CORRELATED DEFAULT RISK

Sanjiv R. Das, Laurence Freed,
Gary Geng, Nikunj Kapadia

Date: August 2002.

Key words and phrases. Correlated default, hazard rates, credit risk.

Sanjiv Das is at Santa Clara University, Laurence Freed is at Moody's Investors Services, Gary Geng is at Paloma Partners, CT., and Nikunj Kapadia is at the University of Massachusetts, Amherst. We are extremely thankful for many constructive suggestions from Gurdip Bakshi, N. Chidambaran, Rong Fan, Gifford Fong, John Knight, N. R. Prabhala, Jun Pan, Dmitry Pugachevsky, Shuyan Qi, Rangarajan Sundaram and Suresh Sundaresan. Participants at various seminars also provided useful feedback: at the AIMR talks in Tokyo, Singapore and Sydney, the AIMR Research Foundation Workshop in Toronto, Risk Conference in Boston, BGI in San Francisco, the Credit Conference at Carnegie-Mellon University, Institutional Investors Fixed Income Forum in Boston. We are especially grateful to Darrell Duffie and Ken Singleton, whose insights into this work resulted in many improvements and additional analyses. The first author gratefully acknowledges support from the Price Waterhouse Cooper's Risk Institute, the Dean Witter Foundation, and a Research Grant from Santa Clara University. We are also grateful to Gifford Fong Associates, and Moody's Investors Services for data and research support for this paper. Please address all correspondence to Professor Sanjiv Das, Dean Witter Fellow & Associate Professor, Santa Clara University, Leavey School of Business, Dept of Finance, 208 Kenna Hall, Santa Clara, CA 95053-0388. Email: `srdas@scu.edu`.

Correlated Default Risk

Abstract

Using a comprehensive and unique data set of default probabilities from Moody's, we examine correlations between default risk for over 7,000 US public firms. This is the first paper to empirically document the correlation structure both in the time-series and in the cross-section across almost all US non-financial firms. We find that default probabilities of issuers vary substantially over time, and are positively correlated. Moreover, the correlations across firms also varies over time systematically, in a manner that is related to an economy-wide level of default risk. Joint default risk increases as the default risk in the economy increases. Our results also suggest that the magnitude of joint default depends on the quality of issuers; highest quality issuers have higher default correlations than medium grade firms. The results in this paper are informative for managers of credit portfolios, and for structured products such as collateralized debt obligations.

1. Introduction

One of the most important open questions in financial markets concerns the behavior of *default correlation*, that is, the manner in which across-issuer likelihood of default co-varies over time. An understanding of this correlation is of obvious importance to banks, traders, and other operatives in corporate-debt markets from the standpoint of overall portfolio management as well as for the pricing/hedging of such instruments as Collateralized Debt Obligations (CDOs). It is also of considerable interest to regulators; indeed, the Basel Committee on Banking Supervision has repeatedly urged a move towards a system in which the capital adequacy requirements of banks are based on the overall credit-riskiness of the banks' aggregate *portfolios*. Yet, in contrast to the substantial theoretical and empirical literature on default risk at the level of the *individual* issuer,[1] the research on default correlations is virtually empty. It is not even known whether joint default risk is affected by systematic factors.

Our paper takes the first substantial step in this direction. We provide a comprehensive empirical investigation of default correlations for over 7,000 US firms, using a database of issuer-level default probabilities from Moody's. This database, covering the period 1987-2000, provides a unique opportunity to investigate joint default risk across almost all US non-financial public firms. Our approach of using firm-level default probabilities enables a comprehensive investigation of joint default risk in both the cross-section and time series, and allows us to account for both time-varying default probabilities and correlations between defaults.

Our analysis confirms several intuitive features one might associate with the behavior of default risk such as its variation over the business cycle. More importantly from the standpoint of using the results, it provides *quantitative* expression to this behavior. We also identify several new and important features of default correlation that have not received much comment in the literature, such as, for example, the differences between the behavior of default correlation in high-rated firms and that in lower-rated firms. In particular, we uncover evidence that default risk has a systematic component. A more detailed description of our main findings follows.

First, confirming intuition, we show that default probabilities of individual firms of a given rating class vary substantially over time. This time-variation is linked directly to the economic environment. Default probabilities are low, on average, in the benign credit

[1]For theoretical modeling of default and credit risk, see the structural models of Merton (1974), Longstaff and Schwartz (1995), Collin-Dufresne and Goldstein (2001), and reduced form models of Duffie and Singleton (1999), Jarrow and Turnbull (1995), Das and Tufano (1996), Jarrow, Lando and Turnbull (1997), Madan and Unal (1998), and Das and Sundaram (2000), among others.

environment of 1993-97 and show steep increases in the periods corresponding to the 1991-92 recession, as well as the 1999-2000 period that foreshadowed the most recent recession. For example, the median default probability of Baa and Ba firms increases from 0.77% in the period 1/94-4/97 to 1.17% in the period 6/97-10/00. Default probabilities of higher quality firms show an even steeper increase.

Second, we provide extensive evidence that there are systematic components to default risk. Specifically, we find that correlations between changes in default intensities are non-negative on average even after autocorrelation in default intensities is taken into account. We show that this conclusion is robust across time-periods and sub-groups of firms. Principal component analysis confirms the presence of a large first principal component; for example, over the most recent sub-period in our sample, the first principal component explains more than 50% of the variation in changes in default intensities for high grade firms. The first two principal components combined explain a substantial portion of the variation across all sub-periods and rating classifications.

Third, we find that correlations between default intensities of individual firms are not stable over time, but vary in tune with the business cycle. Correlations increase with the level of default risk in the economy. Thus, joint default risk is directly linked to the business cycle, and this evidence provides additional support for the finding that default risk is not fully diversifiable. In addition, that correlations are not stable over time is of particular importance for the portfolio modeling of credit risk.

Fourth, we document *cross-sectional* differences across rating classes. We find that firms of the highest credit quality have the highest correlations, and that this is especially so in time periods when the overall level of default is high. This empirical finding indicates that firms of high credit quality are especially exposed to an economy wide factor. Moreover, an analysis of the principal components affecting default intensities suggests structural differences across quality groups. From a credit-risk modeling standpoint, this implies it may be necessary to fit reduced form models separately across rating classes.[2]

Fifth, to further understand how the systematic component of default affects joint default risk, we model time-varying default probabilities and correlations jointly in a parsimonious statistical model. To account for our observation that default probabilities vary with the business cycle, we allow the economy wide default risk to be regime-dependent. We find strong support for a two-regime model (a high default regime and a low default regime) with the former having a mean default level more than twice that of the latter. Importantly,

[2]This finding complements that of Duffee (1999). Duffee finds that estimated parameters of risk-neutral default intensity processes are unstable across rating classes.

we find that each regime shows a different correlation structure: default correlations are, on average, higher in the high default regime as compared with the low default regime.

Our results, in summary, establish that both default probabilities and default correlations time-vary in relation with an economy-wide factor and provide a quantitative description of the extent of this time-variation. These results provide testable implications for the dynamics of credit spreads where systematic default risk is priced (Jarrow, Lando and Yu (2001), El-Karoui and Martellini (2002)). First, credit spreads change over time because changes in default probabilities affect expected credit loss. Second, credit spreads may also vary because variation in the systematic component of default intensities may affect the risk premia in credit spreads. Because correlations between defaults increase with the level of default risk, both expected loss and risk premia increase simultaneously. Thus, credit spreads will systematically change more than would be indicated by changes in expected loss. The presence of time-varying correlation linked to an economy-wide factor may potentially explain why it has been difficult to relate structural models to changes in credit spreads (Collin-Dufresne, Goldstein and Martin (2001)).

Our findings have direct implications for the management of default risk on a portfolio basis, as well as the pricing of Collaterized Debt Obligations and other securities whose cash flows are determined by risk at the portfolio level. In particular, our findings indicate that clustering of defaults in a portfolio during a downturn in the economy occurs because of an increase in default probabilities *and* an increase in correlation between defaults. Ignoring time-variation of correlations may significantly understate portfolio default risk. That the correlation structure varies over time also implies that a non-linear model is required to describe evolution and pricing of default risk across firms (for example, see recent work by Collin-Dufresne, Goldstein and Helwege (2002)).

Our analysis complements that of Zhou (2001). Within the context of a structural model of default, Zhou (2001) considers the implications of correlated asset returns on correlations between defaults. Our empirical findings are consistent with his conclusions that default correlations are positive and time-varying. In addition, our empirical finding that default correlations time-vary with an economy wide factor is potentially consistent with structural models that explicitly account for time-varying and asymmetric asset correlations.

The rest of this paper is as follows. In Section 2, we provide the framework for our empirical work. Section 3 describes Moody's database of default probabilities. In Section 4, we investigate correlation between default intensities both in the time-series and cross-section. Section 5 shows that a parsimonious regime-shifting model is useful for jointly describing time-variation in default intensities and time-variation in correlations between default intensities. Section 6 examines the issue of conditional independence of default times. The last section offers conclusions and extensions.

2. Understanding Default Correlations

In this section, we present a framework for our empirical investigation. For a firm i, we define a stopping time τ_i. For a given horizon T, default occurs if the stopping time $\tau_i \leq T$. We define the probability of default, $PD_i(t,T)$, at time t as,

$$PD_i(t,T) = Pr[\tau_i \leq T]. \tag{1}$$

With appropriate assumptions, this setup is consistent both with reduced form, hazard rate models (Jarrow and Turnbull (1995), Madan and Unal (1998), Duffie and Singleton (1999)), as well as with structural models of Merton (1974), Longstaff and Schwartz (1995), and Collin-Dufresne and Goldstein (2001). In the former, the probability of default is determined by an exogenously specified instantaneous default intensity or hazard rate, $\lambda_i(t)$. Given the stochastic process followed by $\lambda_i(t)$, $PD_i(t,T)$ can be calculated as

$$PD_i(t,T) = E_t\left[1 - \exp\left(-\int_t^T \lambda_i(\tau)d\tau\right)\right], \tag{2}$$

where E_t is the expectations operator under the physical measure. In the context of structural models, the value of the firm is modeled as a stochastic process, and $PD_i(t,T)$ can be derived as an output of the model. In reduced form models, $PD_i(t,T)$ and $\lambda_i(t)$ are directly the primitive modeling variables. Below, we follow the spirit of the reduced form literature. Thus, view the default process as a Cox or doubly stochastic Poisson process [see Lando (1994)] in the following sense. First, the default intensity, $\lambda_i(t)$, is a stochastic process. Second, conditioned on the level of default intensity, the default event itself is a random event. The process for $\lambda_i(t)$ provides a complete description of the default risk for firm i.

Given two firms, $i = 1, 2$, our object of interest is the probability of joint default, $Pr(\tau_1 \leq T, \tau_2 \leq T)$. The doubly stochastic assumption implies that defaults are independent given the time t default intensity. In this case, the probability of joint default will be determined by the correlation structure of the default intensities, $\lambda_i(t)$. Define the correlation between shocks to $\lambda_1(t)$ and $\lambda_2(t)$ as ρ_{12}. Given an estimate of ρ_{12}, and a model for the default intensity processes, we can compute $Pr(\tau_1 \leq T, \tau_2 \leq T)$ either analytically or numerically. As an illustration, suppose the joint dynamics of $\lambda_i(t)$, $i = 1, 2$, are defined by,

$$\lambda_i(t) = \alpha_i + \beta_i \lambda_i(t-1) + e_i(t),$$

where $\text{corr}(e_1, e_2) = \rho_{12}$. We can now estimate $Pr(\tau_1 \leq T, \tau_2 \leq T)$, using a Monte Carlo simulation with $\lambda_i(t)$ as the stochastic jump arrival intensity of a Poisson process whose first jump occurs at default. For example, let $\alpha_1 = \alpha_2 = 0.012\%$, $\beta_1 = \beta_2 = 0.94$ and $\lambda_1(0) = \lambda_2(0) = 0.2\%$, (calibrating to recently observed parameters of high grade firms, see our Tables 2 and 5 below). For $\rho_{12} \in (-0.6, -0.2, 0.0, 0.2, 0.6)$, probability of joint default over a 5 year horizon, $Pr(\tau_1 \leq 5, \tau_2 \leq 5)$, is 0.64%, 1.56%, 1.80%, 2.36% and 2.96%, respectively. The correlation between default intensities has a substantial effect on the joint probability of default.[3]

Default intensities may be correlated because of the presence of systematic factors affecting the probability of default. For example, in Duffee (1999), the risk-free rate is the systematic factor affecting default intensities of all firms. In addition to having the risk-free rate, Bakshi, Madan and Zhang (2001) also include the long-term mean of the risk-free rate as an additional factor. Additional systematic factors that drive the value of the firm and its volatility will also affect default intensities.

We summarize below the implications of existing theoretical and empirical work for ρ_{ij}.

(1) *Default correlations are positive.* Existing theoretical and empirical research suggests that most factors that are determinants of default probabilities are positively correlated, including asset values and asset volatilities. Moreover, changes in debt-equity ratios are also correlated across firms in the economy. This suggests that we should expect default correlations to be positive. Zhou (2001) provides an example of a structural model where positive correlations between asset values imply positive correlation of defaults. Empirical evidence that credit spreads across bonds are positively correlated is also consistent with a positive correlation between default intensities under the physical measure.

(2) *Default correlations will vary in the cross-section.* For the same level of asset correlation, Zhou (2001) notes that firms with a higher debt level are more highly correlated. Controlling for debt level, cross-sectional variation will, however, be related to the systematic factors affecting the default intensities. In the absence of evidence that indicates default intensities are equally sensitive to systematic factors across the range of firms, especially within distinct credit classes, it appears appropriate to hypothesize that default correlations vary in the cross-section.

[3]The Monte Carlo simulation, done with 10,000 simulations, assumes that defaults are independent when conditioned on λ_t (also see Das, Fong and Geng (2001)). This assumption is not necessary for the analysis of the paper. If defaults are not independent, joint default risk will be affected by the actual incidence of defaults, in addition to the correlation structure between the default intensity processes.

(3) *Default correlations are time-varying.* Within the class of structural models, for fixed credit quality, time-variation in asset correlations and time-variation in asset volatilities both imply that default correlations will also vary. In particular, it has been noted that changes in asset correlations are related to changes in asset returns. Ang and Chen (2001) show that equity correlations for negative equity return are larger than for positive returns, and especially so for large negative returns. Because default probabilities increase with decline in asset values, we expect to observe variation in default correlation that is linked to the level of the default probabilities, or to the state of the economy.

The primary focus of this paper is to investigate the presence and properties of the correlation structure between default intensities. In particular, we are interested in understanding any systematic risk affecting default intensities.

3. The Data: Default Probabilities

This research takes advantage of the unique data resources of Moody's Investors Service. Moody's, through its subsidiary Moody's Risk Management Services ("MRMS"), collects and maintains extensive databases of default events, default recoveries and other related information on rated and unrated, public and private, companies. MRMS regularly publishes default and recovery rate studies based on this database. In particular, MRMS uses this data to calibrate a model of short-term default risk for both public and private firms. The output from their *RiskCalc* model is an estimate of a one-year default probability for an individual issuer.

Details of the RiskCalc model and its econometric fit are described in Falkenstein (2000), Sobehart, Stein, Mikityanskaya and Li (2000), and Falkenstein, Ibarra, Kocagil, and Sobehart (2001). The model uses as input firm-specific information: (a) company financial statement information (including leverage, profitability, and liquidity measures), (b) equity market information, in particular, equity volatility, and (c) Moody's ratings, if available.

The data used in this paper consists of a monthly time series of probabilities of default ("PDs") for US non-financial public firms over the period January 1987 to October 2000. Although the potential universe may include as many as 12,000 firms, the actual number of firms for which a PD is available is much smaller and varies by observation date, typically because of insufficient data, defaults or mergers. We employ two sample sets of firms in our empirical work. First, we construct a sample of firms with continuous data over the entire sample period. This results in a sample of 1705 firms. Second, we sub-divide our sample period into four sub-periods, and include all firms that have continuous observations within any of these sub-periods. The choice of four sub-periods is motivated by the trade-off between having sufficient time series observations for estimation of correlation matrices, yet ensuring that the period is relatively homogeneous, so that differences in correlation matrices across sub-periods may be observed. The sub-periods are of almost equal length, and comprise the periods, 1/87-4/90, 5/90-12/93, 1/94-4/97, and 5/97-10/00 (the first three periods are of 42 months and the last of 40 months). Finally, allowing all firms with continuous data in any sub-period to be included in the sample reduces a bias towards higher quality firms, as higher rated firms are more likely to have continuous observations of default probabilities. The total number of unique firms included in this sample is 7,363.

For our primary analysis, we group firms by credit rating. Each firm that has a rating is classified by Moody's into one of 6 rating classes, in order of descending credit quality. In addition, there is a seventh rating class for firms that are unrated, but yet have a Moody's PD. For tractability, we divide all rated firms into three groups: High Grade, Medium Grade, and Low Grade, each comprising of two rating classes. Thus, high grade firms combine rating classes 1 and 2, medium grade combine rating classes 3 and 4, and low grade combine rating classes 5 and 6. A firm is classified into a rating class according to its average rating over the sub-period. Panel A of Table 1 provides the description of the classification by rating class. For robustness of some of the results, we also form an alternative grouping by industry sector. There are ten sectors, classified by SIC codes. These are described in Panel B of Table 1.

Table 2 provides the descriptive statistics of the cross-sectional distribution of mean default probability for the firms in our sample, classified by credit rating and industry sector. The total number of unique firms in our sample range from 3,202 to 5,170 over the sub-periods. The average number of firms in the high, medium and low rated groups are 215, 405 and 184, respectively. As the universe of firms rated by Moody's is relatively small, the vast majority of the firms in our sample are unrated, ranging from 2724 in sub-period I to 4142 in sub-period IV. As might be expected, the mean PD increases monotonically as the average rating declines, across each sub-period. For instance, the mean PD in sub-period IV is 0.23%, 1.17% and 5.65% for high, medium and low grade firms, respectively. The mean PD of unrated firms is in the range of 1.63% to 2.45%, higher than the average PD of medium-rate firms, indicating that if these firms were rated, they are likely to be in lower rating groups.

When classified by industry, more than half the firms in our sample belong to the manufacturing sector, Sector 4. The sector with the least number of firms is Sector 1 (agriculture, forestry and fishing) with as few as 14 firms in sub-period II. Although there is variation across PDs within the cross-section, this variation is small compared with that across credit classes, indicating that each sector has a mix of rating classes, with a substantial number of firms in the lower rating categories. Firms in sectors 1 and 5 have the lowest average PD, while firms in sector 10 have the highest. The high default risk of sector 10 arises from firms of SIC group 99 (firms of industries that cannot be classified into any of the other SIC groups).

Comparing PDs across sub-periods, we see that there is considerable time-variation. On average, PDs are the lowest in sub-period III, and highest in sub-period IV. This pattern is remarkably consistent across every classification (the only exception being sector 10), suggesting that the time-variation in PDs is driven by a common economic factor that affects all firms. Figure 1 plots the monthly time series of default probability, averaged over all firms that had data available for that particular month in our sample period. The figure confirms the observation from Table 2 that PDs exhibit time-variation across the average firm in the economy. Moreover, PDs seem to lie in two distinct regimes, i.e., for some periods, they tend to be low, and in others, they spike up dramatically. Casual observation suggests that the spikes can be linked to economic events. The sharp increase in PDs in sub-period I corresponds to the aftermath of the 1987 stock market crash, and that in sub-period II corresponds to the 1990-91 recession. Finally, the sharp increase in sub-period IV can be linked to the increased volatility of the equity markets and subsequent downturn in both equity values and economic growth.

How does the mean default probability relate to the actual number of defaults in the economy? Over our sample period, data from Moody's indicates that there were a total of 831 individual issuer defaults, with the number of defaults in each month of our sample ranging from zero to a maximum of 19. Figure (2) plots the mean default probability against the number of actual defaults in the economy for each month of our sample period. It is evident that the mean default probability is related to the number of defaults in the economy. More formally, in a Poisson regression, we consider the mean default probability in month $t-1$ as an explanatory variable for defaults in month t. We estimate the model, $\ln \mathrm{E}[NUMDEF(t)] = \beta_0 + \beta_1 MEANPD(t-1) + \beta_2 NUMDEF(t-1)$, where $NUMDEF(t)$ represents the number of defaults and $MEANPD(t)$ the mean default probability in month t, respectively. The lagged value of $NUMDEF$ is included as a proxy for the business cycle, and provides a simple alternative model to evaluate the importance of $MEANPD$ in determining defaults. The estimated coefficient β_1 is significant with a p-value of 0.00, with the model having a reasonable fit with a pseudo-R^2 of 20%. A likelihood ratio test for the hypothesis that $\beta_1 = 0$ is likewise rejected with a p-value of 0. Thus, at the aggregate economy level, there is evidence that the mean default probability helps explain the incidence of defaults.

We use the time-series of PDs to estimate correlations between default risk. As in previous research, we first transform the 1-year default probability for firm i to the corresponding "default intensity", λ_i,

$$\lambda_i = -\ln(1 - PD_i), \tag{3}$$

where λ_i is interpreted as the constant default intensity that corresponds to a 1-year default probability, $PD_i \in [0, 1)$. Given the observed time-series of default intensities, $\lambda_i(t)$, the objective is to infer the correlations between unexpected *changes* in $\lambda_i(t)$ and $\lambda_j(t)$. Here, we make a distinction between correlation of levels of the default intensity, and unexpected changes in the levels of the default intensity. The former includes any known dynamics; for instance, if there is "stickiness" in ratings over a business cycle, it would get reflected in correlation between levels of default intensities. The latter reflects any co-movement that remains after modeling the time-series behavior of the default intensity. This allows direct comparison of our results to reduced form models.

4. Default Correlations

In this section, we examine both the cross-sectional and time series properties of the correlation between default intensities. We also test the main implications of structural and reduced form models of default - that default correlations are likely to be, on average, positive, time-varying and different in the cross-section.

4.1. **Correlations in Default Risk.**

We begin by examining how correlations between default probabilities vary within sub-groups of firms, where the expected default intensity is constant over the time-period. Over each of the four sub-periods, we estimate the correlation matrix by modeling the default probability as

$$\lambda_i(t) = \bar{\lambda}_i + \epsilon_i(t), \tag{4}$$

where $\bar{\lambda}_i$ is the average default intensity over the sample period. The correlation between the default intensities of firm i and j, ρ_{ij}, is then computed as the correlation between ϵ_i and ϵ_j over this period.

Table 3 reports the average of the pair-wise correlations of firms within each industry sector and credit class, over each sub-period of the sample period. For n firms, the average correlation is calculated as the arithmetic average over the $n(n-1)/2$ pair-wise correlations. A firm is included in each sub-period only if it has a complete time-series of observations over this sub-period. Firms that do not have a Moody's rating are not included in the classification by rating class.

Panel A reports the statistics for the quality groups. Three observations can be immediately made. First, across all credit classes and in every period, default correlations are, on average, positive. The mean correlation ranges from 0.00 to 0.62 across periods and credit classes. The table also reports the fraction of the estimated pair-wise correlations that are positive. The results mirror those of the mean; except for sub-period III, the fraction of positive correlations exceed 0.50 by a large margin. There are also clear cross-sectional differences across the credit classes. In particular, averaging over the sub-periods, the highest correlation amongst the rating classes is between firms with the highest ratings. Except for sub-period III, the positive correlations are consistent with the main implication of structural models that positive correlations between firms in the economy is likely to result in positive correlations between default probabilities.

Second, the average correlation within each group varies over the sub-periods. For the high and medium grade firms, there appears to be a clear pattern, correlations decline from period I to III and then increase in period IV. In particular, across all groups, period III has the lowest average correlation, which is close to zero. From the previous section, recall that period III is also the period with the lowest average default intensity. We will investigate this further below. Overall, these results suggest that there may be commonality in the factors that affects correlations across the groups.

Finally, there are cross-sectional differences in the time-variation in correlation across credit classes. The median correlation ranges from 0.05 to 0.58 for high rated firms, showing a steep decline in sub-period III and a sharp subsequent increase. In contrast, the range of correlations is 0.02-0.35 and 0.00-0.29 for medium and low rated firms, respectively. High rated firms are have the highest correlations on average, and show the greatest time-variation.

In all of the above results, firms were classified by rating. Although results are likely to be less sharp when we compare firms of potentially very different rating (and leverage) groups, such results can still serve as a robustness test. We, therefore, classify all firms into industry sectors by SIC code. Here, we also include non-rated firms. Panel B of Table 3 reports the results. Average correlations are predominantly positive; across all sub-groups and sub-periods, there is only one instance where the median correlation is negative. Correlations are time-varying; moreover, in 8 of the 9 groups, the minimum average correlation is in period III. The exception is sector 10. Although, on average, sectors differ in the average correlation, in period III, the difference is minimal. For example, the range of the median correlations across all sectors is 0.15 to 0.54 for sub-period I, and -0.04 to 0.08 for sub-period III. The two sectors with the highest median correlation are agriculture (sector 1) and the transportation sector (sector 5). Not surprisingly, firms in sector 9, that is populated mainly with firms of SIC code 99 (non-classifiable industries) have the least correlation, and show the least time-variation.

Overall, these results indicate that default correlations are positive, time-varying in a systematic manner, and different across the cross-section. Interestingly, high quality firms have the highest correlations. Although, correlations are time-varying for all firms, there are cross-sectional differences; firms of the highest quality exhibit the greatest time-variation. The period where the median correlation between default probabilities across all rating classes is low (close to zero) also corresponds to a period where default probabilities are also low. In the sections that follow, we investigate these results in more detail, check robustness of results to modeling assumptions, and suggest a means of modeling time-variation in correlations.

4.2. **Correlations with Dynamic Default Probabilities.** The correlations in the previous section were computed under the assumption that there is no time-variation in default probabilities. Are the conclusions of the previous section robust when we allow for time-varying default probabilities? Following Duffie and Singleton (1999) and the empirical evidence in Duffee (1999), we model the default intensity as a discrete time AR(1) process,

$$\lambda_i(t) = \alpha_i + \beta_i \lambda_i(t-1) + \tilde{\epsilon}_i(t). \tag{5}$$

The correlation between firms i and j is the correlation between $\tilde{\epsilon}_i$ and $\tilde{\epsilon}_j$, and takes into account the dynamics of the default intensity.

Table 4 reports a summary of the parameter estimates for equation (5) across each of the 4 sub-periods, and all firms in the sample set. The median estimate of β_i across all periods and rating classes is in the range of 0.90 to 0.94, with over 99% of the estimates significant. The variation in α_i in both the cross-section and time-series is similar to the variation in the average PD in Table 2.

Table 5 reports the median correlations across each classification over the four sub-periods, as well as the fraction of correlations that are positive. Panel A reports the results for classification by rating, and Panel B for classification by sector. Comparing with Table 3, we can observe that the median correlation for each group is lower. For instance, in sub-period I, the median correlation across high-grade firms reduces from 0.54 to 0.37, for medium grade-firms from 0.35 to 0.23, and for low grade firms from 0.17 to 0.16. The same conclusion holds across all sector classes; for example, the median correlation for firms in the manufacturing sector (Sector 5) reduces from 0.42 to 0.13.

Although the magnitudes of the correlation are different, the variation in the median correlation in the cross-section and time-series is consistent with that observed earlier in

Table 3. Default correlations are predominantly positive. For the rating groups, the median correlation is in the range of 0.01 to 0.37. Similarly, the fraction of positive correlations is in the range of 52% to 89%. Results across industry sectors are similar, although of much lower magnitude.

Second, the median correlation time-varies over the sub-periods; in particular, it is much lower in sub-period III, as compared with sub-period I or IV. Third, the cross-sectional variation across groups is lowest in period III. For instance, the cross-sectional variation in median correlations across rating groups in sub-period I ranges from 0.16 to 0.37. In contrast, the range in sub-period III is 0.00 to 0.02. The same conclusion can also be drawn from the industry sectors.

The results of both Table 3 and Table 5 raise two economic questions. First, what can explain the pattern of time-variation in correlation, in particular, the coincidence of low default intensities with low correlations between default intensities in sub-period III? A possible explanation may stem from the observation that equity correlations are asymmetric (Longin and Solnik (2001) and Ang and Chen (2001)), which we consider in more detail in the sections that follows. A second question that arises is whether the cross-sectional variation across credit classes implies structural differences in how underlying factors affect default intensities within each credit class. This hypothesis is also motivated by the observation that correlations for firms within industry sectors are consistently lower (as sectors combine firms of different rating classes). We investigate this question of structural difference across rating classes in more detail below.

Consider an analysis of the principal components of the residuals in equation (5). Table 6 reports the fraction of the variance of the residual, explained by the first, two, and four principal components respectively. Results are reported for all credit classes, and, as a robustness test, for 5 of the 9 industry sectors.

The results confirm the earlier observations. First, a few principal components explain a substantial portion of the variance of the residual variance. Across sub-period I, the first two factors explain 95.20%, 29.45%, 34.98% of the variance for high, medium and low grade bonds, respectively. Similarly, across the sectors, the first two factors explain from 22.78% (for sector 9) to 91.49% (for sector 1). There is clear evidence of common factors affecting default intensities for each credit class. Second, the effect of these common components is time-varying. For example, for high grade firms, the amount of variance explained varies from 27.08% in sub-period III to 95.20% in sub-period I. Finally, for 5 of the 8 groups, the fraction of variance explained by the first few principal components is the lowest in sub-period III.

Figure 3 graphs the fraction of the variance explained in $\tilde{\epsilon}_i$ by up to 10 principal components for the three credit classes. Results are shown across each of the four sub-periods. Both Figure 3 and Table 5 indicate that there are substantial differences in correlation structure across rating classes. In particular, there is a substantial difference in correlations between high grade firms compared with the others, and this is more marked in periods I and IV.[4] To verify if this is consistent with differing factors across rating classes, we compute the correlation matrices for the first principal component for each of the three grades of firms. The results are presented in Table 7. The correlation coefficients between the first principal factor across rating classes shows wide variation over the periods. In periods II and III, all three grades have their principal components positively correlated, but the magnitude of correlation is much lower in period III. In periods I and IV, the first principal component of the lowest rating class is negatively correlated with that of the highest rating class. The first principal component of medium rated firms also shows unstable correlations with that of both the other rating categories, being positive in only three of the four periods. For robustness, we also considered the correlation between the second and third principal components of each rating class with similar results.

The results indicate that, across the quality groups, there are structural differences in how underlying factors affect default intensities. This finding has implications for the relative spreads between high grade and other firms. In particular, it suggests that reduced form models should be fitted separately across rating class, or allow for a credit-class related factor. This conclusion is consistent with the analysis of Duffee (1999), who finds that parameter estimates of the risk-neutral default intensity process differ across credit classes.

Although the preceding analysis has uncovered differences across credit classes, it has little to say about the commonality of the factors affecting default intensities across credit. Understanding whether or not there is a systematic factor that affects *all* default intensities is especially relevant to understanding how default risk is priced. We explore this further below.

4.3. **Asymmetric Correlation of Default Probabilities.** The results of the previous section indicated that the median correlation between firms is time-varying, and, in particular, across every rating or industry group, the median correlation is lowest in period III. Recent research in the equity market provides support for the hypothesis that correlations between default intensities are time-varying. It has been often noted that equity returns show an asymmetry in correlations, such that correlations between individual equities is

[4]The pattern of correlations of default risk, where high-grade debt has more covariation than low-grade debt is also evidenced in the correlation of spreads, as in Exhibit 3 of Brown (2001).

greater for negative returns than upside moves. Ang and Chen (2001) document that the correlation of individual firm's equity returns with the market return increases with negative market returns, and especially so when there are large negative returns. Increased correlation in extreme returns has also been documented by Longin and Solnik (2001), and modeled in Das and Uppal (2000) (see also Glosten, Jagannathan and Runkle (1993)). As negative market returns also result in a higher probability of default, this suggests an increase in default intensities are likely to be linked to an increase in the correlation between default intensities. Thus, asymmetry in correlations in the equity markets would also suggest an asymmetry in correlations in the credit markets, such that the correlation between default intensities increases with an increase in default probabilities. The evidence that correlations are lowest in period III, January 1994 to April 1997, which also has the lowest average default probabilities, supports this hypothesis.

To further investigate the hypothesis that there is asymmetry of correlations in the credit markets, we follow Longin and Solnik (2001) and Ang and Chen (2001) and construct *exceedance-correlation* graphs. The graph is created as follows. Using all the firms in a rating class, we compute the total default intensity (Λ_t) at each point in time (t), i.e. $\Lambda_t = \sum_{i=1}^{N} \lambda_i(t), \ \forall t$. We then normalize the index Λ_t by subtracting its mean from each observation and dividing by its standard deviation to construct the standardized total default intensity, $\bar{\Lambda}_t$. Within each rating class, we define the exceedance correlation between firms $i = 1, 2$, $\bar{\rho}_{12}(\nu)$ $[\bar{\rho}_{12}(-\nu)$ $]$ as the correlation when $\bar{\Lambda}_t \geq \nu$ $[\bar{\Lambda}_t \leq -\nu]$,

$$\begin{aligned} \bar{\rho}_{12}(\nu) &= \operatorname{corr}(\tilde{\epsilon}_1, \tilde{\epsilon}_2 \mid \bar{\Lambda}_t \geq \nu), \\ \bar{\rho}_{12}(-\nu) &= \operatorname{corr}(\tilde{\epsilon}_1, \tilde{\epsilon}_2 \mid \bar{\Lambda}_t \leq -\nu), \end{aligned} \tag{6}$$

where $\tilde{\epsilon}_i$ is defined in equation (5). Using exceedance levels from the set $\nu \in \{0, 0.5, 1.0, 1.5\}$, we compute pair-wise exceedance correlations between firms within each credit class, and graph the mean exceedance correlation against ν. The conditioning on $\bar{\Lambda}_t$ allows us to link any asymmetry in correlations to the economy as a whole, i.e., we can answer the question: Do defaults get more correlated when the default risks in the economy as a whole are higher? The exceedance graphs provides a visual representation of asymmetry. In the absence of asymmetry, $\bar{\rho}_{12}(\nu) = \bar{\rho}_{12}(-\nu)$, and the exceedance graph is symmetric.

Figure 4 presents the exceedance correlation plots for our 3 rating groups (high, medium and low). The x-axis plots the exceedance levels and the y-axis the median correlation, $\bar{\rho}_{12}$, in each credit class. We can make the following observations. First, across every credit class, the levels of correlations are different for the left and right sides of the graph - the graphs are strongly asymmetric. In particular, the correlations are higher for positive exceedance levels as opposed to negative, i.e., $\bar{\rho}_{12}(+0) > \bar{\rho}_{12}(-0)$. Thus, correlations between default intensities are higher when default intensities are rising. Second, we can also observe that there are cross-sectional differences across credit grades in the level of asymmetry. In particular, both the mean correlation as well as the degree of asymmetry is greatest for high grade firms. This also indicates that the correlations for high grade firms are likely to show greater time-variation in correlation, consistent with results reported in Table 5. Third, as ν increases in absolute magnitude, one expects $\bar{\rho}$ to decrease (Longin and Solnik (2001)). The correlations do not decline to zero for large positive ν, suggesting an *increase* in correlations for extreme increases in the overall default level of the economy.

The empirical evidence presented in this section provides support for two observations. First, there is substantial time-variation in default probabilities that is linked to a systematic economy-wide factor - the overall default level of the economy. Second, correlations are asymmetric. Default correlations decline as the overall level of default risk in the economy declines. These conclusions support those from Tables 3 and Table 5, providing an explanation why period III, representing the low volatility bull period of the 1990's, is characterized by low default correlations.

The qualitative graphical analysis is informative, but does not provide a formal description of the time-variation in correlations. In particular, how should we jointly model time-varying default intensities and correlations? In the next section, we show how time variation in correlations and default intensities can be jointly modeled. The model allows us to construct formal statistical tests on whether the correlation structure differs across regimes, and thus provide support for asymmetric, time-varying correlations, driven by an economy-wide factor.

5. Modeling Time Varying Default Probabilities and Correlation

5.1. **Determining Regimes for Default Intensities.** The empirical evidence of the previous section suggests that default correlations are time-varying, such that correlations increase with the economy-wide probability of default. Moreover, such time-variation can be motivated by empirical research in equity markets that indicates equity correlations increase in down markets. This research also suggests how asymmetric correlation in the credit market can be modeled. Ang and Chen (2001) show that a regime shifting model is the most successful in capturing asymmetry in equity correlations. Regimes in equity correlations would then suggest regimes in default correlations, where the regimes are related to the level of default risk in the economy. A Hamilton-type (1989) regime-shifting model for modeling default risk in the economy can also be motivated by the observation that, economically, one would expect the overall level of default in the economy to be related to the business cycle. Finally, Figure 1 provides support for two distinct regimes, where default probabilities are relatively higher in one regime versus the other.

We estimate a two-regime model for default risk in the economy. Let the average (across all issuers) default intensity be denoted by $\lambda(t) = \frac{1}{N}\sum_{i=1}^{N}\lambda_i(t)$, which follows a suitable discretization of an AR(1) continuous-time model, that depends on the prevailing regime, $k_t \in \{1,2\}$, at time t (we will suppress the subscript on k_t below):

$$\begin{aligned}\lambda(t) - \theta^k &= \beta^k[\lambda(t-1) - \theta^k] + v^k \tilde{z}_t \\ \text{Regimes } k_t &= \begin{cases} 1 \\ 2 \end{cases}\end{aligned}$$

where $\tilde{z}_t$ has a standard normal distribution. k_t follows a Markov chain with a transition matrix,

$$\text{Transition Matrix} = \begin{pmatrix} q_1 & 1-q_1 \\ 1-q_2 & q_2 \end{pmatrix}$$

where q_1 and q_2 are the transition probabilities, for time $s > t$, $q_1(s,t) = Pr(k_s = 1 \mid k_t = 1)$ and $q_2(s,t) = Pr(k_s = 2 \mid k_t = 2)$. Changes from one regime to the other are generated via the transition matrix. The mean reversion rate, $1-\beta^k$, the mean default rate, θ^k, and the volatility v^k are regime dependent.

The model is estimated, following Hamilton. Consider the transition density,

$$f[\lambda(s) \mid \lambda(t)] = \exp\left(\frac{-(\lambda(s) - \lambda(t) - (1-\beta^k)(\theta^k - \lambda(t))\,)^2}{2(v^k)^2}\right)\frac{1}{\sqrt{2\pi(v^k)^2}} \tag{7}$$

for time $s > t$. We next write the Markov probabilities at time t in logit form as follows:

$$q_j(s,t) = \frac{\exp\left(a^k + b^k r_t\right)}{1+\exp(a^k + b^k r_t)}, \quad k_t = 1,2. \tag{8}$$

Here, we allow for the possibility that the transition probabilities may vary over time with the short rate, r_t, or be constant, i.e. $b_k = 0$. As we expect regimes, if any, to be correlated with the business cycle, the short-rate is a natural candidate for determining transition probabilities. The time-varying ex-ante probabilities for states 1 and 2, $[p_1(s), p_2(s)]$ are specified in terms of the ex-post probabilities from the previous period, denoted $[\widehat{p}_1(t), \widehat{p}_2(t)]$, and the Markov chain parameters $(q_1(s,t), q_2(s,t))$:

$$\begin{aligned}p_1(s) &= q_1\widehat{p}_1(t) + (1-q_2)\widehat{p}_2(t) \\ p_2(s) &= 1 - p_1(s).\end{aligned}$$

The mixture distribution is then given by:

$$f[\lambda(s)] = p_1(s).f[\lambda(s) \mid \lambda(t), k_t = 1] + p_2(s).f[\lambda(s) \mid \lambda(t), k_t = 2]. \tag{9}$$

Updating of the ex-post probabilities follows Bayes' Theorem:

$$\begin{aligned} \widehat{p}_1(s) &= \frac{p_1(s).f[\lambda(s) \mid \lambda(t), k_t = 1]}{f[\lambda(s)]} \\ \widehat{p}_2(s) &= \frac{p_2(s).f[\lambda(s) \mid \lambda(t), k_t = 2]}{f[\lambda(s)]}. \end{aligned}$$

Estimates of the parameters are obtained by maximizing the log-likelihood function

$$\{\beta^1, \beta^2, \theta^1, \theta^2, v^1, v^2, a^1, a^2\}^* = \arg\max \sum_{t=1}^{T} \log[f(\lambda_t)]. \tag{10}$$

Maximum likelihood estimates for the model are presented in Table 8. Panel A provides the estimates under the assumption of constant transition probabilities, $b^k \equiv 0$. The mean level of default in regime 1 (1.77%) is more than twice that of regime 2 (0.74%). The volatility, v^k, is also higher in regime 1 versus regime 2 (0.07% vs. 0.04%). The probability parameter (a^k) results in high values of q_k, suggesting strong persistence in each regime. The probability of remaining in regime 1 is lower (0.8765) than the probability of remaining in regime 2 (0.9692). Raising the transition matrix to the power of infinity provides the long-run stable probabilities of each regime, which are roughly 20% in the regime 1 and 80% in regime 2. All estimates are statistically significant. Panel B reports the estimates where the transition probabilities are allowed to be state-dependent, as a function of the interest rate. The coefficient, b^k, is insignificant, and, therefore, results are similar to those in Panel A.

The parameter estimates across the two regimes identify regime 1 (2) as a regime with high (low) mean default rates and volatility. In Figure 5, the probability of being in regime 1 is presented for the time period spanned by the data set. The probabilities track the patterns originally observed in Figure 1, indicating, in particular, the existence of high/low default rate regimes. From here onwards, we will identify regime 1 as the "High" regime, and regime 2 as the "Low" regime, and use the identified regimes to consider differences between regimes in the default intensities and correlations across rating classes and industry sectors.

5.2. **Regime-Dependent Parameters for Rating Classes.** In the previous section, we determined the two regimes within our sample period in terms of the process for the average level of default in the economy. Before, we consider how correlations vary across these two regimes within each rating class, it is instructive to look at how the parameters of the intensity process vary across these two regimes for the average firm in each class. We identify a specific time period as being of regime 1 (High) or 2 (Low), if the probability of that regime is more than 0.5 (see Figure 5). Of the 166 month period from January 1987 to October 2000, 30 months are identified as being in regime 1.

Using the defined regimes, we estimate the parameters for each rating class in each regime, again modeling the default intensity as, (although we use i to index either the rating class or the individual issuer, the context should be clear).

$$\lambda_i^k(t) \;=\; \lambda_i^k(t-1) + (1-\beta_i^k)\left[\theta_i^k - \lambda_i^k(t-1)\right] + v_i^k \epsilon_i(t) \qquad k \in \{1,2\}. \tag{11}$$

Assuming a standard normal distribution for $\epsilon(t)$, the parameters are estimated by maximum likelihood,

$$\max_{\Omega} \sum_{t=1}^{T} \log F[\lambda_i^k(t)], \tag{12}$$

where

$$F[\lambda_i^k(t)] = \mathbf{1}_1(z) \times f[\lambda_i^1(t)] + \mathbf{1}_2(z) \times f[\lambda_i^2(t)], \tag{13}$$

and $f[\lambda_i^k]$ is the conditional density based on the discrete stochastic process above. $\mathbf{1}_k,\ \ k = \{1,2\}$ are indicator functions based on the regimes determined in the previous subsection. The results of this estimation are reported in Tables 9 for the average firm in the class (the intensity for the average firm is now determined by averaging over all the intensities of all firms within the class i at t).

Results, reported in Table 9, show clearly that the mean default rate (θ) is different in the two regimes for every class. Mean default rates increase by 227%, 361% and 202%, for the high, medium and low grade firms, respectively. Volatilities also increases in the high regime, and as credit quality declines. Difference in the volatility between the two regimes differs by credit class, with the lowest credit class showing little difference across regimes. For each rating class, the rate of mean reversion, $1-\beta_i^k$, is also different across the two regimes. Although not reported, results are similar for the average firm in each industry sector - in the high default regime, the mean default rate and volatility are higher than in the low default regimes (for example, the mean default rate for the average firm in sector 5 increases from 0.68% to 1.47%, and the volatility from 0.05% to 0.08%).

Overall, the regime-dependent parameters for the average firm in each class shows similar patterns as those identified earlier for the overall default level of the economy. Thus, the regimes identified as “High” and “Low” in the previous section also provide a reasonable description for each of the credit classes.[5] This provides additional support for the argument

[5]In principle, we could identify separate regimes for each class. Such regimes would, however, be more difficult to interpret economically.

that there is an economy wide factor, likely related to the business cycle, that affects default probabilities across all firms.

5.3. **Differences in correlations across regimes.** Having empirically established that default intensities are time-varying and may be modeled as being regime-dependent, we now consider whether correlations vary across these two regimes, and if so, how they are different. Evidence from the previous section that correlations are asymmetric suggest that correlations are higher in the high default regime. As in the previous section, we divide the entire time period into the two previously identified regimes, "High" and "Low", and for each firm, i, we estimate equation (11).

For each class of firms, we compute the covariance matrix for each of the two regimes, and test for differences across regimes. We begin with a formal statistical test for differences in the covariance matrix between each regime (for each of the classes). Let S^k, $k = \{1, 2\}$ be the covariance matrix for each regime k. The null hypothesis is that the covariance matrices are equal across the two regimes, i.e. $S^1 = S^2$. To construct the test, define

$$C^k = \frac{S^k}{n^k - 1},$$

where n^k, $k = \{1, 2\}$, are the number of observations used to compute each covariance matrix. We also compute the pooled within covariance matrix as follows:

$$C = \frac{\sum_k (n^k - 1) C^k}{n - m}, \quad k = \{1, 2\}$$

where $n = \sum_k n^k$, and $m = 2$ (since we are testing for the equality of only 2 covariance matrices). The null hypothesis of equality of the covariance matrices is tested using a modified log-likelihood ratio statistic L,

$$\begin{aligned}
L &= M \times h, \\
M &= (n - m) \ln |C| - \sum_k (n^k - 1) \ln |C^k|, \\
h &= 1 - \frac{2p^2 + 3p - 1}{6(p + 1)(m - 1)} \left(\sum_k \frac{1}{n^k - 1} - \frac{1}{n - m} \right),
\end{aligned}$$

where p is the dimension of the covariance matrix, and L is distributed $\chi^2[p(p+1)(m-1)/2]$. If L exceeds the critical value, then we reject the hypothesis that the covariance matrices are equal.

Results are presented in Table 10 on a random sample of 30 firms in each rating category. The sample is limited to 30 firms as the number of months in the shorter regime - regime 1 - are thirty. We can easily reject the hypothesis that covariance matrices are equal across the two regimes for all three credit classes. The evidence supports the argument that correlations are time-varying and regime-dependent, where the regimes are determined by the average default risk of the economy as a whole. Similar results are obtained for alternative samples.

Next, we consider how the correlations are different across the two regimes. Using the covariance matrix estimated for each regime, we compute pair-wise correlations for both regimes. For tractability, for each rating class, we randomly select 100 issuers (except for the low grade firms where fewer firms were available with continuous data), and plot each pairwise correlation (we chose 100 so as to make the picture clearer and less overcrowded). These plots, one for each regime, are presented in Figures 6-8. The plots are color coded to highlight the differences in the correlations across regimes. The figure represents a surface plot of the pairwise correlations across all issuers, resulting in a triangular display. The upper triangle depicts the correlation structure in the high regime, and the lower triangle presents the same for the low regime. To make the differences between regimes more easily discernible, issuers are sorted based on the sum of their correlations with every other issuer. Thus, *dark* areas represent firms that have a large correlation with each other, and may be thought of as "hot spots" in a portfolio context.

A comparison of the upper and lower triangles in the graphs suggests a noticeable difference between the two regimes. The dark areas are greater for the high default regime, as compared with the low default regime, across all credit classes. Just as significantly, there are cross-sectional differences across the rating classes. In particular, high grade firms show higher correlations in both regimes, as well as a greater increase in correlations from the low regime to the high regime. These results are consistent with those reported earlier in Table 3 and Table 5, where period III may be interpreted as being in the low default regime.

To verify whether correlations are, indeed, higher in the high default regime, consider the fraction of variance explained by 1 (2) principal factors in each regime, across each credit class. The results are reported in Table 11. The results show that the fraction of the variance explained is much higher in the high default regime as compared to the low default regime. In the high default regime, the fraction of the variance explained by the first principal component is about 61%, 55% and 69% for high grade, medium grade and low grade firms, respectively. In contrast, the fraction explained by the first principal component is 32%, 34% and 35%, respectively.

In the above analysis, the regime-dependent parameters were estimated for each firm separately, identifying the residual $\epsilon_i(t)$. The residuals were then used to estimate the correlation structure for reach regime. As an alternative, we could estimate the regime-shifting model as a multivariate system across two firms at a time, allowing for regime-dependence correlation. Specifically, we estimate for each issuer i,

$$\lambda_i^k(t) - \theta_i^k = \beta_i^k \left[\lambda_i^k(t-1) - \theta_i^k\right] + v_i^k \epsilon_i^k(t), \tag{14}$$

where $k \in 1, 2$, represents the two regimes. For any pair of issuers, i and j, the parameters can be estimated by maximizing the following log-likelihood function:

$$\max_{\Omega} \sum_{t=1}^{T} \log F[\lambda_i^k(t), \lambda_j^k(t)], \tag{15}$$

where

$$F[\lambda_i^k(t), \lambda_j^k(t)] = \mathbf{1}_1(z) \times f[\lambda_i^1(t), \lambda_j^1(t)] + \mathbf{1}_2(z) \times f[\lambda_i^2(t), \lambda_j^2(t)] \tag{16}$$

and $f[\lambda_i^k(t), \lambda_j^k(t)]$ is the conditional density based on the discrete stochastic process above. We assume the residuals $\epsilon_i^k(t)$ and $\epsilon_j^k(t)$ have a jointly normal distribution conditional on k. Thus, $f[\lambda_i^k(t), \lambda_j^k(t)]$ is given as follows:

$$f[\lambda_i^k(t), \lambda_j^k(t)] = \frac{1}{2\pi v_i^k v_j^k \sqrt{(1-(\rho_{ij}^k)^2)}} \exp^{-q/2} \tag{17}$$

where

$$q = \frac{1}{1-(\rho_{ij}^k)^2}(x^2 - 2\rho_{ij}^k xy + y^2), \tag{18}$$

and

$$\begin{aligned} x &= \frac{\lambda_i^k(t) - \theta_i^k - \beta_i^k(\lambda_i^k(t-1) - \theta_i^k)}{v_i^k} \\ y &= \frac{\lambda_j^k(t) - \theta_j^k - \beta_j^k(\lambda_j^k(t-1) - \theta_j^k)}{v_j^k}. \end{aligned}$$

Notice that in the above model, the correlation ρ_{ij}^k is regime-dependent.

By also estimating a restricted model where the correlation is assumed to be the same across both regimes, we can use a standard likelihood ratio test to verify that the estimated correlation are different across regimes. To estimate both the unrestricted and restricted models, we choose a random sample of 100 firms across the entire sample. The resultant sample has 30, 62 and 8 firms of high grade, medium grade and low grade, respectively. We calculate the values of the log-likelihood function for both the restricted model and the unrestricted model. The LR-test statistics is defined as two times the difference of the values of log-likelihood functions from the unrestricted and restricted models, and has an asymptotic chi-square distribution with 1 degree of freedom.

The estimated results again verify that the high default regime has higher correlations than the low default regimes. The median correlation in the high default regime across the 4,950 estimated correlations is 0.17, as compared with a median of 0.08 in the low default regime. The likelihood ratio test rejects the hypothesis of equal correlation across regimes at the 95% significance level for 23.58% of the sample. This result provides additional support for the hypothesis that default correlations are different across regimes, and that, on average, default intensities tend to be more highly correlated in the high default regime.

In summary, the results of this section provide support for the argument that both default probabilities and default correlations are regime-dependent, and that correlations increase in the high default regime. This regime dependence can create time-variation in default correlations, as well as an asymmetric relation between default probabilities and correlations, as noted in Section 4.3. In particular, the results in this section indicates that the time-varying correlations are driven by economy-wide factor, the level of default risk in the economy. The presence of time-varying correlation indicates that a non-linear model is required to describe evolution of default risk across firms.

6. Conclusions and Implications

Understanding default risk at the level of the portfolio has implications for the pricing and management of credit risk for all credit-sensitive instruments. In this paper, we provide the first look at how default risk is correlated across several thousand US firms, using a unique default database from Moody's. Our main results and their implications, can be summarized as follows:

- Default correlations are, on average, positive, irrespective of whether firms are classified by average default level, credit rating, or industry. This empirical finding is more general than the implication from structural models that, for fixed leverage, default probabilities are positively correlated.
- Our findings indicate that there are cross-sectional differences in how default intensities are correlated across quality grades. On average, default intensities of high grade firms are the most highly correlated. An examination of the principal components of default intensities provides additional support for the hypothesis that there are structural differences in how common factors affect default across credit classes.
- Default correlations are time-varying in tandem with an economy wide default factor. We provide evidence that indicates that this time variation may be modeled in a regime-shifting framework, where the two regimes are characterized as a high

default and low default regimes. Default correlations and joint default risk is higher in the high default regime as compared with the low default regime.

What are the pricing implications of our empirical findings? As default intensities directly affect credit spreads, much of our empirical findings have similar implications for correlation between credit spreads. Evidence that a systematic, economy-wide default factor may affect variation in correlation has additional implications for how default risk is priced within the context of an equilibrium model. Our findings suggest that the risk premium component of credit spreads may systematically vary such that it increases when default risk increases. If so, this finding provides a potential explanation for the result of Collin-Dufresne, Goldstein and Martin (2001) that there is an additional systematic factor in credit spreads that cannot be explained by factors proxying solely for default loss.

Our findings also have direct implications for the management of credit risk on a portfolio basis, as well as the pricing of CDOs and other securities, whose cash flows are directly determined by risk at the portfolio level. How default risk is correlated has implications for how pools of risky debt are rated, and we provide a framework that may be of use for developing newer rating methodologies. Our framework and analysis can, for instance, be applied to the determination of an actuarially fair insurance rate for bank deposits.

Our research provides motivation for resolving other open questions in understanding firm and economy-wide defaults. In particular, an important assumption in many versions of reduced form models is that defaults, conditioned on the default intensity, are independent. The assumption has implications for the determination of joint default risk and future empirical research will hopefully shed light on its applicability. Although our analysis sheds light on joint default risk, much remains to be learn of its implications on the evolution of credit spreads. In particular, given the presence of a systematic factor driving default risk, it would be of interest to determine the relation between the statistical default intensity and the risk-neutral intensity, especially in light of the recent results of Jarrow, Lando and Yu (2001).

References

[1] Ang, A., and J. Chen (2001). "Asymmetric Correlation of Equity Portfolios," forthcoming *Journal of Financial Economics*.

[2] Bakshi, G., D. Madan, and F. Zhang (2001), "Investigating the Sources of Default Risk: Lessons from Empirically Evaluating Credit Risk Models," working paper, University of Maryland.

[3] Brown, D, (2001). "An Empirical Analysis of Credit Spread Innovations," *Journal of Fixed Income*, Sept, 9-27.

[4] Collin-Dufresne, P., and B. Solnik (2001). "On the Term Structure of Default Premia in the Swap and Libor Markets," *The Journal of Finance*, June 2001,

[5] Collin-Dufresne, P., B. Goldstein, and S. Martin (2001). "The Determinants of Credit Spread Changes," forthcoming, *Journal of Finance*.

[6] Collin-Dufresne, P., and B. Goldstein (2001). "Do Credit Spreads Reflect Stationary Leverage Ratios?," *The Journal of Finance*, forthcoming.

[7] Crosbie, P. (1999), "Modeling Default Risk", *KMV Corporation*.

[8] Das, S., G. Fong., and G. Geng (2001). "The Impact of Correlated Default Risk on Credit Portfolios," Journal of Fixed Income, December 2001, v11(3), 9-19.

[9] Das, S., and R. Sundaram (2000). "A Discrete-Time Approach to Arbitrage-Free Pricing of Credit Derivatives," *Management Science*, v46(1), 46-62.

[10] Das, S., and R. Uppal (2000). "International Portfolio Choice with Systemic Risk," working paper, London Business School.

[11] Duffie, D., and M. Huang (1996). "Swap Rates and Credit Quality" *Journal of Finance*, v51, 921-950.

[12] Duffie, D. and Singleton, K (1997). "An Econometric Model of the Term Structure of Interest Rate Swap Yields," *Journal of Finance*, Vol. 52(4), (September), 1287-1321.

[13] Duffie, D. and Singleton, K (1999). "Modeling Term Structures of Defaultable Bonds," *Review of Financial Studies*, Vol. 12(4), (Special), 687-720.

[14] Duffie, Darrel, J., and Kenneth J. Singleton (1999). " Simulating Correlated Defaults," working paper, Stanford University, Graduate School of Business.

[15] Duffee, G (1999). "Estimating the price of default risk," *Review of Financial Studies* 12, 197-226.

[16] Engle, R., and K. Kroner (1995). "Multivariate Simultaneous Generalized ARCH," *Econometric Theory*, v11, 122-150.

[17] Falkenstein, E., (2000). "RiskCalc^{TM} For Private Companies: Moody's Default Model," Moody's Investors Service, Global Credit Research, Rating Methodology, May.

[18] Falkenstein, E., E. Ibarra, A. Kocagil, and J. Sobehart, (2001). "RiskCalc^{TM} Public - Europe," Moody's Investors Service, Global Credit Research, Rating Methodology, March.

[19] Glosten, L., R. Jagannathan, and R. Runkle (1993). "On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks," *Journal of Finance*, v48(5), 1770-1801.

[20] Gluck, J., (2001a). "2000 CDO Review/2001 Preview: Growth of Synthetics Overshadows Troubles in High-Yield Debt Market," Moody's Investors Service, Structured Finance Special Report, January.

[21] Gluck, J., (2001b). "First Quarter 2001 Global CDO Review: Healthy Pace of Activity Across All CDO Subsectors," Moody's Investors Service, Structured Finance Special Report, April.

[22] Hamilton, J.D. (1989), "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle," *Econometrica*, 57, 357-384.

[23] Harris, G., (2001). "Credit Rating Migration of CDO Notes, 1996-2000," Moody's Investors Service, Structured Finance Special Report, April 27.

[24] Jarrow, R and S. Turnbull. (1995) "Pricing Derivatives on Financial Securities subject to Credit Risk," *Journal of Finance* v50[1] (March), 53-86.

[25] Jarrow, R., D. Lando, and S. Turnbull (1997). "A Markov Model for the Term Structure of Credit Spreads," *Review of Financial Studies*, v10, 481-523.

[26] Jarrow, R., D. Lando, and F. Yu (2001), "Default Risk and Diversification: Theory and Applications," Working paper, Cornell University.

[27] Jones E. P., S. P. Mason and E. Rosenfeld (1984). "Contingent Claims Analysis of Corporate Capital Structures: An Empirical Investigation," *Journal of Finance*, 39, 611:627.

[28] Lando, David (1994). "Three essays on contingent claims pricing," Ph.D. thesis, Cornell University.

[29] Lando, David (1998). "On rating transition analysis and correlation," *Credit Derivatives: Applications for Risk Management, Investment and Portfolio Optimization*, Risk Publications, 147-155.

[30] Longin, F., and B. Solnik (1995). "Is the Correlation in International Equity Returns Constant: 1960-1990?" *Journal of International Money and Finance*, v14, 3-26.

[31] Longin, F., and B. Solnik (2001). "Extreme Correlation of International Equity Markets," *Journal of Finance*, v56, 649-676.

[32] Madan, Dilip and H. Unal, (1998). "Pricing the Risks of Default," *Review of Derivatives Research*, v2 (2/3), 121-160. 449-470.

[33] Merton, R. (1974). "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates," *Journal of Finance*, v29, 449-470.

[34] Longstaff, F., and Schwartz, E. (1995), "A Simple Approach to Valuing Risky Fixed and Floating Rate Debt," Journal of Finance, Vol. 51[3] (September), 789-819.

[35] 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.

[36] Zhou, C. (2001). "An Analysis of Default Correlations and Multiple Defaults," *Review of Financial Studies* 14, 555-576.

TABLE 1. **Rating and Sector Classifications**

Panel A provides the description of each rating class in terms of Moody's ratings. Panel B provides the description of each industry sector by broad SIC groups. Sector 8 corresponds to financial, insurance and real estate and is not represented in the sample.

Panel A: Rating Classification

Rating Class	Moody's Rating	Credit Group
1	Aaa, Aa1, Aa2, Aa3	High Grade
2	A1, A2, A3	High Grade
3	Baa1, Baa2, Baa3	Medium Grade
4	Ba1, Ba2, Ba3	Medium Grade
5	B1, B2, B3	Low Grade
6	Caa1, Caa2, Caa3, Ca, C	Low Grade
7	Not rated	-

Panel B: Sector Classification

SIC Sector	Description
1	Agriculture, Forestry, and Fishing
2	Mining
3	Construction
4	Manufacturing
5	Transportation, Communications, Electric, etc.
6	Wholesale Trade
7	Retail Trade
9	Services
10	Public Administration

Table 2. **Descriptive Statistics of Default Probabilities**

The table report the median default probabilities for US non-financial public firms, sorted by rating and sector. The sample period of January 1987 to October 2000 is divided into 4 sub-periods: 1/87-4/90, 5/90-12/93, 1/94-4/97, 5-97-10/00. The median default probability for each group is then calculated by averaging monthly observations of Moody's default probabilities over each sub-period and over all firms within that group. Sector and rating groups are as defined in Table 1. Number of firms are shown in brackets.

Group	Sub-Period I (%)	Sub-Period II (%)	Sub-Period III (%)	Sub-Period IV (%)
High Grade	0.09	0.09	0.07	0.23
	{213}	{223}	{210}	{215}
Medium Grade	0.75	0.77	0.49	1.17
	{344}	{337}	{403}	{535}
Low Grade	4.56	5.01	3.29	5.65
	{130}	{125}	{203}	{278}
Not Rated	1.66	1.89	1.63	2.45
	{2724}	{2517}	{3183}	{4142}
Sector 1	1.59	1.99	1.21	2.03
	{15}	{14}	{21}	{20}
Sector 2	2.04	1.80	1.36	2.78
	{245}	{215}	{243}	{276}
Sector 3	2.43	2.39	1.92	2.47
	{39}	{42}	{51}	{73}
Sector 4	1.50	1.68	1.38	2.09
	{1788}	{1687}	{2083}	{2633}
Sector 5	1.17	1.48	1.35	2.40
	{357}	{353}	{402}	{491}
Sector 6	1.75	2.13	1.87	2.87
	{179}	{169}	{204}	{270}
Sector 7	1.74	1.98	1.79	2.92
	{236}	{216}	{307}	{396}
Sector 9	1.80	2.14	1.99	2.76
	{509}	{472}	{645}	{964}
Sector 10	1.99	3.53	3.56	3.83
	{43}	{34}	{43}	{47}

TABLE 3. **Descriptive Statistics of Average Correlations**

The table reports the (i) median and (ii) fraction of pairwise correlation, ρ_{ij} that are positive for firms belonging to the group. The correlation is defined as the correlation between e_i and e_j, from the regression

$$\lambda_i(t) = \bar{\lambda}_i + e_i(t),$$

where $\lambda_i(t)$ is the default intensity of firm i in month t. Panel A reports the results by credit class, and Panel B by SIC industry sector.

Panel A

Group	**Period I**		**Period II**		**Period III**		**Period IV**	
	Median	$\rho_{ij} > 0$	Median	$\rho_{ij} > 0$	Median	$\rho_{ij} > 0$	Median	$\rho_{ij} > 0$
High Grade	0.62	0.86	0.18	0.63	0.05	0.54	0.58	0.90
Medium Grade	0.35	0.73	0.20	0.63	0.02	0.52	0.35	0.74
Low Grade	0.17	0.61	0.19	0.61	0.00	0.50	0.29	0.69

Panel B

Group	**Period I**		**Period II**		**Period III**		**Period IV**	
	Median	$\rho_{ij} > 0$	Median	$\rho_{ij} > 0$	Median	$\rho_{ij} > 0$	Median	$\rho_{ij} > 0$
Sector 1	0.54	0.67	0.14	0.54	-0.04	0.47	0.40	0.70
Sector 2	0.16	0.61	0.04	0.51	0.00	0.49	0.50	0.70
Sector 3	0.29	0.65	0.25	0.61	0.02	0.49	0.23	0.66
Sector 4	0.29	0.66	0.12	0.57	0.00	0.49	0.21	0.61
Sector 5	0.42	0.73	0.17	0.62	0.01	0.51	0.15	0.64
Sector 6	0.32	0.67	0.14	0.62	0.00	0.49	0.20	0.64
Sector 7	0.40	0.69	0.14	0.60	0.02	0.51	0.13	0.63
Sector 9	0.17	0.58	0.06	0.52	0.00	0.51	0.13	0.55
Sector 10	0.15	0.63	0.04	0.48	0.08	0.51	0.10	0.53

Table 4. **Summary of Parameter Estimates for AR(1) Model for Default Intensity**

The table reports the results of the OLS regression,

$$\lambda_i(t) = \alpha_i + \beta_i \lambda_i(t-1) + \tilde{\epsilon}_i(t),$$

where $\lambda_i(t)$ is the default intensity for firm i for month t. The regression is estimated in each of 4 sub-periods for every firm that has continuous observations in each sub-period. The table reports the first quartile, median, and third quartiles for α_i and β_i across all firms in each credit class. N is the number of firms. Over 99% of the estimates for β_i were significant.

		$100 * \alpha_i$			β_i			
		25%	Median	75%	25%	Median	75%	N
High Grade	Period I	0.0021	0.0045	0.0086	0.91	0.94	0.96	213
	Period II	0.0022	0.0051	0.0112	0.86	0.91	0.94	223
	Period III	0.0029	0.0049	0.0105	0.85	0.90	0.93	210
	Period IV	0.0059	0.0125	0.0216	0.89	0.93	0.97	215
Medium Grade	Period I	0.0115	0.0233	0.0729	0.87	0.91	0.94	344
	Period II	0.0084	0.0184	0.0491	0.86	0.91	0.94	337
	Period III	0.0063	0.0154	0.0352	0.85	0.90	0.94	403
	Period IV	0.0212	0.0441	0.1170	0.88	0.92	0.96	535
Low Grade	Period I	0.0488	0.1489	0.4087	0.84	0.91	0.95	130
	Period II	0.0226	0.0930	0.2390	0.88	0.92	0.96	125
	Period III	0.0254	0.1034	0.2797	0.87	0.93	0.98	203
	Period IV	0.0860	0.2257	0.4680	0.90	0.94	0.98	278

Table 5. **Correlations under Time-Varying Default Probabilities**

The table reports the (i) median and (ii) fraction of pairwise correlation, ρ_{ij} that are positive for firms belonging to the group. The correlation is defined as the correlation between $\tilde{\epsilon}_i$ and $\tilde{\epsilon}_j$, from the regression

$$\lambda_i(t) = \alpha_i + \beta_i \lambda_i(t-1) + \tilde{\epsilon}_i(t),$$

where $\lambda_i(t)$ is the default intensity of firm i in month t.

Panel A

Group	**Period I**		**Period II**		**Period III**		**Period IV**	
	Median	$\rho_{ij} > 0$	Median	$\rho_{ij} > 0$	Median	$\rho_{ij} > 0$	Median	$\rho_{ij} > 0$
High Grade	0.37	0.89	0.10	0.53	0.01	0.52	0.11	0.70
Medium Grade	0.23	0.82	0.10	0.68	0.02	0.55	0.08	0.67
Lower Grade	0.16	0.80	0.07	0.66	0.02	0.54	0.08	0.68

Panel B

Group	**Period I**		**Period II**		**Period III**		**Period IV**	
	Median	$\rho_{ij} > 0$	Median	$\rho_{ij} > 0$	Median	$\rho_{ij} > 0$	Median	$\rho_{ij} > 0$
Sector 1	0.19	0.70	0.06	0.64	0.00	0.50	0.00	0.50
Sector 2	0.09	0.68	0.00	0.51	0.01	0.52	0.10	0.66
Sector 3	0.16	0.80	0.05	0.64	0.02	0.54	0.07	0.65
Sector 4	0.17	0.78	0.05	0.61	0.00	0.50	0.06	0.61
Sector 5	0.13	0.71	0.06	0.63	0.00	0.51	0.05	0.60
Sector 6	0.16	0.77	0.06	0.63	0.00	0.51	0.05	0.60
Sector 7	0.19	0.77	0.07	0.63	0.01	0.50	0.05	0.60
Sector 9	0.11	0.71	0.04	0.59	0.00	0.51	0.04	0.59
Sector 10	0.09	0.69	0.01	0.52	0.00	0.49	0.03	0.58

Table 6. **Principal Components Analysis**

For each credit class, the table reports the fraction of the variance of the residual in equation (5),

$$\lambda_i(t) = \alpha_i + \beta_i\lambda_i(t-1) + \tilde{\epsilon}_i(t),$$

that is explained by the first one, and two principal components, respectively.

Panel A

Group	**Sub-Period I**	**Sub-Period II**	**Sub-Period III**	**Sub-Period IV**
High Grade	55.01	24.70	15.28	51.07
	95.20	36.44	27.08	69.91
Medium Grade	16.34	27.27	16.70	12.04
	29.45	41.83	29.92	23.33
Low Grade	22.64	19.67	11.10	15.45
	34.98	32.80	19.53	25.45

Panel B

Group	**Sub-Period I**	**Sub-Period II**	**Sub-Period III**	**Sub-Period IV**
Sector 1	77.93	58.11	41.88	32.75
	91.49	78.13	63.09	51.73
Sector 3	40.24	33.59	17.51	21.25
	64.98	53.41	31.55	40.94
Sector 5	21.46	14.15	12.00	13.45
	35.51	27.25	22.06	24.13
Sector 7	18.50	19.09	10.13	11.83
	32.10	30.56	19.99	21.08
Sector 9	14.76	15.85	10.39	12.39
	22.78	24.11	18.43	21.94

Table 7. **Correlations between principal components**

This table presents the correlations between the first principal component of each rating grade, i.e. high, medium and low grades. We present the correlation matrices for each subperiod. The principal components are computed from the residuals of the AR(1) model.

Period	Rating	Rating		
		High	Medium	Low
Period 1	High	1.0000	0.3980	-0.0735
	Medium	0.3980	1.0000	-0.7640
	Low	-0.0735	-0.7640	1.0000
		High	Medium	Low
Period 2	High	1.0000	0.7671	0.5725
	Medium	0.7671	1.0000	0.5914
	Low	0.5725	0.5914	1.0000
		High	Medium	Low
Period 3	High	1.0000	0.0520	0.1852
	Medium	0.0520	1.0000	0.1811
	Low	0.1852	0.1811	1.0000
		High	Medium	Low
Period 4	High	1.0000	-0.5736	-0.1475
	Medium	-0.5736	1.0000	0.5113
	Low	-0.1475	0.5113	1.0000

Table 8. **Estimation of the Regime Switching model**

The table reports estimates from a two-regime model for default probabilities. Panel A is for the case of a constant transition matrix, and Panel B for the case where transition probabilities vary with the short rates.

Panel A: Constant Transition Probabilities

Parameters	Estimates	t-statistics
β^1	0.78	17.37
β^2	0.92	57.48
θ^1	1.77%	18.43
θ^2	0.74%	15.61
v^1	0.07%	6.57
v^2	0.04%	15.28
a^1	1.96	3.39
a^2	3.45	6.59

Panel B: Time-varying Transition Probabilities

Parameters	Estimates	t-statistics
β^1	0.78	17.21
β^2	0.92	51.94
θ^1	1.76%	17.03
θ^2	0.74%	16.33
η^1	0.07%	7.61
η^2	0.04%	15.47
a^1	3.46	0.87
a^2	5.14	2.56
b^1	-0.26	-0.39
b^2	-0.29	-0.91

Table 9. **Regime Dependent Parameters within each Rating Category**

For each rating category, i, the table reports the maximum likelihood estimates of the model,

$$\lambda_i^k(t) = \lambda_i^k(t-1) + (1-\beta_i^k)[\theta_i^k - \lambda_i^k(t-1)] + v_i^k \epsilon_i(t), \quad k \in \{1,2\}.$$

across two regimes. The regimes are based on the estimation in Panel A of Table 8. The t-statistics are reported in parenthesis.

Rating Class	Parameters					
	β_i^1	β_i^2	θ_i^1 (%)	θ_i^2 (%)	v_i^1 (%)	v_i^2 (%)
High	0.97	0.99	0.25	0.11	0.03	0.01
	(82.00)	(49.32)	(5.67)	(3.77)	(7.18)	(14.80)
Medium	0.86	0.98	1.48	0.41	0.09	0.05
	(18.64)	(49.37)	(9.00)	(2.54)	(9.37)	(17.34)
Low	0.85	0.96	4.81	2.38	0.12	0.12
	(8.92)	(7.77)	(8.93)	(7.77)	(8.21)	(16.07)

Table 10. **Statistical differences in covariance matrices**

This table presents the chi-square statistics for the differences in covariance matrices in the high and low default regimes. The test was performed on a randomly chosen group of 30 firms from each rating class.

Rating Class	Test-Statistics	Degrees of Freedom	p-value
High Grade	985.69	465	0.0000
Medium Grade	657.65	465	0.0000
Low Grade	904.75	465	0.0000

TABLE 11. **Principal Component Analysis Across Regimes**

Across each regime and for each credit class, the table reports the fraction of the variance in ϵ_i from equation (11), explained by the first one, and the first two principal components, respectively.

Rating Class	Low Default Regime	High Default Regime
High Grade	32.02	61.37
	68.07	78.86
Medium Grade	33.87	55.22
	61.08	78.85
Low Grade	45.27	69.66
	64.37	85.48

Table 12. **Defaults Conditional on PDs**

Linear regression - Dependent variable: $NUMDEF$		
Ind. Var	Coefficient	T-Statistic
Intercept (β_0)	-6.5613	-3.67
$AVGEDF$ (β_1)	667.3942	5.94
$REGIME$ (β_2)	1.8603	2.04
$R^2 = 0.4532$	F-stat (pval) = 0.0000	

AR(1) Model - Dependent variable: $NUMDEF$		
Ind. Var	Coefficient	T-Statistic
Intercept (β_0)	-6.9003	-3.17
$AVGEDF$ (β_1)	693.0343	5.15
$REGIME$ (β_2)	1.4011	1.42
ρ	0.2757	3.61
$R^2 = 0.4912$	Durbin-Watson = 2.08	

FIGURE 1. **Historical PDs**

We plot average Moody's default probabilities across the issuers in the sample on a monthly basis. The data reflects periods of high default intensity and periods of low intensity.

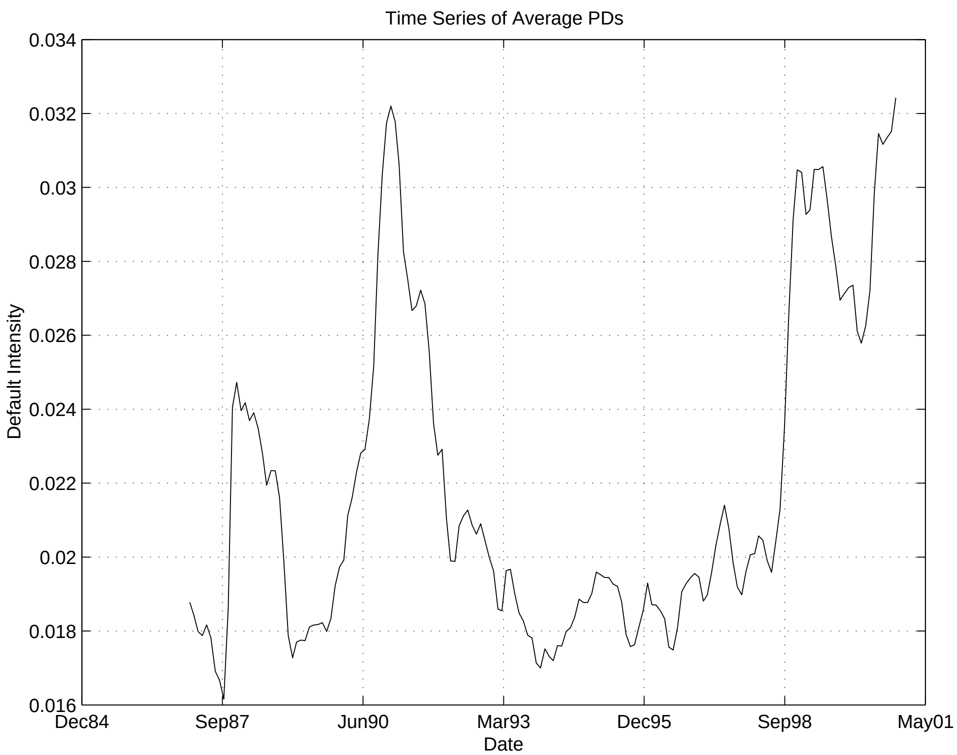

Figure 2. **Relationship of Actual Default to Lagged PDs**

We scatter plot actual defaults against the average Moody's default probabilities across the issuers in the sample on a monthly basis. The data reflects the strong positive relationship between defaults and default intensity. We plot the regression line as well to demonstrate the positive relationship.

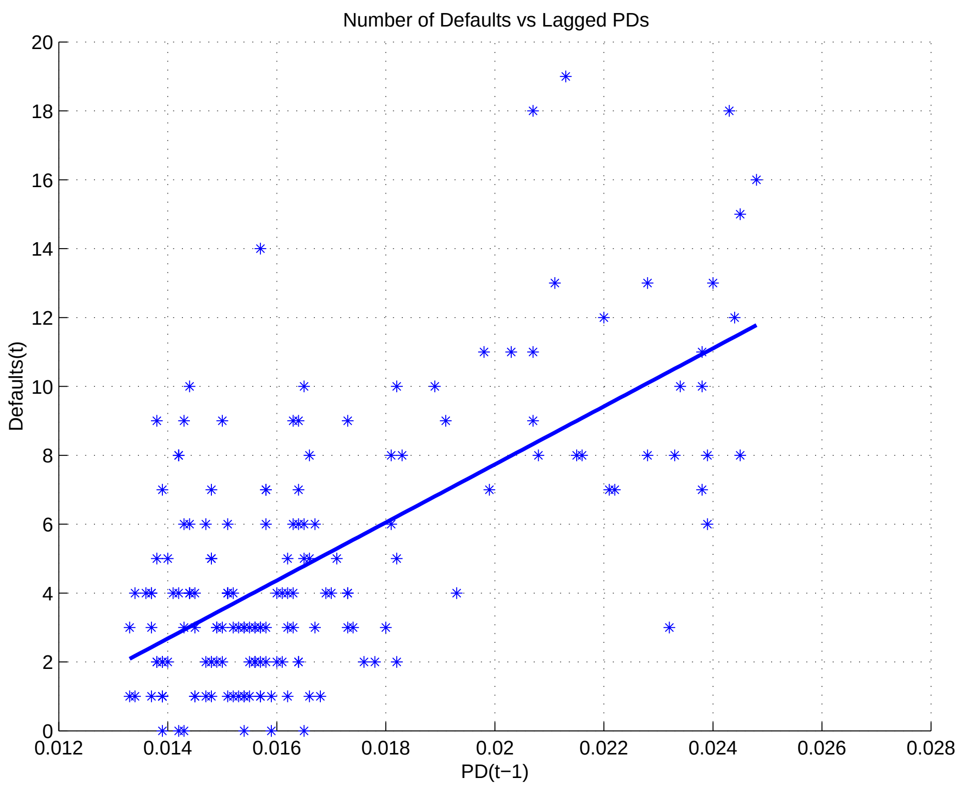

FIGURE 3. **Principal Components in Residuals**

This figure graphs the fraction of the variance of $\epsilon_i(t)$ of equation 5 that is explained by up to the first 10 principal components. Each panel, comprising 3 plots represents one of four sub-periods, and shows the results across each of three credit classes, High Grade, Medium Grade and Low Grade. The upper left panel is for period I, the upper right for period II, bottom left for period III and finally, the bottom right panel is for period IV.

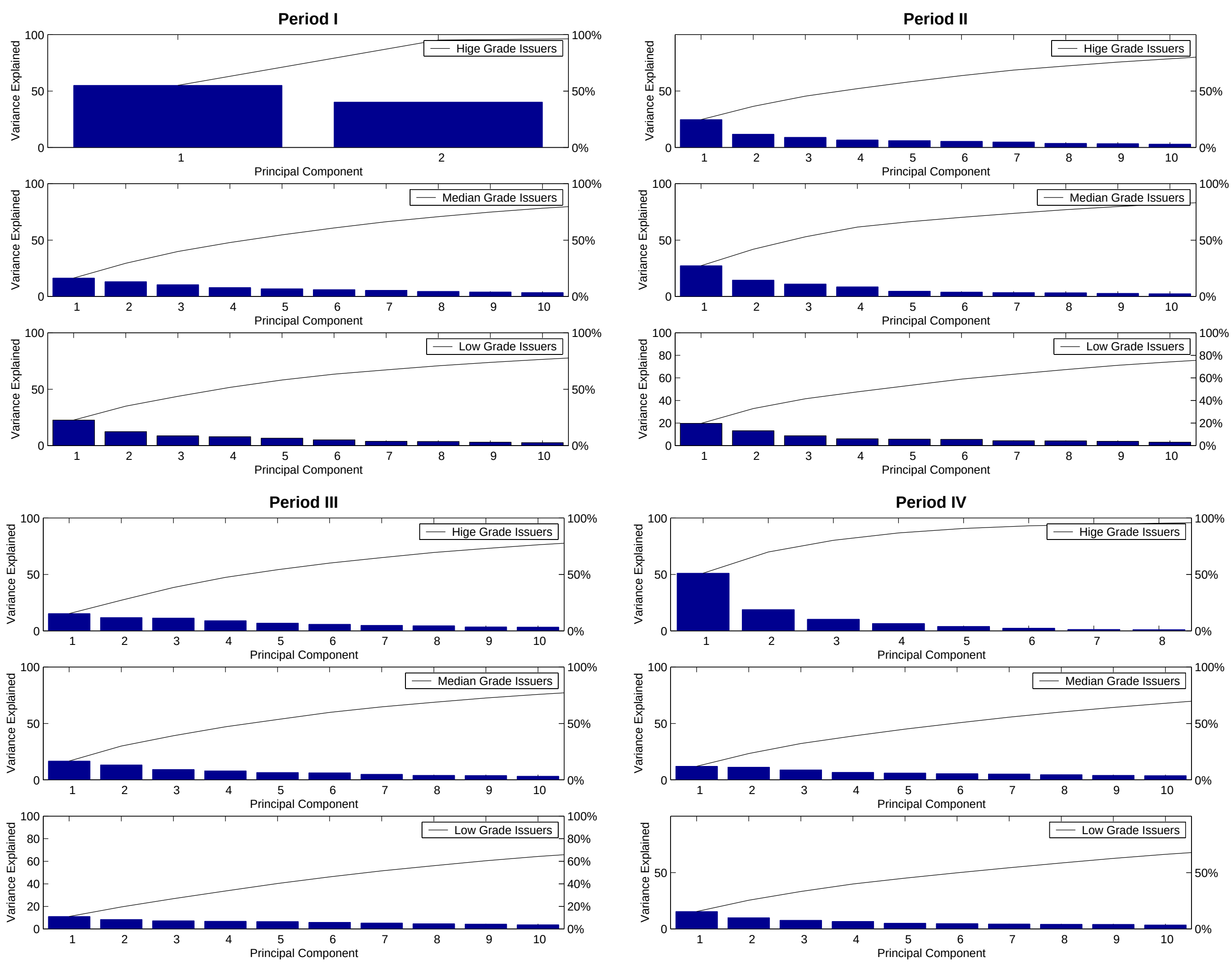

Figure 4. **Asymmetric Correlation**

This figure graphs the exceedance correlations for the different rating classes in our study. The high, medium and low rating classes are as defined in Table 1.

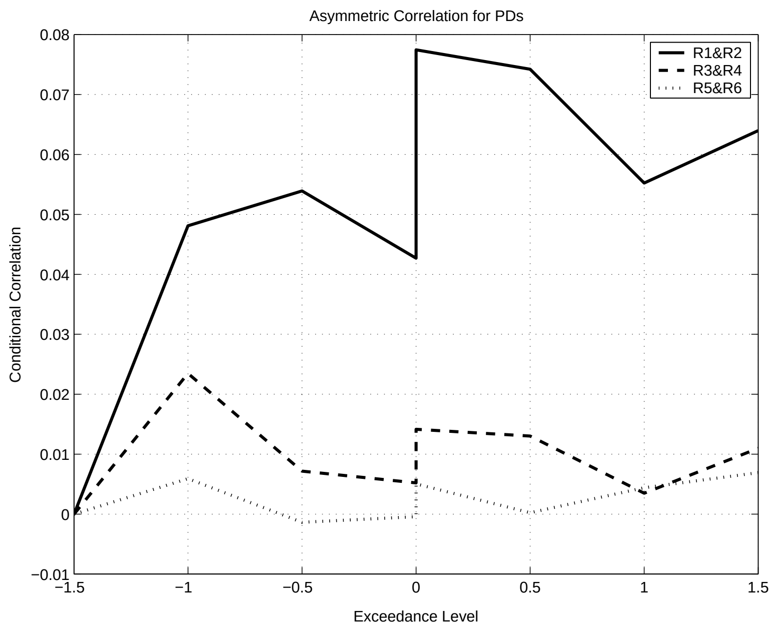

Figure 5. **Time series of regime probabilities**

This figure presents the time series plot of probability of being in the high default intensity regime. The probability of being in the low intensity regime is one minus the high intensity probabilities, since there are only two regimes in the estimated model. The upper panel presents the results when the transition probabilities are assumed to be fixed. The lower panel shows probabilities when the transition matrix varies with the state vector z. The state vector comprises the short-term interest rate r.

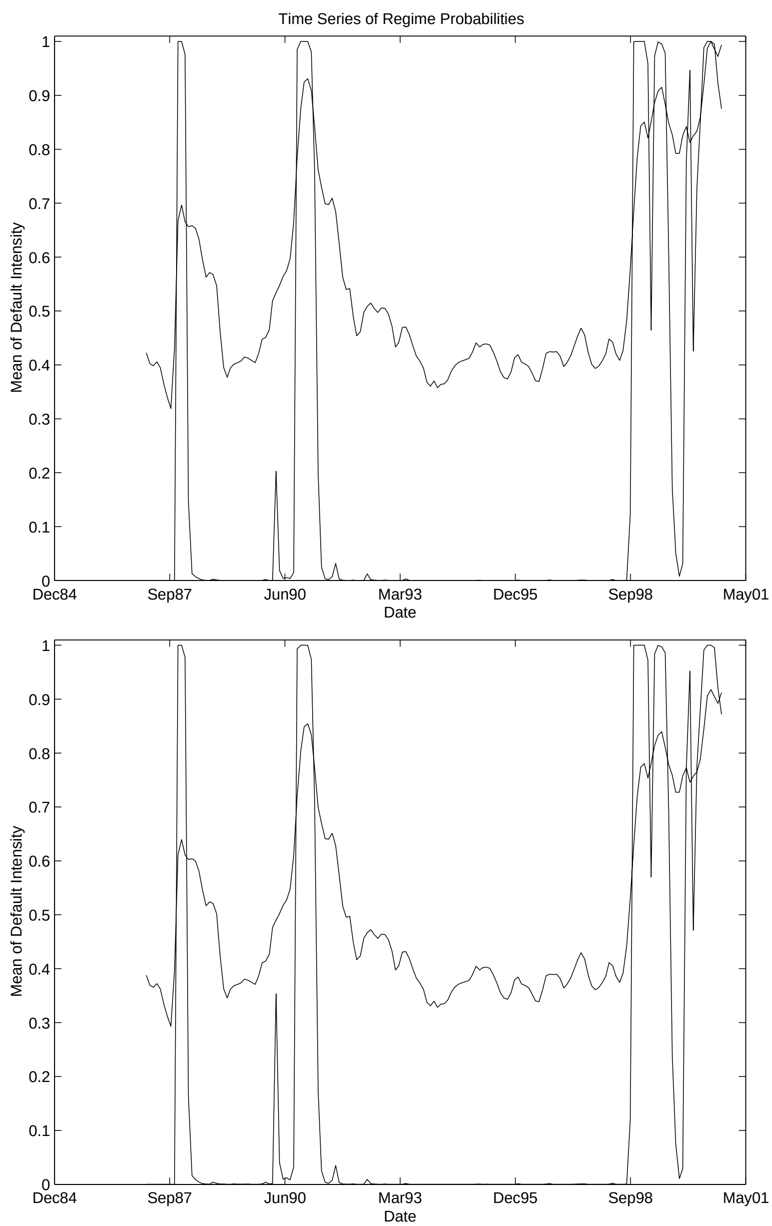

FIGURE 6. **Correlation plots for differences in regimes**

This figure presents correlation plots for high rated firms for a random selection of 100 issuers in each class. The upper triangle depicts the correlations in the high default regime, and the lower triangle shows the correlations in the low default regime. A comparison of upper and lower triangles reflects differences in the correlation structure across the two regimes.

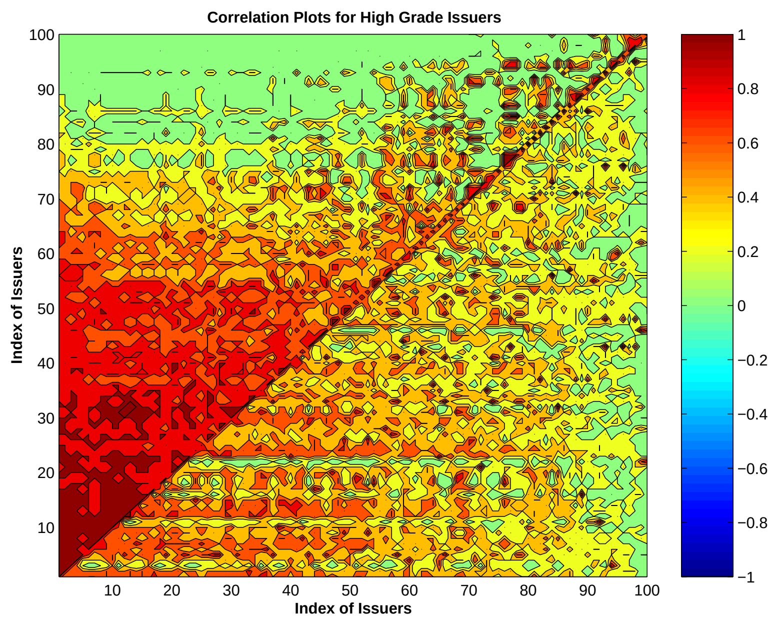

Figure 7. **Correlation plots for differences in regimes**

This figure presents correlation plots for medium rated firms for a random selection of 100 issuers in each class. The upper triangle depicts the correlations in the high default regime, and the lower triangle shows the correlations in the low default regime. A comparison of upper and lower triangles reflects differences in the correlation structure across the two regimes.

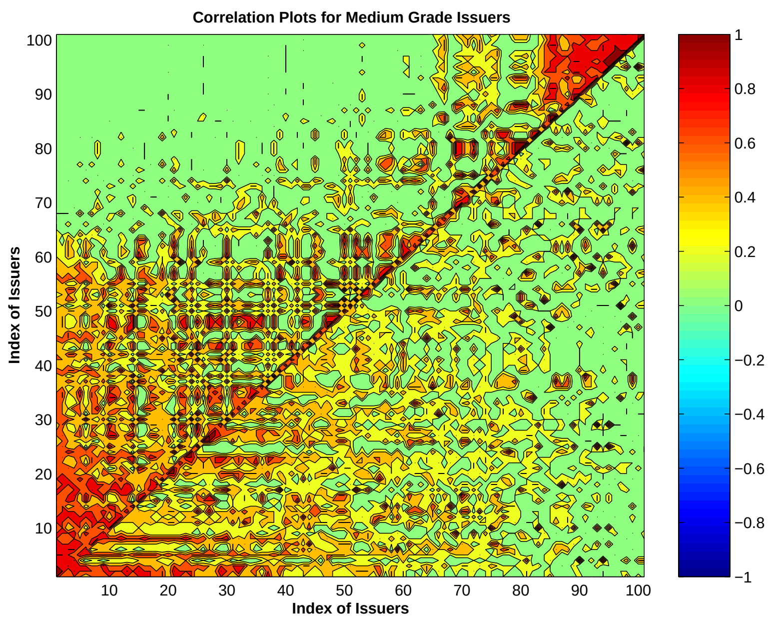

FIGURE 8. **Correlation plots for differences in regimes**

This figure presents correlation plots for low rated firms for a random selection of 100 issuers in each class. The upper triangle depicts the correlations in the high default regime, and the lower triangle shows the correlations in the low default regime. A comparison of upper and lower triangles reflects differences in the correlation structure across the two regimes.