AN EMPIRICAL ANALYSIS OF THE PRICING OF COLLATERALIZED DEBT OBLIGATIONS

Francis A. Longstaff*
Arvind Rajan**

Abstract. We use the information in collateralized debt obligation (CDO) prices to study market expectations about how corporate defaults cluster or are correlated across firms. We find that a three-factor portfolio credit model allowing for firm-specific, industry, and economywide default events explains virtually all of the time-series and cross-sectional variation in an extensive data set of CDX index tranche prices. These tranches are priced as if losses of 0.4, 6, and 35 percent of the portfolio occur with expected frequencies under the risk-neutral measure of 1.2, 41.5, and 763 years, respectively. On average, 65 percent of the CDX spread is due to firm-specific default risk, 27 percent to clustered industry or sector default risk, and 8 percent to catastrophic or systemic default risk. Recently, however, firm-specific default risk has begun to play a larger role.

Current version: January 2007.

*Allstate Professor of Insurance and Finance, UCLA Anderson School and the NBER. **Managing Director, Relative Value Trading, Citigroup. We are grateful for the valuable comments and suggestions received from Vineer Bhansali, Pierre Collin-Dufresne, Keith Crider, Sanjiv Das, Darrell Duffie, Youseff Elouerkhaoui, Bjorn Flesaker, Kay Giesecke, Ben Golub, Mitch Janowski, Holger Kraft, David Lando, L. Sankarasubrahmanian, Alan Shaffran, David Shelton, Jure Skarabot, Ken Singleton, Ryoichi Yamabe, and seminar participants at UC Berkeley, the Federal Reserve Bank of New York, IXIS Capital Markets, Harvard Business School/Harvard Department of Economics, the NBER Summer Institute, New York University, Pimco, UCLA, the University of Michigan, and Yale University. We are particularly grateful for the comments and suggestions of the editor Campbell Harvey and an anonymous referee and associate editor. We also thank Guarav Bansal, Yuzhao Zhang, and Xiaolong Cheng for capable research assistance. All errors are our responsibility.

1. INTRODUCTION

In its most basic form, a collateralized debt obligation (CDO) is simply a financial claim to the cash flows generated by a portfolio of debt securities or, equivalently, a basket of credit default swaps (CDS contracts). Thus, CDOs can be viewed as the credit-market counterparts to the familiar collateralized mortgage obligations (CMOs) actively traded in secondary mortgage markets. Since its inception in the mid 1990s, the market for CDOs has become one of the most-rapidly-growing financial markets ever. Industry sources estimate the size of the CDO market at the end of 2006 to be nearly $2 trillion, representing more than a 30 percent increase over the prior year.[1] Recently, CDOs have been in the spotlight because of the May 2005 credit crisis in which downgrades of Ford's and General Motors' debt triggered a wave of large CDO losses among many credit-oriented hedge funds and major Wall Street dealers. Despite the growing importance of this market, however, relatively little research on CDOs has appeared yet in the academic literature.

It is important to stress that CDOs are not merely the "latest" in a series of exotic instruments introduced by Wall Street and of interest only to hedge funds, speculators, and arbitrageurs. Rather, CDOs represent a fundamental new class of securities and are important to a much wider set of investors. Furthermore, CDOs are important to researchers since they may in fact provide a near-ideal "laboratory" for studying a number of fundamental issues in financial economics. For example, CDOs provide a unique opportunity to identify the *joint* distribution of default risk across firms since CDOs are claims against a portfolio of debt, information that cannot be inferred from the *marginal* distributions associated with single-name credit instruments. The joint distribution, however, is crucial to understanding the extent to which credit risk is diversifiable in the economy. Furthermore, if the market expects defaults to cluster in some way, this clearly has implications for the behavior of the corresponding stocks—clustered default risk in bond markets necessarily implies related nondiversifiable event risk in the equity market. As another example, the pricing of senior CDO tranches opens a new window on the upper tail of the distribution of potential credit losses in the economy. This information is essential in understanding the systemic risk faced by financial institutions, the possibility of contagion across business and credit cycles,

[1] A key driver of the growth in CDO market is the parallel growth in the credit derivatives market which the International Swaps and Derivatives Association (ISDA) estimates reached $26 trillion notional in mid-2006.

and the risk of "credit crunches" and liquidity crises in the capital markets.

CDO-like structures are emerging as a major new type of financial vehicle and/or "virtual" institution.[2] In particular, the CDO structure can be viewed as an efficient special purpose vehicle for making illiquid assets tradable, creating new risk-sharing and insurance opportunities in financial markets, and completing markets across credit states of the world. Specifically, CDO-like structures are now used not only for corporate bonds and loans, but also for less-liquid and more-private assets such as home-equity loans, credit-card receivables, commercial mortgages, auto loans, student loans, equipment leases, trade receivables, small business loans, private equity, emerging market local assets, and even the "intellectual" property rights of rock stars.[3] To illustrate how these vehicles can complete markets, imagine a scenario where the economy is in such a deep depression that there are no longer any investment grade bonds in the market. By tranching a portfolio of high-yield bonds in a way that transfers virtually all default risk to junior tranches, it would still be possible for a senior CDO tranche to carry a AAA rating. Thus, this vehicle allows for risk sharing among agents across extreme states of the world. Finally, observe that a CDO could also be viewed as a "synthetic bank" in the sense that its assets consist of loans and its liabilities run the gamut from near-riskless senior debt to highly leveraged equity. The key distinction, however, is that the "synthetic" CDO bank may not engage in the same type of monitoring activities as actual banks. Thus, a comparison of CDO equity and bank stocks could provide insights into the delegated-monitoring role of financial intermediaries.[4]

This paper represents a first attempt to understand the economic structure of default risk across firms using information from the CDO market. Specifically, we use the prices of standardized tranches on the CDX credit index to infer the market's expectations about the way in which default events cluster across firms. The study is based on an extensive proprietary data set recently made available to us by Citigroup. To our knowledge, this is the first large-scale empirical study of this market. Thus, our initial responsibility is to provide an introduction to the CDO market and describe the fundamental characteristics of index tranche pricing.

Motivated by recent research by Collin-Dufresne, Goldstein, and Martin (2001), Elton, Gruber, Agrawal, and Mann (2001), Eom, Helwege, and Huang (2004), Longstaff, Neis, and Mithal (2005), and others that corporate credit spreads are driven by firm-specific factors as well as broader economic forces, we develop a simple multi-factor portfolio credit model for pricing CDOs. Our framework has some features in common with Duffie and Gârleanu (2001) who allow for three types of default events

[2]For a discussion of the role of tranching in markets with asymmetric information, see Demarzo (2005).

[3]For example, see Richardson (2005) for a discussion of the "Bowie" bonds.

[4]For example, see Diamond (1984).

in their framework: idiosyncratic or firm-specific defaults, industrywide defaults in a specific sector of the economy, and economywide defaults affecting virtually every industry and sector. Rather than focusing on the individual "quantum" or "zero-one" states of default for each firm and aggregating up to the portfolio level, however, our framework takes a "statistical mechanics" approach by modeling portfolio credit losses directly. Specifically, we allow portfolio losses to occur as the realizations of three separate Poisson processes, each with a different jump size and intensity process.[5] Thus, realizations of the three Poisson events could potentially (although not necessarily) be mapped into the three types of default events in the Duffie and Gârleanu framework. We take the model to the data by fitting it to the CDX index spread and the prices of the 0−3, 3−7, 7−10, 10−15, and 15−30 percent CDX index tranches for each date during the sample period.

We first address the issue of how many factors are actually needed to explain the pricing of CDOs. To do this, we also estimate one-factor and two-factor versions of the model. These one-factor and two-factor models restrict the degree of potential default clustering relative to the three-factor model, but are both economically viable alternatives to the three-factor model. We then use a likelihood ratio approach to test whether the two-factor model has significant incremental explanatory power relative to the one-factor model and, likewise, whether the three-factor model has significant incremental explanatory power relative to the two-factor model. We find that the three-factor model significantly outperforms the two-factor model which, in turn, significantly outperforms the one-factor model. These results provide the first direct evidence that the market expects defaults for the firms in the CDX index to cluster (correlated defaults).

Focusing on the three-factor results, we find that the estimated jump sizes for the three Poisson processes are on the order of 0.4, 6, and 35 percent, respectively. Since there are 125 firms in the CDX index, the jump size of 0.4 percent for the first Poisson process can be interpreted as the portfolio loss resulting from the default of a single firm, given a 50 percent recovery rate ($1/125 \times 0.50 = 0.004$). The jump size of 6 percent for the second Poisson process can be viewed as an event in which, say, 15 firms default together. Since this represents roughly 10 percent of the firms in the portfolio, one possible interpretation of this event could be that of a major crisis that pushes an entire industry or sector into financial distress. For expositional purposes, we will generally refer to a realization of the second Poisson process as an industry event. Clearly, however, there are many other possible interpretations. For example, this type of event could just as easily involve clustered defaults among firms with similar accounting ratios, currency or raw materials exposures, firm age, firm

[5]In independent work, Giesecke and Goldberg (2005) put forward an interesting approach to modeling multiname credit risk that also has many similarities to ours. Their approach is called the top-down approach.

size, etc.[6] Finally, the 35 percent jump size for the third Poisson process could be viewed as a catastrophic or systemic event that wipes out the majority of firms in the economy. Our analysis indicates that all three types of credit risk are anticipated by the market.

We also estimate the probabilities or intensities of the three Poisson events (under the risk-neutral pricing measure). On average, the expected time until an idiosyncratic or firm-specific default is 1.2 years, the expected time until a clustered industry default crisis is 41.5 years, and the expected time until a catastrophic economywide default event is 763 years.[7]

In an effort to understand whether clustering in default risk is in fact linked to industry, we perform a principal components analysis of changes in the CDS spreads for the individual firms in the CDX index. We find that there is a dominant first factor driving spreads across all industries. This is consistent with the view that there is a pervasive economywide component to credit. Moving beyond this first factor, however, we find that the second, third, fourth, etc. principal components are significantly related to specific industries or groups of industries. Thus, there is some evidence that default clustering occurs in ways that have some relation to industry. On the other hand, we repeat the principal components analysis using stock returns for the individual firms in the CDX index and find that the second, third, fourth, etc. principal components for stock returns are much more strongly related to industry than is the case for the CDS spreads. Thus, there are intriguing differences in the cross-sectional structure of stock returns and credit spreads for the firms in the CDX index.

Using the intensity estimates, we decompose the level of the CDX index spread into its three components. We find that on average, firm-specific default risk represents only 64.6 percent of the total CDX index spread, while clustered industry or sector and economywide default risks represent 27.1 and 8.3 percent of the index spread, respectively. Thus, the risk of industry or economywide financial distress accounts for more than one-third of the default risk in the CDX portfolio. Recently, however, idiosyncratic default risk has played a larger role. By definition, the same decomposition holds for the default spreads of the typical firm in the index. Thus, the event that a firm defaults can be partitioned into three mutually exclusive subevents: the event

[6]I am grateful to the associate editor for this insight.

[7]An expected time of 763 years may seem unrealistically long, but it is important to observe that there has never been a credit event in U.S. history—not even during the U.S. Civil War or the Great Depression—in which more than 50 percent of the firms in the economy defaulted or went bankrupt. On the other hand, there are numerous documented economic collapses and sovereign defaults in erstwhile safe countries over the past centuries, suggesting that a non-zero probability is appropriate to attach to such an event (see Kindleberger (2005)).

that only the firm defaults, the event that the firm and a number of other firms in the same industry or sector or similar grouping default together, and the catastrophic event in which the majority of firms in the economy default together. We show that time variation in the components of the spread explains changes over time in the standard deviation of credit spreads across firms. This is consistent with the view that these components may be capturing economic effects that drive time variation in the correlations of credit risk across firms.

We then examine how well the model captures the pricing of individual index tranches. Even though tranche spreads are often measured in hundreds or even thousands of basis points, the root-mean-squared error (RMSE) of the three-factor model is typically on the order of only two to three basis points, which is well within the typical bid-ask spreads in the market. Thus, virtually all of the time-series and cross-sectional variation in index tranche prices is captured by the model. We find that the largest pricing errors occur shortly after the inception of the CDX index and tranche market, but decrease rapidly after several weeks. Thus, despite some early mispricing, the evidence suggests that the CDX index tranche market quickly evolved.

Finally, we link the portfolio-level information obtained from the fitted model to the cross-sectional structure of credit spreads for individual firms. Specifically, we model the joint distribution of individual firm defaults using a multivariate correlated Bernoulli distribution and solve for the implied default correlation across firms. The implied default correlation ranges from roughly 0.05 to 0.10 during the sample period, but reaches a high of about 0.13 around the May 2005 credit crisis. We also find that the volatility of the loss distribution for the portfolio is an increasing function of the dispersion or cross-sectional variability in credit spreads for the individual firms in the index.

There is a rapidly growing literature on credit derivatives and correlated defaults.[8] This paper contributes to this primarily theoretical literature by presenting a new approach to modeling portfolio default losses, conducting the first extensive empirical analysis of pricing in the CDO markets, and providing the first direct estimates of the nature and degree of default clustering across firms expected by market participants.

The remainder of this paper is organized as follows. Section 2 provides an introduction to the CDO market. Section 3 describes the data used in the study. Section 4 presents the three-factor portfolio credit model. Section 5 applies the model to the

[8]Important recent contributions in this area include Duffie and Gârleanu (2001), Hull and White (2003), Giesecke (2004), Jorion and Zhang (2005), Longstaff, Mithal, and Neis (2005), Das, Duffie, Kapadia, and Saita (2005), Das, Freed, Geng, and Kapadia (2005), Saita (2005), Yu (2005a, 2005b), Giesecke and Goldberg (2005), Frey and Backhaus (2005), Schönbucher (2005), Bakshi, Madan, and Zhang (2004), Duffie, Saita, and Wang (2006), Duffie, Eckner, Horel, and Saita (2006), Errais, Giesecke, and Goldberg (2006), and many others.

valuation of index tranches. Section 6 reports the results from the empirical analysis. Section 7 summarizes the results and makes concluding remarks.

2. AN INTRODUCTION TO CDOs

CDOs have become one of the most important new financial innovations of the past decade. It is easiest to think of a CDO as a portfolio containing certain debt securities as assets, and multiple claims in the form of issued notes of varying seniority. The liabilities are serviced using the cash flows from the assets, as in a corporation. Although CDOs existed in various forms previously, it was only in the mid-1990s that they began to be popular. Over subsequent years, issuance experienced rapid growth. For example, the during the first three quarters of 2006, issuance was $322 billion, representing nearly a 102 percent increase over the same period during 2005.[9] The assets securitized by cash CDOs have broadened to include investment-grade bonds, high yield bonds, emerging market securities, leveraged loans, middle market loans, trust preferred securities, asset-backed securities, commercial mortgages, and even previously issued CDO tranches. CDOs play an increasingly important economic role in the markets, allowing portfolio credit risk and return to be redistributed according to demand among investors with different levels of risk aversion.[10]

Over the past few years, the technology of cash CDOs has merged with the technology of credit derivatives to create the so-called synthetic CDO, which is the main focus of this paper. Synthetic CDOs differ from cash CDOs in that the portfolios that provide the cash flow to service their liabilities consist of credit default swaps rather than bonds or other cash securities. The majority of synthetic securities are based on corporate credit derivatives, and tend to be simpler (and hence more amenable to economic modeling).

2.1 An Example of a Stylized CDO

To build up understanding of a full-fledged synthetic CDO, we will begin with a simpler example based on a $100 million investment in a diversified portfolio of five-year par corporate bonds. Imagine that a financial institution (CDO issuer) sets up this portfolio, which consists of 100 separate bonds, each with a market value of $1 million, and each issued by a different firm. Imagine also that every bond in the portfolio is rated BBB and has a coupon spread over Treasuries of 100 basis points.

[9]To put these numbers in perspective, we note that according to the Securities Industry and Financial Markets Association, the total issuance of corporate bonds and agency mortgage-backed securities during 2005 was $703.2 and $966.1 billion, respectively.

[10]For additional insights into the CDO market, see the excellent discussions provided by Duffie and Gârleanu (2001), Duffie and Singleton (2003), Roy and Shelton (2007), and Rajan, McDermott, and Roy (2007).

The CDO issuer can now sell five-year claims against the cash flows generated by the portfolio. These claims are termed CDO tranches and are constructed to vary in credit risk from very low (senior tranches) to low (junior or mezzanine tranches) to very high (the "equity" tranche).

Let us illustrate a typical CDO structure by continuing the example. First, imagine that the CDO issuer creates a so-called equity tranche with a total notional amount of three percent of the total value of the portfolio ($3 million). By definition, this tranche absorbs the first three percent of any defaults on the entire portfolio. In exchange, this tranche may receive a coupon rate of, say, 2500 basis points above Treasuries. If there are no defaults, the holder of the equity tranche earns a high coupon rate for five years and then receives back his $3 million notional investment. Now assume that one of the 100 firms represented in the portfolio defaults (and also that there is zero recovery in the event of default). In this case, the equity tranche absorbs the $1 million loss to the portfolio and the notional amount of the equity tranche is reduced to $2 million. Thus, the equity tranche holder has lost one-third of his investment. Going forward, the equity tranche investor receives the 2500 basis point coupon spread as before, only now only on his $2 million notional position. Now assume that another two firms default. In this case, the equity tranche absorbs the additional losses of $2 million, the notional amount of the equity tranche investor's position is completely wiped out, and the investor receives neither coupons nor principal going forward. Because a three-percent loss in the portfolio translates into a 100-percent loss for the equity tranche investor, we can view the equity tranche investor as being leveraged 33 1/3 to 1. However, unlike an investor who leverages by borrowing, the equity tranche investor has no liability beyond a three-percent portfolio loss, a condition referred to as "non-recourse" leverage.

Now imagine that the CDO issuer also creates a junior mezzanine tranche with a total notional amount of four percent of the total value of the portfolio ($4 million). This tranche absorbs up to four percent of the total losses on the entire portfolio after the equity tranche has absorbed the first three percent of losses. For this reason, this tranche is designated the 3−7 percent tranche. In exchange for absorbing these losses, this tranche may receive a coupon rate of, say, 300 basis points above Treasuries. If total credit losses are less than three percent during the five-year horizon of the portfolio, then the 3−7 percent investor earns the coupon rate for five years and then receives back his $4 million notional investment. If total credit losses are greater than or equal to seven percent of the portfolio, the total notional amount for the 3−7 percent investor is wiped out.

The CDO issuer follows a similar process in creating additional mezzanine, senior mezzanine, and even super-senior mezzanine tranches. A typical set of index CDO tranches might include the 0−3 percent equity tranche, and 3−7, 7−10, 10−15, 15−30, and 30−100 percent tranches. The initial levels 3, 7, 10, 15, and 30 percent at which losses begin to accrue for the respective tranches are called attachment points or subordination levels. Note that the total notional valuation of all the tranches equals the \$100 million notional of the original portfolio of corporate bonds. In addition, the total coupon payments to all tranches must equal the total coupon flow from the entire underlying portfolio. For example, if the coupon spreads on the $0-3$, $3-7$, $7-10$, $10-15$, $15-30$, and $30-100$ tranches were 2500, 300, 50, 30, 20, and 10, respectively, then the weighted-average coupon spread of $.03 \times 2500 + .04 \times 300 + .03 \times 50 + .05 \times 30 + .15 \times 20 + .70 \times 10 = 100$ basis points would equal the total coupon spread of the overall portfolio.

An interesting aspect of the CDO creation process is that because each tranche has a different degree of credit exposure, each tranche would have its own credit rating. For example, the super-senior $30-100$ percent tranche can only suffer credit losses if total losses on the underlying portfolio exceed 30 percent of the total notional amount. Since this is highly unlikely, this super-senior tranche would typically have a AAA rating. In contrast, the highly-subordinated $3-7$ percent mezzanine tranche might carry a below-investment-grade rating as a reflection of its junior status in the "capital structure" of the overall portfolio. In general, the credit rating of a tranche should be similar to those for ordinary corporate bonds with similar coupon spreads. For example, a $7-10$ percent tranche with a coupon spread of 50 basis points will likely have a rating similar to corporate bonds with comparable coupon spreads. If an issuer purchases the underlying bond portfolio, and then sells off each individual tranche, the issuer no longer has any net economic exposure to credit events in the portfolio (this is referred to as selling the entire capital structure). This example also illustrates one of the most important economic functions that the CDO tranching process fulfills. Specifically, the tranching process allows securities of any credit rating to be created. Thus, the CDO process can serve to complete the financial market by creating high-credit-quality securities that might not otherwise exist in the market.

2.2 Synthetic CDOs

To take advantage of the wide availability of credit derivatives, credit markets have recently introduced CDO structures known as synthetic CDOs. This type of structure has become very popular and the total notional amount of synthetic CDO tranches is growing rapidly. A synthetic CDO is economically similar to a cash CDO in most respects. The principal difference is that rather than there being an underlying portfolio of corporate bonds on which tranches are based, the underlying portfolio is actually a basket of credit default swap contracts. Recall that a CDS contract functions as an insurance contract in which a buyer of credit protection makes a fixed payment each quarter for some given horizon such as five years.[11] If there is a default on the underlying reference bond during that period, however, then the buyer of protection

[11] As with any swap contract, however, CDS contracts carry the small additional risk of a counterparty default. In reality, this risk can be largely mitigated by the posting of collateral between swap counterparties.

is able to give the defaulted bond to the protection seller and receive par (the full face value of the bond).[12] Thus, the first step in creating a synthetic CDO is to define the underlying basket of CDS contracts.

2.3 Credit Default Indexes and Index Tranches

In this study, we focus on CDOs with cash flows tied to the most liquid U.S. corporate credit derivative index, the DJ CDX North American Investment Grade Index. This index is managed by Dow Jones and is based on a liquid basket of CDS contracts for 125 U.S. firms with investment grade corporate debt. The CDX index itself trades just like a single-name CDS contract, with a defined premium based on the equally-weighted basket of its 125 constituents. The individual firms included in the CDX basket are updated and revised ("rolled") every six months in March and September, with a few downgraded and illiquid names being dropped and new ones taking their places. CDX indexes are numbered sequentially. Thus, the index for the first basket of 125 firms was designated the CDX NA IG 1 index in 2003, the index for the second basket of 125 firms the CDX NA IG 2 index, etc., and so on up to CDX NA IG 7 in September 2006, of which the first five series comprise the data set analyzed in this paper. While there is considerable overlap between successive CDX NA IG indexes, there can occasionally be significant changes across index rolls. For example, the CDX NA IG 4 index (beginning in March 2005) includes Ford and General Motors while the CDX NA IG 5 index (beginning in September 2005) does not since the debt for these firms dropped below investment grade in May 2005.

Index CDO tranches have also been issued, each tied to a specific CDX index. The attachment points of these CDO tranches are standardized at 3, 7, 10, 15, and 30 percent, exactly as in the example above. To illustrate how an index CDO tranche works, let us use the numbers from our previous example—but instead of bonds, using the CDX index with a market premium of 50 basis points consisting of 125 CDS contracts of $1 million each (index notional of $125 million). Suppose the equity tranche holder receives a coupon spread of 2500 basis points. If a default occurs in one of the 125 index names, however, the equity tranche holder faces the same cash outflows as the protection seller in a CDS contract on the defaulting firm. Specifically, the equity tranche investor would have a cash outflow of $1 million dollars if the recovery of principal (the value of the debt obligation delivered by the protection buyer) is zero percent, and the notional amount on which future coupon cash flows are based would be reduced by $1 million (exactly as in the cash CDO example). In a similar manner, the equity index tranche investor would bear the entire losses of any subsequent defaults (up to her notional exposure) just as if she were the protection seller in CDS contracts on the defaulting firms. In essence, we can view the cash

[12]This aspect of the design of the contract means that the protection buyer can be compensated for his losses relatively quickly; the protection buyer does not need to wait until the end of the bankruptcy and recovery process.

outflows to the various index tranche investors as the equivalent of being the protection seller on the first three to default, on the fourth through seventh to default, the eighth through tenth to default, etc. Thus, the initial coupon spread on each tranche is held fixed over time (but applied to the remaining notional amount within each tranche).

Since these instruments are structured as credit default swaps, when investors "buy" a synthetic index tranche from a counterparty, they are selling protection on that tranche. Their counterparty has bought protection on the same tranche from them. This highlights a convenient feature of these index tranches—that is, a dealer need not create and sell the entire capital structure of tranches to investors; rather investors are free to synthetically create and trade (sell or buy) individual index tranches (single-tranche index CDOs) according to their needs. For example two investors can trade protection on the 3−7 percent tranche of the DJ CDX 5 index with each other without anyone having to create the 0−3 percent, 7−10 percent and other tranches. As observed earlier, the losses on an $N{-}M$ percent tranche are zero if the total losses on the underlying portfolio are less than N. On the other hand, the total losses on the tranche are 1.00 or 100 percent if the total losses on the underlying portfolio equals or exceeds M. For underlying portfolio losses between N and M, tranche losses are linearly interpolated between zero and one. Because of this, the losses on a $N{-}M$ percent tranche can be viewed intuitively as a call spread on the total losses of the underlying portfolio. This intuition will be formalized in a later section. Just as an option has a "delta", that is, an equivalent exposure to the underlying, the tranche has a delta with respect to its underlying index.

The simplicity of construction, the liquidity of the underlying CDX indexes, the standardization of attachment points, the availability of tranche pricing models, and the freedom from creating the full capital structure have all contributed to a great increase in the volume of trading in index tranches. Therefore, in the past few years, index tranches have become liquidly traded and quotes and data for them are available from many dealers in the market daily.

2.4 Other Synthetic CDOs

Although index tranches are the most liquid synthetic tranches, a synthetic tranche can be based on any portfolio. A tranche created with a specific non-index portfolio, and with customized attachment points, e.g. 5−8 percent, is called a bespoke tranche. Investors use bespoke tranches to buy or sell specific slices of protection on specific portfolios that they wish to express views on. For example, an investor selling protection on the 0−4 percent equity tranche of a BB rated portfolio might be expressing a bullish point of view on the likelihood of defaults in his favorite basket of high-yield credits, while another investor buying protection on a 5−10 percent basket of 75 investment-grade names might be buying "out-of-the-money" catastrophe protection on a portfolio she owns (since the likelihood of five percent or more of default-related losses in an investment grade portfolio in five years is quite small). While the results in this paper are based on index tranche data, the analysis can equally well be applied to most bespoke CDO tranches. Finally, there are also full capital structure synthetic CDOs, created when demand exists for all parts of the capital structure. Provided a CDO observes the simple type of structure we specified in the example, a model such as the one in this paper may be used to price its tranches.[13]

3. THE DATA

CDOs are a relatively new financial innovation and have only recently begun to trade actively in the markets. Because of this, it has been difficult for researchers to obtain reliable CDO pricing data. We were fortunate, however, to be given access by Citigroup to one of the most extensive proprietary data sets of CDO index and tranche pricing data in existence.[14]

The data consist of daily closing values for the five-year CDX NA IG index (CDX index for short) for the period from October 2003 to October 2005. As discussed earlier, the underlying basket of 125 firms in the index is revised every March and September. Thus, the index data is actually for the five individual indexes denoted CDX i, $i = 1, 2, 3, 4$, and 5. CDX 1 covers October 20, 2003 to March 19, 2004; CDX 2 covers March 22, 2004 to September 22, 2004; CDX 3 covers September 23, 2004 to March 18, 2005; CDX 4 covers March 21, 2005 to September 19, 2005; CDX 5 covers September 20, 2005 to October 18, 2005. This data set covers virtually the entire history of the CDX index through 2005. Data is missing for some days during the earlier part of the sample. We omit these days from the sample, leaving us with a total of 435 usable daily observations for the two-year sample period. For the primarily descriptive purposes of this section, we report summary statistics based on the continuous series of the on-the-run CDX index (rather than reporting statistics separately for the individual CDX series).

In addition to the index data, we also have daily closing quotation data for the 0−3, 3−7, 7−10, 10−15, and 15−30 percent tranches on the CDX index. The pricing data for most tranches are in terms of the basis point premium paid to the CDO investor for absorbing the losses associated with the individual tranches. Thus, a price of 300 for the 3−7 percent tranche implies that the tranche investor would receive a premium of 300 basis points per year paid quarterly on the remaining balance in exchange for absorbing the default losses from three to seven percent on the CDX

[13]The analysis in this paper, however, may not apply directly to certain other types of portfolio derivative products, for example Nth-to-default baskets, CDO-squareds, and cash CDOs, which have more granular compositions, more complex structures, or more difficult-to-model cash flows and rules, respectively.

[14]Although the data set we were given access to is proprietary, data for standardized CDX index tranches are now available on the Bloomberg system and other commercial sources.

index. The exception is the market convention for the equity tranche (the 0−3 percent tranche) which is generally quoted in terms of points up front. Specifically, a price of 50 for this tranche means that an investor would need to receive \$50 up front per \$100 notional amount, plus a premium of 500 basis points per year paid quarterly on the remaining balance, to absorb the first three percent of losses on the CDX index. Rather than using this market convention, however, we convert the points up front into spread equivalents to facilitate comparison with the pricing data for the other tranches.

In addition to the CDX index and tranche data, we also collect daily New York closing data on 3-month, 6-month, 12-month Libor rates, and on 2-year, 3-year, 5-year, 7-year, and 10-year swap rates. The Libor data is obtained from the Bloomberg system. The swap data is obtained from the Federal Reserve Board's web site. From this Libor spot rate and swap par rate data, we use a standard cubic spline approach to bootstrap zero-coupon curves that will be used throughout the paper to discount cash flows.[15] Since the same zero-coupon curve is used to discount both legs of the CDO contract, however, the results are largely insensitive to the decision to discount using the Libor-swap curve; the results are virtually identical when the bootstrapped Treasury curve is used for discounting cash flows.

Table 1 provides summary statistics for the levels and first differences of the index and tranche data. As shown, the average values of the spreads are monotone decreasing in seniority (attachment point). The average spread for the 0−3 percent equity tranche is 1758.87 basis points (which translates into an average number of points up front of 39.34). This spread is many times larger than the average spread for the junior mezzanine 3−7 percent tranche, reflecting that the expected losses for the equity tranche are much higher than for more senior tranches. Similar comparisons hold for all the other tranches. Figure 1 plots the time series of tranche spreads for the various attachment points. The correlations indicate that while these spreads have a high level of correlation with each other, there is also considerable independent variation.

4. THE MODEL

Motivated by these aspects of the data, as well as by the mounting evidence in the literature that credit spreads are driven by idiosyncratic as well as broader market factors, we develop a simple multifactor portfolio credit model for valuing CDO index tranches in this section.[16] Although developed independently, our framework comple-

[15]See Longstaff, Mithal, and Neis (2005) for a more detailed discussion of this bootstrapping algorithm.

[16]Evidence about the multifactor nature of credit risk is provided by Collin-Dufresne, Goldstein, and Martin (2001), Elton, Gruber, Agrawal, and Mann (2001), Eom, Hel-

ments important recent theoretical work on "top down" portfolio credit modeling by Giesecke and Goldberg (2005) and others.[17]

To explain the logic behind this model, it is useful to first make a short digression into the modeling of stock index options (such as options on the S&P 100). In theory, one could model S&P 100 index options by specifying the dynamics of each of the 100 firms in the index and then evaluating a 100-dimensional expectation. In reality, of course, such an approach would be cumbersome and impractical. Rather, the standard approach to valuing index options is to take a more "macro" perspective and model the dynamics of the index directly.

To date, most modeling of CDOs has likewise been done at an individual firm level. Typically, practitioners model the losses on, say, the 125-firm portfolio underlying the CDX index by first simulating the dynamics of each firm, checking whether each firm is in the "quantum" or "zero-one" state of default, and then aggregating losses over the entire portfolio. As discussed earlier, however, losses on the tranches are simple functions of the total losses on the underlying portfolio. Thus, the distribution of total portfolio losses represents a "sufficient statistic" for valuing tranches (just as the distribution of the stock index is sufficient for pricing stock index options). Accordingly, rather than modeling individual "quantum" defaults, we will follow a "statistical mechanics" based approach of modeling the distribution of total portfolio losses directly.

In doing this, it is important to stress that we are not implying that individual firm-level information about default status is unimportant. In fact, for many types of credit derivatives (such as credit default swaps or first-to-default swaps on small baskets of firms), individual firm default status is essential in defining the cash payoffs. Rather, we suggest that for many other types of credit-related contracts that are tied to larger portfolios, the "reduced-form" approach of modeling portfolio-level losses directly may provide important advantages such as simplicity, transparency, and tractability with little loss in our ability to capture the underlying economics. In general, the smaller the single-name risk concentration in a portfolio, the more applicable is the aggregate loss approach taken here.

Let L_t denote the total portfolio losses on the CDX portfolio per \$1 notional amount. By definition, $L_0 = 0$. To model the dynamic evolution of L_t we assume

$$\frac{dL_t}{1 - L_t} = \bar{\gamma}_1 \; dN_{1t} \;+\; \bar{\gamma}_2 \; dN_{2t} \;+\; \bar{\gamma}_3 \; dN_{3t}, \tag{1}$$

where $\bar{\gamma}_i = 1 - e^{-\gamma_i}$, $i = 1, 2, 3$, where γ_1, γ_2, and γ_3 are nonnegative constants defining

wege, and Huang (2004), Longstaff, Neis, and Mithal (2005), and many others.

[17] Also see recent papers by Giesecke (2004), Schönbucher (2005), and Sidenius, Piterbarg, and Andersen (2005).

jump sizes, and where N_{1t}, N_{2t}, and N_{3t} are independent Poisson processes. Note that for small values of γ_i, the jump size $\bar{\gamma}_i$ is essentially just γ_i. Thus, for expositional simplicity, we will take a slight liberty and generally refer to the parameters γ_1, γ_2, and γ_3 simply as jump sizes. Integrating Equation (1) and conditioning on time-zero values (a convention we adopt throughout the paper) gives the general solution for L_t

$$L_t = 1 - e^{-\gamma_1 N_{1t}} \quad e^{-\gamma_2 N_{2t}} \quad e^{-\gamma_3 N_{3t}}. \tag{2}$$

From this equation, it can be seen that the economic condition $0 \leq L_t \leq 1$ is satisfied for all t. Furthermore, since N_{1t}, N_{2t}, and N_{3t} are nondecreasing processes, the intuitive requirement that total losses be a nondecreasing function of time is also satisfied.[18]

These dynamics imply that there are three factors at work in generating portfolio losses. To illustrate how these factors affect total losses, assume that the three jump sizes are 0.01, 0.10, and 0.50, respectively, and that there is zero recovery in the event of a default. When a jump in the first Poisson process occurs, the portfolio experiences a one-percent loss. Thus, a realization of the first Poisson process could be viewed as an isolated default affecting only one firm. In contrast, a realization of the second Poisson process results in a portfolio loss of 10 percent. Thus, this event could be interpreted as the impact of a major event that decimates the ranks of a group of firms (possibly in a specific sector or industry). Similarly, when a jump in the third Poisson process occurs, 50 percent of the remaining firms in the portfolio default, corresponding to the realization of some catastrophic event affecting the entire economy. Thus, the model could be viewed as allowing for both idiosyncratic or firm-specific default as well as for the broader systemic risk of multiple defaults within a sector or throughout the economy.

The fact that the model incorporates both firm-specific and multiple-firm defaults allows the model to capture the notion of default correlation. To see this, consider the hypothetical situation in which the intensities of the second and third Poisson processes were zero. In this case, only idiosyncratic or uncorrelated defaults could occur. At the other extreme, consider the situation where the intensities of the first and second Poisson processes are zero. In this situation, only highly correlated systemic defaults could occur in the economy. In between these two extremes, a full spectrum of possible default correlations could arise based on the relative magnitudes of the intensities of

[18]Because portfolio losses on the CDX portfolio consist of payments to protection buyers, it is difficult to imagine a realistic set of circumstances where total losses would actually decline over time. It would not be sufficient for a defaulted bond to subsequently increase in value since the cash flow to the protection buyer is made at a specific point in time. Rather, it would require that a protection payment made on a CDS contract be reversed and rebated back to the protection seller.

the Poisson processes.

The intensities of the three Poisson processes are designated λ_{1t}, λ_{2t}, and λ_{3t}, respectively. To complete the specification of the general model, we assume that the dynamics for the intensity processes are given by,

$$d\lambda_{1t} = (\alpha_1 - \beta_1 \ \lambda_{1t}) \ dt \ + \ \sigma_1 \ \sqrt{\lambda_{1t}} \ dZ_{1t}, \tag{3}$$

$$d\lambda_{2t} = (\alpha_2 - \beta_2 \ \lambda_{2t}) \ dt \ + \ \sigma_2 \ \sqrt{\lambda_{2t}} \ dZ_{2t}, \tag{4}$$

$$d\lambda_{3t} = (\alpha_3 - \beta_3 \ \lambda_{3t}) \ dt \ + \ \sigma_3 \ \sqrt{\lambda_{3t}} \ dZ_{3t}, \tag{5}$$

where Z_{1t}, Z_{2t}, and Z_{3t} are standard independent Brownian motion processes. These dynamics insure that the intensities for the three Poisson processes are always nonnegative. Furthermore, the mean-reverting nature of the intensities allows the model to potentially capture expected migrations in the credit quality of the underlying portfolio. Specifically, we would anticipate that over time, the lowest credit-quality firms would tend to exit the portfolio sooner, resulting in an expected downward trend in the value of λ. This trend could be reflected in the model in the situation where the initial value of λ was above its long-run mean value of α/β.[19] Since these intensities are stochastic, it is clear from the previous discussion that this framework allows default correlations to vary over time. Although we present analytical results for the general case implied by Equations (3) through (5) in this section, the empirical results to be presented later are actually based on the special case of the model where $\alpha_i = \beta_i = 0$ for all i.

To value claims that depend on the realized losses on a portfolio, we first need to determine the distribution of L_t. From Equation (2), L_t is a simple function of the values of the three Poisson processes. Thus, it is sufficient to find the distributions for the individual Poisson processes, since expectations of cash flows linked to L_t can be evaluated directly with respect to the distributions of N_{1t}, N_{2t}, and N_{3t}.

Since many of the following results are equally applicable to each of the three Poisson processes, we will simplify notation whenever possible by dropping the subscripts 1, 2, and 3 when we present generic results and the interpretation is clear from context. Standard results imply that, conditional on the path of λ_t, the probability of $N_T = i$, $i = 0, 1, 2, \ldots$, can be expressed as

$$\frac{\exp\left(-\int_0^T \lambda_t \ dt\right) \left(\int_0^T \lambda_t \ dt\right)^i}{i!}. \tag{6}$$

[19]We are very grateful to the referee for pointing this out.

Let $P_i(\lambda, T)$ denote $i!$ times the probability that $N_T = i$, conditional on the current (the time-zero unsubscripted) value of λ. Thus,

$$P_i(\lambda, T) = E\left[\exp\left(-\int_0^T \lambda_t \; dt\right)\left(\int_0^T \lambda_t \; dt\right)^i\right]. \tag{7}$$

For $i = 0$, the Appendix shows that this expression is easily solved in closed form from results in Cox, Ingersoll, and Ross (1985). For $i > 0$, the results in Karlin and Taylor (1981), pp. 202-204 can be used to show that $P_i(\lambda, T)$ satisfies the recursive partial differential equation,

$$\frac{\sigma^2 \lambda}{2} \; \frac{\partial^2 P_i}{\partial \lambda^2} + (\alpha - \beta\lambda) \; \frac{\partial P_i}{\partial \lambda} - \lambda \; P_i + i \; \lambda \; P_{i-1} = \frac{\partial P_i}{\partial T}. \tag{8}$$

The Appendix shows that this partial differential equation for $P_i(\lambda, T)$ has the following (poly-affine) closed-form solution,

$$P_i(\lambda, T) = A(T) \; e^{-B(T)\lambda} \; \sum_{j=0}^{i} C_{i,j}(T) \; \lambda^j, \tag{9}$$

where

$$A(T) = \exp\left(\frac{\alpha(\beta - \xi)T}{\sigma^2}\right)\left(\frac{2\xi}{\beta + \xi - (\beta - \xi)e^{-\xi T}}\right)^{\frac{2\alpha}{\sigma^2}}, \tag{10}$$

$$B(T) = \frac{2\xi(\beta + \xi)}{\sigma^2(\beta + \xi - (\beta - \xi)e^{-\xi T})} - \frac{\beta + \xi}{\sigma^2}, \tag{11}$$

and $\xi = \sqrt{\beta^2 + 2\sigma^2}$. The first $C_{i,j}(T)$ function is $C_{0,0}(T) = 1$. The remaining $C_{i,j}(T)$ functions are given as solutions of the recursive system of first-order ordinary differential equations,

$$\frac{dC_{i,i}}{dt} = i \; C_{i-1,i-1} \; - (\sigma^2 B(t) + \beta) \; i \; C_{i,i}, \tag{12}$$

$$\frac{dC_{i,j}}{dt} = i \; C_{i-1,j-1} \; - (\sigma^2 B(t) + \beta) \; j \; C_{i,j} \; + (j+1) \; (\alpha + j\sigma^2/2) \; C_{i,j+1}, \tag{13}$$

$$\frac{dC_{i,0}}{dt} = \alpha \; C_{i,1}, \tag{14}$$

where $1 \leq j \leq i-1$. These differential equations are easily solved numerically subject to the initial condition that $C_{i,j}(0) = 0$ for all $i > 0$.

With these solutions, the expectation of an arbitrary function $F(L_t)$ of the portfolio losses (satisfying appropriate regularity conditions of course) can be calculated directly by the expression

$$E[\ F(L_t)\ ] = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\ \frac{P_{1,i}(\lambda_1,t)}{i!}\ \frac{P_{2,j}(\lambda_2,t)}{j!}\ \frac{P_{3,k}(\lambda_3,t)}{k!}\ F(L_t). \tag{15}$$

Although the summations range from zero to infinity, only the first few terms generally need to be evaluated since the remainder are negligible.

5. VALUING TRANCHES

Given the solutions for the Poisson probabilities, it is now straightforward to value securities with cash flows tied to the realized credit losses of an underlying portfolio such as the CDX index. In practice, cash flows for the CDX index and its tranches occur discretely (typically on a quarterly frequency) and these contracts can be valued by applying Equation (15) to each discounted cash flow and then summing over all cash flows. To build intuition about these contracts, however, this section provides formal expressions for the values of the CDX index spread and the spreads on its tranches using the convenient assumption that cash flows are paid continuously. Observe, however, that these formal expressions are simply illustrative; the empirical results presented in the paper are based on the actual discrete cash flows and are valued using Equation (15) to evaluate expectations under the risk-neutral measure.

Let $D(t)$ denote the present value (as of time zero) of a zero-coupon riskless bond with a maturity date of t. Turning first to the CDX index, recall that a protection buyer pays a fixed annuity of c on the remaining balance of the CDX portfolio in exchange for protection against default losses as they occur. Thus, the value of the premium leg can be expressed as

$$c\ E\left[\int_0^T\ (1-L_t)\ e^{-\left(\int_0^t r_s ds\right)}\ dt\right]. \tag{16}$$

This expression reflects that the fixed annuity is received continuously on the remaining notional amount of the index, $1 - L_t$. The exponential term in the integral is the discount factor applied to the fixed annuity payments. For simplicity, we will assume that the riskless interest rate is independent of the Poisson and intensity processes. Similarly, the value of the protection leg is given by

$$E\left[\int_0^T \ e^{-\left(\int_0^t r_s ds\right)} \ dL\right], \tag{17}$$

where dL represents the change in the total cumulative losses on the portfolio, which is just the instantaneous realized loss. Since L_t is a nondecreasing process, the integral in the above equation represents a standard Riemann-Stieltjes integral. The Appendix provides a closed-form solution for the value of c implied by the two expressions above.

As an aside, we note that a much more convenient expression for the value of c can be obtained by slightly modifying the definition of the cash flows received from the CDX index. Specifically, by assuming that the cash flows from both legs of the CDX contract are paid on the unamortized balance (rather than the amortized balance), the value of c can be expressed as

$$c = \frac{\int_0^T \ D(t) \ (\bar{\gamma}_1 \ E[\lambda_{1t}] + \ \bar{\gamma}_2 \ E[\lambda_{2t}] \ + \bar{\gamma}_3 \ E[\lambda_{3t}]) \ dt}{\int_0^T \ D(t) \ dt}. \tag{18}$$

where the expected value of an intensity process is given by

$$E[\lambda_t] = e^{-\beta t}\lambda + \left(\frac{\alpha}{\beta}\right)(1 - e^{-\beta t}). \tag{19}$$

Since the same amortization is applied to both legs on the contract, this simplification has very little effect on the value of c. In the special case where the intensity processes follow martingales ($\alpha = \beta = 0$), Equation (18) reduces further to give

$$c = \bar{\gamma}_1 \ \lambda_1 \ + \ \bar{\gamma}_2 \ \lambda_2 \ + \ \bar{\gamma}_3 \ \lambda_3. \tag{20}$$

Thus, the CDX index spread becomes a simple linear combination of the current jump intensities for the three Poisson processes in this special case.

Turning now to the valuation of tranches, we observe that the total losses on an individual $N{-}M$ percent tranche can be modeled as a call spread on the underlying state variable L_t. Specifically, the total losses V_t on a $N{-}M$ tranche can be expressed as

$$V_t = \frac{1}{M - N}\big(\max(0, L_t - N) - \max(0, L_t - M)\big), \tag{21}$$

where N and M are denoted in decimal form. This expression reflects that if the total loss on the underlying portfolio L_t is less than N, then the loss on the tranche V_t is zero. If L_t is midway between N and M, the total loss on the tranche V_t is 0.50 or 50 percent. If L_t equals or exceeds M, the total loss on the tranche V_t equals 1.00 or 100 percent. As with the total losses on the underlying portfolio, V_t is a nondecreasing function of time.

As with the index, an investor in a index tranche receives a fixed annuity of h on the remaining balance $1 - V_t$ of the tranche, in exchange for compensating the protection buyer for the losses dV on the tranche. Thus, the value of the premium leg of a $N{-}M$ percent tranche is given formally by,

$$h\ E\left[\int_0^T\ (1 - V_t)\ e^{-\left(\int_0^t r_s ds\right)}\ dt\right]. \tag{22}$$

Similarly, the value of the protection leg of the $N{-}M$ percent tranche is given by

$$E\left[\int_0^T\ e^{-\left(\int_0^T r_s ds\right)}\ dV\right], \tag{23}$$

where this integral is once again a Riemann-Stieltjes integral. Setting the value of the two legs equal to each other and solving for the value of the tranche spread h gives,

$$h = \frac{\int_0^T\ D(t)\ E[dV]}{\int_0^T\ (1 - E[V_t])\ D(t)\ dt}. \tag{24}$$

The expectation $E[V_t]$ in this expression is easily evaluated by substituting the closed-form solutions for the Poisson probabilities into Equation (15). Although dV is itself expressible as a linear combination of call options on L_t, it is generally much more convenient to evaluate $E[dV]$ numerically by solving for the difference between $E[V_{t+\epsilon}]$ and $E[V_t]$ for some suitably small value of ϵ.

6. EMPIRICAL ANALYSIS

In this section, we estimate the model using the times series of CDX index values and the associated index tranche prices. We then examine how the model performs and explore the economic implications of the results.

6.1 The Empirical Approach

To make the intuition behind the results more clear, we focus on a simple special case of the model in which each of the intensity processes follows a martingale. Thus, we assume that the α and β parameters in Equations (3) through (5) are zero. As we will show, even this simplified specification allows us to fit the data with a very small

RMSE (and only marginal improvements would be possible by estimating the general case of the model).[20]

In this specification, there are six parameters that need to be estimated: the three jump size parameters γ_1, γ_2, and γ_3, and the three volatility parameters σ_1, σ_2, and σ_3. In addition, the values of the three intensity processes need to be estimated for each date. Our approach in estimating the model will be to solve for the parameter and intensity values that best fit the model to the data. In doing this, we estimate the model separately for each of the five CDX indexes. The reason for this is that there are slight differences in the composition of the individual indexes, potentially resulting in minor differences in parameter values.

Let us illustrate the estimation approach with the specific example of the CDX 1 index. The CDX 1 index was the on-the-run index from October 20, 2003 to March 20, 2004, and there are 65 observations for this index in the data set. To estimate the model, the algorithm first picks a trial set of values for the six γ and σ parameters. Then for each of the 65 days in the sample period (and conditional on the parameter values), we solve for the three values of the intensity process in the following way. First, we require that the values of the intensity processes fit exactly the market value of the CDX index spread. The other identification conditions are provided by requiring that the values of the intensity processes minimize the sum of squared errors between the market and model-implied tranche spreads for the $0-3$, $3-7$, $7-10$, $10-15$, and $15-30$ percent tranches. Thus, we are using six market prices (the index and five tranche spreads) to identify the three intensity processes for each date (conditional on the parameter values). Once we have estimated the values of the three intensity processes for each of the 65 days, we then evaluate the sum of squared errors for the five tranches over all 65 days. Finally, we iterate the process over different sets of parameter values until we find a single set of parameter values that minimizes the sum of squared errors over all 65 days in the sample period, given that the intensity values are chosen to match the CDX index exactly and minimize the sum of squared errors for the five tranches on each date. This algorithm is essentially nonlinear least squares and has been widely used in the finance literature in similar types of applications.[21] The optimization methodology we use is a direct search polytope algorithm that does not use the gradient or Hessian of the objective function. While this algorithm displays robust convergence properties for a variety of starting values, its direct search

[20]These parameter restrictions imply that the intensity process is absorbed at zero if it reaches zero. Thus, a more robust specification might allow for a small positive value for α. In actuality, however, the implied intensity values are generally many standard deviations away from zero. Thus, this technical consideration likely has little effect on the estimation results.

[21]See Longstaff, Mithal, and Neis (2005), Liu, Longstaff, and Mandell (2006), and many others.

nature (which keeps trying parameter values far removed from the current minimizing parameter vector in order to avoid local minima) admittedly makes it very slow to converge.[22] As a result, some of the optimizations for longer time series such as CDX4 take more than 12 hours of CPU time to complete. Clearly, more efficient optimization algorithms could reduce the computational time significantly. We use a similar procedure to estimate the model for the other CDX indexes.

It is important to note that parameter values and intensities are estimated for the risk-neutral pricing measure (not the objective or historical measure). A more general analysis might be to follow an approach such as Duffie and Singleton (1997) or Pan and Singleton (2005) in which the parameters under both the objective and risk-neutral measures—and therefore also the market prices of risk—are estimated by maximum likelihood. To estimate these market prices of risk accurately, however, one would likely need to have more than just the two years of data available to us given the well-known difficulty in estimating the drift of a diffusion process (and, therefore, market prices of risk). Accordingly, we limit our analysis to the estimation of parameters and intensity processes under the pricing or risk-neutral measure.

6.2 Testing for the Number of Factors

One of the fundamental issues that needs to addressed at the outset is the question of how many factors are actually needed in pricing CDOs. So far, we have presented a three-factor version of the model. Clearly, however, the model could be adapted to allow only one or two factors by simply setting the values of two or one of the intensity processes to zero.[23] In this section, we explore the issue of how many factors are needed by testing whether the two-factor version has incremental explanatory power relative to the one-factor version, and then whether the three-factor version has incremental explanatory power relative to the two-factor version. The testing methodology is based on straightforward chi-square tests of differences in the log sum of squared pricing errors.

These tests for the number of factors needed to price tranches also provide insight into an issue that is of fundamental importance in credit markets—default correlation. To see this, imagine the hypothetical situation where defaults are purely idiosyncratic and independent of each other. Furthermore, assume that the value of the intensity process for each firm in the CDX is a constant 0.01, and that the default of any firm

[22]As robustness checks for the results, we use a variety of starting values for the optimization. For example, we use starting values ranging from 0.0001 to 0.05 for γ_1, from 0.0001 to 0.20 for γ_2, and 0.0001 to 0.75 for γ_3 (and similarly for the σ_1, σ_2, and σ_3 parameters). The convergence results are robust to the choice of starting parameters.

[23]Similarly, the model could be extended to allow for four (or even more) factors by adding an intensity process in the obvious way.

results in a loss of 1/125 of the value of the portfolio (equal weights). Recall that the sum of independent Poisson realizations is itself a Poisson random variate with intensity equal to the sum of the intensities of the individual independent Poisson processes. In this independent default case, the total losses on the portfolio could be modeled as the realization of a single Poisson process with intensity of $125 \times 0.01 = 1.25$, and where each event results in a jump of 1/125. This means that only one factor would be needed to capture CDO pricing if defaults were in fact independent. Thus, a one-factor model is certainly an economically meaningful alternative model to the three-factor model developed earlier in the paper. On the other hand, if defaults are not independent and some correlation or clustering in default occurs, it would no longer be possible to represent total portfolio losses as the realization of a single Poisson event. Rather, we would expect to find that more than one Poisson process is needed to explain the cross section of tranche spreads across different attachment points. Thus, finding that a one-factor model is not sufficient to explain CDO pricing would represent direct evidence that the market expects correlation or clustering in the defaults of CDX firms.

Table 2 presents summary statistics for the pricing errors obtained by estimating one-factor, two-factor, and three-factor versions of the model. In each case, the values of the intensity processes are chosen to match the CDX index spread exactly. In the two-factor and three-factor models, the RMSE of the difference between market and model implied spreads for the five index tranches is also minimized. The table reports the RMSEs for each of the 0−3, 3−7, 7−10, 10−15, and 15−30 percent tranches individually, as well as the RMSE computed over all tranches. Table 2 also reports the p-values for the chi-square tests of the two-factor vs. one-factor and three-factor vs. two-factor specifications. In the one-factor specification, we estimate the two parameters γ_1 and σ_1, as well as N values of λ_1, where N is the number of days in the sample. In the two-factor specification, we estimate the four parameters γ_1, γ_2, σ_1, and σ_2 as well as N values each for λ_1 and λ_2. Thus, the one-factor specification is nested within the two-factor specification by imposing $N + 2$ restrictions; the chi-square statistic has $N + 2$ degrees of freedom. Similarly for the test of the three-factor vs. two-factor specification.

As shown in Table 2, the RMSEs for the one-factor version of the model are very large across all of the tranches. The overall RMSEs range from about 30 to 41 basis points. Increasing the number of factors to two results in a significant reduction in the RMSEs, both overall and across tranches. Typically, the overall RMSE for the two-factor version of the model is between about 5 and 14 basis points. For each CDX index, the incremental explanatory power of the two-factor version relative to the one-factor version is highly statistically significant. Note that this test is much more relevant than a simple comparison of the nonlinear least squares R^2s which are all high, primarily because of the extreme degree of variation in the data (some spreads are in thousands while others are in single digits).

The three-factor version of the model results in very small RMSEs. With the exception of the CDX 1 index, the overall RMSEs are all on the order of two to three basis points. In fact, the RMSE for CDX 5 is actually less than one basis point. Again, with the exception of the CDX 1 index, the incremental explanatory power of the three-factor version relative to the two-factor model is highly significant. Thus, the three-factor model provides a very close fit to the data. Accordingly, we will report results based on the three-factor version of the model throughout the remainder of the paper.

6.3 The Parameter Estimates

Table 3 reports the parameter estimates obtained from the three-factor model along with their asymptotic standard errors (Gallant (1975)). Focusing first on the estimates of the jump sizes, the table shows that there is a strong uniformity across the different CDX indexes. In particular, the jump sizes associated with the first Poisson process are in a tight range from 0.00387 to 0.00469. Since each firm in the CDX index has a weight of $1/125 = 0.008$ in the index, a jump size of, say, 0.004 is completely consistent with the interpretation that a jump in the first Poisson process represents the idiosyncratic default of an individual firm, where the implicit recovery rate for the firm's debt is 50 percent. If we adopt this interpretation, then the implied recovery rates implied by the estimated jump sizes are 56.6, 51.4, 50.3, 48.4, and 58.6 percent for the individual CDX indexes, respectively.[24]

The jump sizes for the second Poisson process are also very uniform across the CDX indexes, ranging from roughly 0.052 to 0.066. These values are consistent with the interpretation of the second Poisson process reflecting a major event in a specific sector or industry. As one way of seeing this, observe that virtually every broad industry classification is represented in the CDX index. In particular, the CDX index includes firms in the consumer durables, nondurables, manufacturing, energy, chemicals, business equipment, telecommunications, wholesale and retail, finance and insurance, health care, utilities, and construction industries. If we place the CDX firms into these 12 broad industry categories, then this implies that there are $125/12 = 10.42$ firms per category. Assuming a 50 percent recovery rate, a major event that resulted in the loss of an entire industry would lead to a total loss for the index of $10.42/125 \times 0.50 = 0.042$, which is on the order of magnitude of the jump size estimated for the second Poisson process. We note again, however, that a number of alternatives to this industry-event-risk interpretation could be equally valid.

The estimated jump sizes for the third Poisson process display somewhat more

[24]Historical recovery rates on corporate debt vary based on macroeconomic conditions, the seniority of the debt, the nature of the default, the rating of the issuer, and many other factors. For the senior unsecured debt referenced by the CDX indexes, the normal range of recovery between 1981 and the present has ranged from 20 to 70 percent according to Moody's (for example, see Gupton (2005)).

variation than for the other two Poisson processes, with values ranging from about 0.17 to 0.52. The average value across all five indexes is about 0.35. Again assuming a 50 percent recovery rate, a jump size of 0.35 associated with a realization of the third Poisson process can be given the interpretation as a major economic shock to the entire economy in which as many as 70 percent of all firms default of their debt. This is clearly a nightmare scenario that is difficult to imagine occuring. Potential examples of such a scenario might include nuclear war, a worldwide pandemic, or a severe and sustained economic depression. We note that the latter would need to be much more severe than any the U.S. has yet experienced, but has been observed elsewhere in a number of instances during the two-millenium-long experience of sovereign defaults and collapses in ancient Rome, Germany, Russia, and many other states (see Winker (1999)).

Turning now to the estimates of the volatility parameters, Table 3 shows that the volatility estimates of each of the three intensity processes are generally of the same order of magnitude. Specifically, with the exception of the first CDX index, the volatility parameters range from roughly 0.10 to 0.30 across all three processes and across all the CDX indexes. Recall from Equation (20) that the CDX index spread can be approximated as a linear combination of the three intensity processes. Thus, the volatility of the CDX index spread could be linked to these volatility parameters.

It is important to stress that these parameters are all estimated in sample, which leaves open the usual issue of how the model would fit out of sample. The fact that many of the estimated parameters are similar across different CDX indexes, however, provides some indirect support that the out-of-sample performance of the model might not be unreasonable. Finally, we note that the standard errors for a few of the parameters are large relative to the parameter estimates, particularly for the CDX 1 results and for the estimates of σ_3. In general, however, most of the other parameters appear to be reasonably precisely estimated.

6.4 The Intensity Processes

6.4.1 The time series

Fig. 2 plots the time series of the estimated values of the three intensity processes. Table 4 presents summary statistics for these estimated values. Again, the estimated intensities are all under the risk-neutral measure.

As shown, the first intensity process λ_1 ranges from roughly 0.50 to 1.50 during the sample period. For the majority of the sample period, this process takes values between 0.60 to 0.90 and displays a high level of stability. During the credit crisis of May 2005, however, this intensity process spiked rapidly to a value of 1.52, but then declined to just over 1.00 by the middle of June 2005. Thus, this spike was relatively short lived. Given the average value of λ_1 during the sample period, the expected (risk-neutral) waiting time until a firm-specific default is 1.16 years.

The second intensity process λ_2 ranges from a high of about 0.04 to a low of about 0.01 during the sample period. The value of this process is generally declining throughout the period. During the credit crisis, the value of this process doubled from about 0.015 to just over 0.030. After the crisis, the value of this process continued to decline. This suggests that the market implied probability of a major industry or sector crisis declined significantly during the past several years. Put another way, the expected waiting time for this type of event declined from roughly 28 years to 125 years during the sample period. The average (risk-neutral) waiting time for a realization of the second Poisson process is 41.5 years during the sample period.

The third intensity process λ_3 has more apparent variability across CDX indexes than do the other two intensity processes. In particular, the value of this process increases rapidly for the CDX 1 index, but then generally takes lower values for the other four CDX indexes. The apparent discontinuity in this process as it rolls from CDX 1 to 2 is probably related to the higher standard errors of the estimated parameters for CDX 1; the estimated parameters and values of the intensity processes for CDX 1 are likely much noisier than for the other indexes. As with the second intensity process, the third intensity process essentially doubles around the time of the credit crisis. The average value for this intensity process throughout the entire sample period is 0.00131. Thus, the implied risk-neutral probability of a catastrophic meltdown scenario is very small with an expected (risk-neutral) waiting time of about 763 years on average.

To illustrate the implications of the results for the risk-neutral portfolio loss distribution, Figure 3 plots the time series of loss distributions implied by the model. Specifically, the distributions shown are for total portfolio losses at the five-year horizon and are truncated to show only values ranging from zero to 16 percent (the probabilities for larger losses are visually difficult to distinguish from zero). Furthermore, to make the distributions easier to visualize, they are plotted as continuous functions. In reality, of course, the portfolio loss distribution is discrete. As shown, the distribution of portfolio losses is multimodal and displays considerable time-series variation.

6.4.2 Interpreting the factors

Although we have referred to the three Poisson processes as being consistent with idiosyncratic, industry or sector, and economywide credit events, respectively, it is important to stress that we have provided no direct evidence supporting this interpretation. Intuitively, it seems reasonable to think of the first Poisson variable as an idiosyncratic credit event given that its realization maps into a portfolio loss of roughly 0.004. Similarly, it also seems natural to interpret the third Poisson event as a serious credit event affecting a large fraction of firms throughout the economy. In contrast, however, the second Poisson process need not necessarily be an industry or sector event. In fact, it could just as easily represent a default event for a subset of firms related by a variety of other firm attributes.

One possible way to explore the economic role played by the second Poisson process is by examining its implications for the factor structure of credit spreads for individual firms. To see this, imagine that all credit risk was purely idiosyncratic and that the correlation of credit spread changes across firms was zero. This is clearly not the case since the average correlation of daily credit spread changes across firms in the CDX index is 0.245. Similarly, imagine that all credit risk was economywide. In this polar extreme case, all credit spread correlations would be one, which is again easily rejected by the data. Now imagine that credit risk was a blend of both idiosyncratic and economywide risk, where the relative proportion varies across firms. In this case, credit spreads would be cross-sectionally correlated, but factor analysis would reveal that there was one common factor driving credit spreads—the remaining variation in credit spreads would be purely idiosyncratic.

With these preliminaries, now consider the more realistic case corresponding to the model estimated in this paper in which there is idiosyncratic and economywide credit risk, but also clustered default risk for subsets of firms related to the second Poisson process. Assume that this default clustering occurs across firms in a way that has nothing to do with their industry or sector. In this scenario, a principal components analysis would reveal a common economywide component driving individual firm credit spreads, and then a number of other common factors affecting specific subsets of the firms, but unrelated to industry grouping.

To explore this interpretation, we do the following. First, we map each of the 125 firms in the CDX indexes into one of the 12 Fama-French industry categories. Averaging over all five CDX indexes, 5.92 percent are in consumer nondurables, 3.36 percent in consumer durables, 10.24 percent in manufacturing, 5.44 percent in energy, 3.20 percent in chemicals, 6.40 percent in business equipment, 8.00 percent in telecommunications, 5.60 percent in utilities, 11.68 percent in wholesale/retail, 3.20 percent in healthcare, 22.24 percent in finance, and 14.72 percent classified as other. Next, we extract out time series of CDS spreads for the 94 firms that are present in the CDX indexes throughout the sample period and also have traded stock. We then compute the correlation matrix of daily credit spread changes for these firms and perform a principal components analysis. Finally, we regress the principal components (the corresponding eigenvectors) on industry dummy variables for each firm. Clearly, if default clustering is unrelated to industry, then these dummy variables should not have cross-sectional explanatory power for the second, third, fourth, etc. principal components.

The results provide an number of interesting insights into the cross-sectional structure of credit risk. Credit risk is obviously not purely idiosyncratic; the first principal component explains more than 27 percent of the variation in credit spreads across firms. On the other hand, idiosyncratic risk appears to be the dominant nature of individual firm credit spreads. Specifically, the next five principal components only explain an incremental 5.1, 4.5, 3.5, 3.1, and 2.8 percent, respectively. Furthermore, eight principal components are required before more than 50 percent of the variation in credit spreads is explained.

Table 5 reports the results from the cross-sectional regression of the principal component weights on the industry dummy variables. As shown, the first principal component is consistent with the interpretation of an economywide credit variable affecting the majority of firms; all 12 of the industry dummy variables are highly significant. Although not shown, the regression coefficients for the industry dummy variables are remarkably uniform, ranging from about 0.08 to 0.11. Thus, the first factor can be viewed as a "parallel shift" in the credit spreads of all firms.

Moving beyond the first principal component, we can now test whether the default clustering in subsets of firms is related to industry categories. Recall that if the default clustering reflected by the second Poisson process has nothing to do with industry, then these principal components should be orthogonal to the industry dummy variables. In actuality, however, there appears to be a significant relation between many of the principal components and the industry dummies. For example, four of the industry dummy variables are significant for the second principal component and the corresponding adjusted R^2 is 0.419. The four significant industries are the manufacturing, energy, finance, and other industries. Similarly, the energy, telecommunications, and finance industry dummies are significant for the third principal component and the adjusted R^2 is 0.263. Interestingly, for the fourth through eighth principal components, only one or two of the industry dummy variables are significant at the five or ten percent level. Thus, were the R^2s for the regressions higher, there would be the possibility of almost a one-to-one mapping between these principal components and a specific industry or pair of industries.

It is important to provide some caveats at this point. For example, a number of the significant industry dummy variables have negative signs. Clearly, these negative signs muddy the interpretation of the relation between principal components and specific industries. Furthermore, the adjusted R^2s for many of the principal components are not high, indicating that industry grouping may only be a small part of the total picture in explaining the default clustering being captured by the second Poisson process. Despite these caveats, however, these results provide at least some evidence that after extracting the common economywide factor in individual firm credit spreads, industry does play a significant role in explaining the clustering of credit risk in subsets of firms. Thus, the interpretation that the second Poisson event reflects the default of firms in an industry, or a small group of industries, may not be completely without foundation.

To understand better how these results fit into a broader economic perspective, we also repeat the same exercise using daily stock return data for the same 94 firms and sample period. The results are shown in Table 6.

There are many similarities between the results for credit spreads and those for the stock returns for these 94 firms. For example, the first principal component for stock returns explains about 26.9 percent of the variation in returns, which is almost the same amount explained by the first principal component for credit spreads. As before, however, the second, third, fourth, etc. principal components explain only small proportions of the total variation, suggesting that much of the variability in stock returns is idiosyncratic.

The regression results indicate that the industry dummy variables have significant explanatory power for the stock return principal components. Similar to the results in Table 5, the first principal component loads on all 12 of the Fama-French industry dummy variables, consistent with the usual view of the first factor in stock returns being related to the market. Where the results differ from those for credit spreads is in the implications for the other principal components. For example, the industry dummy variables explain more than 82 percent of the variation in the loadings for the second principal component. This is a much higher proportion than in the credit spread results. Furthermore, eight of the 12 industry dummies have significant explanatory power for the second stock return principal component. Similar results hold for the third, fourth, fifth, etc. principal components: principal components for the stock returns are much more related to industry than is the case for credit spreads. Furthermore, few, if any, of the stock return principal components can be linked to one or two industries; stock return principal components seem to be related to broader subsets of firms in the economy than is the case for credit spreads. These results are intriguing and argue for a more in-depth comparison of the cross-sectional structure of credit spread changes and that of stock returns for the corresponding firms than we are able to provide in this paper.

6.5 CDX Index Spread Components

There are several ways in which the marginal impact of each type of default risk on the overall CDX index spread can be evaluated, each of which gives very similar results. One particularly intuitive way of doing this is simply to adopt the slightly modified definition of CDX index contract cash flows discussed in Section 5. Recall that this implies that the CDX index spread can be expressed as a linear combination of the three intensity processes. Thus, we can decompose the CDX index spread into three distinct components to measure the approximate overall economic impact of idiosyncratic, industry, and economywide default risks. In particular, the idiosyncratic component of the CDX index spread is given by $\bar{\gamma}_1 \, \lambda_1$, the industry component is given by $\bar{\gamma}_2 \, \lambda_2$, and the economywide component is given by $\bar{\gamma}_3 \, \lambda_3$. Recall from Equation (20) that the sum of these three components approximates the value of the CDX index spread.

Table 7 reports summary statistics for these three components. As shown, idiosyncratic default risk accounts for about two-thirds of the total value of the CDX index spread across the different indexes. Interestingly, however, the percentage of the CDX index spread due to idiosyncratic default risk has increased steadily throughout the sample period. In particular, the percentage has increased from about 58 percent for the first two indexes to more than 82 percent for the CDX 5 index.

The portion of the CDX index spread due to industry or sector default risk declined significantly during the sample period. In particular, the portion due to industry default risk is about 33 percent for the CDX 1 index, but only about 10 percent for the CDX 5 index. This decline reflects the dramatic decrease in the value of the second intensity process λ_{2t} during the sample period.

Economywide default risk accounts for an average of about 7 to 12 percent of the total CDX index value. There is no clear trend in the size of this component during the sample period. We note, however, that this component takes its largest value for the CDX 4 index which spans the period during the May 2005 credit crisis.

In summary, these results indicate that idiosyncratic default risk constitutes the majority of the CDX index spread. The combined effects of industry and economywide risk, however, are also significant and have represented more than 40 percent of the total CDX index spread at times. Even though the probabilities of industrywide or catastrophic economywide default events are much smaller than the probability of an isolated idiosyncratic default, the economic impact of these types of events is much more severe. Thus, industry and economywide default risks have a disproportionately larger influence on the value of the overall CDX index.

Since the CDX index spread is an average of the default spreads for the underlying 125 firms in the index, the decomposition of the index into the three types of default events is equally applicable to the typical firm in the index. Thus, our results imply that the event of default for the typical firm can be partitioned into three mutually exclusive events: the event that only the firm defaults, the event that the firm and many of the other firms in its industry or sector default together, and the event that the firm and the majority of all firms in the economy default. On average, the probabilities of the three types of default events are 64.6, 27.1, and 8.3 percent, respectively, of the total default probability for the typical firm in the index.

Finally, we explore the extent to which variation in the size of these components affects the cross-sectional distribution of credit spreads. For example, if the portion of individual credit spreads due to the economywide factor increases, we would anticipate that credit spreads across firms would tend to be more similar or coherent. Similarly for an increase in the industry component. To explore this, we calculate the cross-sectional standard deviation of credit spreads each day for the 94 firms which appear in the CDX index throughout the sample period. We then regress this time series of cross-sectional standard deviations on the industry and economywide component percentages. The results from this regression are

$$\begin{array}{lllll} \text{Std. Deviation}_t & = 26.741 & -\ 7.247\ \text{Ind\%}_t & -\ 33.688\ \text{Econ\%}_t & +\ \epsilon_t \\ & \quad (14.28) & \quad (-1.79) & \quad (-3.26) & \end{array}$$

$R^2 = 0.028$, $N = 434$, and t-statistics in parentheses. Thus, as the industry and economywide components become larger, credit spreads across firms become more homogeneous, consistent with the interpretation that the factors capture some of the commonalities in the economic structure of credit spreads.

6.6 The Time Series of RMSEs

Although Table 2 reports summary RMSE statistics for the three-factor model, it is also interesting to examine the time series variation in the ability of the model to capture market tranche spreads more closely. Accordingly, Figure 4 plots the time series of daily RMSEs obtained by fitting the model to the five tranches.

As shown, the ability of the model to match market tranche spreads increased significantly during the sample period. Initially, some of the RMSEs are as large as 19 basis points. The RMSEs decline rapidly, however, and are on the order of five basis points by early 2004. By mid 2004, the RMSEs decline further and hover around two basis points for most of the sample period. The only exception is around the May 2005 credit crisis when the RMSE increases slightly to about five basis points. After the crisis, however, the RMSEs decline rapidly and reach values below one basis point near the end of the sample period. The small spikes in the RMSEs at the beginning and end of each index series are potentially due to the effects of investors rolling positions from tranches based on the previous index to tranches based on the new on-the-run index.

These results provide evidence consistent with the view that while the fledging index tranche market may have experienced some inconsistencies in the relative pricing of individual tranches, the market matured rapidly and pricing errors were quickly arbitraged away as learning occured. In fact, this market now appears to be very consistent in terms of the relative pricing of tranches. In particular, the fact that the RMSEs increased to only about five basis points during the credit crisis of May 2005 makes a strong case that the model captures the underlying economics of the market.

6.7 Pricing Errors

We turn next to the pricing errors, defined as the difference between the model implied spreads and the market spreads for the various CDX index tranches, and examine their properties. Table 8 presents summary statistics and reports t-statistics for the significance of the average pricing errors.

As shown, the pricing errors from the three factor model are surprisingly small across all indexes and tranches. In particular, the average pricing errors are all within 10 basis points of zero and most are within one or two basis points of zero. Recall from Table 1 that the average sizes for the $0-3$, $3-7$, $7-10$, $10-15$, and $15-30$ percent tranche spreads are about 1759, 240, 82, 34, and 12 basis points, respectively. Thus, average pricing errors of only a few basis points are extremely small in percentage terms as well.

The pricing errors for the equity $0-3$ percent tranches are particularly small. With the exception of the CDX 1 index, the average pricing errors for the equity tranche are all within 1.5 basis points of zero. Pricing errors this small would clearly be well within the bid-ask spread for these securities. Even the average pricing error of −7 basis points for the CDX 1 index is well within one percent of the average size of the 0−3 percent tranche spread.

The pricing errors for the junior mezzanine 3−7 percent tranche are uniformly small across all the indexes. In most cases, the average pricing errors are inside of a single basis point. Again, this is remarkably small in terms of the average size of the spreads for this tranche. The pricing errors for the mezzanine 7−10 percent tranche are the smallest of all of the tranches. The pricing errors for the senior 10−15 percent tranche are also generally small. For the CDX 3, 4, and 5 indexes, the average pricing errors for this tranche are all less than one basis point in magnitude.

Finally, the pricing errors for the senior 15−30 percent tranche are also generally small. Since the average spread on this tranche is only about 12 basis points, however, the percentage pricing errors on this tranche are probably the largest among all of the tranches. Despite this, these average pricing errors are generally well within the two to three basis points bid-ask spread for this tranche observed in the market.

Table 8 also reports the t-statistics for the mean (correcting for the persistence in the pricing errors). As shown, the only statistically significant mean pricing errors occur for the CDX 1 index. Specifically, the pricing errors for the 0−3 percent equity and 15−30 percent senior tranches have significant means for the CDX 1 index. Thus, these results are consistent with the interpretation that there may have been arbitrage opportunities across CDX tranches during the early stages of the market.

While the results about the size of the pricing errors are encouraging, it is important to acknowledge that the model is rejectable. For example, most of the first-order serial correlation coefficients are very large, indicating that there is a high degree of persistence in the pricing errors. Thus, in principle, it might be possible to construct a trading rule that exploited model mispricings. On the other hand, the fact that the errors are typically much smaller than realistic transaction costs for trading these tranches argues against the potential profitability of such a trading rule.

6.8 Linking Firm-Level and Portfolio-Level Information

In this paper, we have focused on the implications of the data for the distribution of default losses for a large portfolio of credit-sensitive contracts. Ideally, we would like to be able to use portfolio-level information to infer something about the economic nature of credit risk at the individual-firm level.

One way to do this is to solve for the default correlation among individual firms implied by the data. The event that firm i has defaulted by time T can be characterized is a simple binary or Bernoulli variate with probability $\pi_i = 1 - e^{-\xi T}$, where ξ is a firm-specific constant. With this structure, it is straightforward to show that the value of the premium leg for a five-year quarterly-pay CDS contract on the firm is

$$\frac{s}{4} \sum_{t=1}^{20} D(t/4)\ e^{-\xi t/4}, \tag{25}$$

and the value of the protection leg for the CDS contract is

$$\frac{w}{4} \sum_{t=1}^{20} D(t/4)\ \xi\ e^{-\xi t/4}, \tag{26}$$

where s is the CDS premium, and w is the write-down fraction on the firm's debt in the event of a default. Thus, the value of ξ is given immediately from the CDS premium for the firm by the relation $\xi = s/w$.

The joint distribution of losses on the 125 firms in the CDX index is a multivariate correlated Bernoulli distribution. As discussed in Marshall and Olkin (1985), Park, Park, and Shin (1996), Lunn and Davies (1998), and many others, this distribution is very difficult to characterize in either closed-form or via simulation. To solve for the implied correlation, however, we do not need to evaluate the joint density explicitly. In particular, assume that the event of a default of any firm translates into a loss fraction of 0.004 for the CDX portfolio. Given this structure, the variance of the loss distribution for the CDX portfolio at horizon T is given by

$$0.004^2 \sum_{i=1}^{125}\ \pi_i\ (1-\pi_i) + 0.004^2 \sum_{i=1}^{125} \sum_{\substack{j=1 \\ j \neq i}}^{125} \rho_{ij} \sqrt{\pi_i\ \pi_j\ (1-\pi_i)\ (1-\pi_j)}. \tag{27}$$

where ρ_{ij} is the pairwise correlation coefficient for the Bernoulli variates for firms i and j. To solve for an implied correlation it is necessary to place some additional structure on the correlations. For simplicity, we assume that ρ_{ij} is constant for all i, and j, $i \neq j$. With this assumption, it is now straightforward to solve for the variance of the portfolio loss distribution implied by CDO prices, set it equal to the above expression, and then solve for the implied default correlation.[25]

Figure 5 plots the time series of the implied correlation estimates. The implied correlation typically ranges from about 0.05 to 0.10 during the sample period. The average value and standard deviation of the implied correlation during the sample period are 0.0835 and 0.0177, respectively. Around the credit crisis of May 2005, however, the implied correlation increases to about 0.13, but then rapidly declines.

[25]The data for the CDS levels of the individual firms in the CDX indexes are also provided by Citigroup.

The lowest values for the implied correlation occur near the end of the sample period and are in the neighborhood of 0.04.

The expression in Equation (27) also provide insights about how the cross-sectional structure of credit risk affects the portfolio loss distribution. From this expression, it is immediately clear that the variance of the loss distribution is an increasing function of the individual pairwise correlations. This is completely consistent with the usual portfolio intuition that as correlations increase, the portfolio is less-well diversified, resulting in a higher portfolio variance.

What is less clear, however, is how cross-sectional dispersion or heterogeneity in the individual firm CDS levels (holding fixed the average value) affects the distribution of portfolio losses. This question is an important from an economic standpoint given some evidence that the cross-sectional distribution of credit spread tends to expand and contract in response to changes in macroeconomic variables.[26] Thus, an economy in which all firms have a credit spread of 100 basis points may be very different from one in which half of all firms have a credit spread of 50 basis points, and the other half have a credit spread of 150 basis points (even though the CDX index is the same in both cases).

From Equation (27), it can be seen that if default risk is uncorrelated, then the variance of portfolio losses is a decreasing function of dispersion (since the variance is then a concave function of the CDS levels). On the other hand, it is easily seen that if some correlations ρ_{ij} are negative, then the variance of portfolio losses could be an increasing function of dispersion. Thus, the sign of the relation cannot be determined on purely analytical grounds.

To determine the sign of the relation empirically, we first calculate the standard deviation of CDS premia for the 125 firms in the CDX index for each day in the sample period, which we term the dispersion measure. We then regress the standard deviation of the portfolio loss distribution implied by the estimated parameters of the model on the CDX index and the cross-sectional dispersion measure. The regression results are

$$\begin{array}{rlllll} \text{Std. Deviation}_t & = 0.0089 & + 3.3763\ \text{CDX}_t & + 0.8002\ \text{Dispersion}_t & + \epsilon_t \\ & \quad (10.59) & \quad (18.95) & \quad (10.30) & \end{array}$$

$R^2 = 0.681$, $N = 435$, and t-statistics in parentheses. As shown, both the CDX index spread and the dispersion measure have significant explanatory power for the implied standard deviation of the loss distribution. The sign of the coefficient for the dispersion measure is positive, indicating that greater dispersion or heterogeneity in CDS levels tends to increase the volatility of the loss distribution.

[26] For example, see Kwan (1996), Elton, Gruber, Agrawal, and Mann (2001), Collin-Dufresne, Goldstein, and Martin (2001), and Wang and Zhang (2006).

7. CONCLUSION

This paper uses the information in the prices of synthetic CDX index tranches to study the market's expectations about how corporate defaults cluster in various economic environments—the cross-sectional structure of default risk. To do this, we first develop a new portfolio credit model in which three type of Poisson events generate portfolio credit losses. Using an extensive data set of CDX index and tranche spreads, we then estimate the model and evaluate its performance.

A number of interesting results emerge from this analysis. For example, we find that one-factor and even two-factor models are insufficient to explain the relative pricing of CDX index tranches in the market. In contrast, a three-factor model that allows for jumps of approximate sizes 0.004, 0.06, and 0.35 explains virtually all of the cross-sectional and time-series variation in the index tranche data. Assuming a 50 percent recovery rate, a jump of 0.004 has the clear interpretation of an idiosyncratic default of a single firm out of the 125 firms in the CDX index ($0.50 \times 1/125 = 0.004$). Similarly, a jump size of 0.06 has the interpretation of a default event in which roughly 10 percent of the firms in the index default together. This could be viewed as an event in which an entire industry or sector experiences financial distress. Finally, a jump size of 0.35 represents the realization of a financial catastrophe in which more than 50 percent of the firms in the economy default on their debt. The results indicate that under the risk-neutral measure, the average expected time until a realization of these three types of events is 1.2, 41.5, and 763 years, respectively.

The results provide a number of insights into the important issue of default clustering. In particular, we find that the market expects significant clustering to occur. We show that roughly one-third of the value of the default spread for the typical firm in the CDX index is due to events in which multiple firms default together.

These results have a number of important economic implications. For example, they suggest that a significant portion of corporate credit risk may not be diversifiable. This has immediate implications for portfolio choice, the cost of corporate debt capital, and the systemic risk of financial institutions. Furthermore, since correlated default risk necessarily translates into correlated shocks to the stock values of the corresponding firms, these results may also have implications for the extreme risks being priced in equity markets.

APPENDIX

There are several ways in which the partial differential equation for P_i can be derived. For example, the approach outlined in Karlin and Taylor (1981) pp. 202-204 leads directly to the partial differential equation. To provide an alternative approach, recall that

$$P_i = E_t \left[\exp \left(-\int_0^T \lambda_s \; ds \right) \left(\int_0^T \lambda_s \; ds \right)^i \right]. \qquad (A1)$$

Let

$$H_t = \int_0^t \lambda_s \; ds. \qquad (A2)$$

This implies

$$dH_t = \lambda_t \; dt. \qquad (A3)$$

Now, rewrite P_i as

$$P_i = E_t \left[\exp \left(-\int_0^t \lambda_s \; ds - \int_t^T \lambda_s \; ds \right) \left(\int_0^t \lambda_s \; ds + \int_t^T \lambda_s \; ds \right)^i \right], \qquad (A4)$$

$$= E_t \left[\exp \left(-H_t - \int_t^T \lambda_s \; ds \right) \left(H_t + \int_t^T \lambda_s \; ds \right)^i \right]. \qquad (A5)$$

From this expression, P_i can be expressed explicitly as a function of λ_t, H_t, and $\tau = T - t$. An application of Itô's Lemma gives

$$dP_i = (\alpha - \beta\lambda) \; \frac{\partial P_i}{\partial \lambda} \; dt + \sigma \; \sqrt{\lambda} \; \frac{\partial P_i}{\partial \lambda} \; dZ + \frac{\sigma^2 \lambda}{2} \; \frac{\partial^2 P_i}{\partial \lambda^2} \; dt$$
$$- \frac{\partial P_i}{\partial \tau} \; dt - \lambda \; P_i \; dt + i \; \lambda \; P_{i-1} \; dt. \qquad (A6)$$

Since P_i is a martingale, however, the expected value of $dP_i = 0$. Thus,

$$\frac{\sigma^2 \lambda}{2} \frac{\partial^2 P_i}{\partial \lambda^2} + (\alpha - \beta\lambda) \frac{\partial P_i}{\partial \lambda} - \lambda \ P_i \ + i \ \lambda_t \ P_{i-1} = \frac{\partial P_i}{\partial \tau}, \tag{A7}$$

which is Equation (8) (at $t = 0$). For $i = 0$, the boundary condition is $P_0(\lambda, 0) = 1$. For all other i, the boundary condition is $P_i(\lambda, 0) = 0$.

For the case $i = 0$, the solution to the partial differential equation is identical to that provided by Cox, Ingersoll, and Ross (1985) in obtaining zero-coupon bond prices. The solution for this case can be expressed as shown in Equation (9) with $i = 0$. For $i > 0$, we conjecture that the solution is of the form shown in Equation (9). Differentiating the conjectured solution for $P_i(\lambda, T)$, substituting into Equation (10), and collecting terms in the powers of λ leads to the system of first-order differential equations shown in Equations (12), (13), and (14). This system can be solved recursively following in the order $C_{1,1}$, $C_{1,0}$, $C_{2,2}$, $C_{2,1}$, $C_{2,0}$, $C_{3,3}$, $C_{3,2}$, $C_{3,1}$, $C_{3,0}$, etc. Thus, for each i, we solve for $C_{i,j}$, where j runs backwards from i to 0.

Turning now to the derivation of the CDX index value in Section 5, we observe that conditional on the path of the intensity process, the expectation of $\exp(-\gamma N_t)$ is

$$E\left[e^{-\gamma N_t} \mid \ \cdot \ \right] = \exp\left(-\bar{\gamma} \ \int_0^t \lambda_s \ ds\right), \tag{A8}$$

This follows from the definition of the moment generating function for Poisson variates. Similarly, the conditional expectation of dN_t is $\lambda_t \ dt$. Since N_t and dN_t are conditionally independent (but not unconditionally), the conditional expected value of terms involving the product $e^{-\gamma N_t} \ dN_t$ is simply the product of these two conditional expectations. From Equation (A8), one minus the expected loss conditional on the paths of the intensity processes is given by

$$1 - E[L_t \mid \cdot] = \exp(-\int_0^t \bar{\gamma}_1 \ \lambda_{1s} + \bar{\gamma}_2 \ \lambda_{2s} + \bar{\gamma}_3 \ \lambda_{3s} \ ds). \tag{A9}$$

Using the results above in conjunction with Equations (16) and (17) implies that the CDX index spread can be expressed formally as

$$c = \frac{\int_0^T D(t)\ E\left[\exp(-\int_0^t \bar{\gamma}_1 \lambda_{1s} ds + \bar{\gamma}_2 \lambda_{2s} ds + \bar{\gamma}_3 \lambda_{3s} ds)(\bar{\gamma}_1 \lambda_{1t} + \bar{\gamma}_2 \lambda_{2t} + \bar{\gamma}_3 \lambda_{3t})\right]\ dt}{\int_0^T D(t)\ (1 - E[L_t])\ dt}. \tag{A10}$$

To evaluate this expression, we need to provide solutions for the unconditional moments that appear in the numerator and denominator. The unconditional expectation for the portfolio loss $E[L_t]$ is

$$E[L_t] = 1 - E[e^{-\gamma_1 N_{1t}}]\ E[e^{-\gamma_2 N_{2t}}]\ E[e^{-\gamma_3 N_{3t}}], \tag{A11}$$

where

$$E[e^{-\gamma N_t}] = E\left[\exp\left(-\bar{\gamma} \int_0^t \lambda_s\ ds\right)\right] = \hat{A}(t) e^{-\hat{B}(t)\lambda}, \tag{A12}$$

and where $\hat{A}(t)$ and $\hat{B}(t)$ are the same as in Equations (10) and (11) except that ξ is replaced by the constant $\sqrt{\beta^2 + 2\sigma^2 \bar{\gamma}}$. This result is obtained by a simple extension of the approach used in Cox, Ingersoll, and Ross (1985) in valuing zero-coupon bonds. These results imply that

$$1 - E[L_t] = \hat{A}_1(t)\hat{A}_2(t)\hat{A}_3(t) \exp(-\hat{B}_1(t)\lambda_1 - \hat{B}_2(t)\lambda_2 - \hat{B}_3(t)\lambda_3), \tag{A13}$$

where the subscripted values of $\hat{A}(t)$ and $\hat{B}(t)$ correspond to the respective Poisson processes in the obvious way.

Multiplying out the expression in the numerator of Equation (A10) leads to products of terms of the form given in Equation (A12) and also terms of the form,

$$E\left[\exp\left(-\hat{\gamma} \int_0^t \lambda_s\ ds\right) \lambda_t\right]. \tag{A14}$$

These latter expectations, however, can be evaluated directly from the results in Duffie, Pan, and Singleton (2000). Specifically, the expectation in Equation (A14) has the form $\tilde{A}(t) \exp(-\hat{B}(t)\lambda)(1 + \tilde{C}(t))$ where $\hat{B}(t)$ is as defined above in Equation (A12), $\tilde{C}(t)$ is the solution to the following ordinary differential equation,

$$\frac{d\tilde{C}}{dt} = -(\sigma^2 \hat{B}(t) + \beta) \tilde{C}(t) - \alpha \tilde{C}(t)^2, \quad (A15)$$

with boundary condition $\tilde{C}(0) = 1$, and $\tilde{A}(t)$ is given as

$$\ln(\tilde{A}(t)) = \alpha \int \tilde{C}(s) - \hat{B}(s) \, ds, \quad (A16)$$

with boundary condition $\tilde{A}(0) = 1$. From the expressions for these unconditional expectations in Equations (A12) and (A14), we can then solve for the unconditional expectation in the numerator of Equation (A10). Specifically, this expectation is given by

$$\bar{\gamma}_1 \, \tilde{A}_1(t)e^{-\hat{B}_1(t)\lambda_1}(1 + \tilde{C}_1(t))\hat{A}_2(t)e^{-\hat{B}_2(t)\lambda_2}\hat{A}_3(t)e^{-\hat{B}_3(t)\lambda_3}$$

$$+ \bar{\gamma}_2 \, \hat{A}_1(t)e^{-\hat{B}_1(t)\lambda_1}\tilde{A}_2(t)e^{-\hat{B}_2(t)\lambda_2}(1 + \tilde{C}_2(t))\hat{A}_3(t)e^{-\hat{B}_3(t)\lambda_3}$$

$$+ \bar{\gamma}_3 \, \hat{A}_1(t)e^{-\hat{B}_1(t)\lambda_1}\hat{A}_2(t)e^{-\hat{B}_2(t)\lambda_2}\tilde{A}_3(t)e^{-\hat{B}_3(t)\lambda_3}(1 + \tilde{C}_3(t)). \quad (A17)$$

Similarly, the expectation in the denominator of Equation (A10) is given from Equation (A13). Substituting these expressions into Equation (A10) gives the closed-form solution for the CDX index spread c.

REFERENCES

Bakshi, Gurdip, Dilip Madan, and Frank Zhang, 2006, Investigating the role of systematic and firm-specific factors in default risk: Lessons from empirically evaluating credit risk models, *Journal of Business* 79, 1955-1988.

Collin-Dufresne, Pierre, Robert S. Goldstein, and J. Spencer Martin, 2001, The determinants of credit spread changes, *Journal of Finance* 56, 2177-2207.

Cox, John, Jonathan Ingersoll, Jr., and Stephen Ross, 1985, A theory of the term structure of interest rates, *Econometrica* 53, 385-407.

Das, Sanjiv, Darrell Duffie, Nikunj Kapadia, and Leandro Saita, 2005, Common failings: How corporate defaults are correlated, *Journal of Finance*, forthcoming.

Das, Sanjiv, Laurence Freed, Nikunj Kapadia, and Gary Geng, 2006, Correlated default risk, *Journal of Fixed Income* Fall, 7-32.

Demarzo, Peter, 2005, The pooling and tranching of securities: A model of informed intermediation, *Review of Financial Studies* 18, 1-35.

Diamond, Douglas, 1984, Financial intermediation and delegated monitoring, *Review of Economic Studies* 60, 383-414.

Duffie, Darrell, Andreas Eckner, Guillaume Horel, and Leandro Saita, 2006, Frailty correlated default, Working paper, Stanford University.

Duffie, Darrell, and Nicolae Gârleanu, 2001, Risk and valuation of collateralized debt obligations, *Financial Analysts Journal* 57 (January/February), 41-59.

Duffie, Darrell, Jun Pan, and Kenneth J. Singleton, 2000, Transform analysis and asset pricing for affine jump-diffusions, *Econometrica* 68, 1343-1376.

Duffie, Darrell, Leandro Saita, and Ke Wang, 2006, Multi-period corporate default prediction with stochastic covariates, *Journal of Financial Economics*, forthcoming.

Duffie, Darrell, and Kenneth J. Singleton, 1997, An econometric model of the term structure of interest-rate swap yields, *Journal of Finance* 52, 1287-1321.

Duffie, Darrell, and Kenneth J. Singleton, 1999, Modeling term structures of defaultable bonds, *Review of Financial Studies* 12, 687-720.

Duffie, Darrell, and Kenneth J. Singleton, 2003, *Credit Risk: Pricing, Measurement, and Management*, (Princeton University Press, Princeton, NJ).

Elton, Edwin J., Martin J. Gruber, Deepak Agrawal, and Christopher Mann, 2001, Explaining the rate spread on corporate bonds, *The Journal of Finance* 56, 247-277.

Eom, Young Ho, Jean Helwege, and Jing-zhi Huang, 2004, Structural models of corporate bond pricing: An empirical analysis, *Review of Financial Studies* 17, 499-544.

Errais, Eymen, Kay Giesecke, and Lisa Goldberg, 2006, Pricing credit from the top down with affine point processes, Working paper, Stanford University.

Frey, Rudiger, and Jochen Backhaus, 2005, Portfolio credit risk models with interacting default intensities: A Markovian approach, Working paper, Department of Mathematics, University of Leipzig.

Gallant, A. Ronald, 1975, Nonlinear regression, *The American Statistician* 29, 73-81.

Giesecke, Kay, 2004, Correlated default with incomplete information, *Journal of Banking and Finance* 28, 1521-1545.

Giesecke, Kay, and Lisa Goldberg, 2005, A top down approach to multi-name credit, Working paper, Cornell University.

Gupton, Greg M., 2005, Advancing loss given default prediction models: How the quiet have quickened, *Economic Notes* 34, 185-230.

Hull, John, and Alan White, 2003, Valuation of a CDO and a n-th to default CDS without Monte Carlo simulation, Working paper, University of Toronto.

Jorion, Philippe, and Gaiyan Zhang, 2005, Intra-industry credit contagion: Evidence from credit default swap and equity markets, Working paper, University of California at Irvine.

Karlin, Samuel, and Howard Taylor, 1981, A second course in stochastic processes, Academic Press Inc, Orlando, Florida.

Kindleberger, Charles P., 2005, *Manias, Panics, and Crashes: A History of Financial Crises, 5th Ed.*, John Wiley & Sons, Hoboken, NJ.

Kwan, Simon, 1996, Firm-specific information and the correlation between individual stocks and bonds, *Journal of Financial Economics* 40, 63-80.

Liu, Jun, Francis A. Longstaff, and Ravit E. Mandell, 2006, The market price of risk in interest rate swaps: The roles of default and liquidity risks, *The Journal of Business* 79, 2337-2359.

Longstaff, Francis A., Sanjay Mithal, and Eric Neis, 2005, Corporate yield spreads: Default risk or liquidity? New evidence from the credit default swap market, *Journal of Finance* 60, 2213-2253.

Lunn, A. Daniel, and Stephen J. Davies, 1998, A note on generating correlated binary variables, *Biometrika* 85, 487-490.

Marshall, Albert, and Ingram Olkin, 1985, A family of bivariate distributions generated by the bivariate Bernoulli distribution, *Journal of the American Statistical Association* 80, 332-338.

Pan, Jun, and Kenneth J. Singleton, 2005, Default and recovery implicit in the term structure of sovereign CDS spreads, Working paper, Stanford University.

Park, Chul G., Taesun Park, and Dong W. Shin, 1996, A simple method for generating correlated binary variates, *The American Statistician* 50, 306-310.

Rajan, Arvind, Glen McDermott, and Ratul Roy, 2007, *The Structured Credit Handbook*, John Wiley & Sons, Hoboken, NJ.

Richardson, Karen, 2005, Bankers hope for a reprise of 'Bowie Bonds', *Wall Street Journal*, August 23.

Roy, Ratul, and David Shelton, 2007, Trading credit tranches, in *The Structured Credit Handbook*, Arvind Rajan, Glen McDermott, and Ratul Roy, eds., John Wiley & Sons, Hoboken, NJ.

Saita, Leandro , 2005, The puzzling price of corporate default risk, Working paper, Stanford University.

Schönbucher, Philipp , 2005, Portfolio losses and the term structure of loss transition rates: A new methodology for the pricing of portfolio credit derivatives, Working paper, Department of Mathematics, ETH Zürich.

Sidenius, Jakob, Vladimir Piterbarg, and Leif Andersen, 2005, A new framework for dynamic credit loss modeling, Working paper, Royal Bank of Scotland.

Wang, Ashley, and Gaiyan Zhang, 2006, Institutional equity investment, asymmetric information, and credit spreads, Working paper, University of California at Irvine.

Winkler, Max, 1999, *Foreign Bonds: An Autopsy*, Beard Books.

Yu, Fan, 2005a, Correlated defaults in intensity-based models, *Mathematical Finance*, forthcoming.

Yu, Fan, 2005b, Default correlation in a reduced-form model, *Journal of Investment Management* 3(4), 33-42.

Table 1

Summary Statistics for the Levels and First Differences of the CDX North American Investment Grade Index and Index Tranche Spreads. This table reports summary statistics for the market spreads and the daily change in the spreads (spreads measured in basis points) for the indicated time series. Correlations shown in the top panel are correlations of levels; correlations shown in the bottom panel are correlations of first differences. Results are reported for the combined on-the-run time series. The sample period is from October 2003 to October 2005.

	Correlations													
	0-3	3-7	7-10	10-15	15-30	Mean	Std. Dev.	Min.	Med.	Max.	Skew.	Kurt.	Serial Corr.	N
CDX Index	0.809	0.668	0.607	0.672	0.880	54.52	6.95	41.00	54.50	78.00	0.25	−0.44	0.977	435
0−3 Tranche		0.183	0.068	0.217	0.711	1758.87	245.30	1282.52	1780.28	2914.55	1.05	3.28	0.978	435
3−7 Tranche			0.960	0.968	0.610	240.07	85.66	103.50	232.00	411.00	0.13	−1.33	0.992	435
7−10 Tranche				0.933	0.540	82.27	36.73	24.75	74.75	158.00	0.22	−1.23	0.992	435
10−15 Tranche					0.674	34.43	14.49	11.88	29.50	66.50	0.39	−1.16	0.989	435
15−30 Tranche						11.54	2.89	5.00	11.88	20.75	0.03	−0.42	0.962	435
Δ CDX Index	0.896	0.872	0.761	0.750	0.730	−0.02	1.51	−7.00	0.00	7.00	0.02	3.86	0.013	434
Δ 0−3 Tranche		0.762	0.648	0.662	0.651	0.48	51.63	−221.86	0.27	285.15	0.52	6.38	0.081	434
Δ 3−7 Tranche			0.879	0.840	0.773	−0.55	10.88	−54.00	0.00	40.50	−0.44	4.24	−0.028	434
Δ 7−10 Tranche				0.861	0.751	−0.20	4.68	−24.50	0.00	19.50	0.07	3.30	−0.079	434
Δ 10−15 Tranche					0.764	−0.09	2.14	−8.50	0.00	9.50	0.17	3.00	−0.067	434
Δ 15−30 Tranche						−0.02	0.79	−3.00	0.00	4.50	0.39	4.10	−0.043	434

Table 2

Root Mean Squared Errors (RMSE) from Model Fitting and Tests of the Number of Factors. This table reports the RMSEs for the individual CDX index tranches resulting from fitting the indicated models, where the RMSE is calculated from the pricing errors for the individual tranche. The table also reports the overall RMSE which is calculated from the pricing errors for all five of the tranches. All RMSEs are measured in basis points. The p-value is for the test of n vs. $n-1$ factors. N denotes the number of observations for the indicated CDX index. The sample period is from October 2003 to October 2005.

Number of Factors	Index	Tranche RMSE 0−3	3−7	7−10	10−15	15−30	Overall RMSE	p-Value	R^2	N
One Factor	CDX 1	13.13	78.55	17.78	40.15	13.44	41.11	–	0.9924	65
	CDX 2	26.90	54.02	47.08	47.50	13.37	40.73	–	0.9931	108
	CDX 3	39.21	65.48	41.22	24.75	9.29	40.55	–	0.9887	118
	CDX 4	31.81	39.76	47.65	23.28	12.38	33.35	–	0.9966	127
	CDX 5	49.58	29.92	26.37	13.83	6.21	29.25	–	0.9965	17
Two Factors	CDX 1	9.89	20.24	8.46	6.73	9.82	12.01	0.000	0.9993	65
	CDX 2	5.60	17.17	5.12	5.89	3.02	8.90	0.000	0.9997	108
	CDX 3	3.78	5.77	3.28	1.68	3.07	4.74	0.000	0.9998	118
	CDX 4	22.40	14.30	9.95	7.39	10.85	13.98	0.000	0.9994	127
	CDX 5	2.24	1.95	7.13	6.99	0.66	4.67	0.000	0.9999	17
Three Factors	CDX 1	11.01	4.62	8.68	10.37	8.31	8.88	0.997	0.9996	65
	CDX 2	2.52	1.31	3.04	6.08	4.68	3.90	0.000	0.9999	108
	CDX 3	0.99	0.76	3.36	2.80	2.06	2.23	0.000	1.0000	118
	CDX 4	1.18	1.19	4.52	2.27	2.46	2.63	0.000	1.0000	127
	CDX 5	0.29	0.28	1.10	0.61	0.46	0.63	0.000	1.0000	17

Table 3

Parameter Estimates. This table reports the parameter estimates for the indicated CDX indexes. The jump size parameters are the parameters γ_1, γ_2, and γ_3 in the model. The volatility parameters are the σ_1, σ_2, and σ_3 parameters in the model. Asymptotic standard errors are in parentheses and are computed as in Gallant (1975). The sample period is from October 2003 to October 2005.

	Jump Size Parameters			Volatility Parameters			
Index	First	Second	Third	First	Second	Third	N
CDX 1	0.00453	0.06093	0.17091	0.03016	0.10847	0.00098	65
	(0.00820)	(0.00215)	(0.00925)	(3.98870)	(0.11198)	(2.18805)	
CDX 2	0.00411	0.06498	0.35104	0.20854	0.19569	0.14246	108
	(0.00064)	(0.00020)	(0.11342)	(0.04624)	(0.00542)	(0.13661)	
CDX 3	0.00402	0.06621	0.35347	0.11955	0.16863	0.18110	118
	(0.00165)	(0.00027)	(0.10936)	(0.20686)	(0.00729)	(0.10325)	
CDX 4	0.00387	0.05260	0.51615	0.14003	0.25083	0.16539	127
	(0.00007)	(0.00081)	(0.26476)	(0.01756)	(0.00530)	(0.23667)	
CDX 5	0.00469	0.05628	0.33801	0.17315	0.27763	0.29674	17
	(0.00072)	(0.00509)	(0.10590)	(0.5473)	(0.06554)	(0.06784)	

Table 4

Summary Statistics for the Estimated Intensity Processes. This table reports summary statistics for the first, second, and third intensity processes for the indicated CDX indexes. Values for the intensity processes represent estimates under the risk-neutral pricing measure. The sample period is from October 2003 to October 2005.

Intensity Process	Index	Mean	Standard Deviation	Minimum	Median	Maximum	N
First	CDX 1	0.72561	0.04280	0.64748	0.73229	0.80964	65
	CDX 2	0.85419	0.02374	0.79102	0.85175	0.90252	108
	CDX 3	0.76570	0.04105	0.68886	0.75330	0.86144	118
	CDX 4	1.02303	0.15058	0.73804	1.00283	1.51795	127
	CDX 5	0.81643	0.02452	0.79326	0.81125	0.86468	17
Second	CDX 1	0.03063	0.00307	0.02336	0.03108	0.03560	65
	CDX 2	0.03489	0.00450	0.02367	0.03632	0.04309	108
	CDX 3	0.02115	0.00411	0.01464	0.01999	0.03077	118
	CDX 4	0.01639	0.00669	0.00841	0.01395	0.03216	127
	CDX 5	0.00869	0.00065	0.00790	0.00864	0.00978	17
Third	CDX 1	0.00263	0.00118	0.00029	0.00261	0.00498	65
	CDX 2	0.00092	0.00040	0.00001	0.00091	0.00188	108
	CDX 3	0.00093	0.00013	0.00062	0.00095	0.00119	118
	CDX 4	0.00136	0.00048	0.00043	0.00127	0.00244	127
	CDX 5	0.00102	0.00013	0.00087	0.00101	0.00131	17

Table 5

Principal Components Analysis and Industry Regression Results for Changes in Individual Firm CDS Spreads. This table reports the incremental and cumulative fraction explained by the indicated principal components for the correlation matrix of daily changes in individual firm CDS spreads for the subsample of 94 firms that appeared in the CDX index throughout the sample period and have traded stock. Also reported are t-statistics and adjusted R^2s from the cross-sectional regression of the principal component loadings or weights (the eigenvectors) on dummy variables for the 12 Fama-French industry groups. An asterisk denotes significance at the five-percent level.

Princ Comp	Fraction Explain	Cumul Explain	Non-Dura	Dura	Manuf	Energy	Chems	BusEq	Telcom	Utils	Shops	Health	Finan	Other	Adj. R^2
1	0.272	0.272	8.55*	5.73*	16.02*	10.59*	8.73*	11.72*	6.70*	11.66*	9.76*	5.75*	14.73*	16.56*	0.929
2	0.051	0.323	-0.81	1.75	-2.79*	-2.00*	-0.78	-0.56	0.16	-0.24	0.58	1.04	7.67*	-2.17*	0.419
3	0.045	0.368	0.84	-1.49	-0.77	-2.77*	-0.25	-0.13	1.96*	1.57	3.80*	0.36	-3.59*	1.13	0.263
4	0.035	0.403	-0.51	-0.26	0.06	0.50	-0.67	2.63*	-0.90	-0.11	-0.70	1.37	-0.92	-1.13	0.011
5	0.031	0.434	-1.91	-0.24	-0.19	0.05	-0.43	-0.35	0.70	-1.02	1.04	-1.03	0.37	1.11	-0.033
6	0.028	0.462	-0.02	0.55	-1.10	0.47	0.15	-0.43	0.85	-0.47	0.64	2.06*	0.36	0.93	-0.039
7	0.023	0.485	0.07	2.23*	-1.24	-0.48	1.23	-0.48	1.20	-2.32*	2.26*	0.05	-0.53	-0.30	0.083
8	0.022	0.507	-1.24	-0.18	0.34	0.16	-0.98	1.75	-3.21*	1.20	4.29*	-1.55	-1.17	-1.29	0.234

Table 6

Principal Components Analysis and Industry Regression Results for Individual Firm Stock Returns. This table reports the incremental and cumulative fraction explained by the indicated principal components for the correlation matrix of individual firm stock returns for the subsample of 94 firms that appeared in the CDX index throughout the sample period and had traded stock. Also reported are t-statistics and adjusted R^2s from the cross-sectional regression of the principal component loadings or weights (the eigenvectors) on dummy variables for the 12 Fama-French industry groups. An asterisk denotes significance at the five-percent level.

Princ Comp	Fraction Explain	Cumul Explain	Non-Dura	Dura	Manuf	Energy	Chems	BusEq	Telcom	Utils	Shops	Health	Finan	Other	Adj. R^2
1	0.269	0.269	14.21*	11.32*	25.42*	11.41*	16.34*	16.26*	8.98*	17.34*	17.79*	11.03*	29.34*	26.48*	0.965
2	0.055	0.324	-3.03*	-1.61	1.08	17.70*	1.07	-3.20*	-2.35*	9.13*	-5.05*	-2.88*	-3.77*	-1.34	0.821
3	0.034	0.358	2.81*	-1.23	-4.91*	-2.51*	-5.16*	-1.83	-0.02	9.29*	-0.24	-0.13	6.63*	-3.33*	0.671
4	0.025	0.383	-0.91	-0.32	-0.83	4.18*	-0.94	0.02	1.45	-4.66*	-2.09*	0.92	4.96*	-1.83	0.403
5	0.024	0.407	-0.35	-0.52	3.69*	-3.37*	2.72*	-0.04	-0.75	2.04*	-5.32*	0.55	0.26	-0.93	0.364
6	0.024	0.431	3.18*	0.15	-1.30	2.32*	-1.72	2.99*	1.56	0.95	1.95	3.65*	-4.36*	-0.93	0.374
7	0.019	0.450	-2.24*	-0.48	-2.82*	-0.71	-0.80	-0.90	-1.22	1.87	0.92	-1.04	1.08	3.12*	0.179
8	0.019	0.469	-0.66	-0.84	-0.50	-1.46	1.63	-3.04*	2.70*	2.12*	-0.25	2.65*	-0.84	-0.51	0.195

Table 7

Percentage of the Spread Due to Different Types of Default Risk Events. This table reports summary statistics for the percentages of the total CDX index spread due to the indicated Poisson processes. The percentage for the first Poisson process is given from the ratio of its jump size times its intensity ($\bar{\gamma}_1 \ \lambda_1$) to the total CDX index value, and similarly for the other two Poisson processes. Sample period is from October 2003 to October 2005.

Poisson Process	Index	Mean	Standard Deviation	Minimum	Median	Maximum	N
First	CDX 1	58.83	2.24	53.36	58.43	63.59	65
	CDX 2	57.70	2.38	54.11	56.85	64.44	108
	CDX 3	64.19	2.26	59.08	64.85	67.54	118
	CDX 4	71.56	4.37	61.57	72.91	80.01	127
	CDX 5	82.17	1.41	79.23	82.79	83.50	17
Second	CDX 1	33.34	2.79	29.34	32.19	40.24	65
	CDX 2	37.02	3.17	29.02	38.00	41.99	108
	CDX 3	28.88	2.98	23.65	28.20	34.82	118
	CDX 4	15.92	7.11	7.81	13.24	30.16	127
	CDX 5	10.48	0.69	9.83	10.12	11.71	17
Third	CDX 1	7.83	3.03	0.91	7.83	13.31	65
	CDX 2	5.28	2.32	0.06	5.27	10.82	108
	CDX 3	6.93	1.18	4.07	6.86	9.77	118
	CDX 4	12.52	3.77	5.04	12.26	19.97	127
	CDX 5	7.35	0.84	6.38	7.29	9.43	17

Table 8

CDX Index Tranche Pricing Errors. This table reports summary statistics for the pricing errors for the indicated CDX index tranches. The t-statistic for the mean is corrected for first-order serial correlation. Pricing errors are measured in basis points. The sample period is from October 2003 to October 2005.

Tranche	Index	Mean	Standard Deviation	t-Statistic for the Mean	Serial Correlation	N
0−3 Tranche	CDX 1	−7.00	8.56	−1.66	0.896	65
	CDX 2	−1.27	2.18	−1.33	0.916	108
	CDX 3	−0.19	0.97	−0.26	0.984	118
	CDX 4	−0.20	1.17	−0.55	0.850	127
	CDX 5	0.02	0.29	0.32	−0.019	17
3−7 Tranche	CDX 1	−3.21	3.34	−1.98	0.893	65
	CDX 2	−0.51	1.21	−1.16	0.880	108
	CDX 3	−0.21	−.73	−0.55	0.947	118
	CDX 4	0.00	1.19	0.01	0.894	127
	CDX 5	−0.03	0.28	−0.42	0.110	17
7−10 Tranche	CDX 1	−2.60	8.34	−0.66	0.887	65
	CDX 2	0.03	3.06	0.04	0.801	108
	CDX 3	0.61	3.32	0.30	0.968	118
	CDX 4	−0.27	4.53	−0.15	0.914	127
	CDX 5	0.16	1.12	0.51	0.152	17
10−15 Tranche	CDX 1	−6.62	8.05	−1.67	0.896	65
	CDX 2	−2.07	5.74	0.57	0.967	108
	CDX 3	−0.45	2.77	−0.22	0.984	118
	CDX 4	−0.27	2.26	−0.53	0.752	127
	CDX 5	−0.76	0.62	−0.31	0.486	17
15−30 Tranche	CDX 1	−8.09	1.88	−9.13	0.885	65
	CDX 2	−2.98	3.63	−1.66	0.936	108
	CDX 3	−0.83	1.90	−0.64	0.975	118
	CDX 4	−0.62	2.39	−0.72	0.892	127
	CDX 5	0.06	0.46	0.45	0.238	17

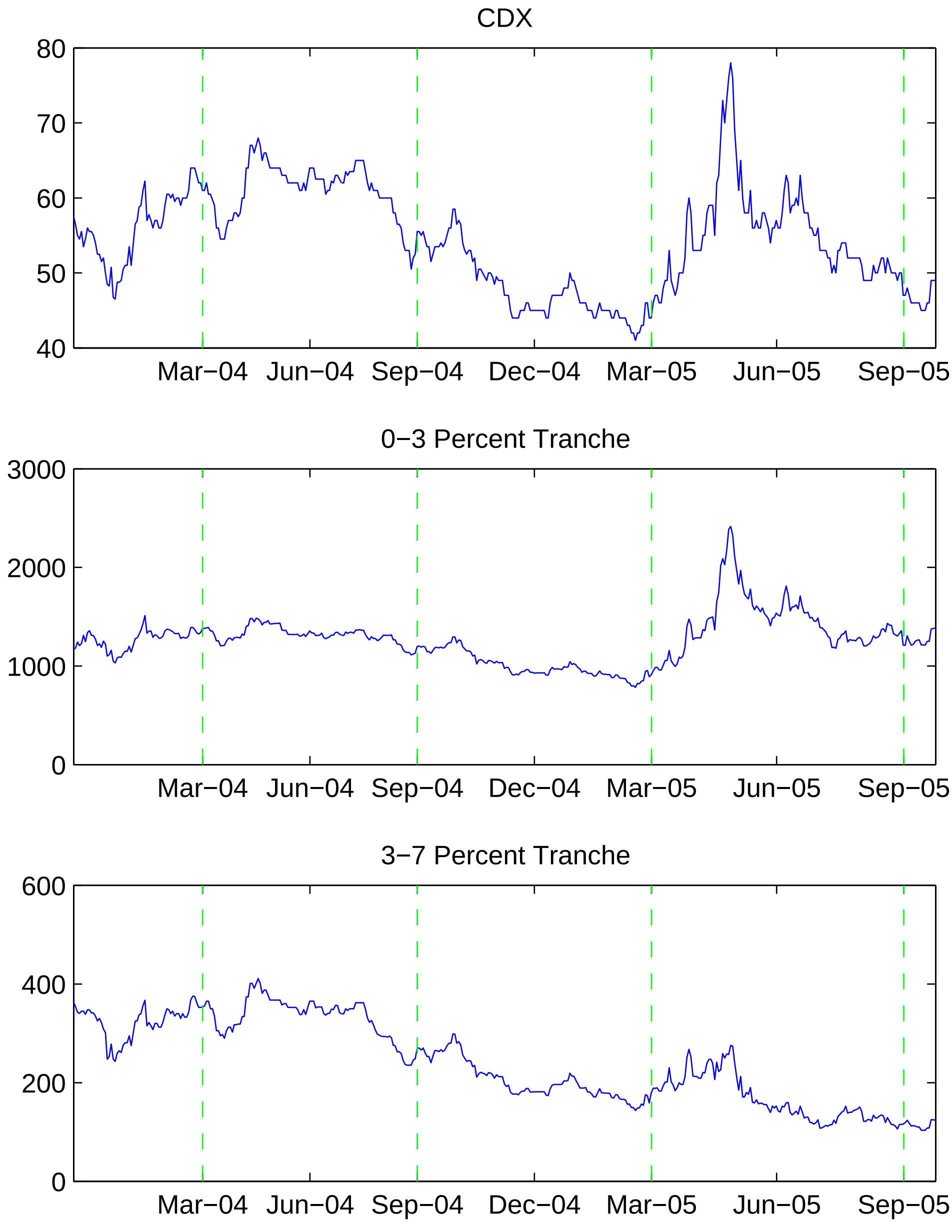

Fig. 1. CDX Index and Tranche Spreads. This figure graphs the time series of the CDX index and its tranche spreads for the October 2003 to October 2005 sample period. Spreads are in basis points. The vertical division lines denote the roll from one CDX index to the next.

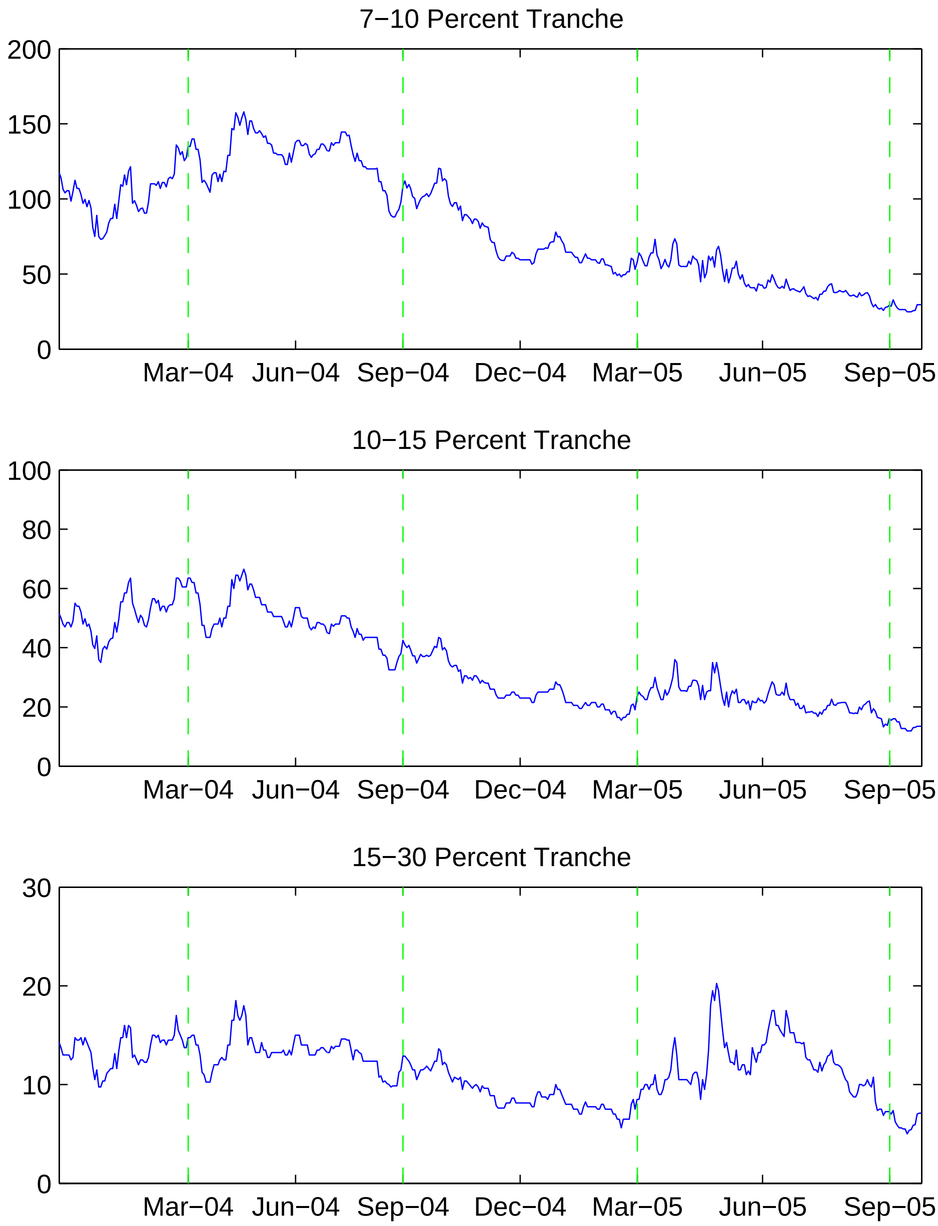

Fig. 1. Continued

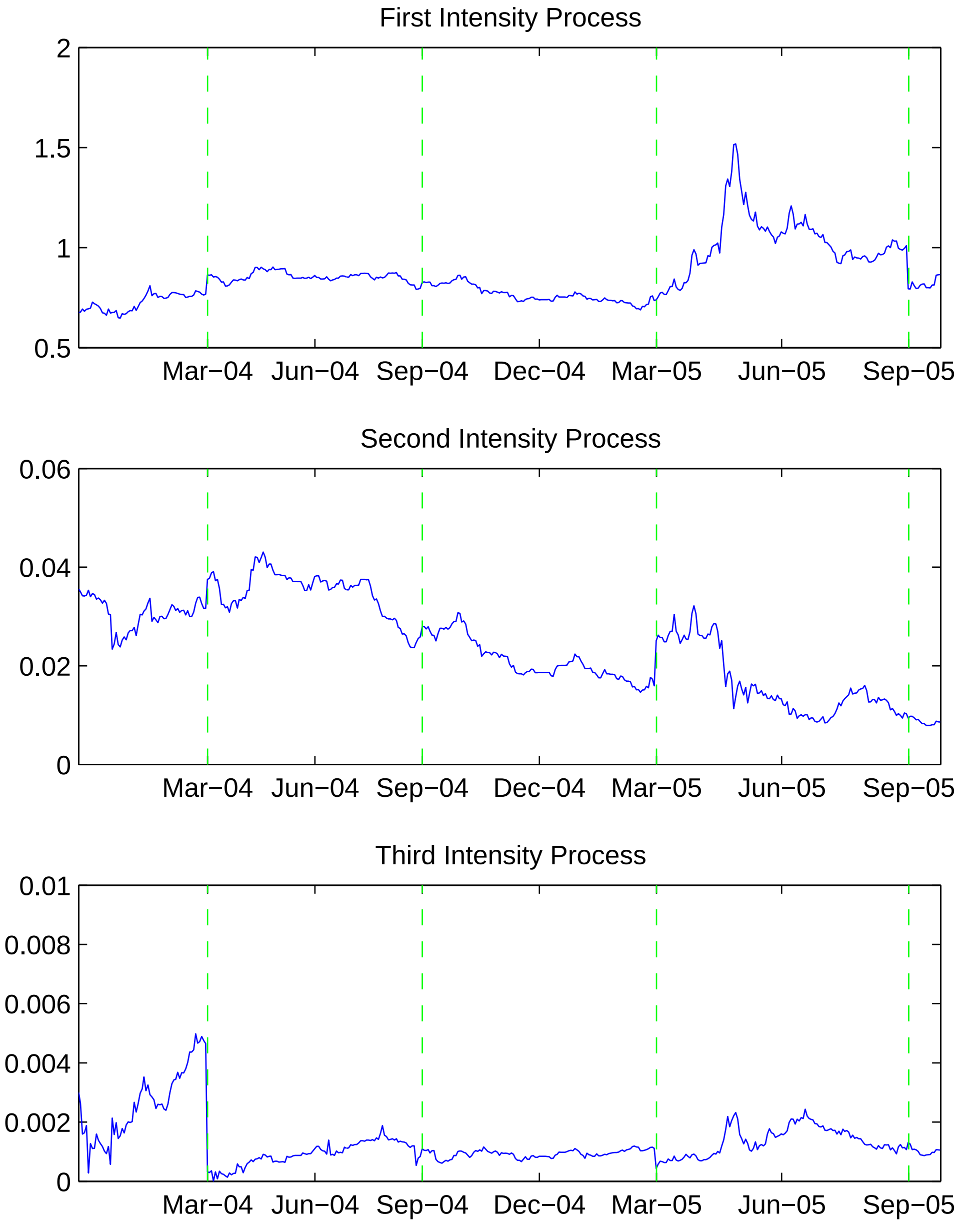

Fig. 2. Intensity Processes. This figure graphs the estimated intensity processes. The vertical division lines denote the roll from one CDX index to the next.

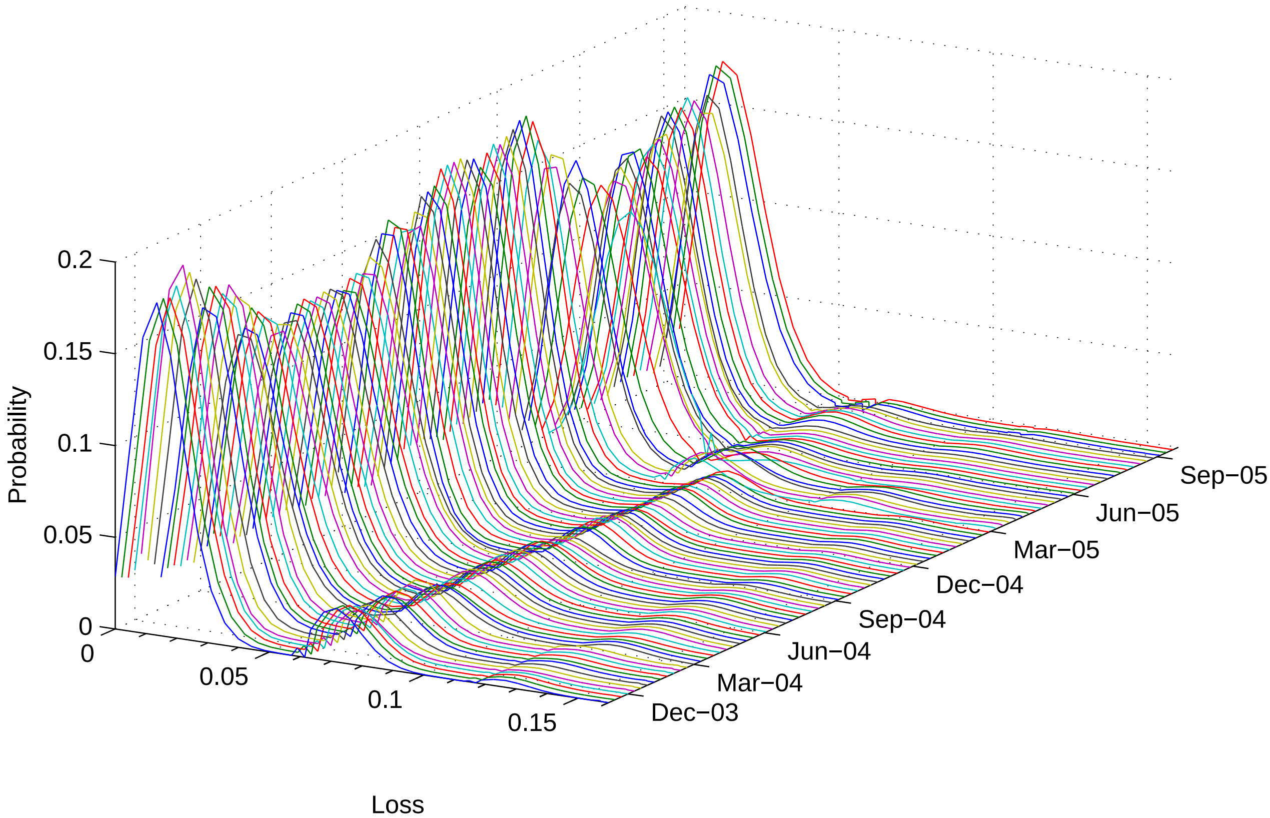

Fig. 3. Loss Distribution Functions. This figure graphs the loss distribution implied by the fitted three-factor model for losses ranging from 0.00 to 0.16. The actual loss distribution is discrete, but is approximated by a continuous function in the graph.

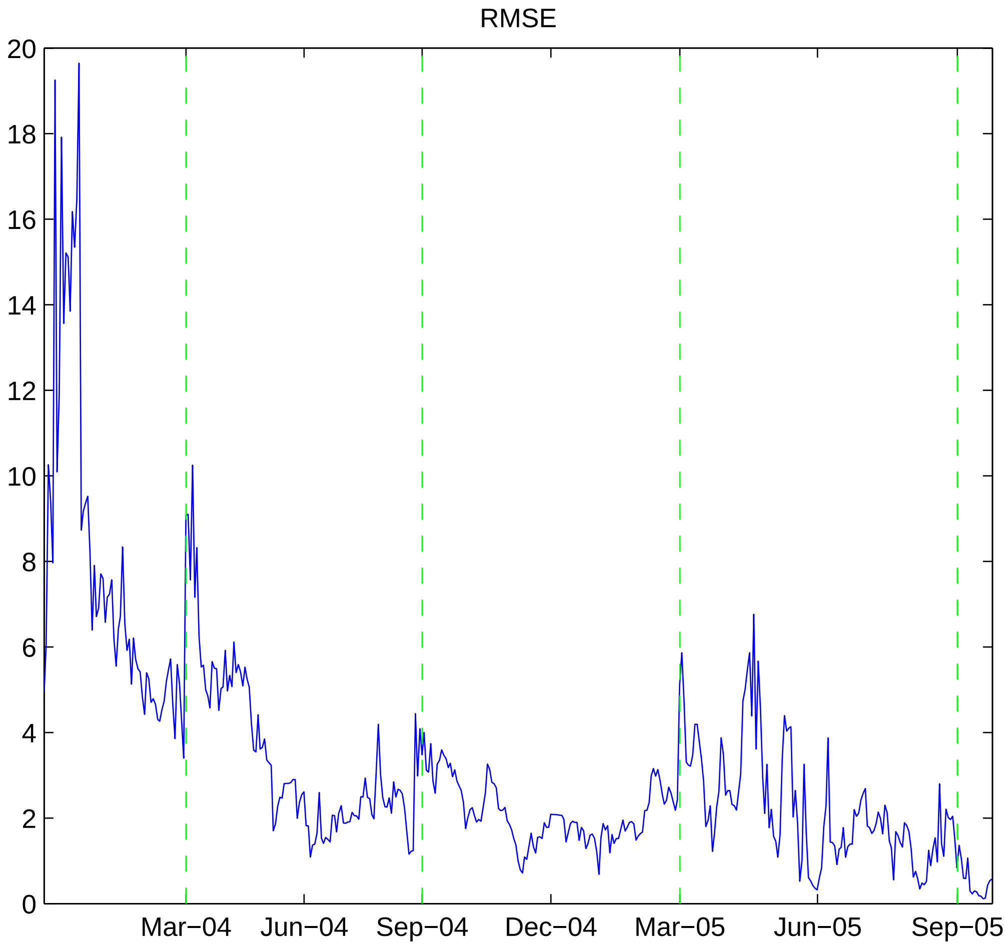

Fig. 4. Time Series of RMSEs. This figure graphs the time series of daily root mean squared error (RMSE) from the fitting of the five CDX tranches. RMSEs are measured in basis points. The vertical division lines denote the roll from one CDX index to the next.

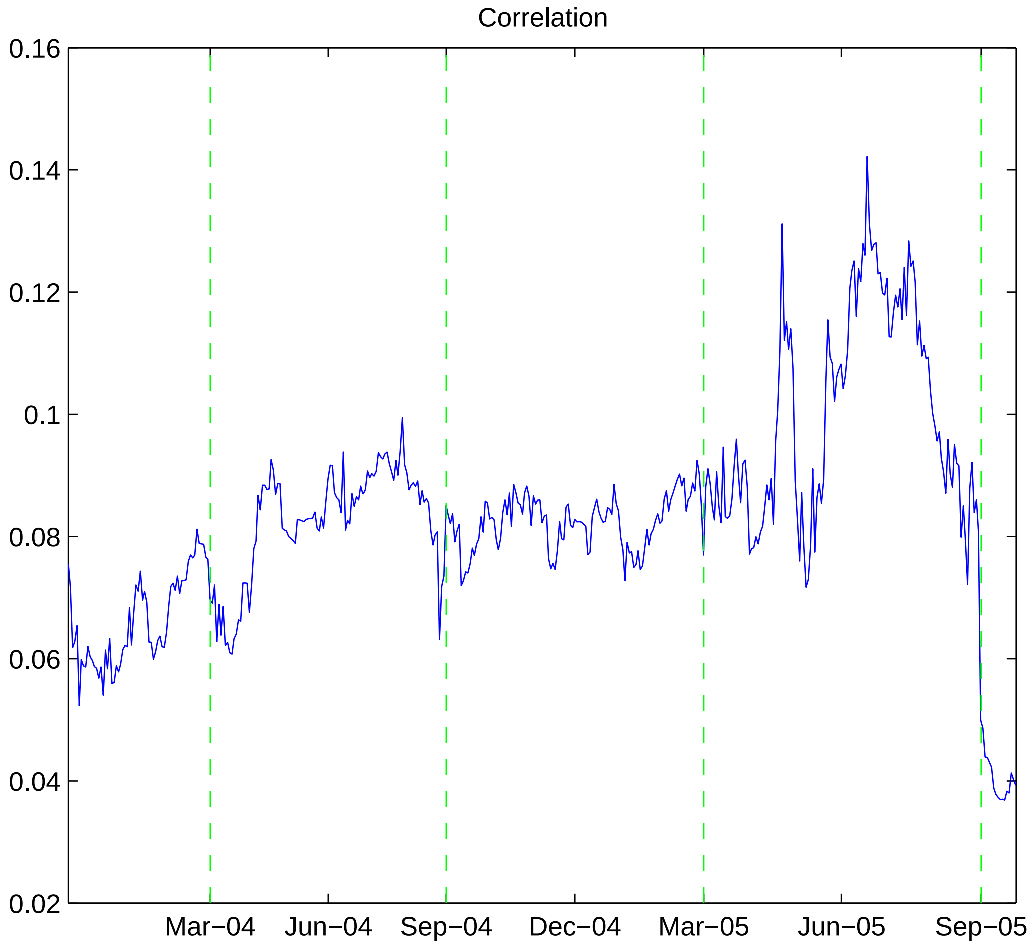

Fig. 5. Implied Default Correlations. This figure graphs the implied default correlation among firms in the CDX index. These default correlations are given by solving for the single correlation parameter that sets the variance of the sum of the default losses for the individual firms in the CDX index, where the joint distribution of default events is distributed as a multivariate correlated Bernoulli distribution, equal to the implied variance of the distribution of portfolio losses after five years estimated by fitting the three-factor model to the data. The vertical division lines denote the roll from one CDX index to the next.