This mean there are  $N$  equations  $A_N = B_N$  available, but  $2N$  parameters:  $\{\lambda_j, \phi_j\}, j = 1, 2, \dots, N$  to be inferred. Hence the system of equations is not sufficient to result in an identification of all the required parameters. In the development of this paper, we consider various identification restrictions that may be imposed so as to extract default and/or recovery information from the model.

### 2.3.1 Constant Recovery Rates

In practice, a common assumption is to fix the recovery rate to a known constant. If we impose the condition that  $\phi_j \equiv \phi, \forall j$ , it eliminates  $N$  parameters, leaving only the  $N$  default intensities,  $\lambda_j$ . Now, we have  $N$  equations with as many parameters, which may be identified in a recursive manner using bootstrapping. To establish ideas, we detail some of the bootstrapping procedure.

Starting with the one-period ( $N = 1$ ) default swap, with a premium  $C_1$  per annum, we equate payments on the swap as follows:

$$\begin{aligned} A_1 &= B_1 \\ C_1 h S(T_0)D(T_1) &= (1 - e^{-\lambda_1 h}) D(T_1)(1 - \phi) \\ C_1 h &= (1 - e^{-\lambda_1 h}) (1 - \phi) \end{aligned}$$

This results in an identification of  $\lambda_1$ , which is:

$$\lambda_1 = -\frac{1}{h} \ln \left[ \frac{1 - \phi - C_1 h}{1 - \phi} \right], \quad (6)$$

which also provides the survival function for the first period, i.e.  $S(T_1) = \exp(-\lambda_1 h)$ .

We now use the 2-period default swap to extract the intensity for the second period. The premium for this swap is denoted as  $C_2$ . We set  $A_2 = B_2$  and obtain the following equation which may be solved for  $\lambda_2$ .

$$C_2 h \sum_{j=1}^2 S(T_{j-1})D(T_j) = \sum_{j=1}^2 S(T_{j-1}) \left(1 - e^{-\lambda_j h}\right) D(T_j)(1 - \phi) \quad (7)$$

Expanding this equation, we have

$$\begin{aligned} & C_2 h \{S(T_0)D(T_1) + S(T_1)D(T_2)\} \\ &= S(T_0) \left(1 - e^{-\lambda_1 h}\right) D(T_1)(1 - \phi) \\ & \quad + S(T_1) \left(1 - e^{-\lambda_2 h}\right) D(T_2)(1 - \phi) \end{aligned}$$

Re-arranging this equation delivers the value of  $\lambda_2$ , i.e.

$$\lambda_2 = -\frac{1}{h} \ln[L_1/L_2] \quad (8)$$

$$\begin{aligned} L_1 &\equiv S(T_0) \left(1 - e^{-\lambda_1 h}\right) D(T_1)(1 - \phi) \\ & \quad + S(T_1)D(T_2)(1 - \phi) - C_2 h [D(T_1) + S(T_1)D(T_2)] \end{aligned} \quad (9)$$

$$L_2 \equiv S(T_1)D(T_2)[1 - \phi] \quad (10)$$

and we also note that  $S(T_0) = 1$ .

In general, we may now write down the expression for the  $k$ th default intensity:

$$\lambda_k = \frac{\ln \left[ \frac{S(T_{k-1})D(T_k)(1-\phi) + \sum_{j=1}^{k-1} G_j - C_k h \sum_{j=1}^k H_j}{S(T_{k-1})D(T_k)(1-\phi)} \right]}{-h} \quad (11)$$

$$G_j \equiv S(T_{j-1}) \left(1 - e^{-\lambda_j h}\right) D(T_j)(1 - \phi) \quad (12)$$

$$H_j \equiv S(T_{j-1})D(T_j) \quad (13)$$

Thus, we begin with  $\lambda_1$  and through a process of bootstrapping, we arrive at all  $\lambda_j, j = 1, 2, \dots, N$ .

### 2.3.2 Time-Dependent Recovery Rates

The analysis in the previous section is easily extended to the case where recovery rates are different in each period, i.e. we are given a vector of  $\phi_j$ s. The bootstrapping procedure

remains the same, and the general form of the intensity extraction equation becomes (for all  $k$ )

$$\lambda_k = \frac{-1}{h} \ln \left[ \frac{S(T_{k-1})D(T_k)(1 - \phi_j) + \sum_{j=1}^{k-1} G_j - C_k h \sum_{j=1}^k H_j}{S(T_{k-1})D(T_k)(1 - \phi_j)} \right] \quad (14)$$

$$G_j \equiv S(T_{j-1}) \left(1 - e^{-\lambda_j h}\right) D(T_j)(1 - \phi_j) \quad (15)$$

$$H_j \equiv S(T_{j-1})D(T_j) \quad (16)$$

This is the same as equation (11) where the constant  $\phi$  is replaced by maturity-specific  $\phi_j$ s.

## 2.4 Two-way Dependence of Intensity and Recovery

In this section, we present the heart of the paper, namely the approach that leads to identification of recovery rates separately from default intensities. We just showed that given  $N$  values of the recovery rate,  $\phi_j$ , one for each period  $j$ , we may employ bootstrapping to determine  $N$  values of default intensity,  $\lambda_j$ . Alternatively, we no longer impose the recovery rate exogenously; rather, we use a form of functional dependence between intensity and recovery to obtain “implied” recovery rates. This dependence must come from additional theoretical structure, backed up by the injection of data to complete the identification of recovery rates.

### 2.4.1 General Approach

Both, conditional default probabilities and recovery rate term structures may be extracted jointly if a function connecting the two is supplied. The function may be linear, or take complex nonlinear forms. We may write this function generally as

$$\phi(T) = \phi[\lambda(T)] = g[\lambda(T); \theta] \quad (17)$$

where both  $\phi, \lambda \in R^N$  are term structure vectors, and  $\theta$  is a parameter set. Without a functional connection (or as in the previous section, where  $\phi$  was assumed constant), many pairs of  $(\phi, \lambda)$  may be consistent with the same set of spreads; in short, precise identification is infeasible.

Given  $N$  CDS spreads it is not possible to extract  $N$  parameters for default probability ( $\lambda_j, j = 1 \dots N$ ) and another  $N$  parameters for recovery ( $\phi_j, j = 1 \dots N$ ). However, if we write each  $\phi_j$  as a function of the  $\lambda_j$ s, then we eliminate  $N$  parameters, and the system becomes perfectly identified, i.e. we obtain a system of  $N$  equations in  $N$  unknowns. It remains to establish a functional relationship between recovery rates and default probabilities.

Since we are using CDS spreads as input data (in a reduced-form model framework), we chose to use the structural model framework to obtain our identification condition, thereby exploiting both models. In this setting, the recovery rate at horizon  $T$  may be written as the ratio of the value of the firm  $V$  to debt face value  $F$ , if the firm is not solvent, i.e. when

$V(T) < F$ . Let the density function for firm value be  $f[V(T)]$ , then noting that

$$\begin{aligned} f[V(T)] &= f[V(T) \mid V(T) < F] \times \text{Prob}[V(T) < F] \\ \iff f[V(T) \mid V(T) < F] &= \frac{f[V(T)]}{\text{Prob}[V(T) < F]}, \end{aligned}$$

we may write the recovery rate (as a fraction of  $F$ ):

$$\begin{aligned} \phi(T) &= E \left\{ \frac{V(T)}{F} \mid V(T) < F \right\} \\ &= \frac{1}{F} E \{ V(T) \mid V(T) < F \} \\ &= \frac{1}{F} \int_0^F V(T) f[V(T) \mid V(T) < F] dV(T) \\ &= \frac{1}{F} \int_0^F V(T) \frac{f[V(T)]}{\text{Prob}[V(T) < F]} dV(T) \\ &= \frac{1}{F} \frac{1}{\text{Prob}[V(T) < F]} \int_0^F V(T) f[V(T)] dV(T) \end{aligned}$$

where  $E\{\cdot\}$  is the expectation under the risk-neutral measure, and the fourth line in the equation above contains the critical substitution. This formula contains two familiar components:

1. The term  $\text{Prob}[V(T) < F] \equiv \lambda(T)$  is the risk-neutral probability of default. Since this appears in the denominator, it imposes a negative correlation between recovery rates and default probabilities, and is a characteristic of structural models of default.
2. The term  $\int_0^F V(T) f[V(T)] dV(T)$  is the value of an undiscounted asset-or-nothing put option, i.e. the forward value of the option. We will exploit this in the empirical application that follows.

Hence, we may write the recovery rate as follows:

$$\phi(T) \equiv g[\lambda(T); \theta] = \frac{1}{\lambda(T)} \times \frac{[\text{Forward value of asset-or-nothing put}]}{F} \quad (18)$$

We note here that it is entirely feasible to determine recovery rates in the structural model itself, once the parameters of the process for the firm value  $V$  are known. In fact, both the recovery rate and default probability are separately identifiable in structural models. However, we would also like to ensure that the recovery rate and default probability are consistent with the term structure of CDS spreads. Therefore, rather than take the default probability from the structural model as the final answer, we use the proportionality relationship between recovery rates and default probability from the structural model in equation (18) to determine the term structures of recovery rates and default probabilities in

the reduced-form model framework. Essentially, we take the relationship between recovery and default probability from the structural model framework, and then use CDS spreads to pin down default probabilities (and also recovery rates, through the functional relationship from the structural model).

There are many ways in which we may compute the solution to the  $N$  equations in  $N$  unknowns for the  $\lambda_j$ s. (Note that once the  $\lambda_j$ s are known, equation (18) gives the values of the  $\phi_j$ s). In the interest of computational speed, we developed a simple approach that solves the equation system iteratively, pinning down both, recovery rates and default probabilities using a *fixed-point* argument, which is computationally efficient. Starting with a guess vector for the term structure of recovery, ( $\phi$ ), we compute the vector of hazard rates ( $\lambda$ ) consistent with CDS spreads using equation (14) in Section 2.3.2. Next, we compute  $\phi$  using the specified function stated generally in equations (17) or (18). We repeat the extraction of  $\lambda$  based on the new  $\phi$ , and again apply the function to get another  $\phi$  vector. Within a few iterations, the system converges to a stable set of  $\{\lambda, \phi\}$ , both consistent with the term structures of forward rates and CDS spreads. This fixed-point algorithm is very fast, and calibrates the curves within seconds. We point out here that this approach is just a numerically efficient way to implement what is essentially a bootstrap approach. We could just as well have implemented the model by solving for the default probability and recovery rate for each maturity one by one, starting from the shortest maturity and proceeding on by incrementing the horizon one period at a time.

The usefulness of this overall approach for implied volatility stems from the following features. First, *fast* computation, which is valuable to traders. Second, the model allows for a *theoretical* structure under the risk-neutral measure for separate identification of recovery rates. Three, the identification function is *dynamic* and adjusts over time, since it is implicitly a function of the state variables (stock price and volatility). We now specialize this general approach to the Merton (1974) model that we will use to implement our implied recovery algorithm.

## 2.4.2 Identification with the Merton model

The approach we will take in this paper is to use the Merton (1974) model for the theoretical structure, which is implemented using market prices of equity and the volatility of equity. The idea is as follows. In the Merton model, the risk-neutral probability of default over horizon  $T$  is given as

$$\lambda(T) = N[-d_2] \quad (19)$$

$$d_2 = d_1 - \sigma\sqrt{T} \quad (20)$$

$$d_1 = \frac{\ln(V/F) + (r + \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}} \quad (21)$$

where  $V \equiv V(0)$  and  $\sigma$  are the unobserved values of the firm and its volatility respectively. The risk free rate is  $r$  and the maturity of zero-coupon debt of face value  $F$  is  $T$ . As shown

in Merton (1974), we may extract the unknown values  $V$  and  $\sigma$  from the observable price of equity  $E$  and its volatility  $\sigma_E$ , using the following pair of equations:

$$E = V N(d_1) - F e^{-rT} N(d_2), \quad \sigma_E = \sigma N(d_1) \frac{V}{E} \quad (22)$$

Recall that in the Merton model, the first equation defines the price of equity as a call option on the value of the firm. The second equation expresses equity volatility as a function of firm volatility. These equations may be solved simultaneously to imply the values of  $\{V, \sigma\}$ .

The recovery rate in the Merton model is the value of the firm  $V$ , when  $V(T) < F$ . Hence, *conditional* on the value of the firm falling below the face value of debt at maturity, debt holders receive the assets of the firm, in the value of  $V(T)$ . The forward price of an ‘‘asset-or-nothing’’ put option, which is an option that pays the value of the asset when  $V(T) < F$ , i.e., when the underlying asset is less than the strike at maturity is as follows:

$$e^{rT} V N[-d_1]$$

Then, by applying equation (18), we have the recovery rate

$$\phi(T) = \frac{e^{rT}}{N[-d_2]} \frac{1}{F} E[V(T) \mid V(T) < F] = e^{rT} \frac{V}{F} \frac{N[-d_1]}{N[-d_2]}$$

We may also write this as

$$\phi(T) = e^{rT} \frac{V}{F} \frac{N[-d_1]}{\lambda(T)} \equiv g[\lambda(T); \theta] \quad (23)$$

where  $\theta = \{V, F, T, r, \sigma\}$ . This equation portrays the inverse relationship between recovery rates and default probabilities for each maturity  $T$ . We use this relationship to undertake the iterative algorithm for implied recovery rates. The algorithm is described next.

### 2.4.3 Algorithm

For the purposes of the empirical analyses that follow, we used a specific implementation version of the model that is explained here. The implied recovery algorithm is an iterative one, similar to that described earlier in the paper. The specific details are as follows:

1. Given the stock price in the market ( $S$ ), the stock volatility ( $\sigma_S$ ), and debt per share ( $D$ ), use the Merton (1974) model in equation (22) to solve for the value of the firm  $V$ , and the volatility of firm value  $\sigma$ .
2. The other inputs to the model are the forward curve  $f$  and the CDS spread curve  $c$ . Also choose the time step  $h$  in the model (we used  $h = 1/2$  year) and CDS spreads and forward rates to a maximum maturity of 5 years.

3. For each maturity  $T$ , compute the default probability  $\lambda(T)$  using equation (19) and the recovery rate  $\phi(T)$  using equation (23). Regress  $\ln[\phi(T)]$  on an intercept and  $\ln[\lambda(T)]$ . This loglinear relationship between  $\phi(T)$  and  $\lambda(T)$ , justified by the form of equation (18), provides the identification function for our iterative algorithm. The regression relationship provides an approximation to the true function and aides computation. We show in section 2.4.4 that the term structures of recovery rates are highly robust to the various specifications of the relationship between recovery rates and default probabilities.
4. Finally, using the forward curve, spread curve and a starting value for  $\phi(T)$  (we used an initial flat curve  $\phi(T) = 0.4, \forall T$ ), iterate between (i) finding  $\lambda(T)$  from a given curve  $\phi(T)$  using equation (14), and (ii) finding  $\phi(T)$  from  $\lambda(T)$  using the loglinear regression relationship (in point 3 above). The system stabilizes rapidly within a few iterations.

The procedure above exploits information from two markets, the equity market and the credit default swap market. (It also uses the forward curve of risk less rates, thereby incorporating information from the interest rate markets as well). To summarize, there are two steps: (a) Use information from the equity markets through the Merton model to obtain an identification relationship between recovery rates and default rates. (b) We then use this identification condition coupled with information in the CDS market to decompose spreads into their default probability and recovery components based on the application of the iterative fixed-point algorithm.

### *Example*

As an example, we applied our algorithm to data for Sun Microsystems (Ticker: SUNW) on May 16, 2006. We used the spread curve from `CreditGrades` and the current market data on stock price, volatility, debt per share and interest rates to imply the forward curve of recoveries and default probabilities. The results are shown in Figure 1. For short maturities, the forward recovery rates are high, then decline over time, before ticking back up slightly at the five-year horizon. This also shows a negative relationship with forward default probabilities.

Next, we demonstrate the flexibility of this approach by looking at two other approaches. One is denoted the maximal recovery rate model of Chan-Lau (2003), which is a special case of our algorithm; two, we show how the framework may be applied directly to piggyback on the empirical work of Altman, Brady, Resti and Sironi (2005). These examples are only used to show the robustness of the algorithm. In the remainder of the paper we will use the richer Merton algorithm in Section 2.4.2 for our empirical work.

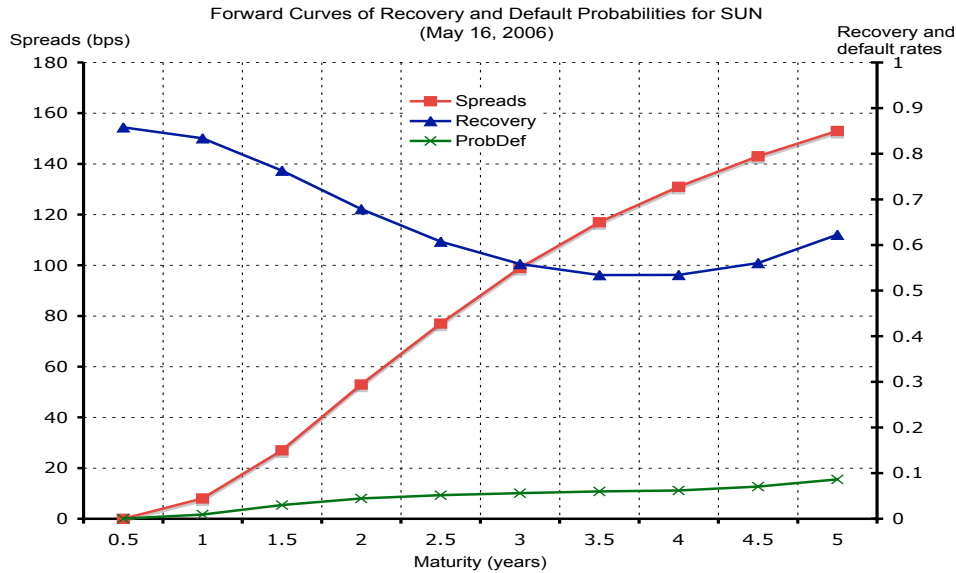


Figure 1: Implied risk-neutral forward recovery rates and default probabilities for Sun Microsystems on May 16, 2006. These are computed from the application of the algorithm in section 2.4.3. The default swap spread curve from `CreditGrades` is also shown. The stock price for SUNW was \$4.75, the forward curve was assumed flat at 5.43%, and the asset volatility was taken to be 0.46. Debt per share is 1.11.

#### 2.4.4 Robustness

For comparison, we also present two other simple approaches that may be adopted, but that are not as theoretically driven as the approach just covered, nor do they result in better identification. The first one is the idea of computing bounds on recovery rates. Whereas this does not provide a sharp estimate, it requires no additional identification structure or data. The second is to use an empirical identification condition, not a theoretical one as in the Merton model. Both approaches are in the same class of forward-looking recovery estimators. What is seen from these robustness comparisons is that our calibration approach does well in extracting forward implied recovery rates irrespective of the choice of identification function.

##### *Comparison to Maximal/Minimal Recovery Rates*

This notion of maximal recovery rates was introduced in Chan-Lau (2003). The paper's motivation was to obtain an upper bound on the constant recovery rate  $\phi$  given a set of default swap premia for various maturities.

Fixing a given recovery level  $\phi$ , the bootstrapping equation (14) is used to determine a vector of default intensities  $\lambda_j$ . Of course,  $\phi$  is an assumed value, and may be increased progressively. There is a level  $\phi_H$  beyond which elements in the derived default intensity vector may be unacceptably high, i.e. implying default probabilities that exceed 1; this value is denoted as the *maximal* recovery rate, i.e. an upper bound on the implied recovery rate. Likewise, we may define  $\phi_L$ , the *minimal* recovery rate, as being the lowest we can set the recovery rate without inferring unacceptable default intensities from the bootstrap

procedure. Therefore, using this scheme, we obtain upper and lower bounds.

As an example, consider the plot in Figure 2. The figure has two graphs. The upper graph shows the inputs: the term structure of forward rates, and the CDS spread curve. The lower graph shows the output: the implied hazard rates for increasing levels of the flat recovery rate, from an application of equation 14. In this plot, we assumed a flat term structure of recovery rates and then derived, using the mathematics of section 2.3.1, the term structure of hazard rates. At  $\phi = \{0.2275, 0.4550, 0.6825\}$ , the implied term structure of hazard rates is stable, and well within an acceptable range. When we raise  $\phi$  to 0.91, the conditional default probability for the 5-year maturity explodes beyond 1.0. Hence, we find that the maximal feasible recovery rate occurs at 0.91. At higher levels of recovery, the conditional probability of default needs to exceed 1 in order to be consistent with observed spreads. Therefore, hazard rates appear to be insensitive for a range of assumed recovery rates before experiencing a sharp move above acceptable bounds.

Implied recovery rates in the maximal-minimal framework tend to populate a very broad range. If the range of acceptable recovery rates that are consistent with the existing CDS spread term structure is too wide, it does not offer adequate identification of recovery rates. Furthermore, this approach assumes a flat term structure of recovery which may be too simple.

In the next section we explore a different approach using the regressions in Altman, Brady, Resti and Sironi (2005) to determining implied recovery that delivers well-identified implied hazard rate and recovery term structures in a simultaneous way.

### *Function-based Recovery with Regressions*

While the Merton (1974) model is easy to implement, it offers a relationship between recovery rates and default probabilities that is more complex and harder to implement than that of a simple regression model. For comparison, we implement the model framework using identification functions based on the work of Altman, Brady, Resti and Sironi (2005). In their paper, the authors investigated historical default rates and recovery rates in corporate bonds with a view to understanding their dependence, using data over the period 1982-2002. They are able to explain a significant amount of the variation in recovery rates ( $\phi$ ) using default rates ( $\lambda$ ) alone. We examine four models (the regression functions from their paper are reproduced here):

1. Linear model:  $\phi = 0.51 - 2.61\lambda$ ,  $R^2 = 0.51$ .
2. Quadratic model:  $\phi = 0.61 - 8.72\lambda + 54.8\lambda^2$ ,  $R^2 = 0.63$ .
3. Logarithmic model:  $\phi = 0.002 - 0.113 \ln(\lambda)$ ,  $R^2 = 0.65$ .
4. Power model:  $\phi = 0.138/\lambda^{0.29}$ ,  $R^2 = 0.65$ .

These functions fit the historical record well, and provide a basis for illustrative regression-based implementation of the model. Essentially, the theoretical relation in equation (23)

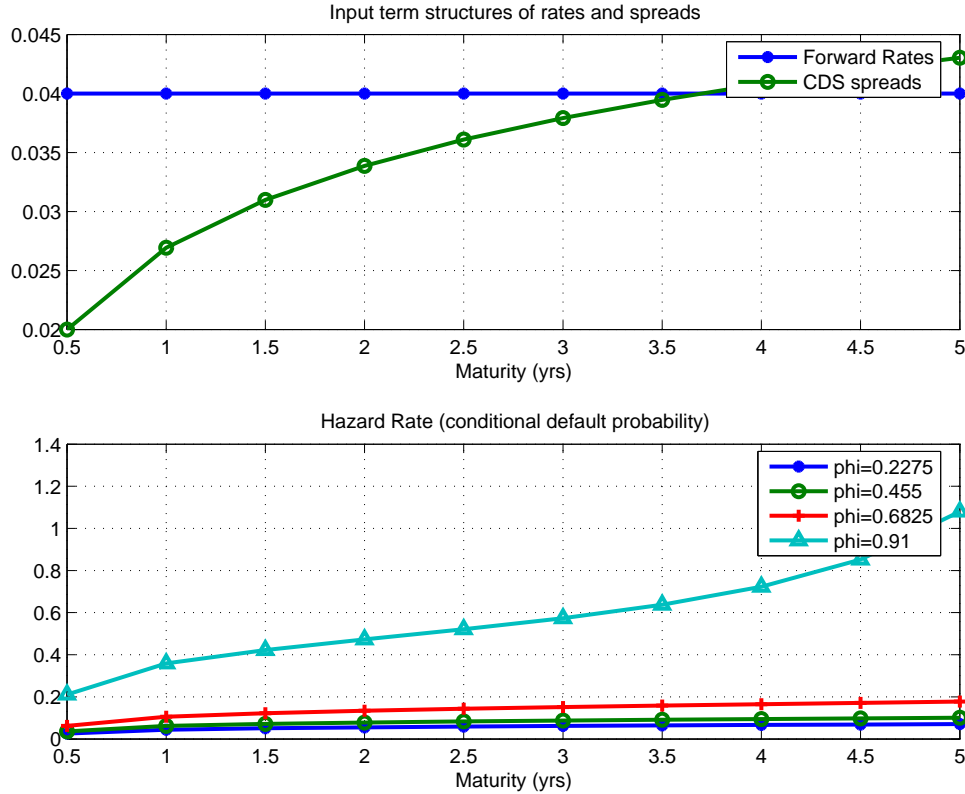


Figure 2: Fitted hazard rate term structures for the maximal recovery model of Chan-Lau (2003). The initial forward curve is given by the following equation  $f = 0.04$ , i.e. a flat curve. The initial CDS spread curve is given by the equation  $c = 0.02 + 0.01 \ln(t)$ . We present the below the inputs, and resultant hazard rates, for a range of fixed recovery rates:  $\{0.2275, 0.4550, 0.6825, 0.9100\}$ . The conditional probability of default blows up at the last level of recovery rate (the maximal recovery rate).

is replaced with one of the empirical regressions shown above. Assuming different forward curves and CDS spread curves, we generated term structures of hazard rates and forward recovery rates for all four functions. The analyses are portrayed in Figures 3 and 4. In each of these figures, there are three graphs. The top plot presents the input term structures of forward risk-free interest rates and CDS spreads. The middle plot shows the four hazard rate term structures implied by each of the functions used in the model, and the lower plot presents the corresponding implied forward recovery term structures.

First, we note that the implied  $\lambda$  and  $\phi$  vectors are relatively insensitive to the choice of empirical function. On average, the difference between term structures extracted from the four different functions is less than 5%.

Second, comparing across the figures, we notice that the linear model tends to deliver hazard and recovery term structures that are further away from that derived from the other three models. If we restrict attention only to the three nonlinear models, the difference in results is reduced. Intuitively, this may be related to the poorer fit of the linear model in the work of Altman, Brady, Resti and Sironi (2005) (in terms of  $R^2$ ). Notice also that the dispersion across the recovery term structures from the different functional forms used is

small in comparison to the historical standard deviation of recovery rates (see Altman and Fanjul (2004)). This suggests that our implied recovery rate term structures may be quite insensitive to the precise choice of identification function; such robustness bodes well for the practical use of the model.

Third, the linear model returns the highest term structures irrespective of input forward curves of interest rates and spreads. It is thus most likely to run up against the upper bounds encountered in the maximal recovery approach. Intuitively, the imposed linearity is likely to prevent this model from providing as good a fit as the other three functional choices.

In sum, the implied fitting of both, hazard and recovery term structures is possible using simple models of the relation between hazard rates and recovery. We see that different functional forms result in similar term structures, almost invariant in shape, with minor variation in level. These fitted hazard and recovery term structures may then be used to price various derivative contracts. We explored two identification models, (a) theoretical identification using the Merton (1974) model, (b) empirical identification using the regression relationships based on the historical record from the work of Altman, Brady, Resti and Sironi (2005). We also compared these to the approach in Chan-Lau (2003). In the next section, we undertake a large-scale empirical study, implementing identification with the structural model of Merton (1974).

## 3 Empirical Analysis of the Merton-based Identification Approach

### 3.1 Data

We use data from **CreditMetrics** on CDS spreads for 3,130 distinct firms for the period from January 2000 to July 2002. The data comprised an indicative spread curve with maturities from 1 to 10 years derived from market data using their model. We used only the curve from 1 to 5 years, and incorporated half-year time steps into our analysis by interpolating the half-year spread levels from 1 to 5 years. The data was of daily frequency for each firm; we averaged the spread curve across all days within a month to obtain a single spread curve per firm for the month. The database also contained stock prices, debt per share, and 1000-day historical equity volatilities, that we used for the part of our algorithm where the Merton model is invoked. Forward interest rate term structures are computed for each day by bootstrapping Treasury yields. This yield data is obtained from the Federal Reserve Historical data repository that is available on their web site. We ran the iterative algorithm on each firm's month-averaged spread curve to extract the implied *forward* recovery rates for maturities from 0.5 to 5 years. We used the 5-year forward recovery rate as an estimate of the asymptotic recovery rate for each firm-month observation.

There are a few firms for which the algorithm is unable to converge. Many times, this was because the spreads were very small (less than 10 basis points), as in cases where the firm has almost no debt; these are uninteresting cases. After dropping these cases, we are

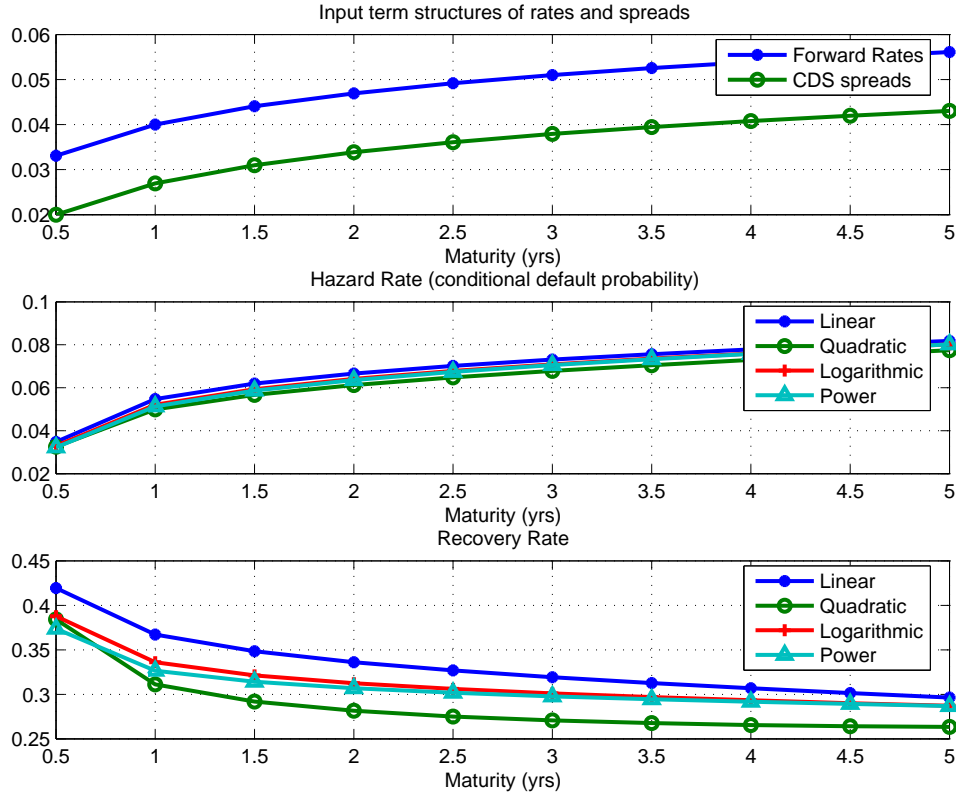


Figure 3: Fitted hazard rate and recovery term structures for the model of Altman, Brady, Resti and Sironi (2005). The fit is undertaken using the four regressions in the paper presented in Figure 1 therein. The paper contains four models: (a) Linear  $\phi = 0.51 - 2.61\lambda$ , (b) Quadratic  $\phi = 0.61 - 8.72\lambda + 54.8\lambda^2$ , (c) Logarithmic  $\phi = 0.002 - 0.113 \ln(\lambda)$ , (d) Power  $\phi = 0.138/\lambda^{0.29}$ . The initial forward curve is given by the equation  $f = 0.04 + 0.01 \ln(t)$ , and the initial CDS spread curve is given by the equation  $c = 0.02 + 0.01 \ln(2t)$ . We present the below the inputs (top graph), hazard rates (middle graph), and the forward recovery rates (lower graph) in three consecutive panels.

left with 84,187 firm-month observations. We remind the reader that the output of the algorithm results in *risk-neutral* implied recovery rates. We also obtain the risk-neutral forward default probabilities as a by-product of extracting the implied recovery rates, as it is part of the iterative algorithm.

### 3.2 Descriptive statistics

Table 1 presents a summary of the data generated by the algorithm. We sorted firms into quintiles by expected default frequency (EDF, or probability of default, provided by **CreditMetrics**; note that these are under the physical probability measure). The table shows the implied risk-neutral recovery rate in each of five quintiles sorted by EDF. Mean recovery rates within quintiles range from 83.3% in the lowest EDF quintile to 47.9% in the highest EDF quintile (most firms that default come from the highest EDF quintile, and therefore, the value is close to the market assumed 50%). Hence, recovery rates are inversely proportional to expected default rates, consistent with the findings of Altman and Fanjul

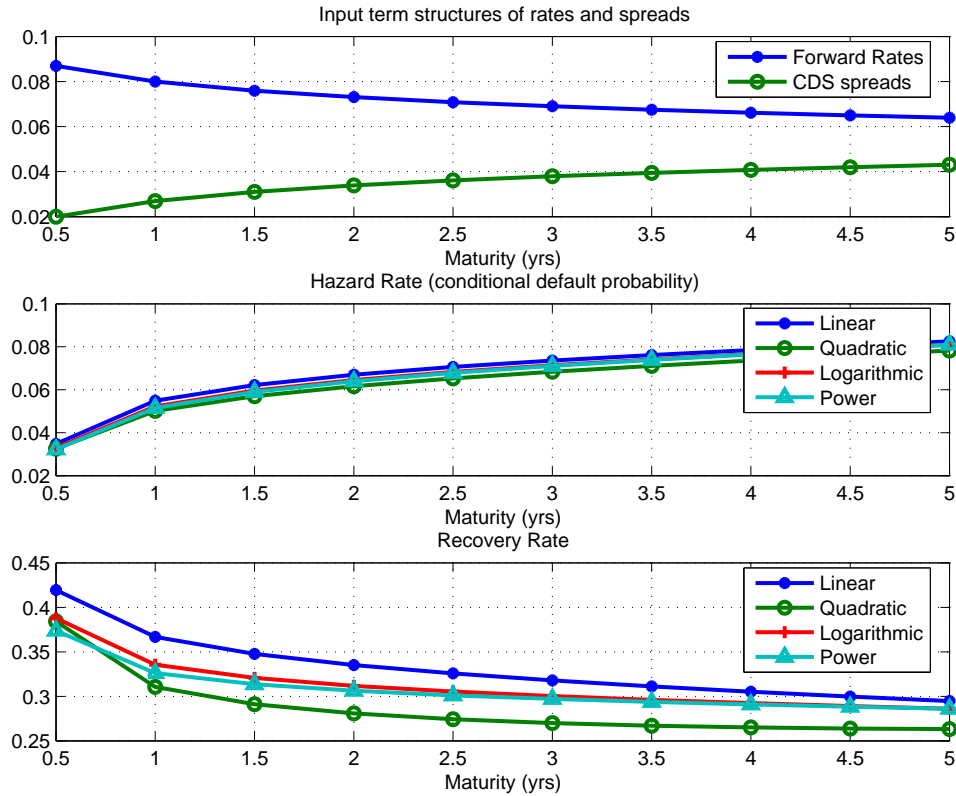


Figure 4: Fitted hazard rate and recovery term structures for the model of Altman, Brady, Resti and Sironi (2005). The fit is undertaken using the four regressions in the paper presented in Figure 1 therein. The paper contains four models: (a) Linear  $\phi = 0.51 - 2.61\lambda$ , (b) Quadratic  $\phi = 0.61 - 8.72\lambda + 54.8\lambda^2$ , (c) Logarithmic  $\phi = 0.002 - 0.113 \ln(\lambda)$ , (d) Power  $\phi = 0.138/\lambda^{0.29}$ . The initial forward curve is given by the equation  $f = 0.08 - 0.01 \ln(t)$ , and the initial CDS spread is given by the equation  $c = 0.02 + 0.01 \ln(2t)$ . We present the below the inputs (top graph), hazard rates (middle graph), and the forward recovery rates (lower graph) in three consecutive panels.

(2004). The standard deviation of absolute recoveries increases as the EDF increases, varying from 7% to 14%. We notice from the interquartile range that recovery rate ranges across adjacent EDF quintiles tend to overlap substantively.

Risk-neutral default probabilities are positively related to EDFs. We see that as we go from the lowest EDF to the highest EDF quintile, the mean forward risk-neutral default probability within each quintile increases from 1.7% to 20.1%. The variation in default probability within each quintile is about 3%. Median and mean recovery rates and default probabilities are similar. Figure 5 shows the time series of recovery rates and default probabilities averaged cross-sectionally (equally-weighted) each month. The inverse relationship between recovery rates and default probabilities is quite evident.

The monthly cross-sectional correlation over time between default probabilities and recovery rates is presented in Figure 6. This correlation becomes increasingly negative over time in the sample, coincident with increasing default risk in the economy, i.e. the absolute magnitude of these negative correlations increases over the sample period. These correlations are of the same magnitude as those uncovered by Chava, Stefanescu and Turnbull (2006).

Table 1: Descriptive statistics of implied recovery rates. The data is sorted by expected default frequency (EDF, or probability of default, provided by **CreditGrades**). The table shows the implied recovery rate and probability of default in each of five quintiles sorted by EDF.

<i>Panel A: Recovery Rate (<math>\phi</math>)</i>						
<i>Quintile</i>	N	Mean	Std	Q1	Median	Q3
1-Lowest EDF	16836	0.83276	0.07234	0.78338	0.84257	0.885
2	16839	0.77728	0.09941	0.69617	0.79045	0.855
3	16837	0.71063	0.11518	0.61367	0.71129	0.806
4	16838	0.63468	0.12754	0.53062	0.62107	0.743
5-Highest EDF	16837	0.47889	0.14394	0.37485	0.4587	0.579

<i>Panel B: Risk-Neutral Probability of Default (<math>\lambda</math>)</i>						
<i>Quintile</i>	N	Mean	Std	Q1	Median	Q3
1-Lowest EDF	16836	0.01695	0.01129	0.01	0.01	0.021
2	16839	0.07412	0.03149	0.05082	0.06838	0.091
3	16837	0.13957	0.03933	0.11093	0.13279	0.163
4	16838	0.1832	0.02981	0.16390	0.18259	0.203
5-Highest EDF	16837	0.20066	0.03837	0.19172	0.20423	0.217

This pattern of correlation suggests that there may be additional contagion risk in defaults showing up in recovery levels. It is known that default probability correlations increase with higher default risk levels in the economy, as pointed out by Das, Freed, Geng, and Kapadia (2001), and that contagion effects are prevalent, as evidenced in Das, Duffie, Kapadia and Saita (2004).

The fact that the correlation between recovery and default becomes more negative as default increases, implies that contagion effects also spread to losses given default. That is, when default risk levels increase in the economy, the losses given default are likely to be higher, implying a souring of the resale market for the firm's assets. Thus, the contagion effects of increasing default risk on credit portfolios may be even more adverse than previously known.

### 3.3 Principal components analysis of recovery rates

After extracting the recovery rates for each firm, we then sorted firms by their time-series averaged EDF into ten decile portfolios. For each decile we obtain the portfolio recovery rate, assuming equal weights for each firm. The time series of all ten portfolios is subjected to principal components analysis. Figure 7 shows the percentage variation explained by each component. There is one main component, which accounts for 78.9% of the common variation across the firms. The second component is much smaller, explaining only 8.2% of common variation. Thus, for parsimony, a single factor model might suffice to explain the variation across time in implied risk-neutral forward recovery rates.

After extracting the principal components, we generated the time series for the first two

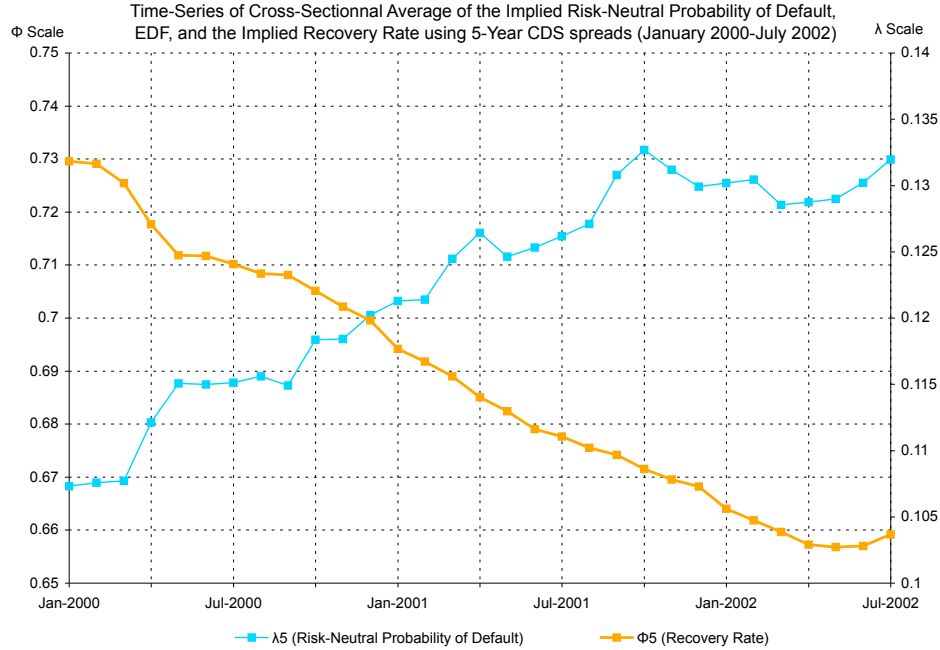


Figure 5: Cross-sectional average of the implied recovery rate ( $\phi$ ) and the corresponding risk-neutral probability ( $\lambda$ ) for all firms in the sample for the period from January 2000 to July 2002. For each month and each firm, we computed the implied recovery rate using the Merton-based algorithm for each half-year forward period upto 5 years. The plot above uses the implied forward recovery rates for the last period, which proxies for the asymptotic recovery rate in the model. Recovery rates are implied using information from both, the equity and credit default swap markets. Full details of the algorithm used are presented in section 2.4.2. The recovery rate time series slopes downwards, and the default probability series slopes upwards, clearly evidencing the negative correlation between the two.

components, and then plotted them in Figure 8. The main component, as expected, will track the time series of the recovery rate closely, given that it explains a substantial proportion of the common variation in recovery rates. In the next subsection, we will analyze various possible variables with a view to identifying these components.

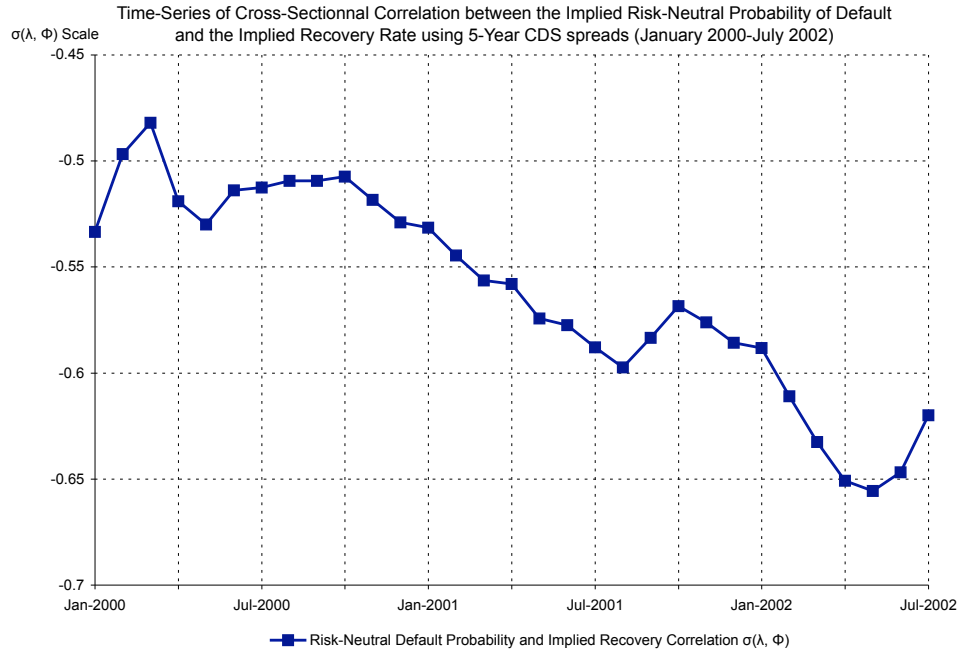


Figure 6: Cross-sectional correlation of the implied recovery rate ( $\phi$ ) and the corresponding risk-neutral probability ( $\lambda$ ) for all firms in the sample for the period from January 2000 to July 2002. For each month and each firm, we computed the implied recovery rate and default probability using the Merton-based iterative algorithm for each half-year forward period upto 5 years. The plot above uses the implied forward recovery rates for the last period, which proxies for the asymptotic recovery rate in the model. Full details of the algorithm used are presented in section 2.4.2. Correlations between recovery and default rates are computed in the cross-section for each month. The correlations become more negative (increases in absolute sign) as default risk in the economy increases.

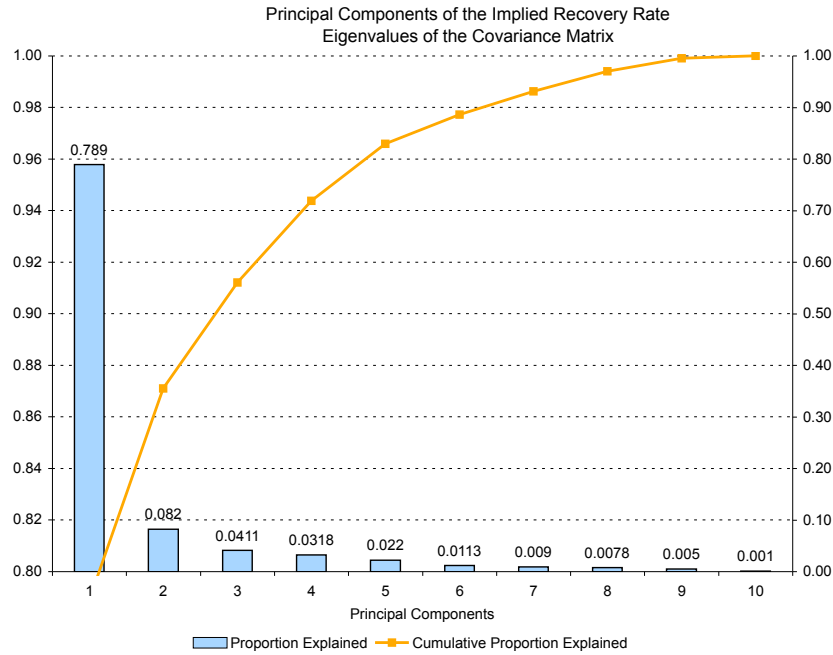


Figure 7: Eigenvalues of the common variation in the implied recovery rate ( $\phi$ ) for ten decile portfolios of firms in the sample for the period from January 2000 to July 2002. Firms were sorted into deciles based on their EDF levels.

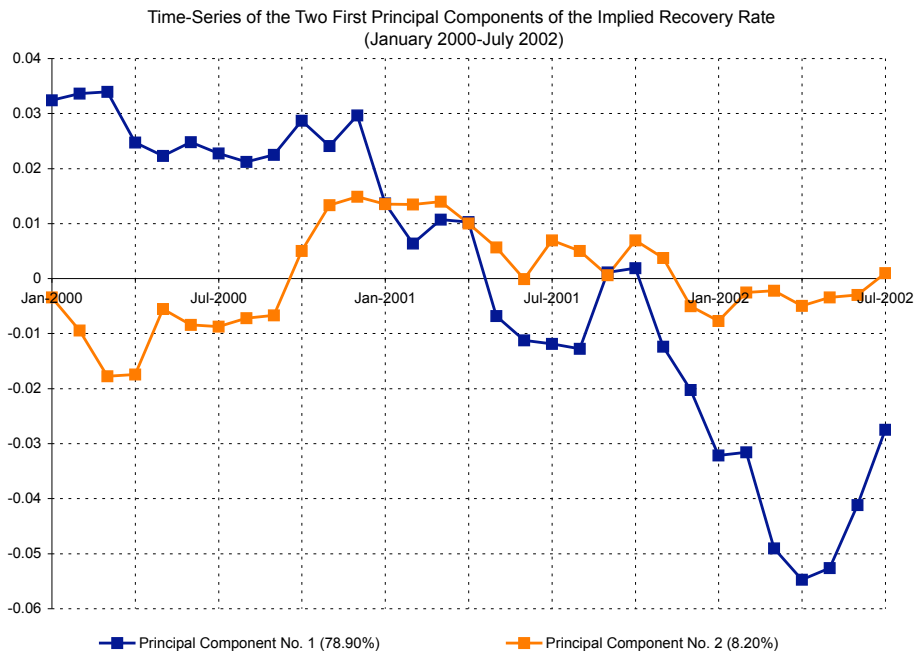


Figure 8: Time series of the first two principal components of the implied recovery rate ( $\phi$ ) for ten decile portfolios of firms in the sample for the period from January 2000 to July 2002. Firms were sorted into deciles based on their EDF levels.

Table 2: Correlation of principal components with market variables, and their descriptive statistics. We report the correlations of the first two principal components on various independent variables for the period January 2000 to July 2002 (p-values below correlations). The risk free rate is the one-month rate taken from the Fama French database, along with the excess market return, small minus big (SMB), high minus low (HML) and up minus down (UMD). The annualized percentage implied volatility of the S&P500 (VIX) is from Yahoo Finance. The inflation (IR) rate is the percentage change in CPI. The CPI index is the consumer price index from the Bureau of Labor Statistics - the CPI-Urban Consumers, All Items. “PCn” stands for the  $n$ -th principal component.

<i>Panel A: Correlations</i>										
	PC1	PC2	RF	IR	VIX	MKTRF	HML	SMB	UMD	
PC1	1									
PC2	0	1								
Risk Free Rate	0.90799	0.10318	1							
	0.0001	0.5807								
Inflation Rate	0.15537	-0.22401	0.23681	1						
	0.4039	0.2257	0.1996							
Implied Volatility of S&P 500 (VIX)	0.15886	0.41938	-0.00398	-0.25005	1					
	0.3933	0.0188	0.9831	0.1749						
Excess Market Return	0.08968	-0.14097	-0.00112	0.04913	-0.45252	1				
	0.6314	0.4494	0.9952	0.793	0.0106					
High-Minus-Low	0.0265	0.1003	0.18529	-0.06813	0.09844	-0.5406	1			
	0.8875	0.5914	0.3183	0.7157	0.5983	0.0017				
Small-Minus-Big	-0.05211	0.10226	-0.12695	0.07014	-0.18503	0.2789	-0.69743	1		
	0.7807	0.5841	0.4962	0.7077	0.319	0.1287	0.0001			
Up-Minus-Down	-0.12084	-0.10828	-0.2041	0.09437	-0.02499	-0.27649	-0.05409	0.3063	1	
	0.5173	0.562	0.2708	0.6136	0.8938	0.1321	0.7726	0.0938		
<i>Panel B: Descriptive Statistics</i>										
	Mean	Std	Q1	Median	Q3					
<i>Risk Free Rate</i>	0.0034	0.0015	0.0015	0.0039	0.0048					
<i>Implied Volatility of S&amp;P 500</i>	24.0645	4.1604	21.09	23.65	26.20					
<i>Inflation Rate</i>	0.0022	0.0030	0.0	0.0023	0.0045					
<i>Excess Market Return</i>	-0.0164	0.0542	-0.0622	-0.0213	0.0274					
<i>High-Minus-Low</i>	0.0202	0.0607	-0.0130	0.0285	0.0609					
<i>Small-Minus-Big</i>	0.0079	0.0690	-0.0373	0.0034	0.0430					
<i>Up-Minus-Down</i>	0.0145	0.0877	-0.0470	0.0218	0.0683					

### 3.4 Relating principal components to market variables

After extracting the time series of the principal components, we then related them to various market variables which are known to be relevant for the credit markets. In Table 2 we present the computed correlations of the first two principal components with these market factors. The table also shows the descriptive statistics of the market variables.

The first principal component  $PC1$ , is positively correlated with the level of the 1-month risk free interest rate ( $RF$ ). The second principal component  $PC2$ , displays its highest and only significant positive correlation with the implied volatility of the S&P500 index ( $VIX$  index). Therefore, the evidence suggests that term structure level and equity market volatility are important determinants of the level of recovery rates in the economy. These two variables can be seen to be orthogonal to each other from the correlation table (See also Trück, Harpainter and Rachev (2005), who find that business cycle variables explain recovery rates over time.. Duffie, Saita and Wang (2005) find strong corroborative evidence that the level of interest rates is an important determinant of default risk. This supports our finding that interest rate levels primarily determine recovery rates. Further, it is well known that the distance-to-default ( $DTD$ ) measure, widely used for default prediction, is a volatility-adjusted metric of leverage. Duffie, Saita and Wang (2005) also find  $DTD$  to be highly significant in explaining defaults, and therefore, the evidence here that volatility impacts recovery rates corroborates similar other findings.

Table 3 presents the results of a regression of each principal component on the market variables. These market factors are able to explain 90.7% of the variation in the first principal component. The risk free rate, equity volatility, excess market return, and the momentum factor are all highly significant. Unlike the first principal component, the second one is most significantly related to implied volatility of S&P 500 ( $VIX$ ). It is also strongly related to the  $SMB$  and  $HML$  factors of the Fama-French model. These various factors explain 35.7% of the variation in the second principal component.

This analysis suggests that the first component is most likely the rate of interest (corroborating the business cycle factor noted in Schuermann (2004)), and the second one is market volatility, and is what we would expect from the raw correlations. We may also verify whether these variables provide the highest economic impact by measuring the impact on the principal component value of a one standard deviation change in the market variable. This is done by multiplying the standard deviation of the variable from Table 2 by the regression coefficient from Table 3 (we do not report these calculations in order to keep the exposition brief). The risk free rate has the highest impact on the first principal component and the implied volatility of the S&P 500 has the greatest economic impact on the second principal component.

### 3.5 Analyzing default probabilities

Although the focus of this paper is on implied recovery rates, for completeness of analysis, we also present here the results of a principal components analysis of implied default proba-

Table 3: Explaining the principal components of recovery rates using market variables. We report the results of regressions of the first two principal components, our dependent variables on various independent variables for the period January 2000 to July 2002. The risk free rate is the one-month rate taken from the Fama French database, along with the excess market return, small minus big (SMB), high minus low (HML) and up minus down (UMD). The annualized percentage implied volatility of the S&P500 (VIX) is from Yahoo Finance. The inflation rate is the percentage change in CPI. The CPI index is the consumer price index. The standard errors and t-statistics that are reported are corrected for heteroskedasticity using the White (1980) approach. T-statistics are reported below the coefficient values. “PCn” stands for the  $n$ -th principal component.

	<i>PC1</i>	<i>PC2</i>
<i>Intercept</i>	-0.098	-0.029
	-8.89	-3.46
<i>Risk Free Rate</i>	17.633	0.440
	16.11	0.64
<i>Inflation Rate</i>	-0.329	-0.323
	-0.61	-0.62
<i>Implied Volatility of S&amp;P 500</i>	0.002	0.001
	4.32	3.44
<i>Excess Market Return</i>	0.111	0.028
	2.64	0.79
<i>High-Minus-Low</i>	-0.020	0.075
	-0.64	2.73
<i>Small-Minus-Big</i>	-0.040	0.076
	-1.14	2.94
<i>Up-Minus-Down</i>	0.049	-0.018
	2.23	-0.81
R-Square	90.7%	35.7%
N	31	31

bilities. As with recovery rates, we formed deciles of firms sorted by EDF and then compiled the time series average of each decile for default probabilities that were extracted using our algorithm.

In contrast to recovery rates, we find that default probabilities may be projected onto at least three, if not four principal components, as displayed in Figure 9. The largest principal component explains only 41.9% of the common variation, whereas for recovery rates, the main component explained as much as 78.9% of variation. The second principal component accounts for 30.4% of common variation (versus 8.2% for recovery rates). Thus, recovery rates appear to be simpler to explain than default probabilities. The third and fourth principal components explain 11.7% and 6.8% of the variation respectively, accounting for a reasonable portion of variation. Together, all four factors explain 90.9% of the common variation. The times series of the four components is displayed in Figure 10.

To gain an understanding of which factors the principal components relate to, we examine the correlations of the components with our market factors, as shown in Table 4. We also

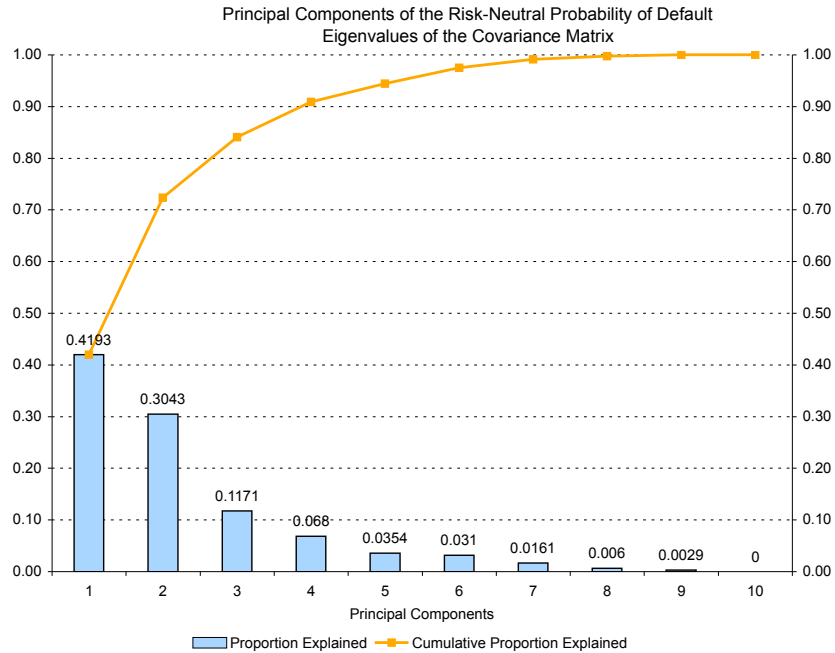


Figure 9: Eigenvalues of the common variation in the implied default probability ( $\lambda$ ) for ten decile portfolios of firms in the sample for the period from January 2000 to July 2002. Firms were sorted into deciles based on their EDF levels.

regress the four principal components on the market variables, reporting these results in Table 5. Taken together, based on absolute size of correlations and significance levels in the regressions, the evidence is suggestive that the first component is the term structure. The second component is most correlated with the level of the term structure (risk free rate) and equity market volatility (VIX). Excluding the risk free rate, the third principal component appears most correlated with the HML factor in the Fama-French data set and with the inflation rate. Finally, the fourth component relates most significantly to the excess return on the market (the S&P500).

Recall that for recovery rates, the first two components were the level of the term structure and market volatility respectively. Overall, therefore, the first and second components of default probability appear to be similar to those for recovery rates, but as we will see their importance is reversed. We may verify this by computing the correlation between the principal components of recovery rates and default probabilities. In confirmation, the lower half of Table 4 shows the correlations between the principal components of default probability and of recovery rates, which are exactly as expected based on the regression evidence. The first principal component of recovery is most correlated with the second one for default probabilities, and vice versa. The third and fourth principal components of default probability are not correlated with the principal components of recovery; hence there are additional components in default probability that are not present in recovery rates.

We may also verify whether these variables provide the highest economic impact by measuring the impact on the principal component value of a one standard deviation change in the market variable. This is done by multiplying the standard deviation of the variable

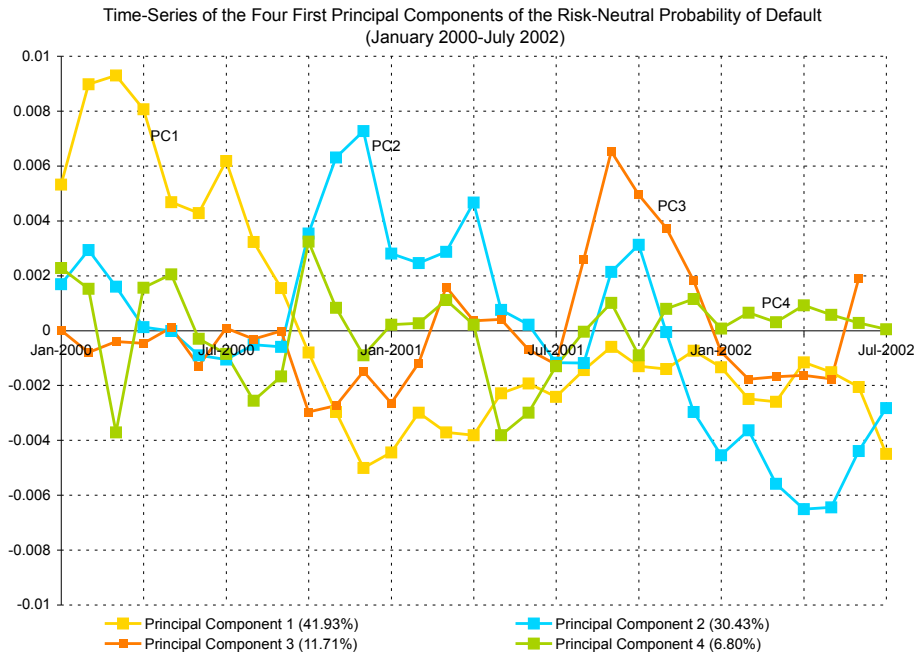


Figure 10: Time series of the first two principal components of the implied default probability ( $\lambda$ ) for ten decile portfolios of firms in the sample for the period from January 2000 to July 2002. Firms were sorted into deciles based on their EDF levels.

from Table 2 by the regression coefficient from Table 5; the results confirm the higher impact of the risk free rate and equity market volatility on variation in default probabilities.

## 4 Concluding Comments

Our understanding of recovery rates on default is improving, and the literature in this area is growing. Extraction of recovery rates separately from hazard rates has been explored in prior literature in both, the realm of structural models and reduced-form ones. Many of these models extract a single recovery rate, or require a long time series for calibration. Our goal in this paper is to use the term structure of CDS spreads with equity market information on a given day to extract both, term structures of hazard rates and recovery rates simultaneously, using a single functional assumption. Hence, we are able not only to extract the entire term structure of recovery, but to do so using a single day's CDS term structure and no time series information, resulting in a bootstrapping model that enables construction of a panel data set of forward implied recovery rates and a technique that is of practical value for economists and traders. The fitting approach exploits a fixed-point algorithm that makes the calibration extremely fast, and does not rely on any sort of minimization or statistical optimization. A range of functional choices from the literature are used in experiments, showing that the approach provides robust term structures of recovery.

Our empirical implementation uses CDS spreads and adopts the Merton (1974) model for the identification condition between recovery rates and default probabilities, though any

Table 4: Correlation of principal components of default probabilities with market variables. We report the correlations of the first four principal components on various independent variables for the period January 2000 to July 2002 (p-value below the correlations). The risk free rate is the one-month rate taken from the Fama French database, along with the excess market return, small minus big (SMB), high minus low (HML) and up minus down (UMD). The annualized percentage implied volatility of the S&P500 (VIX) is from Yahoo Finance. The inflation rate is the percentage change in CPI. The CPI index is the consumer price index. The  $p$ -values are reported below the correlations. “PC $n$ ” stands for the  $n$ -th principal component. The lower panel of the table shows the correlation of the principal components between that of the default probability and those of the recovery rates.

	<i>PC1</i>	<i>PC2</i>	<i>PC3</i>	<i>PC4</i>
<i>Risk Free Rate</i>	0.3740	0.7199	-0.3563	-0.0247
	0.0382	0.0001	0.0491	0.8951
<i>Inflation Rate</i>	0.2602	0.0103	-0.3817	-0.1240
	0.1574	0.9560	0.0341	0.5062
<i>Implied Volatility of S&amp;P 500 (VIX)</i>	-0.2292	0.4793	0.3067	0.2708
	0.2149	0.0064	0.0933	0.1407
<i>Excess Market Return</i>	0.1469	0.0326	0.3457	-0.3176
	0.4305	0.8619	0.0568	0.0817
<i>High-Minus-Low</i>	-0.0783	0.0427	-0.4720	-0.0208
	0.6755	0.8195	0.0073	0.9114
<i>Small-Minus-Big</i>	-0.0296	-0.0051	0.1086	0.1288
	0.8743	0.9783	0.5609	0.4900
<i>Up-Minus-Down</i>	0.0665	-0.1754	-0.1945	-0.0095
	0.7225	0.3454	0.2944	0.9594
<i>Recovery PC1</i>	0.51481	0.79783	-0.06862	-0.01421
	0.0030	0.0001	0.7138	0.9395
<i>Recovery PC2</i>	-0.79862	0.50876	-0.08700	0.07477
	0.0001	0.0035	0.6417	0.6893

other structural model may be used. Using data from January 2000 to July 2002, we compute the daily recovery rates for each of 3,130 firms. After forming decile-based portfolios, we extract the principal components. We find that there are two main components that explain the variation in recovery rates over time, and we are able to identify these components with the interest rate term structure (accounting for 78.9% of the common variation) and equity market volatility (8.2%) respectively. In a regression framework, market variables are able to explain the variation in the first and second components very well, with  $R^2$ s of 90.7% and 35.7% respectively. A similar decomposition for default probabilities results in four factors accounting for more than 90% of common variation. The first two factors are again identified with the level of interest rates (accounting for 41.9% of common variation) and equity market volatility (30.4%), and regressions containing market variables also support these factors.

The goals of this paper are methodological and empirical; and we leave a rich menu

Table 5: Explaining the principal components of default probabilities using market variables. We report the results of regressions of the first four principal components, our dependent variables on various independent variables for the period January 2000 to July 2002. The risk free rate is the one-month rate taken from the Fama French database, along with the excess market return, small minus big (SMB), high minus low (HML) and up minus down (UMD). The annualized percentage implied volatility of the S&P500 (VIX) is from Yahoo Finance. The inflation rate is the percentage change in CPI. The CPI index is the consumer price index. The standard errors and t-statistics that are reported are corrected for heteroskedasticity using the White (1980) approach. T-statistics are reported below the coefficient values. “PCn” stands for the  $n$ -th principal component.

	<i>PC1</i>	<i>PC2</i>	<i>PC3</i>	<i>PC4</i>
<i>Intercept</i>	0.0017	-0.0197	-0.0027	-0.0010
	0.41	-8.69	-1.26	-0.45
<i>Risk Free Rate</i>	1.1503	1.6770	-0.3669	0.0071
	3.19	11.91	-2.07	0.04
<i>Inflation Rate</i>	0.1130	-0.0140	-0.1789	-0.0470
	0.49	-0.20	-2.39	-0.54
<i>Implied Volatility of S&amp;P 500 (VIX)</i>	-0.0002	0.0006	0.0002	0.0000
	-1.32	6.81	2.57	0.41
<i>Excess Market Return</i>	0.0009	0.0274	0.0130	-0.0143
	0.05	3.96	1.61	-1.78
<i>High-Minus-Low</i>	-0.0233	0.0147	-0.0167	-0.0031
	-1.54	2.34	-2.40	-0.50
<i>Small-Minus-Big</i>	-0.0201	0.0134	-0.0071	0.0065
	-1.38	3.11	-1.45	1.38
<i>Up-Minus-Down</i>	0.0106	0.0016	-0.0020	-0.0041
	0.98	0.66	-0.60	-1.70
<i>R</i> <sup>2</sup>	28.4%	86.8%	62.9%	22.3%
<i>N</i>	31	31	31	31

of extensions for follow up work. For one, the credit default swap market is still quite new, and good data has only recently become available, circa 2000. To draw more extensive conclusions and to study the dynamics of recovery rates, will require a longer time series than we have available. It may be feasible to undertake the same study with a time series of credit spreads from the bond market. Two, using a cross-sectional panel of firms, we may study the determinants of recovery rates using firm-level and industry-level variables, complementing the work of Acharya, Bharath and Srinivasan (2003), Chava and Jarrow (2004), and Chava, Stefanescu and Turnbull (2006). Three, using a different functional form than the Merton model, we may apply the technique to sovereign debt in work complementary to that of Pan and Singleton (2005). Four, we may extend this approach to relationships between default and recovery where the process for recovery is triggered by the default event as in Guo, Jarrow and Zeng (2005). The methodology developed here to extract forward recovery term structures will provide financial economists with a better understanding of loss given default, and provide traders with better estimates of recovery rates for credit derivative models. It will also enable regulators to set loss given default (LGD) levels more precisely

for the implementation of Basel II (see Schuermann (2004)).

## References

- Acharya, Viral., Sreedhar Bharath, and Anand Srinivasan (2003). Does Industry-wide Distress Affect Defaulted Firms? - Evidence from Creditor Recoveries, working paper, London Business School.
- Altman, Edward., and G. Fanjul (2004). "Defaults and Returns in the High-Yield Bond Market: Analysis through 2003," NYU Salomon Center, working paper.
- Altman, Edward., Andrea Resti, and Andrea Sironi (2005). "Default Recovery rates in Credit Risk Modeling: A Review of the Literature and Empirical Evidence," *Journal of Finance Literature* 1, 21-45.
- Altman, Edward., Brooks Brady, Andrea Resti, and Andrea Sironi (2005). "The Link between Default and Recovery Rates: Theory, Empirical Evidence and Implications," *Journal of Business* 78, 2203-2228.
- Bakshi, Gurdip., Dilip Madan, and Frank Zhang (2001). "Recovery in Default Risk Modeling: Theoretical Foundations and Empirical Applications," working paper, University of Maryland.
- Berd Arthur, M. (2005). "Recovery Swaps," *Journal of Credit Risk* 1(3), 1-10.
- Chan-Lau, Jorge A. (2003). "Anticipating Credit Events using Credit Default Swaps, with an Application to Sovereign Debt Crisis," IMF working paper.
- Chava, Sudheer., and Robert Jarrow (2004). "Bankruptcy Prediction with Industry Effects," *Review of Finance* 8, 537-569.
- Chava, Sudheer., C. Stefanescu, and Stuart M. Turnbull (2006). "Modeling Unexpected Loss with Unobservable Heterogeneity," working paper, University of Houston.
- Christensen, Jens (2005). "Joint Estimation of Default and Recovery Risk: A Simulation Study," Working paper, Copenhagen Business School.
- Das, Sanjiv., and Rangarajan Sundaram (2002). "A Simple Model for Pricing Securities with Equity, Interest-Rate, and Default Risk," working paper, Santa Clara University.
- Das, Sanjiv., Laurence Freed, Gary Geng, and Nikunj Kapadia (2001). "Correlated Default Risk," working paper, Santa Clara University.
- Das, Sanjiv., Darrell Duffie, Nikunj Kapadia, and Leandro Saita (2004). "Common Failings: How Corporate Defaults are Correlated," forthcoming *Journal of Finance*.
- Duffie, Darrell (1999). "Credit Swap Valuation," *Financial Analysts Journal*, January-February, 73-87.
- Duffie, D., and K. Singleton (1999) "Modeling Term Structures of Defaultable Bonds," *Review of Financial Studies* 12, pp. 687-720.

- Duffie, Darrell, Saita, Leandro, and Ke Wang (2005). "Multi-Period Corporate Default Prediction with Stochastic Covariates," Working Paper, Stanford University.
- Finger, Christopher., Vladimir Finkelstein, Jean-Pierre Lardy, George Pan, Thomas Ta, and John Tierney (2002). "CreditGrades Technical Document", RiskMetrics.
- Guo, Xin., Robert Jarrow, and Yan Zeng (2005). "Modeling the Recovery Rate in a Reduced Form Model," Working paper, Cornell University.
- Jarrow, R.A., and S.M. Turnbull (1995). "Pricing Derivatives on Financial Securities Subject to Credit Risk", *Journal of Finance* 50(1), 53-85.
- Jarrow, R.A., David Lando, and S.M. Turnbull (1997). "A Markov model for the term structure of credit risk spreads," *Review of Financial Studies* 10, 481-523.
- Karoui, Lotfi (2005). "Modeling the Term Structure of Defaultable Bonds under Recovery Risk," Working paper, McGill University.
- Leland, Hayne E. (1994). "Corporate Debt Value, Bond Covenants and Optimal Capital Structure," *Journal of Finance* 49(4), 1213-1252.
- Leland, Hayne., and Klaus Toft (1996). "Optimal Capital Structure, Endogenous Bankruptcy, and the Term Structure of Credit Spreads," *Journal of Finance* 51(3), 987-1019.
- Longstaff, Francis A., and Eduardo Schwartz (1995). "A Simple Approach to Valuing Risky Fixed and Floating Rate Debt," *Journal of Finance* 50(3), 789-819.
- Madan, Dilip., and Haluk Unal (1995). "Pricing the Risks of Default," *Review of Derivatives Research* 2, 121-160.
- Merton, Robert C. (1974). "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates," *The Journal of Finance* 29, 449-470.
- Pan, Jun., and Ken Singleton (2005). "Default and Recovery Implicit in the Term Structure of Sovereign CDS Spreads," working paper, Stanford University.
- Schuermann, Til (2004). "What do we know about Loss Given Default?" Working paper, Federal Reserve Bank of New York, forthcoming in *Credit Risk Models and Management*, 2nd Edition.
- Sobehart, J., R. Stein, V. Mikityanskaya, and L. Li, (2000). "Moody's Public Firm Risk Model: A Hybrid Approach To Modeling Short Term Default Risk," Moody's Investors Service, Global Credit Research, Rating Methodology, March.
- Trück, Stefan., Stefan Harpaiter, and Svetlozar Rachev (2005). "A Note on Forecasting Aggregate Recovery Rates with Macroeconomic Variables," working paper, Institut für Statistik und Mathematische Wirtschaftstheorie, Universität Karlsruhe.

White, Halbert, (1980). “A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity”, *Econometrica* 48(4), 817-838.

Zhang, Frank (2003). “What Did the Credit Market Expect of Argentina Default? Evidence from Default Swap Data,” working paper, Federal Reserve Board.